LECTURE II - STOCHASTICS

John Gough
Aberystwyth

CIRM, 16-20 April 2018

Discrete Time White Noise

- We start with a discrete time model for noise. Suppose we have a sequence $\xi_1, \xi_2, \xi_3, \cdots$ of independent random variables occurring each Δt seconds and with

$$\langle \xi_k \rangle = 0, \quad \langle \xi_k^2 \rangle = 1.$$

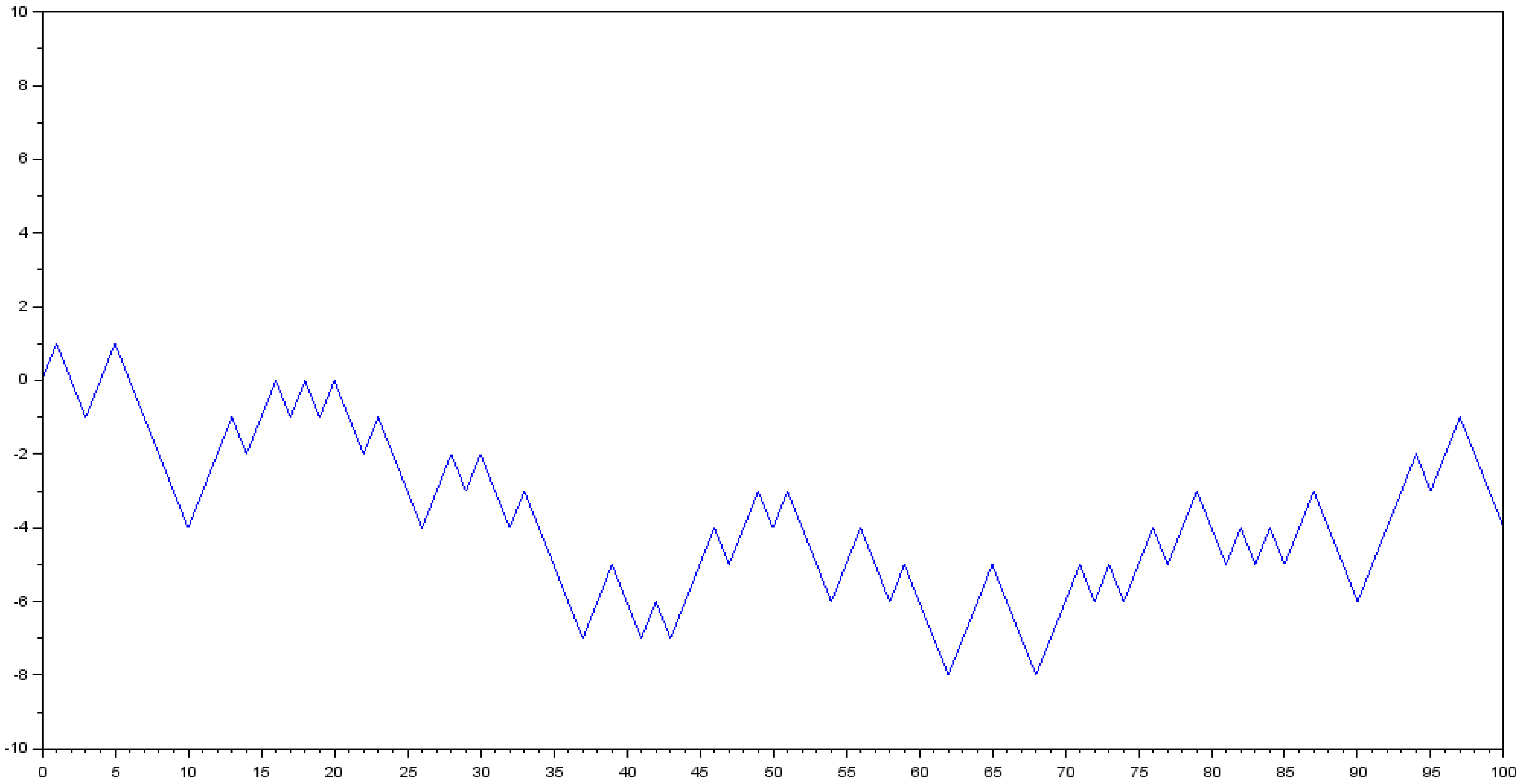
- A random walk, $X_1, X_2, X_3, \cdots$, is given by

$$X_n = \sum_{k=1}^{n} \xi_k$$

and we have

$$\langle X_n \rangle = 0, \quad \langle X_n^2 \rangle = n.$$

Typical Random Walks



De Moivre – Laplace Limit

- For time $t > 0$ fixed, let $N(t)$ be the largest integer less than or equal to $t/\Delta t$. Introduce the rescaled variable

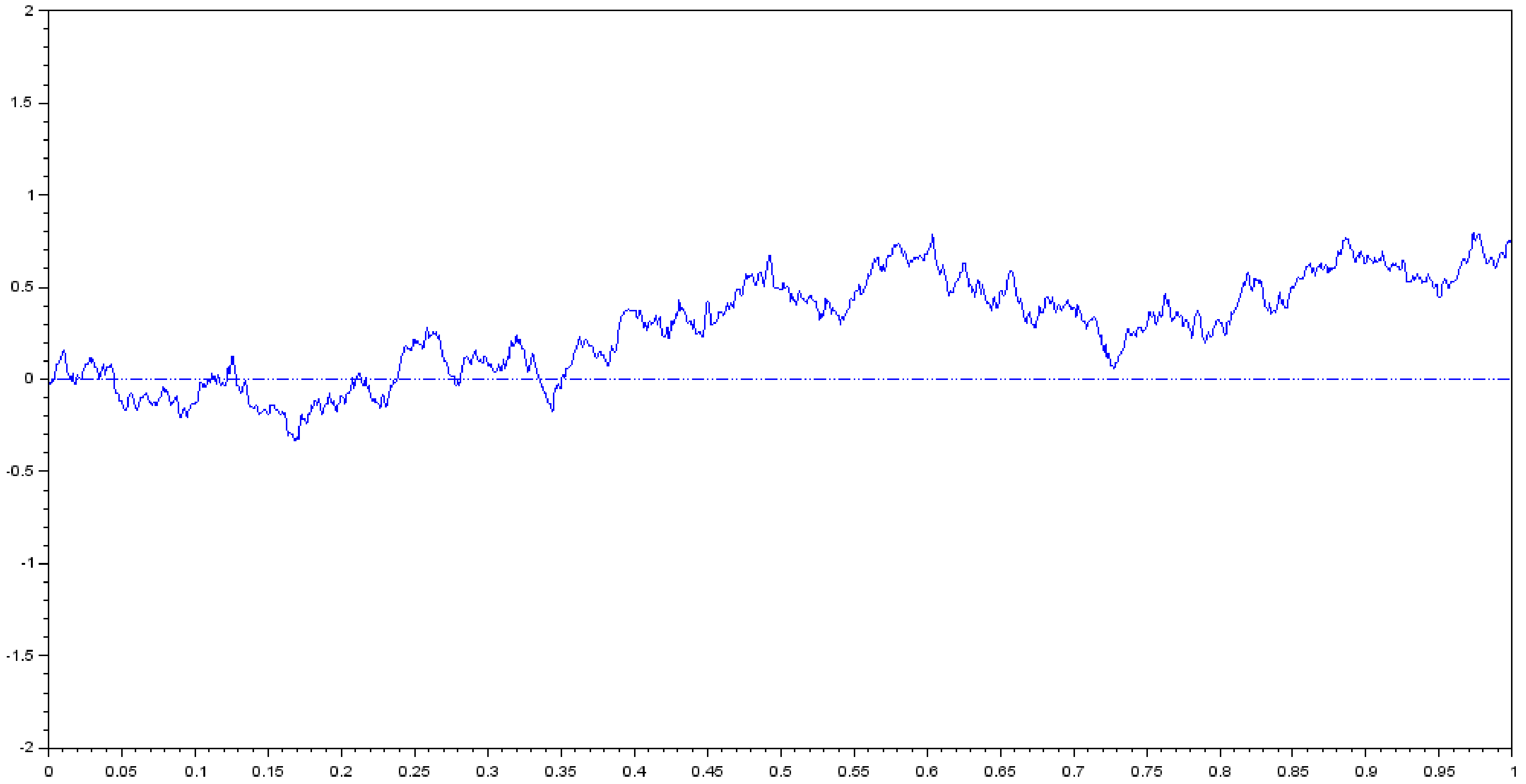
$$W_{\text{approx}}(t) = \sqrt{\Delta t} \sum_{k=0}^{N(t)} \xi_k .$$

- By a central limit, $W_{\text{approx}}(t)$ converges to a limit variable $W(t)$ which is Gaussian with

$$\langle W(t) \rangle = 0, \quad \left\langle W(t)^2 \right\rangle = t.$$

  The family $\{W(t) : t \geq 0\}$ obtained this way is called a **Wiener process**.

Wiener Paths



White Noise (Continuous Time)

Wiener showed that the limit process has sample paths that are almost always continuous and nowhere differentiable. Let us look at the approximate derivative

$$\dot{W}_k = \frac{\Delta W_k}{\Delta t} = \frac{W_{k+1} - W_k}{\Delta t} = \frac{\sqrt{\Delta t}\,\xi_{k+1}}{\Delta t},$$

then

$$\left\langle \dot{W}_k \right\rangle = 0, \quad \left\langle \dot{W}_k^2 \right\rangle = \frac{1}{\Delta t}$$

so the variance of $\dot{W}_k$ blows up as $\Delta t \to 0$. Formally, one may consider **white noise** to be the limit process $\dot{W}(t)$ which is Gaussian and δ-correlated:

$$\left\langle \dot{W}(t) \right\rangle = 0, \quad \left\langle \dot{W}(t)\, \dot{W}(s) \right\rangle = \delta(t - s).$$

Randomized Dynamics

- Let us start with the **1st order ODE**

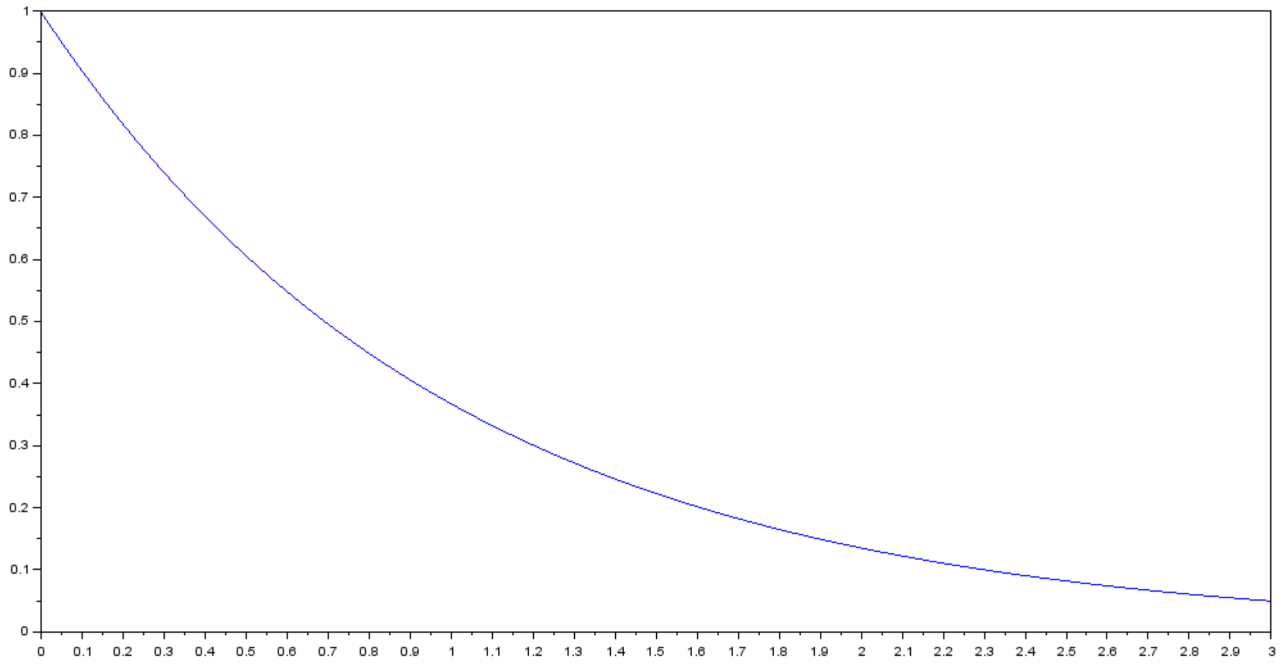
$$\dot{X}(t) = v(X(t)), \quad X(0) = x_0.$$

To solve this numerically we use a time step Δt as before and consider the discrete time iteration

$$X_{k+1} = X_k + v(X_k)\,\Delta t, \quad X_0 = x_0,$$

then $X_{\text{approx}}(t) = X_{N(t)}$ should converge to $X(t)$ as $\Delta t \to 0$.

Simulating an ODE



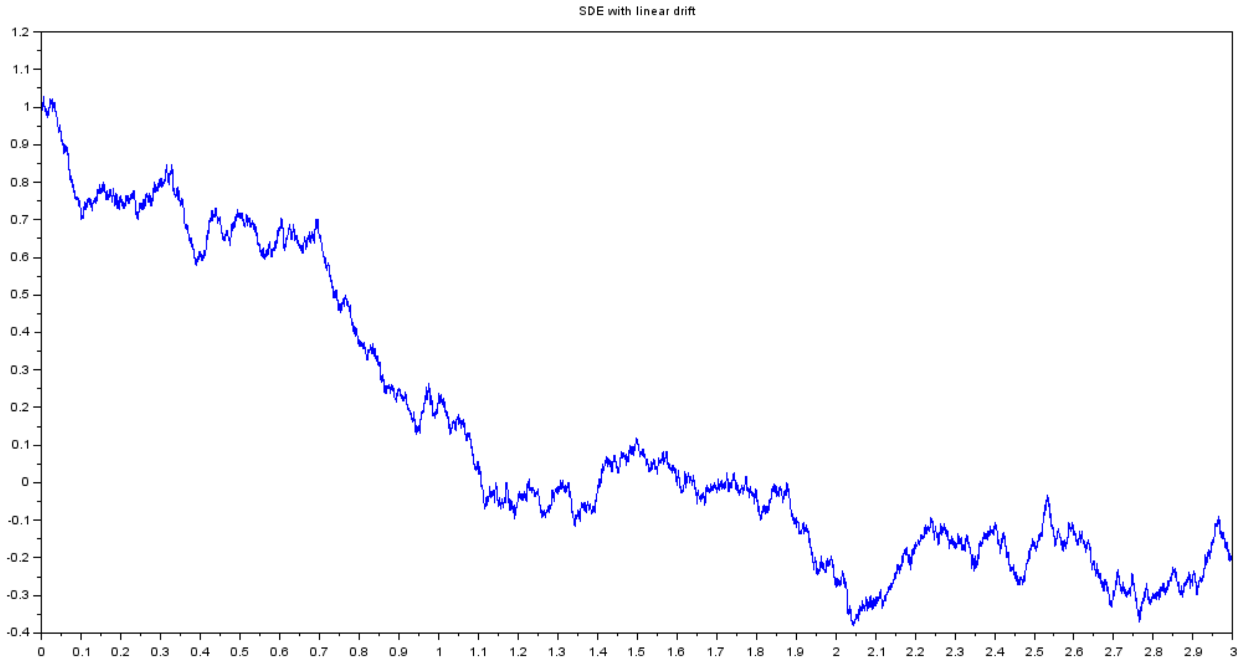
An approximation to the solution of the ODE $\dot{\mathbf{X}} = -\mathbf{X}$ with $\mathbf{X}(0)=1$.

We now try and add some noise:

$$\dot{X}(t) = v(X(t)) + \sigma \dot{W}(t), \quad X(0) = x_0.$$

This time we have the approximation scheme

$$X_{k+1} = X_k + v(X_k)\,\Delta t + \sigma\sqrt{\Delta t}\,\xi_{k+1}, \quad X_0 = x_0.$$



σ = 0.30

Ito form of the SDE

So far so good! But if we want to make σ depend on $X(t)$ then we need to be more precise. We interpret the SDE

$$dX(t) = v(X(t))\,dt + \sigma(X(t))\,dW(t), \quad X(0) = x_0,$$

to have *future pointing differentials*, that is

$$dX(t) \equiv X(t+dt) - X(t),$$

and is approximated by the scheme

$$X_{k+1} = X_k + v(X_k)\,\Delta t + \sigma(X_k)\sqrt{\Delta t}\xi_{k+1}, \quad X_0 = x_0.$$

The issue is that, while

$$\langle \sigma\left(X_k\right)\xi_{k+1}\rangle = \langle \sigma\left(X_k\right)\rangle\langle \xi_{k+1}\rangle = 0,$$

we have

$$\langle \sigma\left(X_k\right)\xi_k\rangle \propto \sqrt{\Delta t}$$

and so we would get a different limit process if we used ξ_k in the iteration rather than ξ_{k+1}.

Markov Processes

- To specify a **stochastic process** we must specify all its multi-time pdfs

$$\rho\left(x_n, t_n; \cdots; x_1, t_1\right)$$

for $X\left(t_1\right) = x_1, \cdots, X\left(t_n\right) = x_n$ for each $n \geq 0$.

- The process is **Markov** if we have

$$\rho\left(x_n, t_n; \cdots; x_1, t_1\right) = T(x_n, t_n | x_{n-1}, t_{n-1}) \cdots T(x_2, t_2 | x_1, t_1)\, \rho(x_1, t_1),$$

where whenever $t_n > t_{n-1} > \cdots > t_1$.

Transition Mechanism

- These give the conditional probabilities

$$\text{Prob}\{x \leq X(t) \leq x + dx | X(t_0) = x_0\} = T(x, t | x_0, t_0)\, dx,$$

for $t > t_0$.

- We have the propagation rule (Chapman-Kolmogorov equation)

$$\int T(x, t | x_1, t_1)\, T(x_1, t_1 | x_0, t_0)\, dx_1 = T(x, t | x_0, t_0),$$

for all $t > t_1 > t_0$.

Wiener's construction

- The Wiener process is a Markov process which starts at the origin at time zero, and has the transition mechanism

$$
\begin{aligned}
T\left(x, t \middle| x_0, t_0\right) &= \frac{1}{\sqrt{2\pi\left(t - t_0\right)}} e^{-\frac{\left(x - x_0\right)^2}{2\left(t - t_0\right)}}, \\
\rho\left(x, 0\right) &= \delta_0\left(x\right).
\end{aligned}
$$

- The transition mechanism is the Green's function for the heat equation

$$
\frac{\partial}{\partial t}\rho = \frac{1}{2}\frac{\partial^2}{\partial x^2}\rho.
$$

The Ito differential *dW(t)*

- This is not a true differential! The square of *dW(t)* is not negligible. Instead,

$$dW(t)\,dW(t) = dt.$$

- So, for instance, we have the Ito formula

$$\begin{aligned} dg(W(t)) &= g'(W(t))\,dW(t) + \frac{1}{2}g''(W(t))\,dW(t)^2 + \cdots \\ &= g'(W(t))dW(t) + \frac{1}{2}g''(W(t))dt. \end{aligned}$$

Functional Integration (Feynman & Kac)

Indeed, we have

$$\rho\left(x_n, t_n; \cdots; x_1, t_1\right) dx_n \cdots dx_1 \propto e^{-\sum_k \frac{(x_k - x_{k-1})^2}{2(t_k - t_{k-1})}} dx_n \cdots dx_1.$$

Formally, we may introduce a limit "path integral" with probability measure on the space of paths

$$\mathbb{P}^t_{\text{Wiener}}\left[d\mathbf{w}\right] = e^{-S_{\text{Wiener}}[\mathbf{w}]}\mathcal{D}\mathbf{w}.$$

where we have the action

$$S_{\text{Wiener}}\left[\mathbf{w}\right] = \int_0^t \frac{1}{2}\dot{w}\left(\tau\right)^2 d\tau.$$

Diffusions

- Let's return to the Ito SDE

$$dX(t) = v(X(t))\,dt + \sigma(X(t))\,dW(t), \quad X(0) = x_0.$$

- We have the stochastic differential expansion

$$\begin{aligned} dg(X(t)) &= g'(X(t))\,dX(t) + \frac{1}{2}g''(X(t))\,dX(t)^2 + \cdots \\ &= \Big[v(X(t))g'(X(t)) + \frac{1}{2}\sigma(X(t))^2 g''(X(t))\Big]dt + \sigma(X(t))g'(X(t))\,dW(t). \end{aligned}$$

Diffusions

- Averaging gives $$\langle dg(X(t))\rangle = \langle \mathcal{L} g(X(t))\rangle \, dt$$

where the generator of the diffusion is $$\mathcal{L} = v(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma(x)^2\frac{\partial^2}{\partial x^2}.$$

- Alternatively, this may be formulated as the Fokker-Planck equation

$$\frac{\partial}{\partial t}\rho(x,t) = \mathcal{L}^\star \rho(x,t) = -\frac{\partial}{\partial x}[v(x)\rho(x)] + \frac{1}{2}\frac{\partial}{\partial x^2}[\sigma(x)^2\rho(x)].$$

Filtering

- Suppose that we have a system described by a process $\{X(t) : t \geq 0\}$. We obtain information by observing a related process $\{Y(t) : t \geq 0\}$.

$$\begin{aligned} dX &= v(X)\,dt + \sigma(X)\,dW \quad \text{(stochastic dynamics)}, \\ dY &= h(X)\,dt + dZ \quad \text{(Noisy observations)}. \end{aligned}$$

  Here we assume that the dynamical noise W and the observational noise Z are independent Wiener processes.

- Estimate unknown *X(t)* using the observations ***y*** = *{ y(s) : 0 $\leq s \leq$ t}.*

We'll cheat a bit and use Path integrals

The joint probability of both X and Y up to time t is

$$\mathbb{P}^t_{X,Y}\left[d\mathbf{x}, d\mathbf{y}\right] = e^{-S_{X,Y}[x,y]}\mathcal{D}\mathbf{x}\mathcal{D}\mathbf{y},$$

where

$$\begin{aligned} S_{X,Y}\left[\mathbf{x},\mathbf{y}\right] &= S_X\left[\mathbf{x}\right] + \int_0^t \frac{1}{2}\left[\dot{y} - h\left(x\right)\right]^2 d\tau \\ &= S_X\left[\mathbf{x}\right] + S_{\text{Wiener}}[\mathbf{y}] - \int_0^t \left[h\left(x\right)\dot{y} - \frac{1}{2}h\left(x\right)^2\right] d\tau. \end{aligned}$$

Or,

$$\mathbb{P}^t_{X,Y}\left[d\mathbf{x}, d\mathbf{y}\right] = \mathbb{P}^t_X\left[d\mathbf{x}\right]\mathbb{P}^t_{\text{Wiener}}\left[d\mathbf{y}\right]\,\lambda\left(\mathbf{y}|\mathbf{x}\right).$$

where the **Kallianpur-Streibel likelihood** is

$$\lambda\left(\mathbf{y}|\mathbf{x}\right) = e^{\int_0^t\left[h(x)dy(\tau)-\frac{1}{2}h(x)^2d\tau\right]}.$$

The distribution for $X\left(t\right)$ given observations $\mathbf{y} = \left\{y\left(\tau\right) : 0 \leq \tau \leq t\right\}$ is then

$$\rho^{\text{post}}_t\left(x_t|\mathbf{y}\right) = \frac{\int_{x(0)=x_0}^{x(t)=x_t}\lambda\left(\mathbf{y}|\mathbf{x}\right)\mathbb{P}^t_X\left[d\mathbf{x}\right]}{\int_{x(0)=x_0}\lambda\left(\mathbf{y}|\mathbf{x}'\right)\mathbb{P}^t_X\left[d\mathbf{x}'\right]}.$$

The Filter

Let us write $\rho_t(x)$ for $\rho_t^{\text{post}}(x|\{Y(\tau) : 0 \le \tau \le t\})$. This is the pdf for $X(t)$ conditioned on the past observations $\{Y(\tau) : 0 \le \tau \le t\}$.

The estimate for $f(X(t))$ will be the **filter**

$$\pi_t(f) = \int \rho_t(x) f(x)\, dx = \frac{\int \sigma_t(x) f(x) dx}{\int \sigma_t(x') dx'}$$

where the non-normalized σ_t satisfies the **Duncan-Mortensen-Zakai equation**

$$d\sigma_t(x) = \mathcal{L}^* \sigma_t(x)\, dt + h(x)\sigma_t(x)\, dY(t).$$

The Filter Equations

The estimate for $f(X(t))$ will be the **filter**

$$d\pi_t(f) = \pi_t(\mathcal{L}f)\,dt + \{\pi_t(fh) - \pi_t(f)\pi_t(h)\}dI(t),$$

where the **innovations process** is defined as

$$dI(t) = dY(t) - \pi_t(h)\,dt.$$