AN INTRODUCTION TO STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

John B. WALSH

The general problem is this. Suppose one is given a physical system governed by a partial differential equation. Suppose that the system is then perturbed randomly, perhaps by some sort of a white noise. How does it evolve in time? Think for example of a guitar carelessly left outdoors. If $u(x,t)$ is the position of one of the strings at the point x and time t, then in calm air $u(x,t)$ would satisfy the wave equation $u_{tt} = u_{xx}$. However, if a sandstorm should blow up, the string would be bombarded by a succession of sand grains. Let $\dot{W}_{xt}$ represent the intensity of the bombardment at the point x and time t. The number of grains hitting the string at a given point and time will be largely independent of the number hitting at another point and time, so that, after subtracting a mean intensity, $\dot{W}$ may be approximated by a white noise, and the final equation is

$$u_{tt}(x,t) = u_{xx}(x,t) + \dot{W}(x,t),$$

where $\dot{W}$ is a white noise in both time and space, or, in other words, a two-parameter white noise.

One peculiarity of this equation - not surprising in view of the behavior of ordinary stochastic differential equations - is that none of the partial derivatives in it exist. However, one may rewrite it as an integral equation, and then show that in this form there is a solution which is a continuous, though non-differentiable, function.

In higher dimensions - with a drumhead, say, rather than a string - even this fails: the solution turns out to be a distribution, not a function. This is one of the technical barriers in the subject: one must deal with distribution-valued solutions, and this has generated a number of approaches, most involving a fairly extensive use of functional analysis.

Our aim is to study a certain number of such stochastic partial differential equations, to see how they arise, to see how their solutions behave, and to examine some techniques of solution. We shall concentrate more on parabolic equations than on hyperbolic or elliptic, and on equations in which the perturbation comes from something akin to white noise.

In particular, one class we shall study in detail arises from systems of branching diffusions. These lead to linear parabolic equations whose solutions are generalized Ornstein-Uhlenbeck processes, and include those studied by Ito, Holley and Stoock, Dawson, and others. Another related class of equations comes from certain neurophysiological models.

Our point of view is more real-variable oriented than the usual theory, and, we hope, slightly more intuitive. We regard white noise $\dot{W}$ as a measure on Euclidean space, $W(dx, dt)$, and construct stochastic integrals of the form $\int f(x,t)dW$ directly, following Ito's original construction. This is a two-parameter integral, but it is a particularly simple one, known in two-parameter theory as a "weakly-adapted integral". We generalize it to include integrals with respect to martingale measures, and solve the equations in terms of these integrals.

We will need a certain amount of machinery: nuclear spaces, some elementary Sobolev space theory, and weak convergence of stochastic processes with values in Schwartz space. We develop this as we need it.

For instance, we treat SPDE's in one space dimension in Chapter 3, as soon as we have developed the integral, but solutions in higher dimensions are generally Schwartz distributions, so we develop some elementary distribution theory in Chapter 4 before treating higher dimensional equations in Chapter 5. In the same way, we treat weak convergence of $\underline{\underline{S}}'$-valued processes in Chapter 6 before treating the limits of infinite particle systems and the Brownian density process in Chapter 8.

After comparing the small part of the subject we can cover with the much larger mass we can't, we had a momentary desire to re-title our notes: "An Introduction to an Introduction to Stochastic Partial Differential Equations";

which means that the introduction to the notes, which you are now reading, would be the introduction to "An Introduction ... ", but no. It is not good to begin with an infinite regression. Let's just keep in mind that this is an introduction, not a survey. While we will forego much of the recent work on the subject, what we do cover is mathematically interesting and, who knows? Perhaps even physically useful.

CHAPTER ONE

WHITE NOISE AND THE BROWNIAN SHEET

Let $(E,\underline{\underline{E}},\nu)$ be a σ-finite measure space. A white noise based on ν is a random set function W on the sets $A \in \underline{\underline{E}}$ of finite ν-measure such that

(i) $W(A)$ is a $N(0,\nu(A))$ random variable;

(ii) if $A \cap B = \phi$, then $W(A)$ and $W(B)$ are independent and

$W(A \cap B) = W(A) + W(B)$.

In most cases, E will be a Euclidean space and ν will be Lebesgue measure. To see that such a process exists, think of it as a Gaussian process indexed by the sets of $\underline{\underline{E}}$: $\{W(A),\ A \in \underline{\underline{E}},\ \nu(A) < \infty\}$. From (i) and (ii) this must be a mean-zero Gaussian process with covariance function C given by

$$C(A,B) = E\{W(A)\ W(B)\} = \nu(A \cap B).$$

By a general theorem on Gaussian processes, if C is positive definite, there exists a Gaussian process with mean zero and covariance function C. Now let $A_1,\dots,A_n$ be in $\underline{\underline{E}}$ and let $a_1,\dots,a_n$ be real numbers.

$$\sum_{i,j} a_i a_j\ C(A_i,A_j) = \sum_{i,j} a_i a_j \int I_{A_i}(x)\ I_{A_j}(x)dx$$

$$= \int (\sum_i a_i\ I_{A_i}(x))^2 dx \geq 0.$$

Thus C is a positive definite, so that there exists a probability space $(\Omega,\underline{\underline{F}},P)$ and a mean zero Gaussian process $\{W(A)\}$ on $(\Omega,\underline{\underline{F}},P)$ such that W satisfies (i) and (ii) above.

There are other ways of defining white noise. In case $E = R$ and $\nu =$ Lebesgue measure, it is often described informally as the "derivative of Brownian motion". Such a description is possible in higher dimensions too, but it involves the Brownian sheet rather than Brownian motion.

Let us specialize to the case $E = R^n_+ = \{(t_1,\dots,t_n):\ t_i \geq 0,\ i=1,\dots,n\}$ and $\nu =$ Lebesgue measure. If $t = (t_1,\dots,t_n) \in R^n_+$, let $(0,t] = (0,t_1]\times\cdots\times(0,t_n]$. The Brownian sheet on R^n_+ is the process $\{W_t,\ t \in R^n_+\}$ defined by $W_t = W\{(0,t]\}$. This is a mean-zero Gaussian process. If $s = (s_1,\dots,s_n)$ and $t = (t_1,\dots,t_n)$, its covariance function is

(1.1) $$E\{W_s W_t\} = (s_1 \wedge t_1) \cdots (s_n \wedge t_n)$$

If we regard $W(A)$ as a measure, W_t is its distribution function. Notice that we can recover the white noise in $\mathbf{R}_+^n$ from W_t, for if $\mathbf{R}$ is a rectangle, $W(R)$ is given by the usual formula (if $n = 2$ and $0 \leq u \leq s$, $0 \leq v \leq t$, $W((u,v),(s,t)] = W_{st} - W_{sv} - W_{ut} - W_{uv}$). If A is a finite union of rectangles, $W(A)$ can be computed by additivity, and a general Borel set A of finite measure can be approximated by finite unions of rectangles A_n in such a way that

$$E\{(W(A) - W(A_n))^2\} = \nu(A - A_n) + \nu(A_n - A) \to 0 .$$

Interestingly, the Brownian sheet was first introduced by a statistician, J. Kitagawa, in 1951 in order to do analysis of variance in continuous time. To get an idea of what this process looks like, let's consider its behavior along some curves in $\mathbf{R}_+^2$, in the case $n = 2$, ν = Lebesgue measure.

1). W vanishes on the axes. If $s = s_o > 0$ is fixed, $\{W_{s_o t}, t \geq 0\}$ is a Brownian motion, for it is a mean-zero Gaussian process with covariance function $C(t,t') = s_o(t \wedge t')$.

2). Along the hyperbola $st = 1$, let

$$X_t = W_{e^t, e^{-t}}.$$

Then $\{X_t, -\infty < t < \infty\}$ is an Ornstein-Uhlenbeck process, i.e. a strictly stationary Gaussian process with mean zero, variance 1, and covariance function

$$C(s,t) = E\{W_{e^s, e^{-s}} W_{e^t, e^{-t}}\} = e^{-|s-t|}.$$

3). Along the diagonal the process $M_t = W_{tt}$ is a martingale, and even a process of independent increments, although it is not a Brownian motion, for these increments are not stationary. The same is true if we consider W along <u>increasing</u> paths in $\mathbf{R}_+^2$.

4). Just as in one parameter, there are scaling, inversion, and translation transformations which take one Brownian sheet into another.

<u>Scaling:</u> $$A_{st} = \frac{1}{ab} W_{a^2 s, b^2 t}.$$

<u>Inversion:</u> $$C_{st} = st\, W_{\frac{1}{s} \frac{1}{t}} ; \quad D_{st} = s\, W_{\frac{1}{s} t} .$$

<u>Translation by (s_o,t_o):</u> $$E_{st} = W_{s_o+s, t_o+t} - W_{s_o+s, t_o} - W_{s_o, t_o+t} + W_{s_o t_o}.$$

Then A, C, D, and E are Brownian sheets, and moreover, E is independent of $F^*_{s_o t_o} = \sigma\{W_{uv} : u \le s_o \text{ or } v \le t_o\}$.

The easiest way to see this in the case of A, C and D is to notice that they are all mean zero Gaussian processes with the right covariance function. In the case of E, we can go back to white noise, and notice that $E_{st} = W((s_o,s] \times (t_o,t])$. The result then follows immediately from the properties (i) and (ii).

5). Another interesting transformation is this: let $U_{st} = e^{-s-t} W_{e^{2s},e^{2t}}$. Then $\{U_{st}, -\infty<s, t<\infty\}$ is an <u>Ornstein-Uhlenbeck sheet</u>. This is a stationary Gaussian process on $\mathbb{R}^2$ with covariance function $E\{U_{st}U_{uv}\} = e^{-|u-s|-|t-v|}$. If we look at U along any line, we get a one-parameter Ornstein-Uhlenbeck process. That is, if $V_s = U_{s,a+bs}$, then $\{V_s, -\infty<s<\infty\}$ is an Ornstein-Uhlenbeck process.

SAMPLE FUNCTION PROPERTIES

The Brownian sheet has continuous paths, but we would not expect them to be differentiable - indeed, nowhere-differentiable processes such as Brownian motion can be embedded in the sheet, as we have just seen. We will see just how continuous they are. This will give us an excuse to derive several beautiful and useful inequalities, beginning with an elegant result of Garsia, Rodemich, and Rumsey.

Let $\Psi(x)$ and $p(x)$ be positive continuous functions on $(-\infty,\infty)$ such that both Ψ and p are symmetric about 0, p(x) is increasing for x > 0 and p(0) = 0, and Ψ is convex with $\lim_{x\to\infty} \Psi(x) = \infty$. If R is a cube in $\mathbb{R}^n$, let e(R) be the length of its edge and |R| its volume. Let R_1 be the unit cube.

<u>THEOREM 1.1</u>. If f is a measurable function on R_1 such that

(1.2)
$$\int_{R_1}\int_{R_1} \Psi\left(\frac{f(y)-f(x)}{p(|y-x|/\sqrt{n})}\right) dx\, dy = B < \infty,$$

then there is a set K of measure zero such that if $x,y \in R_1 - K$

(1.3)
$$|f(y) - f(x)| \le 8 \int_0^{|y-x|} \Psi^{-1}\left(\frac{B}{u^{2n}}\right) dp(u).$$

If f is continuous, (1.3) holds for all x and y.

PROOF. If $Q \subset R_1$ is a rectangle and $x,y \in Q$, then $|y-x| \leq \sqrt{n}\, e(Q)$. Ψ is increasing, so that (3.2) implies

$$\int_Q\int_Q \Psi\left(\frac{f(y)-f(x)}{p(e(Q))}\right)dxdy \leq B. \tag{1.4}$$

Let $Q_0 \supset Q_1 \supset \cdots$ be a sequence of subcubes of R_1 such that

$$p(e(Q_j)) = \frac{1}{2}\, p(e(Q_{j-1})).$$

For any cube Q, let $f_Q = \frac{1}{|Q|}\int_Q f(x)dx$.

Since Ψ is convex

$$\Psi\left(\frac{f_{Q_j}-f_{Q_{j-1}}}{p(e(Q_{j-1}))}\right) \leq \frac{1}{|Q_{j-1}|}\int_{Q_{j-1}} \Psi\left(\frac{f_{Q_j}-f(x)}{p(e(Q_{j-1}))}\right)dx$$

$$\leq \frac{1}{|Q_{j-1}||Q_j|}\int_{Q_{j-1}}\int_{Q_j} \Psi\left(\frac{f(y)-f(x)}{p(e(Q_{j-1}))}\right)dx\,dy$$

By (1.4) this is

$$\leq \frac{B}{|Q_{j-1}||Q_j|}.$$

If Ψ^{-1} is the (positive) inverse of Ψ

$$|f_{Q_j}-f_{Q_{j-1}}| \leq p(e(Q_{j-1}))\,\Psi^{-1}\left(\frac{B}{|Q_j||Q_{j-1}|}\right). \tag{1.5}$$

Now $p(e(Q_{j-1})) = 4|(p(e(Q_{j+1})) - p(e(Q_j))|$, so this is

$$= 4\Psi^{-1}\left(\frac{B}{|Q_j||Q_{j-1}|}\right)|p(e(Q_{j+1})) - p(e(Q_j))|.$$

Ψ^{-1} increases so if $e(Q_{j+1}) \leq u \leq e(Q_j)$, then $|Q_{j-1}||Q_j| \geq u^{2n}$ and

$$\Psi^{-1}\left(\frac{B}{|Q_{j-1}||Q_j|}\right) \leq \Psi^{-1}\left(\frac{B}{u^{2n}}\right).$$

Set $v_j = e(Q_j)$. Then from (1.5)

$$|f_{Q_j}-f_{Q_{j-1}}| \leq 4\int_{v_j}^{v_{j-1}} \Psi^{-1}\left(\frac{B}{u^{2n}}\right)dp(u). \tag{1.6}$$

Sum this over j:

$$\limsup |f_{Q_j}-f_{Q_0}| \leq 4\int_0^{v_0} \Psi^{-1}\left(\frac{B}{u^{2n}}\right)dp(t). \tag{1.7}$$

By the Vitali theorem, if x is not in some null set K, then $f_{Q_j} \to f(x)$ for any sequence Q_j of cubes decreasing to $\{x\}$. If x and y are in $R_1 - K$, and if Q_0 is the smallest cube containing both, then, since $v_0 \leq |y-x|$,

$$|f(x) - f_{Q_0}| \leq 4\int_0^{|y-x|}\Psi^{-1}\left(\frac{B}{u^{2n}}\right)dp(u).$$

The same inequality holds for y, proving the theorem.

This is purely a real-variable result - f is deterministic - but it can often be used to estimate the modulus of continuity of a stochastic process. One usually computes $E\{B\}$ to show $B < \infty$ a.s. Everything hinges on the choice of Ψ and p. If we let Ψ be a power of x, we get a venerable result of Kolmogorov.

COROLLARY 1.2 (Kolmogorov). Let $\{X_t, t \in R_1\}$ be a real-valued stochastic process. Suppose there are constants $k > 1$, $K > 0$ and $\varepsilon > 0$ such that for all $s,t \in R_1$

$$E\{|X_t - X_s|^k\} \leq K|t-s|^{n+\varepsilon} .$$

Then

(i) X has a continuous version;

(ii) there exist constants C and γ, depending only on n, k, and ε, and a random variable Y such that with probability one, for all $s, t \in R_1$

$$|X_t - X_s| \leq Y\, |t-s|^{\varepsilon/k}\left(\log \frac{\gamma}{|t-s|}\right)^{2/k}$$

and

$$E\{Y^k\} \leq CK;$$

(iii) if $E\{|X_t|^k\} < \infty$ for some t, then

$$E\{\sup_{t\in R_1} |X_t|^k\} < \infty.$$

PROOF. We will apply Theorem 1.1 to the paths of X. We will use s and t instead of x and y, and the function f(x) is replaced by the sample path $X_t(\omega)$ for a fixed ω. Choose $\Psi(x) = |x|^k$ and $p(x) = |x|^{\frac{2n+\varepsilon}{k}} \left(\log \frac{\gamma}{|x|}\right)^{2/k}$. If $\gamma = \sqrt{n}\, e^{\frac{k}{n}}$, p will be increasing on $(0, \sqrt{n})$. Notice that the quantity B in (1.2) is now random, for it depends on ω. Let us take its expectation. By Fubini's Theorem

$$E\{B\} = n^{n+\frac{\varepsilon}{2}} \int_{R_1}\int_{R_1} \frac{E\{|X_t - X_s|^k\}}{|t-s|^{2n+\varepsilon}\log^2\left(\frac{\gamma}{|t-s|}\right)}\, ds\, dt$$

$$\leq n^{n+\frac{\varepsilon}{2}} K \int_{R_1}\int_{R_1} \frac{ds\ dt}{|t-s|^n \log^2(\gamma|t-s|^{-1})}$$

If t is fixed, the integral over R_1 with respect to s is dominated by the integral over the ball of radius $\sqrt{n}$ centered at t, since the ball contains R_1, whose diameter is $\sqrt{n}$. Let σ_n be the area of the unit sphere in R^n. Then integrate in polar coordinates:

$$\leq n^{n+\frac{\varepsilon}{2}} \sigma_n K \int_0^{\sqrt{n}} \frac{r^{n-1}dr}{r^n(\log \frac{\gamma}{2})^2}$$

$$= n^{n+1+\frac{\varepsilon}{2}} \frac{\sigma_n}{k} K.$$

If we integrate by parts twice in (1.3), we get

$$|X_t - X_s| \leq 8B^{1/k}[\,|t-s|^{\frac{\varepsilon}{k}}(\log \gamma|t-s|^{-1})^{\frac{2}{b}}(1 + \frac{2}{\varepsilon} n)]$$

$$+ \frac{4n}{k\varepsilon} \int_0^{|t-s|} \log^{\frac{2}{k}-1}(\frac{\gamma}{u})u^{\frac{\varepsilon}{k}-1} du.$$

The integral is dominated by $|t-s|^{\varepsilon/k}(\log \gamma|t-s|^{-1})^{2/k}$ for small enough values of $|t-s|$ - and for all $|t-s| \leq \sqrt{n}$ if $k \geq n$ - so that for a suitable constant A, we have

$$\leq 8AB^{1/k}\ |t-s|^{\varepsilon/k}(\log \gamma|t-s|^{-1})^{2/k}.$$

Then we take $Y = 8AB^{1/k}$, proving (i) and (ii).

To see (iii), just note that if $s, t \in R_1$, then $|t-s| \leq \sqrt{n}$ so that

$$\sup_t |X_t| \leq |X_{t_0}| + Y\, n^{\varepsilon/2k}(\log \frac{\gamma}{\sqrt{n}})^{2/k}.$$

Since X_{t_0} and Y are in L^k, so is $\sup_t |X_t|$.

Q.E.D.

The great flexibility in the choice of Ψ and p is useful, but it has its disadvantages: one always suspects he could have gotten a better result had he chosen them more cleverly. For example, if we take p(x) to be $|x|^{\frac{2n+\varepsilon}{k}}(\log\frac{\gamma}{|x|})^{1/k}(\log\log\frac{\gamma}{|x|})^{2/k}$, we can improve the modulus of continuity of X to $Y|t-s|^{\varepsilon/k}(\log \gamma|t-s|^{-1})^{1/k}(\log\log \gamma|t-s|^{-1})^{2/k}$, and so on.

If we apply Theorem 1.1 to Gaussian processes, we get the following result.

<u>COROLLARY 1.3</u>. Let $\{X_t, t\in R_1\}$ be a mean zero Gaussian process, and set

$$p(u) = \max_{|s-t|\leq|u|\sqrt{n}} E\{|X_t-X_s|^2\}^{1/2}.$$

If $\int_0^1 (\log \frac{1}{u})^{1/2} dp(u) < \infty$, then X has a continuous version whose modulus of continuity $\Delta(\delta)$ satisfies

(1.8) $$\Delta(\delta) \leq C \int_0^\delta (\log \frac{1}{u})^{1/2} dp(u) + Yp(\delta)$$

where C is a universal constant and Y is a random variable such that $E\{\exp(Y^2/256)\} \leq \sqrt{2}$.

PROOF. Let $\Psi(x) = e^{x^2/4}$. Note that $U_{st} \overset{\text{def}}{=} \dfrac{X_t - X_s}{p(|t-s|/\sqrt{n})}$ is Gaussian with mean zero and variance $\sigma^2(s,t) \leq 1$. We can use the Gaussian density to calculate directly that

$$E\{B\} = \int_{R_1}\int_{R_1} E\{\exp(\tfrac{1}{4} U_{st})\} ds\ dt \leq \sqrt{2}.$$

Thus $B < \infty$ a.s. Now $\Psi^{-1}(u) = \sqrt{4 \log \frac{1}{u}}$ so Theorem 1.1 implies that if $|s-t| \leq \delta$, and if s,t are not in some exceptional null-set,

(1.9) $$|X_t - X_s| \leq 16|\log B|^{1/2} p(\delta) + 16\sqrt{2n} \int_0^\delta |\log \frac{1}{u}|^{1/2} dp(u).$$

It now follows easily that X has a continuous version and that $\Delta(\delta)$ is bounded by the right hand side of (1.9), proving (1.8), with $C = 16\sqrt{2n}$ and $Y = 16|\log B|^{1/2}$. Q.E.D.

This theorem usually gives the right order of magnitude for $\Delta(\delta)$, but it does not always give the best constants.

To apply this to the Brownian sheet, note that if $s = (s_1,\ldots,s_n)$, $t = (t_1,\ldots,t_n)$, then $E\{(W_t - W_s)^2\} \leq \sum_{i=1}^n |t_i - s_i| \leq \sqrt{n}\ |t-s|$. Thus $p(u) = \sqrt{nu}$. If $\rho(\delta) = 16\sqrt{2} \int_0^\delta |\frac{\log \frac{1}{u}}{2u}|^{1/2} du$ then Corollary 1.3 gives

PROPOSITION 1.4. W has a continuous version with modulus of continuity

(1.10) $$\Delta(\delta) \leq n\rho(\delta) + Y\sqrt{\delta}$$

where Y is a random variable with $E\{e^{Y^2/16}\} < \infty$. Moreover, with probability one, for all $t \in R_1$ simultaneously:

$$(1.11)\qquad \limsup_{|h|\to 0} \frac{W_{t+h} - W_h}{\sqrt{2|h| \log 1/|h|}} \leq 16\sqrt{2n}$$

Here (1.11) follows from (1.10) on noticing that $\frac{\rho(|h|)}{\sqrt{2h \log 1/|h|}} \to 16\sqrt{2}$ as $h \to 0$. The constants are not best possible. Orey and Pruitt have shown that the right-hand side of (1.11) is $\sqrt{n}$.

This gives the modulus of continuity of W_t. There is also a law of the iterated logarithm.

THEOREM 1.5. (i) $\limsup_{s,t\to\infty} \frac{W_{st}}{\sqrt{4st \log \log st}} = 1$ a.s.

(ii) $\limsup_{s,t\to 0} \frac{W_{st}}{\sqrt{4st \log \log \frac{1}{st}}} = 1$ a.s..

We will not prove this except to remark that (ii) is a direct consequence of (i) and the fact that $st\, W_{\frac{1}{s}\frac{1}{t}}$ is a Brownian sheet.

SOME REMARKS ON THE MARKOV PROPERTY

In order to keep our notation simple, let us consider only the case $n = 2$, so that the Brownian sheet becomes a two-parameter process W_{st}. We first would like to examine the analogue of the strong Markov property of Brownian motion: that Brownian motion restarts after each stopping time. We don't have stopping times in our set-up, but we can define stopping points. Let $(\Omega, \underline{\underline{F}}, P)$ be a probability space. Recall that $\underline{\underline{F}}^*_t = \sigma\{W_s : s_i \leq t_i$ for at least one $i=1,2.\}$ A random variable $T = (T_1, T_2)$ with values in $\mathbf{R}^2_+$ is a weak stopping point if the set

$$\{T_1 < t_1, T_2 < t_2\} \in \underline{\underline{F}}^*_t .$$

The main example we have in mind is this: let $\underline{\underline{F}}^1_{t_1} = \sigma\{W_{s_1 s_2} : s_1 \leq t_1\}$. If T_2 is a stopping time relative to the filtration $(\underline{\underline{F}}^2_{t_2})$ and if $S_1 \geq 0$ is measurable relative to $\underline{\underline{F}}^2_{T_2}$, then (S_1, T_2) is a weak stopping point.

For a weak stopping point T, set

$$\underline{\underline{F}}^*_T = \{A \in \underline{\underline{F}} : A \cap \{T_1 < t_1, T_2 < t_2\} \in \underline{\underline{F}}^*_t\}.$$

This is clearly a σ-field . Set

$$W_t^T = W((T,T+t]), \qquad t \in \mathbb{R}_+^2 ,$$

where the mass of the (random) rectangle (T,T+t] is computed from W_s by the usual formula.

THEOREM 1.6. Let T be a finite weak stopping point. Then the process $\{W_t^T, t \in \mathbb{R}_+^n\}$ is a Brownian sheet, independent of $\underline{\underline{F}}_T^*$.

PROOF. We approximate T from above as follows:

Write $T = (T^1,\ldots,T^n)$ and define $T_m = (T_m^1,\ldots,T_m^n)$ by $T_m^i = j\,2^{-m}$ if $(j-1)2^{-m} \leq T^i < j\,2^{-m}$. Let $\{r_i\}$ be any enumeration of the lattice points $(j_1 2^{-m},\cdots,j_n 2^{-m})$, and note that $\{T_m = r_i\} \in \underline{\underline{F}}_{r_i}^*$ for all i. Now for each t, $W(T,T+t] = \lim W(T_m,T_m+t]$, by continuity of W. For any set $A \in \underline{\underline{F}}_T^* \subset \underline{\underline{F}}_{T_m}^*$, and any Borel set B

$$P\{W(T_m,T_m+t] \in B;\ A\}$$
$$= \sum_i P\left\{W(r_i,r_i+t] \in B;\ A \cap \{T_m = r_i\}\right\}.$$

But $A \cap \{T_m = r_i\} \in \underline{\underline{F}}_{r_i}^*$ so, by the independence property of white noise this is

$$= \sum_i P\{W(r_i,r_i+t] \in B\}\ P\left\{A \cap \{T_m = r_i\}\right\}$$
$$= P\{W_t \in B\}P\{A\}.$$

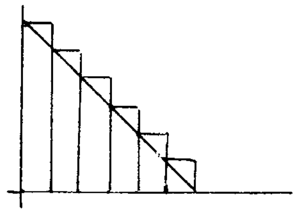
Thus, for each m, $\{W(T_m,T_m+t]\}$ is a Brownian sheet, independent of $\underline{\underline{F}}_T^*$. The same is therefore true of $\{W(T,T+t]\}$ in the limit. Q.E.D.

Notice that this is merely a random version of the translation property given at the beginning of this chapter.

A second, quite different type of Markov property is this. For any set $D \subset \mathbb{R}_+^2$ let $\underline{\underline{G}}_D = \sigma\{W_t,\ t\in D\}$, and let $\underline{\underline{G}}_D^* = \bigcap_{\varepsilon>0} \underline{\underline{G}}_{D^\varepsilon}$ where D^ε is an open ε-neighborhood of D. We say W satisfies <u>Lévy's Markov property</u> for D if $\underline{\underline{G}}_D$ and $\underline{\underline{G}}_{D^c}$ are conditionally independent given $\underline{\underline{G}}_{\partial D}^*$, where ∂D is the boundary of D. We say W satisfies <u>Lévy's sharp Markov property</u> if $\underline{\underline{G}}_D$ and $\underline{\underline{G}}_{D^c}$ are conditionally independent given $\underline{\underline{G}}_{\partial D}$.

It is rather easy to show W satisfies Lévy's sharp Markov property relative to a rectangle; this follows from the independence property of white noise. With slightly more work, one can show this also holds for finite unions of rectangles. It is more surprising to learn that it does not hold for all sets D. Indeed, consider the following example.

<u>EXAMPLE</u>. Let D be the triangle with corners at (0,0), (1,0) and (0,1). Since W_t vanishes on the axes, $\underline{\underline{G}}_{\partial D} = \{W_{s,1-s},\ 0\leq s\leq 1\}$. Let us notice that W(D) is measurable with respect to both $\underline{\underline{G}}_D$ and $\underline{\underline{G}}_{D^c}$.



Call the above union of rectangles D_n. The mass of each rectangle is given in terms of the W_{t_j}, where the t_j are the corners of the rectangles, and hence is $\underline{\underline{G}}_{D^c}$-measurable. Since $W(D_n) \to W(D)$, $W(D) \in \underline{\underline{G}}_{D^c}$. A similar argument shows $W(D) \in \underline{\underline{G}}_D$. (Moreover if V_ε is an ε-neighborhood of D, $W(D_n) \in \underline{\underline{G}}_{V^\varepsilon}$ for large n, hence $W(D) \in \underline{\underline{G}}^*_{\partial D}$ too). On the other hand, we claim W(D) is <u>not</u> measurable with respect to the sharp boundary field, and Lévy's sharp Markov property does not hold. We need only show $\hat{D} \overset{\text{def}}{=} E\{W(D)|\underline{\underline{G}}_{\partial D}\}$ is not equal to W(D).

Since all the random variables are Gaussian, $\hat{D}$ will be a linear combination of the $W_{s,1-s}$, determined by

$$E\{(W(D)-\hat{D})\, W_{s,1-s}\} = 0, \quad \text{all } 0 \leq s \leq 1.$$

Let

$$\hat{D} = 2\int_0^1 W_{u,1-u}\,du.$$

Then

$$E\{(W(D)-\hat{D})W_{s,1-s}\} = s(1-s)-2\int_0^s (1-s)u\,du - 2\int_0^1 s(1-u)du$$
$$= 0.$$

But $\hat{D} \neq W(D)$. Indeed, $E\{\hat{D}\,W(D)\} = 2\int_0^1 u(1-u)du = \frac{1}{3}$, while $E\{W(D)^2\} = |D| = \frac{1}{2}$.

Thus W does not satisfy Lévy's sharp Markov property relative to D. It is not hard to see that it satisfies Lévy's Markov property, i.e. that the germ field $\underline{\underline{G}}^*_{\partial D}$ is a splitting field. In fact this is always the case: W_t satisfies Lévy's Markov property relative to all Borel sets D, but we will not prove it here.

THE PROPAGATION OF SINGULARITIES

The Brownian sheet is far from rotation invariant, even far away from the axes. Any unusual event has a tendency to propagate parallel to the axes. Let us look at an example of this.

We need two facts about ordinary Brownian motion. Let $\{B_t, t\geq 0\}$ be a standard Brownian motion.

1.) For any fixed t, $\limsup_{h\downarrow 0} \frac{B_{t+h}-B_t}{\sqrt{2h \log\log 1/h}} = 1$ a.s.

2.) For a.e. ω, there exist uncountably many t for which

$$\limsup_{h\downarrow 0} \frac{B_{t+h}(\omega)-B_t(\omega)}{\sqrt{2h \log\log 1/h}} = \infty .$$

The first fact is well-known, and the second a consequence of the fact that the exact modulus of continuity of Brownian motion is $\sqrt{2h \log 1/h}$, not $\sqrt{2h \log\log 1/h}$.

Indeed, Lévy showed that if $d(h) = (2h \log 1/h)^{1/2}$, then for any $\varepsilon > 0$ and $a < b$, there exist $a < s < t < b$ for which $|B_t - B_s| > (1-\varepsilon)\, d(t-s)$. Thus we can choose $s_1 < t_1$ in (a,b) such that $|B_{t_1} - B_{s_1}| > \frac{1}{2} d(t_1 - s_1)$. Having chosen $s_1,\ldots,s_n$ and $t_1,\ldots,t_n$, choose s_n' such that $s_n' \in (s_n,t_n)$, $s_n' \leq s_n + 2^{-n}$ and such that $|B_{t_n} - B_s| > \frac{n}{n+1} d(t_n - s)$ for all $s \in (s_n, s_n')$, which we can do by continuity. Next, choose $s_{n+1} < t_{n+1}$ in (s_n, s_n') for which $|B_{t_{n+1}} - B_{s_{n+1}}| > \frac{n+1}{n+2} d(t_{n+1} - s_{n+1})$. Now let $s_0 \in \bigcap_n [s_n, t_n]$. If $h_n = t_n - s_0$, $|B_{s_0+h_n} - B_{s_0}| > \frac{n}{n+1} d(h_n)$. this shows that there is dense set of points s for which

$$\limsup_{h\downarrow 0} \frac{|B_{s+h}-B_s|}{\sqrt{2h \log 1/h}} = 1,$$

which is more than we claimed.

One can see there are uncountably many such s by modifying the construction slightly. In the induction step, just break each interval (s_n, s_n') into three parts, throw away the middle third, and operate with each of the two remaining parts as above. See Orey and Taylor's article for a detailed study of these singular points.

PROPOSITION 1.7. Fix s_0. Then, with probability one,

$$\limsup_{h\downarrow 0} \frac{W_{s_0+h,t} - W_{s_0 t}}{\sqrt{2h \log\log 1/h}} = \sqrt{t}, \tag{1.12}$$

simultaneously for all $t > 0$.

PROOF. (1.12) holds a.s. for each fixed t by the law of the iterated logarithm, hence it holds for a.e. t by Fubini. We must show it holds for all t. Set

$$L_t = \limsup_{h\downarrow 0} \frac{W_{s_0+h,t} - W_{s_0 t}}{\sqrt{2h \log\log 1/h}} .$$

It is easy to see that L_t is well-measurable relative to the fields $\underline{\underline{F}}_t^2 = \underline{\underline{F}}_{(0,\infty)\times(0,t]}$ for it is a measurable function of $W_{\cdot,t}$ which, being continuous and adapted to $(\underline{\underline{F}}_t^2)$, is itself well-measurable. By Meyer's section theorem, if $P\{\exists t \ni L_t \neq \sqrt{t}\} > 0$, there exists a finite stopping time T (relative to the $\underline{\underline{F}}_t^2$) such that $P\{L_T \neq \sqrt{T}\} > 0$.

But now let $B_{st} = W_{s,T+t} - W_{sT}$. B is again a Brownian sheet (apply Theorem 1.5 to the weak stopping point (0,T)) so that if $\delta > 0$ we have

$$|L_{T+\delta} - L_T| \leq \limsup \frac{|B_{s_0+h,\delta} - B_{s_0,\delta}|}{\sqrt{2h \log\log 1/h}} = \sqrt{\delta}.$$

It follows that if $L_T(\omega) \neq \sqrt{T}$, then $L_{T+\delta} \neq \sqrt{T+\delta}$ for small enough δ, i.e. $L_t \neq \sqrt{t}$ for a set of positive Lebesgue measure, a contradiction.

PROPOSITION 1.8. Fix $t_0 > 0$ and let $S \geq 0$ be a random variable which is $\underline{\underline{F}}_{t_0}^2$ measurable. Suppose that

$$\limsup_{h\downarrow 0} \frac{W_{S+h,t} - W_{S,t}}{\sqrt{2h \log\log 1/h}} = \infty \quad \text{a.s.} \tag{1.13}$$

for $t = t_0$. Then (1.13) also holds for all $t \geq t_0$. If S is $\sigma\{W_{st_0}, s \geq 0\}$ - measurable, then (1.13) holds for all $t \geq 0$.

PROOF. We can assume without loss of generality that $t_0 = 1$. Set $B_{st} = W((S,1),(S+s, 1+t)]$. Note that $(S,1)$ is a weak stopping point, so B_{st} is a Brownian sheet. By Proposition 1.7, it satisfies the log log law for all $t > 0$. Thus, if $t' = 1 + t$

$$\limsup_{h\downarrow 0} \frac{W_{S+h,t'} - W_{S,t'}}{\sqrt{2h \log\log 1/h}} \geq \limsup_{h\downarrow 0} \frac{W_{S+h,1} - W_{S,1}}{\sqrt{2h \log\log 1/h}} - \limsup_{h\downarrow 0} \frac{B_{ht} - B_{0t}}{\sqrt{2h \log\log 1/h}} = \infty - \sqrt{t} = \infty .$$

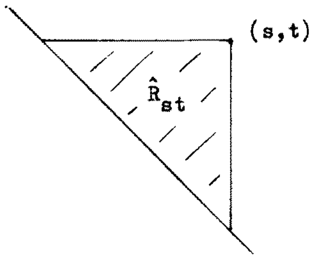
This proves (1.13) for all $t \geq 1$. Suppose $S \in \sigma\{W_{s1}, s \geq 0\}$. To see that (1.13) follows for $t < 1$ as well, set $\hat{W}_{st} = tW_{s\frac{1}{t}}$. Then $\hat{W}_{st}$ is a Brownian sheet, and $\hat{W}_{s1} = W_{s1}$ for all s. Clearly $S \in \sigma\{\hat{W}_{s1}, s \geq 0\}$. Thus $\hat{W}$ satisfies (1.13) for all $t > 1$, which implies that W satisfies (1.13) at $\frac{1}{t}$ for all $t > 1$. Q.E.D.

REMARKS. If we call a point at which the law of the iterated logarithm fails a singular point, the above proposition tells us that such singularities propagate vertically. By symmetry, there are singularities of the same type propagating horizontally. One can visualize these propagating singularities as wrinkles in the sheet.

THE BROWNIAN SHEET AND THE VIBRATING STRING

It is time to connect the Brownian sheet with our main topic, stochastic partial differential equations: the Brownian sheet gives the solution to a vibrating string problem.

Let us first modify the sheet as follows. Let D be the half plane $\{(s,t) : s + t \geq 0\}$ and put $\hat{R}_{st} = D \cap (-\infty,s] \times (-\infty,t]$. If W is a white noise, define $\hat{W}_{st} = W(\hat{R}_{st})$.



Then $\hat{W}$ is not a Brownian sheet: instead of vanishing on the coordinate axes, it vanishes on $\{s + t = 0\}$. However, it is easily seen that $W_{st} \overset{\text{def}}{=} \hat{W}_{st} - \hat{W}_{so} - \hat{W}_{ot}$, $s,t \geq 0$ is a Brownian sheet, and that the processes $s \to \hat{W}_{so}$ and $t \to \hat{W}_{ot}$ are independent continuous processes of independent increments. We can use this to read off many of the properties of $\hat{W}$ from those of W. In particular, the singularities of $\hat{W}$ propagate exactly like those of W.

Now let us put the sheet back in the closet for the moment and let us consider a vibrating string driven by white noise. One can imagine a guitar left outdoors during a sandstorm. The grains of sand hit the strings continually but irregularly. The number of grains hitting a portion dx of the string during a time ds will be essentially independent of those hitting a different portion dy during a time dt. Let W(dx,dt) be the (random) measure of the number hitting in (dx,dt), centered by subtracting the mean. Then W will be essentially a white noise, and we expect the position V(t,x) of the string to satisfy the inhomogeneous wave equation driven by a white noise. In order to avoid worrying about boundary conditions, we will assume that the string is infinite, and that it is initially at rest. Thus V should satisfy

$$(1.14)\quad \begin{cases} \dfrac{\partial^2 V}{\partial t^2}(x,t) = \dfrac{\partial^2 V}{\partial x^2}(x,t) + W(dx,dt), & t > 0,\ -\infty < x < \infty \\ V(x,0) = \dfrac{\partial V}{\partial t}(x,0) = 0, & -\infty < x < \infty. \end{cases}$$

Putting aside questions of the existence and uniqueness of solutions to (1.14), let us recall how to solve it when the driving term is a smooth function. If f(t,x) is smooth and bounded, then the solution to

$$\begin{cases} \dfrac{\partial^2 V}{\partial t^2} = \dfrac{\partial^2 V}{\partial x^2} + f \\ V(x,0) = \dfrac{\partial V}{\partial t}(x,0) = 0 \end{cases}$$

is given by

$$V(x,t) = \frac{1}{2} \int_0^t \int_{x+s-t}^{x+t-s} f(y,s)\,dy\,ds, \tag{1.15}$$

which can be checked by differentiating. Now let us rotate coordinates by 45°. Let $u = (s-y)/\sqrt{2}$, and $v = (s+y)/\sqrt{2}$, and set $\hat{V}(u,v) = V(y,s)$, $\hat{f}(u,v) = f(y,s)$. Then (1.15) implies that

$$\hat{V}(u,v) = \frac{1}{2} \int_0^v \int_{-v}^u \hat{f}(u',v')\,du'dv',$$

or

$$\hat{V}(u,v) = \frac{1}{2} \int_{\hat{R}_{uv}}\int \hat{f}(u',v')\,du'dv'. \tag{1.16}$$

By a slight act of faith, we see that the solution of (1.14) should be given by (1.15), with f dy ds replaced by W(dy,ds), or, in the form (1.16), that

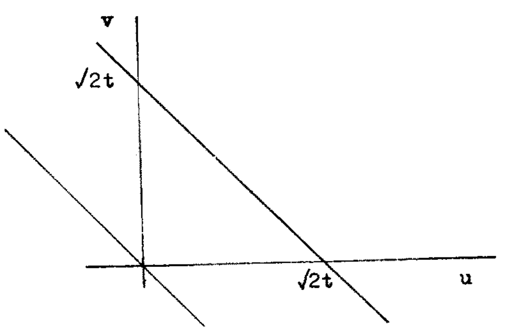
$$V(u,v) = \frac{1}{2} \int_{D \cap R_{uv}}\int dW$$

or, finally, that

$$\hat{V}(u,v) = \frac{1}{2} \hat{W}_{uv}, \quad u,v \geq 0,$$

where $\hat{W}$ is the modified Brownian sheet defined above.

We can conclude that the shape of the vibrating string at time t is just the cross-section of the sheet $\frac{1}{2}\hat{W}$ along the -45° line $u + v = \sqrt{2t}$.



This gives us a complete representation of the solution of (1.14), and it also gives us an interpretation of the propagating singularities. Like W, the sheet $\hat{W}$ has singularities propagating parallel to the axes of the uv-plane. In the xt-plane these propagate along the lines $x = c + t$ and $x = c - t$ respectively. Thus these propagating singularities correspond to travelling waves which move along the string with velocity one, the speed of propagation in the equation (1.14). In general, the tendency of unusual events to propagate parallel to the axes of the Brownian sheet can be understood as the propagation of waves in our vibrating string.

It also explains the rather puzzling failure of the Brownian sheet to satisfy the sharp Markov property. In fact, the initial conditions for the vibrating string involve not only the position V, but also the velocity $\frac{\partial V}{\partial t}$, and in order to calculate the velocity, one must know V in some neighborhood. This is exactly why we needed the germ field $\underline{\underline{G}}^{*}_{\partial D}$ for the Markov property in the above example. A more delicate analysis of the Markov property would show that the minimal splitting field is in fact made up of the values of V and its derivative on the boundary.

Exercise 1.1. For each fixed x, show that there exists a standard Brownian motion $\{B_s, s \geq 0\}$ such that $V(x,t) = \frac{1}{2} B(t^2)$, all $t \geq 0$. Show that with probability one, $\lim_{t \downarrow 0} \frac{1}{t} V(x,t)$ does not exist, while $\lim_{t \downarrow 0} \int_0^t \frac{1}{s} V(x,s)ds = 0$.
Discuss the initial condition $\frac{\partial V}{\partial t}(x,0) = 0$.

WHITE NOISE AS A DISTRIBUTION

One thinks of white noise on $\mathbf{R}$ as the derivative of Brownian motion. In two or more parameters, white noise can be thought of as the derivative of the Brownian sheet, and this can be made rigorous.

The Brownian sheet W_{st} is nowhere-differentiable in the ordinary sense, but its derivatives will exist in the sense of Schwartz distributions. Thus define

$$\dot{W}_{st} = \frac{\partial^2 W_{st}}{\partial s \, \partial t};$$

that is, if $\phi(s,t)$ is a test function with compact support in $\mathbf{R}^2_+$, $\dot{W}$ is the distribution

$$\dot{W}(\phi) = \int_{\mathbf{R}^2_+}\int W_{uv} \frac{\partial^2(\phi)}{\partial u \partial v}(u,v)\, dudv.$$ If we may anticipate

the introduction of stochastic integrals, let us note that this is almost surely

$$= \iint \phi \, dW.$$

Formally, if $\phi(u,v) = I_{\{0\leq u\leq s,\ 0\leq v\leq b\}}$ then

$$\dot{W}(\phi) = \int_0^s \int_0^t \frac{\partial^2 W}{\partial u \partial v} \, du \, dv = W_{st} = \iint \phi \, dW,$$

but it takes some work to make it rigorous. We leave it as an exercise.

If we regard the "measure" W as a distribution, then certainly $W(\phi) = \iint \phi \, dW$. In other words $\dot{W}(\phi) = W(\phi)$ so that $\dot{W}$ and W are the same distribution.

Note that in $\mathbf{R}^n$, $\dot{W}$ would be the n^{th} mixed partial:

$$\dot{W} = \frac{\partial^n}{\partial t_1 \cdots \partial t_n} W_{t_1 \cdots t_n} .$$

CHAPTER TWO

MARTINGALE MEASURES

We will develop a theory of integration with respect to martingale measures. We think of them as white noises, but we treat them differently. Instead of considering set functions on R^{d+1} with all coordinates treated symmetrically, we break off one coordinate to play the role of "time" and think of the remaining coordinates as "space".

Let us begin with some remarks on random set functions and vector-valued measures. Let $(E,\underline{\underline{E}})$ be a Lusin space, i.e. a measurable space homeomorphic to a Borel subset of the line. (This includes all Euclidean spaces and, more generally, all Polish spaces.)

Suppose $U(A,\omega)$ is a function defined on $\underline{\underline{A}} \times \Omega$, where $\underline{\underline{A}} \subset \underline{\underline{E}}$ is an algebra, and such that $E\{U(A)^2\} < \infty$, $A \in \underline{\underline{A}}$. Suppose that U is finitely additive: if $A \cap B = \phi$, $A, B \in \underline{\underline{A}}$, then $U(A \cup B) = U(A) + U(B)$ a.s.

In most interesting cases U will not be countably additive if we consider it as a real-valued set function. However, it may become countably additive if we consider it as a set function with values in, say, $L^2(\Omega,\underline{\underline{F}},P)$. This is the case, for instance, with white noise. Let $\|U(A)\|_2 = E\{U^2(A)\}^{1/2}$ be the L^2-norm of $U(A)$.

We say U is <u>σ-finite</u> if there exists an increasing sequence $(E_n) \subset \underline{\underline{E}}$ whose union is E, such that for all n

(i) $\underline{\underline{E}}_n \subset \underline{\underline{A}}$, where $\underline{\underline{E}}_n = \underline{\underline{E}}|_{E_n}$;

(ii) $\sup\{\|U(A)\|_2 : A \in \underline{\underline{E}}_n\} < \infty$.

Define a set function μ by

$$\mu(A) = \|U(A)\|_2^2.$$

A σ-finite additive set function U is countably additive on $\underline{\underline{E}}_n$ (as an L^2-valued set function) iff

$$A_j \in \underline{\underline{E}}_n,\ \forall n,\ A_j \downarrow \phi \Rightarrow \lim_{j\to\infty} \mu(A_j) = 0 \tag{2.1}$$

If U is countably additive on $\underline{\underline{E}}_n$, $\forall n$, we can make a trivial further extension: if $A \in \underline{\underline{E}}$, set $U(A) = \lim_{n\to\infty} U(A \cap E_n)$ if the limit exists in L^2, and let $U(A)$ be undefined otherwise. This leaves U unchanged on each $\underline{\underline{E}}_n$, but may change its value on some sets $A \in \underline{\underline{E}}$ which are not in any $\underline{\underline{E}}_n$. We will assume below that all our countably additive set functions have been extended in this way. We will say that such a U is a <u>σ-finite L^2-valued measure</u>.

<u>DEFINITION</u>. Let $(\underline{\underline{F}}_t)$ be a right continuous filtration. A process $\{M_t(A), \underline{\underline{F}}_t, t \geq 0, A \in \underline{\underline{A}}\}$ is a <u>martingale measure</u> if

(i) $M_o(A) = 0$;

(ii) if $t > 0$, M_t is a σ-finite L^2-valued measure;

(iii) $\{M_t(A), \underline{\underline{F}}_t, t \geq 0\}$ is a martingale.

<u>Exercise 2.1</u>. Let $\nu_t(A) = \sup\{E\{M_t(B)^2\}: B \subset A, B \in \underline{\underline{E}}\}$. Show that $t \to \nu_t(A)$ is increasing. Conclude that for each T, the same family (E_n) works for all M_t, $t \leq T$.

It is not necessary to verify the countable additivity for all t; one t will do, as the following exercise shows.

<u>Exercise 2.2</u>. If N is a σ-finite L^2-valued measure and $(\underline{\underline{F}}_t)$ a filtration, show that

$$M_t(A) = E\{N(A)|\underline{\underline{F}}_t\} - E\{N(A)|\underline{\underline{F}}_0\}$$

is a martingale measure.

Note: One commonly gets such an M_t by first defining it for a small class of sets and then constructing the L^2-valued measure from these. This is, in fact, exactly what one does when constructing a stochastic integral, although the fact that the result is a vector-valued measure is usually not emphasized. In the interest of a speedy development, we will assume that the L^2-measure has already been constructed. Thus we know how to integrate over dx for fixed t - this is the Bochner integral - and over dt for fixed sets A - this is the Ito integral. The problem facing us now is to integrate over dx and dt at the same time.

There are two rather different classes of martingale measures which have been popular, orthogonal martingale measures and martingale measures with a nuclear covariance.

DEFINITION. A martingale measure M is orthogonal if, for any two disjoint sets A and B in $\underline{\underline{A}}$, the martingales $\{M_t(A),\ t \geq 0\}$ and $\{M_t(B),\ t \geq 0\}$ are orthogonal.

Equivalently, M is orthogonal if the product $M_t(A)M_t(B)$ is a martingale for any two disjoint sets A and B. This is in turn equivalent to having $\langle M(A),\ M(B)\rangle_t$, the predictable process of bounded variation, vanish.

DEFINITION. A martingale measure M has nuclear covariance if there exists a finite measure η on $(E,\underline{\underline{E}})$ and a complete ortho-normal system (ϕ_k) in $L^2(E,\underline{\underline{E}},\eta)$ such that $\eta(A) = 0 \Rightarrow \mu(A) = 0$ for all $A \in \underline{\underline{E}}$ and

$$\sum_k E\{M_t(\phi_k)^2\} < \infty,$$

where $M_t(\phi_k) = \int \phi_k(x)M_t(dx)$ is a Bochner integral.

The canonical example of an orthogonal martingale measure is a white noise. If W is a white noise on $E \times \mathbb{R}_+$, let $M_t(A) = W(A \times [0,t])$. This is clearly a martingale measure, and if $A \cap B = \phi$, $M_t(A)$ and $M_t(B)$ are independent, hence orthogonal. Any martingale measure derived from a white noise this way will also be called a white noise.

If $(E,\underline{\underline{E}},\eta)$ is a finite measure space, if $f \in L^2(E,\underline{\underline{E}},\eta)$ and if $\{B_t,\ t \geq 0\}$ is a standard (one-dimensional) Brownian motion, then the measure defined by

$$M_t(A) = B_t \int_A f(x)\eta(dx)$$

has nuclear covariance, since for any CONS (ϕ_k), $\sum_k E\{M_t^2(\phi_k)\} = \|f\|_2^2$. More generally, if $B^1, B^2, \ldots$ are iid standard Brownian motions and if $a_1, a_2, \ldots$ are real numbers such that $\sum a_k^2 < \infty$, then

$$M_t(A) = \sum_k a_k B_t^k \int_A \phi_k(x)\eta(dx)$$

has nuclear covariance.

Note that it is only in exceptional cases, such as when E is a finite set, that a white noise will have nuclear covariance.

WORTHY MEASURES

Unfortunately, it is not possible to construct a stochastic integral with respect to all martingale measures - we will give a counter-example at the end of the chapter - so we will need to add some conditions. These are rather strong, and, though sufficient, are doubtless not necessary. However, they are satisfied for both orthogonal martingale measures and those with a nuclear covariance.

Let M be a σ-finite martingale measure. By restricting ourselves to one of the E_n, if necessary, we can assume that M is finite. We shall also restrict ourselves to a fixed time interval $[0,T]$. The extension to infinite measures and the interval $[0,\infty]$ is routine.

DEFINITION. The covariance functional of M is

$$\bar{Q}_t(A,B) = \langle M(A), M(B)\rangle_t.$$

Note that $\bar{Q}_t$ is symmetric in A and B and biadditive: for fixed A, $\bar{Q}_t(A,\cdot)$ and $\bar{Q}_t(\cdot,A)$ are additive set functions. Indeed, if $B \quad C = \phi$,

$$\begin{aligned}\bar{Q}_t(A, B \quad C) &= \langle M(A), M(B) + M(C)\rangle_t \\ &= \langle M(A), M(B)\rangle_t + \langle M(A), M(C)\rangle_t \\ &= \bar{Q}_t(A,B) + \bar{Q}_t(A,C).\end{aligned}$$

Moreover, by the general theory,

$$|\bar{Q}_t(A,B)| \leq Q_t(A,A)^{1/2}\, Q_t(B,B)^{1/2}.$$

A set $A \times B \times (s,t] \subset E \times E \times R_+$ will be called a rectangle. Define a set function Q on rectangles by

$$Q(A \times B \times (s,t]) = \bar{Q}_t(A,B) - \bar{Q}_s(A,B),$$

and extend Q by additivity to finite disjoint unions of rectangles, i.e. if $A_i \times B_i \times (s_i,t_i]$ are disjoint, $i = 1,\ldots,n$, set

$$(2.2)\qquad Q\Big(\bigcup_{i=1}^{n} A_i \times B_i \times (s_i,t_i]\Big) = \sum_{i=1}^{n} \big(\bar{Q}_{t_i}(A_i,B_i) - \bar{Q}_{s_i}(A_i,B_i)\big).$$

<u>Exercise 2.3</u>. Verify that Q is well-defined, i.e. if

$\Lambda = \bigcup_{i=1}^{n} A_i \times B_i \times (s_i,t_i] = \bigcup_{j=1}^{m} A'_j \times B'_j \times (s'_j,t'_j]$, each representation gives the same value for $Q(\Lambda)$ in (2.5). (Hint: use biadditivity.)

If $a_1,\dots,a_n \in \mathbb{R}$ and if $A_1,\dots,A_n \in \underline{\underline{E}}$ are disjoint, then for any $s < t$

$$(2.3)\qquad \sum_{j=1}^{n}\sum_{j=1}^{n} a_i a_j\, Q(A_i \times A_j \times (s,t]) \geq 0,$$

for the sum is

$$= \sum_{i,j} a_i a_j\big(\langle M(A_i), M(A_j)\rangle_t - \langle M(A_i), M(A_j)\rangle_s\big)$$

$$= \Big\langle \sum_i a_i(M_t(A_i) - M_s(A_i)), \sum_i a_i(M_t(A_i) - M_s(A_i))\Big\rangle \geq 0.$$

<u>DEFINITION</u>. A signed measure $K(dx\,dy\,ds)$ on $\underline{\underline{E}} \times \underline{\underline{E}} \times \underline{\underline{B}}$ is <u>positive</u> <u>definite</u> if for each bounded measurable function f for which the integral makes sense,

$$(2.4)\qquad \int_{E\times E\times R_+} f(x,s)f(y,s)K(dxdyds) \geq 0\quad .$$

For such a positive definite signed measure K, define

$$(f,g)_K = \int_{E\times E\times R_+} f(x,s)g(y,s)K(dxdyds).$$

Note that $(f,f)_K \geq 0$ by (2.4).

<u>Exercise 2.4</u>. Suppose K is symmetric in x and y. Prove Schwartz' and Minkowski's inequalities

$$(f,g)_K \leq (f,f)_K^{1/2}(g,g)_K^{1/2}$$

and

$$(f+g,\ f+g)_K^{1/2} \leq (f,f)_K^{1/2} + (g,g)_K^{1/2}.$$

It is not always possible to extend Q to a measure on $\underline{\underline{E}} \times \underline{\underline{E}} \times \underline{\underline{B}}$, where $\underline{\underline{B}}$ = Borel sets on R_+, as the example at the end of the chapter shows. We are led to the following definition.

DEFINITION. A martingale measure M is worthy if there exists a random σ-finite measure $K(\Lambda,\omega)$, $\Lambda \in \underline{\underline{E}} \times \underline{\underline{E}} \times \underline{\underline{B}}$, $\omega \in \Omega$, such that

(i) K is positive definite and symmetric in x and y;

(ii) for fixed A, B, $\{K(A \times B \times (0,t]),\ t \geq 0\}$ is predictable;

(iii) for all n, $E\{K(E_n \times E_n \times [0,T])\} < \infty$;

(iv) for any rectangle Λ, $|Q(\Lambda)| \leq K(\Lambda)$.

We call K the dominating measure of M.

The requirement that K be symmetric is no restriction. If not, we simply replace it by $K(dx\,dy\,ds) + K(dy\,dx\,ds)$. Apart from this, however it is a strong condition on M. We will show below that it holds for the two important special cases mentioned above: both orthogonal martingale measures and those with nuclear covariance are worthy. In fact, we can state with confidence that we will have no dealings with unworthy measures in these notes.

If M is worthy with covariation Q and dominating measure K, then K + Q is a positive set function. The σ-field $\underline{\underline{E}}$ is separable, so that we can first restrict ourselves to a countable subalgebra of $\underline{\underline{E}} \times \underline{\underline{E}} \times \underline{\underline{B}}$ upon which $Q(\cdot,\omega)$ is finitely additive for a.e. ω. Then K + Q is a positive finitely additive set function dominated by the measure 2K, and hence can be extended to a measure. In particular, for a.e. ω $Q(\cdot,\omega)$ can be extended to a signed measure on $E \times E \times \underline{\underline{B}}$, and the total variation of Q satisfies $|Q|(\Lambda) \leq K(\Lambda)$ for all $\Lambda \in \underline{\underline{E}} \times \underline{\underline{E}} \times \underline{\underline{B}}$. By (2.3), Q will be positive definite.

Orthogonal measures and white noises are easily characterized. Let $\Delta(E) = \{(x,x): x \in E\}$, be the diagonal of E.

PROPOSITION 2.1. A worthy martingale measure is orthogonal iff Q is supported by $\Delta(E) \times R_+$.

PROOF. $Q(A \times B \times [0,t]) = \langle M(A), M(B)\rangle_t$.
If M is orthogonal and $A \cap B = \phi$, this vanishes hence $|Q|[E \times E - \Delta(E)) \times R_+] = 0$, i.e. supp $Q \subset \Delta(E) \times R_+$. Conversely, if this vanishes for all disjoint A and B, M is evidently orthogonal. Q.E.D.

STOCHASTIC INTEGRALS

We are only going to do the L^2-theory here - the bare bones, so to speak. It is possible to extend our integrals further, but since we won't need the extensions in this course, we will leave them to our readers.

Let M be a worthy martingale measure on the Lusin space $(E,\underline{\underline{E}})$, and let Q_M and K_M be its covariation and dominating measures respectively. Our definition of the stochastic integral may look unfamiliar at first, but we are merely following Ito's construction in a different setting.

In the classical case, one constructs the stochastic integral as a process rather than as a random variable. That is, one constructs $\{\int_0^t f\, dB,\ t \geq 0\}$ simultaneously for all t; one can then say that the integral is a martingale, for instance. The analogue of "martingale" in our setting is "martingale measure". Accordingly, we will define our stochastic integral as a martingale measure.

Recall that we are restricting ourselves to a finite time interval $[0,T]$ and to one of the E_n, so that M is finite. As usual, we will first define the integral for elementary functions, then for simple functions, and then for all functions in a certain class by a functional completion argument.

<u>DEFINITION</u>. A function $f(x,s,\omega)$ is <u>elementary</u> if it is of the form

(2.5) $$f(x,s,\omega) = X(\omega) I_{(a,b]}(s)\, I_A(x),$$

where $0 \leq a < t$, X is bounded and $\underline{\underline{F}}_a$-measurable, and $A \in \underline{\underline{E}}$. f is <u>simple</u> if it is a finite sum of elementary functions. We denote the class of simple functions by $\underline{\underline{S}}$.

<u>DEFINITION</u>. The <u>predictable σ-field</u> $\underline{\underline{P}}$ on $\Omega \times E \times R_+$ is the σ-field generated by $\underline{\underline{S}}$. A function is <u>predictable</u> if it is $\underline{\underline{P}}$-measurable.

We define a norm $\|\ \|_M$ on the predictable functions by

$$\|f\|_M = E\{(|f|,\ |f|)_K\}^{1/2}.$$

Note that we have used the absolute value of f to define $||f||_M$, so that

$$(f,f)_Q \le ||f||_M^2.$$

Let $\underline{\underline{P}}_M$ be the class of all predictable f for which $||f||_M < \infty$.

PROPOSITION 2.2. Let $f \in \underline{\underline{P}}_M$ and let $A = \{(x,s): |f(x,s)| \ge \varepsilon\}$.

Then

$$E\{K(A \times E \times [0,T])\} \le \frac{1}{\varepsilon} ||f||_M \, E\{K(E \times E \times [0,T])\}$$

PROOF.

$$\begin{aligned} \varepsilon\, E\{K(A \times E \times [0,T])\} &\le E\{\int |f(x,t)|K(dx\ dy\ dt)\} \\ &= E\{(|f|,\ 1)_K\} \\ &\le E\{(|f|,\ |f|)_K^{1/2} K(E \times E \times [0,T])\} \\ &\le ||f||_M \, E\{K(E \times E \times [0,T])\}^{1/2} \end{aligned}$$

where we have used Schwartz' inequality in two forms (see Exercise 2.4).

Q.E.D.

Exercise 2.5. Use Proposition 2.2 to show $\underline{\underline{P}}_M$ is complete, and hence a Banach space.

PROPOSITION 2.3. $\underline{\underline{S}}$ is dense in $\underline{\underline{P}}_M$.

PROOF. If $f \in \underline{\underline{P}}_M$, let $f_N(x,s) = \begin{cases} f(x,s) & \text{if } |f(x,s)| < N \\ 0 & \text{otherwise} \end{cases}$.

Then $||f-f_N||_M^2 = E\{\int |f(x,s) - f_N(x,s)|\ |f(y,s) - f_N(y,s)|K(dxdyds)\}$ which goes to zero by monotone convergence. Thus the bounded functions are dense. If f is a bounded step function, i.e. if there exist $0 \le t_0 < t_1 < \ldots < t_n$ such that $t \to f(x,t)$ is constant on each $(t_j,\ t_{j+1}]$, then f can be uniformly approximated by simple functions. Thus the simple functions are dense in the step functions. It remains to show that the step functions are dense in the bounded functions.

To simplify our notation, let us suppose that $K(E \times E \times ds)$ is absolutely continuous with respect to Lebesgue measure. [We can always make a preliminary time change to assure this.] If $f(x,s,\omega)$ is bounded and predictable, set

$$f_n(x,s,\omega) = 2^{-n} \int_{(k-1)2^{-n}}^{k2^{-n}} f(x,u,\omega)du \quad \text{if } k2^{-n} \leq s < (k+1)2^{-n}.$$

Fix ω and x. Then $f_n(x,s,\omega) \to f(x,s,\omega)$ for a.e. s by either the martingale convergence theorem or Lebesgue's differentiation theorem, take your choice. It follows easily that $\|f - f_n\|_M \to 0$. Q.E.D.

We can now construct the integral with a minimum of interruption. If $f(x,s,\omega) = X(\omega)\, I_{(a,b]}(s)\, I_A(x)$ is an elementary function, define a martingale measure $f\cdot M$ by

(2.6) $$f\cdot M_t(B) = X(\omega)(M_{t\wedge b}(A \cap B) - M_{t\wedge a}(A \cap B)).$$

<u>LEMMA 2.4</u>. $f\cdot M$ is a worthy martingale measure. Its covariance and dominating measures $Q_{f\cdot M}$ and $K_{f\cdot M}$ are given by

(2.7) $$Q_{f\cdot M}(dx\ dy\ ds) = f(x,s)\ f(y,s)\ Q_M(dx\ dy\ ds)$$

(2.8) $$K_{f\cdot M}(dx\ dy\ dx) = |f(x,s)f(y,s)|K_M(dx\ dy\ ds).$$

Moreover

(2.9) $$E\{f\cdot M_t(B)^2\} \leq \|f\|_M^2 \quad \text{for all } B \in \underline{\underline{E}},\ t \leq T.$$

<u>PROOF</u>. $f\cdot M_t(B)$ is adapted since $X \in \underline{\underline{F}}_a$; it is square integrable, and a martingale. $B \to f\cdot M_t(B)$ is countably additive (in L^2), which is clear from (2.6). Moreover

$$f\cdot M_t(B) f\cdot M_t(C) - \int_{B\times C\times[0,t]} f(x,s)f(y,s)Q_M(dx\ dy\ ds)$$

$$= X^2[(M_{t\wedge b}(A \cap B) - M_{t\wedge a}(A \cap B)(M_{t\wedge b}(A \cap C) - M_{t\wedge a}(A \cap C))$$

$$- \langle M(A \cap B), M(A \cap C)\rangle_{t\wedge b} + \langle M(A \cap B), M(A \cap C)\rangle_{t\wedge a}]$$

which is a martingale. This proves (2.7), and (2.8) follows immediately since $K_{f\cdot M}$ is positive and positive definite. (2.9) then follows easily.

Q.E.D.

We now define $f\cdot M$ for $f \in \underline{\underline{S}}$ by linearity.

<u>Exercise 2.6</u>. Show that (2.7)-(2.9) hold for $f \in \underline{\underline{S}}$.

Suppose now that $f \in \underline{\underline{P}}_M$. By Prop. 2.6 there exist $f_n \in \underline{\underline{S}}$ such that

$\|f-f_n\|_M \to 0$. By (2.9), if $A \in \underline{\underline{E}}$ and $t \leq T$,

$$E\{(f_m \cdot M_t(A) - f_n \cdot M_t(A))^2\} \leq \|f_m - f_n\|_M \to 0$$

as $m, n \to \infty$. It follows that $(f_n \cdot M_t(A))$ is Cauchy in $L^2(\Omega, \underline{\underline{F}}, P)$, so that it converges in L^2 to a martingale which we shall call $f \cdot M_t(A)$. The limit is independent of the sequence (f_n).

<u>THEOREM 2.5</u>. If $f \in \underline{\underline{P}}_M$, then $f \cdot M$ is a worthy martingale measure. It is orthogonal if M is. Its covariance and dominating measures respectively are given by

(2.10) $$Q_{f \cdot M}(dx\ dy\ ds) = f(x,s)f(y,s)\ Q_M(dx\ dy\ ds);$$

(2.11) $$K_{f \cdot M}(dx\ dy\ ds) = |f(x,s)\ f(y,s)|K_M(dx\ dy\ ds).$$

Moreover, if $g \in \underline{\underline{P}}_M$ and $A, B \in \underline{\underline{E}}$, then

(2.12) $$\langle f \cdot M(A),\ g \cdot M(B)\rangle_t = \int_{A \times B \times [0,t]} f(x,s)f(y,s)Q_M(dx\ dy\ ds);$$

(2.13) $$E\{(f \cdot M_t(A))^2\} \leq \|f\|_M^2.$$

<u>PROOF</u>. $f \cdot M(A)$ is the L^2 limit of the martingales $f_n \cdot M(A)$, and is hence a square-integrable martingale. For each n

(2.14) $$f_n \cdot M_t(A)\ f_n \cdot M_t(B) - \int_{A \times B \times [0,t]} f_n(x,s)f_n(y,s)\ Q_M(dx\ dy\ ds)$$

is a martingale. $f_n \cdot M(A)$ and $f_n \cdot M(B)$ each converge in L^2, hence their product converges in L^1. Moreover

$$E\{\Big|\int_{A \times B \times [0,t]} (f_n(x,s)f_n(y,s) - f(x,s)f(y,s))Q_M(dx\ dy\ ds)\Big|\}$$

$$\leq E\{\int_{E \times E \times [0,T]} |f_n(x)\,||f_n(y)-f(y)|K_M(dx\ dy\ ds)\}$$

$$+ E\{\int_{E \times E \times [0,T]} |f_n(x)-f(x)|\ \ |f(y)|K_M(dx\ dy\ ds)\}$$

$$\leq E\{(|f_n|,\ |f-f_n|)_K + (|f-f_n|,\ |f|)_K\}$$

By Schwartz:

$$\leq (\|f_n\|_M + \|f\|_M)\|f_n - f\|_M \to 0\ .$$

Thus the expression (2.14) converges in L^1 to

$$f \cdot M_t(A)\ f \cdot M_t(B) - \int_{A \times B \times [0,t]} f(x,s)f(y,s)\ Q_M(dx\ dy\ ds)$$

which is therefore a martingale. The latter integral, being predictable, must therefore equal $\langle f\cdot M(A), f\cdot M(B)\rangle_t$, which verifies (2.10), and (2.11) follows.

This proves (2.12) in case $g = f$, and the general case follows by polarization. (2.13) then follows from (2.11).

To see that $f\cdot M_t$ is a martingale measure, we must check countable additivity. If $A_n \subset E$, $A_n \downarrow \phi$, then

$$E\{f\cdot M_t(A_n)^2\} \le E\{\int_{A_n\times A_n\times[0,t]} |f(x,s)f(y,s)|K(dx\ dy\ ds)\}$$

which goes to zero by monotone convergence.

If M is orthogonal, Q_M sits on $\Delta(E) \times [0,T]$, hence, by (2.10), so does $Q_{f\cdot M}$. By Proposition 2.4, $f\cdot M$ is orthogonal. Q.E.D.

Now that the stochastic integral is defined as a martingale measure, we define the usual stochastic integrals by

$$\int_{A\times[0,t]} f\ dM = f\cdot M_t(A)$$

and

$$\int_{E\times[0,t]} f\ dM = f\cdot M_t(E).$$

while

$$\int f\ dM = \lim_{t\to\infty} f\cdot M_t(E).$$

When it is necessary we will indicate the variables of integration. For instance

$$\int_{A\times[0,t]} f(x,s)M(dx\ ds) \quad \text{and} \quad \int_0^t \int_A f(x,s)dM_{xs}$$

both denote $f\cdot M_t(A)$.

It is frequently necessary to change the order of integration in iterated stochastic integrals. Here is a form of stochastic Fubini's theorem which will be useful.

Let $(G, \underline{\underline{G}}, \mu)$ be a finite measure space and let M be a martingale with dominating measure K.

<u>THEOREM 2.6</u>. Let $f(x,s,\omega,\lambda)$, $x \in E$, $s \ge 0$, $\omega \in \Omega$, $\lambda \in G$ be a $\underline{\underline{P}} \times \underline{\underline{G}}$-measurable function. Suppose that

(2.15) $$E\{\int_{E\times E\times[0,T]\times G} |f(x,s,\omega,\lambda)\ f(y,s,\omega,\lambda)|K(dx\ dy\ ds)\mu(d\lambda)\} < \infty.$$

Then

(2.16) $\int_G \Big[\int_{E\times[0,t]} f(x,s,\lambda)\, M(dxds)\Big]\mu(d\lambda) = \int_{E\times[0,t]} \Big[\int_G f(x,s,\lambda)\mu(d\lambda)\Big] M(dxds).$

PROOF. If $f(x,s,\omega,\lambda) = X(\omega)\, I_{(a,b]}(s)\, I_A(x) g(\lambda)$, then both sides of (2.16) equal

$$X\big(M_{t\wedge b}(A) - M_{t\wedge a}(A)\big) \int g(\lambda)\mu(d\lambda).$$

Both sides of (2.16) are additive in f, so this also holds for finite sums of such f. If f is $\underline{\underline{P}} \times \underline{\underline{G}}$ - measurable and satisfies (2.16), we can apply an argument similar to the proof of Proposition 2.6 to show that there exists a sequence (f_n) of such functions such that

$$E\{ \int |f(x,s,\lambda) - f_n(x,s,\lambda)||f(y,s,\lambda) - f_n(y,s,\lambda)| K(dx\ dy\ ds)\ \mu(d\lambda)\}$$
$$= \int \|f(\lambda) - f_n(\lambda)\|_M^2\ \mu(d\lambda) \to 0 .$$

We see that the integral in brackets on the right hand side of (2.16) is $\underline{\underline{P}}$-measurable, (Fubini) so that the integral makes sense, providing that $\| \int f(\lambda)\ \mu(d\lambda) \|_M < \infty$.

On the left-hand side we can take a subsequence if necessary to have $\|f(\lambda) - f_n(\lambda)\|_M \to 0$ for η - a.e.λ. This implies that for a.e.λ $\int f_n(\lambda)dM \to \int f(\lambda)dM$ in L^2, hence in measure. Using Fubini's theorem again we see that $\int f_n(\omega,\lambda)dM \to \int f(\omega,\lambda)dM$ in $P \times \mu$-measure, hence the latter integral is measurable in the pair (ω,λ). It follows that $\int f(\lambda)dM$ is μ-measurable for fixed ω, so that the integral on the left-hand side of (2.16) makes sense.

We must show that both sides converge as $n \to \infty$. Set $g_n = f - f_n$.

$$\| \int g_n(\lambda)\mu(d\lambda) \|_M = E\{ \int_{E\times E\times[0,T]} \int_G g_n(x,s,\lambda)\mu(d\lambda)K(dxdyds) \int_G g_n(y,s,\lambda')\mu(d\lambda')\}$$
$$= \int_{G\times G} E\{(g_n(\lambda),\ g_n(\lambda'))_K\}\mu(d\lambda)\mu(d\lambda');$$

by Schwartz, this is

$$\le \int_{G\times G} E\{(|g_n(\lambda)|,\ |g_n(\lambda)|)_K\}^{1/2} E\{(|g_n(\lambda')|,\ |g_n(\lambda')|)_K\}^{1/2}\mu(d\lambda)\mu(d\lambda')$$
$$= \Big(\int_G \|g_n(\lambda)\|_M \mu(d\lambda)\Big)^2$$
$$\le \mu(G) \int_G \|f(\lambda) - f_n(\lambda)\|_M^2 d\lambda$$

which tends to zero by (2.15).

This implies that the right-hand side of (2.16) converges. On the left,

$$E\Big\{ \int_G \Big(\int g_n(x,s,\lambda)M(dxds)\Big)^2 \mu(d\lambda)\Big\}$$

$$= \int_G E\Big\{\Big(\int g_n(x,s,\lambda)M(dxds)\Big)^2\Big\}\mu(d\lambda)$$

$$\leq \int_G \|g_n(\lambda)\|_M^2 \,\mu(d\lambda) \to 0.$$

By choosing a subsequence if necessary, we see that for a.e. ω, $\int_G \Big(\int g_n(x,s,\lambda)M(dxds)\Big)^2 \mu(d\lambda) \to 0$, hence $\int_G (\int f - f_n)dM)d\mu \to 0$, and the left hand side of (2.16) converges too. Q.E.D.

ORTHOGONAL MEASURES

The remainder of this chapter concerns special properties of measures which are orthogonal or have nuclear covariance. We must certify their worthiness, so that the foregoing integration theory applies.

We should admit here that although we are handling a wide class of martingale measures in this chapter, our main interest is really in orthogonal measures. This is not because the theory is simpler - it is only simpler at the beginning - but because the problems which motivated this study involved white noises and related orthogonal measures.

The theory of integration does simplify, at least initially, if the integrator is orthogonal. For instance, the covariance measure Q sits on the diagonal and is positive, so that $Q = K$. Instead of having two measures on $E \times E \times R_+$, we need only concern ourselves with a single measure ν on $E \times R_+$ where $\nu(A \times [0,t]) = Q(A \times A \times [0,t])$, and this leads to several rather pleasant consequences which we will detail below.

The proof that an orthogonal measure M is worthy comes down to finding a good version of the increasing process $\langle M(A)\rangle_t$, one which is a measure in A and is right continuous in t.

We will fix our attention on a fixed time interval $0 \leq t \leq T$, and we continue to assume that $E = E_n$, so that M is finite. Define

$$\mu(A) = E\{M_T(A)^2\} = E\{\langle M(A)\rangle_T\}.$$

1° $\langle M(\cdot)\rangle_t$ is an additive set function,

$$\text{i.e. } A \cap B = \phi \Rightarrow \langle M(A)\rangle_t + \langle M(B)\rangle_t = \langle M(A \cup B)\rangle_t \text{ a.s.}$$

Indeed, $\langle M(A \cup B)\rangle_t = \langle M(A) + M(B)\rangle_t$

$$= \langle M(A)\rangle_t + \langle M(B)\rangle_t + 2\langle M(A), M(B)\rangle_t,$$

and the last term vanishes since M is orthogonal.

2° $A \subset B \Rightarrow \langle M(A)\rangle_t \le \langle M(B)\rangle_t$

$s \le t \Rightarrow \langle M(A)\rangle_s \le \langle M(B)\rangle_s$

3° μ is a σ-finite measure: it must be σ-finite since M_T is, and additivity follows by taking expectations in 1°.

The increasing process $\langle M(A)\rangle_t$ is finitely additive for each t by 1°, but it is better than that. It is possible to construct a version which is a measure in A for each t.

<u>THEOREM 2.7</u>. Let $\{M_t(A), \underline{\underline{F}}_t, 0 \le t \le T, A \in \underline{\underline{E}}\}$ be an orthogonal martingale measure. Then there exists a family $\{\nu_t(\cdot), 0 \le t \le T\}$ of random σ-finite measures on $(E,\underline{\underline{E}})$ such that

(i) $\{\nu_t, 0 \le t \le T\}$ is predictable;

(ii) for all $A \in \underline{\underline{E}}$, $t \to \nu_t(A)$ is right-continuous and increasing;

(iii) $P\{\nu_t(A) = \langle M(A)\rangle_t\} = 1$ all $t \ge 0$, $A \in \underline{\underline{E}}$.

<u>PROOF</u>. We can reduce this to the case $E \subset \mathbf{R}$, for E is homeomorphic to a Borel set $F \subset \mathbf{R}$. Let $h: E \to F$ be the homeomorphism, and define $\bar{M}_t(A) = M_t(h^{-1}(A))$, $\bar{\mu}(A) = \mu(h^{-1}(A))$. If we find a $\bar{\nu}_t$ satisfying the conclusions of the theorem and if $\bar{\nu}_t(\mathbf{R} - F) = 0$, then $\nu_t = \bar{\nu}_t \circ h$ satisfies the theorem. Thus we may assume E is a Borel subset of $\mathbf{R}$.

Since M is σ-finite, there exist $E_n \uparrow E$ for which $\mu(E_n) < \infty$. Then there are compact $K_n \subset E_n$ such that $u(E_n - K_n) < 2^{-n}$. We may also assume $K_n \subset K_{n+1}$ all n. It is then enough to prove the theorem for each K_n. Thus we may assume E is compact in $\mathbf{R}$ and $\mu(E) < \infty$.

Define $F_t(x) = \langle M(-\infty,x]\rangle_t$, $-\infty < x < \infty$.

Then

a) $x \leq x' \Rightarrow F_t(x) \leq F_t(x')$;

$t \leq t' \Rightarrow F_t(x') - F_t(x) \leq F_{t'}(x') - F_{t'}(x)$

b) $$E\{ \sup_{\substack{t\leq T \\ x_1\leq x\leq x_2 \\ x \in Q}} |F_t(x) - F_t(x_1)|\} \leq \mu((x_1,x_2])$$

Indeed (a) follows from 2°. To see (b), note that for fixed $t \leq T$, $F_t(x) - F_t(x_1) \leq \langle M(x_1,x_2]\rangle_t \leq \langle M(x_1,x_2]\rangle_T$ a.s. by 2°. By right continuity this holds simultaneously for all $t \leq T$ and all rational x in $(x_1,x_2]$. But then (b) follows since $E\{\langle M(x_1,x_2]\rangle_T\} = \mu((x_1,x_2])$.

Define $\overline{F}_t(x) = \inf\{F_{t'}(x'):x' > x,\ t' > t,\ x', t' \in Q\}$. This will be the "good" version of F. We claim that $\overline{F}_t$ is the distribution function of a measure.

c) $t_1 \leq t_2$ and $x_1 \leq x_2 \Rightarrow \overline{F}_{t_1}(x_1) \leq \overline{F}_{t_2}(x_2)$;

d) $\overline{F}_t(x)$ is right continuous in the pair (x,t);

e) for fixed x, $P\{\overline{F}_t(x) = F_t(x),\ \text{all } t \leq T\} = 1$.

Indeed, (c) is clear and (d) and (e) follow from the uniform convergence guaranteed by (b). To see (e), for instance, choose rational t_n and x_n which strictly decrease to t and x respectively. Then

$$F_t(x) < \overline{F}_t(x) \leq F_{t_n}(x_n) = F_{t_n}(x) + \left(F_{t_n}(x_n) - F_{t_n}(x)\right).$$

But $F_{t_n}(x) \to F_t(x)$ by right continuity and the term in square brackets tends to zero in probability by (b).

Let ν_t be the distribution on R generated by the distribution function $\overline{F}_t$. Note that ν_t does not charge R-E, for E is compact; and if $(a,b] \subset R - E$, $\nu_t(a,b] = \overline{F}_t(b) - \overline{F}_t(a) = F_t(b) - F_t(a) \leq F_T(b) - F_T(a)$ by (e). This is true simultaneously for all rational t. Since $\overline{F}$ is right continuous we have a.s.

$$0 \leq \sup_{t\leq T} \nu_t(a,b] \leq F_T(b) - F_T(a)$$

and the latter has expectation $\mu\{(a,b]\} = 0$.

Note that $\{\nu_t\colon 0 \leq t \leq T\}$ is predictable, for it is determined by $\overline{F}_t(x)$, $x \in Q$, hence by $F_t(x)$, $x \in Q$ by (e), and the F_t are predictable.

If $t < t'$, $\bar{F}_{t'}(x) - \bar{F}_t(x)$ is a distribution function, so $t \to \nu_t(A)$ is increasing for each A. It is right continuous in t if $A = (0,x]$, some x, or if $A = \mathbf{R}$ (for $E \subset \mathbf{R}$ is compact). Then right continuity for all Borel A follows by a monotone class argument.

To show (iii), note that if $A = (-\infty,x]$,

$$M_t^2(A) - \nu_t(A) \text{ is a martingale,} \tag{2.17}$$

for then $\nu_t(A) = F_t(x) = \langle M(A)\rangle_t$. Let $\underline{\underline{G}}$ be the class of A for which (2.17) holds. $\underline{\underline{G}}$ must contain finite unions of intervals of the form (a,b], and it contains $\mathbf{R}$, for $\nu_t(-\infty,x] = \nu_t(\mathbf{R})$ for large x.
It is closed under complementation, for

$$M_t^2(A^c) - \nu_t(A^c) = M_t^2(\mathbf{R}) - \nu_t(\mathbf{R}) - \left(M_t^2(A) - \nu_t(A)\right) - 2\, M_t(A)M_t(A^c)$$

and each of the terms on the right hand side is a martingale if $A \in \underline{\underline{G}}$. $\underline{\underline{G}}$ is also closed under monotone convergence. If $A_n \uparrow A$, for instance, $M_t(A_n)$ converges in L^2 to $M_t(A)$, and $\nu_t(A_n)$ increases to $\nu_t(A)$, hence the martingale $M_t^2(A_n) - \nu_t(A_n)$ converges in L^1 to $M_t^2(A) - \nu_t(A)$. The latter must therefore be a martingale. The case where $A_n \downarrow A$ follows by complementation. Thus $\underline{\underline{G}}$ contains all Borel sets. Q.E.D.

Now $t \to \nu_t(A)$ is increasing, so that we can define a measure ν on $E \times \mathbf{R}_+$ by defining $\nu(A \times (0,t]) = \nu_t(A)$ and extending it to $\underline{\underline{E}} \times \underline{\underline{B}}$, where $\underline{\underline{B}}$ is the class of Borel subsets of $\mathbf{R}_+$. This gives us the following.

<u>COROLLARY 2.8</u>. Let M be an orthogonal martingale measure. Then there exists a random σ-finite measure $\nu(dxds)$ on $E \times \mathbf{R}_+$ such that $\nu_t(A) = \nu(A \times [0,t])$ for all $A \in \underline{\underline{E}}$, $t > 0$.

We can get the covariance measure Q of M directly from ν. Set $\Delta = \Delta(E) \times \mathbf{R}_+$ where $\Delta(E)$ is the diagonal of $E \times E$ and let $\pi: \Delta \to E \times \mathbf{R}_+$ be defined by $\pi(x,x,t) = (x,t)$. Then we define Q by

$$Q(\Lambda) = \nu(\pi(\Lambda \cap \Delta)), \quad \Lambda \in \underline{\underline{E}} \times \underline{\underline{E}} \times \underline{\underline{B}}.$$

Then

$$\begin{aligned} Q(A \times B \times [0,t]) &= \nu(A \cap B \times [0,t]) \\ &= \langle M(A \cap B)\rangle_t \\ &= \langle M(A), M(B)\rangle_t, \end{aligned}$$

so Q is indeed the covariance measure of M. Since Q is positive and positive definite, we can set K = Q and we have:

COROLLARY 2.9. An orthogonal martingale measure is worthy.

We noted above that a white noise gives rise to an orthogonal martingale measure. It is easy to characterize white noises among martingale measures.

PROPOSITION 2.10. Let M be an orthogonal martingale measure, and suppose that for each $A \in \underline{\underline{E}}$, $t \to M_t(A)$ is continuous. Then M is a white noise if and only if its covariance measure is deterministic.

PROOF. Rather than use Q, let us use the measure ν of Corollary 2.8, which is equivalent.

If M is a white noise on $E \times \mathbb{R}_+$ based on a measure μ, it is easy to see that $\nu = \mu$, so ν is deterministic.

Conversely, if M is orthogonal and if ν is deterministic, then for $B \in \underline{\underline{E}}$, both $M_t(B)$ and $M_t^2(B) - \nu(B \times [0,t])$ are martingales.

To show M is a white noise, we must show it gives disjoint sets independent Gaussian values. One can see it is sufficient to show the following: if $B_1,\ldots,B_n$ are disjoint sets, then $\{M_t(B_1),\ t \geq 0\},\ldots,$ $\{M_t(B_n),\ t \geq 0\}$ are independent mean zero Gaussian processes with independent increments. This reduces to the following calculation.

Let

$$N_s = \exp\{i \sum_{j=1}^{n} \lambda_j (M_{t+s}(B_j) - M_s(B_j))\} .$$

By Ito's formula

$$N_s = 1 + \sum_{j=1}^{n} \int_t^{t+s} i\,\lambda_j\, N_u\, dM_u(B_j) - \frac{1}{2} \sum_{j=1}^{n} \lambda_j^2 \int_t^{t+s} N_u \nu(B_j \times du)$$

where we have used the fact that $d\langle M(B_j), M(B_k)\rangle_t$ vanishes by orthogonality if $j \neq k$. Let $f(x) = E\{N_s | \underline{\underline{F}}_t\}$ and note that

$$f(s) = 1 - \frac{1}{2} \sum_{j=1}^{n} \lambda_j^2 \int_t^{t+s} f(u) d\nu(B_j \times du).$$

This has the unique solution

$$f(s) = \prod_{j=1}^{n} e^{-\frac{1}{2}\lambda_j^2 \nu(B_j \times (s,t])},$$

from which we see that the increments $M_{t+s}(B_j) - M_t(B_j)$ are independent of $\underline{\underline{F}}_t$ and of each other, and are Gaussian with mean zero and variance $\nu(B_j \times (s,t])$. Thus M is a white noise based on ν. Q.E.D.

NUCLEAR COVARIANCE

We will develop some of the particular properties of martingale measures with nuclear covariance. In particular, we will show they are worthy.

Suppose M is a martingale measure on $(E,\underline{\underline{E}})$ with nuclear covariance. Then there is a measure η and a CONS (ϕ_n) in $L^2(E,\underline{\underline{E}},\eta)$ such that

(2.18) $$\sum_n E\{M_T(\phi_n)^2\} < \infty.$$

We continue to assume $E = E_n$ for some n, so that M is finite, not just σ-finite.

PROPOSITION 2.11. For each x there exists a square-integrable martingale $\{M_t(x), t \geq 0\}$ such that for a.e. ω, $x \to M_t(x)$ $L^2(E, \underline{\underline{E}}, \eta)$. Moreover

(i) $M_t(A) = \int_A M_t(x)\eta(dx)$, $A \subset E$;

(ii) $Q_M(A \times B \times [0,t]) = \int_{A\times B} \langle M(x), M(y)\rangle_t \, \eta(dx)\eta(dy)$.

Furthermore, there exists a predictable increasing process C_t and a positive predictable function $\sigma(x,t)$ such that $E\{\int_0^T \sigma^2(x,s)dC_s\eta(dx)\} < \infty$ and

(iii) $K_M(A \times B \times [0,t]) = \int_{A\times B\times[0,t]} \sigma(x,s)\sigma(y,s)\eta(dx)\eta(dy)dC_s$.

Finally

(iv) $\sum_n E\{M_t(\phi_n)^2\} = E\{\int_E M_t(x)^2\eta(dx)\}$.

PROOF. We will be rather cavalier in handling null-sets and measurable versions below. We leave it to the reader to supply the details.

The map $\phi \to M_t(\phi)$ is linear, and since $\sum M_t^2(\phi_n) < \infty$ a.s., it is a bounded linear functional on $L^2(E, \underline{\underline{E}}, \eta)$. It thus corresponds to a function which we denote $M_t(x)$, such that for $\phi \in L^2$,

$$M_t(\phi) = \int M_t(x)\phi(x)\eta(dx).$$

In fact
$$M_t(x,\omega) = \sum M_t(\phi_n)\phi_n(x).$$

This series converges in L^2 for a.e. ω hence, taking a subsequence if necessary, it converges in $L^2(\Omega, \underline{\underline{F}}, P)$ for η-a.e. x. For each such x, $M_t(x)$ must be a martingale, being the L^2-limit of martingales. By modifying $M_t(x)$ on a set of η-measure zero, we can assume $M_t(x)$ is a martingale for all x.

Now
$$M_t(A)M_t(B) = \int_{A\times B} M_t(x)M_t(y)\,\eta(dx)\eta(dy)$$

so that

$$\langle M(A), M(B)\rangle_t = \int_{A\times B} \langle M(x), M(y)\rangle_t\, \eta(dx)\eta(dy).$$

This proves (i) and (ii).

$\underline{\underline{E}}$ is separable, so it is generated by a countable sub-algebra $\underline{\underline{A}}$. Let $\underline{\underline{M}}$ be the smallest class of martingales which contains $M_t(A)$, all $A \in \underline{\underline{A}}$, and which is closed under L^2-convergence. Then one can show that there exists an increasing process C_t such that for all $N \in \underline{\underline{M}}$, $d\langle N\rangle_t \ll dC_t$. Consequently, by Motoo's theorem, $\langle N\rangle_t = \int_0^t h(s)dC_s$ for some predictable h. This holds in particular for all the $M_t(\phi_k)$, for these are in $\underline{\underline{M}}$, and hence for the $M_t(x)$.

Furthermore, one can see by polarization that there exists a function $h(x,y,s)$ such that

$$\langle M(x), M(y)\rangle_t = \int_0^t h(x,y,s)dC_s.$$

But since $\langle M(x), M(y)\rangle \le \langle M(x)\rangle^{1/2} \langle M(y)\rangle^{1/2}$, we see that $|h(x,y,s)| \le h(x,x,s)^{1/2}\, h(y,y,s)^{1/2}$. Set $\sigma^2(x,s) = h(x,x,s)$. Then

$$Q_M(A \times B \times [0,t]) \le \int_{E\times E\times[0,t]} \sigma(x,s)\sigma(y,s)\eta(dx)\eta(dy)dC_s$$

which identifies $K(dx\, dy\, ds) = \sigma(x,s)\sigma(y,s)\eta(dx)\eta(dy)dC_s$. This is clearly positive and positive definite. To see that it is finite, write

$$E\{K(E \times E \times [0,T])\} = E\Big\{ \int_0^T \Big[\int_E \sigma(x,s)\eta(dx)\Big]^2 dC_s\Big\}$$

$$\le \eta(E)\, E\Big\{ \int_E \int_0^T \sigma^2(x,s)dC_s\eta(dx)\Big\}$$

$$= \eta(E)\, E\{\int_E M_t(x)^2\, \eta(dx)\} < \infty.$$

Finally, to see (iv), note that

$$\sum E\{M_t(\phi_n)^2\} = \sum E\{(\int M_t(x)\phi_n(x)\eta(dx))^2\}$$
$$= E\{\sum_n M_n^2(t)\}$$

where $M_n(t) = \int M_t(x)\phi_n(x)\eta(dx)$. By the Plancharel theorem, this is:

$$= E\{\int M_t(x)^2 \eta(dx)\} .$$

Q.E.D.

REMARK. Note from (iv) of the proposition that the sum in (2.18) does not depend on the particular CONS (ϕ_n).

Exercise 2.7. Suppose M has nuclear covariance on $L^2(E, \underline{\underline{E}}, \eta)$. Let f be predictable and set

$$k^2(x) = E\{\int_0^T f^2(x,s)\, d < M(x)>_s\}.$$

Show that if $\int k^2(x)\, \eta(dx) < \infty$, then

(i) $\|f\|_M < \infty$ (so $f\cdot M$ is defined);

(ii) $f\cdot M_t(x) \overset{\text{def}}{=} \int_0^t f(x,s)dM_s(x)$ exists as an Ito integral for η-a.e. x;

(iii) $f\cdot M_t(A) = \int_A f\cdot M_t(x)\, \eta(dx)$;

(iv) for any CONS (ϕ_n) in $L^2(E, \underline{\underline{E}}, \eta)$ and $t \leq T$,

$$\sum_n E\{f\cdot M_t(\phi_n)^2\} \leq \int_E k^2(x)\eta(dx).$$

so that $f\cdot M$ has nuclear covariance.

AN EXAMPLE

We will construct D. Bakry's example of a martingale measure which is not an integrator.

Let U be a random variable, uniformly distributed on [0,1]. Let $s_n = 1 - 1/n$, n = 1,2,... and define a filtration $(\underline{\underline{F}}_t)$ as follows.

$$\underline{\underline{F}}_{s_n} = \sigma\{[2^n U]\} \qquad ([n] = \text{greatest integer} \leq n)$$

$$\underline{\underline{F}}_t = \underline{\underline{F}}_{s_n} \quad \text{if} \quad s_n \leq t < s_{n+1}$$

and define

$$M_t(A) = P\{U \in A | \underline{\underline{F}}_t\}, \quad A \subset [0,1], \quad t \geq 0.$$

If $K = [2^n U]$, let $J_n = [K\, 2^{-n}, (K+1)2^{-n})$ and put $H_n = [K2^{-n}, (K+1/2)2^{-n})$ and $L_n = [(K+1/2)2^{-n}, (K+1)2^{-n})$. Then $J_n = H_n \cup L_n$ and all three are $\underline{\underline{F}}_{s_n}$-measurable. Note that $U \in J_n$ for all n.

Then M is a martingale measure and

(i) if $t < 1$, then $M_t(dx) = 2^n I_{J_n}(x)dx$;

(ii) if $t \geq 1$, then $M_t(A) = I_A(U)$.

If $t < 1$ and $s_n \leq t < s_{n+1}$, then J_n is $\underline{\underline{F}}_{s_n}$-measurable and the conditional distribution of U given $\underline{\underline{F}}_t = \underline{\underline{F}}_{s_n}$ is uniform on J_n, which implies (i), while if $t \geq 1$, U is $\underline{\underline{F}}_t$ measurable, which implies (ii).

Thus $M_t(\cdot)$ is a (real-valued) measure of total mass one, not just an L^2-measure. However there exist bounded predictable f for which $\int f(x,s)dM$ does not exist.

Set

$$f(x,t) = \begin{cases} 2 & \text{if } x \in H_n \text{ and } s_n < t \leq s_{n+1} \\ -2 & \text{if } x \in L_n \text{ and } s_n < t \leq s_{n+1} \\ 0 & \text{otherwise} \end{cases}$$

Then f is adapted and $t \to f(x,t)$ is left continuous, so f is predictable. $t \to M_t$ is constant on each $[s_n, s_{n+1})$, and it jumps at each s_n. f is a sum of simple functions so, if $f \cdot M$ exists, we can compute it directly.

$$\int_{[0,1]\times(s_n,s_{n+1}]} f(x,t)dM = 2(M_{s_{n+1}}(H_n) - M_{s_n}(H_n)) - 2(M_{s_{n+1}}(L_n) - M_{s_n}(L_n)) .$$

Now by (i) this is

$$= I_{\{J_{n+1} = H_n\}} - I_{\{J_{n+1} = L_n\}}.$$

Since J_{n+1} is either H_n or L_n, with probability 1/2 each, the above is either 1 or -1, with probability 1/2 each. But now if $\int_{[0,1]\times(0,1)} f\, dM$ exists, it must equal $\sum_n \int_{[0,1]\times(s_n,s_{n+1}]} f\, dM$ which diverges since each term has absolute value 1.

Evidently the dominating measure K does not exist. To see why, let us calculate the covariance measure Q. Let $N_t = M_t(dx)M_t(dy)$. For $t = s_{n+1}$,

$$Q(dx, dy, \{s_{n+1}\}) = \Delta\langle M(dx), M(dy)\rangle_{s_{n+1}} = N_{s_{n+1}} - E\{N_{s_{n+1}} | \underline{\underline{F}}_{s_n}\}$$

$$= [4^n I_{\{x,y \in J_{n+1}\}} - 4^{n+1} P\{x,y \in J_{n+1} | \underline{\underline{F}}_{s_n}\}]dxdy.$$

If x and y are both in H_n or both in L_n, $P\{x,y \in J_{n+1} | \underline{\underline{F}}_{s_n}\} = 1/2$ by (i). If one is in H_n and one in L_n, they can't both be J_{n+1}, (which equals H_n or L_n) and the conditional expectation vanishes. Thus it is

$$= 4^n[I_{\{x,y \in J_{n+1}\}} - 2I_{\{x,y \in H_n\}} - 2I_{\{x,y \in L_n\}}]dxdy.$$

The term in brackets is ±1 on the set $\{x,y \in J_n\}$ and zero off. Thus

$$|Q(dx\ dy \times \{s_{n+1}\})| = 4^n I_{\{x,y \in J_n\}} dxdy$$

and hence

$$|Q([0,1] \times [0,1] \times \{s_{n+1}\})| = 4^n \int_0^1 \int_0^1 I_{J_n}(x) I_{J_n}(y) dxdy \ .$$
$$= 1$$

Thus, if K exists, it must dominate Q and

$$K([0,1] \times [0,1] \times (0,1)) \geq \sum_n 1 = \infty.$$

CHAPTER THREE

EQUATIONS IN ONE SPACE DIMENSION

We are going to look at stochastic partial differential equations driven by white noise and similar processes. The solutions will be functions of the variables x and t, where t is the time variable and x is the space variable. There turns out to be a big difference between the case where x is one-dimensional and the case where $x \in \mathbf{R}^d$, $d \geq 2$. In the former case the solutions are typically, though not invariably, real-valued functions. They will be non-differentiable, but are usually continuous. On the other hand, in $\mathbf{R}^d$, the solutions are no longer functions, but are only generalized functions.

We will need some knowledge of Schwartz distributions to handle the case $d \geq 2$, but we can treat some examples in one dimension by hand, so to speak. We will do that in this chapter, and give a somewhat more general treatment later, when we treat the case $d \geq 2$.

THE WAVE EQUATION

Let us return to the wave equation of Chapter one:

$$(3.1)\qquad \begin{cases} \dfrac{\partial^2 V}{\partial t^2} = \dfrac{\partial^2 V}{\partial x^2} + \dot{W}, & t > 0,\ x \in \mathbf{R}; \\ V(x,0) = 0, & x \in \mathbf{R} \\ \dfrac{\partial V}{\partial t}(x,0) = 0, & x \in \mathbf{R}. \end{cases}$$

White noise is so rough that (3.1) has no solution: any candidate for a solution will not be differentiable. However, we can rewrite it as an integral equation which will be solvable. This is called a weak form of the equation.

We first multiply by a C^∞ function $\phi(x,t)$ of compact support and integrate over $\mathbf{R} \times [0,T]$, where $T > 0$ is fixed. Assume for the sake of argument that $V \in C^{(2)}$.

$$\int_0^T \int_{\mathbf{R}} \left[V_{tt}(x,t) - V_{xx}(x,t)\right]\phi(x,t)\,dx\,dt = \int_0^T \int_{\mathbf{R}} \phi(x,t)\dot{W}(x,t)\,dx\,dt.$$

Integrate by parts twice on the left-hand side. Now ϕ is of compact support in x, but it may not vanish at t = 0 and t = T, so we will get some boundary terms:

$$\int_0^T \int_R V(x,t)[\phi_{tt}(x,t)-\phi_{xx}(x,t)]\,dxdt + \int_R [\phi(x,\cdot)V_t(x,\cdot)\Big|_0^T-\phi_t(x,\cdot)V(x,\cdot)\Big|_0^T]\,dx$$

$$= \int_0^T \int_R \phi(x,t)\dot{W}(x,t)\,dxdt.$$

If $\phi(x,T) = \phi_t(x,T) = 0$, the boundary terms will drop out because of the initial conditions. This leads us to the following.

DEFINITION. We say that V is a weak solution of (3.1) providing that V(x,t) is locally integrable and that for all T > 0 and all C^∞ functions $\phi(x,t)$ of compact support for which $\phi(x,T) = \phi_t(x,T) = 0, \forall x$, we have

$$\int_0^T \int_R V(x,t)[\phi_{tt}(x,t) - \phi_{xx}(x,t)]\,dxdt = \int_0^T \int_R \phi\,dW. \tag{3.2}$$

The above argument is a little unsatisfying; it indicates that if V satisfies (3.1) in some sense, it should satisfy (3.2), while it is really the converse we want. We leave it as an exercise to verify that if we replace $\dot{W}$ by a smooth function f in (3.1) and (3.2), and if V satisfies (3.2) and is in $C^{(2)}$, then it does in fact satisfy (3.1).

THEOREM 3.1. There exists a unique continuous solution to (3.2), namely $V(x,t) = \frac{1}{2}\hat{W}\left(\frac{t-x}{\sqrt{2}}, \frac{t+x}{\sqrt{2}}\right)$, where $\hat{W}$ is the modified Brownian sheet of Chapter One.

PROOF. Uniqueness: if V_1 and V_2 are both continuous and satisfy (3.2), then their difference $U = V_2 - V_1$ satisfies

$$\iint U(x,t)[\phi_{tt}(x,t) - \phi_{xx}(x,t)]\,dxdt = 0$$

Let f(x,t) be a C^∞ function of compact support in $R \times (0,T)$. Notice that there exists a $\phi \in C^\infty$ with $\phi(x,T) = \phi_t(x,T) = 0$ such that $\phi_{tt} - \phi_{xx} = f$. Indeed, if $C(x,t;\, x_o,t_o)$ is the indicator function of the cone $\{(x,t):\, t<t_o, x_o+t-t_o< x < x_o+t_o-t\}$, then

$$\phi(x_o,t_o) = \int_R \int_0^T f(x,t)C(x,t;\, x_o,t_o)\,dxdt.$$

Thus

$$\iint U(x,t)\ f(x,t)\ dxdt = 0,$$

so $U = 0$ a.e., hence $U \equiv 0$ a.e.

To show existence, let us rotate coordinates by 45°. Let $u = (t-x)/\sqrt{2}$, $v = (t+x)/\sqrt{2}$ and set $\hat{\phi}(u,v) = \phi(x,t)$ and $\hat{W}(dudv) = W(dxdt)$. Note that $\phi_{tt} - \phi_{xx} = 2\hat{\phi}_{uv}$. Define $\hat{R}(u,v;u_o,v_o) = I_{\{u\leq u_o,v\leq v_o\}}$. The proposed solution is

$$\hat{V}(u,v) = \frac{1}{2}\iint_{\{u'+v'>0\}} \hat{R}(u',v';u,v)\hat{W}(du'dv').$$

Now V satisfies (3.2) iff the following vanishes:

$$(3.3)\quad \iint_{\{u'+v'>0\}} \Big[\iint_{\{u+v>0\}} \frac{1}{2}\hat{R}(u,v;u',v')\hat{W}(dudv)\Big]2\hat{\phi}_{uv}(u',v')du'dv' - \iint_{\{u+v>0\}} \hat{\phi}(u,v)\hat{W}(dudv).$$

We can interchange the order of integration by the stochastic Fubini's theorem of Chapter Two:

$$= \iint_{\{u+v>0\}} \Big[\int_v^\infty\int_u^\infty \hat{\phi}_{uv}(u',v')du'dv' - \hat{\phi}(u,v)\Big]\hat{W}(dudv).$$

But the term in brackets vanishes identically, for $\hat{\phi}$ has compact support.

QED

The literature of two-parameter processes contains studies of stochastic differential equations of the form

$$(3.4)\qquad d\hat{V}(u,v) = f(\hat{V})d\hat{W}(u,v) + g(\hat{V})dudv$$

where $\hat{V}$ and $\hat{W}$ are two parameter processes, and $d\hat{V}$ and $d\hat{W}$ represent two-dimensional increments, which we would write $\hat{V}(dudv)$ and $\hat{W}(dudv)$. These equations rotate into the non-linear wave equation

$$V_{tt}(x,t) = V_{xx}(x,t) + f(V)\dot{W}(x,t) + g(V)$$

in the region $\{(x,t): t > 0, -t<x<t\}$. One specifies Dirichlet boundary conditions. Because of the special nature of the region, it is enough to give V on the boundary; one does not have to give V_t as well.

We have of course just finished solving the linear case ($f\equiv 1$, $g\equiv 0$). We will not pursue this any further here but we will treat a non-linear parabolic equation in the next section.

AN EXAMPLE ARISING IN NEUROPHYSIOLOGY

Let us look at a particular parabolic SPDE. The general type of equation has many applications, but this particular example came up in connection with a study of neurons. These nerve cells are the building blocks of the nervous system, and operate by a mixture of chemical, biological and electrical properties, but in this particular mathematical oversimplifcation they are regarded as long, thin cylinders, which act much like electrical cables. If such a cylinder extends from 0 to L, and if we only keep track to the x coordinate, we let V(x,t) be the electrical potential at the point x and time t. This potential is governed by a system of non-linear PDE's, called the Hodgkin-Huxley equations, but in certain ranges of values of V, they are well-approximated by the cable equation $V_t = V_{xx} - V$. (The variables have been scaled to make the coefficients all equal to one.)

The surface of the neuron is covered with synapses, thru which it receives impulses of current. If the current arriving at (x,t) is F(x,t), the system will satisfy the inhomogeneous PDE

$$V_t = V_{xx} - V + F.$$

Even if the system is at rest, the odd random impulse will arrive, so that F will have a random component. The different synapses are more-or-less independent, and there are an immense number of them, so that one would expect the impulses to arrive according to a Poisson process. The impulses may be of different amplitudes and even of different signs (impulses can be either "excitatory" or "inhibitory").

We thus expect that F can be written as $F = \overline{F} + \Pi$, where $\overline{F}$ is deterministic and Π is a compound Poisson process, centered so that it has mean zero. Since the equation is linear, its solution will be the sum of the solutions of the PDE $V_t = V_{xx} - V + \overline{F}$, and of the SPDE $V_t = V_{xx} - V + \Pi$. We can study the two separately, and, since the first is familiar, we will concentrate on the latter.

The impulses are generally small, and there are many of them, so that in fact Π is very nearly a white noise $\dot{W}$. This leads us to study the SPDE $V_t = V_{xx} - V + \dot{W}$.

One final remark. The response of the neuron to a current impulse may depend on the local potential, so that instead of $\dot{W}$, we have a term $f(V)\dot{W}$ in the above equation. f is often assumed to have the form $f(V) = V-V_o$, where V_o is a constant.

Let W be a white noise on a probability space $(\Omega,\underline{\underline{F}},P)$, let $(\underline{\underline{F}}_t)$ be a filtration such that W_t is adapted and such that, if $A \subset [t,\infty) \times R$, $W(A)$ is independent of $\underline{\underline{F}}_t$. Consider

$$(3.5)\qquad \begin{cases} \dfrac{\partial V}{\partial t} = \dfrac{\partial^2 V}{\partial x^2} - V + f(V,t)\dot{W}, & t > 0,\ 0 < x < L; \\ \dfrac{\partial V}{\partial x}(0,t) = \dfrac{\partial V}{\partial x}(L,t) = 0, & t > 0; \\ V(x,0) = V_o(x),\ t > 0. & \end{cases}$$

We assume that V_o is $\underline{\underline{F}}_o$-measurable and that $E\{V_o(x)^2\}$ is bounded, and that f satisfies a uniform Lipschitz condition, so that there exists a constant K such that

$$(3.5a)\qquad \begin{cases} |f(y,t) - f(x,t)| \le K|y-x|, \\ |f(y,t)| \le K(1+t)(1+|y|) \end{cases}$$

for all $x,y \in [0,L]$ and $t > 0$.

The homogeneous form of (3.5) is called the <u>cable equation</u>. We have specified reflecting boundaries for the sake of concreteness, but there is no great difficulty in treating other boundary conditions.

The Green's function for the cable equation can be gotten by the method of images. It is given by

$$G_t(x,y) = \frac{e^{-t}}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} \left[\exp\left(-\frac{(y-x-2nL)^2}{4t}\right) + \exp\left(-\frac{(y+x-2nL)^2}{4t}\right)\right].$$

We won't need to use this explicitly. We will just need the following facts, which can be seen directly:

$$(3.6)\qquad \int_0^L G_s(x,y)G_t(y,z)\,dy = G_{s+t}(x,z), \text{ and } G_t(x,y) = G_t(y,x);$$

for each $T > 0$ there is a constant C_T such that

$$(3.7)\qquad G_t(x,y) \le \frac{C_T}{\sqrt{t}} \exp\left(-\frac{|y-x|^2}{4t} - t\right).$$

Define $G_t(\phi,y) = \int_0^L G_t(x,y)\phi(x)\,dx$ for any function ϕ on $[0,L]$ for which the integral exists. Then $G_t(x,y)$ satisfies the homogeneous cable equation (i.e. it satisfies (3.5) with $f \equiv 0$) except at $t = 0$, and $G_o(\phi,y) = \phi(y)$. After integrating by parts, we have

(3.8) $$G_t(\phi,y) = \phi(y) + \int_0^t G_s(\phi''-\phi;y)ds$$

for all test functions ϕ for which $\phi'(0) = \phi'(L) = 0$.

Once again we pose the problem in a weak form.

Let $\phi \in C^\infty(R)$, with $\phi'(0) = \phi'(L) = 0$. Multiply (3.5) by $\phi(x)$ and integrate over both variables:

$$\int_0^L V(x,t)\phi(x)dx = \int_0^L V_o(x)\phi(x)dx + \int_0^t\int_0^L (\frac{\partial^2 V}{\partial x^2} - V)(x,s)\phi(x)dsdx$$
$$+ \int_0^t\int_0^L f(V(x,s),s)\phi(x)W(dxds).$$

Integrate by parts over x and use the boundary conditions on V and ϕ to get the following weak form of (3.5):

For each $\phi \in C^\infty(R^n)$ of compact support such that $\phi'(0) = \phi'(L) = 0$,

(3.9) $$\int_0^L (V(x,t)-V_o(x))\phi(x)dx = \int_0^t\int_0^L V(x,s)(\phi''(x)-\phi(x))dxds+\int_0^t\int_0^L f(V(x,s),s)\phi(x)W(dxds)$$

<u>Exercise 3.1</u>. (3.9) can be extended to smooth functions $\psi(x,t)$ of two variables which satisfy $\frac{\partial\psi}{\partial x}(0,t) = \frac{\partial\psi}{\partial x}(L,t) = 0$ for each t. Show that (3.9) implies that

$$\int_0^L \left[V(x,t)\ \psi(x,t) - V_o(x)\psi(x,0)\right]dx$$

(3.10) $$= \int_0^t\int_0^L V(x,s)(\frac{\partial^2\psi}{\partial x^2} - \psi + \frac{\partial\psi}{\partial t})(x,s)\ dxds$$
$$+ \int_0^t\int_0^L f(V(x,s),s)\psi(x,s)W(dx\ ds)\ .$$

<u>Exercise 3.2</u>. Show that (3.5) and (3.9) are equivalent if things are smooth, i.e. show that if V_0 and $\dot{W}$ are smooth functions and if $V \in C^{(2)}$, then (3.9) implies (3.5).

<u>THEOREM 3.2</u>. There exists a unique process $V = \{V(x,t),\ t\geq 0,\ 0<x<L\}$ which is L^2-bounded on $[0,L] \times [0,T]$ for any T and which satisfies (3.9) for all $t \geq 0$. If

$V_o(x)$ is bounded in L^p for some $p \geq 2$, then $V(x,t)$ is L^p-bounded on $[0,L]\times[0,T]$ for any T.

PROOF. We will suppress the dependence of f on s, and write f(x) rather than f(x,s).

Uniqueness: a solution of (3.9) must satisfy (3.10), so fix t and let $\phi(y,s) = G_{t-s}(\phi,y)$. Then $\phi(y,t) = \phi(y)$ and, by (3.8), $\phi_{xx} - \phi + \phi_s = 0$. Thus (3.10) becomes

$$\int_0^L V(x,t)\phi(x)dx = \int_0^L V_o(y)G_t(\phi,y)dy + \int_0^t\int_0^L f(V(y,s))G_{t-s}(\phi,y)W(dyds).$$

Let us refine this. $E\{V^2(x,t)\}$ is bounded in $[0,L]$, so for a.e. ω $V^2(x,t)$ will be integrable with respect to x (Fubini's theorem). Let ϕ approach a delta function, e.g. take ϕ of the form $(2\pi n)^{-1/2}\exp(-\frac{(y-x)^2}{2n})$ and let $n \to \infty$. The above equation will tend to

$$V(x,t) = \int_0^L V_o(y)G_t(x,y)dy + \int_0^t\int_0^L f(V(y,s))G_{t-s}(x,y)W(dyds) \tag{3.11}$$

a.s. for a.e. pair (t,x). (To see this, apply Lebesgue's differentiation theorem to the left hand side, and note that $G_t(\phi,y) \to G_t(x,y)$ on the right.)

If V_1 and V_2 both satisfy (3.11), let $U = V_2 - V_1$, and define $F(x,t) = E\{U^2(x,t)\}$ and $H(t) = \sup_x F(x,t)$, which is finite by hypothesis. Then from (3.11)

$$F(x,t) = \int_0^t\int_0^L E\{(f(V_2(y,s)) - f(V_1(y,s))^2\}G_{t-s}^2(x,y)dy\,ds$$

$$\leq K^2 \int_0^t\int_0^L F(y,s)G_{t-s}^2(x,y)dy\,ds$$

by (3.5a). Thus

$$H(t) \leq K^2 \int_0^t H(s) \int_0^L G_{t-s}^2(x,y)dy\,ds$$

$$\leq K^2C\int_0^t H(s)\frac{ds}{\sqrt{t-s}}$$

by (3.7). Iterate this:

$$\leq (K^2C)^2\int_0^t\int_0^s H(u)\frac{duds}{\sqrt{(s-u)(t-s)}}\,.$$

Interchange the order in the integral, noting that if $t - u = b$,

$$\int_u^t \frac{ds}{\sqrt{(t-s)(s-u)}} = \int_0^b \frac{dv}{\sqrt{v(b-v)}} \leq 2\int_0^{b/2} \sqrt{\frac{2}{b}}\, \frac{dv}{\sqrt{v}} = 4$$

so that

$$H(t) \leq 4K^4C^2 \int_0^t H(s)\,ds.$$

Iterating this, we see it is

$$\leq \frac{(4K^4C^2)^{n+1}}{n!} \int_0^t H(s)(t-s)^n ds$$

which tends to zero. Thus $H = 0$, so with probability one, $V^1 = V^2$ a.e.

To prove existence, we take a hint from (3.1) and define

$$(3.12) \qquad \begin{cases} V^o(x,t) = \int_0^L V_0(y)G_t(x,y)dy \\ V^{n+1}(x,t) = V^o(x,t) + \int_0^t\int_0^L f(V^n(y,s))G_{t-s}(x,y)W(dyds) \end{cases}$$

Let $p \geq 2$ and suppose that $\{V_0(y),\ 0\leq y\leq L\}$ is L^p bounded. We will show that V^n converges in L^p to a solution V. Define

$$F_n(x,t) = E\{|V^{n+1}(x,t) - V^n(x,t)|^p\}$$

and

$$H_n(t) = \sup_x F_n(x,t).$$

From (3.12)

$$F_n(x,t) = E\{|\int_0^t\int_0^L (f(V^n(y,s)) - f(V^{n-1}(y,s)))G_{t-s}(x,y)W(dyds)|^p\}.$$

We are trying to find the pth moment of a stochastic integral. We can bound this in terms of the associated increasing process by Burkholder's inequality.

$$\leq C_pE\{(\int_0^t\int_0^L |f(V^n(y,s)) - f(V^{n+1}(y,s))|^2G^2_{t-s}(x,y)dyds)^{p/2}\}$$

$$\leq C_pK\ E\{|\int_0^t\int_0^L (V^n(y,s) - V^{n-1}(y,s))^2G^2_{t-s}(x,y)dyds|^{p/2}\}$$

where we have used (3.5a), the Lipschitz condition on f. We can bound this using Hölders inequality. To see how to choose the exponents, note from (3.7) that if $0 < r < 3$,

$$(3.13) \qquad \int_0^L G_t^r(x,y)dy \leq Ce^{-tr}t^{-r/2}\int_{-\infty}^{\infty} e^{-\frac{ry^2}{2t}}\, dy \leq C'e^{-tr}\, t^{\frac{1-r}{2}},$$

which is integrable in t over the interval $(0,\infty)$. Thus we must keep the exponent of G under 3. Set $q = \frac{p}{p-2}$ and choose $0 \le \varepsilon \le 1$ to be strictly between $1 - \frac{3}{p}$ and $\frac{3}{2} - \frac{3}{p}$ ($\varepsilon = 0$ if $p = 2$ and $\varepsilon = 1$ if $p \ge 6$). Then

$$F_n(x,t) \le C\Big(\int_0^t\int_0^L G_s^{2\varepsilon q}(x,y)\Big)^{p/2q} \int_0^t\int_0^L E\{|V^n(y,s)-V^{n-1}(y,s)|^p\}G_{t-s}^{(1-\varepsilon)p}(x,y)\,dyds\ .$$

In this case $2\varepsilon q < 3$, so the first factor is bounded; by (3.13) the expression is

$$\le C\int_0^t H_{n-1}(s)\ (t-s)^a ds,$$

where $a = \frac{1}{2}(1+\varepsilon p-p) > -1$, and C is a constant.

Thus

$$H_n(t) \le C\int_0^t H_{n-1}(s)(t-s)^a ds,\quad t \ge 0 \tag{3.14}$$

for some $a > -1$ and $C > 0$. Notice that if H_{n-1} is bounded on an interval $[0,T]$, so is H_n.

$$H_o(t) \le \sup_x C_p E\Big\{\Big|\int_0^t f(V^o(x,s))^2 G_{t-s}(x,y)dyds\Big|^{p/2}\Big\}.$$

But $V^o(x,s)$ is L^p-bounded since $V_0(y)$ is, hence so is $f(V^o(x,s))$ by (3.5a). An argument similar to the above shows $H_0(t)$ is bounded on $[0,T]$.

Thus the H_n are all finite. We must show they tend to zero quickly. This follows from:

<u>LEMMA 3.3</u>. Let $\{h_n(t),\ n=0,1,\dots\}$ be a sequence of positive functions such that h_0 is bounded on $[0,T]$ and, for some $a > 1$ and constant C_1,

$$h_n(t) \le C_1\int_0^t h_{n-1}(s)(t-s)^a ds,\quad n = 1,2,\dots\ .$$

Then there is a constant C and an integer $k > 1$ such that for each $n \ge 1$ and $t \in [0,T]$,

$$h_{n+mk}(t) \le C^m\int_0^t h_n(s)\ \frac{(t-s)}{(m-1)!}ds,\ m = 1,2,\dots\ . \tag{3.15}$$

Let us accept the lemma for the moment. It applies to the H_n, and implies that for each n, $\sum_{m=0}^{\infty} (H_{n+mk}(t))^{1/p}$ converges uniformly on compacts, and therefore so does $\sum_{n=0}^{\infty} (H_n(t))^{1/p}$. Thus $V^n(x,t)$ converges in L^p, and the convergence is uniform in $[0,L] \times [0,T]$ for any $T > 0$. In particular, V^n converges in L^2. Let $V(x,t) = \lim V^n(x,t)$.

It remains to show that V satisfies (3.9). (Note that it is easy to show that V satisfies (3.11) - this follows from (3.12). However, we would still have to show that (3.11) implies (3.9), so we may as well show (3.9) directly.)

Consider

$$(3.16) \qquad \int_0^L (V^n(x,t)-V_0(x))\phi(x)dx - \int_0^t\int_0^L V^n(x,s)[\phi''(x)-\phi(x)]dx\,ds - \int_0^t\int_0^L f(V^{n-1}(y,s))\phi(y)W(dyds).$$

By (3.12) this is

$$= \int_0^L\int_0^t\int_0^L f(V^{n-1}(y,s))G_{t-s}(x,y)W(dyds)\phi(x)dx$$

$$+ \int_0^L \Big(\int_0^L G_t(x,y)V_0(y)dy - V_0(x)\Big)\phi(x)dx$$

$$- \int_0^t\int_0^L \Big[\int_0^L G_s(x,y)V_0(y)dy + \int_0^u\int_0^L f(V^{n-1}(y,s))G_{u-s}(x,y)W(dyds)\Big](\phi''(x)-\phi(x))dxdu$$

$$- \int_0^t\int_0^L f(V^{n-1}(y,s))\phi(y)W(dyds).$$

Integrate first over x and collect terms:

$$= \int_0^t\int_0^L f(V^{n-1}(y,s))\Big[G_{t-s}(\phi,y) - \int_s^t G_{u-s}(\phi''-\phi,y)ds - \phi(y)\Big]W(dyds)$$

$$- \int_0^L \Big[G_t(\phi,y) - \phi(y) - \int_0^t G_u(\phi''-\phi,y)du\Big]V_0(y)dy.$$

But this equals zero since both terms in square brackets vanish by (3.8). Thus (3.16) vanishes for each n. We claim it vanishes in the limit too.

Let $n \to \infty$ in (3.16). $V^n(x,s) \to V(x,s)$ in L^2, uniformly in $[0,L]\times[0,T]$ for each $T > 0$, and, thanks to the Lipschitz conditions, $f(V^{n-1}(y,s))$ also converges uniformly in L^2 to $f(V(y,s))$.

It follows that the first two integrals in (3.16) converge as $n \to \infty$. So does the stochastic integral, for

$$E\Big\{\Big(\int_0^t\int_0^L (f(V(y,s)) - f(V^{n-1}(y,s))\phi(y)W(dyds)\Big)^2\Big\}$$

$$\leq K \int_0^t\int_0^L E\{(V(y,s) - V^{n-1}(y,s))^2\}\ \phi(y)dyds$$

which tends to zero. It follows that (3.16) still vanishes if we replace V^n and V^{n-1} by V. This gives us (3.9).

Q.E.D.

We must now prove the lemma.

PROOF (of Lemma 3.3). If $a \geq 0$ take $k = 1$ and $C = C_1$. If $-1 < a < 0$,

$$h_n(t) \leq C_1^2 \int_0^t h_{n-2}(u) \left(\int_u^t (t-s)^a (s-u)^a ds \right) du.$$

If $a = -1 + \varepsilon$, the inner integral is bounded above by

$$2\left(\frac{2}{t-u}\right)^{1-\varepsilon} \int_0^{1/2(t-u)} \frac{dv}{v^{1-\varepsilon}} \leq \frac{4}{\varepsilon} (t-u)^{2\varepsilon-1}$$

so

$$h_n(t) \leq C_1 \int_0^t h_{n-1}(s) \frac{ds}{(t-s)^{1-\varepsilon}} \leq \frac{4}{\varepsilon} C_1^2 \int_0^t h_{n-2}(s) \frac{ds}{(t-s)^{1-2\varepsilon}} .$$

If $2\varepsilon \geq 1$ we stop and take $k = 2$ and $C = \frac{4}{\varepsilon} C_1^2$. Otherwise we continue

$$\leq \frac{16}{\varepsilon^2} C_1^4 \int_0^t h_{n-4}(s) \frac{ds}{(t-s)^{1-4\varepsilon}}$$

until we get (t-s) to a positive power. When this happens, we have

$$h_n(t) \leq C \int_0^t h_{n-k}(s) ds.$$

But now (3.15) follows from this by induction. Q.E.D.

In many cases the initial value $V_0(x)$ is deterministic, in which case $V(x,t)$ will be bounded in L^p for all p. We can then show that V is actually a continuous process, and, even better, estimate its modulus of continuity.

COROLLARY 3.4. Suppose that $V_0(x)$ is L^p-bounded for all $p > 0$. Then for a.e. ω, $(x,t) \to V(x,t)$ is a Hölder continuous function with exponent $\frac{1}{4} - \varepsilon$, for any $\varepsilon > 0$.

PROOF. A glance at the series expansion of G_t shows that it can be written

$$G_t(x,y) = g_t(x,y) + H_t(x,y)$$

where

$$g_t(x,y) = (4\pi t)^{-1/2} e^{-\frac{(y-x)^2}{4t} - t},$$

$H_t(x,y)$ is a smooth function of (t,x,y) on $(0,L) \times (0,L) \times (-\infty, \infty)$, and H vanishes if $t \leq 0$. By (3.11)

$$V(x,t) = \int_0^L V_0(y)G_t(x,y)dy + \int_0^t\int_0^L f(V(y,s))H_{t-s}(x,y)W(dyds)$$
$$+ \int_0^t \int_0^L f(V(y,s))g_{t-s}(x,y)W(dyds).$$

The first term on the right hand side is easily seen to be a smooth function of (x,t) on $(0,L) \times (0,\infty)$. The second term is basically a convolution of W with a smooth function H. It can also be shown to be smooth; we leave the details to the reader.

Denote the third term by U(x,t). We will show that U is Hölder continuous by estimating the moments of its increments and using Corollary 1.4. Now

$$E\{|U(x+h,\ t+k) - U(x,t)|^n\}^{1/n} \le E\{|U(x+h,\ t+k) - U(x,\ t+k)|^n\}^{1/n}$$
$$+ E\{|U(x,t+k) - U(x,t)|^n\}^{1/n}.$$

We will estimate the two terms separately. The basic idea is to use Burkholder's inequality to bound the moments of each of the stochastic integrals. Replacing t+k by t, we see that

$$E\{|U(x+h,t) - U(x,t)|^n\} \le C\ E\{|\int_0^t \int_0^L f^2(V(y,s))(g_{t-s}(x+h,y) - g_{t-s}(x,y))^2 dyds|^{\frac{n}{2}}\}.$$

Apply Hölders inequality with $p = n/2$, $q = \frac{n}{n-2}$:

$$\le C_n\ E\{ \int_0^t\int_0^L f(V(y,s))^2 dyds\}\ [\int_0^t\int_0^L |g_s(x+h,y)-g_s(x,y)|^{2q}dyds]^{\frac{n}{2q}}.$$

The expectation is finite for any n by (3.5a) and Theorem 3.2. Letting C be a constant whose value many change from line to line we have

$$\le C\ [\int_0^t\int_{-\infty}^{\infty} \frac{e^{-2qs}}{(4\pi s)^q}\ |e^{\frac{(y+h)^2}{4s}} - e^{-\frac{y^2}{4s}}|^{2q} dy\ ds]^{\frac{n}{2q}}.$$

If we let $y = hz$ and $s = h^2v$, we can see this is

$$\le C[h^{3-2q}\int_0^\infty\int_0^\infty v^{-q}|e^{-\frac{(z+1)^2}{4v}} - e^{-\frac{z^2}{4v}}|^{2q}\ dv]^{\frac{n}{2q}}.$$

The integral converges if $q < 3/2$, i.e. if $n > 6$, so that this is

$$= C\ h^{\frac{n}{2}-1}.$$

The first term of (3.15) is thus bounded by $C\ h^{\frac{1}{2}-\frac{1}{n}}$.

Similarly,

$$E\{|U(x,t+k) - U(x,t)|^n\}^{1/n} \le C_n\, E\{|\int_0^t\int_0^L f(V(y,s))^2|g_{t+k}(x,y) - g_t(x,y)|^2 dyds|^{\frac{n}{2}}\}^{\frac{1}{n}}$$

$$+ C_n E\{|\int_t^{t+k}\int_0^L f(V(y,s))^2|g_{t+k}(x,y)|^2 dyds|^{\frac{n}{2}}\}^{\frac{1}{n}}.$$

The first expectation on the right is bounded by

$$C\, E\{\int_0^t\int_0^L f(V(y,s))^n dy\, ds\}^{1/n}[\int_0^t\int_{-\infty}^{\infty} |(s+k)^{-1/2}e^{-\frac{y^2}{4(s+k)}} - s^{-1/2}e^{-\frac{y^2}{4s}}|^{2q}]^{\frac{1}{2q}}$$

The expectation above is finite. If we set $s = ku$, $y = \sqrt{k}\, z$, this becomes

$$\le C[k^{3/2-q}\int_0^{\infty}\int_{-\infty}^{\infty} \left|\frac{e^{-\frac{-z^2}{4(u+1)}}}{\sqrt{u+1}} - \frac{e^{-\frac{z^2}{4u}}}{\sqrt{u}}\right|^{2q} dzdu]^{1/2q}.$$

The integral converges if $q < 3/2$, i.e. if $n > 6$, so the expansion is

$$= C\, k^{\frac{1}{4}-\frac{2}{n}}.$$

Finally, the second expectation on the right is bounded by

$$C_n E\{\int_t^{t+k}\int_0^L f^n(V(y,s))dyds\}^{1/n}[\int_0^k\int_0^L g_s(x,y)^2 dyds]^{1/2q}$$

We have seen that $E\{f^n(V(y,s))\}$ is bounded so this is

$$\le C\, k^{1/n}[\int_0^k\int_{-\infty}^{\infty} s^{-q}\, e^{-\frac{qy^2}{s}}\, dy\, ds]^{1/2q}.$$

We can do the integral explicitly to get

$$= C\, k^{\frac{1}{4}-\frac{2}{n}}.$$

Putting these bounds together in (3.16) we see that

$$E\{|U(x+h,\, t+k) - U(x,t)|^n\}^{1/n} \le C[h^{\frac{1}{4}-\frac{1}{n}} + 2k^{\frac{1}{4}-\frac{2}{n}}]$$

$$\le C(\sqrt{h^2+k^2})^{\frac{1}{4}-\frac{4}{n}}.$$

We can choose n as large as we please, so that the result follows from Kolmogorov's Theorem (Corollary 1.4).

Q.E.D.

The uniqueness theorem gives us the Markov property of the solution, exactly as it does in the classical case. We omit the proof.

<u>THEOREM 3.5</u>. The process $\{V(\cdot,t),\ t\geq 0\}$, considered as a process taking values in $C[0,L]$, is a diffusion process.

Consider the more general equation

$$(3.5b)\qquad \begin{cases} \dfrac{\partial V}{\partial t} = \dfrac{\partial^2 V}{\partial x^2} + g(V,t) + f(V,t)\dot{W}, & t > 0,\ 0 < x < L; \\ \dfrac{\partial V}{\partial x}(0,t) = \dfrac{\partial V}{\partial x}(L,t) = 0, & t > 0; \\ V(x,0) = V_0(x),\ 0 < x < L. \end{cases}$$

<u>Exercise 3.3</u>. Find the weak form of (3.5b).

<u>Exercise 3.4</u>. Show that if both f and g satisfy the Lipschitz condition (3.5a) then Theorem 3.2 holds. In particular, (3.5b) has a unique weak solution.

(Hint. The Green's function is as before except that there is no factor e^{-t}, and the Picard iteration formula (3.12) becomes

$$V^{n+1}(x,t) = V^0(x,t) + \int_0^t\int_0^L G_{t-s}(x,y)[g(V^n(y,s))dyds + f(V^n(y,s))W(dyds)].$$

The proof of Theorem 3.2 then needs only a little modification. For instance, for uniqueness, let

$$F(x,t) = 2\int_0^t\int_0^L G^2_{t-s}(x,y)[(f(V_2(y,s)) - f(V_1(y,s)))^2 + Lt(g(V_2(y,s) - g(V_1(y,s)))^2]dyds,$$

show that $E\{|V_2(x,t) - V_1(x,t)|^2\} \leq F(x,t)$,

and conclude that $H(t) \leq K^2C \int_0^t H(s)\dfrac{ds}{\sqrt{t-s}}$.

In order to prove existence, define F_n and H_n as in the proof and note that, once (3.14) is established, the rest of the proof follows nearly word by word. In order to prove (3.14), first show that

$$F_n(x,t) \leq 2^pK\,E\{|\int_0^t\int_0^L |V^n(y,s) - V^{n-1}(y,s)|G_{t-s}(x,y)dyds|^p\}$$
$$+ 2^pC_pK\,E\{|\int_0^t\int_0^L |V^n(y,s) - V^{n-1}(y,s)|^2G_{t-s}(x,y)dyds|^{p/2}\},$$

and then apply Hölder's inequality to each term as in the proof to deduce (3.14).)

<u>Exercise 3.5</u>. Show that Corollary 3.5 also holds for solutions of (3.5b), so that the weak solution of (3.5) is Hölder continuous.

In case f is <u>constant</u>, there is a direct relation between the solutions of (3.5) and (3.5b).

<u>Exercise 3.6</u>. Let $f(x,t) = \sigma$ be constant, and let U and V be solutions of (3.5) and (3.5b) respectively. Show that $V = U + u$, where $u(x,t)$ is the solution of the PDE

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + g(U(x,t) + u(x,t),t); \\ u(0,t) = u(L,t) = 0; \\ u(x,0) = 0. \end{cases}$$

Thus write V explicitly in terms of U.

(The point is that once U is known, one can fix ω and solve this as a classical non-stochastic PDE; the solution u can be written in terms of the Green's function.)

The technique of Picard iteration works for the non-linear wave equation, too. Consider

$$\text{(3.1a)} \qquad \begin{cases} \dfrac{\partial^2 V}{\partial t^2} = \dfrac{\partial^2 V}{\partial x^2} + g(V,t) + f(V,t)\dot{W}, & t > 0,\ x \in \mathbf{R}; \\ V(x,0) = V_0(x), & x \in \mathbf{R}; \\ \dfrac{\partial V}{\partial t}(x,0) = U_0(x), & x \in \mathbf{R}. \end{cases}$$

In this case we let $V^0(x,t)$ be the classical solution of the homogeneous wave equation with initial position $V_0(x)$ and velocity $U_0(x)$ - which we can write explicitly - and define

$$V^{n+1}(x,t) = V^0(x,t) + \int_0^t \int_{\mathbf{R}} C(x,t;y,s)[g(V^{n-1}(y,s))dyds + f(V^{n-1}(y,s))W(dyds)]$$

where $C(x,t;\, y,s) = \begin{cases} 1 & \text{if } s \le t \text{ and } |y-x| \le t-s, \\ 0 & \text{otherwise}. \end{cases}$

Then C is the indicator function of the light cone.

The Picard iteration is in fact easier than it was for the cable equation since C is bounded. We leave it as an exercise for iteration enthusiasts to carry out.

Exercise 3.7. (a) Write (3.1a) in a weak form analogous to (3.2).

(b) Show that if both f and g satisfy the Lipschitz conditions (3.5a), then (3.1a) has a unique weak solution, which has a Hölder continuous version.

THE LINEAR EQUATION

Let us now consider the linear equation ($f \equiv$constant). This is relatively easy to analyze because the solution is Gaussian and most questions can be answered by computing covariances. (The case $f(x) = ax + b$ is linear too, but the solution, which now involves a product $V\dot{W}$, is no longer Gaussian. This case is often referred to as semi-linear).

The solution can be expanded in eigenfunctions. Assume $L = \pi$, so that (3.5) becomes

(3.17)
$$\begin{cases} V_t = V_{xx} - V + \dot{W}, & 0 < x < \pi,\ t > 0, \\ V_x(0,t) = V_x(\pi,t) = 0, & t > 0; \\ V(x,0) = 0, & 0 < x < \pi. \end{cases}$$

The eigenfunctions and eigenvalues of (3.17) are

$$\phi_0 \equiv 1/\sqrt{\pi}, \quad \phi_k(x) = \sqrt{\frac{2}{\pi}} \cos kx \quad k = 1,2,\ldots$$

$$\lambda_k = k^2 + 1,\ k = 0,1,2,\ldots\ .$$

The Green's function can be expanded in the ϕ_k:

$$G_t(x,y) = \sum_{k=0}^{\infty} \phi_k(x)\phi_k(y)e^{-\lambda_k t}.$$

For each fixed x, this converges in $L^2[0,\pi]\times[0,T]$ as a function of (y,t).

Thus the unique solution of (3.17) is, by (3.11)

(3.17a)
$$\begin{aligned} V_t &= \int_0^t \int_0^\pi G_{t-s}(x,y)W(dyds) \\ &= \int_0^t \int_0^\pi \sum_{k=0}^{\infty} \phi_k(x)\phi_k(y)e^{-\lambda_k(t-s)} W(dyds). \end{aligned}$$

We can interchange order since the series converges in L^2:

$$= \sum_{k=0}^{\infty} \left(\int_0^t \int_0^\pi \phi_k(y) e^{-\lambda_k(t-s)} W(dyds)\right)\phi_k(x).$$

Exercise 3.8. Define $B_t^k = \int_0^t \int_0^\pi \phi_k(y) W(dyds$

and $A_t^k = \int_0^t \int_0^\pi \phi_k(y) e^{-\lambda_k(t-s)} W(dyds).$

(i) Show that $B^1, B^2, \ldots$ are iid standard Brownian motions, and that

$$A_t^k = \int_0^t e^{-\lambda_k(t-s)} dB_s^k.$$

(ii) Show that A^k satisfies $dA_t^k = dB_t^k - \lambda_k A_t^k dt$.

The processes A^k are familiar, for they are Ornstein-Uhlenbeck processes with parameter λ_k - abbreviated OU(λ_k) - which are mean zero Gaussian Markov processes. They are independent.

Thus we have

$$V(x,t) = \sum_{k=0}^{\infty} A_t^k \phi_k(x) \tag{3.17b}$$

where A^k is an OU(λ_k) process and $A^0, A^1, A^2, \ldots$ are independent. Recall the following well-known facts about Ornstein-Uhlenback processes.

PROPOSITION 3.6. Let $\{A_t, t \geq 0\}$ be an OU(λ) process with $A_0 = 0$. A is a mean zero Gaussian process with covariance function

(i) $E\{A_{s+t}A_t\} = \frac{e^{-\lambda s}}{2\lambda}\left[1 - e^{-2\lambda t}\right];$

(ii) $E\{(A_{s+t} - A_t)^2\} = \frac{1}{\lambda}(1 - e^{-\lambda s}) - \frac{1}{2\lambda} e^{-2\lambda t}(1 - e^{-\lambda s})^2.$

PROPOSITION 3.7. If $0 \leq s \leq t$ and $0 \leq x, y \leq \pi$,

(i) $E\{(V(y,t) - V(x,t))^2\} \leq 4|y-x|;$

(ii) $E\{(V(x,t) - V(y,s))^2\} \leq \frac{4}{\pi}\sqrt{t-s}.$

Before proving this, let us see what we can learn from it. From (i) and (ii),

$$\max\{E\{(V(x,t) - V(y,s))^2\} : |t-s|^2 + |y-x|^2 \leq 2^{-1/2}u\} \leq Cu^{1/4}$$

Thus, let $p(u) = Cu^{1/4}$ in Corollary 1.3 to get

THEOREM 3.8. V has a version which is continuous in the pair (x,t). For T > 0 there is a constant C and random variable Y such that the modulus of continuity $\Delta(\delta)$ of $\{V(x,t),\ t\le T)\}$ satisfies

$$\Delta(\delta) \le Y\delta^{1/4} + C\delta^{1/4}\sqrt{\log 1/\delta}, \quad 0 \le \delta \le 1.$$

When we compare Theorem 3.8 with Corollary 3.4, we see that the moduli of continuity are substantially the same in the linear and non-linear cases. The paths are essentially Hölder (1/4).

We will need the following lemma.

LEMMA 3.9. (i) $\sum_{k=1}^{\infty} \frac{(\phi_k(y)-\phi_k(x))^2}{2\lambda_k} \le 4|y-x|$

(ii) $\sum_{k=1}^{\infty} \frac{1-e^{-\lambda_k t}}{2\lambda_k} \le 1 \wedge \sqrt{t}.$

PROOF. $\phi_k(x) = \sqrt{2}\cos kx$ so

$$(\phi_k(y) - \phi_k(x))^2 \le 2(4\wedge k^2(y-x)^2) .$$

Since $\lambda_k > k^2$,

$$\sum_{k=1}^{\infty} \frac{(\phi_k(y)-\phi_k(x))^2}{2\lambda_k} \le \int_1^{\infty} \frac{4}{u^2} \wedge (y-x)^2 du$$

$$\le \int_1^{2|y-x|^{-1}} (y-x)^2 du + 4 \int_{2|y-x|^{-1}}^{\infty} \frac{du}{u^2}$$

$$\le 4|y-x|.$$

The second series is handled the same way, using $1 - e^{-\lambda_k t} \le 1 \wedge (1+k^2)t$. We leave the details to the reader.

PROOF (of Proposition 3.7). We prove (i), for (ii) is similar. By (3.17b) and Propositon 3.6,

$$E\{(V(y,t) - V(x,t))^2\} = E\{[\sum_{k=0}^{\infty} A_t^k(\phi_k(y) - \phi_k(x))]^2\}$$

$$\le \sum_{k=0}^{\infty} \frac{1}{2\lambda_k} (\phi_k(y) - \phi_k(x))^2$$

$$\le 4|y-x|$$

by the Lemma. Q.E.D.

When we compare Theorem 3.8 with Corollary 3.4, we see that the moduli of continuity are substantially the same in the linear and non-linear cases. The paths are essentially Hölder (1/4).

Note that V is rougher than either Brownian motion or the Brownian sheet. Both of the latter are Hölder continuous of exponent 1/2, while V has exponent 1/4. We can ask if this is due to bad behavior in t, or in x, or in combination. The following exercises eliminate x as a suspect.

Exercise 3.9. Expand Brownian motion B_t in the eigenfunctions ϕ_k to see

$$B_t = \frac{1}{\sqrt{3}}\,\xi_0 + \sum_{k=1}^{\infty} \frac{1}{k}\,\xi_k\phi_k(t), \quad 0 < t < \pi, \tag{3.18}$$

where the ξ_k are iid N(0,1).

Exercise 3.10. Write $V(x,t) = \frac{1}{\sqrt{2}} \sum_{k=0}^{\infty} \frac{\sqrt{2\lambda_k}\,A_k(t)}{\sqrt{\lambda_k}}\,\phi_k(x)$ and compare with (3.18). Conclude that

$$V(x,t) = \frac{1}{\sqrt{2}}\,B_x + R_x$$

where $\{B_x,\ 0\leq x<\pi\}$ is a Brownian motion and R is twice-differentiable.

Evidently $x \to V(x,t)$ will have the same local behavior as Brownian motion so that t must be the culprit responsible for the bad behavior of the paths. One striking exhibit of this is the following, according to which $t \to V(x,t)$ has non-trivial quartic (i.e. fourth power) variation, whereas Brownian motion has quadratic variation but zero quartic variation. Define

$$Q_n(t) = (V(0,t) - V(0,[2^n t]2^{-n}))^4 + \sum_{j=1}^{[nt]} (V(0,j2^{-n}) - V(0,(j-1)2^{-n}))^4 .$$

THEOREM 3.10. For a.e. ω, $Q_n(t,\omega)$ converges uniformly on compacts to a limit Qt, where $Q > 0$ is a constant.

THE BARRIER PROBLEM

There is one open problem which deserves mention here, because it is an important question for the neuron, which this equation is meant to describe. That is the problem of finding the distribution of the first hitting time of a given level.

The neuron collects electrical impulses, and when the potential at a certain spot - called the soma and represented here as x = 0 - passes a fixed level, called the barrier, the neuron fires and transmits an impulse, the action potential, through the nervous system. The generation of the action potential comes from non-linearities not present in the cable equation, so the cable equation is valid only until the first time τ that $V(0,t)$ exceeds the barrier. However, it can still be used up to τ, and we can ask the question, "what is the distribution of τ?"

We will describe the problem in a bit more detail and show how it is connected with a first-hitting problem for infinite-dimensional diffusions.

Set $\lambda > 0$ and put

$$\tau = \inf\{t > 0 : V(0,t) > \lambda\}.$$

One can show $\tau < \infty$ a.s. and that its moments are finite, and that it even has some exponential moments. Write

$$V(0,t) = \frac{1}{\sqrt{\pi}} A_t^o + \sqrt{\frac{2}{\pi}} \sum_{k=1}^{\infty} A_t^k .$$

Now $V(\cdot,t)$ is a Markov process, but $V(0,t)$ is not, so that the method of studying τ by reducing the problem to a question in differential equations can't be applied directly. However, note that $\{A_t^o, t \geq 0\}$ is a diffusion, and, moreover, if

$$A_\infty(t) = (A_t^0, A_t^1, \ldots)$$

and

$$A_N(t) = (A_t^o, \ldots, A_t^N),$$

then A_N is a diffusion in R^{N+1}, and A_∞ is a diffusion in ℓ_∞. Define

$$\tau_N = \inf\{t\colon A_t^o + \sqrt{2} \sum_{k=1}^{N} A_t^k > \sqrt{\pi}\lambda\}$$

for $N = 0,1,\ldots,\infty$; $(\tau_\infty = \tau)$.

Let H_λ be the half space

$$H_\lambda = \{x \in \ell_\infty : x_0 + \sqrt{2} \sum_{1}^{\infty} x_k > \sqrt{\pi}\lambda\}.$$

Then τ is the first hitting time of H_λ by the infinite-dimensional diffusion A. Since the components of A are independent $OU(\lambda_k)$ processes, we can write down its infinitesimal generator and recast the problem in terms of PDE's, at least formally.

Let us see how we would find the expected value of τ, for instance. Suppose F is a smooth function on ℓ_∞ depending on only finitely many coordinates. Then A has the generator $\mathbf{G}$:

$$\mathbf{G}F(x) = \sum_{k=0}^{\infty} \left(\frac{1}{2} \frac{\partial^2 F}{\partial x_k^2} - \lambda_k \frac{\partial F}{\partial x_k}\right)(x).$$

Let us proceed purely formally - which means that we will ignore questions about the domain of $\mathbf{G}$ and won't look too closely at the boundary values - and set

$$F(x) = E\{\tau | A(0) = x\}, \quad x \in \ell_\infty .$$

Then F should satisfy

$$(3.19) \begin{cases} \text{(i)} \quad \mathbf{G}F = -1 & \text{in } \ell_\infty - H_\lambda \\ \text{(ii)} \quad F = 0 & \text{on } \partial H_\lambda \\ \text{(iii) F is the smallest positive function satisfying (i) and (ii).} \end{cases}$$

Then F(0) is the solution to our problem.

Now (3.19) would hold rigorously for a diffusion in $\mathbf{R}^N$, and in particular, it does hold for each of the $A_N(t)$. But, rigor aside, we can not solve (3.19). We can solve the finite-dimensional analogue. For N = 0 we can solve it in closed form and for small N, we can solve it numerically, but even this becomes harder and harder as N increases. (In this context, N = 1 is a large number, N = 2 is immense, and N = 3 is nearly infinite.)

We have the following:

<u>THEOREM 3.11.</u> $\lim_{N\to\infty} E\{\tau_N\} = E\{\tau\}$.

This might appear to solve the problem, but, in view of the difficulty of finding $E\{\tau_N\}$, one must regard the problem of finding $E\{\tau\}$ as open, and the problem of finding the exact distribution of τ as essentially unattempted.

HIGHER DIMENSIONS

Let us very briefly pose the analogous problem in $\mathbf{R}^2$ and see why the above methods fail.

Consider

$$(3.20)\qquad \begin{cases} \frac{\partial V}{\partial t}(x,y;t) = \left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} - V\right)(x,y;t) + \dot{W}_{xyt} \\ \frac{\partial V}{\partial x}(0,y,t) = \frac{\partial V}{\partial x}(\pi,y,t) = \frac{\partial V}{\partial y}(x,0;t) = \frac{\partial V}{\partial y}(x,\pi;t) = 0 \\ V(x,y,0) = 0 \end{cases}$$

The problem separates, and the eigenfunctions are

$$\phi_{jk}(x,y) = \phi_j(x)\phi_k(y)$$

where the ϕ_j are the eigenfunctions of (3.17), and the eigenvalues are

$$\lambda_{jk} = 1 + j^2 + k^2.$$

Proceeding as before, set

$$A_t^{jk} = \int_0^t \int_0^L e^{-\lambda_{jk}(t-s)} \phi_j(x)\phi_k(y)\, W(dxdyds).$$

The A^{jk} are independent $OU(\lambda_{jk})$ processes, as before, and the solution of (3.19) should be

$$(3.21)\qquad V(x,y,t) = \sum_{j,k=0}^{\infty} A^{jk}(t)\phi_j(x)\phi_k(y).$$

The only problem is that the series on the right hand side does not converge. Indeed, choose, say, $x = y = 0$ and note that for $t = 1$ and large j and k, that $\sqrt{\frac{\pi}{2}}\, A^{jk}(t)$ is essentially a $N(0, \frac{1}{\lambda_{jk}})$ random variable by Proposition 3.6. Thus (3.21) converges iff $\sum_{j,k} \frac{1}{\lambda_{jk}} < \infty$. But $\sum_{j,k} \frac{1}{1+j^2+k^2}$ diverges!

One can check that the representation of V as a stochastic integral analogous to (3.17a) also diverges.

However - and there is a however, or else this course would finish right here - we can make sense of (3.21) as a Schwartz distribution. Let ψ be a C^∞ function of compact support in $(0,\pi) \times (0,\pi)$, and write

$$V(\psi,t) = \iint \psi(x,y)\, V(x,y,t)dxdy$$
$$= \sum_{j,k=0}^{\infty} A_{jk}(t)\, \hat{\psi}_{jk}$$

where $\hat{\psi}_{jk} = \int_0^\pi \int_0^\pi \phi_{jk}(x,y)\psi(x,y)dxdy$.

Now the first integral makes no sense, but the sum does, since $\hat{\psi}_{jk}$ tends to zero faster than $j^2 + k^2$ as j and k go to ∞ - a well-known fact of Fourier series - so $\sum \frac{\hat{\psi}_{jk}^2}{1+j^2+k^2} < \infty$. Thus $V(\psi,t)$ makes sense for any test function ψ, and we can use this to define V as a Schwartz distribution rather than as a function.

CHAPTER FOUR

DISTRIBUTION-VALUED PROCESSES

If M_t is a martingale measure and ϕ a test function, put

$$M_t(\phi) = \int_0^L \int_0^t \phi dM.$$

Then M is clearly additive:

$$M_t(a\phi + b\psi) = aM_t(\phi) + bM_t(\psi) \quad \text{a.s.}$$

The exceptional set may depend on a,b,ϕ, and ψ however. We cannot say a priori that $\phi \to M_t(\phi)$ is a continuous linear functional, or even that it is a linear functional. In short, M_t is not yet a distribution. However, it is possible to construct a regular version of M which is. This depends on the fact that spaces of distributions are nuclear spaces.

Let us recall some things about nuclear spaces; we will consider only the simplest setting, which is already sufficient for our purposes.

A norm $\| \ \|$ on a vector space E is <u>Hilbertian</u> if $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$, $x,y \in E$. The associated inner product is

$$\langle x,y\rangle = \frac{1}{4}(\|x+y\| - \|x-y\|),$$

so that $(E,\| \ \|)$ is a pre-Hilbert space.

If $\| \ \|_1$ and $\| \ \|_2$ are Hilbertian norms, we say $\| \ \|_1$ is <u>HS weaker than</u> $\| \ \|_2$, and write $\| \ \|_1 \underset{HS}{<} \| \ \|_2$, if

(4.1) $$\sup\Big\{ \sum_k \|e_k\|_1^2 : (e_k) \text{ is } \| \ \|_2\text{-ortho-normal}\Big\} < \infty .$$

(HS stands for Hilbert-Schmidt, for (4.1) is equivalent to the injection map of $(E,\| \ \|_2) \to (E,\| \ \|_1)$ being Hilbert-Schmidt.)

If E is separable relative to $\| \ \|_2$, we can use the Gram-Schmidt procedure to construct a complete ortho-normal basis (f_k) for $(E,\| \ \|_2)$. In this case (4.1) is equivalent to $\sum_k \|f_k\|_1^2 < \infty$.

Let E be a vector space and let $\| \ \|_0 < \| \ \|_1 < \| \ \|_2 < \cdots$ be a sequence of Hilbertian norms on E such that

(i) E is separable with respect to $\| \ \|_n$, all n;

(ii) for each m, there exists n > m such that $\| \ \|_m \underset{HS}{<} \| \ \|_n$.

For each n let $e_{n1}, e_{n2}, \ldots$ be a complete ortho-normal system (CONS) in $(E, \| \ \|_n)$. Let E'_n be the dual of $(E, \| \ \|_n)$ with dual norm $\| \ \|_{-n}$ defined by

$$\|f\|_{-n}^2 = \sum_{k=1}^{\infty} f(e_{nk})^2, \quad f \in E'_n. \tag{4.2}$$

It will be clear shortly why we use $-n$ as an index. Meanwhile, note that $E'_m \subset E'_n$ if $m < n$ for, since $\| \ \|_m \leq \| \ \|_n$, any linear functional on E which is continuous relative to $\| \ \|_m$ is also continuous relative to the larger norm $\| \ \|_n$. Note also that E'_n is a Hilbert space; denote $H_{-n} = E'_n$. For $n = 0,1,2,\ldots$ let H_n be the completion of E relative to $\| \ \|_n$.

Then H_{-n} is the dual of H_n. (We identify H_0 with its dual H_{-0}, but we do <u>not</u> identify H_n and H_{-n}. In fact we have:

$$\cdots \supset H_{-2} \supset H_{-1} \supset H_0 \supset H_1 \supset H_2 \supset \cdots$$

Let us give E the toplogy determined by the $\| \ \|_n$. A neighborhood basis of 0 is $\{x : \|x\|_n < \varepsilon\}$, $n = 0,1,2,\ldots$, $\varepsilon > 0$.

Let E' be the dual of E. Then $E' = \bigcup_n H_{-n}$. To see this, suppose $f \in E'$. Then there is a neighborhood G of zero such that $|f(y)| < 1$ if $y \in G$. Thus there is a member of the basis such that $\{x : \|x\|_n < \varepsilon\} \subset G$. For $\delta > 0$, if $\|x\|_n < \varepsilon\delta$, then $|f(x)| < \delta$. This implies that $\|f\|_{-n} \leq 1/\varepsilon$, i.e. $f \in H_{-n}$. Conversely, if $f \in H_{-n}$, it is a linear functional on E, and it is continuous relative to $\| \ \|_n$, and hence continuous in the topology of E.

<u>Note</u>: The argument above also proves the following more general statement: Let F be a linear map of E into a metric space. Then F is continuous iff it is continuous in one of the norms $\| \ \|_n$.

We give E' the strong topology: a set $A \subset E$ is <u>bounded</u> if it is bounded in each norm $\| \ \|_n$, i.e. if $\{\|x\|_n, x \in A\}$ is a bounded set for each n. Define a semi-norm

$$p_A(f) = \sup\{|f(x)| : x \in A\}.$$

The strong topology is generated by the semi-norms $\{p_A : A \subset E \text{ is bounded}\}$. Now E is not in general normable, but its topology is compatible with the metric

$$d(x,y) = \sum_n 2^{-n}(1 + \|y-x\|_n)^{-1}\|y-x\|_n,$$

and we can speak of the completeness of E. If E is complete, then $E = \bigcap_n H_n$.

(Clearly $E \subset \bigcap_n H_n$, and if $x \in \bigcap_n H_n$, then for each n there is $x_n \in E$ such that

$\|x-x_n\|_n < 2^{-n} \Rightarrow \|x-x_n\|_m < 2^{-n}$ $m < n$. Thus $d(x,x_n) < 2^{-n+1}$.)

If E is complete, it is called a <u>nuclear space</u>, and we have

$$E' = \bigcup_n H_{-n} \supset \ldots \supset H_{-1} \supset H_0 \supset H_1 \supset H_2 \supset \ldots \supset \bigcap_n H_n = E$$

where H_n is a Hilbert space relative to the norm $\| \ \|_n$, $-\infty < n < \infty$, E is dense in H_n, H_{-n} is dual to H_n, and for all m there exists $n > m$ such that $\| \ \|_m \underset{HS}{<} \| \ \|_n$.

We may not often use the following explicitly in the sequel, but it is one of the fundamental properties of the spaces H_n.

<u>Exercise 4.1</u>. Suppose $\| \ \|_m \underset{HS}{<} \| \ \|_n$. Then the closed unit ball in H_n is compact in H_m. (Hint: show it is totally bounded.)

REGULARIZATION

Let E be a nuclear space as above. A stochastic process $\{X(x), x \in E\}$ is a <u>random linear functional</u> if, for each $x,y \in E$ and $a,b, \in R$,

$$X(ax + by) = aX(x) + bX(y) \text{ a.s.}$$

<u>THEOREM 4.1</u>. Let X be a random linear functional on E which is continuous in probability in $\| \ \|_m$ for some m. If $\| \ \|_m \underset{HS}{<} \| \ \|_n$, then X has a version which is in H_{-n} a.s. In particular, X has a version with values in E'.

Convergence in probability is metrizable, being compatible with the metric

$$|\!|\!| X(x) |\!|\!| \overset{\text{def}}{=} E\{|X(x)| \wedge 1\}.$$

If X is continuous in probability on E, it is continuous in probability in $\| \ \|_m$ for some m by our note. There exists n such that $\| \ \|_m \underset{HS}{<} \| \ \|_n$. Thus we have

<u>COROLLARY 4.2</u>. Let X be a random linear functional which is continuous in probability on E. Then X has a version with values in E'.

PROOF (of Theorem 4.1). Let (e_k) be a CONS in $(E,\| \|_n)$. We will first show that $\sum X(e_k)^2 < \infty$.

For $\varepsilon > 0$ there exists $\delta > 0$ such that $|||X(x)||| < \varepsilon$ whenever $\|x\|_m < \delta$. We claim that

$$\operatorname{Re} E\{e^{iX(x)}\} \geq 1 - 2\varepsilon - 2\varepsilon\delta^{-2}\|x\|_m^2 .$$

Indeed, the left-hand side is greater than $1 - \frac{1}{2} E\{X^2(x)\wedge 4\}$, and if $\|x\|_m \leq \delta$,

$$E\{X^2(x)\wedge 4\} \leq 4E\{|X(x)|\wedge 1\} \leq 4\varepsilon,$$

while if $\|x\|_m > \delta$,

$$E\{X^2(x)\wedge 4\} \leq \|x\|_m^2\delta^{-2}E\{X^2(\delta x/\|x\|_m)\wedge 4\} \leq 4\ \varepsilon\ \delta^{-2}\|x\|_m^2.$$

Let us continue the trickery by letting $Y_1,Y_2,\ldots$ be iid $N(0,\sigma^2)$ r.v. independent of X, and set $x = \sum_{k=1}^{N} Y_k e_k$. Then

$$\operatorname{Re} E\{e^{iX(x)}\} = E\Big\{\operatorname{Re}E\{\exp\big[i \sum_1^N Y_k X(e_k)\big]|X\}\Big\}.$$

But if X is given, $\sum Y_k X(e_k)$ is conditionally a mean zero Gaussian r.v. with variance $\sigma^2\sum X^2(e_k)$, and the above expectation is its characteristic function:

$$= E\{e^{-\frac{\sigma^2}{2}\{\sum_{k=1}^{N} X^2(e_k)\}}\}.$$

On the other hand, it also equals

$$E\{\operatorname{Re} E\{e^{i\sum Y_k X(e_k)}|Y\}\} \geq 1 - 2\varepsilon - 2\delta^{-2}\varepsilon E\{\|x\|_m^2\}$$

$$= 1- 2\varepsilon - 2\delta^{-2}\varepsilon \sum_{j,k=1}^{N} E\{Y_jY_k\} < e_j,e_k>_m$$

$$= 1 - 2\varepsilon - 2\delta^{-2}\ \varepsilon\ \sigma^2 \sum_{k=1}^{N} \|e_k\|_m^2.$$

Thus

$$E\{e^{-\frac{\sigma^2}{2}\sum_{k=1}^{N} X^2(e_k)}\} \geq 1 - 2\varepsilon - 2\delta^{-2}\varepsilon\sigma^2 \sum_{k=1}^{N} \|e_k\|_m^2.$$

Let $N \to \infty$ and note that the last sum is bounded since $\| \|_m \underset{HS}{<} \| \|_n$. Next let $\sigma^2 \to 0$ to to see that

$$P\{\sum_{k=1}^{\infty} X^2(e_k) < \infty\} \geq 1 - 2\varepsilon.$$

Let $\Omega_1 = \{\omega: \sum_k X^2(e_k,\omega) < \infty\}$. Then $P\{\Omega_1\} = 1$. Define

$$Y(x,\omega) = \begin{cases} \sum_k \langle x,e_k\rangle_n X(e_k) & \text{if } \omega \in \Omega_1. \\ 0 & \text{if } \omega \in \Omega - \Omega_1. \end{cases}$$

The sum is finite by the Schwartz inequality, so Y is well-defined. Moreover, $Y \in H_{-n}$ with norm

$$\|Y\|_{-n} = \sum_k Y^2(e_k) = \sum_k X^2(e_k) < \infty$$

Finally, $P\{Y(x) = X(x)\} = 1$, $x \in E$. Indeed, let $x_N = \sum_{k=1}^{N} \langle x,e_k\rangle_n e_k$. Clearly $X(x_N) = Y(x_N)$ on Ω_1, and $\|x-x_N\|_m \leq \|x-x_N\|_n \to 0$. Thus

$$\begin{aligned} Y(x) &= \lim Y(x_N) \\ &= \lim X(x_N) = X(x). \end{aligned}$$

Note: We have followed some notes of Ito in this proof. The tricks are due to Sazanov and Yamazaki.

EXAMPLES

Let us see what the spaces E and H_n are in some special cases.

EXAMPLE 1. Let $G \subset \mathbb{R}^d$ be a bounded domain and let $E_0 = \underline{\underline{D}}(G)$ be the set of C^∞ functions of compact support in G. Let $\| \ \|_0$ be the usual L^2-norm on G and set

$$\|\phi\|_n^2 = \|\phi\|_0^2 + \sum_{1\leq|\alpha|\leq n} \|D^\alpha\phi\|_0^2,$$

where α is a multi-index of length $|\alpha|$, and D^α is the associated partial derivative operator. Let E be the completion of E in the topology induced by the norms $\| \ \|_n$.

In this case $H_0 = L^2(G)$ and H_n is the classical Sobolev space (often denoted $W_0^{n,2}(G)$). By Maurin's theorem, $\| \ \|_m \underset{HS}{<} \| \ \|_n$ if $n > m + d/2$. H_n consists of all L^2-functions whose partials of order n or less are all in L^2. (These are derivatives in the sense of distributions. However, if $n > d/2$ the functions will be continuous; for larger n, they will be differentiable in the usual sense, and $\bigcap_n H_n$ will consist of C^∞ functions.

The spaces H_{-n} - duals to the H_n - consist of derivatives: $f \in H_{-n}$ iff there exist $f_\alpha \in L^2$ such that

$$f = \sum_{|\alpha| \leq n} D^{\alpha} f_{\alpha}.$$

EXAMPLE 1a. If we let $G = \mathbf{R}^d$ in Example 1, we can use the Fourier transform to define the H_n in a rather neat way. Let $E = \underline{\underline{S}}(\mathbf{R}^d)$. If $u \in E$, define the Fourier transform $\hat{u}$ of u by

$$\hat{u}(\xi) = \int_{\mathbf{R}^d} e^{-2\pi i x \cdot \xi} u(x)\,dx.$$

If u is a tempered distribution, i.e. if $u \in \underline{\underline{S}}'(\mathbf{R}^d)$, we can define $\hat{u}$ - as a distribution - by $\hat{u}(\phi) = u(\hat{\phi})$, $\phi \in E$.

Define a norm on E by

$$\|u\|_t = \int_{\mathbf{R}^d} (1 + |\xi|^2)^t \, |\hat{u}(\xi)|^2 d\xi \tag{4.2}$$

and let H_t be the completion of E in the norm $\| \ \|_t$.

If u is a distribution whose Fourier transform $\hat{u}$ is a function, then $u \in H_t$ iff (4.2) is finite. The space $H_0 = L^2$ by Plancharel's theorem. For $t > 0$ the elements of H_n are functions. For $t < 0$ they are in general distributions. It can be shown that if t is an integer, say $t = n$, the norms $\| \ \|_n$ defined here and in Example 1 are equivalent, and the spaces H_n in the two examples are identical. Note that (4.2) makes sense for all real t, positive or negative, integer or not, and $\| \ \|_s \underset{HS}{<} \| \ \|_t$ if $t > s + d/2$.

EXAMPLE 2. Let $E = \underline{\underline{S}}(\mathbf{R}^d)$, the Schwartz space of rapidly decreasing functions. Let

$$g_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}$$

and set

$$h_k(x) = (\pi^{1/2} 2^k k!)^{-1/2} g_k(x) e^{-x^2/2}.$$

Then $g_0, g_1, \ldots$ are the Hermite polynomials, and $h_0, h_1, \ldots$ are the Hermite functions. The latter are a CONS in $L^2(\mathbf{R}^d)$.

Let $q = (q_1, \ldots, q_d)$ where the q_i are non-negative integers, and for $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$, set

$$h_q(x) = h_{q_1}(x_1) \cdots h_{q_d}(x_d).$$

Then $h_q \in \underline{\underline{S}}(\mathbf{R}^d)$, and they form a CONS in $L^2(\mathbf{R}^d)$. If $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$, let $\hat{\phi}_q = \langle\phi, h_q\rangle$ and write

$$\phi = \sum_q \hat{\phi}_q h_q$$

Define

$$\|\phi\|_n^2 = \sum_q (2|q|+d)^n \hat{\phi}_q^2$$

where $|q|^2 = q_1^2 + \cdots + q_d^2$. One can show $\|\phi\|_n < \infty$ if $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$. Let H_n be the completion of E in $\| \ \|_n$. Note that this makes sense for negative n - in fact for all real n - and

$$\|\phi\|_{-n}^2 = \sum_q (2|q| + d)^{-n} \hat{\phi}_q^2 \quad .$$

H_{-n} is dual to H_n under the inner product

$$\langle\psi, \phi\rangle = \sum_q \hat{\psi}_q \hat{\phi}_q \quad .$$

The Hilbert-Schmidt ordering is easily verified in this example, since the functions $e_q = (2|q|+d)^{-n/2} h_q$ are a CONS under $\| \ \|_n$, and if $m < n$

$$\sum_q \|e_q\|_m = \sum_q (2|q|+d)^{(n-m)}$$

which is finite if $n > m + d/2$. Thus $\| \ \|_m \underset{HS}{<} \| \ \|_n$ if $n > m + \frac{d}{2}$.

<u>EXAMPLE 3</u>. Let us look at an example which is specifically linked to a differential operator.

Let M be a smooth compact d-dimensional differentiable manifold with a smooth (possibly empty) boundary. Let dx be the element of area, and let L be a self-adjoint uniformly strongly elliptic second order differential operator with smooth coefficients, and smooth homogeneous boundary conditions.

$-L$ has a CONS of smooth eigenfunctions $\{\phi_n\}$ with eigenvalues $\{\lambda_n\}$. The eigenvalues satisfy $\sum_j (1+\lambda_j)^{-p} < \infty$ if $p > \frac{d}{2}$.

Let E_0 be the set of f of the form $f(x) = \sum_{j=1}^N \hat{f}_j \phi_j(x)$, where the $\hat{f}_j$ are constants. For each integer n, positive or negative, define

$$\|f\|_n = \sum_j (1 + \lambda_j)^n \hat{f}_j^2 .$$

Note that $\| \ \|_m \underset{HS}{<} \| \ \|_n$ if $n > m + d/2$. Indeed, set $e_j = (1 + \lambda_j)^{-n/2} \phi_j$. The e_j form a CONS relative to $\| \|_n$, and

$$\sum_j \|e_j\|_m^2 = \sum_j (1 + \lambda_j)^{m-n} < \infty.$$

Let H_n be the completion of $(E_0, \| \|_n)$. If $f \in H_n$, we can represent f by the formal series

$$f = \sum_j \hat{f}_j \phi_h,$$

where $\sum (1 + \lambda_j)^n \hat{f}_j^2 = \|f\|_n < \infty$. Then H_n and H_{-n} are dual under the inner product $\langle f,g \rangle = \sum_j \hat{f}_j \hat{g}_j$.

Finally, let E be the completion of E_0 in the topology determined by the $\| \|_n$. The ϕ_j are smooth, so that the elements of H_n will be differentiable for large n. Since $E = \bigcap_n H_n$, E will consist of C^∞ functions. Note that if $f \in H_n$, $Lf \in H_{n-2}$ since $Lf = \sum_j \lambda_j \hat{f}_j \phi_j$.

The similarity of Examples 2 and 3 is more than superficial. In fact Example 2 corresponds to the operator $L = -\Delta + |x|^2$.

<u>EXAMPLE 4</u>. At the start of the chapter we raised the question of whether or not a martingale measure could be regarded as a distribution. Let us consider this in the setting of, say, Example 1. Let G be a bounded open set in $\mathbf{R}^d$ with a smooth boundary, let $\underline{\underline{D}}(G)$ be the space of test functions on G, and let M be a worthy martingale measure on G with dominating measure K and let $\mu_t(A\times B) = E\{K(A\times B\times[0,t]\}$. Assume μ_t is finite.

If $\phi \to 0$ in $\underline{\underline{D}}(G)$, $\sup|\phi(x)| \to 0$, hence $E\{M_t^2(\phi)\} = \int \phi(x)\phi(y)\mu_t(dxdy) \to 0$. (Careful! This is not trivial; we have used Sobolev's Theorem.) It follows that $\phi \to M_t(\phi)$ is continuous in probability on $\underline{\underline{D}}(G)'$. By Corollary 4.2, M_t has a version with values in $\underline{\underline{D}}(G)$.

M actually lives in a Sobolev space of negative index. To see why, note that if $n > d/2$, H_n embeds in $C_b(G)$ by the Sobolev embedding theorem and $\| \|_n \underset{HS}{<} \| \|_{2n}$ by Maurin's theorem. By Theorem 4.1, M_t has a version with values in H_{-2n}. In particular, $M_t \in H_{-d-2}$, and if d is odd, we have $M_t \in H_{-d-1}$. (A more delicate analysis here would show that, locally at least, $M_t \in H_{-n}$ for any $n > d/2$.)

<u>Exercise 4.2</u>. Show that under the usual hypotheses (i.e. right continuous filtration, etc.) that the process M_t, considered as a process with values in $\underline{\underline{D}}(G)$, has a right continuous version. Show that it is also right continuous in the appropriate Sobolev space.

Even pathology has its degrees. The martingale measure M_t will certainly not be a differentiable or even continuous function, but it is not infinitely bad. According to the above, it is at worst a derivative of order d + 2 of an L^2 function or, using the embedding theorem again, a derivative of order $\frac{3}{2}$ d + 3 of a continuous function. Thus a distribution in H_{-n} is "more differentiable" than a distribution in H_{-n-1}, and the statement that M does indeed take values in a certain H_{-n} can be regarded as a regularity property of M.

In the future we will discuss most processes as having values in $\underline{\underline{D}}(G)'$ or another relevant nuclear space, and put off the task of deciding which H_{-n} is appropriate until we discuss the regularity of the process. As a practical matter, it is usually easier to do it this way; for it is often much simpler to verify that a process is distribution-valued than to verify it lives in a given Sobolov space...and as an even more practical matter, we shall usually leave even that to the reader.

CHAPTER FIVE

PARABOLIC EQUATIONS IN $\mathbf{R}^d$

Let $\{M_t, \underline{\underline{F}}_t, t \geq 0\}$ be a worthy martingale measure on $\mathbf{R}^d$ with covariation measure $Q(dx\ dy\ ds) = d\langle M(dx), M(dy)\rangle_s$ and dominating measure K. Let $\mu(\Lambda) = E\{K(\Lambda)\}$. Assume that for some $p > 0$ and all $T > 0$

$$\int_{\mathbf{R}^d \times [0,T]} \frac{1}{(1+|x|^p)(1+|y|^p)} \mu(dx\ dy\ ds) < \infty.$$

Then $M_t(\phi) = \int_{\mathbf{R}^d \times [0,T]} \phi(x) M(dx\ ds)$ exists for each $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$.

Let L be a uniformly elliptic self-adjoint second order differential operator with bounded smooth coefficients. Let T be a differential operator on $\mathbf{R}^d$ of finite order with bounded smooth coefficients. (Note that T and L operate on x, not on t). Consider the SPDE

$$\text{(5.1)} \qquad \begin{cases} \dfrac{\partial V}{\partial t} = LV + T\dot{M} \\ V(x,0) = 0 \end{cases}$$

We will clearly need to let V and M have distribution values, if only to make sense of the term $T\dot{M}$. We will suppose they have values in the Schwartz space $\underline{\underline{S}}'(\mathbf{R}^d)$.

We want to cover two situations: the first is the case in which (5.1) holds in $\mathbf{R}^d$. Although there are no boundary conditions as such, the fact that $V_t \in \underline{\underline{S}}'(\mathbf{R}^d)$ implies a boundedness condition at infinity.

The second is the case in which D is a bounded domain in $\mathbf{R}^d$, and homogeneous boundary conditions are imposed on ∂D.

(There is a third situation which is covered - formally at least - by (5.1), and that is the case where T is an integral operator rather than a differential operator. Suppose, for instance, that $Tf(x) = g(x) \int f(y)h(y)dy$ for suitable functions g and h. In that case, $TM_t(x) = g(x)M_t(h)$. Now $M_t(h)$ is a real-valued martingale, so that (5.1) can be rewritten

$$\begin{cases} dV_t = LV\ dt + g\,dM_t(h) \\ V(x,0) = 0 \end{cases}$$

This differs from (5.1) in that the driving term is a one-parameter martingale rather than a martingale measure. Its solutions have a radically different behavior from those of (5.1) and it deserves to be treated separately.)

Suppose (5.1) holds on $\mathbf{R}^d$. Integrate it against $\phi \varepsilon \underline{\underline{S}}(\mathbf{R}^d)$, and then integrate by parts. Let T^* be the formal adjoint of T. The weak form of (5.1) is then

$$(5.2) \qquad V_t(\phi) = \int_0^t V_s(L\phi)ds + \int_0^t \int_{\mathbf{R}^d} T^*\phi(x)M(dxds), \quad \phi \in \underline{\underline{S}}(\mathbf{R}^d).$$

Notice that when we integrate by parts, (5.2) follows easily for ϕ of compact support, but in order to pass to rapidly decreasing ϕ, we must use the fact that V and $T\dot{M}$ do not grow too quickly at infinity.

In case D is a bounded region with a smooth boundary, let B be the operator $B = d(x)D_N + e(x)$, where D_N is the normal derivative on ∂D, and d and e are in $C^\infty(\partial D)$. Consider the initial-boundary-value problem

$$(5.3) \qquad \begin{cases} \dfrac{\partial V}{\partial t} = LV + T\dot{M} & \text{on } D \times [0,\infty); \\ BV = 0 & \text{on } \partial D \times [0,\infty); \\ V(x,0) = 0 & \text{on } D. \end{cases}$$

Let $C^\infty(D)$ and $C_0^\infty(D)$ be respectively the set of smooth functions on D and the set of smooth functions with compact support in D. Let $C^\infty(\overline{D})$ be the set of functions in $C^\infty(D)$ whose derivatives all extend to continuous functions on $\overline{D}$. Finally, let

$$\underline{\underline{S}}_B = \{\phi \in C^\infty(\overline{D}) \colon B\phi = 0 \text{ on } \partial D\}.$$

The weak form of (5.3) is

$$(5.4) \qquad V_t(\phi) = \int_0^t V_s(L\phi)ds + \int_0^t \int_D T^*\phi(x)M(dxds), \quad \phi \in \underline{\underline{S}}_B.$$

This needs a word of explanation. To derive (5.4) from (5.3), multiply by ϕ and integrate formally over $D \times [0,t]$ - i.e. treat $T\dot{M}$ as if it were a differentiable function - and then use a form of Green's theorem to throw the derivatives over on ϕ. This works on the first integral if both V and ϕ satisfy the boundary condition. Unless T is of zeroth order, it may not work for the second, for $\dot{M}$ may not satisfy the boundary conditions. (It does work if ϕ has compact support in D, however.) Nevertheless, <u>the equation we wish to solve is (5.4), not (5.3).</u>

The requirement that (5.4) hold for all ϕ satisfying the boundary conditions is essentially a boundary condition on V.

The above situation, in which we regard the integral, rather than the differential equation as fundamental, is analogous to many situations in which physical reasoning leads one directly to an integral equation, and then mathematics takes over to extract the partial differential equation. See the physicists' derivations of the heat equation, Navier-Stokes equation, and Maxwell's equation, for instance.

As in the one-variable case, it is possible to treat test functions $\phi(x,t)$ of two variables.

Exercise 5.1. Show that if V satisfies (5.4) and if $\phi(x,t)$ is a smooth function such that for each t, $\phi(\cdot,t) \in \underline{\underline{S}}_B$, then

$$V_t(\phi(t)) = \int_0^t V_s(L\phi(s) + \frac{\partial\phi}{\partial s}(s))ds + \int_0^t \int_D T^*\phi(x,s)M(dxds). \tag{5.5}$$

Let $G_t(x,y)$ be the Green's function for the homogeneous differential equation. If $L = \frac{1}{2}\Delta$, $D = \mathbf{R}^d$, then

$$G_t(x,y) = (2\pi t)^{-d/2} e^{-\frac{|y-x|^2}{2t}}.$$

For a general L, $G_t(x,y)$ will still be smooth except at $t = 0$, $x = y$, and its smoothness even extends to the boundary: if $t > 0$, $G_t(x,\cdot) \in C^\infty(\overline{D})$. It is positive, and for $\tau > 0$,

$$G_t(x,y) \leq Ct^{-d/2} e^{-\frac{|y-x|^2}{\delta t}}, \quad x, y \in D, \ 0 \leq t \leq \tau, \tag{5.6}$$

where $C > 0$ and $\delta > 0$. (C may depend on τ). This holds both for $D = \mathbf{R}^d$ and for bounded D. If $D = \mathbf{R}^d$, $G_t(x,\cdot)$ is rapidly decreasing at infinity by (5.6), so it is in $\underline{\underline{S}}(\mathbf{R}^d)$. Moreover, for fixed y, $(x,t) \to G_t(x,y)$ satisfies the homogeneous differential equation plus boundary conditions. Define $G_t(\phi,y) = \int_D G_t(x,y)\phi(x)dx$.

Then if ϕ is smooth, $G_o\phi = \phi$. This can be summarized in the integral equation:

$$G_{t-s}(\phi,y) = \phi(y) + \int_s^t G_{u-s}(L\phi,y)du, \quad \phi \in \underline{\underline{S}}_B \tag{5.7}$$

The smoothness of G then implies that if $\phi \in C^\infty(\overline{D})$, then $G_t(\phi,\cdot) \in \underline{\underline{S}}_B$. In case $D = R^d$, then $\phi \in \underline{\underline{S}}(R^d)$ implies that $G_t(\phi,\cdot) \in \underline{\underline{S}}(R^d)$.

THEOREM 5.1. There exists a unique process $\{V_t, t>0\}$ with values in $\underline{\underline{S}}'(R^d)$ which satisfies (5.4). It is given by

$$V_t(\phi) = \int_0^t \int_{R^d} T^* G_{t-s}(\phi,y) M(dyds). \tag{5.8}$$

The result for a bounded region is similar except for the uniqueness statement.

THEOREM 5.2 There exists a process $\{V_t, t \geq 0\}$ with values in $\underline{\underline{S}}'(R^d)$ which satisfies (5.5). V can be extended to a stochastic process $\{V_t(\phi), t \geq 0, \phi \in \underline{\underline{S}}_B\}$; this process is unique. It is given by

$$V_t(\phi) = \int_0^t \int_D T^* G_{t-s}(\phi,y) M(dy\ ds), \ \phi \in \underline{\underline{S}}_B. \tag{5.9}$$

PROOF. Let us first show uniqueness, which we do by deriving (5.9).

Choose $\psi(x,s) = G_{t-s}(\phi,x)$, and suppose that U is a solution of (5.4). Consider $U_s(\psi(s))$. Note that $U_0(\psi(0)) = 0$ and $U_t(\psi(t)) = U_t(\phi)$. Now $G_{t-s}(\phi,\cdot) \in \underline{\underline{S}}_B$, so we can apply (5.5) to see that

$$U_t(\phi) = U_t(\psi(t)) = \int_0^t U_s(L\psi(s) + \frac{\partial \psi}{\partial s}(s))ds + \int_0^t \int_D T^* \psi(x,s) M(dx\ ds).$$

But $L\psi + \frac{\partial \psi}{\partial s} = 0$ by (5.7) so this is

$$= \int_0^t \int_D T^* \ \psi(x,s)\ M(dx\ ds)$$

$$= \int_0^t \int_D T^* G_{t-s}(\phi,x) M(dx\ ds) = V_t(\phi).$$

Existence: Let $\phi \in \underline{\underline{S}}_B$ and plug (5.9) into the right hand side of (5.4):

$$\int_0^t \Big[\int_0^s \int_D T^* G_{s-u}(L\phi,y)\ M(dy\ du)\Big] ds + \int_0^t \int_D T^* \phi(y) M(dy,du)$$

$$= \int_0^t \int_D \Big[\int_u^t T^* G_{s-u}(L\phi,y) ds + T^* \phi(y)\Big] M(dy\ du).$$

Note that $T^* G_{s-u}(L\phi,y)$ and $T^*\phi(y)$ are bounded, so the integrals exist. By (5.7) this is

$$= \int_0^t \int_D T^* G_{t-u}(\phi,y) M(dy\ du)$$

$$= V_t(\phi).$$

by (5.9). This holds for any $\phi \in \underline{\underline{S}}_B$, but (5.9) also makes sense for ϕ which are not in $\underline{\underline{S}}_B$. In particular, it makes sense for $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$ and one can show using Corollary 4.2 that V_t has a version which is a random tempered distribution.

This proves Theorem 5.2. The proof of Theorem 5.1 is nearly identical; just replace D by $\mathbf{R}^d$ and $\underline{\underline{S}}_B$ by $\underline{\underline{S}}(\mathbf{R}^d)$. Q.E.D.

AN EIGENFUNCTION EXPANSION

We can learn a lot from an examination of the of the case $T \equiv 1$. Suppose D is a bounded domain with a smooth boundary. The operator $-L$ (plus boundary conditions) admits a CONS $\{\phi_j\}$ of smooth eigenfunctions with eigenvalues λ_j. These satisfy

(5.10) $$\sum_j (1+\lambda_j)^{-p} < \infty \quad \text{if } p > d/2.$$

(5.11) $$\sup_j \|\phi_j\|_\infty^2 (1+\lambda_j)^{-p} < \infty \quad \text{if } p > d/2.$$

Let us proceed formally for the moment. We can expand the Green's function:

$$G_t(x,y) = \sum_j \phi_j(x)\phi_j(y)e^{-\lambda_j t}.$$

If ψ is a test function

$$G_t(\psi,y) = \sum_j \hat{\psi}_j \phi_j(y) e^{-\lambda_j t}$$

where $\hat{\psi}_j = \int_D \psi(x)\phi_j(x)dx$, so by (5.9)

$$V_t(\psi) = \int_0^t \int_D \sum_j \hat{\psi}_j \phi_j(y) e^{-\lambda_j (t-s)} M(dyds).$$

Let

$$A_j(t) = \int_0^t \int_D \phi_j(y) e^{-\lambda_j (t-s)} M(dyds).$$

Then

$$V_t(\phi) = \sum \hat{\phi}_j A_j(t). \tag{5.12}$$

This will converge for $\phi \in \underline{\underline{S}}_B$, but we will show more. Let us recall the spaces H_n introduced in Ch. 4, Example 3. H_n is isomorphic to the set of formal eigenfunction series

$$f = \sum_{j=1}^{\infty} a_j \phi_j$$

for which

$$\|f\|_n^2 = \sum a_j^2 (1+\lambda_j)^n < \infty.$$

We see from (5.12) that $V_t \sim \sum_j A_j(t)\phi_j$.

<u>PROPOSITION 5.3</u>. Let V be defined by (5.12). If $n > d$, V_t is a right continuous process in H_{-n}; it is continuous if $t \to M_t$ is. Moreover, V is the solution of (5.4) with $T \equiv 1$. If M is a white noise based on Lebesgue measure then V is a continuous process in H_{-n} for any $n > d/2$.

<u>PROOF</u>. We first bound $E\{\sup_{t\le\tau} A_j^2(t)\}$. Let $X_j(t) = \int_0^t \int_D \phi_j(x)\, M(dxds)$ and note that, as in Exercise 3.3,

$$\begin{aligned} A_j(t) &= \int_0^t e^{-\lambda_j(t-s)}\, dX_j(s). \\ &= X_t - \int_0^t \lambda_j e^{-\lambda_j(t-s)} X_j(s)ds \end{aligned}$$

where we have integrated by parts in the stochastic integral. Thus

$$\begin{aligned} \sup_{t\le\tau}|A_j(t)| &\le \sup_{t\le\tau} |X_j(t)|(1+\int_0^t \lambda_j e^{-\lambda_j(t-s)} ds) \\ &\le 2 \sup_{t\le\tau} |X_j(t)|. \end{aligned}$$

Thus

$$\begin{aligned} E\{\sup_{t\le\tau} A_j^2(t)\} &\le 4E\{\sup X_j^2(t)\} \\ &\le 16E\{X_j^2(\tau)\} \end{aligned}$$

by Doob's inequality. This is

$$(5.13) \qquad = 16 \int_{D\times D\times[0,\tau]} \phi_j(x)\phi_j(y)\mu(dx\,dy\,ds)$$
$$\leq 16\ \mu(D\times D\times[0,\tau])\|\phi_j\|_\infty^2$$
$$\leq C(1+\lambda_j)^p$$

for some constant C and $p > d/2$ by (5.11).

Thus

$$E\{\sum_j \sup_{t\leq\tau} A_j^2(t)(1+\lambda_j)^{-n}\} \leq C\sum_j (1+\lambda_j)^{-n+p}.$$

By (5.10) this is finite if $n - p > d/2$ or, remembering that $p > d/2$, if $n > d$.

Then, clearly,

$$\|V_t\|_{-n}^2 = \sum A_j^2(t)(1+\lambda_j)^{-n}$$

is a.s. finite, hence $V_t \in H_{-n}$. Moreover, if $s > 0$,

$$\|V_{t+s}-V_t\|_{-n}^2 = \sum (A_j(t+s)-A_j(t))^2(1+\lambda_j)^{-n}.$$

The summands are right continuous, and they are continuous if M is. The sum is dominated by

$$4\sum_j \sup_{t\leq\tau} A_j^2(t)(1+\lambda_j)^{-n} < \infty$$

for a.e.ω. Now $A_j(s) \to A_j(t)$ as $s \downarrow t$, hence $\|V_s-V_t\|_{-n} \to 0$ as $s \downarrow t$. If M is continuous, so is A_j, and we can let $s \uparrow t$ to see V is also left continuous, hence continuous.

If M is a white noise the integral in (5.13) reduces to $\int \phi_j^2(x)dxds$. Since the ϕ_j are orthonormal this is just τ. This means we can take $p = 0$ and $n > d/2$ in the remainder of the argument.

<u>REMARKS</u>. The conditions on the Sobolev spaces in Proposition 2.3 can be improved. For instance, if M is a white noise based on Lebesgue measure, V will be continuous in H_{-n} for every $n > d/2 - 1$, not just for $n > d/2$. The same proof shows it, once one improves the estimate of $E\{\sup_{t\leq\tau} A_j^2(t)\}$. In this case, A_j is an $OU(\lambda_j)$ process and one can show that this quantity is bounded by a constant times $\lambda_j^{-1}\log \lambda_j$.

Once we know that V_t is actually a solution of (5.4), the uniqueness result implies that V also satisfies (5.9).

Exercise 5.2. Verify that V (defined by (5.12)) satisfies (5.4).

Exercise 5.3. Treat the case $D = R^d$ using the Hermite expansion of Example 2, Ch.4.

The spaces H_n above are analogous to the classical Sobolev spaces, but they don't explicitly involve derivatives. Here is a result which relates the regularity of the solution directly to differentiability.

THEOREM 5.4. Suppose M is a white noise based on Lebesgue measure. Then there exists a real-valued process $U = \{U(x,t): x \in D, t \geq 0\}$ which is Hölder continuous with exponent $1/4 - \varepsilon$ for any $\varepsilon > 0$ such that if $D^{d-1} = \frac{\partial^{d-1}}{\partial x_2, \ldots, \partial x_d}$, then

$$V_t = D^{d-1} U_t.$$

Note. This is of course a derivative in the weak sense. A distribution Q is the weak αth derivative of a function f if for each test function ϕ,

$$Q(\phi) = (-1)^{|\alpha|} \int f(x) D^\alpha \phi(x)\,dx.$$

If we let H_0^n denote the classical Sobolev space of Example 1 Ch.4, this implies that V_t can be regarded as a continuous process in H_0^{-d+1}.

Note. One must be careful in comparing the classical Sobolev spaces H_0^n with related but not identical spaces H_n of Example 3 in Chapter 4. Call the latter H_3^n for the moment. Theorem 5.4 might lead one to guess that V is in H_3^{-d+1}, but in fact, it is a continuous process in H_3^{-n} for any $n > d/2$ by Proposition 5.3. This is a much sharper result if $d \geq 3$.

This gives us an idea of the behavior of the solution of the equation

$$\frac{\partial V}{\partial t} = LV + \dot{M}.$$

Suppose now that T is a differential operator and suppose both T and L have constant coefficients, so TL = LT. Apply T to both sides of the SPDE:

$$\frac{\partial}{\partial t}(TV) = LTV + T\dot{M}$$

i.e. U = TV satisfies $\frac{\partial U}{\partial t} = LU + T\dot{M}$. Of course, this argument is purely formal, but the following exercise makes it rigorous.

Exercise 5.4. Suppose T and L commute. Let U be the solution of (5.4) for a general T with bounded smooth coefficients and let V be the solution for $T \equiv 1$. Verify that if we restrict U and V to the space $\underline{\underline{D}}(D)$ of C^∞ functions of compact support in D, that U = TV.

Exercise 5.5. Let V solve

$$\begin{cases} \frac{\partial V}{\partial t} = \frac{\partial^2 V}{\partial x^2} - V + \frac{\partial}{\partial x} \dot{W}, & 0 < x < \pi,\ t > 0; \\ \frac{\partial V}{\partial x}(0,t) = \frac{\partial V}{\partial x}(\pi,t) = 0, & t > 0; \\ V(x,0) = 0, & 0 < x < \pi. \end{cases}$$

Describe $V(\cdot,t)$ for fixed t. (Hint: use Exercises 5.4 and 3.5.)

REMARKS. 1. Theorem 5.2 lacks symmetry compared to Theorem 5.1. V_t exists as a process in $\underline{\underline{S}}(R^d)$ but must be extended slightly to get uniqueness, and this extension doesn't take values in $\underline{\underline{S}}'(R^d)$. It would be nicer to have a more symmetric statement, on the order of "There exists a unique process with values in such and such a space such that ...". One can get such a statement, though it requires a litle more Sobolev space theory and a little more analysis to do it. Here is how.

Let $\| \ \|_n$ be the norm of Example 1, Chapter 4. Let H_B^n be the completion of $\underline{\underline{S}}_B$ in this norm. If n is large enough, one can show that V_t is an element of $(H_B^n)' \overset{\text{def}}{=} H_B^{-n}$. Theorem 5.2 can then be stated in the form: there exists a unique process V with values in H_B^{-n} which satisfies (5.4) for all $\phi \in H_B^n$.

2. Suppose that T is the identity and consider (5.4). Extend V to be a distribution on $D \times R_+$ as follows. If $\phi = \phi(x,t)$ is in $C_0^\infty(D \times (0,\infty))$, let

$$V(\phi) = \int_0^\infty V_s(\phi(s))ds \quad \text{and} \quad T\dot{M}(\phi) = \int_0^\infty \int_D T^* \phi(x,s)\, M(dxds).$$

Then Corollary 4.2 implies that for a.e. ω, V and $T\dot{M}$ define distributions on $D \times (0,\infty)$. Now consider (5.5). For large t, the left-hand side vanishes, for ϕ has compact support. The right-hand side then tells us that $V(L\phi + \frac{\partial \phi}{\partial s}) + TM(\phi) = 0$ a.s. In other words, for a.e. ω, the distribution $V(\cdot,\omega)$ is a distribution solution of the (non-stochastic) PDE

$$\frac{\partial \mu}{\partial t} - L\mu = T\dot{M} .$$

Thus Theorem 5.1 follows from known non-stochastic theorems on PDE's. If T is the identity, the same holds for Theorem 5.2. In general, the translation of (5.4) or (5.5) into a PDE will introduce boundary terms. Still, we should keep in mind that the theory of distribution solutions of deterministic PDE's has something to say about SPDE's.

CHAPTER SIX

WEAK CONVERGENCE

Suppose E is a metric space with metric ρ. Let $\underline{\underline{E}}$ be the class of Borel sets on E, and let (P_n) be a sequence of probability measures on $\underline{\underline{E}}$. What do we really mean by "$P_n \to P_o$"? This is a non-mathematical question, of course. It is asking us to make an intuitive idea precise. Since our intuition will depend on the context, it has no unique answer. Still, we might begin with a reasonable first approximation, see how it might be improved, and hope that our intuition agrees with our mathematics at the end.

Suppose we say:

$$\text{"}P_n \to P_o \text{ if } P_n(A) \to P_o(A), \text{ all } A \in \underline{\underline{E}}.\text{"}$$

This looks promising, but it is too strong. Some sequences which should converge, don't. For instance, consider

PROBLEM 1. Let $P_n = \delta_{1/n}$, the unit mass at $1/n$, and let $P_o = \delta_o$. Certainly P_n ought to converge to P_o, but it doesn't. Indeed $0 = \lim P_n\{0\} \neq P_o\{0\} = 1$. Similar things happen with sets like $(-\infty,0]$ and $(0,1)$.

CURE. The trouble occurs at the boundary of the sets, so let us smooth them out. Identify a set A with its indicator function I_A. Then $P(A) = \int I_A dP$. We "smooth out the boundary of A" by replacing I_A by a continuous function f which approximates it, and ask that $\int f dP_n \to \int f dP$. We may as well require this for all f, not just those which approximate indicator functions.

This leads us to the following. Let C(E) be the set of bounded real valued continuous functions on E.

DEFINITION. We say P_n converges weakly to P, and write $P_n \Rightarrow P$, if, for all $f \in C(E)$,

$$\int f dP_n \to \int f dP.$$

PROBLEM 2. Our notion of convergence seems unconnected with a topology.

CURE. We prescribe two definitions:

$$\underline{\underline{P}}(E) \overset{\text{def}}{=} \{P: P \text{ is a probability measure on } \underline{\underline{E}}\}.$$

A fundamental system of neighborhoods is given by sets of the form

$$\{P \in \underline{\underline{P}}(E): |\int f_i dP - \int f_i dP_o| < \varepsilon,\ i=1,\ldots,n\},\quad f_i \in C(E),\ i = 1,\ldots,n.$$

This notion of convergence may not appear to fill our needs - for we shall be discussing convergence of processes, rather than of random variables - but it is in fact exactly what we need. The reason why it is sufficient is itself extremely interesting, and we shall go into it shortly, but let us first establish some facts.

The first, which gives a number of equivalent characterizations of weak convergence, is sometimes called the Portmanteau Theorem.

THEOREM 6.1. The following are equivalent

(i) $P_n \Rightarrow P$;

(ii) $\int f dP_n \to f dP$, all bounded uniformly continous f;

(iii) $\int f dP_n \to \int f dP$, all bounded functions which are continuous, P-a.e.;

(iv) $\lim\sup P_n(F) \leq P(F)$, all closed F;

(v) $\lim\inf P_n(G) \geq P(G)$, all open G;

(vi) $\lim P_n(A) = P(A)$, all $A \in \underline{\underline{E}}$ such that $P(\partial A) = 0$.

Let E and F be metric spaces and $h : E \to F$ a measurable map. If P is a probability measure on E, then Ph^{-1} is a probability measure on F, where $Ph^{-1}(A) = P(h^{-1}(A))$.

THEOREM 6.2 If $h : E \to F$ is continuous (or just continuous P-a.e.) and if $P_n \Rightarrow P$ on E, then $P_n h^{-1} \Rightarrow Ph^{-1}$ on F.

Let $P_1, P_2, \ldots$ be a sequence in $\underline{\underline{P}}(E)$. When does such a sequence converge? Here is one answer. Say that a set $K \subset \underline{\underline{P}}(E)$ is relatively compact if every sequence in A has a weakly convergent subsequence. (This should be "relatively sequentially compact," but we follow the common usage.)

Then (P_n) converges weakly if

(i) there exists a relatively compact set $K \subset \underline{\underline{P}}(E)$ such that $P_n \in K$ for all n.

(ii) the sequence has at most one limit point in $\underline{\underline{P}}(E)$.

Since (i) guarantees at least one limit point, (i) and (ii) together imply convergence.

If this condition is to be useful - and it is - we will need an effective criterion for relative compactness. This is supplied by Prohorov's Theorem.

DEFINITION. A set $A \subset \underline{\underline{P}}(E)$ is tight if for each $\varepsilon > 0$ there exists a compact set $K \subset E$ such that for each $P \in A$, $P\{K\} > 1 - \varepsilon$.

THEOREM 6.3. If A is tight, it is relatively compact. Conversely, if E is separable and complete, then if A is relatively compact, it is tight.

Let us return to the question of the suitability of our definition of weak convergence.

PROBLEM 3. We are interested in the behavior of processes, not random variables, so this all seems irrelevant.

CURE. We already know the solution to this. We just have to stand back far enough to recognize it. We often define a process canonically on a function space: if Ω is a space of, say, right continuous functions on $[0,\infty)$, then a process $\{X_t: t>0\}$ can be defined on Ω by $X_t(\omega) = \omega(t)$, $\omega \in \Omega$, for ω, being an element of Ω, is itself a function. X is then determined by its distribution P, which is a measure on Ω. But this means that we are regarding the whole process as a single random variable. The random variable simply takes its values in a space of functions.

With this remark, the outline of the theory becomes clear. We must first put a metric on the function space Ω in some convenient way. The above definition will then apply to measures on Ω.

The Skorokhod space $\underline{\underline{D}} = \underline{\underline{D}}([0,1],E)$ is a convenient function space to use. It is the space of all functions $f : [0,1] \to E$ which are right-continuous and have left limits at each $t \in (0,1]$. We will metrize $\underline{\underline{D}}$. The metric is a bit tricky. It is much like a sup-norm, but the presence of jump discontinuities forces a modification.

First, let Λ be the class of strictly increasing, continuous maps of $[0,1]$ onto itself. If $\lambda \in \Lambda$, then $\lambda(0) = 0$ and $\lambda(1) = 1$. Define

$$\|\lambda\| = \sup_{0 \le s < t \le 1} \left|\log \frac{\lambda(t)-\lambda(s)}{t-s}\right|, \quad \lambda \in \Lambda.$$

(We may have $\|\lambda\| = \infty$. We don't worry about that.) If $\|\lambda\|$ is small, λ must be close to the identity.

Next we define a distance on $\underline{\underline{D}}$ by

$$d_o(f,g) = \inf\{\|\lambda\| + \sup_t \rho(f(t),g(\lambda(t))) : \lambda \in \Lambda\}.$$

The functions λ should be considered as time-changes. The reason we need them can be seen by considering $f(t) = I_{[0,1/2+\varepsilon]}(t)$ and $g(t) = I_{[0,1/2]}$ t). Both have a single jump of size one, and if ε is small, the jumps nearly coincide, and we would like $d(f,g)$ to be small. Note that $\sup_t |f(t)-g(t)| = 1$ however. The time-change allows us to move the jump of g to coincide with that of f. After doing this, we see that $\sup_t |f(t)-g(\lambda(t))|$ vanishes. (For an exercise, let $\lambda(t) = (1+4\varepsilon)t - 4\varepsilon t^2$ and show that $d(f,g) \le 4\varepsilon/1-4\varepsilon$.)

Let us collect a few miscellaneous facts.

THEOREM 6.4. (i) d_o is a metric on $\underline{\underline{D}}$. If E is a complete separable metric space, so is $\underline{\underline{D}}$.

(ii) $d_o(f_n,f) \to 0$ iff there exist $\lambda_n \in \Lambda$ such that $\|\lambda_n(t)-t\|_\infty \to 0$ and

$$\sup_t \rho(f(t), f_n(\lambda_n(t)))_\infty \to 0.$$

(iii) $C([0,1],E)$ is a closed subspace of $\underline{\underline{D}}$.

We have to be able to characterize compact sets in $\underline{\underline{D}}$ if we want to apply Prohorov's Theorem. Remembering the Arzela-Ascoli Theorem, this should have something to do with equicontinuity. Let us introduce a "modulus of continuity" which is tailored for right continuous processes.

DEFINITION. For $f \in \underline{\underline{D}}$, let

$$w(\delta,f) = \inf_{\{t_i\}} \max_i \sup_{t_i \le s<t<t_{i+1}} \rho(f(s),f(t)), \tag{6.1}$$

where the infimum is over all finite partitions $0 = t_0 < t_1 < \cdots < t_n = 1$ such that $t_i - t_{i-1} \geq \delta$, for all $i \leq n-1$.

This discounts the large jumps of f, for one can place the partition points there. The effect is to find the modulus of continuity between the jumps. If f is continuous, however, this reduces to the ordinary modulus of continuity.

The counterpart of the Arzéla-Ascoli theorem is:

THEOREM 6.5. (Arzéla-Ascoli theorem for $\underline{\underline{D}}$). Let E be a complete separable metric space. A set A has compact closure in $\underline{\underline{D}}$ iff

(i) for each rational $t \in [0,1]$ there is a compact set $K_t \subset E$ such that if $f \in A$, then $f(t) \in K_t$, all $t \in \mathbb{Q} \cap [0,1]$;

(ii) $\lim\limits_{\delta \to 0} \sup\limits_{f \in A} w(f,\delta) = 0$.

Note: If E is locally compact, the compact sets K_t of (i) can be chosen to be independent of t.

Let $(P_n) \subset \underline{\underline{P}}(D)$ be weakly convergent. In order to identify its limit, one often checks the convergence of the finite-dimensional distributions. If $h(x_1,\ldots,x_n)$ is a bounded continuous function on $E \times \cdots \times E$, and if $0 \leq t_0 \leq \cdots \leq t_n \leq 1$, define a function H on $\underline{\underline{D}}$ by

$$H(\omega) = h(\omega(t_1),\ldots,\omega(t_n)), \quad \omega \in \underline{\underline{D}}.$$

We say that the finite dimensional distributions converge if, for each n and each $0 \leq t_1 \leq \cdots \leq t_n \leq 1$, there exists a measure $\mu_{t_1 \ldots t_n}$ on $E \times \cdots \times E$ such that for each such h and H

$$\int H dP_n \to \int h d\mu_{t_1 \ldots t_n}.$$

PROPOSITION 6.6. A sequence $(P_n) \subset \underline{\underline{P}}(D)$ converges weakly iff

(i) (P_n) is tight;

(ii) the finite-dimensional distributions converge.

All of this is convenient to describe in the language of processes. If $X_n = \{X_n(t),\ 0\le t\le 1\}$ is a sequence of processes, we say that (X_n) converges weakly if the corresponding distributions (P_n) on $\underline{\underline{D}}$ converge weakly. Similarly, (X_n) is tight if the (P_n) are.

To show the weak convergence of (X_n), we must show (Prop. 6.5) that

(a) (X_n) is tight;

(b) the finite-dimensional distributions of the X_n converge weakly.

Of the two, tightness is often the most difficult to show. It is useful to have easily-checkable criteria. The following theorem, due primarily to Aldous, gives two criteria which are useful in case there are martingales present.

ALDOUS' THEOREM

Let E be a complete separable metric space and let $\underline{\underline{D}} = \underline{\underline{D}}([0,1],E)$. Let ρ be a bounded metric on E. Let $(\underline{\underline{F}}_t)$ be the canonical filtration on $\underline{\underline{D}}$, i.e. $\underline{\underline{F}}_t = \sigma\{\omega(s),\ s\le t,\ \omega\in\underline{\underline{D}}\}$. Let $\underline{\underline{T}}$ be the class of finite-valued stopping times T such that $T \le 1$.

If X is a process defined canonically on $\underline{\underline{D}}$, let

$$\mu(\delta,X) = \sup_{T\in\underline{\underline{T}}} E\{\rho(X_{T+\delta},X_T)\}$$

$$\nu(\delta,X) = \sup_{\alpha\le\delta} \mu(\alpha,X).$$

Results such as Prohorov's Theorem can be extended to some non-metrizable spaces. Here is one such extension due to Le Cam.

THEOREM 6.7. Let E be a competely regular topological space such that all compact sets are metrizable. If (P^n) is a sequence of probability measures on E which is tight, then there exists a subsequence (n_k) and a probability measure Q such that $P^{n_k} \Rightarrow Q$.

THEOREM 6.8. Let (X_n) be a sequence of processes with paths in $\underline{\underline{D}}$. Suppose that for each rational $t \in [0,1]$ the family of random variables $\{X_n(t),\ n=1,2,\dots\}$ is tight.

Then either of the following conditions implies that (X_n) is tight in $\underline{\underline{D}}$.

(a) (Aldous). For every sequence (T_n,δ_n) where $T_n \in \underline{\underline{T}}$ and $\delta_n \to 0$, $\delta_n > 0$,

$$\rho(X_n(T_n+\delta_n),\ X_n(T_n)) \to 0 \text{ in probability.}$$

(b) (Kurtz). There exists $p > 0$ and processes $\{A_n(\delta), 0<\delta<1\}$, $n = 1,2,\ldots$ such that

$$\text{(i)} \quad E\{\rho(X_n(t+\delta),\ X_n(t))^p \mid \underline{\underline{F}}_t\} \le E\{A_n(\delta) \mid \underline{\underline{F}}_t\}$$

and

$$\text{(ii)} \quad \lim_{\delta\to 0} \limsup_{n\to\infty} E\{A_n(\delta)\} = 0.$$

We will follow a proof of T. Kurtz. Most of the work is in establishing the following three lemmas.

<u>LEMMA 6.9</u>. (a) is equivalent to

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \nu(\delta,X_n) = 0 . \tag{6.2}$$

<u>PROOF</u>. If (6.2) holds, (a) follows on noticing that $\delta \to \nu(\delta,X_n)$ is increasing for each n, and that $E\{\rho(X_{T_n+\delta_n},\ X_{T_n})\} \le \nu(\delta_n,X_n)$.

Conversely, if (6.2) does not hold, there is an $\varepsilon > 0$ such that for each n_o and $\delta_o > 0$ there exists $n \ge n_o$ and $\delta_n \le \delta_o$ such that $\nu(\delta_n,\ X_n) > \varepsilon/2$. Then $E\{\rho(X_n(T_n+\delta_n'),X(T_n))\} > \varepsilon/2$ for some $T_n \in \underline{\underline{T}}$ and $\delta_n' \le \delta_n$. Since ρ is bounded, this implies (a) does not hold. Q.E.D.

<u>LEMMA 6.10</u>. If $T_1 \le T_2 \in \underline{\underline{T}}$ and $T_2 - T_1 \le \delta$, then

$$\text{(i)} \quad E\{\rho(X_{T_1},X_{T_2})\} \le \frac{2}{\delta}\int_0^{2\delta} \mu(u,X)du \le 4\nu(2\delta,X);$$

$$\text{(ii)} \quad \nu(\delta,X) \le \frac{2}{\delta}\int_0^{2\delta} \mu(u,X)du.$$

<u>PROOF</u>. Clearly (i)=>(ii) (let $T_2=T_1+\alpha$). To prove (i), use the triangle inequality:

$$\rho(X_{T_1},X_{T_2}) \le \frac{1}{\delta}\int_0^{\delta} (\rho(X_{T_1},X_{T_2+v}) + \rho(X_{T_2+v},X_{T_2}))dv .$$

if $v \le \delta$, $T_2 + v = T_1 + u$ for some $u \le 2\delta$:

$$\le \frac{1}{\delta}\int_0^{2\delta} \rho(X_{T_1},X_{T_1+u})du + \frac{1}{\delta}\int_0^{\delta} \rho(X_{T_2+v},X_{T_2})dv$$

Take expectations of both sides to see that

$$E\{\rho(X_{T_1},X_{T_2})\} \le + \frac{1}{\delta}\int_0^{2\delta}\mu(u,X)du + \frac{1}{\delta}\int_0^{\delta}\mu(v,X)dv$$

which implies (i). Q.E.D.

LEMMA 6.11. (a) implies that

(6.3) $$\lim_{\delta\to 0}\limsup_{n\to\infty} E\{w(\delta,X_n)\} = 0 .$$

PROOF. Fix n, $\varepsilon > 0$ and $0 < \delta < 1$. Set $S_o = 0$ and define

$$S_{k+1} = \inf\{t>S_k : \rho(X_t,X_{S_k}) \ge \varepsilon\} \wedge 1$$

If $t \in [S_k,S_{k+1})$, $\rho(X_t,X_{S_k}) < \varepsilon$ so $\rho(X_s,X_t) < 2\varepsilon$ if $s,t \in [S_k,S_{k+1})$. If the (S_k) form a partition of mesh $\ge\delta$, $w(\delta,X) \le 2\varepsilon$.

Fix K and notice that if $S_K = 1$, there is some $J < K$ for which $S_J < 1 = S_{J+1}$, and if $S_k - S_{k-1} > \delta$ for all $k \le J$, $\{S_o,\ldots,S_j,1\}$ forms a partition of mesh $\ge\delta$, hence $w(\delta,X) \le 2\varepsilon$. In any case, $w(\delta,X) \le 1$. Thus, a very rough estimate of $w(\delta,X)$ is

$$w(\delta,X) \le 2\varepsilon + I_{\{S_k-S_{k-1}\le \delta \text{ some } k\le J,\ J\le K;\ S_K=1\}}.$$

Let $T_k = S_{k+1} \wedge (S_k+\delta)$:

$$\le 2\varepsilon + \sum_{k=0}^{K-1}\frac{1}{\varepsilon}\rho(X_{T_k},X_{S_k}) + I_{\{S_K<1\}}$$

where we use the fact that if $S_{k+1} - S_k < \delta$, $T_k = S_{k+1}$ and $\rho(X_{T_k},X_{S_k}) \ge \varepsilon$ so the sum is $\ge \varepsilon$.

Thus

$$E\{w(\delta,X_n)\} \le 2\varepsilon + \frac{1}{\varepsilon}\sum_{k=0}^{K-1}E\{\rho(X_n(T_k),X_n(S_k))\} + P\{S_K<1\}$$
$$\le 2\varepsilon + \frac{4K}{\varepsilon}\nu_n(2\delta,X_n) + P\{S_K<1\}$$

by Lemma 6.10.

The final probability is independent of δ, and it converges to zero as $K \to \infty$. In fact, the convergence is even uniform in n. To see this, fix δ_o and note

$$P\{S_K<1\} = P\{e^{1-S_K} > 1\} \le E\{e^{1-S_K}\}$$

$$\leq e\, E\{e^{-\sum_1^n (S_{k+1}-S_k)}\}$$

$$= e\, E\{\prod_{k=0}^{K-1} e^{-(S_{k+1}-S_k)}\}$$

$$\leq e \prod_{k=0}^{K-1} E\{e^{-K(S_{k+1}-S_k)}\}^{1/K}$$

by Hölder's inequality. The integrand is bounded by $e^{-K\delta_o}$ on $\{S_{k+1}-S_k>\delta_o\}$ and by one otherwise, so this is

$$\leq e \prod_{k=0}^{K-1} (e^{-K\delta_0} + P\{S_{k+1}-S_k<\delta_0\})^{1/K}$$

$$\leq e \prod_{k=0}^{K-1} (e^{-K\delta_o} + \frac{4}{\varepsilon}\nu(2\delta_0,X_n))^{1/K}$$

by Lemma 6.10.

Thus

$$E\{w(\delta,X_n)\} \leq 2\varepsilon + \frac{4K}{\varepsilon}\nu(2\delta,X_n) + e^{1-K\delta_0} + \frac{4e}{2\varepsilon}\nu(2\delta_0,X_n).$$

Let $n \to \infty$ and $\delta \to 0$ in that order and use (a). Then let $K \to \infty$, $\delta_0 \to 0$ and finally $\varepsilon \to 0$ to get (6.3).

PROOF (of Theorem 6.8). We first show that (b) => (6.2), which implies (a) by Lemma 6.9, and then we show that (a) implies tightness.

If $0 \leq \alpha \leq \delta$ and $p > 0$

$$E\{\rho^p(X_n(T+\alpha),X_n(T))\} \leq E\{E\{\rho^p(X_n(T+\alpha),X_n(T))|\underline{\underline{F}}_T\}\}$$

$$\leq E\{E\{A_n(\alpha)|\underline{\underline{F}}_T\}$$

$$= E\{A_n(\alpha)\}$$

hence $\mu(\alpha,X_n) \leq E\{A_n(\alpha)\}$. By the lemma

$$\nu(\delta,X_n) \leq \frac{2}{\delta}\int_0^{2\delta} E\{A_n(u)\}du .$$

Let $n \to \infty$ and $\delta \to 0$ and use Fatou's lemma and (ii) to see that (6.2) holds.

To prove tightness, note that $w(\delta,X_n) \downarrow 0$ as $\delta \downarrow 0$, so Lemma 6.11 implies

$$\limsup_{\delta\to 0}\ \sup_n E\{w(\delta,X_n)\} = 0 .$$

Let (t_j) be an ordering of $\mathbb{Q} \cap [0,1]$ and let $\varepsilon \to 0$. For each j there is a compact $K_j \subset E$ such that

$$P\{X_n(t_j) \in K_j\} > 1 - \varepsilon/2^{j+1}.$$

Choose $\delta_k \downarrow 0$ such that $\sup_n E\{w(\delta_k,X_n)\} \le \frac{\varepsilon}{k2^{k+1}}$. Thus $\sup_n P\{w(\delta_k,X_n) > \frac{1}{k}\} \le \varepsilon/2^{k+1}$. Let $\Lambda \subset \underline{\underline{D}}$ be

$$\Lambda = \{\omega \in \underline{\underline{D}} : \omega(t_k) \in K_k,\ w(\omega,\delta_k) \le \frac{1}{k},\quad k=1,2,\ldots\}.$$

Now $\lim_{\delta\to 0} \sup_{\omega\in\Lambda} w(\delta,\omega) = 0$. Thus Λ has a compact closure in $\underline{\underline{D}}$ by Theorem 6.5. Moreover

$$\begin{aligned} P\{X_n \in \Lambda\} &\ge 1 - \sum_k P\{X_n(t_k) \in K_k\} \\ &\quad - \sum_k P\{w(\delta_k,X_n) > 1/k\} \\ &\ge 1 - \varepsilon/2 - \varepsilon/2 = 1 - \varepsilon, \end{aligned}$$

hence (X_n) is tight.

MITOMA'S THEOREM

The subject of SPDE's involves distributions in a fundamental way. We will need to know about the weak convergence of processes with values in $\underline{\underline{S}}'$. Since $\underline{\underline{S}}'$ is not metrizable, the preceeding theory does not apply directly.

However, weak convergence of distribution-valued processes is almost as simple as that of real-valued processes. According to a theorem of Mitoma, in order to show that a sequence (X^n) of processes tight, one merely needs to verify that for each ϕ, the real-valued processes $(X^n(\phi))$ are tight.

Rather than restrict ourselves to $\underline{\underline{S}}'$, we will use the somewhat more general setting of Chapter Four. Let

$$E' = \bigcup_n H_n \supset \cdots \supset H_{-1} \supset H_0 \supset H_1 \supset \cdots \supset \bigcap_n H_n = E$$

where H_n is a separable Hilbert space with norm $\|\ \|_n$, E is dense in each H_n, $\|\ \|_n \le \|\ \|_{n+1}$ and for each n there is a $p > n$ such that $\|\ \|_n \underset{HS}{<} \|\ \|_p$. E has the topology determined by the norms $\|\ \|_n$, and E' has the strong topology which is determined by the semi norms

$$p_A(f) = \sup\{|f(\phi)|,\ \phi \in A\}$$

where A is a bounded set in E.

Let $\underline{\underline{D}}([0,1],E')$ be the space of E'-valued right continuous functions which have left limits in E', and let $C([0,1],E')$ be the space of continuous E'-valued functions. $C([0,1],H_n)$ and $\underline{\underline{D}}([0,1],H_n)$ are the corresponding spaces of H_n-valued functions.

If $f,g \in \underline{\underline{D}}([0,1],E')$, let

$$d_A(f,g) = \inf\{\|\lambda\| + \sup_t p_A(f(t)-g(\lambda(t))),\ \lambda \in \Lambda\},$$

and

$$\bar{d}_A(f,g) = \sup_t p_A(f(t)-g(t)).$$

Give $\underline{\underline{D}}([0,1],E')$ (resp. $C([0,1],E')$) the topology determined by the d_A (resp. $\bar{d}_A$) for bounded A E. They both become complete, separable, completely regular spaces. The $\underline{\underline{D}}([0,1],H_n)$ have already been defined, for H_n is a metric space.

We will need two "moduli of continuity". For $\omega \in \underline{\underline{D}}([0,1],E')$, $\phi \in E$, set

$$w(\delta,\omega;\phi) = \inf_{\{t_i\}} \max_i \sup_{t_i \le s < t < t_{i+1}} |\langle \omega(t) - \omega(s), \phi\rangle|$$

where the infimum is over all finite partitions $0 = t_0 < t_1 < \cdots < t_n = 1$ such that $t_i - t_{i-1} \geq \delta$ for $i = 1,\ldots,n-1$.

Similarly, for $\omega \in \underline{\underline{D}}([0,1],H_n)$, let

$$w_n(\delta,\omega) = \inf_{\{t_i\}} \max_i \sup_{t_i \le s < t < t_{i+1}} \|\omega(t)-\omega(s)\|_n .$$

We will define w and w_n on $C([0,1],E')$ as the ordinary moduli of continuity.

There is a metatheorem which says that anything that happens in E' already happens in one of the H_n. The regularity theorem of Chapter Four is one instance of this. Here is another.

<u>THEOREM 6.12</u>. Let A be compact in $\underline{\underline{D}} = \underline{\underline{D}}([0,1],E')$ (resp. $C([0,1],E')$). Then there is an n such that A is compact in $\underline{\underline{D}}([0,1],H_{-n})$ (resp. $C([0,1],H_{-n})$.

<u>PROOF</u>. We will only prove this for $C([0,1],E')$. The proof for $\underline{\underline{D}}$ involves the same ideas, but is considerably more technical.

Let $\phi \in E$ and map $C([0,1],E') \to C([0,1],\mathbb{R})$ by $\omega \to \{\langle\omega(t),\phi\rangle : 0 \le t \le 1\}$. It is easy to see that this map is continuous, so that the image of A is compact in $C([0,1],\mathbb{R})$. By the Arzela-Ascoli Theorem (see Theorem 6.5)

(i) $\sup_{\omega \varepsilon A} \sup_{t \varepsilon [0,1]} |\langle\omega(t),\phi\rangle| < \infty$, and

(ii) $\lim_{\delta \to 0} \sup_{\omega \varepsilon A} w(\delta,\omega,\phi) = 0.$

From (i), we see that the linear functionals $F_{\omega t}$ defined by $F_{\omega t}(\phi) = \langle\omega(t),\phi\rangle$ are bounded at each ϕ, and hence equicontinuous by the Banach-Steinhaus theorem. Thus there is a neighborhood V of zero such that $\phi \varepsilon V \Rightarrow |F_{\omega t}(\phi)| < 1$, all $\omega \varepsilon A$, $t \varepsilon [0,1]$. V contains a basis element, say $\{\phi : \|\phi\|_m < \varepsilon\}$.

Thus if $c = 1/\varepsilon$,

$$|\langle\omega(t),\phi\rangle| \leq c\|\phi\|_m, \quad \text{all } \phi \varepsilon E. \tag{6.4}$$

There exists $p > m$ such that, if (e_j) is a CONS relative to $\| \ \|_p$, $\sum_j \|e_j\|_m^2 = \ell < \infty$. Then

$$\sup_{\omega \varepsilon A} \sup_t \|\omega(t)\|_{-p}^2 = \sup_{\omega \varepsilon A} \sup_t \sum_j \langle\omega(t),e_j\rangle^2 \leq c^2\ell.$$

Set $K = \{\omega \varepsilon H_{-p} : \|\omega\|_{-p} \leq c^2\ell\}$. There exists $n > p$ such that $\| \ \|_p \underset{HS}{<} \| \ \|_n$, so K will be compact in H_{-n} by Exercise 4.1. Thus

(i') $\omega \varepsilon A \Rightarrow \omega \varepsilon K$ and K is compact in $C([0,1],H_{-n})$.

Moreover, (6.4) implies that

$$\sup_{\omega \varepsilon A} w(\delta,\omega,e_j) \leq 2c\|e_j\|_m,$$

so

$$\lim_{\delta \to 0} \sup_{\omega \varepsilon A} w_{-m}(\delta,\omega) = \lim_{\delta \to 0} \sup_{\substack{0 \leq s < t \leq 1 \\ |t-s| \leq \delta}} \Big(\sum_j \langle\omega(t)-\omega(s),e_j\rangle^2\Big)^{1/2}.$$

The sum is bounded by $2c^2\sum_j \|e_j\|_m^2 < \infty$, so that it converges uniformly in s, t and δ, and it is

$$\leq \lim_{\delta \to 0} \Big(\sum_j \sup_{\omega \varepsilon A} w(\delta,\omega,e_j)^2\Big)^{1/2}$$

$$= \Big(\sum_j \lim_{\delta \to 0} \sup_{\omega \varepsilon A} w(\delta,\omega,e_j)^2\Big)^{1/2}$$

$$= 0.$$

Since $\| \ \|_{-n} \leq \| \ \|_{-m}$,

(ii') $\lim_{\delta \to 0} \sup_{\omega \varepsilon A} w_{-n}(\delta,\omega) = 0$

Thus A is compact in $C([0,1],H_{-n})$ by the Arzela-Ascoli theorem. Q.E.D.

THEOREM 6.13 (Mitoma). Let $\{X_t^n, 0 \leq t \leq 1\}$, $n = 1,2,\ldots$ be a sequence of processes whose sample paths are in $\underline{D}([0,1],E')$ a.s. Then the sequence (X^n) is tight iff for each $\phi \in E$, the sequence of real-valued processes $\{X_t^n(\phi), t \geq 0\}$ $n = 1,2,\ldots$ is tight in $D([0,1], R)$.

All the work in the proof comes in establishing the following lemma.

LEMMA 6.14. Suppose $(X^n(\phi))$ is tight for each $\phi \in E$. Then for each $\varepsilon > 0$ there exist p and $M > 0$ such that

$$\sup_n P\{\sup_{0 \leq t \leq 1} \|X_t^n\|_{-p} > M\} < \varepsilon.$$

PROOF. We will do this in stages.

1) Let $\varepsilon > 0$. We claim that there exist an m and $\delta > 0$ such that

$$\|\phi\|_m < \delta \Rightarrow \sup_n |\!|\!| \sup_t |X_t^n(\phi)| |\!|\!| < \varepsilon \tag{6.5}$$

where $|\!|\!| X |\!|\!| = E\{|X|_\wedge 1\}$.

To see this, consider the function

$$F(\phi) = \sup_n |\!|\!| \sup_t X_t(\phi) |\!|\!| , \quad \phi \in E.$$

Then

(i) $F(0) = 0$;

(ii) $F(\phi) \geq 0$ and $F(\phi) = F(-\phi)$;

(iii) $|a| < |b| \Rightarrow F(a\phi) \leq F(b\phi)$;

(iv) F is lower-semi-continuous on E;

(v) $\lim_{n\to\infty} F(\phi/n) = 0$.

Indeed (i)-(iii) are clear. If $\phi_j \to \phi$ in E, $X_t^n(\phi_j) \to X_t^n(\phi)$ in L^0, hence $|X_t^n(\phi_j)|\wedge 1 \to |X_t^n(\phi)|\wedge 1$ in probability, and $\liminf_j [\sup_t |X_t^n(\phi_j)|\wedge 1] \geq \sup_t |X_t^n(\phi_j)|\wedge 1$ a.s. Thus

$$\begin{aligned} F(\phi) = \sup_n E\{\sup_t |X_t^n(\phi)|\wedge 1\} &\leq \sup_n \liminf_j E\{\sup_t |X_t^n(\phi)|\wedge 1\} \\ &\leq \liminf_j \sup_n E\{\sup_t |X_t^n(\phi)|\wedge 1\} \\ &= \liminf_j F(\phi_j), \end{aligned}$$

proving (iv).

To see (v), note that $(X_t^n(\phi))$ is tight, so, given ϕ and $\varepsilon > 0$ there exists an M such that $P\{\sup_t |X_t^n(\phi)| > M\} < \varepsilon/2$.

Choose k large enough so that $M/k < \varepsilon/2$. Then

$$\begin{aligned} F(\phi/k) &= \sup_n E\{\sup_t |X_t^n(\phi/k)| \wedge 1\} \\ &\leq \sup_n [P\{\sup_t |X_t^n(\phi/k)| > M\} + \tfrac{M}{k}] \\ &< \varepsilon. \end{aligned}$$

Let $V = \{\phi: F(\phi) \leq \varepsilon\}$. V is a closed (by (iv)), symmetric (by (ii)), absorbing (by (v)) set. We claim it is a neighborhood of 0. Indeed, $E = \bigcup_n nV$, so by the Baire category theorem, one, hence all, of the nV must have a non-empty interior. In particular, $\frac{1}{2}V$ does. Then $V \subset \frac{1}{2}V - \frac{1}{2}V$ must contain a neighborhood of zero. This in turn must contain an element of the basis, say $\{\phi: \|\phi\|_m < \delta\}$. This proves (6.5).

The next stages of the argument use the same techniques used in proving Theorem 4.1. (We called them "tricks" there. Now we are using them a second time, we call them "techniques".) However, the presence of the supremum over t makes this proof more delicate than the other.

2) We claim that for all n

$$\text{(6.6)} \qquad \operatorname{Re} E\{\sup_t (1-e^{iX_t^n(\phi)})\} \leq 2\varepsilon(1 + \|\phi\|_m^2/\delta^2), \quad \phi \in E.$$

Indeed, $\operatorname{Re}(1-e^{iX_t^n(\phi)}) = 1 - \cos X_t^n(\phi) \leq \frac{1}{2}(X_t(\phi)^2 \wedge 2)$. If $\|\phi\|_m < \delta$, then

$$\begin{aligned} \operatorname{Re} E\{\sup_t (1-e^{iX_t^n(\phi)})\} &\leq \tfrac{1}{2} E\{\sup_t X_t^n(\phi)^2 \wedge 2\} \\ &\leq 2E\{\sup_t |X_t^n(\phi)| \wedge 1|\} \\ &= 2|\!|\!|\sup_t |X_t^n(\phi)|\,|\!|\!| \leq 2\varepsilon. \end{aligned}$$

On the other hand, if $\|\phi\|_m \geq \delta$, replace ϕ by $\psi = \delta\phi/\|\phi\|_m$:

$$\begin{aligned} \operatorname{Re} E\{1-e^{iX_t(\phi)}\} &\leq \tfrac{1}{2} E\{\frac{\|\phi\|_m^2}{\delta^2} X_t^n(\psi)^2 \wedge 2\} \\ &\leq \tfrac{1}{2} \frac{\|\phi\|^2}{\delta^2} E\{X_t^n(\psi)^2 \wedge 2\} \\ &\leq 2\varepsilon \|\phi\|_m^2/\delta^2 \end{aligned}$$

since $\|\phi\|_m \le \delta$.

3) We claim that for $M > 0$ there is $p \ge m$ and constant K such that for all n,

(6.7) $$E\{1 - e^{-1/M \sup_t \|X_t^n\|_{-p}^2}\} \le 2\varepsilon(1 + K/M\delta^2).$$

Indeed, there exists $p > m$ such that $\| \ \|_m \underset{HS}{<} \| \ \|_p$. Let (e_j) be a CONS in E, relative to $\| \ \|_n$. Let $Y_1, Y_2, \ldots$ be iid $N(0, \frac{2}{M})$ random variables, and put $\phi = \sum_1^N Y_j e_j$.

$$\operatorname{Re} E\{\sup_t (1 - e^{iX_t^n(\phi)})\} = \operatorname{Re} E\{E\{\sup_t (1 - e^{iX_t^n(\phi)}) | Y\}\}$$
$$\le E\{2\varepsilon(1 + \|\phi\|_m^2/\delta^2)\}.$$

But $E\{\|\phi\|_m^2\} = E\{\sum_1^N Y_j^2 \|e_j\|_m^2\} \le \frac{2}{M} \sum_1^\infty \|e_j\|_m^2 \overset{def}{=} \frac{K}{M} < \infty$. On the other hand,

$$\operatorname{Re} E\{\sup_t (1 - e^{iX_t^n(\phi)})\} = \operatorname{Re} E\{\sup_t (1 - e^{i \sum_1^N Y_j X_t^n(e_j)})\}$$
$$\ge \operatorname{Re} E\{\sup_t E\{(1 - e^{i\sum_1^N Y_j X_t(e_j)}) | X_t\}\}.$$

But, given X_t, the conditional distribution of $\sum Y_j X_t(e_j)$ is $N(0, \frac{2}{M} \sum_1^N X_j^n(e_j)^2)$. We know the characteristic function of a normal random variable, so we see this is

$$= E\{\sup_t (1 - e^{-\frac{1}{M} \sum_1^N X_t^n(e_j)^2})\}.$$

Let $N \to \infty$. $\sum X_t^n(e_j)^2 = \|X_t^n\|_{-p}$, so we can combine these inequalities to get (6.7).

4) If $\sup_t \|X_t\|_{-p}^2 > M$, $1 - e^{-\frac{1}{M} \sup_t \|X_t\|_{-p}^2} > \frac{e-1}{e}$, so

$$\frac{e-1}{e} P\{\sup_t \|X_t\|_{-p}^2 > M\} \le E\{1 - e^{-\frac{1}{M} \sup_t \|X_t\|_{-p}^2}\}$$
$$\le 2\varepsilon(1 + K/M\delta^2).$$

If $M \ge K/\delta^2$, then

$$P\{\sup_t \|X_t\|_{-p}^2 > M\} \le \frac{4e}{e-1} \varepsilon.$$

Q.E.D.

PROOF (of Theorem 6.13). Fix $\varepsilon > 0$. Choose M and p as in Lemma 6.14. With probability $1-\varepsilon$, X_t lies in $B = \{x: \|x\|_{-p} \le M\}$ for all t. There exists $q > p$ such that $\| \|_p \underset{HS}{<} \| \|_q$, hence $\| \|_{-q} \underset{HS}{<} \| \|_{-p}$. Then A is compact in H_{-q}. Let $K \subset \underline{\underline{D}}([0,1],E')$ be the set $\{\omega : \omega(t) \in A, 0 \le t \le 1\}$.

Let (e_j) be a CONS relative to $\| \|_q$. Since $(X^n(e_j))$ is tight by hypothesis, there exists a compact set $K_j \subset \underline{\underline{D}}([0,1], R)$ such that $P\{X^n_\bullet(e_j) \in K_j\} \ge 1 - \varepsilon/2^j$ for all n. Let K_j' be the inverse image of K_j in $\underline{\underline{D}}([0,1],E')$ under the map $\omega \to \{\langle\omega(t),e_j\rangle : 0 \le t \le 1\}$. By the Arzela-Ascoli Theorem,

$$\lim_{\delta\to 0} \sup_{\omega\in K_j'} w(\delta,\omega;e_j) = 0;$$

moreover

$$P\{X^n_\bullet \in K_j'\} \ge 1 - \varepsilon/2^j.$$

Set $K' = K \cap \bigcap_j K_j'$. Then

$$P\{X^n_\bullet \in K'\} \ge 1 - \varepsilon - \sum \varepsilon/2^j = 1 - 2\varepsilon.$$

Now,

$$\lim_{\delta\to 0} \sup_{\omega\in K'} w(\delta,\omega,H_{-q}) = \lim_{\delta\to 0} \sup_{\omega\in K'} \Big(\sum_j \inf_{\{t_i\}} \max_i \sup_{t_i\le s<t<t_{i+1}} \langle\omega(t)-\omega(s),e_j\rangle^2\Big)^{1/2}$$
$$\le \lim_{\delta\to 0} \sup_{\omega\in K'} \Big(\sum_j w(\delta,\omega,e_j)^2\Big)^{1/2}.$$

But if $\omega \in K'$, $\|\omega(t)\|_{-p} \le M$ so the sum is dominated by $2M \sum_j \|e_j\|_p^2 < \infty$ (since $\| \|_p \underset{HS}{<} \| \|_q$). Thus we can go to the limit inside the sum. It is

$$\le \Big(\sum_j \lim_{\delta\to 0} \sup_{\omega\in K'} w(\delta,\omega;e_j)^2\Big)^{1/2}$$
$$= 0.$$

A is compact in H_{-q}, so Theorem 6.5 tells us that K' is relatively compact in $\underline{\underline{D}}([0,1],H_{-q})$. The inclusion map of H_{-q} into E' is continuous, so that K' is also relatively compact in $\underline{\underline{D}}([0,1],E')$, and hence (X^n) is tight.

Q.E.D.

This brings us to the convergence theorem.

THEOREM 6.15. Let (X^n) be a sequence of processes with paths in $\underline{\underline{D}}([0,1],E')$. Suppose

(i) for each $\phi \in E$, $(X^n(\phi))$ is tight;

(ii) for each $\phi_1,\ldots,\phi_p$ in E and $t_1,\ldots,t_p \in [0,1]$, the distribution of $(X^n_{t_1}(\phi_1),\ldots,X^n_{t_p}(\phi_p))$ converges weakly on $\mathbf{R}^p$.

Then there exists a process X^o with paths in $\underline{\underline{D}}([0,1],E')$ such that $X^n \Rightarrow X^o$.

PROOF. (X^n) is tight by Theorem 6.13. The space $\underline{\underline{D}}([0,1],E')$ is completely regular and each compact subset is in some H_{-n} and is therefore metrizable. By Theorem 6.7, some subsequence converges weakly. But (ii) shows that there is only one possible limit. We conclude by the usual argument that the whole sequence converges.

Q.E.D.

Note: The index p of Lemma 6.14 may depend on ε. If it does not, and if $\| \ \|_p \underset{HS}{<} \| \ \|_q$, then (X^n) then will be tight in $\underline{\underline{D}}([0,1],H_{-q})$, and we get the following.

COROLLARY 6.16. Suppose that the hypotheses of Theorem 6.15 hold. Let $p < q$ and suppose $\| \ \|_p \underset{HS}{<} \| \ \|_q$. Suppose that for $\varepsilon > 0$ and $M > 0$ there exists $\delta > 0$ such that for all n

$$P\{\sup_t |\langle X^n_t,\phi\rangle| > M\} \leq \varepsilon \text{ if } \|\phi\|_p \leq \delta.$$

Then (X^n) converges weakly in $\underline{\underline{D}}([0,1],H_{-q})$.

CHAPTER SEVEN

APPLICATIONS OF WEAK CONVERGENCE

Does the weak convergence of a sequence of martingale measures imply the weak convergence of their stochastic integrals? That is, if $M^n \Rightarrow M$, does $f\cdot M^n \Rightarrow f\cdot M$? Moreover, do the convolution integrals - which give the solutions of SPDE's - also converge?

We will give the beginnings of the answers to these questions in this chapter. We will show that the answer to both is yes, _if_ one is willing to impose strong hypotheses on the integrands. Luckily these conditions are satisfied in many cases of interest.

We will confine ourselves to measures on $\mathbf{R}^d$ and on sub-domains of $\mathbf{R}^d$, where we have already discussed the theory of distributions. We will view martingale measures as distribution-valued processes, so that weak convergence means convergence in distribution on the Skorokhod space $\underline{\underline{D}} = \underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(\mathbf{R}^d)\}$.

Our martingale measures may have infinite total mass, but we will require that they not blow up too rapidly at infinity.

Fix $p_0 > 0$ and define $h_0(x) = (1 + |x|^{p_0})^{-1}$, $x \in \mathbf{R}^d$. If M is a worthy martingale measure with dominating measure K, define an increasing process k by

$$k(t) = \int_{\mathbf{R}^{2d}\times[0,t]} h_0(x)\, h_0(y)\, K(dx\, dy\, ds), \tag{7.1}$$

and

$$\gamma(\delta) = \sup_{t\leq 1} (k(t+\delta) - k(t)) \tag{7.2}$$

We will assume throughout this chapter that $E\{k(1)\} < \infty$. Note that this means that for any $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$, $M_t(\phi)$ is defined for all $t \leq 1$, since ϕ tends to zero at infinity faster than h_0. Thus M_t is a tempered distribution (Corollary 4.2).

For a function f on $\mathbf{R}^d$, define

$$\|f\|_\infty = \sup |f(x)|,$$
$$\|f\|_h = \|f\, h_0^{-1}\|_\infty .$$

Note that $\|\phi\|_h < \infty$ for any $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$, and, moreover, if $\|f\|_h < \infty$, then $M_t(f)$ is defined.

LEMMA 7.1. Let $S \leq T$ be stopping times for M, and let $f \in \underline{\underline{P}}_M$. Then

(7.3) $$\langle f\cdot M(E)\rangle_T - \langle f\cdot M(E)\rangle_S \leq \int_S^T \|f(s)\|_h^2 \, dk(s).$$

Consequently, if $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$,

(7.4) $$\langle f\cdot M(\phi)\rangle_T - \langle f\cdot M(\phi)\rangle_S \leq \|\phi\|_h^2 \int_S^T \|f(s)\|_\infty^2 \, dk(s).$$

PROOF. The left-hand side of (7.3) is

$$\int_{\mathbf{R}^{2d}\times(S,T]} f(x,s)f(y,s)K(dx\,dy\,ds) = \int_{\mathbf{R}^{2d}\times(S,T]} f(x,s)h_0^{-1}(x)f(y,s)h_0^{-1}(y)h_0(x)h_0(y)K(dx\,dy\,ds)$$
$$\leq \int_{\mathbf{R}^{2d}\times(S,T]} \|f(s)\|_h^2 \, h_0(x)h_0(y)\, K(dx\,dy\,ds)$$
$$= \int_S^T \|f(s)\|_h^2 \, dk(s).$$

Then (7.4) follows since $\|f(s)\phi\|_\infty \leq \|\phi\|_h \|f(s)\|_\infty$. Q.E.D.

LEMMA 7.2. Let T be a predictable stopping time. If k is a.s. continuous at T then for any bounded Borel set $A \subset \mathbf{R}^d$ and $f \in \underline{\underline{P}}_M$, $P\{f\cdot M(A)$ is continuous at $t\} = 1$.

PROOF. The graph $[T]$ of T is predictable, hence so is $f(x,t)\, I_{[T]}(t)I_A(x)$. By (7.3), if f is bounded

$$E\{(f\cdot M_T(A) - f\cdot M_{T-}(A))^2\} \leq E\{\int \|f(t)I_{[T]}(t)\|_h^2 \, dk(t)\}$$
$$= E\{\|f(T)\|_h^2(k(T) - k(T-))\}$$
$$= 0.$$

The result follows for all $f \in \underline{\underline{P}}_M$ by approximation. Q.E.D.

Let $M^0, M^1, M^2, \ldots$ be a sequence of worthy martingale measures and let k_n and γ_n, $n = 0,1,2,\ldots$ be the corresponding quantities defined in (7.1) and (7.2).

PROPOSITION 7.3. Suppose that

(7.5) $$\lim_{\delta\to 0} \limsup_{n\to\infty} E\{\gamma_n(\delta)\} = 0.$$

Then the sequence (M^n) is tight on $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(\mathbf{R}^d)\}$.

PROOF. By Mitoma's Theorem, we need only show that $(M^n(\phi))$ is tight for each $\phi \in \underline{\underline{S}}'(R^d)$. By (7.4)

$$E\{(M^n_{t+\delta}(\phi) - M^n_t(\phi))^2|\underline{\underline{F}}_t\} \le \|\phi\|_h^2 \, E\{k_n(t+\delta) - k_n(t)|\underline{\underline{F}}_t\}$$
$$\le \|\phi\|_h^2 \, E\{\gamma_n(\delta)|\underline{\underline{F}}_t\}.$$

The result now follows from Theorem 6.8(b). Q.E.D.

COROLLARY 7.4. Suppose (M^n) is a sequence of worthy martingale measures satisfying (7.5). Let $f_n \in \underline{\underline{P}}_{M^n}$ be a sequence of uniformly bounded functions. Then $(f_n \cdot M^n)$ is tight.

PROOF. $$K_{f_n \cdot M^n}(dx\ dy\ ds) = |f_n(x,s)f_n(y,s)| \ K_{N^n}(dx\ dy\ ds)$$
$$\le b^2 \ K_{M^n}(dx\ dy\ ds)$$

where b is the uniform bound for the f_n. Then the $f_n \cdot M^n$ satisfy (7.5) and the result follows from Proposition 7.3. Q.E.D.

WEAK CONVERGENCE OF STOCHASTIC INTEGRALS

In order to talk of the convergence of a sequence of stochastic integrals $f \cdot M^n$, we must be able to define the integrand f for each of the M^n. We can do this by defining all of the M^n on the same probability space. The most convenient spate for this is the Skorokhod space $\underline{\underline{D}}$. Thus we will define all our martingale measures canonically on $\underline{\underline{D}}$, so that once we define f(x,t,w) on $\mathbf{R}^d \times \mathbf{R}^+ \times \underline{\underline{D}}$, we can define all the $f \cdot M^n$.

The stochastic integral is not in general a continuous function on $\underline{\underline{D}}$, so that it is not always true that $M^n \Rightarrow M$ implies that $f \cdot M^n \Rightarrow f \cdot M$, even for classical martingales. Two examples, both of real valued martingales, will illustrate some of the pitfalls.

Example 7.1. Define $M^n_t = \begin{cases} 0 & \text{if } t < 1 + 1/n \\ X & \text{if } t \ge 1 + 1/n \end{cases}$ and $M_t = \begin{cases} 0 & \text{if } t < 1 \\ X & \text{if } t \ge 1 \end{cases}$, where $X = \pm 1$

with probability 1/2 each. Let $f(t) = I_{[0,1]}(t)$. Then $M^n \Rightarrow M$, but $f \cdot M^n \equiv 0$, $n = 1,2,\ldots$ while $f \cdot M_t = X\, I_{[1,\infty)}(t)$, so the integrals don't converge to the right limit.

<u>Example 7.2</u>. Let B_t be a standard real-valued Brownian motion from zero and put $M_t = M_0 + B_t$ and $M_t^n = M_0^n + B_t$, where M_0 is a random variable uniformly distributed on $[0,1]$, independent of B, and M_0^n is uniformly distributed on $\{1/n, 2/n, \ldots, 1\}$. Define these canonically on $\underline{\underline{D}}\{[0,1],\mathbb{R}\}$, and let $f(t,\omega) = I_{\{\omega(0) \in \mathbb{Q}\}}$ (i.e. $f \equiv 1$ or $f \equiv 0$ depending on whether the initial value of the martingale is rational or not.) Then $M^n \Rightarrow M$, but $f \cdot M^n = B_t$ for all n while $f \cdot M = 0$ a.s., so once again the integrals don't converge to the right limit.

<u>REMARKS</u> 1. In Example 7.1, the integrand was deterministic, and the trouble came from the jumps of M. In Example 7.2, the integrand was simply a badly discontinuous function of ω on $\underline{\underline{D}}$. Although it might seem that the trouble in Example 7.2 comes because the distribution of M is orthogonal to the distribution of the M^n, one can modify it slightly to make the distributions of the M^n and M all equivalent, and still get the same result.

2. It is easy to get examples in which the $f \cdot M^n$ do not just converge to the wrong limit but fail to converge entirely. Just replace every second M^n in either of the examples above by the limit martingale M.

Let M be a martingale measure defined canonically on $\underline{\underline{D}}$, relative to a probability measure P. Let k be defined by (7.1). Let $\mathcal{P}_S(M)$ be the class of functions f on $\mathbb{R}^d \times \mathbb{R}_+ \times \underline{\underline{D}}$ of the form

$$f(x,t,\omega) = \sum_{n=1}^{N} a_n(\omega) I_{(s_n,t_n]}(t)\phi_n(x), \tag{7.6}$$

where $0 \le s_n < t_n$, $\phi_n \in C^\infty(\overline{\mathbb{R}}^d)$, and a_n is bounded and $\underline{\underline{F}}_{s_n}$-measurable, $n = 1,\ldots,N$, such that

(i) a_n is continuous P-a.s. on $\underline{\underline{D}}$;

(ii) $t \rightsquigarrow f(x,t,\omega)$ is continuous at each point of discontinuity of k.

If f is given by (7.6) and if $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$ and $\omega \in \underline{\underline{D}}$, define

$$F_t(\omega)(\phi) = \sum_{n=1}^{N} a_n(\omega) I_{(s_n, t_n]}(t) \, (\omega_{t \wedge t_n}(\phi_n \phi) - \omega_{t \wedge s_n}(\phi_n \phi)).$$

Note that $\omega_t(\cdot)$ is a distribution and $\phi_n \phi \in \underline{\underline{S}}(\mathbf{R}^d)$, so $\omega_t(\phi_n \phi)$ is defined. The map $\omega \to \{F_t(\omega),\ t \geq 0\}$ maps $\underline{\underline{D}} \to \underline{\underline{D}}$. It is continuous at ω if ω is a point of continuity of each of the a_n and if $t \to \omega_t$ is itself continuous at each of the s_n and t_n. Thus if $f \in \mathcal{P}_S(M)$, f is P -a.e. continuous in ω. Moreover, if M is canonically defined on $\underline{\underline{D}}$, then

$$F_t(\omega)(\phi) = f\cdot M_t(\phi).$$

By the continuity theorem (Theorem 6.1 (iii)) we have:

<u>PROPOSITION 7.5</u>. Let $M^0, M^1, \ldots$ be a sequence of worthy martingale measures such that $M^n \Rightarrow M^0$. If $f \in \mathcal{P}_S(M^0)$ then $(M^n, f\cdot M^n) \Rightarrow (M^0, f\cdot M^0)$.

This is too restrictive to be of any real use, so we must extend the class of f. We will do this by approximating more general f by f in $\mathcal{P}_S(M^0)$ and using the fact that L^2 convergence implies convergence in distribution.

The class of f we can treat depends on the sequence (M^n), and, in particular, on the sequence (K_n) of dominating measures. The more we are willing to assume about the K_n, the better description we can give of the class of f. We will start with minimal assumptions on the K_n. This will allow us to give a simple treatment which is sufficient to handle the convergence of the solutions of the SPDE's in Chapter 5. We will then give a slightly deeper treatment under stronger hypotheses on the K_n.

<u>DEFINITION</u>. Let $\overline{\mathcal{P}}_S(M)$ be the closure of $\mathcal{P}_S(M)$ in the norm

$$|||\, f \,||| = \sup_{\substack{\omega \in \underline{\underline{D}} \\ 0 \leq t \leq 1}} \|f(t,\omega)\|_\infty .$$

<u>Note</u>: We will suppress variables from our notation when possible. Thus, we write $\|f(t,\omega)\|$ in place of $\|f(\cdot,t,\omega)\|$.

If $M^n \Rightarrow M^0$, we are interested not only in the convergence of $f\cdot M^n$ to $f\cdot M^0$, but also in the joint convergence of $(M^n, f\cdot M^n)$ to $(M^0, f\cdot M^0)$. We may also want to know about the simultaneous convergence of a number of martingale measures and integrals. This means we will want to state a convergence theorem for martingale measures with values in R^m, and integrands which are r×m matrices. Weak convergence in this context is convergence in the Skorokhod space $\underline{\underline{D}}^m = \underline{\underline{D}}\{[0,1], S'(R^{md})\}$.

If $M = (M^1,\dots,M^n)$ and if $f = (f_{ij})$ is matrix valued, we will say $f \in \overline{\mathcal{P}}_S(M)$ (resp. $f \in \underline{\underline{P}}(M)$) if for each i, j, $f_{ij} \in \overline{\mathcal{P}}_S(M^j)$, (resp. $f_{ij} \in \underline{\underline{P}}(M^j)$).

<u>PROPOSITION 7.6.</u> Let $M^0, M^1, M^2,\dots$ be a sequence of R^m-valued worthy martingale measures and suppose that each coordinate satisfies (7.5). Let $f^n(x,t,\omega) = (f^n_{ij}(x,t,\omega))$ be r×m matrices such that $f^n \in \underline{\underline{P}}(M^n)$, n = 1,2,..., and $f^0 \in \overline{\mathcal{P}}_S(M^0)$. Suppose that $M^n \Rightarrow M^0$ (on $\underline{\underline{D}}^m$) and that $\| |f^n_{ij} - f^0_{ij}\| | \to 0$. Then $f^n\cdot M^n \Rightarrow f^0\cdot M^0$ on $\underline{\underline{D}}^m$.

<u>REMARK.</u> We can choose f^n of the form $\binom{I}{F^n}$ where I is the m×m identity and F^n is diagonal with $F^n_{jj} = f^n_j$. Then $f\cdot M = (M^{n1},\dots,M^{nm}, f^n_1\cdot M^{n1},\dots,f^n_m\cdot M^{nm})$ and the proposition gives the joint convergence of the M^{nj} and the $f^n_j\cdot M^{nj}$.

<u>PROOF.</u> Each coordinate of $(f^n\cdot M^n)$ is tight by Corollary 7.4, so $(f^n\cdot M^n)$ is tight. To show it converges, we need only show convergence of the distributions of $(f^n\cdot M^n_{t_1},\dots,f^n\cdot M^n_{t_k})$, where $t_1 \le t_2 \le \dots \le t_k$ are continuity points of M^0. Note that this vector can be realized by taking another, larger matrix f and looking only at t = 1. For instance, if $t_1 < t_2 \le 1$, $(f\cdot M^n_{t_1}, f\cdot M^n_{t_2}) = \tilde f\cdot M^n_1$, where $\tilde f = (\tilde f_1, \tilde f_2)$ and

$$\tilde f_i(x,t) = \begin{cases} f(x,t), & t \le t_i \\ 0 & t > t_i, \end{cases} \quad i = 1,2.$$

Since $f \in \overline{\mathcal{P}}_S(M^0)$ and the t_i are continuity points of M^0, $\tilde f_i \in \overline{\mathcal{P}}_S(M^0)$.

By Mitoma's theorem, then, it is enough to check the convergence of $f^n\cdot M_t(\phi)$ for a single test function ϕ and for t = 1. By Theorem 6.1, it is enough to show that

$$E\{h(f\cdot M^n_1(\phi))\} \to E\{h(f\cdot M^0_1(\phi))\}$$

for any uniformly continuous bounded h on R^r.

Choose $f_{ij}^{\varepsilon} \in \mathcal{P}_S(M^0)$ such that $\| |f_{ij}^0 - f_{ij}^{\varepsilon}| \| < \varepsilon 2^{-n}$ for each i, j and let $f^{\varepsilon} = (f_{ij}^{\varepsilon})$. Then

$$\begin{aligned}
&|E\{h(f^n \cdot M_1^n(\phi))\} - E\{h(f^0 \cdot M_1^0(\phi))\}| \\
&\leq E\{|h(f^n \cdot M_1^n(\phi)) - h(f^0 \cdot M_1^n(\phi))|\} \\
&+ E\{|h(f^0 \cdot M_1^n(\phi)) - h(f^{\varepsilon} \cdot M_1^n(\phi))|\} \\
&+ |E\{h(f^{\varepsilon} \cdot M_1^n(\phi))\} - E\{f^{\varepsilon} \cdot M_1^0(\phi)\}| \\
&+ E\{|h(f^{\varepsilon} \cdot M_1^0(\phi)) - h(f^0 \cdot M_1^0(\phi))|\} \\
&= E_1 + E_2 + E_3 + E_4.
\end{aligned}$$

Since h is uniformly continuous, given $\rho > 0$ there exists $\delta > 0$ such that $|h(y) - h(x)| < \rho$ if $|y - x| < \delta$. Then

$$\begin{aligned}
E_1 &\leq \rho + 2\|h\|_{\infty} P\{|f^n \cdot M^n(\phi) - f^0 \cdot M^n(\phi)| > \delta\} \\
&\leq \rho + 2\|h\|_{\infty} \delta^{-2} E\{|(f^n - f^0) \cdot M^n(\phi)|^2\} \\
&\leq \rho + 2\|h\|_{\infty} \delta^{-2} \|\phi\|_h^2 \sum_{i,j} \| |f_{ij}^n - f_{ij}^0| \|^2 E\{\gamma^n(1)\}
\end{aligned}$$

by Lemma 7.1. Now it follows from (7.5) that $E\{\gamma_n(1)\}$ is bounded by, say, C, so that $\limsup_{n \to \infty} E_n \leq \rho$. Moreover, the same type of calculation, with $f^n - f^0$ replaced by $f^{\varepsilon} - f^0$, gives

$$E_2 \leq \rho + 2mr \; \| |h| \| \; \delta^{-2} \; \|\phi\|_h^2 \, C \, \varepsilon,$$

and E_4 satisfies the same inequality since the calculation is valid for $n = 0$. Finally $E_3 \to 0$ by Proposition 7.5. Since ρ and ε are arbitrary, we conclude that $E_1 + \ldots + E_4 \to 0$. Q.E.D.

Let us now consider the SPDE (5.4). Its solution is

$$(7.7) \qquad \begin{aligned}
V_t(\phi) &= \int_{D\times[0,t]} T^* G_{t-s}(\phi, y) \, M(dy\, ds) \\
&= \int_{D\times[0,t]} T^* \left(\int_s^t G_{u-s}(L\phi, y) du + \phi(y) \right) M(dy\, ds) \\
&= M_t(T^*\phi) + \int_0^t \left[\int_0^u \int_D T^* G_{u-s}(L\phi, y) \, M(dy\, ds) \right] du,
\end{aligned}$$

where we have used (5.7), the fundamental equation of the Green's function, and then changed the order of integration.

<u>LEMMA 7.7</u>. $E\{\sup_{t\leq 1} V_t(\phi)^2\} \leq (8 \|T^*\phi\|_h^2 + 2 \sup_{t\leq 1} \|T^* G_t(L\phi)\|_h^2) E\{k(1)\}$.

PROOF. $\sup_t V_t(\phi)^2 \le 2 \sup_t M_t^2(T^*\phi) + 2 \sup_t \left[\int_0^t \left(\int_0^u \int_D T^*G_{u-s}(L\phi)M(dy\ ds)\right)^2 du\right]$

by Schwartz. By Doob's inequality

$$E\{\sup_t V_t(\phi)^2\} \le 8\ E\{M_1^2(T^*\phi)\} + 2 \int_0^1 E\left\{\left[\int_0^u \int_D T^*G_{u-s}(L\phi)M(dy\ ds)\right]^2\right\} du,$$

and the conclusion follows by Lemma 7.1.

Q.E.D.

Let $M^0, M^1, M^2, \ldots$ be a sequence of worthy martingale measures and define V^n by (7.7).

PROPOSITION 7.8. If (M^n) satisfies (7.5), then (V^n) is tight. If, in addition, $M^n \Rightarrow M^0$, then $(M^n, V^n) \Rightarrow (M^0, V^0)$.

PROOF. In order to prove (V^n) is tight, it is enough by Mitoma's theorem to show that $(V_t^n(\phi))$ is tight for any $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$. The (M^n) are tight by Proposition 7.3, so we must show that the sequence (U^n) defined by

$$U_t^n = \int_0^t \left[\int_0^u \int_D T^*\ G_{u-s}(L\phi,y)M(dy\ ds)du\right.$$

is tight. Let

$$S_n = \sup_{u\le 1} \left| \int_0^u \int_D T^*\ G_{u-s}(L\phi,y)M(dy\ ds)\right|.$$

If τ_n is a stopping time for M^n, $\delta_n > 0$, and if $\tau_n + \delta_n \le 1$, then

$$|U^n_{\tau_n+\delta_n} - U^n_{\tau_n}| \le \delta_n S_n.$$

By Lemma 7.7, $E\{S_n^2\} \le C\ E\{k_n(1)\}$. By (7.5), this expectation is bounded in n, so $U^n_{\tau_n+\delta_n} - U^n_{\tau_n} \to 0$ in L^1, hence in probability. Moreover, $E\{(U_t^n)^2\} \le E\{S_n^2\}$, so that (U_t^n) is tight for fixed t. By Aldous' Theorem, (U^n) is tight.

We need only check the convergence of the finite-dimensional distributions in order to see that $(M^n, V^n) \Rightarrow (M^0, V^0)$. But for each t, $T^*G_{t-s}(\phi,y)$ is deterministic, is in $\underline{\underline{S}}(\mathbf{R}^d)$ as a function of y, and is a continuous function of s for $s \ne t$. Thus if t is not a discontinuity point of M^0, $(s,y) \to T^*G_{t-s}(\phi,y)$ is in $\overline{\mathcal{P}}(M^0)$. The finite-dimensional convergence now follows from Proposition 7.6.

Q.E.D.

If ϕ is in the Sobolev space $H_p(D)$ of Example 1, Chapter 4, and if $p > d/2$, then there is a constant C such that $\|\phi\|_\infty \leq C\|\phi\|_{H_p}$. Since h_0 is bounded away from zero on the bounded domain D, $\|\phi\|_h \leq C\|\phi\|_{H_p}$ for another constant C. If α is the order of T (and hence of T*) and if $\phi \in H_{p+\alpha+2}$ for all $0 \leq t \leq 1$, then $\|T^*G_t(L\phi)\|_h \leq C\|\phi\|_{H_{p+\alpha+2}}$ for all $0 \leq t \leq 1$ (and in particular for $t = 0$) so that Lemma 7.7 implies that

$$E\{\sup_{t\leq 1} V_t^n(\phi)^2\} \leq C(\|\phi\|_{H_{p+\alpha+2}})E\{k_n(1)\}.$$

Suppose that $q > p + 2 + d/2$. Now $E\{k_n(1)\}$ is bounded if (7.5) holds, and $H_{p+\alpha+2} \underset{HS}{\lessdot} H_q$. By Corollary 6.16, $V^n \Rightarrow V$, as processes with values in H_{-q}. Thus

COROLLARY 7.9. If (M^n) is a sequence of worthy martingale measures which satisfies (7.5) and if $M^n \Rightarrow M$, then $V^n \Rightarrow V^0$ in $\underline{\underline{D}}\{[0,1], H_{-(d+\alpha+3)}\}$.

AN APPLICATION

In many applications - the neurophysiological example of Chapter 3, for instance - the driving noise is basically impulsive, of a Poisson type, but the impulses are so small and so closely spaced that, after centering, they look very much like a white noise. The following results show that for some purposes at least, one can approximate the impulsive model by a continuous model driven by a white noise. One might think of this as a diffusion approximation.

Let us return to the setting of Chapter 5. Let D be a bounded domain in $\mathbf{R}^d$ with a smooth boundary, and consider the initial-boundary value problem (5.3) with two changes: we will allow an initial value given by a measure on $\mathbf{R}^d$, and we will replace the martingale measure M by a Poisson point process Π.

Let Π^n be a sequence of time-homogeneous Poisson point processes on D with characteristic measures μ_n. (Recall that this means that Π^n is a random σ-finite signed measure on $D\times[0,\infty)$ which is a sum of point masses. If $A \subset D$ is Borel and $K \subset \mathbf{R}$ is compact with $0 \notin K$, let $N_t^n(A\times K)$ be the counting process: $N_t^n(A\times K)$ is the number of points in $A\times[0,t]$ whose masses are in K. Then $\{N_t^n(A\times K), t \geq 0\}$ is a

Poisson process with parameter $\mu_n(A\times K)$, and if $A_1\times K_1 \cap A_2\times K_2 = \phi$, then $N^n(A_1\times K_1)$ and $N^n(A_2\times K_2)$ are independent.) Let

$$m_n(A) = \int_{D\times\mathbf{R}} r\,\mu_n(dx\,dr)$$

and

$$\sigma_n(A) = \int_{D\times\mathbf{R}} r^2\mu_n(dx\,dr)$$

be the mean and intensity measures, respectively, of Π^n. For $\delta > 0$, let

$$Q_n(\delta) = \int_D r^{2+\delta}\mu_n(dx\,dy).$$

Let L, T and B be as in (5.3), let ν_n be a finite measure on D and consider the initial-boundary value problem

$$\text{(7.8)}\qquad \begin{cases} \dfrac{\partial V}{\partial t} = LV + T\dot{\Pi}^n \\ BV = 0 \quad \text{on } \partial D \\ V_0 = \nu_n \end{cases}$$

Note that $M_t^n(A) \overset{\text{def}}{=} \Pi^n(A \times [0,t]) - t\,m_n(A)$ is an orthogonal martingale measure. The solution to (7.8) is, by Theorem 5.2,

$$\text{(7.9)}\qquad V_t^n(\phi) = \int_D G_t(\phi,y)\nu_n(dy) + \int_{D\times[0,t]} G_{t-s}(\phi,y)m_n(dy)ds + \int_{D\times[0,t]} T^*G_{t-s}(\phi,y)M^n(dy\,ds).$$

<u>THEOREM 7.10</u>. Suppose that there exist finite (signed) measures ν, m, and σ^2 on D such that $\nu_n \Rightarrow \nu$, $m_n \Rightarrow m$, and $\sigma_n \Rightarrow \sigma$, in the sense of weak convergence of measures on D. Suppose further that for some $\delta > 0$, $Q_n(\delta) \to 0$. Then there exists a white noise W on $D\times[0,t)$, based on $d\sigma\,dt$, such that $(M^n, V^n) \Rightarrow (W, V)$, where V is defined by

$$\text{(7.10)}\qquad V_t(\phi) = \int G_t(\phi,y)\nu(dy) + \int_{D\times[0,t]} G_{t-s}(\phi,y)m(dy)ds + \int_{D\times[0,t]} T^*G_{t-s}(\phi,y)W(dy\,ds)$$

<u>PROOF</u>. The first two terms of (7.9) are deterministic and an elementary analysis shows that they converge uniformly in t, $0 \le t \le 1$, to the corresponding terms of (7.10). (Indeed, $y \to G_t(\phi,y)$ is continuous so the integrals with respect to ν_n and m_n converge; one gets the requisite uniform convergence by noticing that the same holds for the integrals of $\frac{\partial G_t}{\partial t}(\phi,y)$.)

It remains to show that the third term converges weakly. Note that the sequences $(M_n(D))$ and $(\sigma_n(D))$, being convergent, are bounded by, say, K. Thus if $0 < s < t \leq 1$:

$$\langle M^n(D)\rangle_t - \langle M^n(D)\rangle_s = (t-s)\sigma_n(D) \leq K(t-s).$$

It follows that (M^n) is of class (ρ, K^ρ) for all $\rho > 0$ and that (7.5) holds, for $\gamma(\delta) = K\delta$. Thus (M^n) is tight by Proposition 7.3. We claim that $M^n \Rightarrow W$. Note that once we establish this claim, the theorem follows by Proposition 7.8.

It is enough to show that the finite-dimensional distributions of $M^n(\phi)$ tend to those of $W(\phi)$ for one ϕ, and, since $M_t^n(\phi)$ and $W_t(\phi)$ are processes of stationary independent increments, we need only check the convergence for one value of t, say t = 1. Thus we have reduced the theorem to a special case of the classical central limit theorem.

Since we know the characteristic functions for the M^n explicitly, we can do this by a direct calculation rather than applying, say, Liapounov's central limit theorem.

Write $e^{ix}-1 = ix - \frac{1}{2}x^2 + f(x)$. Then certainly $|f(x)| \leq 2x^2$ and $\lim_{x\to 0} f(x)/x^2 = 0$. Let

$$\phi_n(\lambda) = \log E\{e^{i\lambda M_1^n(\phi)}\}$$

$$= \log E\{e^{i\lambda \Pi_1^n(\phi)}\} - i\lambda\, m_n(\phi),$$

where $\Pi_1^n(\phi) = \int_{D\times[0,1]} \phi(x)\Pi^n(dx\, ds)$. Let us also write $m_n(\phi) = \int \phi(x)m_n(dx)$, $\sigma_n(\phi^2) = \int \phi^2(x)\sigma_n(dx)$. From the properties of Poisson processes the above is

$$= \lambda \int_{D\times[0,1]} (e^{i\,\phi(x)r}-1)\,\mu_n(dx\, dr) - i\lambda\, m_n(\phi)$$

$$= i\lambda \int_{D\times[0,1]} \phi(x)\, r\,\mu_n(dx\, dr) - i\,\lambda\, m_n(\phi)$$

$$- \frac{1}{2}\lambda^2 \int_{D\times[0,1]} \lambda^2(x)r^2\mu_n(dx\, dr) + \int_{D\times[0,1]} f(\lambda\phi(x)r)\mu_n(dx\, dr)$$

$$= -\frac{1}{2}\lambda^2\sigma_n(\phi^2) + \int_{D\times[0,1]} f(\lambda\,\phi(x)r)\mu_n(dx\, dr)\ .$$

The first term converges to $-\frac{1}{2}\lambda^2\sigma(\phi^2)$ as $n\to\infty$ since $\sigma_n \Rightarrow \sigma$. We claim the second term tends to zero.

Choose $\varepsilon > 0$ and let $\eta > 0$ be such that if $|x| < \eta$, then $|f(x)| \leq \varepsilon x^2$. The second integral is bounded by

$$\varepsilon \int_{\{\lambda r\|\phi\|_\infty < \eta\}} \lambda^2\phi^2(x)r^2\mu_n(dx\ dr) + \int_{\{\lambda r\|\phi\|_\infty \geq \eta\}} \eta^{-\delta}(\lambda\ \phi^2(x)\ |r|)^{2+\delta}\mu_n(dx\ dr)$$

$$\leq \varepsilon\ \lambda^2\sigma_n(\phi^2) + \eta^{-\delta}\lambda^{2+\delta}\ \|\phi\|_\infty^{2+\delta}Q_n(\delta).$$

Since $(\sigma_n(\phi^2))$ is bounded and $Q_n(\delta) \to 0$, we conclude that

$$\phi_n(\lambda) \to -\frac{\lambda^2}{2}\ \sigma_n(\phi^2)$$

which is the log characteristic function of $W_1(\phi)$.

Q.E.D.

REMARKS. Let α be the degree of T. By Corollary 7.9, we see that $V^n \Rightarrow V$ in $\underline{\underline{D}}\{[0,1],\ H_{-(d+\alpha+3)}\}$. One can doubtless improve the exponent of the Sobolev space by a more careful analysis.

AN EXTENSION

Propositions 7.6 and 7.8 are sufficient for many of our needs, but they are basically designed to handle deterministic integrands, and we need to extend them if we are to handle any reasonably large class of random integrands.

We will look at the case in which the functions $k_n(t)$ of (7.1) are all absolutely continuous. The treatment unfortunately becomes more complicated. Any reader without a morbid interest in Hölder's inequalities should skip this section until he needs it.

Let M be a worthy martingale measure on R^d. Set

$$\Delta M^*(\phi) = \sup_{t\leq 1}|M_t(\phi) - M_{t-}(\phi)| \text{ and } \Delta M^* = \sup_{A\subset R^d}\Delta M_t^*(I_A h_0).$$

Note that for any $f \in \underline{\underline{P}}_M$ such that $\|f(t)\|_h \leq 1$ for all t, we have

$$\sup_{t,A}|f\bullet M_t(A) - f\bullet M_{t-}(A)| \leq 2\ \Delta M^*.$$

We say M has $\underline{L^p\text{-dominated}}$ jumps if $\Delta M^* \in L^p$.

Let us recall Burkholder's inequalities for the predictable square function $\langle M(\phi)\rangle_t$.

THEOREM 7.11. (Burkholder-Davis-Gundy) Suppose Φ is a continuous increasing function on $[0,\infty)$ with $\Phi(0) = 0$, such that for some constant c, $\Phi(2x) \leq c\ \Phi(x)$, for all $x \geq 0$. Then

(i) there exists a constant C such that

$$E\{\Phi(\sup_{s\leq t} |M_s(\phi)|)\} \leq CE\{\Phi(\langle M(\phi)\rangle_t^{1/2})\} + CE\{\Phi(\Delta M^+(\phi))\};$$

(ii) if Φ is concave

$$E\{\Phi(\sup_{s\leq t} |M_s(\phi)|^2\} \leq 5\ E\{\Phi(\langle M(\phi)\rangle_t^{1/2})\}.$$

We will usually apply this to $\Phi(x) = |x|^p$ for $p > 0$. If $0 < p < 1$, we are in case (ii).

In what follows, $M^0, M^1, M^2, \ldots$ is a sequence of worthy martingale measures and $k_0, k_1, k_2, \ldots$ are the increasing processes of (7.1). We will assume, as we may, that all the M^n are defined canonically on $\underline{\underline{D}}$ and we will use P^n and E^n for the distribution and expectation relative to M^n. When there is no danger of confusion we will simply write P and E respectively.

<u>DEFINITION</u>. Let $p > 0$, $K > 0$. The sequence (M^n) is of <u>class (p,K)</u> if for all n there exists a random variable X_n on $\underline{\underline{D}}$ such that

(7.11a) $\quad k_n(t) - k_n(s) \leq X_n |t-s| \quad$ if $0 \leq s \leq t \leq 1$, $n = 0,1,2,\ldots$;

(7.11b) $\quad E^n\{X_n^p\} \leq K$, $n = 0,1,2,\ldots$.

If the M^n are m-dimensional, we say that (M^n) is of class (p,K) if for each n there exists an X_n satisfying (ii) such that (i) holds for each coordinate. (In particular, each coordinate is itself of class (p,K).)

<u>REMARK</u>. If (M^n) is of class (p,K), than each M^n is quasi-left continuous, i.e. M^n has no predictable jumps. This is immediate from Lemma 7.2.

Just as in Proposition 7.6, we want to close $\mathcal{P}_S(M^0)$ in some suitable norm. In this case, the norm depends on the sequence $M = (M^n)$.

<u>DEFINITION</u>. $\bar{\bar{\mathcal{P}}}_S(M)$ is the class of $f \in \bigcap_{n=0}^{\infty} \underline{\underline{P}}_{M^n}$ such that for each $\varepsilon > 0$ there exists $f^\varepsilon \in \mathcal{P}_S(M^0)$ such that

(7.12) $$\sup_n E^n\{1 \wedge \int_0^1 \|f(s) - f^\varepsilon(s)\|_\infty^2 ds\} \leq \varepsilon.$$

Note that $Y \to E\{1 \wedge |Y|\}$ gives a distance compatible with convergence in probability, so that the above says that f can be approximated in P^n-probability for each n, and that the approximation is uniform in n. It is this uniformity which is important. Without it, the condition is trivial, as the following exercise shows.

Exercise 7.1. Let $f \in \bigcap_n \underline{\underline{P}}_{M^n}$. Then for any N and $\varepsilon > 0$ there exists $f^{\varepsilon,N} \in \bigcap_n \mathcal{P}_S(M^n)$ such that

$$E^N\{\int_0^1 \|f(s) - f^{\varepsilon,N}(s)\|_\infty^2 ds\} \leq \varepsilon.$$

PROPOSITION 7.12. Let M^n, $n = 0,1,2,\ldots$ be a sequence of m-dimensional worthy martingale measures which is of class (p,k) for some $p > 0$, $K > 0$. For each n, let $f^n(x,t,\omega) = (f^n_{ij}(x,t,\omega))$ be an $r \times m$ matrix such that $f^n \in \underline{\underline{P}}_{M^n}$. We suppose that there exists an $\varepsilon > 0$ such that $E^n\{|\int_0^1 \|f^n_{ij}(s)\|_\infty^{2+\varepsilon} ds|^\varepsilon\}$, $n = 0,1,2,\ldots$ is bounded.

(i) Then $(f^n \cdot M^n)$ is tight on $\underline{\underline{D}}^r$;

(ii) if, further, $M^n \Rightarrow M^0$ in $\underline{\underline{D}}^m$, and if

a) $f^0_{ij} \in \bar{\bar{\mathcal{P}}}_S(M)$, all i,j;

b) $\lim_{n\to\infty} E^n\{|\int_0^1 \|f^n_{ij}(s) - f^0_{ij}(s)\|_\infty^2 ds|^\varepsilon\} = 0$,

then $f^n \cdot M^n \Rightarrow f^0 \cdot M^0$ on $\underline{\underline{D}}^r$.

PROOF. If the hypotheses hold for a given ε, they also hold for any $\varepsilon' < \varepsilon$, so we may assume that $\varepsilon \leq p \wedge 1$. Thus if X_n, $n = 0,1,2,\ldots$ are the random variables of (7.11), then $E^n\{X^\varepsilon_n\} \leq K^{\varepsilon/p}$ by Jensen's inequality. Notice also that for $f \in \underline{\underline{P}}_{M^n}$,

$$(7.13)\qquad |f^n \cdot M^n(\phi)|^\varepsilon = |\sum_i (\sum_j f^n_{ij} \cdot M^{nj}(\phi))^2|^{\varepsilon/2} \leq \sum_{i,j} |f^n_{ij} \cdot M^{nj}(\phi)|^\varepsilon$$

since $\varepsilon \leq 1$.

Let T_n be a stopping time for M^n and let $\delta_n > 0$ be such that $T_n + \delta_n \leq 1$. By Burkholder's inequality, (7.13), and (7.4), for any test function ϕ

$$E^n\{|f^n \cdot M^n_{T_n+\delta_n}(\phi) - f^n \cdot M^n_{T_n}(\phi)|^\varepsilon\} \leq 5\|\phi\|_h^\varepsilon \sum_{i,j} E^n\{|\int_{T_n}^{T_n+\delta_n} \|f^n_{ij}(s)\|_\infty^2 dh_n(s)|^{\frac{\varepsilon}{2}}\}$$

$$\leq 5 \|\phi\|_h^\varepsilon \sum_{i,j} E^n\{ \int_{T_n}^{T_n+\delta_n} \|f_{ij}^n(s)\|^2 ds)^{\frac{\varepsilon}{2}}\}.$$

Apply Hölder's inequality to the integral with $p = 1 + \varepsilon/2$, $q = 1 + 2/\varepsilon$, then apply Schwartz' inequality:

$$\leq 5 \|\phi\|_h^\varepsilon \quad E^n\{X_n^\varepsilon\}^{1/2} \sum_{i,j} E^n\{| \int_0^1 \|f_{ij}^n(s)\|^{2+\varepsilon} ds \quad |^{2\varepsilon/(2+\varepsilon)}\}^{1/2} \delta^{\varepsilon^2/(4+\varepsilon)}.$$

Now $\frac{2\varepsilon}{2+\varepsilon} < \varepsilon$ so the expectations are bounded, and the above is then

$$\leq C\delta_n^{\varepsilon^2/(4+\varepsilon)}$$

for some constant C. Thus, if $\delta_n \to 0$, $f^n \cdot M_{T_n+\delta_n}^n(\phi) - f^n \cdot M_{T_n}^n(\phi) \to 0$ in probability. Take $T_n = 0$ and $\delta_n = t$ to see that for each t the sequence $(f^n \cdot M_t^n(\phi))$ of random variables is bounded in L^ε, hence tight. Aldous' theorem then implies that the sequence $(f^n \cdot M^n(\phi))$ of real-valued processes is tight, and Mitoma's theorem implies that $(f^n \cdot M^n)$ is tight on $\underline{\underline{D}}^r$.

As in Proposition 7.5 we need only show that $(f^n \cdot M_1^n(\phi))$ converges weakly to prove (ii). Let h be bounded and uniformly continuous on $\mathbf{R}^r$. Then if $\gamma > 0$ there exist $f_{ij}^\gamma \in \mathcal{P}_S(M^0)$ such that

(7.14) $$E^n\{|1 \wedge \int_0^1 \|f_{ij}^\gamma(s) - f_{ij}^0(s)\|^2 ds\} < \gamma, \; n = 0,1,2,\ldots .$$

Now

$$\begin{aligned} &|E^n\{h(f^n \cdot M_1^n(\phi))\} - E^0\{h(f^0 \cdot M_1^0(\phi))\}| \\ &\leq E^n\{|h(f^n \cdot M_1^n(\phi)) - h(f^0 \cdot M_1^n(\phi))|\} + E^n\{|h(f^0 \cdot M_1^n(\phi)) - h(f^\gamma \cdot M_1^n(\phi))|\} \\ &\quad + |E^n\{h(f^\gamma \cdot M_1^n(\phi))\} - E^0\{h(f^\gamma \cdot M_1^0(\phi))\}| \\ &\quad + E^0\{|h(f^\gamma \cdot M_1^0(\phi)) - h(f^0 \cdot M_1^0(\phi))|\} \\ &\stackrel{\text{def}}{=} E_1 + E_2 + E_3 + E_4. \end{aligned}$$

Fix $\rho > 0$ and let $\eta > 0$ be such that $|h(y) - h(x)| < \rho$ if $|y-x| < \eta$. Then

$$\begin{aligned} E_1 &\leq \rho + 2 \|h\|_\infty P\{|(f^n - f^0) \cdot M_1^n(\phi)| \geq \eta\} \\ &\leq \rho + 2\|h\|_\infty \eta^{-\varepsilon} E^n\{|(f^n - f^0) \cdot M_1^n(\phi)|^\varepsilon\}. \end{aligned}$$

By (7.13) and Burkholder's inequality (Theorem 7.11 (ii)) this is

$$\leq \rho + 10 \|h\|_\infty \eta^{-\varepsilon} \sum_{ij} E^n \{| \int_0^1 \|f_{ij}^n(s) - f_{ij}^0(s)\|_\infty^2 dk_n(s)|^{\varepsilon/2}\}$$

$$\leq \rho + 10 \|h\|_\infty \eta^{-\varepsilon} \sum_{ij} E^n\{X_n^{\varepsilon/2} | \int_0^1 \|f^n(s) - f^0(s)\|_\infty^2 ds|^{\varepsilon/2}\}$$

$$\leq \rho + 10\,\|h\|_\infty \eta^{-\varepsilon} E^n\{X_n^\varepsilon\}^{1/2} \sum_{ij} E^n\{|\int_0^1 \|f^n(s) - f^0(s)\|_\infty^2 ds|^\varepsilon\}^{1/2}$$

where we have used the fact that (M^n) is of class (p,K) and Schwartz' equality. But $E^n\{X_n^\varepsilon\}$ is bounded, and the last expectation tends to zero, so $\limsup_{n\to\infty} E_1 \leq \rho$.

Similarly,

$$E_2 \leq \rho + 2\,\|h\|_\infty \eta^{-\varepsilon}\, E\{1 \wedge |(f^\gamma - f^0) \cdot M_1^n(\phi)|^\varepsilon\}.$$

Now $\Phi(x) = 1 \wedge |x|^\varepsilon$ is concave so by Burkholder's inequality and (7.4)

$$\leq \rho + 10\,\|h\|_\infty \eta^{-\varepsilon} \sum_{ij} E^n\{1 \wedge (\int_0^1 \|f^\gamma(s) - f^0(s)\|_\infty^2\, dk_n(s)^{\varepsilon/2}\}$$

$$\leq \rho + 10\,\|h\|_\infty \eta^{-\varepsilon}\, E^n\{X_n^\varepsilon\}^{1/2} \sum_{ij} E^n\{(1 \wedge \int_0^1 \|f^\gamma(s) - f^0(s)\|ds)^\varepsilon\}^{1/2}.$$

Apply Jensen's inequality to this last expectation, and use (7.14):

$$\leq \rho + 10mr\ \|h\|_\infty K^{\varepsilon/2} \eta^{-\varepsilon} \gamma^{\varepsilon/2}$$

This is valid for $n = 0$ too, so E_4 has the same bound. Finally, $E_3 \to 0$ as $n \to \infty$ by Proposition 7.5. Since ρ and γ can be made as small as we wish, it follows that $E_1 + \ldots + E_4 \to 0$. Thus $f^n \cdot M_1^n(\phi) \Rightarrow f^0 \cdot M_1^0(\phi)$ and, by Mitoma's theorem, we are done.

Q.E.D.

We now look at the convolution integrals. In order to prove convergence of distribution-valued processes such as V^n of (7.9), we usually prove that the real-valued processes $(V_n(\phi))$ converge weakly and appeal to Mitoma's theorem. This means that we only need to deal with the convergence of real-valued processes.

In the interest of simplicity - relative simplicity, that is - we will limit ourselves to the case where the integrand does not depend on n. The extension to the case where it does depend on n requires an additional condition on the order of (ii) of Proposition 7.12, but is relatively straightforward. We will be dealing with processes of the form

$$U_t^n = \int_{R^d \times [0,t]} g(x,s,t) M^n(dx\, ds).$$

Let $M^0, M^1, \ldots$ be a family of worthy martingale measures of class (p,K) for some $p > 0$, $K > 0$. Let $g(x,s,t,\omega)$ be a function on $R^d \times \{(s,t): 0 \leq s \leq t\} \times \underline{\underline{D}}$ such that

(i) for each t_0, $g(\cdot,\cdot,t_0,\cdot)I_{\{s\le t_0\}} \in \bigcap_{n=0}^{\infty} \underline{\underline{P}}_{M^n}$;

(ii) there exist functions Y_n on $\underline{\underline{D}}$ and positive constants β, q, and C such that if $0 \le s \le t \le t'$,

a) $\|g(s,t') - g(s,t)\|_h \le Y_n(\omega)\ |t'-t|^{\beta}$ P^n-a.s., all n;

b) $\|g(s,t')\|_h \le Y_n(\omega)$ P^n-a.s., all n;

c) $E^n\{Y_n^q\} \le C$, all n.

<u>DEFINITION</u>. If g satisfies (i) and (ii) above we say g is of <u>Hölder class</u> (β,q,C) relative to (M^n).

<u>THEOREM 7.13</u>. Let (M^n) be of class (p,K), where $p > 2$ and $K > 0$. Let g be of Hölder class (β,q,C), where $\frac{1}{p} \le \beta \le 1$ and $q > \frac{2p}{\beta p-1}$. Suppose further that the jumps of M^n are L^{2p}-dominated, uniformly in n. Then

(i) $\{U_t^n,\ 0 \le t \le 1\}$ has a version which is right continuous and has left limits;

(ii) there exists a constant A, depending only on p, q, C and K such that

$$E\{\sup_{t\le 1} |U_t^n|^r\} \le A \quad \text{if } 1 \le r \le \frac{2pq}{2p+q}\ ;$$

(iii) if $t \to M_t^n$ is continuous, U^n is Hölder continuous. Moreover there exists a random variable Z_n with $E\{Z_n^r\} < \infty$ such that if $r = \frac{2pq}{2p+q}$,

$$|U_{t+s}^n - U_t^n| \le Z_n\ s^{\frac{1}{2}\wedge\beta - \frac{1}{r}}\left(\log \frac{2}{s}\right)^{2/r},\quad 0 \le s,\ t \le 1;$$

(iv) the family (U^n) is tight on $\underline{\underline{D}}\{[0,t],\ \mathbb{R}\}$.

(v) Suppose further that $M^n \Rightarrow M^0$ on $\underline{\underline{D}}$ and that for a dense set of $t \in [0,t]$, $g(x,s,t)I_{\{s\le t\}} \in \bar{\bar{\mathcal{P}}}_S(M)$. Then $(U^n, M^n) \Rightarrow (U^0, M^0)$.

<u>PROOF</u>. By replacing C and K by max(C,K) if necessary, we may assume C = K. By enlarging K further, we may assume that $E\{|\Delta M^{n*}|^{2p}\} \le K$ for all n.

Set $r = 2pq(2p+q)^{-1}$; since $q > 2p(\beta q-1)^{-1}$, $\beta^{-1} \le p < r < 2p$. If X_n and Y_n are the random variables from (7.11) and from the definition of Hölder class (β,q,C) respectively, then

$$E\{X_n^{r/2} Y_n^r\} \le K. \tag{7.15}$$

Indeed the left-hand side is bounded by

$$E\{X_n^p\}^{r/2p} E\{Y_n^q\}^{\frac{2p-r}{2q}} \le K^{r/2p} K^{\frac{2p-r}{2p}}.$$

Let us extend the domain of definition of g to include values of $s \geq t$ by setting

$$g(x,s,t) = g(x,s,t \vee s),\ 0 \le s \le 1,\ 0 \le t \le 1.$$

Then $t \to g(x,s,t)$ is constant on $[0,s]$, so g remains predictable in s and still satisfies (i) and (ii). Define

$$V_t^n = \int_{E\times[0,1]} g(x,s,t)M^n(dx\ ds).$$

Let us first show that V^n has a continuous version for which $\sup_t V_t^n \in L^r$.

Note that $\Delta((g(t') - g(t))\cdot M^n)$ is bounded by

$2 \sup_s \|g(s,t') - g(s,t)\|_h \Delta M^{n*} \le 2Y \Delta M^{n*}(t'-t)^\beta$. By (7.4) and Burkholder's inequality (Thm 7.11 (ii))

$$E\{|V_{t'}^n - V_t^n|^r\} = E\{|\int_{E\times[0,1]} (g(x,s,t') - g(x,s,t))M^n(dx\ ds)|^r\}$$

$$\le C_r E^n\{|\int_0^1 \|g(s,t') - g(s,t)\|_h^2 dk_n(s)|^{r/2} + |2Y\Delta M^{n*}(t'-t)^\beta|^r\}$$

$$\le C_r E^n\{Y^r(X_n^{r/2} + (2\Delta M^{n*})^r\}(t'-t)^{\beta r}.$$

Then (7.15) and a similar calculation with $X_n^{1/2}$ replaced by ΔM^{n*} shows this is

$$\le Cr\ K\ (1 + 2^r)(t'-t)^{\beta r}.$$

By the same argument

$$E^n\{|V_t^n|^r\} \le C_r K(1 + 2^r).$$

Since $\beta r > 1$, Corollary 1.2 implies that V^n has a continuous version. More exactly, there exists a random variable Z_n and a constant A', which does not depend on n, such that for $0 < \gamma < \beta - 1/r$

$$\sup_{0\le t\le t'\le 1} |V_{t'}^n - V_t^n| \le Z_n |t'-t|^\gamma$$

and

$$E\{\sup_{t\le 1} |V_t^n|^r\} \le A',\quad E\{Z_n^r\} \le A'.$$

Now by the general theory the optional projection of V^n will be right continuous and have left limits. But

$$(7.16) \qquad E\{V_t^n | \underline{\underline{F}}_t\} = \int_{R\times[0,t]} g(x,s,t)M^n(dx\,ds) = U_t^n \quad P^n\text{-a.s.},$$

so the optional projection of V is a version of U, and U indeed does have a right continuous version.

Note that

$$|U_t^n| \le E\{|V_t^n| | \underline{\underline{F}}_t\} \le E\{\sup_s |V_s^n| \; |\underline{\underline{F}}_t\} \overset{\text{def}}{=} S_t \quad .$$

Thus

$$E\{ \sup_t |U_t^n|^r\} \le E\{ \sup_t S_t^r\} \le \frac{r^2}{r-1} E\{S_1^r\}$$

by Doob's inequality. But this is

$$\le \frac{r^2}{r-1} A' = A.$$

which proves (ii).

Let us skip (iii) for the moment, and prove (iv). Let $\delta_n > 0$ and let T_n be a stopping time for M^n such that $T_n + \delta_n \le 1$.

$$\begin{aligned} E\{|U^n_{T_n+\delta_n} - U^n_{T_n}|\} &= |E\{V^n_{T_n+\delta_n} | \underline{\underline{F}}_{T_n+\delta_n}\} - E\{V^n_{T_n} | \underline{\underline{F}}_{T_n}\}| \\ &\le E\{|V^n_{T_n+\delta_n} - V_{T_n}| \; |\underline{\underline{F}}_{T_n+\delta_n}\} + |E\{V^n_{T_n} | \underline{\underline{F}}_{T_n+\delta_n}\} - E\{V^n_{T_n} | \underline{\underline{F}}_{T_n}\}| \\ &= E\{|V^n_{T_n+\delta_n} - V^n_{T_n}| \underline{\underline{F}}_{T_n+\delta_n}\} + |\int_{R\times(T_n+\delta_n]} g(x,s,T_n)M^n(dx\,ds)| \end{aligned}$$

Similarly,

$$E\{|U^n_{T_n+\delta_n} - U^n_{T_n}|^2\}^{1/2} \le E\{|V^n_{T_n+\delta_n} - V^n_{T_n}|^2\}^{1/2} + E\{ \int_{T_n}^{T_n+\delta_n} \|g(s,T_n)\|_\infty^2 dk_n(s)\}^{1/2} \; .$$

We can estimate the increment of V^n by (ii) and use the hypotheses on g and k_n in the second term:

$$\le \delta^\gamma E^n\{Z_n^2\}^{1/2} + C_r E^n\{Y^2 X_n\}^{1/2} \delta_n^{1/2} \; .$$

Since both expectations are bounded, we see that if $\delta_n \to 0$, then $U^n_{T_n+\delta_n} - U^n_{T_n} \to 0$ in L^2, hence in probability. Since the U^n are bounded in L^r by (ii), the family (U^n) is tight by Aldous' theorem. This proves (iv).

If now $M^n \Rightarrow M^0$, then Proposition 7.12 implies that the finite-dimensional distributions of U_t^n converge to those of U_t^0. Since the (U^n) are tight, this implies that $U^n \Rightarrow U$.

It remains to prove (iii). Suppose M^n is continuous, and compute

$$E\{|U^n_{t+s} - U^n_t|^r\}^{1/r} \le E\{\int_{R^d\times[0,t]} (g(x,u,t+s) - g(x,u,t))M^n(dx\,du)|^r\}^{1/r}$$

$$+ E\{|\int_{R^d\times(t,t+s]} g(x,u,t+s)M(dx\,du)|^r\}^{1/r}.$$

By Burkholder's inequality and (7.4) - and this time M is continuous, so $\Delta M^* \equiv 0$ - this is

$$\le C_r E\{|\int_0^t \|g(u,t+s) - g(u,t)\|_h^2 dk_n(u)|^{r/2}\}^{1/r}$$

$$+ C_r E\{|\int_t^{t+s} \|g(u,t+s)\|_h^2 dk_n(u)|^{r/2}\}^{1/r}.$$

Since g is of Hölder class (β,q,K) and M^n is of class (p,K):

$$\le C_r\ E\{Y^r X_n^{r/2}\}^{1/r}(s^\beta t^{1/2} + s^{1/2}).$$

By (7.14) we conclude that there is a constant B such that for $0 \le s \le 1$

$$E\{|U^n_{t+s} - U^n_t|^r\} \le B\, s^{\frac{r}{2}\wedge\beta r}.$$

But $\beta r > 1$ and $r/2 > p/2 > 1$ so that (iii) follows by Corollary 1.2.

Q.E.D.

CHAPTER EIGHT

THE BROWNIAN DENSITY PROCESS

Imagine a large but finite number - say 6×10^{23} - of Brownian particles B^1, $B^2, \ldots, B^N$ diffusing through a region $D \subset R^d$. Consider the density of particles at a point x at time t. This is of course just the number of particles per unit volume. We can approximate it by counting the number in, say, a small cube centered at x and dividing by the volume. But is this a good approximation? Not really. If we let the cube shrink to the point $\{x\}$, the limit will either be zero, if there is no particle at x, or infinity, if there is one. So we can't take a limit. We'll have to stick to finite sizes of cubes. Since there seems to be no reason to prefer one size cube to another, perhaps we should do it for <u>all</u> cubes. Once that is admitted, one might ask why we should restrict ourselves to cubes, and whether other forms of averaging might be equally relevant. For instance, if ϕ is a positive function of compact support such that $\int \phi(x)dx = 1$, one could define the ϕ-average as

$$\eta_t(\phi) = \sum_{i=1}^{N} \phi(B_t^i). \tag{8.1}$$

Finally, we might as well go all the way and compute (8.1) for all test functions ϕ. This will give us a Schwartz distribution. In short we will describe the density of particles by the Schwartz distribution (8.1).

It is usually easier to deal with a continuous process than a discrete process such as this, so we might let $N \to \infty$, re normalize, and see if the process η_t goes to a limit. It does, and this limit is what we call the <u>Brownian density process</u>.

From what we have said above, one might think that we should describe the particle density by a measure, rather than a distribution, for (8.1) defines both. It is only when we take the limit as $N \to \infty$ that the reason for the choice becomes clear. In general, the limit will be a pure distribution, not a measure.

We will add one complication: we will consider branching Brownian motions, rather than just Brownian motions. The presence of branching gives us a more interesting class of limit processes. However, aside from that, we shall operate in the simplest possible setting. It is possible to generalize to branching diffusions, or even branching Hunt processes, but most of the ideas needed in the general case are already present in this elementary situation.

A branching Brownian motion with parameter μ can be described as follows. A particle performs Brownian motion in some region of $\mathbf{R}^d$. In a time interval (t, t+h) the particle has a probability $\mu h + o(h)$ of branching. If it does branch, it either splits into two identical particles, or it dies out, with probability 1/2 each. If it splits, the two daughters begin their life at the branching point. They continue independently as Brownian motions until the time they themselves branch, and so on.

If we start with a single particle and let N_t be the total number of particles at time t, then N_t is an ordinary branching process, and also a martingale. Note that we are in the critical case, $E\{N_t\} \equiv 1$.

We are going to assume that the initial distribution of particles is a Poisson point process on $\mathbf{R}^d$ of parameter λ: the number of particles in a set A is a Poisson random variable with parameter $\lambda|A|$, and the numbers in disjoint sets are independent. This means that the initial number of particles is infinite, but that will not bother us; the Poisson initial distribution makes things easier rather than harder.

Let us first give an explicit construction of branching Brownian motion. There are numerous constructions in the literature, most of which are more sophisticated than this, which is done entirely by hand, but it sets up the process in a useful form.

Let $\underline{\underline{A}}$ be the set of all multi indices, i.e. of strings of the form $\alpha = n_1 n_2 \cdots n_k$ where the n_j are non-negative integers. Let $|\alpha|$ be the length of α. We provide $\underline{\underline{A}}$ with the <u>arboreal ordering</u>: $m_1 \cdots m_p \prec n_1 \cdots n_q$ iff $p \le q$ and $m_1 = n_1, \ldots, m_p = n_p$. If $|\alpha| = p$, then α has exactly p-1 predecessors, which we shall denote respectively by $\alpha-1, \alpha-2, \ldots, \alpha-|\alpha|+1$. That is, if $\alpha = 2341$, then $\alpha-1 = 234$, $\alpha-2 = 23$ and $\alpha-3 = 2$.

Let Π^λ be a Poisson point process on $\mathbf{R}^d$ of parameter λ. The probability that any two points of Π^λ lie exactly the same distance from the origin is zero, so that we can order them by magnitude. Thus the initial values can be denoted by

$$\{X^\alpha(0),\ \alpha \in \underline{\underline{A}},\ |\alpha| = 1\}.$$

Define three families

$$\{B_t^\alpha,\ t \geq 0,\ \alpha \in \underline{\underline{A}}\},\quad \{S_t^\alpha,\ \alpha \in \underline{\underline{A}}\}\quad \text{and}\quad \{N^\alpha,\ \alpha \in \underline{\underline{A}}\},$$

where the B^α are independent standard Brownian motions in $\mathbf{R}^d$ with $B_0^\alpha = 0$, all α; the S^α are i.i.d. exponential random variables with parameter μ, which will serve as lifetimes; and the N^α are i.i.d. random variables with $P\{N^\alpha = 0\} = P\{N^\alpha = 2\} = 1/2$. The families (B^α), (S^α), (N^α), and $(X^\alpha(0))$ are independent.

The <u>birth time</u> $\beta(\alpha)$ of X^α is

$$\beta(\alpha) = \begin{cases} \sum_{1}^{|\alpha|-1} S^{\alpha-j} & \text{if } N^{\alpha-j} = 2,\ j = 1,\dots,|\alpha|-1 \\ \infty & \text{otherwise} \end{cases}.$$

The <u>death time</u> $\zeta(\alpha)$ of X^α is

$$\zeta(\alpha) = \beta(\alpha) + S^\alpha.$$

Define $h^\alpha(t) = I_{\{\beta(\alpha) \leq t < \zeta(\alpha)\}}$, which is the indicator function of the lifespan of X^α.

If $\alpha = n_1 n_2 \cdots n_p$ (so $\alpha - p+1 = n_1$) then the <u>birthplace</u> $X^\alpha(\beta(\alpha))$ is

$$X^\alpha(\beta(\alpha)) = X^{n_1}(0) + \sum_{i=1}^{|\alpha|-1} (B^{\alpha-i}_{\zeta(\alpha-i)} - B^{\alpha-i}_{\beta(\alpha-i)}).$$

Now let ∂ - the cemetary - be a point disjoint from $\mathbf{R}^d$, and put

$$X^\alpha(t) = \begin{cases} \partial & \text{if } \quad t < \beta(\alpha) \text{ or } t \geq \zeta(\alpha) \\ X^\alpha(\beta(\alpha)) + \int_0^t h^\alpha(s)\,dB_s^\alpha & \text{otherwise.} \end{cases}$$

Note that since $h^\alpha = 1$ between β and ζ, X^α is a Brownian motion on the interval $[\beta(\alpha), \zeta(\alpha))$, and $X^\alpha = \partial$ outside it. We make the convention that any function f on $\mathbf{R}^d$ is extended to $\mathbf{R}^d \cup \{\partial\}$ by $f(\partial) = 0$.

Finally, define

$$\eta_t(\phi) = \sum_{\alpha \in \underline{\underline{A}}} \phi(X_t^\alpha) \tag{8.2}$$

for any ϕ on $\mathbf{R}^d$ for which the sum makes sense.

The branching process η_t^n starting at the single point $X^n(0)$ is constructed from the X^α for $n \prec \alpha$. That is,

$$\eta_t^n(\phi) = \sum_{\substack{\alpha \in \underline{\underline{A}} \\ n \prec \alpha}} \phi(X_t^\alpha); \quad N_t^n = \sum_{\substack{\alpha \in \underline{\underline{A}} \\ n \prec \alpha}} h^\alpha(t). \tag{8.3}$$

Note that N_t^n is the number of particles alive at time t, and it is a classical branching process of the critical case - we leave it to the reader to verify that this does indeed follow from our construction - hence $E\{N_t^n\} = 1$. In particular, $\eta_t^n(\phi)$ is finite for all t.

Furthermore, a symmetry argument shows that, given $h^\alpha(t) = 1$ (so X^α is alive at t),

$P\{X_t^\alpha \in A | X^n(0) = x\} = P\{x + B_t \in A\}$, where $n \prec \alpha$. (This relies on the fact that all birth and death times are independent of the B_t^α.) Thus

$$E\{\sum_{n\prec\alpha} \phi(X_t^\alpha) | X^n(0) = x\} = \sum_{n\prec\alpha} E\{\phi(X_t^\alpha) | X^n(0) = x, h^\alpha(t) = 1\} P\{h^\alpha(t) = 1 | X^n(0) = x\}$$

$$= E\{\phi(B_t + x)\} \sum_{n\prec\alpha} P\{h^\alpha(t) = 1\}$$

$$= E\{\phi(B_t + x)\}.$$

Suppose now that $\phi \geq 0$. We can integrate over the initial values, which are Poisson(λ), to see that

$$E\{\sum_\alpha \phi(X_t^\alpha)\} = \int \lambda \, dx \, E\{\phi(x + B_t)\}$$

$$= \lambda \iint (2\pi t)^{-d/2} e^{-\frac{|y-x|^2}{2t}} \phi(y) dx \, dy .$$

Since Lebesgue measure is invariant this is

$$= \lambda \int \phi(x) dx .$$

Thus, for any positive ϕ,

$$E\{\eta_t(\phi)\} = \lambda \int \phi(x) dx. \tag{8.4}$$

This makes it clear that $\eta_t(\phi)$ makes sense for any integrable ϕ.

THE FUNDAMENTAL NOISES

There are two distinct sources of randomness for our branching Brownian motion. The first comes from the diffusion, and the second comes from the branching. We will treat these separately.

Recall that λ is the parameter of the initial distribution and μ is the branching rate. Define, for a Borel set $A \subset R^d$,

$$(8.5) \quad \begin{cases} W(A \times (0,t]) = \lambda^{-1/2} \sum_{\alpha \in \underline{\underline{A}}} \int_0^t h^\alpha(s) I_{\{X_s^\alpha \in A\}} dB_s^\alpha \\ \\ Z(A \times (0,t]) = (\lambda\mu)^{-1/2} \sum_{\alpha \in \underline{\underline{A}}} (N^\alpha - 1) I_{\{\zeta(\alpha) \leq t, \, X^\alpha(\zeta(\alpha)-) \in A\}} \end{cases}$$

(If $\mu = 0$, $Z \equiv 0$).

These take some explanation. In the expression for W, $h^\alpha \equiv 1$ when X^α is alive so $dB^\alpha = dX^\alpha$, and the integral gives us the total amount that X^α diffuses while in A; the sum over α gives the total diffusion of all the particles while they are in A. Thus W keeps track of the diffusion and ignores the branching.

Z does just the opposite. N^α is the size of family foaled by X^α when it branches, at time $\zeta(\alpha)$. $N^\alpha - 1$ counts +1 if X^α splits in two, -1 if it dies without progeny. Thus $Z(A \times (0,t])$ is the number of births minus the number of deaths occurring in A up to time t.

Let us put

$$W_t(A) = W(A \times (0,t]); \qquad Z_t(A) = Z_t(A \times (0,t]).$$

PROPOSITION 8.1 W and Z are orthogonal martingale measures. W is continuous and R^d -valued while Z is purely discontinuous and real-valued. Moreover,

(i) $\langle W(A)\rangle_t = \langle Z(A)\rangle_t I = \frac{1}{\lambda}\left(\int_0^t \eta_s(A)ds\right)I;$

(ii) $\langle W_{ij}(A), Z(C)\rangle_t = 0$ for all Borel A and C of finite measure and all $i, j \leq n$;

(iii) both W and Z have the same mean and covariance as white noises (based on Lebesgue measure).

Note. W_t has values in R^d, so $\langle W\rangle_t$ is a $d\times d$ matrix whose ijth component is $\langle W^i, W^j\rangle_t$.

PROOF. We must first show W and Z are square integrable. This done, the rest is easy. First look at the branching diffusion starting from one of the $X^n(0)$. Define

$$W_t^n(A) = \lambda^{-1/2} \sum_{n<\alpha} \int_0^t h^\alpha(s)\, I_{\{X_s^\alpha \in A\}}\, dB_s^\alpha.$$

Since the B^α are independent, hence orthogonal,

$$\langle W^n(A)\rangle_t = \lambda^{-1}\left(\sum_{n<\alpha} \int_0^t h^\alpha(s)\, I_{\{X_s^\alpha \in A\}}\, ds\right)I$$

$$= \lambda^{-1}\left(\int_0^t \eta_s^n(A)ds\right)I .$$

The interchange is justified since the sum is dominated by $\int_0^t N_s^n ds$ which has expectation t. (I is the d×d identity matrix.)

It follows that $E\{|W_t^n(A)|^2\} = E\{|\langle W^n(A)\rangle_t|\} \le td/\lambda$.

Turning to Z, let

$$M_t^\alpha = (N^\alpha - 1)I_{[\zeta(\alpha),\infty)}(t)I_{\{X^\alpha(\zeta(\alpha)-) \in A\}}.$$

Note that this is adapted to the natural σ-fields $(\underline{\underline{F}}_t)$. Since $N^\alpha-1$ has mean zero and is independent of the indicator functions, M^α is a martingale. Noticing that $\zeta(\alpha) = \beta(\alpha) + S^\alpha$, where S^α is exponential(μ), it is easy to see that

$$\langle M^\alpha\rangle_t = \int_0^t h^\alpha(s)I_{\{X_s^\alpha \in A\}}\mu ds.$$

The N^α are independent, so the M^α are orthogonal for different α and

$$\langle Z_t^n(A)\rangle = \frac{1}{\lambda\mu}\sum_{n<\alpha}\langle M^\alpha\rangle = \frac{1}{\lambda}\sum_{n<\alpha}\int_0^t h^\alpha(s)\, I_{\{X_s^\alpha \in A\}}ds$$

$$= \frac{1}{\lambda}\int_0^t \eta_s^n(A)ds.$$

The processes η^n, n = 0,1,2,.. are conditionally independent given $\underline{\underline{F}}_0$, so if m ≠ n

$$\langle W^m(A), W^n(A)\rangle_t = \langle Z^m(A), Z^n(A)\rangle_t = 0.$$

Then

(8.6)
$$\langle W(A)\rangle_t = \frac{1}{\lambda}\sum_n \langle W^n(A)\rangle_t$$

$$= \frac{1}{\lambda}\Big(\sum_n \int_0^t \eta_s^n(A)ds\Big)I$$

$$= \frac{1}{\lambda}\Big(\int_0^t \eta_s(A)ds\Big)I.$$

If A has finite Lebesgue measure, this is finite and even integrable by (8.4). The same reasoning applies to $Z_t(A)$.

Now that we know $W_t(A)$ and $Z_t(A)$ are square-integrable, we can read off their properties directly from the definition. If $A \cap C = \phi$ and if they have finite measure,

$$\langle W(A), W(C)\rangle_t = \sum_{\alpha,\gamma\in\underline{\underline{A}}}\int_0^t h^\alpha(s)h^\gamma(s)\, I_{\{X_s^\alpha \in A,\ X_s^\gamma \in C\}}d\langle B^\alpha, B^\gamma\rangle_s.$$

Now $\langle B^\alpha, B^\gamma\rangle_t = 0$ unless α = γ, and if α = γ, the indicator function vanishes, so $\langle W(A), W(C)\rangle_t = 0$. Similarly, $\langle Z(A), Z(C)\rangle_t = 0$. Furthermore both $A \to W_t(A)$ and $A \to Z_t(A)$ are L^2-valued measures. This is clear for Z from (8.5) and almost clear for W, so we leave the verification to the reader. This proves (i).

If A and C are two sets of finite Lebesgue measure, $W_t(A)$ is a continuous martingale while $Z_t(A)$ is purely discontinuous, so the two are orthogonal, proving (ii).

To see (iii), recall W and Z have mean zero and note that by (i)

$$E\{W_s(A)W_t(C)^T\} = E\{\langle W(A), W(C)\rangle_{s\wedge t}\}I$$
$$= (\frac{1}{\lambda}\int_0^{s\ t} E\{\eta_s(A\cap C)\}ds)I$$
$$= (s\ t)|A\cap C|I$$

by (8.4). Exactly the same calculation holds for Z, except that I = 1 since Z is real-valued. This is the covariance of white noise.

Q.E.D.

Corollary 8.2 If $\phi(x)$ is deterministic

$$E\{(\int_0^t \int_{R^d} \phi(x)W(dxds))^2\} = E\{\int_0^t \int_{R^d} (\phi(x)Z(dxds))^2\} = t\int_{R^d} \phi^2(x)dx .$$

It is clear how to integrate with respect to Z. Here is a fundamental identity for integrals with respect to W.

PROPOSITION 8.3 Let $\phi(x) \in L^2(R^d)$. Then

$$\sum_{\alpha\in\underline{\underline{A}}} \int_0^t \phi(X_s^\alpha)h^\alpha(s)dB_s^\alpha = \sqrt{\lambda}\int_0^t \int_{R^d} \phi(x)\ W(dxds) . \tag{8.7}$$

PROOF. If $\phi(x) = I_A(x)$, this is true by the definition of W, for both sides equal $W_t(A)$. It follows that (8.7) holds for finite sums of such processes, hence for all $\phi \in L^2(R^d)$ by the usual argument. Q.E.D.

AN INTEGRAL EQUATION FOR η

Let ϕ be a real-valued $C^{(2)}$ function on R^d. By convention, $\phi(X_t^\alpha) = 0$ unless $t \in [\beta(\alpha), \zeta(\alpha))$; for t in this interval we can apply Ito's formula:

$$\phi(X_t^\alpha) = \phi(X_\beta^\alpha) + \int_0^t h^\alpha(s)\ \nabla\phi(X_s^\alpha)\cdot dB_s^\alpha + \frac{1}{2}\int_0^t h^\alpha(s)\ \Delta\phi(X_s^\alpha)ds.$$

(If $t > \zeta(\alpha)$, we must subtract $\phi(X^\alpha_{\zeta-})$ from this.) Thus for any test function ϕ:

$$\begin{aligned}\eta_t(\phi) = \eta_0(\phi) &- \sum_{|\alpha|=1} \phi(X^\alpha_{\zeta-}) I_{\{\zeta(\alpha)\le t\}} \\ &+ \sum_{|\alpha|\ge 2} (\phi(X^\alpha_\beta) I_{\{\beta(\alpha)\le t\}} - \phi(X^\alpha_{\zeta-}) I_{\{\zeta(\alpha)\le t\}}) \\ &+ \sum_\alpha \int_0^t h^\alpha(s) \nabla\phi(X^\alpha_s)\cdot dB^\alpha_s \\ &+ \sum_\alpha \frac{1}{2} \int_0^t h^\alpha(s)\, \Delta\phi(X^\alpha_s)\, ds \; .\end{aligned}$$

We can identify all these terms. $\eta_0(\phi)$ is of course the initial value. The next two sums give all the births and deaths, and combine into $\sqrt{\lambda\mu} \int_0^t \int_{\mathbf{R}^d} \phi(x) Z(dxds)$.

Proposition 8.3 applies to the stochastic integral, which equals

$\sqrt{\lambda} \int_0^t \int_{\mathbf{R}^d} \nabla\phi(x)\cdot W(dxds)$. The final integral is $\frac{1}{2} \int_0^t \eta_s(\Delta\phi)ds$. Thus we have proved

<u>PROPOSITION 8.4</u>. Let $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$. Then

(8.8)
$$\eta_t(\phi) = \eta_0(\phi) + \frac{1}{2} \int_0^t \eta_s(\Delta\phi)ds + \sqrt{\lambda\mu} \int_0^t \int_{\mathbf{R}^d} \phi(x) Z(dxds)$$
$$+ \sqrt{\lambda} \int_0^t \int_{\mathbf{R}^d} \nabla\phi(x)\cdot W(dxds) \; .$$

If we check equations (5.3) and (5.4), we see this translates into the SPDE

(8.9)
$$\begin{cases} \dfrac{\partial\eta}{\partial t} = \dfrac{1}{2} \Delta\eta + \sqrt{\lambda\mu}\, \dot{Z} + \sqrt{\lambda}\, \nabla\cdot\dot{W} \\ \eta_0 = \Pi^\lambda \end{cases}$$

If $G_t(x,y) = (2\pi t)^{-d/2} e^{-\frac{|y-x|^2}{2t}}$, Theorem 5.1 tells us the solution of (8.8) and (8.9) is

(8.10)
$$\eta_t(\phi) = \int_{\mathbf{R}^d} G_t(\phi,y)\Pi^\lambda(dy) + \sqrt{\lambda\mu} \int_0^t \int_{\mathbf{R}^d} G_{t-s}(\phi,y) Z(dyds)$$
$$+ \sqrt{\lambda} \int_0^t \int_{\mathbf{R}^d} \nabla G_{t-s}(\phi,y)\cdot W(dyds).$$

Note that in the last integral the ith component of $\nabla G_t(\phi,y)$ is $G_t(\frac{\partial\phi}{\partial x}, y)$ so we will write $\nabla G_t(\phi,y) = G_t(\nabla\phi,y)$ below.

If $\phi \in L^1(\mathbf{R}^d)$, let $\langle\phi\rangle = \int_{\mathbf{R}^d} \phi(x)dx$. From (8.4)

$$E\{\eta_t(\phi)\} = \lambda\langle\phi\rangle, \quad t \geq 0.$$

We want to estimate the higher moments - at least the second moments - of η.

Let $\eta_t^*(\phi) = \sup_{s\leq t} |\eta_s(\phi)|$.

PROPOSITION 8.5. There exist continuous functions $C_1(\phi,t)$ and $C_2(\phi,t)$ such that

$$E\{\eta_t^*(\phi)^2\} \leq \lambda^2 C_1(\phi,t) + \lambda(\mu + 1)C_2(\phi,t) \tag{8.11}$$

PROOF. Note that the three terms on the right-hand side of (8.10) are orthogonal, so, remembering that Z and N have the same covariance as white noise, (Cor. 8.2) we have

$$\begin{aligned} E\{\eta_t(\phi)^2\} &= E\{(\int_{\mathbf{R}^d} G_t(\phi,y)\Pi^\lambda(dy))^2\} \\ &\quad + \lambda\mu \int_0^t \int_{\mathbf{R}^d} G_{t-s}^2(\phi,y)dyds \\ &\quad + \lambda \int_0^t \int_{\mathbf{R}^d} |G_{t-s}(\nabla\phi,y)|^2 dyds \\ &= \lambda(\langle G_t^2(\phi,\cdot)\rangle + \lambda\langle G_t(\phi,\cdot)\rangle^2) + \lambda\mu \int_0^t \langle G_{t-s}^2(\phi,\cdot)\rangle ds \\ &\quad + \lambda \int_0^t \langle|G_{t-s}(\nabla\phi,\cdot)|^2\rangle ds . \end{aligned}$$

Now $\langle G_t^2(\phi,\cdot)\rangle$ is bounded uniformly in t, so the above is bounded by

$$E\{\eta_t(\phi)^2\} \leq \lambda^2 c_1(\phi) + t\lambda(\mu + 1)c_2(\phi) \tag{8.12}$$

for some c_1 and c_2. From (8.8)

$$\begin{aligned} E\{\eta_t^*(\phi)^2\} &\leq 9\, E\{\eta_0(\phi)^2\} + \frac{9}{4} E\{(\int_0^t \eta_s(|\phi|)ds)^2\} \\ &\quad + 9\, E\{\sup_{s\leq t} (\sqrt{\lambda\mu} \int_0^s \int_{\mathbf{R}^d} \phi dZ + \sqrt{\lambda} \int_0^s \int_{\mathbf{R}^d} \nabla\phi \cdot dW)^2\} . \end{aligned}$$

Now

$$\begin{aligned} E\{\eta_0(\phi)^2\} &= E\{(\int \phi(x)\, \Pi^\lambda(dx))^2\} \\ &= \lambda \langle\phi^2\rangle + \lambda^2\langle\phi\rangle^2. \end{aligned}$$

Apply Schwartz' inequality and (8.12) to the second term:

$$E\{(\int_0^t \eta_s(\phi)ds)^2\} \leq t \int_0^t E\{\eta_s^2(\phi)\}ds$$

$$\leq \lambda^2 t^2 c_1(\phi) + \lambda(\mu + 1) \frac{t^3}{2} c_2(\phi).$$

Apply Doob's inequality to see the third term is bounded by

$$36\ E\{(\sqrt{\lambda\mu} \int_0^t \int_{R^d} \phi\ dZ + \sqrt{\lambda} \int_0^t \int_{R^d} \nabla\phi \cdot dW)^2\}\ .$$

The two stochastic integrals are orthogonal, so by Cor.8.2:

$$= 36\ (t\ \lambda\mu\langle\phi^2\rangle + \lambda t\ \langle|\nabla\phi|^2\rangle).$$

Put these three estimates together to get (8.11).

Exercise 8.1. Show η_t^* is bounded in L^p for all $p > 0$. (Hint: By Prop. 8.1 and Burkholder's inequality, η_t $L^p \Rightarrow W$ and Z are in L^{2p}. Use induction on $p = 2^n$ to see η_t L^p for all p, then use (8.8) and Doob's L^p inequality as above.)

THEOREM 8.6. Let (μ_n, λ_n) be sequence of parameter values and let (W^n, Z^n) be the corresponding processes. Let $V^n(dx) = \lambda_n^{-1/2}(\Pi^{\lambda_n}(dx) - \lambda_n dx)$ be the normalized initial measure. If the sequence $((\mu_n+1)/\lambda_n)$ is bounded, then (V^n, W^n, Z^n) is tight on $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(R^{d^2+2d})\}$.

PROOF. We regard V^n as a constant process: $V_t^n \equiv V^n$, in order to define it on $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(R^{d^2+2d})\}$. It is enough to prove the three are individually tight.

By Mitoma's theorem it is enough to show that $(V^n(\phi))$, $(W^n(\phi))$, and $(Z^n(\phi))$ are each tight for each $\phi \in \underline{\underline{S}}(R^d)$.

In the case of V^n, which is constant in t, it is enough to notice that $E\{V^n(\phi)^2\} = \langle\phi^2\rangle$ is uniformly bounded in n.

We will use Kurtz' criterion (Theorem 6.8b) for the other two. Set

$$A_n(\delta) = (\delta d/\lambda_n)\eta_1^{n*}(\phi),\ 0 \leq \delta \leq 1\ .$$

Then $\sup_{t\leq 1} (|\ \langle W^n(\phi)\rangle_{t+\delta} - \langle W^n(\phi)\rangle_t|) \leq A_n(\delta)$ and

$$\sup_{t\leq 1} (\langle Z^n(\phi)\rangle_{t+\delta} - \langle Z^n(\phi)\rangle_t) \leq A_n(\delta)$$

by Proposition 8.1. By Jensen's inequality and (8.11),

$$\lim_{\delta\to 0}\limsup_{n\to\infty} E\{A_n(\delta)\} \le \lim_{\delta\to 0}\limsup_{n\to\infty} \frac{\delta d}{\lambda_n} E\{\eta_1^{n*}(\phi)^2\}^{1/2}$$

$$\le \lim_{\delta\to 0} \delta d(c_1 + \frac{\mu_n+1}{\lambda_n} c_2)$$

$$= 0$$

since $(\mu_n + 1)/\lambda_n$ is bounded. Furthermore

$$E\{|W_t^n(\phi)|^2\} = d\, E\{Z_t^n(\phi)^2\} = td < \phi^2>$$

for all n so that for each t, $(W_t^n(\phi))$ and $(Z_t^n(\phi))$ are tight on R^d and R respectively. By Theorem 6.8 the processes $(W^n(\phi))$ and $(Z^n(\phi))$ are each tight.

Q.E.D.

<u>THEOREM 8.7</u>. If $\lambda_n \to \infty$, $\mu_n\lambda_n \to \infty$, and $\mu_n/\lambda_n \to 0$, then $(V^n, Z^n, W^n) \Rightarrow (V^0, Z^0, W^0)$, where V^0, Z^0 and W^0 are white noises based on Lebesgue measure on R^d, $R^d \times R_+$, and $R^d \times R_+$ respectively; V^0 and Z^0 are real-valued and W^0 has values in R^d. If $\lambda_n \to \infty$ and $\mu_n/\lambda_n \to 0$, $(V^n, W^n) \Rightarrow (V^0, W^0)$.

<u>PROOF</u>. Suppose λ_n is an integer. Modifications for non-integral λ are trivial. To show weak convergence, we merely need to show convergence of the finite-dimensional distributions and invoke Theorem 6.15.

The initial distribution is Poisson (λ_n) and can thus be written as a sum of λ_n independent Poisson (1) point processes.

Let $\hat{\eta}^1$, $\hat{\eta}^2$,... be a sequence of iid copies with $\lambda = 1$, $\mu = \mu_n$. (We have changed notation: these are not the $\hat{\eta}^n$ used in constructing the branching Brownian motion.) Then the branching Brownian motion corresponding to λ_n, μ_n has the same distribution as $\hat{\eta}^1 + \hat{\eta}^2 + \dots + \hat{\eta}^{\lambda_n}$. Define $\hat{V}^1$, $\hat{V}^2,\dots$, $\hat{W}^1$, $\hat{W}^2,\dots$ and $\hat{Z}^1$, $\hat{Z}^2,\dots$ in the obvious way. Then

$$V^n = \hat{V}^1 + \dots + \hat{V}^{\lambda_n}, \quad W^n = \frac{\hat{W}^1+\dots+\hat{W}^{\lambda_n}}{\sqrt{\lambda_n}}, \quad Z^n = \frac{\hat{Z}^1+\dots+\hat{Z}^{\lambda_n}}{\sqrt{\lambda_n}} .$$

We have written everything as sums of independent random variables. To finish the proof, we will call on the classical Lindeberg theorem.

Let $\phi_1, \phi_1', \phi_1'',\dots,\phi_p, \phi_p', \phi_p'' \in \underline{\underline{S}}'(R)$, $t_1 \le t_2 \le \dots \le t_p$. We must show weak convergence of the vector

$$U^n \overset{\text{def}}{=} (V^n(\phi_1),\dots,V^n(\phi_p),\ W^n_{t_1}(\phi'_1),\dots,W^n_{t_p}(\phi'_p),\ Z^n_{t_1}(\phi''_1),\dots,Z^n_{t_p}(\phi''_p))$$

This can be written as a sum of iid vectors, and the mean and covariance of the vectors are <u>independent of n</u> (Prop. 8.1).

It is enough to check the Lindeberg condition for each coordinate. The distribution of V^n does not depend on μ_n, so we leave this to the reader.

Fix i and look at $W^n_{t_i}(\phi'_i) = \lambda^{-1/2} \sum_{k=1}^{\lambda_n} \hat{W}^k_{t_i}(\phi'_i)$. Now $(\hat{W}^k_t(\phi'_i))$ is an $\mathbf{R}^d$-valued continuous martingale, so by Burkholder's inequality

$$E\{|\hat{W}^k_{t_i}(\phi'_i)|^4\} \le c_4\, E\{|\langle \hat{W}^k(\phi_i)\rangle|^2\}$$

$$\le td\, c_4 \int_0^t E\{\hat{\eta}^k_s(\phi_i^2)^2\}ds \quad .$$

Now $t \le 1$ so by Proposition 8.5 with $\lambda = 1$, there is a C independent of k and λ_n such that this is

$$\le C(\mu_n + 1).$$

For $\varepsilon > 0$,

$$E\{|\hat{W}^k_{t_i}(\phi_i)|^4;\ |\hat{W}^k_{t_i}(\phi)|^2 \ge \lambda_n\varepsilon\} \le E\{|\hat{W}^k_{t_i}(\phi_i)|^4\}^{1/2} P\{|\hat{W}^k_{t_i}|^4 \ge \lambda_n^2\varepsilon^2\}^{1/2}$$

by Schwartz. Use Chebyshev with the above bound:

$$\le [C(1+\mu_n)]^{1/2}[C(1+\mu_n)/\lambda_n^2\varepsilon^2]^{1/2}$$

$$\le C(1+\mu_n)/\lambda_n\varepsilon.$$

Thus

$$\sum_{k=1}^{\lambda_n} E\{|\lambda_n^{-1/2}\hat{W}^k_{t_i}(\phi_i)|^2;\ |\lambda_n^{-1/2}\hat{W}^k_{t_i}(\phi_i)|^2 > \varepsilon\}$$

$$= E\{|\hat{W}^1_{t_i}(\phi_i)|^2:\ |\hat{W}^1_{t_i}(\phi_i)|^2 > \lambda_n\,\varepsilon\}$$

$$\le C_3(1+\mu_n)/\lambda_n\varepsilon \to 0.$$

Thus the Lindeberg condition holds for each of the $W^n_{t_i}(\phi'_i)$. The same argument holds for the $Z^n_{t_i}(\phi''_i)$. In this case, while $(\hat{Z}^n_t(\phi''_i))$ is not a continuous martingale, its jumps are uniformly bounded by $(\lambda_n\mu_n)^{-1/2}$, which goes to zero, and we can apply Burkholder's inequality in the form of Theorem 7.11(i). Thus the finite-dimensional distributions converge by Lindeberg's theorem, implying weak convergence.

The only place we used the hypothesis that $\lambda_n \mu_n \to \infty$ was in this last statement, so that if we only have $\lambda_n \to \infty$, $\mu_n/\lambda_n \to 0$, we still have $(V^n, W^n) \Rightarrow (V^0, W^0)$.

Q.E.D.

We have done the hard work and have arrived where we wanted to be, out of the woods and in the cherry orchard. We can now reach out and pick our results from the nearby boughs.

Define, for n = 0, 1, ...

$$R_t^n(\phi) = \int_0^t \int_{\mathbf{R}^d} G_{t-s}(\nabla\phi, y) \cdot W^n(dy\, ds)$$

$$U_t^n(\phi) = \int_0^t \int_{\mathbf{R}^d} G_{t-s}(\phi, y) Z^n(dy\, ds).$$

Recall from Proposition 7.8 that convergence of the martingale measures implies convergence of the integrals. It thus follows immediately from Theorem 8.7 that

COROLLARY 8.8. Suppose $\lambda_n \to \infty$ and $\mu_n/\lambda_n \to 0$. Then

(i) $(V^n, W^n, R^n) \Rightarrow (V^0, W^0, R^0)$;

(ii) if, in addition, $\lambda_n \mu_n \to \infty$,

$$(V^n, W^n, Z^n, R^n, U^n) \Rightarrow (V^0, W^0, Z^0, R^0, U^0).$$

Rewrite (8.10) as

$$\frac{\eta_t(\phi) - \lambda\langle\phi\rangle}{\sqrt{\lambda}} = V(G_t(\phi,\cdot)) + \sqrt{\mu}\, U_t(\phi) + R_t(\phi). \tag{8.13}$$

In view of Corollary 8.6 we can read off all the weak limits for which $\frac{\mu}{\lambda} \to 0$.

THEOREM 8.9 (i) If $\lambda_n \to \infty$ and $\mu_n \to 0$, then $\dfrac{\eta_t^n(\phi) - \lambda_n\langle\phi\rangle}{\sqrt{\lambda_n}}$ converges in $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(\mathbf{R}^d)\}$ to a solution of the SPDE

$$\begin{cases} \dfrac{\partial\eta}{\partial t} = \dfrac{1}{2}\Delta\eta + \nabla\cdot\dot{W} \\ \eta_0 = V^0 \end{cases}.$$

(ii) If $\lambda_n \to \infty$, $\mu_n \to \infty$ and $\mu_n/\lambda_n \to 0$, then $\dfrac{\eta_t^n(\phi) - \lambda_n\langle\phi\rangle}{\sqrt{\lambda_n\mu_n}}$ converges in $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(R^d)\}$ to a solution of the SPDE

$$\begin{cases} \dfrac{\partial\eta}{\partial t} = \dfrac{1}{2}\Delta\eta + \dot{Z} \\ \eta_0 = 0\ . \end{cases}$$

(iii) If $\lambda_n \to \infty$, $\lambda_n\mu_n \to \infty$ and $\mu_n \to c^2 \geq 0$, then $\dfrac{\eta_t^n(\phi) - \lambda_n\langle\phi\rangle}{\sqrt{\lambda_n}}$ converges in $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(R^d)\}$ to a solution of the SPDE

$$\begin{cases} \dfrac{\partial\eta}{\partial t} = \dfrac{1}{2}\Delta\eta + c\dot{Z} + \nabla\cdot\dot{W} \\ \eta_0 = V_0. \end{cases}$$

Theorem 8.9 covers the interesting limits in which $\lambda \to \infty$ and $\mu/\lambda \to 0$. These are all Gaussian. The remaining limits are in general non Gaussian. those in which μ and λ both tend to finite limits are trivial enough to pass over here, which leaves us two cases

(iv) $\lambda \to \infty$ and $\mu/\lambda \to c^2 > 0$;

(v) $\mu \to \infty$ and $\mu/\lambda \to \infty$.

The limits in case (v) turn out to be zero, as we will show below. Thus the only non-trivial, non-Gaussian limit is case (iv), which leads to measure-valued processes.

A MEASURE DIFFUSION

<u>THEOREM 8.10</u> Suppose $\lambda_n \to \infty$ and $\mu_n/\lambda_n \to c^2 > 0$. Then $\frac{1}{\lambda_n}\eta_t^n$ converges weakly in $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(R^d)\}$ to a process $\{\eta_t, t \in [0,1]\}$ which is continuous and has measure-values.

There are a number of proofs of this theorem in the literature (see the Notes), but all those we know of use specific properties of branching processes which we don't want to develop here, so we refer the reader to the references for the proof, and limit ourselves to some formal remarks.

We can get some idea of the behavior of the limiting process by rewriting

(8.13) in the form

$$(8.14)\qquad \frac{1}{\lambda}\,\eta_t(\phi) = \langle\phi\rangle + cU_t(\phi) + \frac{1}{\sqrt{\lambda}}\,(V(G_t(\phi,\cdot)) + \frac{1}{\lambda}\,R_t(\phi)).$$

If (λ_n,μ_n) is any sequence satisfying (iv), $\{(V^n,W^n,Z^n,R^n,U^n,\frac{1}{\lambda_n}\,\eta^n)\}$ is tight by Theorem 8.6 and Proposition 7.8, hence we may choose a subsequence along which it converges weakly to a limit (V, W, Z, R, U, η). From (8.14)

$$(8.15)\qquad \eta_t(\phi) = \langle\phi\rangle + cU_t(\phi) = \langle\phi\rangle + c\int_0^t\int_{R^d} G_{t-s}(\phi,y)Z(dy,ds).$$

In SPDE form this is

$$(8.16)\qquad \begin{cases} \dfrac{\partial\eta}{\partial t} = c\Delta\eta + \dot{Z} \\ \eta_0(dx) = dx \end{cases}$$

We can see several things from this. For one thing, η^n is positive, hence so is η. Consequently, η_t, being a positive distribution, is a measure. It must be non-Gaussian - Gaussian processes aren't positive - so Z itself must be non-Gaussian. In particular, it is not a white noise.

Now η_0 is Lebesgue measure, but if d > 1, Dawson and Hochberg have shown that η_t is purely singular with respect to Lebesgue measure for t > 0. If d = 1, Roelly-Coppoletta has shown that η_t is absolutely continuous.

To get some idea of what the orthogonal martingale measure Z is like, note from Proposition 8.1 that

$$\langle Z^n(A)\rangle = \int_0^t \frac{1}{\lambda_n}\,\eta_s^n(A)ds,$$

which suggests that in the limit

$$\langle Z(A)\rangle_t = \int_0^t \eta_s(A)ds,$$

or, in terms of the measure ν of Corollary 2.8,

$$(8.17)\qquad \nu(dx,ds) = \eta_s(dx)ds.$$

This indicates why the SPDE (8.16) is not very useful for studying η: the statistics of Z are simply too closely connected with those of η, for Z vanishes wherever η does, and η vanishes on large sets - in fact on a set of full Lebesgue measure if $d \geq 2$. In fact, it seems easier to study η, which is a continuous state branching process, than Z, so (8.16) effectively expresses η in terms of a process which is even less understood. This contrasts with cases (i)-(iii), in which Z and W were white noises, processes which we understand rather well.

Nevertheless, there is a heuristic transformation of (8.16) into an SPDE involving a white noise which is worthwhile giving. This has been used by Dawson to give some intuitive understanding of η, but which has never, to our knowledge, been made rigorous. And which, we hasten to add, will certainly not be made rigorous here.

Let W be a real-valued white noise on $R^d \times R_+$. Then (8.17) indicates that Z has the same mean and covariance as Z', where

$$Z'_t(A) = \int_0^t \int_{R^d} \sqrt{\eta_s(y)}\, W(dy\, ds).$$

(If $d = 1$, $\eta_s(dy) = \eta_s(y)dy$, so $\sqrt{\eta_s(y)}$ makes sense. If $d > 1$, η_s is a singular measure, so it is hard to see what $\sqrt{\eta_s}$ means, but let's not worry about it.)

In derivative form, $\dot{Z}' = \sqrt{\eta_s}\,\dot{W}$, which makes it tempting to rewrite the SPDE (8.16) as

(8.18) $$\frac{\partial \eta}{\partial t} = c\Delta\eta + \sqrt{\eta}\,\dot{W}\ .$$

It is not clear that this equation has any meaning if $d \geq 2$, and even if $d = 1$, it is not clear what its connection is with the process η which is the weak limit of the infinite particle system, so it remains one of the curiosities of the subject.

THE CASE $\mu/\lambda \to \infty$

REMARKS. One of the features of Theorem 8.9 is that it allows us to see which of the three sources - initial measure, diffusion, or branching - drives the limit process. In case (i), the branching is negligeable and the noise comes from the initial distribution and the diffusion. In case (ii) the initial distribution washes out completely, the diffusion becomes deterministic and only contributes to the drift term $\frac{1}{2}\Delta\,\eta$, while the noise comes entirely from the branching. In case (iii), all three effects contribute to the noise term. In case (iv), the measure-valued diffusion, we see from (8.16) that the initial distribution and diffusion both become deterministic, while the randomness comes entirely from the branching.

In case (v), which we will analyze now, it turns out that all the sources wash out. Notice that Theorem 8.6 doesn't apply when $\mu/\lambda \to \infty$, and in fact we can't affirm that the family is tight. Nevertheless, η tends to zero in a rather strong way. In fact the unnormalized process tends to zero.

<u>THEOREM 8.11</u>. Let $\lambda \to \infty$ and $\mu/\lambda \to \infty$. Then for any compact set $K \subset R^d$ and $\varepsilon > 0$

(i) $P_{\lambda,\mu}\{\eta_t(K) = 0, \text{ all } t \in [\varepsilon, 1/\varepsilon]\} \to 1$

and, if $d = 1$,

(ii) $P_{\lambda,\mu}\{\eta_t(K) = 0, \text{ all } t \geq \varepsilon\} \to 1$.

Before proving this we need to look at first-hitting times for branching Brownian motions. This discussion is complicated by the profusion of particles: many of them may hit a given set. To which belongs the honor of first entry?

The type of first hitting time we have in mind uses the implicit partial ordering of the branching process - its paths form a tree, after all - and those familiar with two parameter martingales might be interested to compare these with stopping lines.

Suppose that $\{X^\alpha, \alpha \in \underline{\underline{A}}\}$ is the family of processes we constructed at the beginning of the chapter, and let $A \subset \mathbf{R}^d$ be a Borel set. For each α, let $\tau_A^\alpha = \inf\{t > 0: X_t^\alpha \in A\}$, and define T_A^α by

$$T_A^\alpha = \begin{cases} \tau_A^\alpha & \text{if } \tau_A^\beta = \infty \text{ for all } \beta \prec \alpha,\ \beta \neq \alpha; \\ \infty & \text{otherwise} \end{cases} .$$

The time T_A^α is our analogue of a first hitting time. Notice that T_A^α may be finite for many different α, but if $\alpha \prec \beta$, T_A^α and T_A^β can't both be finite. Consider, for example, the first entrance T_E of the British citizenery to an earldom. If an individual - call him α - is created the first Earl of Emsworth, some of his descendants may inherit the title, but his elevation is the vital one, so only T_E^α is finite. On the other hand, a first cousin - call him β - may be created the first Earl of Ickenham; then T_E^β will also be finite.

In general, if $\alpha \neq \beta$ and if T_A^α and T_A^β are both finite, then the descendants of X^α and of X^β form disjoint families. (Why?) By the strong Markov property and the independence of the different particles, the post-T_A^α and post-T_A^β processes are conditionally independent given $X^\alpha(T_A^\alpha)$ and $X^\beta(T_A^\beta)$.

Let P_μ^x be the distribution of the branching Brownian motion with branching rate μ which starts with a single particle X^1 at x. Under P_0^x, then, X_t^1 is an ordinary (non-branching) Brownian motion.

The following result is a fancy version of (8.4). While it is true for the same reason (symmetry), it is more complicated and rates a detailed proof.

PROPOSITION 8.12. Let $\phi(x,t)$ be a bounded Borel function $R^n \times \overline{R}_+$, with $\phi(x,\infty) = 0$, x R^d. For any Borel set $A \subset R^d$

$$E_\mu^x\{ \sum_\alpha \phi(X^\alpha(T_A^\alpha), T_A^\alpha)\} = E_0^x\{\phi(X^1(\tau_A^1),\tau_A^1)\} .$$

PROOF. By standard capacity arguments it is enough to prove this for the case where A is compact and ϕ has compact support in $R^d \times [0,\infty)$. We will drop the subscript and write T^α and τ^α instead of T_A^α and τ_A^α.

Define $u(x,t) = E_0^x\{\phi(X^1_{\tau^1}, t + \tau^1)\}$. Note that $\{u(X^1_{t \wedge \tau^1}, t \wedge \tau^1), t \geq 0\}$ is a martingale, so that we can conclude that $u \in C^{(2)}$ on the open set $A^c \times R_+$ and $\frac{\partial}{\partial t} u + \frac{1}{2}\Delta u = 0$. Thus by Ito's formula

$$\begin{aligned} Y_t &\overset{\text{def}}{=} \sum_\alpha u(X^\alpha_{t \wedge T^\alpha}, t \wedge T^\alpha) \\ &= u(x,0) + \sum_\alpha \int_0^t h^\alpha(s)\, I_{\{s<T^\alpha\}} \nabla u(X_s^\alpha,s) \cdot dB_s^\alpha \\ &\quad + \sum_\alpha u(X^\alpha_{\zeta(\alpha)-}, \zeta(\alpha)) I_{\{\zeta(\alpha) \leq t\ T^\alpha\}} (N^\alpha - 1), \end{aligned}$$

where h^α, B^α, $\zeta(\alpha)$, and N^α are the quantitites used to define the branching process. Note that $Y_t = \eta_t(u(\cdot,t))$ as long as $t < \inf_\alpha T^\alpha$. Since u is bounded, Y has all moments. We claim it is a martingale.

Certainly the stochastic integrals are martingales, hence so is their sum. To see that the second sum is also a martingale, put

$$v_t^\alpha = u(X^\alpha(\zeta(\alpha) -), \zeta(\alpha))(N^\alpha - 1)I_{\{\zeta(\alpha) \leq t \wedge T^\alpha\}}.$$

Then $v_t^\alpha = 0$ if $t < \zeta(\alpha)$, it is constant on $[\zeta(\alpha), \infty)$, and it is identically zero on the set $\{T^\alpha < \zeta(\alpha)\}$. Thus if $s < t$, $E\{v_t^\alpha | \underline{\underline{F}}_t\}$ vanishes on $\{t < \zeta(\alpha)\}$, and equals $v_t^\alpha = v_s^\alpha$ on $\{s \geq \zeta(\alpha), T^\alpha \geq \zeta(\alpha)\}$. One can use the fact that $N^\alpha - 1$ has mean zero and is independent of (X^α, ζ^α) to see that the expectation also vanishes on the set $\{s < \zeta(\alpha), t \geq \zeta(\alpha), T^\alpha \geq \zeta(\alpha)\}$. Thus in all cases $E\{v_t^\alpha | \underline{\underline{F}}_s\} = v_s^\alpha$, proving the claim.

If x is a regular point of A, i.e. if $P_0^x\{\tau^1 = 0\} = 1$, then $u(x,t) = \phi(x,t)$ for all $t \geq 0$. A Brownian motion hitting A must do so at a regular point, so $u(X^\alpha_{T^\alpha}, T^\alpha) = \phi(X^\alpha_{T^\alpha}, T^\alpha)$. (This even holds if $T^\alpha = \infty$, since both sides vanish then.) Thus

$$u(x,0) = E_\mu^x\{\lim Y_t\} = E_\mu^x\{\sum_\alpha \phi(X^\alpha_{T^\alpha}, T^\alpha)\}.$$

Q.E.D.

REMARKS. This implies that the hitting probabilities of the branching Brownian motion are dominated by those of Brownian motion - just take $\phi \equiv 1$ and note that the left hand side of (8.14) dominates $E_\mu^x\{\sup_\alpha \phi(X_{T^\alpha}, T^\alpha)\} = P_\mu^x\{T^\alpha < \infty, \text{ some } \alpha\}$. It also implies that the left hand side of (8.14) is independent of μ.

We need several results before we can prove Theorem 8.11. Let us first treat the case $d = 1$. Let D be the unit interval in $\mathbf{R}^1$ and put

$$H_\mu(x) = P_\mu^x\{T_D^\alpha < \infty, \text{ some } \alpha\}.$$

PROPOSITION 8.14. $H_\mu(x) = \frac{6}{\mu}(x - 1 - \sqrt{6/\mu})^{-2}$ if $x \geq 1$.

PROOF. This will follow once we show that H_μ is the unique solution of

$$\text{(8.19)} \qquad \begin{cases} u'' = \mu u^2 & \text{on } (1,\infty) \\ u(1) = 1 & \\ 0 \leq u \leq 1 & \text{on } (1,\infty), \end{cases}$$

since it is easily verified that the given expression satisfies (8.19).

Let $T = \inf_\alpha T^\alpha$. If $x > 1$, Proposition 8.12 implies

(8.20) $$P_\mu^x\{T < h\} = P_0^x\{T < h\} = o(h) \quad \text{as } h \to 0.$$

Let ζ be the first branching time of the process. Then

$$H_\mu(x) = P_\mu^x\{T \le \zeta \wedge h\} + P_\mu^x\{\zeta > h,\ \zeta \wedge h < T < \infty\}$$
$$+ P_\mu^x\{\zeta \le h,\ \zeta \wedge h < T < \infty\}.$$

The first probability is o(h) by (8.20). Apply the strong Markov property at $\zeta \wedge h$ to the latter two. If $\zeta > h$, there is still only one particle, X^1, alive, so $T = T^1$ and the probability equals $E_\mu^x\{\zeta > h,\ H_\mu(X^1_{\zeta\wedge h})\} + o(h)$, where the o(h) comes from ignoring the possibility that $T < \zeta \wedge h$. X^1 either dies or splits into two independent particles, X^{11} and X^{12}, at ζ, so if $\zeta \le h$, $I_{\{T<\infty\}} = I_{\{T^{11}<\infty\}} + I_{\{T^{12}<\infty\}} - I_{\{T_1^{11}<\infty, T^{12}<\infty\}}$. Since T^{11} and T^{12} are independent given $\underline{\underline{F}}_\zeta$, we see the second term is

$$\frac{1}{2} E_\mu^x\{\zeta \le h,\ 2H_\mu(X^1_{\zeta\wedge h-})\} - \frac{1}{2} E_\mu^x\{\zeta \le h:\ H_\mu^2(X^1_{\zeta\wedge h-})\} + o(h),$$

where we have used the fact that $X_\zeta^{11} = X_\zeta^{12} = X^1_{\zeta-}$.

Add these terms together to see that

$$E_\mu^x\{H_\mu(X^1_{\zeta\wedge h-})\} - H_\mu(x) = \frac{1}{2} E_\mu^x\{\zeta \le h;\ H_\mu^2(X^1_{\zeta\wedge h-})\} + o(h).$$

X^1 is Brownian motion up to ζ, which is exponential (μ), so we can calculate this. Divide by $E_\mu^x\{\zeta \wedge h\} > (1 - e^{-\mu h})/\mu$, let $h \to 0$, and use Dynkin's formula. The left had side tends to $H_\mu''(x)/2$ while, since $H_\mu^2(X^1_{\zeta\wedge h-}) \to H_\mu^2(x)$, the right hand side tends to $\mu\, H_\mu^2(x)/2$. Thus H_μ satisfies (8.19).

To see that the solution of (8.19) is unique, suppose that u_1 and u_2 are both solutions and that $u_1'(1) < u_2'(1)$. Let $v = u_2 - u_1$. Then $v'' = \mu(u_1 + u_2)v$, which is strictly positive on $\{x: v(x) > 0\}$. Thus v' is increasing on $\{v > 0\}$, while $v(1) = 0$ and $v'(1) > 0$. This implies that v' is increasing on $[1,\infty)$, hence $v(x) \to \infty$ as $x \to \infty$, contradicting the fact that $0 \le u_1, u_2 \le 1$. Thus $u_1'(1) = u_2'(1)$, $u_1(1) = u_2(1)$, and the usual uniqueness result for the initial value problem implies that $u_1 \equiv u_2$. Q.E.D.

Moving to the d-dimensional case, let D_r be the ball of radius r centered at 0, let $T = \inf_\alpha T^\alpha_{D_1}$, and put

$$f_\mu(x,t) = P_\mu^x\{T \le t\}.$$

Let $Q(x,r,ds) = P_0^x\{\tau_{D_r}^1 \quad ds\}$ be the distribution of τ_{D_r} for ordinary Brownian motion.

LEMMA 8.15. Let $r > 1$ and let $x, y \in \mathbf{R}^d$ be such that $|x| > r$ and $|y| = r$. Then

$$f_\mu(x,t) \le \int_0^t Q(x,r,ds) f_\mu(y,t-s).$$

PROOF. In order for a particle to reach D_1 from x either it or one of its ancestors must first reach D_r. Now

$$f_\mu(x,t) = 1 - P_\mu^x\{T_{D_1}^\alpha > t, \text{ all } \alpha\}$$

and

$$P_\mu^x\{T_{D_1}^\alpha > t \text{ all } \alpha\} = P_\mu^x\{\text{for all } \alpha\text{: } T_{D_1}^\beta > t \text{ for all } \beta \succ \alpha \text{ if } T_{D_r}^\alpha \le t\}.$$

Let us apply the strong Markov property at $T_{D_r}^\alpha$. Since $|y| = r$ and $f_\mu(y,t)$ is symmetric in y, the conditional probability that the particle - or some descendant - reaches D_1 before t given that $T_{D_r}^\alpha < t$ is $f_\mu(y,t-T_{D_r}^\alpha)$. Since the different post-T^α processes are independent, the above probability equals

$$E_\mu^x\{\prod_\alpha(1 - f_\mu(y,\ t-T_{D_r}^\alpha))\},$$

where $f_\mu(y,\ t-T_{D_r}^\alpha) = 0$ if $T_{D_r}^\alpha > t$. Thus

$$\begin{aligned} f_\mu(x,t) &= 1 - E_\mu^x\{\prod_\alpha (1 - f_\mu(y,\ t-T_{D_r}^\alpha))\} \\ &\le 1 - E_\mu^x\{1 - \sum_\alpha f_\mu(y,\ t-T_{D_r}^\alpha)\} \\ &= E_\mu^x\{\sum_\alpha f_\mu(y,\ t-T_{D_r}^\alpha)\} \\ &= E_0^x\{f_\mu(y,\ \tau_{D_r})\} \end{aligned}$$

by Proposition 8.12. Q.E.D.

This brings us to Theorem 8.11.

PROOF of Theorem 8.11. Suppose without loss of generality that K is the unit ball D_1. Write $\eta_t^\lambda = \bar{\eta}_t^\lambda + \bar{\bar{\eta}}_t^\lambda$ where $\bar{\eta}_t^\lambda$ comes from those initial particles inside D_2, and

$\overline{\overline{\eta}}_t^{\lambda} = \eta_t^{\lambda} - \overline{\eta}_t^{\lambda}$.

Now if we start a single particle from x, $\eta_t^{\lambda}(1)$ is an ordinary branching process hence (see e.g. Harris [28]).

$$P^x\{\eta_t(1) > 0\} \sim \frac{c}{\mu t} \quad \text{as } \mu \to \infty \quad \text{for } t > 0.$$

Thus

$$P_\mu\{\overline{\eta}_t^{\lambda} = 0 \quad \text{all } t > \varepsilon\} = P_\mu\{\overline{\eta}_t^{\lambda}(1) = 0\}$$
$$= \int_{D_2} \frac{c}{\mu\varepsilon} \lambda dx \sim \frac{c}{\varepsilon}\frac{\lambda}{\mu} \to 0.$$

Next

$$P_\mu\{\overline{\overline{\eta}}_t^{\lambda}(D_1) > 0 \quad \text{some } t < 1/\varepsilon\}$$
$$= \int_{R^d - D_2} P_\mu^x\{T_{D_1}^{\alpha} < \tfrac{1}{\varepsilon}\}\lambda\, dx$$
$$\leq \int_{R^d - D_2} \int_0^t Q(x, \tfrac{3}{2}, ds) f_\mu(y, t-s)\lambda ds$$

by Lemma 8.15.

Now in order for a particle to hit D_1, its first coordinate must hit $[-1,1]$, so that $f_\mu(y, t-s) \leq H_\mu(|y|)$. Thus this is

$$\leq \lambda\, H_\mu(\tfrac{3}{2}) \int_{R^d - D_2} Q(x, \tfrac{3}{2}, [0,t])dx\ .$$

This integral is finite - indeed, it is bounded by

$$P_0^x\{\sup_{0 \leq s \leq t} |X_s^1 - x| \leq |x| - \tfrac{3}{2}\} \leq C\, e^{-\frac{(|x|-3/2)^2}{2t}}$$, so by Proposition 8.14, this is

$$= C\, \frac{\lambda}{\mu} \left(\frac{1}{2} + \sqrt{6/\mu}\right)^{-2} \to 0.$$

Putting these together gives (i). In case d = 1,

$$P\{\overline{\overline{\eta}}_t^{\lambda}(D_1) > 0 \text{ some } t \geq 0\} = 2 \int_1^\infty \frac{6}{\mu}\left(x - 1 + \sqrt{\frac{6}{\mu}}\right)^{-2} \lambda\, dx$$
$$\leq 12\, \frac{\lambda}{\mu} \to 0$$

giving (ii).

Q.E.D.

THE SQUARE OF THE BROWNIAN DENSITY PROCESS

Now that we have seen how the Brownian density process can be squeezed out of an infinite particle system, we can't resist the temptation to look at its square. We renormalize during its construction, so that it is actually the square by analogy rather than by algebra, but it is at least closely related.

We return to the system of discrete particles to set up the process and then take a weak limit at the end. We will show that the limiting process satisfies a stochastic partial differential equation whose solution can be written in terms of multiple Wiener integrals. In particular, it is non-Gaussian. The end result is in Theorem 8.18.

Let $\{X^\alpha, \alpha \in \mathbf{N}\}$ be a family of i.i.d. standard (i.e. non-branching) Brownian motions in $\mathbf{R}^d$, with initial distribution given by a Poisson point process Π^λ of parameter λ. Set, as before,

$$\eta_t(\phi) = \sum_\alpha \phi(X_t^\alpha).$$

Now $\eta_t^2(\phi) = \sum_{\alpha,\beta} \phi(X_t^\alpha)\phi(X_t^\beta)$. We will first symmetrize this, then throw away the terms with $\alpha = \beta$, to get a new process, Q_t.

Let $\{\xi^\alpha, \alpha \in \mathbf{N}\}$ be a sequence of i.i.d. random variables, independent of the X^α, such that $P\{\xi^\alpha = 1\} = P\{\xi^\alpha = -1\} = 1/2$. Define

$$\tilde{\eta}_t(\phi) = \sum_{\alpha \in \mathbf{N}} \xi^\alpha \, \phi(X_t^\alpha) \tag{8.21}$$

and, for any function ψ on $\mathbf{R}^d \times \mathbf{R}^d$, set

$$Q_t(\psi) = \lambda^{-1} \sum_{\alpha \neq \beta} \xi^\alpha \xi^\beta \psi(X_t^\alpha, X_t^\beta). \tag{8.22}$$

This is the process of interest. We define it on $\mathbf{R}^{2d}$ rather than $\mathbf{R}^d$,; to see its connection with $\tilde{\eta}^2$, set $\psi(x,y) = \phi(x)\phi(y)$. Then $Q_t(\psi) = \lambda^{-1}(\tilde{\eta}_t^2(\phi) - \tilde{\eta}_t(\phi^2))$.

<u>Notation</u>: Let $D \subset \mathbf{R}^{2d}$ be the set $\{(x,y)\colon x \in \mathbf{R}^d, y \in \mathbf{R}^d, x = y\}$. If μ is a measure on $\mathbf{R}^d$, define a measure $\mu \otimes \mu$ on $\mathbf{R}^{2d}$ by $\mu \otimes \mu(A) = \mu \times \mu(A-D)$, where $A \subset \mathbf{R}^{2d}$, and $\mu \times \mu$ is the product measure on $\mathbf{R}^{2d}$. If we let $\tilde{\Pi}^\lambda = \tilde{\eta}_0$, which is the symmetrized version of Π^λ, then $Q_0 = \tilde{\Pi}^\lambda \otimes \tilde{\Pi}^\lambda$ and $Q_t = \tilde{\eta}_t \otimes \tilde{\eta}_t$.

If we try to write Q in differential form, we would expect that $dQ = \tilde{\eta}_t \times$

$\tilde{\eta}_t + d\tilde{\eta}_t \otimes \tilde{\eta}_t$. This is roughly what happens. To see exactly what happens, tho, we must analyze the system from scratch.

Define $W_t(A) = \lambda^{-1/2} \sum_\alpha \int_0^t I_A(X_s^\alpha)dX_s^\alpha$ as before, and let its symmetrized version be

$$\tilde{W}_t(A) = \lambda^{-1/2} \sum_\alpha \xi^\alpha \int_0^t I_A(X_s^\alpha)dX_s^\alpha.$$

Since there is no branching, (8.10) becomes

$$\eta_t(\phi) = \int_{\mathbf{R}^d} G_t(\phi,y)\eta_0(dy) + \lambda^{1/2} \int_0^t \int_{\mathbf{R}^d} \nabla G_{t-s}(\phi,y)\cdot W(dy\ ds).$$

Its symmetrized version is

$$\text{(8.23)} \qquad \tilde{\eta}_t(\phi) = \int_{\mathbf{R}^d} G_t(\phi,y)\, \tilde{\eta}_0(dy) + \lambda^{1/2} \int_0^t \int_{\mathbf{R}^d} \nabla G_{t-s}(\phi,y)\cdot \tilde{W}(dyds).$$

Let us use ∇_1 and ∇_2 to indicate the gradients $\nabla_1\phi(x,y) = \nabla_x\phi(x,y)$ and $\nabla_2\phi(x,y) = \nabla_y\phi(x,y)$. Similarly we define the Laplacians $\Delta_1 = \Delta_x$ and $\Delta_2 = \Delta_y$, so that the Laplacian on $\mathbf{R}^{2d}$ is $\Delta = \Delta_1 + \Delta_2$.

If $\phi = C^2(\mathbf{R}^{2d})$ has compact support, then by Ito's formula

$$\text{(8.24)} \quad Q_t(\phi) = Q_0(\phi) + \lambda^{-1} \sum_{\alpha\neq\beta} \xi^\alpha\xi^\beta \Big[\int_0^t \nabla_1\phi(X_s^\alpha,X_s^\beta)\cdot dX_s^\alpha + \int_0^t \nabla_2\phi(X_s^\alpha,X_s^\beta)\cdot dX_s^\beta$$

$$+ (2\lambda)^{-1} \sum_{\alpha\neq\beta} \xi^\alpha\xi^\beta \int_0^t \Delta\phi(X_s^\alpha,X_s^\beta)ds.$$

Each of these sums can be identified in terms of η, Q and the martingale measures W and $\tilde{W}$. The last term, for instance, is just $(2\lambda)^{-1}\int_0^t Q_s(\Delta\phi)ds$, while

$$\begin{aligned}
&\sum_{\alpha\neq\beta} \xi^\alpha\xi^\beta \int_0^t \nabla_1\phi(X_s^\alpha,X_s^\beta)\cdot dX_s^\alpha \\
&= \sum_\beta \xi^\beta \int_0^t \xi^\alpha\nabla_1\phi(X_s^\alpha,X_s^\beta)\cdot dX_s^\alpha - \sum_\alpha \int_0^t \nabla_1\phi(X_s^\alpha,X_s^\alpha)\cdot dX_s^\alpha \\
&= \lambda^{1/2} \sum_\beta \xi^\beta \int_0^t \int_{\mathbf{R}^d} \nabla_1\phi(x,X_s^\beta)\cdot d\tilde{W}_{sx} - \lambda^{1/2} \sum_\alpha \int_0^t \int_{\mathbf{R}^d} \nabla_1\phi(x,x)\cdot dW_{sx} \\
&= \lambda^{1/2} \int_0^t \int_{\mathbf{R}^d} \tilde{\eta}_s(\nabla_1\phi(x,\cdot))\cdot d\tilde{W}_{sx} - \lambda^{1/2} \int_0^t \int_{\mathbf{R}^d} \nabla_1\phi(x,x)\cdot dW_{sx}.
\end{aligned}$$

Let $\chi(x) = (\nabla_1\phi)(x,x) + (\nabla_2\phi)(x,x)$. Then (8.24) becomes

$$Q_t(\phi) = Q_0(\phi) + \frac{1}{2}\int_0^t Q_s(\Delta\phi)ds + \lambda^{-1/2} \int_{\mathbf{R}^d\times[0,t]} \tilde{\eta}_s(\nabla_1\phi(x,\cdot))d\tilde{W}_{sx}$$

$$+ \lambda^{-1/2} \int_{\mathbf{R}^d\times[0,t]} \tilde{\eta}_s(\nabla_2\phi(\cdot,y))\cdot d\tilde{W}_{sy} - \lambda^{-1/2}\int_{\mathbf{R}^d\times[0,t]} \chi(x)\cdot dW_{sx}.$$

Let us write $Q = \tilde{Q} + R$ where

$$(8.25)\qquad \tilde{Q}_t(\phi) = Q_0(\phi) + \frac{1}{2}\int_0^t \tilde{Q}_s(\Delta\phi)ds + \lambda^{-1/2}\int_{\mathbf{R}^d\times[0,t]} \tilde{\eta}_s(\nabla_1\phi(x,\cdot)\cdot d\tilde{W}_{sx}$$

$$+ \lambda^{-1/2}\int_{\mathbf{R}^d\times[0,t]} \tilde{\eta}_s(\nabla_2\phi(\cdot,y))\cdot d\tilde{W}_{sy};$$

$$(8.26)\qquad R_t(\phi) = \frac{1}{2}\int_0^t R_s(\Delta\phi)ds + \lambda^{-1/2}\int_{\mathbf{R}^d\times[0,t]} \chi(x)\cdot dW_{sx}.$$

These are integral forms of SPDE's which we can solve by Theorem 5.1. To get them in the form (5.4), define a pair of martingale measures on $\mathbf{R}^{2d}$ by

$$M_t^1(\phi) = \int_{\mathbf{R}^d\times[0,t]}\Big(\int_{\mathbf{R}^d} \phi(x,y)\tilde{\eta}_s(dx)\Big)d\tilde{W}_{sy}$$

and

$$M_t^2(\phi) = \int_{\mathbf{R}^d\times[0,t]}\Big(\int_{\mathbf{R}^d} \phi(x,y)\tilde{\eta}_s(dy)\Big)d\tilde{W}_{sx}.$$

Note that these are worthy martingale measures. The covariance measure for M^1, for instance, is $\tilde{\eta}_s(dx)\tilde{\eta}_s(dx')\delta_y(y')dydy'dsI$, hence its dominating measure is

$$K^1(dx\ dy\ dx'dy'ds) = \eta_s(dx)\eta_s(dx')\delta_y(y')dydy'ds,$$

which is clearly positive definite.

Notice also that the M^i are neither orthogonal nor of nuclear covariance. M^1 and M^2 have values in $\mathbf{R}^d$ (since $\tilde{W}$ does) and, if we abuse notation by writing $\int \nabla_i\phi(x,y)\cdot M_t^i(dx\ dy) = M_t^i(\nabla_i\phi)$, (8.25) becomes

$$(8.27)\qquad \tilde{Q}_t(\phi) = Q_0(\phi) + \frac{1}{2}\int_0^t \tilde{Q}_s(\Delta\phi)ds + \lambda^{-1/2}M^1(\nabla_1\phi) + \lambda^{-1/2}M^2(\nabla_2\phi).$$

By Theorem 5.1

$$\tilde{Q}_t(\phi) = Q_0(G_t\phi) + \int_{\mathbf{R}^d\times[0,t]} \nabla_1 G_{t-s}(\phi)(x,y,s)\cdot M^2(dx\ dy\ ds)$$

$$+ \int_{\mathbf{R}^d\times[0,t]} \nabla_2 G_{t-s}(\phi)(x,y,s)\cdot M^1(dx\ dy\ ds).$$

If we write $M^1(\phi)$ as $\int \eta_s(\phi(\cdot,y))d\tilde{W}^1_{sy}$ this becomes

$$(8.28)\qquad \tilde{Q}_t(\phi) = Q_0(G_t\phi) + \int_{\mathbf{R}^d\times[0,t]} \tilde{\eta}_s(\nabla_1 G_{t-s}\phi(x,\cdot))\cdot d\tilde{W}_{sx}$$

$$+ \int_{\mathbf{R}^d\times[0,t]} \tilde{\eta}_s(\nabla_2 G_{t-s}\phi(\cdot,y))\cdot d\tilde{W}_{sy}.$$

Next, define $N_t(\phi) = \int_0^t \phi(x,x)d\tilde{W}_{sx}$. N is an orthogonal R^d-valued martingale measure on R^{2d}, and (8.26) becomes

$$R_t(\phi) = \frac{1}{2}\int_0^t R_s(\Delta\phi)ds + \lambda^{-1/2}N_t(\nabla_1\phi+\nabla_2\phi).$$

Let $P_t(x,y) = (2\pi t)^{-d/2}e^{-\frac{|y-x|^2}{2t}}$ and let $G_t(x,y;x',y') = P_t(x,y)P_t(x',y')$. Then G is the Green's function on R^{2d} for this problem, so that, by Theorem 5.1 it is

$$= \lambda^{-1/2} \int_{R^d\times[0,t]} (\nabla_1 G_{t-s}\phi(x,y) + \nabla_2 G_{t-s}\phi(x,y)\cdot N(dx\ dy\ ds)$$

or

(8.29) $$R_t(\phi) = \lambda^{-1} \int_{R^d\times[0,t]} \left[\nabla_1 G_{t-s}(\phi)(y,y) + \nabla_2(G_{t-s}(\phi)(y,y)\right]\cdot W(dyds)$$

If we let $\lambda \to \infty$, we will see that $\tilde{Q}$ and R have weak limits, and in fact $R \Rightarrow 0$. Most of the work has already been done. Let us make the dependence on λ explicit, writing W^λ, η^λ, Q^λ etc. Note that W^λ and $\tilde{W}^\lambda$ have the same distribution, and thus both have the same means and covariances as a white noise, independent of λ. They are orthogonal, hence W^λ and $\tilde{W}^\lambda$ must converge weakly to independent white noises by Theorem 8.7. We know about the moments of η from (8.4), (8.12), Proposition 8.5 and Exercise 8.1. In the case of $\tilde{\eta}$,

(8.30) $$E\{\tilde{\eta}_t(\phi)\} = 0, \qquad E\{\tilde{\eta}_t^2(\phi)\} = \lambda\langle\phi^2\rangle,$$

the latter following since the left hand side is

$$E\{\sum_{\alpha,\beta} \xi^\alpha\xi^\beta\phi(X_t^\alpha)\phi(X_t^\beta)\} = E\{\sum_\alpha \phi^2(X_t^\alpha)\}.$$

Once can show as in Exercise 8.1 that $\lambda^{-1/2}\tilde{\eta}^\lambda$ is L^p bounded, independent of λ, for all $p < \infty$.

In order to establish that Q converges weakly, we need to show that $R \Rightarrow 0$ and that the various terms of (8.28) converge weakly. Let us dispatch the easy parts of the convergence argument first.

<u>PROPOSITION 8.16</u>. The processes $\lambda^{-1/2}\tilde{\Pi}^\lambda$, $\lambda^{-1}\tilde{\Pi}^\lambda\otimes\tilde{\Pi}^\lambda$, $\tilde{W}^\lambda$, $\lambda^{-1}\tilde{\eta}^\lambda$, R^λ and Q^λ are tight on the appropriate space $\underline{\underline{D}}\{([0,1],\underline{S}'(R_p)\}$. Moreover, if V_0 and W_0 are independent

white noises on R^d and $R^d \times R_+$ respectively, with values in R and R^d respectively, then, as $\lambda \to \infty$

$$\left(\lambda^{-1/2}\tilde{\Pi}^\lambda,\ \lambda^{-1}\tilde{\Pi}^\lambda \otimes \tilde{\Pi}^\lambda,\ \tilde{W}^\lambda,\ \lambda^{-1/2}\eta^\lambda,\ R^\lambda\right) \Rightarrow (V^0,\ V^0 \otimes V^0,\ W^0,\ \eta, 0),$$

where η is a solution of the SPDE

$$\begin{cases} \frac{\partial \eta}{\partial t} = \frac{1}{2}\Delta\eta + \nabla \cdot W^0 \\ \eta_0 = V^0 . \end{cases}$$

<u>Note</u>. $V^0 \otimes V^0$ is a multiple Wiener integral. We can define $V^0 \otimes V^0(A) = V^0(A_1)V^0(A_2)$ if $A = A_1 \times A_2$ and $A_1 \cap A_2 = \phi$, and extend it to all Borel A by the usual approximation arguments.

<u>PROOF</u>. As remarked above $\tilde{W}^\lambda \Rightarrow W^0$, and we have proved $\lambda^{-1/2}\Pi^\lambda$ converges weakly. We leave it as an exercise to show that $\lambda^{-1/2}\tilde{\Pi}^\lambda \Rightarrow V^0$. If ϕ and ψ are test functions of disjoint support in R^d,

$$\left(\lambda^{-1/2}\Pi^\lambda(\phi),\ \lambda^{-1/2}\Pi^\lambda(\psi)\right) \Rightarrow (V^0(\phi),\ V^0(\psi)).$$

Multiplication is continuous on R^2, so this implies that

$$\lambda^{-1}\ \tilde{\Pi}^\lambda \otimes \tilde{\Pi}^\lambda(\phi,\psi) = \lambda^{-1}\tilde{\Pi}^\lambda(\phi)\tilde{\Pi}^\lambda(\psi) \Rightarrow V^0(\phi)V^0(\psi) = V^0 \otimes V^0(\phi\psi).$$

This holds for finite sums of such functions, hence for all $\psi(x,y) \in \underline{\underline{S}}(R^{2d})$ by approximation.

In view of (8.23), the convergence of $\lambda^{-1/2}\ \tilde{\eta}^\lambda$ to η follow from Proposition 7.8, and the limit is identified as in Theorem 8.9(i).

The tightness of R^λ follows from (8.29) and Proposition 7.8, and, from (8.29)

$$E\{R_t^\lambda(\psi)^2\} \le \frac{c}{\lambda} \to 0,$$

hence $R^\lambda \Rightarrow 0$.

We leave it to the reader to use Proposition 7.6 to show joint convergence of these processes. Q.E.D.

This leaves us the question of the weak convergence of Q or, equivalently, of $\tilde{Q}$. In view of (8.28), it is enough to show the convergence of the stochastic integrals there.

Let $U_t^\lambda = \lambda^{-1/2} \int_{R^d \times [0,t]} \tilde{\eta}_s^\lambda (\nabla_1 G_{t-s}\psi(x,\cdot)) \cdot d\tilde{W}_{xs}^\lambda .$

If $\psi(x,y) = \phi(x)\chi(y)$, then

$$U_t^\lambda = \int_{R^d\times[0,t]} \lambda^{-1/2}\, \tilde{\eta}_s^\lambda(\chi)\, \nabla_1 G_{t-s}\phi(x)\cdot d\tilde{W}_{xs}^\lambda .$$

Let (λ_n) be a sequence tending to infinity.

PROPOSITION 8.17. $\left(\tilde{W}^{\lambda_n}, \lambda^{-1/2}\tilde{\eta}^{\lambda_n}, U^{\lambda_n}\right) \Rightarrow \left(W^0, \eta, U^0\right)$

where

$$U_t^0(\psi) = \int_{R^d\times[0,t]} \eta_s(\nabla_1 G_{t-s}\psi(x,\cdot))\cdot dW_{xs}^0,\ \psi \in \underline{\underline{S}}(R^{2d}).$$

PROOF. We already know from Proposition 8.16 that $\left(\tilde{W}^{\lambda_n}, \lambda^{-1/2}\eta^{\lambda_n}\right) \Rightarrow \left(W^0, \eta\right)$. It remains to treat U. The idea of the proof is the same: the martingale measure converges, hence so does the stochastic integral. However, we can't use Propositions 7.6 and 7.8, for the integrand is not in $\bar{\mathcal{P}}_s$. We will use 7.12 and 7.13 instead.

Define (V^0, W^0) and $(\tilde{V}^{\lambda_n}, \tilde{W}^{\lambda_n})$ canonically on $\underline{\underline{D}} = \underline{\underline{D}}([0,1], \underline{\underline{S}}'(R^{2d}))$, and denote their probability distributions by P^0 and P respectively. By (8.23) we can define $\tilde{\eta}_t^{\lambda_n}$ on $\underline{\underline{D}}$ simultaneously for n = 0,1,2,... as a continuous process. That is, the stochastic integrals in (8.23) are consistent up to sets which are of P^n-measure zero for all $n \geq 0$. Thus we can also define

$$g(x,s,t) = \int_{R^d\times[0,t]} \eta_s(\phi)\nabla P_{t-s}\chi(x)\cdot d\tilde{W}_{xs}^{\lambda_n}$$

for each s,t and x, independent of n. You can do better.

Exercise 8.2. Show $g(\cdot,\cdot,t_0)I_{\{s\leq t_0\}} \in \bar{\bar{\mathcal{P}}}_S(W)$, where W is the sequence $W^0, W^{\lambda_1}, W^{\lambda_2},\dots$,

Hint. It is not quite enough to approximate $\eta_s(\phi)$ by the step function $\eta_{\frac{[ms]}{m}}(\phi)$ where [t] is the greatest integer in t, since it is not clear that this will be a continuous function of ω on $\underline{\underline{D}}$. Go back one step further and approximate $\eta_{\frac{[ms]}{m}}(\phi)$

itself by the integral of a deterministic step function with respect to dW_{xs}. This will be continuous in ω.

Now apply Theorem 7.13 with $p = q = 4$ and $\beta = 1$. Note that g is of Hölder class (1,4,C) for some C. Indeed (i) of the definition was verified in the above exercise; if we take $Y_n(\omega) = \sup_{t \leq 1}\left[\|\nabla P_t \chi\|_\infty + \|P_t \chi\|_\infty\right]$ then parts (a) and (b) of (ii) hold, while (c) follows from the uniform L^p-boundedness of $\lambda^{-1/2}\eta^\lambda$.

The family (W^{λ_n}) is of class (4,K) (see Chapter 7) for some K, and β, p, and q satisfy the necessary inequalities, so that Theorem 7.13 (v) implies that $U^n_t(\phi) \Rightarrow U_t(\phi)$ for ϕ of the form $\chi(x)\phi(y)$. It follows that there is also convergence for finite sums of such functions. Since any $\phi \in \underline{\underline{S}}(\mathbf{R}^{2d})$ can be approximated uniformly and in L^p by such sums, it is easy to see that $U^n(\phi) \Rightarrow U(\phi)$ for all $\phi \in \underline{\underline{S}}(\mathbf{R}^{2d})$. We leave the details to the reader. It now follows from Mitoma's theorem that $U^{\lambda_n} \Rightarrow U^0$ on $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(\mathbf{R}^{2d})\}$.

To show that the triple converges jointly, let $f^n(x,t,\omega)$ be the 3×2 matrix

$$\begin{pmatrix} \phi_1, & P_{t-s}\phi_2, & \eta_s(\phi_3)\nabla P_{t-s}\chi \\ 0 & P_t\phi_2 & 0 \end{pmatrix}^T$$

and let $M^n = (\tilde{W}^{\lambda_n}, V^{\lambda_n})^T$, in which case $(f^n \cdot M^n)_t = \left(\tilde{W}_t^{\lambda_n}(\phi_1), \eta_t^{\lambda_n}(\phi_2), U_t^{\lambda_n}(\phi_3\chi)\right)^T$, and $f^n \cdot M^n \Rightarrow f^0 \cdot M^0$ by Proposition 7.12, which implies joint convergence. Q.E.D

<u>REMARK</u>. We took $p = q = 4$ in the above proof, but as W and η are L^p bounded for all p, we could let p and $q \to \infty$ with $p = q$. In that case, Theorem 7.13(iii) tells us that U is L^p bounded for all p, and Hölder continuous with exponent $\frac{1}{2} - \varepsilon$ for any $\varepsilon > 0$.

The second integral in (8.28) also converges, so, combining Propostions 8.16 and 8.17, we have

<u>THEOREM 8.18</u>. The process Q^λ converges weakly in $\underline{\underline{D}}\{[0,1], \underline{\underline{S}}'(\mathbf{R}^{2d})\}$ to a solution of the SPDE

$$(8.31)\qquad \begin{cases} \dfrac{\partial Q}{\partial t}(x,y) = \dfrac{1}{2}\Delta Q(x,y) + \eta(x)\nabla_2 \cdot W^0_{ys} + \eta(y)\nabla_1 \cdot W^0_{xs} \\ Q_0 = V^0 \otimes V^0 \end{cases}$$

<u>PROOF</u>. There is very little to prove. (8.31) is just the differential form of (8.25) (with $\tilde{W}$ replaced by W^0 and $\tilde{\eta}$ by η) and (8.28) is the unique solution of (8.25). By Propositions 8.16 and 8.17, all the stochastic integrals converge to the right limits, so equation (8.28) is valid for the limiting process, hence so is (8.25).

Q.E.D.

If we plug (8.23) - which remains valid if we replace $\tilde{\eta}_0$ by V^0 and $\tilde{W}$ by W^0 - into (8.25) we get Q in all its glory. For simplicity's sake we will only write it for ψ of the form $\psi(x,y) = \chi(x)\,\phi(y)$.

$$\begin{aligned} Q_t(\psi) &= \int_{R^{2d}-D} P_{t-s}\chi(x)P_{t-s}\phi(y)V^0(dx)V^0(dy) \\ &+ \int_{R^d\times[0,t]} \Big[\int_{R^d} P_s\chi(x)\, V^0(dx) + \int_{R^d\times[0,t]} \nabla P_{s-u}\chi(x)\cdot W^0(dxdu)\Big]\nabla\phi(y)\cdot W^0(dyds) \\ &+ \int_{R^d\times[0.t]} \Big[\int_{R^d} P_s\phi(x)V^0(dx) + \int_{R^d\times[0.t]} \nabla P_{s-u}\phi(x)\cdot W^0(dxdu)\Big]\nabla\chi(y)\cdot W^0(dyds). \end{aligned}$$

The first integral is a classical multiple Wiener integral. The next two could also be called multiple Wiener integrals, as they are iterated stochastic integrals with respect to white noise.

CHAPTER NINE

We have spent most of our time on parabolic equations; non-parabolic equations have made only token appearances, such as at the beginning of these notes when we took a brief glance at the wave equation, which is hyperbolic. It is fitting to end with a brief glance at a token elliptic, Laplace's equation.

We will give one existence and uniqueness theorem for bounded regions, and then see how such equations arise as the limits of parabolic equations. In particular, we will look at the limits of the Brownian density process as $t \to \infty$.

Let D be a bounded domain in $\mathbf{R}^d$ with a smooth boundary. Consider

$$(9.1) \qquad \begin{cases} \Delta u = f & \text{in } D \\ u = 0 & \text{on } \partial D \,. \end{cases}$$

If f is bounded and continuous, the solution to (9.1) is

$$(9.2) \qquad u(y) = \int K(x,y)\, f(x)dx = K(f,y),$$

where K is the Green's function for (9.1). Notice that in particular, $\Delta K(f,y) = f(y)$.

Let M be an L^2-valued measure on $\mathbf{R}^d$ (not a martingale measure, for there is no t in the problem!) Set $Q(A,B) = E\{M(A)M(B)\}$ and suppose that there exists a positive definite measure $\tilde{Q}$ on $\mathbf{R}^d \times \mathbf{R}^d$ such that $|Q(A,B)| \leq \tilde{Q}(A\times B)$ for all Borel A, $B \subset \mathbf{R}^d$. This assures us of a good integration theory. We also assume for convenience that $M(\partial D) = 0$.

Let T be a kth order differential operator on $\mathbf{R}^d$ with smooth coefficients ($0 \leq k < \infty$) and consider the SPDE

$$(9.3) \qquad \begin{cases} \Delta U = T\dot{M} & \text{in } D \\ U = 0 & \text{in } \partial D \,. \end{cases}$$

Let us get the weak form of (9.3). Multiply by a test function ϕ and integrate over $\mathbf{R}^d$, pretending $\dot{M}$ is smooth. Suppose $\phi = 0$ on ∂D. We can then do two integrations by parts to get

$$(9.4) \qquad \int U(x)\Delta\phi(x)dx = \int_D T\,\dot{M}(x)\phi(x)dx.$$

Let T^* be the formal adjoint of T. If T is a zeroth or first order operator, or if ϕ has compact support in D, we can integrate by parts on the right to get

$$U(\Delta\phi) = \int_D T^*\phi(x)\, M(dx). \tag{9.5}$$

Let H_n be the Sobolev space introduced in Example 1, Chapter 4. We say U is an $\underline{H_n \text{ solution}}$ of (9.3) if for a.e.ω, $U(\cdot,\omega)$ takes values in H_{-n} and (9.5) holds for all $\phi \in H_n$. We say U is a <u>weak solution</u> if $U \in \underline{\underline{S}}'(R^d)$ a.e. and if (9.5) holds for all $\phi \in \underline{\underline{S}}_0$, where

$$\underline{\underline{S}}_0 = \{\phi \in \underline{\underline{S}}(R^d) : \phi = 0 \text{ on } \partial D\}.$$

<u>PROPOSITION 9.1</u>. Let k = order of T. If $n > d + k$, then (9.3) has a unique H_n-solution, defined by

$$U(\phi) = \int T^*K(\phi,y)M(dy). \tag{9.6}$$

This also defines a weak solution of (9.3).

<u>PROOF</u>. Uniqueness. By the general theory of PDE, $\phi \in H_n \Rightarrow K(\phi,\cdot) \in H_{n+2} \subset H_n$. If U is any H_n-solution, apply (9.5) to $\psi(y) = K(\phi,y)$:

$$U(\phi) = U(\Delta\psi) = \int_D T^*\psi(y)M(dy) = \int_D T^*K(\phi,y)M(dy).$$

Existence. Define U by (9.6). If $\phi \in C^\infty(R^d)$,

$$\begin{aligned} E\{|U(\phi)|^2\} &= E\{[\int_D T^*K(\phi,y)M(dy)]^2\} \\ &= \int_D \int_D T^* K(\phi,x)T^*K(\phi,y)Q(dx\, dy) \\ &\le C_1 \|T^*K(\phi,\cdot)\|^2_{L^\infty} \end{aligned}$$

where $C_1 = \tilde{Q}\{D \times D\}$. By the Sobolev embedding theorem, if $q > d/2$ this is

$$\le C_2 \|T^*K(\phi,\cdot)\|^2_q .$$

T is a differential operator of order k, hence it is bounded from $H_{q+k} \to H_q$, while K maps $H_{q+k-2} \to H_{q+k}$ boundedly. Thus the above is

$$\le C_4 \|\phi\|^2_{q+k-2}$$

It follows that U is continuous in probability on H_{q+k-2} and, by Theorem 4.1, it is a random linear functional on H_p for any $p > q + k - 2 + d/2$. Fix a $p > d + k - 2$ and let $n = p + 2$. Then $U \in H_{-n}$. (It is much easier to see that $U \in \underline{\underline{S}}'(\mathbf{R}^d)$. Just note that $T^*K(\phi,\cdot)$ is bounded if $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$ and apply Corollary 4.2).

If $\phi \in \underline{\underline{S}}_0$,

$$U(\Delta\phi) = \int_D T^*K(\Delta\phi,y)M(dy)$$

$$= \int_D T^*\phi(y)M(dy).$$

On the other hand, $C_0^\infty \subset \underline{\underline{S}}_0$, and C_0^∞ is dense in all the H_t. U is continuous on H_p so the map $\phi \to \Delta\phi \to U(\Delta\phi)$ of $H_n \to H_p \to \mathbf{R}$ is continuous, while on the right-hand side of (9.5)

$$E\{|\int_{\mathbf{R}^d} T^*\phi dM|^2\} \le C\|T^*\phi\|^2_{L^\infty} \le C\|\phi\|^2_{k+q},$$

which tells us the right-hand side is continuous in probability on H_{k+q}, hence, by Theorem 4.1, it is a linear functional on $H_{d+k} = H_n$. Thus (9.5) holds for $\phi \in H_n$. Q.E.D.

LIMITS OF THE BROWNIAN DENSITY PROCESS

The Brownian density process η_t satisfies the equation

(9.7) $$\frac{\partial\eta}{\partial t} = \frac{1}{2}\Delta\eta + a\,\nabla\cdot\dot{W} + b\dot{Z}$$

where W is a d-dimensional white noise and Z is an independent one-dimensional white noise, both on $\mathbf{R}^d \times \mathbf{R}_+$, and the coefficients a and b are constants. (They depend on the limiting behavior of μ and λ.)

Let us ask if the process has a weak limit as $t \to \infty$. It is not too hard to see that the process blows up in dimensions $d = 1$ and 2, so suppose $d \ge 3$. The Green's function G_t for the heat equation on $\mathbf{R}^d$ is related to the Green's function K for Laplace's equation by

(9.8) $$K(x,y) = -\int_0^\infty G_t(x,y)dt$$

and K itself is given by

$$K(x,y) = \frac{C_d}{|y-x|^{d-2}},$$

where C_d is a constant. The solution of (9.7) is

$$\eta_t(\phi) = \eta_0 G_t(\phi) + a \int_0^t \int_{R^d} \nabla G_{t-s}(\phi,y)\cdot W(dy\, ds) + b \int_0^t \int_{R^d} G_{t-s}(\phi,y) Z(dyds)$$

$$\stackrel{\text{def}}{=} \eta_0 G_t(\phi) + a\, R_t(\phi) + b\, U_t(\phi).$$

R_t and U_t are mean-zero Gaussian processes. The covariance of R_t is

$$E\{R_t(\phi)R_t(\psi)\} = \int_0^t \int_{R^d} (\nabla_x G_{t-s})(\phi,y)\cdot(\nabla_x G_{t-s})(\psi,y)dy\, ds\ .$$

Now $\nabla_x G = -\nabla_y G$; if we then integrate by parts

$$= - \int_0^t \int_{R^d} \Delta_y G_{t-s}(\phi,y) G_{t-s}(\psi,y) dy\, ds$$

$$= - \int_0^t \int_{R^d} G_{t-s}(\Delta\phi,y) G_{t-s}(y,\psi) dy\, ds$$

$$= - \int_0^t G_{2t-2s}(\Delta\phi,\psi) ds$$

$$= - \frac{1}{2} \int_{R^d} \psi(y) \int_0^{2t} G_u(\Delta\phi,y) ds\, dy$$

$$= - \frac{1}{2} \int_{R^d} \psi(y)[-\phi(x) + G_{2t}(x,\phi)]dy$$

by (5.7). Since $d \geq 3$, $G_t \to 0$ as $t \to \infty$ so

(9.9) $$E\{R_t(\phi)R_t(\psi)\} \to \frac{1}{2} \langle\phi,\psi\rangle.$$

The calculation for U is easier since we don't need to integrate by parts:

$$E\{U_t(\phi)U_t(\psi)\} = \int_0^t \int_{R^d} G_{t-s}(\phi,y) G_{t-s}(\psi,y) dy\, ds$$

$$= \frac{1}{2} \int_0^{2t} G_{2t-u}(\phi,\psi) du \to -\frac{1}{2} K(\phi,\psi)$$

as $t \to \infty$. Taking this and (9.9) into account, we see:

<u>PROPOSITION 9.3</u>. Suppose $d \geq 3$. As $t \to \infty$, $\sqrt{2}\, R_t$ converges weakly to a white noise and $\sqrt{2}\, U_t$ converges to a random Gaussian tempered distribution with covariance function

(9.10) $$E\{U(\phi)U(\psi)\} = -K(\phi,\psi).$$

In particular, η_t converges weakly as $t \to \infty$. The convergence is weak convergence of $\underline{\underline{S}}'(R^d)$-valued random variables in all cases.

Exercise 9.1. Fill in the details of the convergence argument.

DEFINITION. The mean zero Gaussian process $\{U(\phi): \phi \in \underline{\underline{S}}(\mathbf{R}^d)\}$ with covariance (9.10) is called the Euclidean free field.

CONNECTION WITH SPDE's

We can get an integral representation of the free field U from Proposition 9.3, for the weak limit of $\sqrt{2}\, U_t$ has the same distribution as

$$\int_0^\infty \int_{\mathbf{R}^d} G_s(\phi,y)Z(dy\, ds).$$

This is not enlightening; we would prefer a representation independent of time. This is not hard to find. Let $\dot{W}$ be a d-dimensional white noise on $\mathbf{R}^d$ (not on $\mathbf{R}^d \times \mathbf{R}_+$ as before) and, for $\phi \in \underline{\underline{S}}(\mathbf{R}^d)$, define

$$U(\phi) = \int_{\mathbf{R}^d} \nabla K(\phi,y)\cdot W(dy). \tag{9.11}$$

If $\phi, \psi \in \underline{\underline{S}}(\mathbf{R}^d)$,

$$\begin{aligned} E\{U(\phi)U(\psi)\} &= \int_{\mathbf{R}^d} \nabla K(\phi,y)\cdot\nabla K(\psi,y)dy \\ &= - \int_{\mathbf{R}^d} K(\phi,y)\Delta K(\psi,y)dy \\ &= - \int_{\mathbf{R}^d} K(\phi,y)\psi(y)dy \\ &= - K(\phi,\psi). \end{aligned}$$

(This shows a posteriori that $U(\phi)$ is defined!) Thus, as $U(\phi)$ is a mean zero Gaussian process, it is a free field.

PROPOSITION 9.4. U satisfies the SPDE

$$\Delta U = \nabla\cdot\dot{W} \tag{9.12}$$

PROOF. $U(\Delta\phi) = \int_{\mathbf{R}} \nabla K(\Delta\phi,y)\cdot W(dy)$

$= \int_{\mathbf{R}} \nabla\phi(y)\cdot W(dy)$

since for $\phi \in S(\mathbf{R}^d)$, $K(\Delta\phi,y) = \phi(y)$. But this is the weak form of (9.12).

Q.E.D.

Exercise 9.1. Convince yourself that for a.e.ω, (9.12) is an equation in distributions.

SMOOTHNESS

Since we are working on $\mathbf{R}^d$, we can use the Fourier transform. Let H_t be the Sobolev space defined in Example 1a, Chapter 4. If u is any distribution, we say $u \in H_t^{loc}$ if for any $\phi \in C_0^\infty$, $\phi u \in H_t$.

PROPOSITION 9.5. Let $\varepsilon > 0$. Then with probability one, $\dot{W} \in H^{loc}_{-d/2-\varepsilon}$ and $U \in H^{loc}_{1-d/2-\varepsilon}$, where U is the free field.

PROOF. The Fourier transform of ϕW is a function:

$$\widehat{\phi W}(\xi) = \int_{\mathbf{R}} e^{-2\pi i\xi\cdot x}\phi(x)W(dx)$$

and

$$E\{|(1 + |\xi|^2)^{t/2}\widehat{\phi W}(\xi)|^2\} = (1 + |\xi|^2)^t \iint \phi(x)\overline{\phi}(y)e^{2\pi i(y-x)\cdot\xi}dydx,$$

so

$$\begin{aligned} E\{\|\phi W\|_t^2\} &= \int(1 + |\xi|^2)^t|\hat{\phi}(\xi)|^2 d\xi \\ &\leq C \int (1 + |\xi|^2)^t d\xi \end{aligned}$$

which is finite if $2t < -d$, in which case $\|\phi W\|_t$ is evidently finite a.s.

Now $\nabla\cdot W \in H^{loc}_{-d/2-1-\varepsilon}$ so, since U satisfies $\Delta U = \nabla\cdot\dot{W}$, the elliptic regularity theorem of PDE's tells us $U \in H^{loc}_{1-\varepsilon-d/2}$. Q.E.D.

THE MARKOV PROPERTY OF THE FREE FIELD

We discussed Lévy's Markov and sharp Markov properties in Chapter One, in connection with the Brownian sheet. They make sense for general distribution-valued processes, but one must first define the σ-fields $\underline{\underline{G}}_D$ and $\underline{\underline{G}}^*_D$. This involves extending the distribution.

Since U takes values on a Sobolev space, the Sobolev embedding theorem tells us that it has a trace on certain lower-dimensional manifolds. But since we want to talk about its values on rather irregular sets, we will use a more direct method.

If μ is a measure on R^d, let us define U(μ) by (9.11). This certainly works if μ is of the form $\mu(dx) = \phi(x)dx$, and it will continue to work if μ is suffciently nice. By the calculation following (9.11), "sufficiently nice" means that

$$\|\mu\|_K^2 = -\int\int \mu(dx)K(x,y)\mu(dy) < \infty. \tag{9.13}$$

Let $\underline{\underline{E}}_+$ be the class of measures on R^d which satisfy (9.13), and let $\underline{\underline{E}} = \underline{\underline{E}}_+ - \underline{\underline{E}}_+$.

If $B \subset R^d$ is Borel, one version of the restriction of U to B would be

$$\{U(\mu):\ \mu \in \underline{\underline{E}},\ \ U(A) > 0,\ \text{all } A \subset R^d - B\}.$$

Of course, this requires that there be measures in $\underline{\underline{E}}$ which sit on B. This will always be true if B has positive capacity, for if B is bounded and has positive capacity, its equilibrium measure has a bounded potential and is thus in E. (For the potential of μ is K(μ,•) and $\|\mu\|_K^2 = \int K(\mu,y)\mu(dy)$.

Thus, define

$$\underline{\underline{G}}_B = \sigma\{U(\mu):\ \mu \in \underline{\underline{E}},\ \mu(A) = 0,\ \text{all } A \subset R^d - B\}$$

$$\underline{\underline{G}}^*_B = \bigcap_{\substack{A \supset B \\ A \text{ open}}} \underline{\underline{G}}_A\ .$$

<u>PROPOSITION 9.6</u>. The free field U satisfies Lévy's sharp Markov property relative to bounded open sets in R^d.

<u>PROOF</u>. This follows easily from the balayage property of K: if $D \subset R^d$ is an open set and if μ is supported by D^c, there exists a measure ν on ∂D such $-K(\nu,y) \leq -K(\mu,y)$ for all y, and $K(\nu,y) = K(\mu,y)$ for all $y \in D$, and all but a set of capacity zero in ∂D. We call ν the <u>balayage of μ on ∂D</u>.

Suppose $\mu \in \underline{\underline{E}}$ and supp $\mu \subset D^c$. If ν is the balayage of μ on ∂D, we claim that

$$E\{U(\mu) | \underline{\underline{G}}_{\bar{D}}\} = U(\nu). \tag{9.14}$$

This will do it since, as $U(\nu)$ is $\underline{\underline{G}}_{\partial D}$-measurable, the left-hand side of (9.14) must be $E\{U(\mu) | \underline{\underline{G}}_{\partial D}\}$.

Note that $\nu \in \underline{\underline{E}}$ (for μ is and $-K(\nu,\cdot) \leq -K(\mu,\cdot)$) so if $\lambda \in \underline{\underline{E}}$, $\mathrm{supp}(\lambda) \subset \bar{D}$,

$$\begin{aligned} E\{(U(\mu) - U(\nu))U(\lambda)\} &= \int [K(\mu,y) - K(\nu,y)]\lambda(dy) \\ &= 0 \end{aligned}$$

since $K(\mu,x) = K(\nu,x)$ on $\bar{D}$, except possibly for a set of capacity zero, and λ, being of finite energy, does not charge sets of capacity zero. Thus the integrand vanishes λ-a.e. But we are dealing with Gaussian processes, so this implies (9.14). Q.E.D.

NOTES

We omitted most references from the body of the text - a consequence of putting off the bibliography till last - and we will try to remedy that here. Our references will be rather sketchy - you may put that down to a lack of scholarship - and we list the sources from which we personally have learned things, which may not be the sources in which they originally appeared. We apologize in advance to the many whose work we have slighted in this way.

CHAPTER ONE

The Brownian sheet was introduced by Kitagawa in [37], though it is usually credited to others, perhaps because he failed to prove the underlying measure was countably additive. This omission looks less serious now than it did then.

The Garsia-Rodemich-Rumsey Theorem (Theorem 1.1) was proved for one-parameter processes in [23], and was proved in general in the brief and elegant article [22], which is the source of this proof. This commonly gives the right order of magnitude for the modulus of continuity of a process, but doesn't necessarily give the best constant, as, for example, in Proposition 1.4. The exact modulus of continuity there, as well as many other interesting sample-path properties of the Brownian sheet, may be found in Orey and Pruitt [49].

Kolmogorov's Theorem is usually stated more simply than in Corollary 1.2. In particular, the extra log terms there are a bit of an affectation. We just were curious to see how far one can go with non-Gaussian processes. Our version is only valid for real-valued processes, but the theorem holds for metric-space valued processes. See for example [44, p.519].

The Markov property of the Brownian sheet was proved by L. Pitt [52]. The splitting field is identified in [59]; the proof there is due to S. Orey (private communication.)

The propagation of singularities in the Brownian sheet is studied in detail in [56]. Orey and Taylor showed the existence of singular points of the Brownian path and determined their Hausdorff dimension in [50]. Proposition 1.7 is due to G. Zimmerman [63], with a quite different proof.

The connection of the vibrating string and the Brownian sheet is due to E. Cabaña [8], who worked it out in the case of a finite string, which is harder than the infinite string we treat. He also discusses the energy of the string.

CHAPTER TWO

In terms of the mathematical techniques involved, one can split up much of the study of SPDE's into two parts: that in which the underlying noise has nuclear covariance, and that in which it is a white noise. The former leads naturally to Hilbert space methods; these don't suffice to handle white noise, which leads to some fairly exotic functional analysis. This chapter is an attempt to combine the two in a (nearly) real variable setting. The integral constructed here may be technically new, but all the important cases can also be handled by previous integrals.

(We should explain that we did not have time or space in these notes to cover SPDE's driven by martingale measures with nuclear covariance, so that we never take advantage of the integral's full generality).

Integration with respect to orthogonal martingale measures, which include white noise, goes back at least to Gihman and Skorohod [25]. (They assumed as part of their definition that the measures are worthy, but this assumption is unnecessary; c.f. Corollary 2.9.)

Integrals with respect to martingale measures having nuclear covariance have been well-studied, though not in those terms. An excellent account can be found in Métivier and Pellaumeil [46]. They handle the case of "cylindrical processes", (which include white noise) separately.

The measure ν of Corollary 2.8 is a Doléans measure at heart, although we haven't put it in the usual form. True Doléans measures for such processes have been constructed by Huang, [31].

Proposition 2.10 is due to J. Watkins [61]. Bakry's example can be found in [2].

CHAPTER THREE

The linear wave and cable equations driven by white and colored noise have been treated numerous times. Dawson [13] gives an account of these and similar equations.

The existence and uniqueness of the solution of (3.5) were established by Dawson [14]. The L^p-boundednes and Hölder continuity of the paths are new. See [57] for a detailed account of the sample path behavior in the linear case and for more on the barrier problem.

The wave equation has been treated in the literature of two-parameter processes, going back to R. Cairoli's 1972 article [9]. The setting there is special because of the nature of the domain: on these domains, only the initial position need be specified, not the velocity.

As indicated in Exercises 3.4 and 3.5, one can extend Theorem 3.2 and Corollary 3.4, with virtually the same proof, to the equation

$$\frac{\partial V}{\partial t} = \frac{\partial^2 V}{\partial x^2} + g(V,t) + f(V,t)\dot{W},$$

where both f and g satisfy Lipschitz conditions. Such equations can model physical systems in which g is potential term. Faris and Jona-Lasinio [19] have used similar equations to model the "tunnelling" of a system from one stable state to another.

We chose reflecting boundary conditions in (3.5) and (3.5b) for convenience. They can be replaced by general linear homogeneous boundary conditions; the important point is that the Green's funciton satisfies (3.6) and (3.7), which hold in general [27].

CHAPTER FOUR

We follow some unpublished lecture notes of the Ito here. See also [24] and [34].

CHAPTER FIVE

The techniques used to solve (5.1) also work when L is a higher order elliptic operator. In fact the Green's function for higher order operators has a lower order pole, so that the solutions are better behaved than in the second-order case.

We suspect that Theorem 5.1 goes back to the mists of antiquity. Ito studies a special case in [33]. Theorem 5.4 and other results on the sample paths of the solution can be found in [58]. See Da Prato [12] for another point of view on these and similar theorems.

CHAPTER SIX

The basic reference on weak convergence remains Billingsley's book [5]. Aldous' theorem is in [1], and Kurtz' criterion is in [42]. We follow Kurtz' treatment here. Mitoma's theorem is proved in [47], but the article is not self-contained. Fouque [21] has generalized this to a larger class of spaces of distributions, which includes the familiar spaces $D(\Omega)$. His proof is close to that of Mitoma.

CHAPTER SEVEN

It may not be obvious from the exposition - in fact we took care to hide it - but the first part of the chapter is designed to handle deterministic integrands. The accounts for its relatively elementary character.

Theorems general enough to handle the random integrands met in practice seem to be delicate; we were surprised to find out how little is known, even in the classical case. Our work in the section "an extension" is just a first attempt in that direction.

Peter Kotelenez showed us the proof of Proposition 7.8. Theorem 7.10 is due to Kallianpur and Wolpert [36]. An earlier, clumsier version can be found in [57]. The Burkholder-Davis-Gundy theorem is summarized in its most highly developed form in [7].

CHAPTER EIGHT

This chapter completes a cycle of results on weak limits of Poisson systems of branching Brownian motion due to a number of authors. "Completes" is perhaps too strong a word, for these point in many directions and we have only followed one: to find all possible weak limits of a certain class of infinite particle systems, and to connect them with SPDE's.

These systems were investigated by Martin-Löf [45] who considered non-branching particles ($\mu = 0$ in our terminology) and by Holley and Stroock [29], who considered branching Brownian motions in $\mathbf{R}^d$ with parameters $\lambda = \mu = 1$; their results look superficially different since, instead of letting μ and λ tend to infinity, they rescale the process in both space and time by replacing x by x/α and t by $\alpha^2 t$. Because of the Brownian scaling, this has the same effect as replacing λ by α^d and μ by α^2, and leaving x and t unscaled. The critical parameter is then $\mu/\lambda = \alpha^{2-d}$, so their results depend on the dimension d of the space. If $d \geq 3$, they find a Gaussian limit (case (ii) of Theorem 8.9), if d = 2 they have the measure-valued diffusion (case (iv)) and if d = 1, the process tends to zero (Theorem 8.11). The case $\mu = 0$, investigated by Martin-Löf and, with some differences, by Ito [33], [34], also leads to a Gaussian limit (Theorem 8.9 (i)).

Gorostitza [26] treated the case where μ is fixed and $\lambda \to \infty$ (Theorem 8.9(iii) if $\mu > 0$). He also gets a decomposition of the noise into two parts, but it is different from ours; he has pointed out [26, Correction] that the two parts are not in fact independent.

The non-Gaussian case (case (iv)) is extremely interesting and has been investigated by numerous authors. S. Watanabe [60] proved the convergence of the system to a measure-valued diffusion. Different proofs have been given by Dawson [13], Kurtz [42], and Roelly-Copoletta [53]. Dawson and Hochberg [15] have looked at the Hausdorff dimension of the support of the measure and showed it is singular with respect to Lebesgue measure if $d \geq 2$. It is absolutely continuous if $d = 1$ (Roelly-Copoletta [53]). A related equation which can be written suggestively as

$$\frac{\partial \eta}{\partial t} = \frac{1}{2} \Delta \eta + \eta(1-\eta)\dot{W}$$

has been studied by Fleming and Viot [20].

The case in which $\mu/\lambda \to \infty$ comes up in Holley and Stroock's paper if $d = 1$. The results presented here, which are stronger, are joint work with E. Perkins and J. Watkins, and appear here with their permission. The noise W of Proposition 8.1 is due to E. Perkins who used it to translate Ito's work into the setting of SPDE's relative to martingale measures (private communication.)

A more general and more sophisticated construction of branching diffusions can be found in Ikeda, Nagasawa, and Watanabe [32]. Holley and Stroock also give a construction.

The square process Q is connected with U statistics. Dynkin and Mandelbaum [17] showed that certain central limit theorems involving U statistics lead to multiple Wiener integrals, and we wish to thank Dynkin for suggesting that our methods might handle the case when the particles were diffusing in time. In fact Theorem 8.18 might be viewed as a central limit theorem for certain U-statistics evolving in time.

We should say a word about generalizations here. We have treated only the simplest settings for the sake of clarity, but there is surprisingly little change if we move to more complex systems. We can replace the Brownian particles by branching diffusions, or even branching Hunt processes, for instance, without changing the character of the limiting process. (Roelly-Copoletta [Thesis, U. of Paris, 1984]). One can treat more general branching schemes. If the family size N has a finite variance, Gorostitza [26] has shown that one gets limiting equations of the form $\frac{\partial \eta}{\partial t} = \frac{1}{2}\Delta\eta + \alpha\eta + \beta Z + \gamma\nabla\cdot W$, where $\alpha = 0$ if $E\{N - 1\} = 0$, so that the only new effect is to add a growth term, $\alpha\eta$.

If $E\{N^2\} = \infty$, however, things do change. For example, η can tend to zero in certain cases when μ/λ has a finite limit. This needs further study.

CHAPTER NINE

The term "random field" is a portmanteau word. At one time or another, it has been used to cover almost any process having more than one parameter - and some one-parameter processes, too. It seems to be used particularly for elliptic systems, though why it should be used more often for elliptic than parabolic or hyperbolic systems is something of a mystery. (As is the term itself, for that matter). At any rate, this chapter is about random fields.

We have used some heavy technical machinery here. Frankly, we were under deadline pressure and didn't have time to work out an easier approach. For Sobolev spaces, see Adams [64]; for the PDE theorems, see Folland [67] and Hormander [30]. The classical potential theory and the energy of measures can be found in Doob [66].

The exponent n of the Sobolev space in Proposition 7.1 can doubtless be improved. If $\dot{M}$ is a white noise, one can bypass the Sobolev embedding in the proof and get $n > k + d/2$ rather than $n > k + d$.

The free field was introduced by Nelson in [48]. He proved the sharp Markov property, and used it to construct the quantum field which describes non-interacting particles. He also showed that it can be modified to describe certain interacting systems.

Rozanov's book [54] is a good reference for Lévy's Markov property. See Evstigneev for a strong Markov property, and Kusuoka [43] for results which also apply to parabolic systems in which, contrary to the claim in [57], one commonly finds that Lévy's Markov property holds but the sharp Markov property does not.

CHAPTER TEN

There is no Chapter Ten in these notes. For some reason that hasn't stopped us from having notes on Chapter Ten. We will use this space to collect some remarks which didn't fit in elsewhere. Since the chapter under discussion doesn't exist, no one can accuse us of digressing.

We did not have a chance to discuss equations relative to martingale measures with a nuclear covariance. These can arise when the underlying noise is smoother than a white noise or, as often happens, it is white noise which one has approximated by a smoothed out version. If one thinks of a white noise, as we did in the introduction, as coming from storm-driven grains of sand bombarding a guitar string, one might think of nuclear covariance noise as coming from a storm of ping-pong balls. The solutions of such systems tend to be better-behaved, and in particular, they often give function solutions rather than distributions. This makes it possible to treat non-linear equations, something rather awkward to do otherwise (how does one take a non-linear function of a distribution?) Mathematically, these equations are usually treated in a Hilbert-space setting. See for instance Curtain and Falb [11], Da Prato [12], and Ichikawa [68].

There have been a variety of approaches devised to cope with SPDE's driven by white noise and related processes. See Kuo [41] and Dawson [13] for a treatment based on the theory of abstract Wiener spaces. The latter paper reviews the subject of SPDE's up to 1975 and has extensive references. Balakrishnan [3] and Kallianpur and Karandikar [35] have used cylindrical Brownian motions and finitely additive measures. See also Métivier and Pellaumail [46], which gives an account of the integration theory of cylindrical processes. Gihman and Skorohod [25] introduced orthogonal martingale measures. See also Watkins [61]. Ustunel [55] has studied nuclear space valued semi-martingales with applications to SPDE's and stochastic flows. The martingale problem method can be adapted to SPDE's as well as ordinary SDE's. It has had succes in handling non-linear equations intractable to other methods. See Dawson [65] and Fleming and Viot [20], and Holley and Stroock [29] for the linear case.

Another type of equation which has generated considerable research is the SPDE driven by a single one-parameter Brownian motion. (One could get such an equation from (5.1) by letting T be an integral operator rather than a differential operator.) An example of this is the Zakai equation which arises in filtering theory. See Pardoux [51] and Krylov and Rosovski [39].

Let us finish by mentioning a few more subjects which might interest the reader: fluid flow and the stochastic Navier-Stokes equation (e.g. Bensoussan and Temam [4]); measure-valued diffusions and their application to population growth (Dawson [65], Fleming and Viot [20]); reaction diffusion equations in chemistry (Kotelenz [38]) and quantum fields (Wolpert [70] and Dynkin [16]).

REFERENCES

[1] Aldous, D., Stopping times and tightness, Ann. Prob. 6 (1978), 335-340.

[2] Bakry, D., Semi martingales à deux indices, Sem. de Prob. XV, Lecture Notes in Math 850, 671-672.

[3] Balakrishnan, A. V., Stochastic bilinear partial differential equations, in Variable Structure Systems, Lecture Notes in Economics and Mathematical Systems 3, Springer Verlag, 1975.

[4] Bensoussan, A. and Temam, R., Equations stochastiques du type Navier-Stokes, J. Fcl. Anal. 13 (1973), 195-222.

[5] Billingsley, P., Convergence of Probability Measures, Wiley, New York, 1968.

[6] Brennan, M. D., Planar semimartingales, J. Mult. Anal. 9 (1979), 465-486.

[7] Burkholder, D. L., Distribution function inequalities for martingales, Ann. Prob. 1 (1973), 19-42.

[8] Cabaña, E., On barrier problems for the vibrating string, ZW 22 (1972), 13-24.

[9] Cairoli, R., Sur une equation differentielle stochastique, C.R. 274 (1972), 1738-1742.

[10] Cairoli, R. and Walsh, J. B., Stochastic integrals in the plane, Acta Math 134 (1975), 111-183.

[11] Curtain, R. F. and Falb, P. L., Stochastic differential equations in Hilbert spaces, J. Diff. Eq. 10 (1971), 434-448.

[12] Da Prato, G., Regularity results of a convolution stochastic integral and applications to parabolic stochastic equations in a Hilbert space (Preprint).

[13] Dawson, D., Stochastic evolution equations and related measure processes, J. Mult. Anal. 5 (1975), 1-52.

[14] Dawson, D., Stochastic evolution equations, Math Biosciences 15, 287-316.

[15] Dawson, D. and Hochberg, K. J., The carrying dimension of a stochastic measure diffusion. Ann. Prob. 7 (1979).

[16] Dynkin, E. B., Gaussian and non-Gaussian random fields associated with Markov processes, J. Fcl. Anal. 55 (1984), 344-376.

[17] Dynkin, E. B. and Mandelbaum A., Symmetric statistics, Poisson point processes, and multiple Wiener integrals, Ann. Math. Stat 11 (1983), 739-745.

[18] Evstigneev, I. V., Markov times for random fields, Theor. Prob. Appl. 22 (1978), 563-569.

[19] Faris, W. G., and Jona-Lasinio, G., Large fluctuations for a nonlinear heat equation with white noise, J. Phys. A: Math, Gen. 15 (1982), 3025-3055.

[20] Fleming, W. and Viot, M., Some measure-valued Markov processes in population genetics theory, Indiana Univ. Journal 28 (1979), 817-843.

[21] Fouque, J-P., La convergence en loi pour les processus à valeurs dans un éspace nucleaire, Ann. IHP 20 (1984), 225-245.

[22] Garsia, A., Continuity properties of Gaussian processes with multidimensional time parameter, Proc. 6th Berkeley Symposium, V.II, 369-374.

[23] Garsia, A., Rodemich, G., and Rumsey, H. Jr., A real variable lemma and the continuity of paths of some Gaussian processes, Indiana U. Math. J. 20 (1970), 565-578.

[24] Gelfand, I. M. and Vilenkin, N. Ya., Generalized Functions, V.4, Academic Press, New York-London 1964.

[25] Gihman, I. I., and Skorohod, A. V., The Theory of Stochastic Processes, III, Springer-Verlag Berlin (1979).

[26] Gorostitza, L., High density limit theorems for infinite systems of unscaled branching Brownian motions, Ann. Prob. 11 (1983), 374-392; Correction, Ann. Prob. 12 (1984), 926-927.

[27] Greiner, P., An asymptotic expansion for the heat equation, Arch. Ratl. Mech. Anal. 41 (1971), 163-218.

[28] Harris, T. E., The Theory of Branching Processes, Prentice-Hall, Englewood Cliffs, N.J., 1963.

[29] Holley, R. and Stoock, D., Generalized Ornstein-Uhlenbeck processes and infinite particle branching Brownian motions, Publ. RIMS Kyoto Univ. 14 (1978), 741-788.

[30] Hormander, L. Linear Partial Differential Operators, Springer Verlag, Berlin, Heidelberg, New York, 1963.

[31] Huang, Zhiyuan, Stochastic integrals on general topological measurable spaces, Z.W. 66 (1984), 25-40.

[32] Ikeda, N., Nagasawa, M., and Watanabe, S., Branching Markov processes, I, II and III, J. Math. Kyoto. Univ. 8 (1968) 233-278, 365-410; 9 (1969), 95-110.

[33] Ito, K., Stochastic anaysis in infinite dimensions; in Stochastic Analysis (A. Friedman and M. Pinsky, eds.) Academic Press, New York, 1980, 187-197.

[34] Ito, K., Distribution-valued processes arising from independent Brownian motions. Math Z. 182 (1983) 17-33.

[35] Kallianpur, G. and Karandikar, R. L., A finitely additive white noise approach to non-linear filtering, J. Appl. Math and Opt. (to appear).

[36] Kallianpur, G. and Wolpert, R., Infinite dimensional stochastic differential equation models for spatially distributed neurons.

[37] Kitagawa, T., Analysis of variance applied to function spaces, Mem. Fac. Sci. Kyusu Univ. 6 (1951), 41-53.

[38] Kotelenz, P., A stopped Doob inequality for stochastic convolution integrals and stochastic evolution equations; (preprint).

[39] Krylov, N. V. and Rosovski, B. L. Stochastic evolution equations, J. Soviet Math. (1981).

[40] Kuo, H. H., Gaussian measures in Banach spaces. Lecture notes in Math. 463, Springer Verlag 1975.

[41] Kuo, H. H., Differential and stochastic equations in abstract Wiener space. J. Fcl. Anal. 12 (1973), 246-256.

[42] Kurtz, T., Approximation of Population Processes CBMS-NSF Reg. Conf. Ser. in Appl. Math. 36, 1981.

[43] Kusuoka, S., Markov fields and local operators, J. Fac. Sci., U of Tokyo 26, (1979), 199-212.

[44] Loève, Probability Theory, Van Nostrand, Princeton, 1963.

[45] Martin-Löf, A., Limit theorems for motion of a Poisson system of independent Markovian particles with high density, Z.W. 34 (1976), 205-223.

[46] Métivier, M. and Pellaumail, J., Stochastic Integration. Academic Press, New York, 1980.

[47] Mitoma, I., Tightness of Probabilities on C([0,1], S') and D([0,1], S'), Ann. Prob. 11 (1983), 989-999.

[48] Nelson, E., The free Markoff field, J. Fcl. Anal 12 (1973), 211-227.

[49] Orey, S. and Pruitt, W., Sample functions of the N-parameter Wiener process, Ann. Prob. 1 (1973), 138-163.

[50] Orey, S. and Taylor, S. J., How often on the Brownian path does the law of iterated logarithm fail? Proc. Lond. Math. Soc. 28 (1974), 174-192.

[51] Pardoux, E., Stochastic PDE and filtering of diffusion processes, Stochastics 3 (1979) 127-167.

[52] Pitt, L. D., A Markov property for Gaussian processes with a multidimensional parameter. Arch. Rat. Mech. Anal 43 (1971), 367-391.

[53] Roelley-Coppoletta, S., Un critère de convergence pour les lois de processus à valeurs mesures. (1984) (preprint).

[54] Rozanov, Yu. A., Markov Random Fields, Springer-Verlag, New York, Heidelberg, Berlin, 1982.

[55] Ustunel, S., Stochastic integration on nuclear spaces and its applications. Ann. Inst. H. Poincaré 28 (1982), 165-200.

[56] Walsh, J. B., Propagation of singularities in the Brownian sheet, Ann. Prob. 10 (1982), 279-288.

[57] Walsh, J. B., A stochastic model of neural response, Adv. Appl. Prob. 13 (1981), 231-281.

[58] Walsh, J. B., Regularity properties of a stochastic partial differential equation, Seminar on Stochastic Processes, 1983, Birkhauser, Boston, (1984), 257-290.

[59] Walsh, J. B., Martingales with a multidimensional parameter and stochastic integrals in the plane. To appear, Springer Lecture Notes in Math.

[60] Watanabe, S., A limit theorem of branching processes and continuous state branching processes, J. Math. Kuoto. Univ. 8 (1968), 141-167.

[61] Watkins, J., A stochastic integral representation for random evolutions. To appear, Ann. Prob.

[62] Zakai, M., On the optimal filtering of diffusion processes, Z.W. 11 (1969), 230-243.

[63] Zimmerman, G. J., Some sample function properties of the two parameter Gaussian process, Ann. Math. Stat. 43 (1972), 1235-1246.

[64] Adams, R. A., Sobolev Spaces, Academic Press, New York, (1975).

[65] Dawson, D. A., Geostochastic calculus, Can. J. Stat. 6 (1978), 143-168.

[66] Doob, J. L., Classical Potential Theory and its Probabilistic Counterpart, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

[67] Folland, G. B., Introduction to Partial Differential Equations, Mathematical Notes 17, Princeton Univ. Press, 1976.

[68] Ichikawa, A., Linear stochastic evolution equations in Hilbert spaces, J. Diff. Eq. 28 (1978), 266-283.

[69] Kotelenz, P., Thesis, U. of Bremen, 1982.

[70] Wolpert, R. L., Local time and a particle picture for Euclidean field theory, J. Fcl. Anal. 30 (1978), 341-357.

I N D E X

A

Aldous' theorem 354
arboreal ordering 386

B

Bakry's example 305
balayage 421
barrier problem 326
birth place 387
birth time 387
branching Brownian motion 401
Brownian density process 385
Brownian sheet 269-270
Burkholder's inequalities 376

C

covariance functional 289

D

death time 387
distribution-valued processus 330
(L^p)-dominated jumps 376
dominating measure 291

E

eigenfunction expansion 343
elementary function 292
Euclidean free field 419

F

finite dimensional distributions convergence 353
first hitting time 401
Fourier transform 335
(Stochastic) Fubini theorem 296

H

Hilbertian norm 330
Hölder classes 381

K

Kolmogorov theorem 273
Kurtz theorem 355

L

Laplace's equation 415
Levy's Markov property 277
Limits of the Brownian density process 417
linear equation 323

M

martingale measure 286
martingale measure (orthogonal) 288
martingale measure (nuclear covariance) 288
martingale measure (worthy) 291
measure diffusion 398
σ-finite L^2-valued measure 287
Mitoma's theorem 358

N

neurophysiology 311

nuclear covariance 303

nuclear space 332

O

Ornstein-Uhlenbeck sheet 271

orthogonal measures 298

P

parabolic equations 339

predictable σ-field 292

R

random linear functional 332

regularization 332

relatively compact set of probability measures 350

S

simple function 292

singular point 281

singularities propagation 279

Sobolev spaces 334

square of Brownian density process 407

stochastic integrals 292

stopping point 276

T

tight set of probability measures 351

V

vibrating string 281

W

wave equation 308

weak convergence 349

weak convergence of stochastic integrals 367

weak convergence in the Gaussian case 393

weak solution (of a PDE) 309

white noise 269-285