The purpose of this lecture is to construct Wiener measure by a method due to Ed Nelson, and then to show how Wiener mesaure can be considered as a generalized Gaussian process in the sense of Gelfand. Then a straighforward computation shows that the derivative of this process is (the restriction to positive time) of "white noise". The details are found in the handout *Wiener.*

Contents

1 The Big Path Space.

Let $\dot{\mathbf{R}}^{\mathbf{n}}$ denote the one point compactification of $\mathbf{R}^n$. Let

$$\Omega := \prod_{0 \le t < \infty} \dot{\mathbf{R}}^{\mathbf{n}} \tag{1}$$

be the product of copies of $\dot{\mathbf{R}}^{\mathbf{n}}$, one for each non-negative t. This is an uncountable product, and so a huge space, but by Tychonoff's theorem, it is compact and Hausdorff. We can think of a point ω of Ω as being a function from $\mathbf{R}_+$ to $\dot{\mathbf{R}}^{\mathbf{n}}$, i.e. as as a "curve" with no restrictions whatsoever.

Let F be a continuous function on the m-fold product:

$$F: \prod_{i=1}^{m} \dot{\mathbf{R}}^{\mathbf{n}} \to \mathbf{R},$$

and let $t_1 \le t_2 \le \cdots \le t_m$ be fixed "times". Define

$$\phi = \phi_{F;t_1,\ldots,t_m} : \ \Omega \to \mathbf{R}$$

by

$$\phi(\omega) := F(\omega(t_1), \ldots, \omega(t_m)).$$

We can call such a function a **finite** function since its value at ω depends only on the values of ω at finitely many points. The set of such functions satisfies our abstract axioms for a space on which we can define integration. Furthermore, the set of such functions is an algebra containing 1 and which separates points, so is dense in $C(\Omega)$ by the Stone-Weierstrass theorem. Let us call the space of such functions $C_{fin}(\Omega)$.

If we define an integral I on $C_{fin}(\Omega)$ then, by the Stone-Weierstrass theorem it extends to $C(\Omega)$ and therefore, by the Riesz representation theorem, gives us a regular Borel measure on Ω. For each $x \in \mathbf{R}^n$ define I_x by

$$I_x(\phi) =$$

$$\int \cdots \int F(x_1, x_2, \ldots, x_m) p(x, x_1; t_1) p(x_1, x_2; t_2 - t_1) \cdots p(x_{m-1}, x_m, t_m - t_{m-1}) dx_1 \ldots dx_m$$

when $\phi = \phi_{F, t_1, \ldots, t_m}$ where

$$p(x, y; t) = \frac{1}{(2\pi t)^{n/2}} e^{-(x-y)^2/2t} \tag{2}$$

(with $p(x, \infty) = 0$) and all integrations are over $\dot{\mathbf{R}}^{\mathbf{n}}$. One checks that this definition is consistent using the semi-group proptery of the heat kernel.

So, for each $x \in \mathbf{R}^n$ we have defined a measure on Ω. We denote the measure corresponding to I_x by pr_x. It is a probability measure in the sense that $\mathrm{pr}_x(\Omega) = 1$.

The intuitive idea behind the definition of pr_x is that it assigns probability

$$\mathrm{pr}_x(E) :=$$

$$\int_{E_1} \cdots \int_{E_m} p(x, x_1; t_1) p(x_1, x_2; t_2 - t_1) \cdots p(x_{m-1}, x_m, t_m - t_{m-1}) dx_1 \ldots dx_m$$

to the set of all paths ω which start at x and pass through the set E_1 at time t_1, the set E_2 at time t_2 etc. and we have denoted this set of paths by E.

2 Paths are continuous with probability one.

This is proved making heavy use of the regularity of the measure together with some estimates involving Gaussian integrals. The conclusion is a famous theorem of Wiener:

Theorem 1 [Wiener.] *The measure* pr_x *is concentrated on the space of continuous paths, i.e.* $\mathrm{pr}_x(\mathcal{C}) = 1$. *In particular, there is a probability measure on the space of continuous paths starting at the origin with*

$$\mathrm{pr}_0(E) = \int_{E_1} \cdots \int_{E_m} p(0, x_1; t_1) p(x_1, x_2; t_2 - t_1) \cdots p(x_{m-1}, x_m, t_m - t_{m-1}) dx_1 \ldots dx_m$$

to the set of all paths ω *which start at* 0 *and pass through the set* E_1 *at time* t_1, *the set* E_2 *at time* t_2 *etc. and we have denoted this set of paths by* E.

2.1 Embedding in $\mathcal{S}'$.

subsectionEmbedding in $\mathcal{S}'$. For convenience we specialize to the case $n = 1$. Let

$$\mathcal{W} \subset \mathcal{C}$$

consist of those paths ω with $\omega(0) = 0$ and

$$\int_0^\infty (1+t)^{-2} w(t) dt < \infty.$$

Proposition 1 [**Stroock**] *The Wiener measure* pr_0 *is concentrated on* $\mathcal{W}$.

Proof by Fubini.

Each element of $\mathcal{W}$ defines a tempered distribution, i.e. an element of $\mathcal{S}'$ according to the rule

$$\langle \omega, \phi \rangle = \int_0^\infty \omega(t)\phi(t) dt.$$

This map from $\mathcal{W}$ to $\mathcal{S}'$ is continuous, hence

the Wiener measure pushes forward to give a measure on $\mathcal{S}'$.

3 Generalities about expectation, variance and characteristic functions.

Let V be a vector space (say over the reals and finite dimensional). Let X be a V-valued random variable. That is, we have some measure space $(M, \mathcal{F}, \mu)$ (which will be fixed and hidden in this section) where μ is a probability measure on M, and $X : M \to V$ is a measurable function. If X is integrable, then

$$E(X) := \int_M X d\mu$$

is called the **expectation** of X and is an element of V.

The function $X \otimes X$ is a $V \otimes V$ valued function, and if it is integrable, then

$$\mathrm{Var}(X) = E(X \otimes X) - E(X) \otimes E(X) = E\left(X - E(X)\right) \otimes \left((X - E(X)\right)$$

is called the **variance** of X and is an element of $V \otimes V$. It is by its definition a symmetric tensor, and so can be thought of as a quadratic form on V^*.

If $A : V \to W$ is a linear map, then AX is a W valued random variable, and

$$E(AX) = AE(X), \qquad \mathrm{Var}(AX) = (A \otimes A)\,\mathrm{Var}(X) \tag{3}$$

assuming that $E(X)$ and $\mathrm{Var}(X)$ exist. We can also write this last equation as

$$\mathrm{Var}(AX)(\eta) = \mathrm{Var}(X)(A^*\eta), \qquad \eta \in W^* \tag{4}$$

if we think of the variance as quadratic function on the dual space.

The function on V^* given by

$$\xi \mapsto E(e^{i\xi\cdot X})$$

is called the **characteristic function** associated to X and is denoted by ϕ_X.

ϕ_X determines $X_*\mu$. In other words, the *law* of the random variable (meaning $X_*\mu$) is determined by its characteristic function.

4 Gaussian measures and their variances.

Let d be a positive integer. We say that N is a **unit** (d-dimensional) **Gaussian random variable** if N is a random variable with values in $\mathbf{R}^d$ with density

$$(2\pi)^{-d/2}e^{-(x_1^2\cdots+x_d^2)/2}.$$

It is clear that $E(N)=0$ and, since

$$(2\pi)^{-d/2}\int x_i x_j e^{-(x_1^2\cdots+x_d^2)/2}dx = \delta_{ij},$$

that

$$\mathrm{Var}(N) = \sum_i \delta_i \otimes \delta_i \tag{5}$$

where $\delta_1,\ldots,\delta_d$ is the standard basis of $\mathbf{R}^d$. We will denote this tensor by I_d.

A V-valued random variable X is called **Gaussian** if (it is equal in law to a random variable of the form)

$$AN+a$$

where

$$A:\mathbf{R}^d \to V$$

is a linear map, where $a\in V$, and where N is a unit Gaussian random variable. Clearly

$$E(X)=a,$$

$$\mathrm{Var}(X) = (A\otimes A)(I_d)$$

or, put another way,

$$\mathrm{Var}(X)(\xi) = I_d(A^*\xi))$$

and hence

$$\phi_X(\xi) = \phi_N(A^*\xi)e^{i\xi\cdot a} = e^{-\frac{1}{2}I_d(A^*\xi)}e^{i\xi\cdot a}.$$

Conversely, any random variable whose charactgeristic function is of this form is a Gaussian. A random variable X is centered Gaussian if and only if $\xi\cdot X$ is a real valued Gaussian random variable with mean zero for each $\xi\in V^*$.

By elementary "functorial" arguments we prove that: Suppose we are given a random variable X with (whose law has) a density proportional to $e^{-S(v)/2}$ where S is a quadratic form which is given as a "matrix" $S = (S_{ij})$ in terms of a basis of V^*. Then $\mathrm{Var}(X)$ is given by S^{-1} in terms of the dual basis of V.

5 Generalized random processes.

Let us say that a probability measure μ on $\mathcal{S}'$ is a **centered Gaussian process** if every $\phi \in \mathcal{S}$, thought of as a function on the probability space $(\mathcal{S}', \mu)$ is a real valued centered random variable; in other words $\phi_*(\mu)$ is a centered Gaussian probability measure on the real line. If we denote this process by Z, then we may write $Z(\phi)$ for the random variable given by ϕ. We clearly have $Z(a\phi + b\psi) = aZ(\phi) + bZ(\psi)$ in the sense of addition of random variables, and so we may think of Z as a rule which assigns, in a linear fashion, random variables to elements of $\mathcal{S}$.

An example is Wiener measure. For each $\phi \in \mathcal{S}$ we get a real valued Gaussian random variable whose variance is given by

$$\int_0^\infty \int_0^\infty \min(s,t)\phi(s)\phi(t)dsdt = 2\int\int_{0\le s\le t} s\phi(s)\phi(t)dsdt. \qquad (6)$$

6 The derivative of Brownian motion is white noise.

If we have generalized random process Z as above, we can consider its derivative in the sense of generalized functions, i.e.

$$\dot{Z}(\phi) := Z(-\dot{\phi}).$$

$\dot{Z}(\phi)$ is a centered Gaussian random variable whose variance is given (according to (6)) by

$$2\int_0^\infty \left(\int_0^t s\dot{\phi}(s)ds\right)\dot{\phi}(t)dt.$$

We can integrate the inner integral by parts to obtain

$$\int_0^t s\dot{\phi}(s)ds = t\phi(t) - \int_0^t \phi(s)ds.$$

Integration by parts now yields

$$\int_0^\infty t\phi(t)\dot{\phi}(t)dt = -\frac{1}{2}\int_0^\infty \phi(t)^2 dt$$

and

$$-\int_0^\infty \left(\int_0^t \phi(s)ds\right)\phi(t)dt = \int_0^\infty \phi(t)^2 dt.$$

We conclude that the variance of $\dot{Z}(\phi)$ is given by

$$\int_0^\infty \phi(t)^2 dt$$

which we can write as

$$\int_0^\infty \int_0^\infty \delta(s-t)\phi(s)\phi(t) ds dt.$$

Notice that now the "covariance function" is the generalized function $\delta(s-t)$. The generalized process (extended to the whole line) with this covariance is called white noise because it is a Gaussian process which is stationary under translations in time and its covariance "function" is $\delta(s-t)$, signifying independent variation at all times, and the Fourier transform of the delta function is a constant, i.e. assigns equal weight to all frequencies.