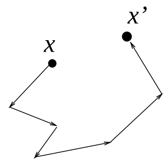
Notes on Itô Calculus

Mini-tutorial by Mitchell Feigenbaum

Notes by Adam Prügel-Bennett

We consider a diffusion process described by the "Green's function"

$$\rho(x', t'|x, t)$$



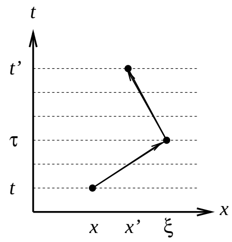
If we consider Brownian motion then this could be interpreted as the probability that a particle starting at a point x at time t moved to a point x' at time t'.

The Green's function has the following *properties*

1. Conserves probability

$$\int \rho(x', t'|x, t)\, \mathrm{d}x' = 1 \tag{1}$$

thus a particle starting at x at time t must go somewhere. Note that the integral over the past positions x does not conserve probability as diffusion is not a phase space conserving process.

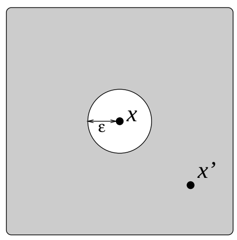


2. The process is Markovian

$$\rho(x', t'|x, t) = \int \rho(x', t'|\xi, \tau)\rho(\xi, \tau|x, t)\, \mathrm{d}\xi, \qquad t < \tau < t' \tag{2}$$

In addition to these two properties the fact that the Green's function describes a diffusion process implies

3. The particles only make small jumps



$$\int_{|x'-x|>\epsilon} \rho(x', t'|x, t)\, \mathrm{d}x' = o(t' - t) \tag{3}$$

so the probability of moving a distance ϵ or greater goes to zero as $t' \to t$, no matter how small we take ϵ.

Note $o(x)$ implies

$$\lim_{x \to 0} \frac{o(x)}{x} \to 0$$

i.e. $o(x)$ goes to zero strictly faster than x. While $O(x)$ implies

$$\lim_{x \to 0} O(x) \to 0$$

4. Sensible Moments Exist. The first two are

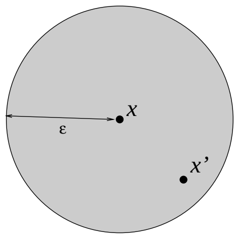
i) Drift

$$\int_{|x'-x|<\epsilon} (x' - x)\, \rho(x', t'|x, t)\, \mathrm{d}x' = a(x, t)\,(t' - t) + o(t' - t) \tag{4}$$

where now ϵ is not necessarily small. Thus the particle can have a net velocity $a(x, t)$.

ii) Diffusion

$$\int_{|x'-x|<\epsilon} (x' - x)^2 \rho(x', t'|x, t)\, \mathrm{d}x' = b(x, t)\,(t' - t) + o(t' - t) \tag{5}$$



The particle diffuses with a diffusion coefficient $b(x, t)$.

Remark 1 *When considering very short times the diffusion appears anomolously faster — that is a particle diffuses much faster than its net velocity which would increase the second moment only as $(t'-t)^2 = o(t'-t)$. Remember however that diffusion results from a large number of very small impulses. For example, in Brownian motion a particle diffuses because of the impacts it receives from the molecules of the liquid.*

Higher moments fall off normally (i.e. as $o(t'-t)$).

From the Green's function we can construct two functions $u(x,t)$ and $v(x,t)$ which satisfy backwards and forwards diffusion equations respectively. We will see that the function $u(x,t)$ is mathematically easier to handle.

To construct $u(x,t)$ we define a ball of final positions $\phi(x)$ which we assume is *uniformly continuous* and *bounded* with a bound B, say. For any $\phi(x)$ we can construct a function $u(x,t)$ through the equation

$$u(x,t) = \int \phi(x')\rho(x',T|x,t)\,\mathrm{d}x' \tag{6}$$

where $t < T$. Note that this is a function going backwards in time. It will of course depend on T and $\phi(x)$.

Remark 2 *Diffusion is an information lossing process — entropy increases and phase space contracts — so a final state can result from many possible initial states. Thus there is in general no well defined inverse for the diffusion process. The function $u(x,t)$ is* not *the distribution of x at an early time. It is simply defined through equation (6).*

The function $u(x,t)$ has the following properties

1. $\boldsymbol{u(x,T) = \phi(x)}$**.** The function $u(x,t)$ is identical to that of the initial blob $\phi(x)$ at time T. This is intuitively obvious but we can prove it rigorously by considering

$$u(x,t) - \phi(x) = \int [\phi(x') - \phi(x)]\rho(x',t'|x,t)\,\mathrm{d}x'$$

thus

$$\begin{aligned}\big|u(x,t) - \phi(x)\big| \;\leq\; & \int\limits_{|x'-x|<\epsilon} |\phi(x') - \phi(x)|\,\rho(x',t'|x,t)\,\mathrm{d}x' \\ & + \int\limits_{|x'-x|>\epsilon} |\phi(x') - \phi(x)|\,\rho(x',t'|x,t)\,\mathrm{d}x'\end{aligned}$$

But by continuity of $\phi(x)$ we have for $|x'-x| < \epsilon$ that $|\phi(x') - \phi(x)| < \delta$ where we can make δ as small as we like by choosing ϵ sufficiently small. Taking this out of the integral and using the fact that $\rho(x',t'|x,t)$ is normalized (property 1) we have that the first integral is smaller than δ. By the boundedness of $\phi(x)$ we know that $|\phi(x') - \phi(x)| \leq 2B$. Taking this out of the integral and using the locality property of $\rho(x',t'|x,t)$ (property 3) we find the second term is $o(t'-t)$. Thus taking $t' \to t$ the second term vanishes so that $|u(x,t) - \phi(x)| = 0$ proving the ascertion.

2. $u(x,t)$ satisfies a diffusion equation running backwards in time. To show this we consider

$$\begin{aligned} u(x,t) &= \int \phi(x')\rho(x',T|x,t)\,\mathrm{d}x' \\ &= \int \phi(x')\rho(x',T|\xi,\tau)\rho(\xi,\tau|x,t)\,\mathrm{d}\xi\,\mathrm{d}x' \\ &= \int u(\xi,\tau)\rho(\xi,\tau|x,t)\,\mathrm{d}\xi \end{aligned}$$

where $t < \tau < T$. We used the Markov property of $\rho(x',t'|x,t)$ (property 2) to obtain the second line and the definition of $u(x,t)$ to obtain the final line.

We can rewriting this as

$$u(x,t-\tau) = \int u(x',t)\rho(x',t|x,t-\tau)\,\mathrm{d}x'. \tag{7}$$

Using the Taylor expansion of $u(x',t)$ around x

$$u(x',t) = u(x,t) + (x'-x)u_x(x,t) + \frac{1}{2}(x'-x)^2 u_{xx}(x+\mu(x'-x),t) \tag{8}$$

where $0 \le \mu \le 1$, subtracting $u(x,t)$ and using the normalization condition and the definition of the moments (properties 1, 4i and 4ii) equation (7) becomes

$$u(x,t-\tau) - u(x,t) = \tau\left[a(x,t)\,u_x(x,t) + \frac{1}{2}b(x,t)\,u_{xx}(x,t)\right]$$

dividing through by τ and taking $\tau \to 0$ we find

$$u_t(x,t) + a(x,t)\,u_x(x,t) + \frac{1}{2}b(x,t)\,u_{xx}(x,t) = 0 \tag{9}$$

This has the form of a diffusion equation going backwards in time.

What happened to the $\mu(x'-x)$ term?
— A. P-B.

We can also define the function

$$v(x,t) = \int \rho(x,t|x',t')v(x',t')\,\mathrm{d}x' \tag{10}$$

where now we integrate over the postions at time t' and consider what happens at *future* times. Because the properties of $\rho(x,t|x',t')$ apply to the future time it is harder to derive the properties of $v(x,t)$, however, we will show that it satisfies a diffusion equation going forward in time. To see this we consider the term

$$I = \int u(x,t)\,\rho(x,t|x',t+\tau)\,v(x',t+\tau)\,\mathrm{d}x'\,\mathrm{d}x = \int u(x,t)\,v(x,t)\mathrm{d}x$$

by the definition of $v(x,t)$. Expand $u(x,t)$ about x, using equation (8) we find

$$\begin{aligned} I &= \int \left[u(x',t) + (x-x')u_x(x',t) + \frac{1}{2}(x-x')^2 u_{xx}(x'+\mu(x-x'),t)\right] \\ &\qquad\qquad \times\rho(x,t|x',t+\tau)\,v(x',t+\tau)\,\mathrm{d}x'\,\mathrm{d}x \\ &= \int \left[u(x',t) - \tau a(x',t)\,u_x(x',t) - \frac{\tau}{2}a(x',t)\,u_{xx}(x',t)\right]\,v(x',t+\tau)\,\mathrm{d}x' \end{aligned}$$

where we have integrated over x and used properties 1, 4i and 4ii. Subtracting the first term and dividing through by $-\tau$

$$\int u(x,t)\left(\frac{v(x,t+\tau)-v(x,t)}{\tau}\right)\mathrm{d}x =$$

$$\int\left[a(x,t)\,u_x(x,t)+\frac{1}{2}a(x,t)\,u_{xx}(x,t)\right]\,v(x,t+\tau)\mathrm{d}x$$

Integrating by parts and taking $\tau\to 0$ we find

$$v_t(x,t) = -\partial_x\left(a(x,t)v(x,t)\right)+\frac{1}{2}\partial_{xx}\left(b(x,t)v(x,t)\right) \tag{11}$$

This is the *Fokker-Planck equation.* It has the usual form of a diffusion equation going forwards in time.

Gaussian Processes

Gaussian diffusion processes are an interesting class of diffusion processes in that they can be solved analytically. We consider a physical observable $x(t)$ which results from N (infinitesimal) random events

$$x_N(t) = \sum_{n=0}^{N-1}\epsilon_n u(t-n\tau)\Theta\left(t-n\tau\right) \tag{12}$$

where $u(t-n\tau)$ is an arbitrary function and ϵ_n are independent indentically distributed random variables with zero mean and variance ϵ (i.e. $\langle\epsilon_n\rangle=0$ and $\langle\epsilon_n\epsilon_m\rangle=\epsilon\delta_{n,m}$, we will assume all higher moments exist although technically only the $2+\delta$ absolute moment needs to exist for some $\delta>0$). We denote this distribution $P(\epsilon_n)$. For example, if u is a constant and ϵ_n is equal to $\pm\epsilon$ with equal probability then the observable $x(t)$ could be interpreted as the accumulative result of a series of N coin tosses, adding $u\epsilon$ for a head and $-u\epsilon$ for a tail.

The α absolute moment of a distribution $P(x)$ is defined by

$$\int|x|^{\alpha}\,P(x)\,\mathrm{d}x.$$

For this system we can calculate the probability distribution, $\rho(x_N(t))$, for the observable, $x_N(t)$, and from this the Green's function $\rho(x,t|x',t')$. To do this we first consider the generalized fourier transform of the distribution

$$\hat\rho_N(k(t)) = \int\rho(x_N(t))\,\mathrm{e}^{-\mathrm{i}\int_0^{Nt}k(t)x_N(t)\,\mathrm{d}t}\,\mathcal{D}x_N(t) \tag{13}$$

Since the $x_N(t)$'s depend on the ϵ_n's we can replace the integral over $x_N(t)$ weighted by $\rho(x_N(t))$ with an integral over ϵ_n weighed by $P(\epsilon_n)$

The logarithm of the fourier transform is a generator for the cumulants, κ_n of a distribution

$$\log[\hat\rho_N(k)] = \sum_{n=1}(-\mathrm{i})\frac{\kappa_n}{N!}$$

$$\hat\rho_N(k(t)) = \prod_{n=0}^{N-1}\int P(\epsilon_n)\,\mathrm{e}^{-\mathrm{i}\int_0^{Nt}k(t)x_N(t)\,\mathrm{d}t}\,\mathrm{d}\epsilon_n$$

Since the ϵ_n's are independent by assumption the integrals all decouple. Thus

$$\hat\rho_N(k(t)) = \prod_{n=0}^{N-1}\left\langle\mathrm{e}^{-\mathrm{i}\epsilon_n I(n)}\right\rangle_{\epsilon_n}$$

where

$$I(n) = \int_{n\tau}^{Nt} k(t) u\,(t - n\tau)\,\mathrm{d}t.$$

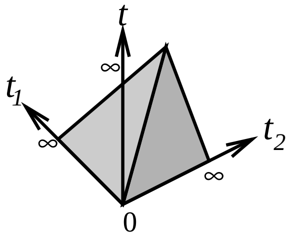
Expanding the exponential

$$\begin{aligned}
\prod_{n=0}^{N-1} \left\langle \mathrm{e}^{-\mathrm{i}\epsilon_n I(n)} \right\rangle_{\epsilon_n} &= \prod_{n=0}^{N-1} \left\langle 1 - \mathrm{i}\epsilon_n I(n) - \frac{\epsilon_n^2}{2} I^2(n) + \mathrm{i}\frac{\epsilon_n^3}{3!} I^3(n) + \cdots \right\rangle_{\epsilon_n} \\
&= \prod_{n=0}^{N-1} \left(1 - \frac{\epsilon^2}{2} I^2(n) + \mathrm{i}\frac{\langle \epsilon_n^3 \rangle}{3!} I^3(n) + \cdots \right) \\
&= \exp\left\{ -\frac{\epsilon^2}{2} \sum_{n=0}^{N-1} I^2(n) + \mathrm{i}\frac{\langle \epsilon_n^3 \rangle}{3!} \sum_{n=0}^{N-1} I^3(n) + \cdots \right\}
\end{aligned}$$

We take $N \to \infty$ and replace the sum by an integral

$$\sum_{n=0}^{N-1} \longrightarrow \frac{1}{\tau} \int_0^\infty \mathrm{d}t$$

to obtain a non-trivial limit we require that ϵ and τ go to zero in such a way that ϵ^2/τ go to a constant, σ^2 say. In this limit all the higher cumulants $\langle \epsilon_n^3 \rangle / \tau$ will go to zero so we are left with



$$\log\left[\hat{\rho}_(k(t))\right] = -\frac{\sigma^2}{2} \int_0^\infty \mathrm{d}t \int_t^\infty \mathrm{d}t_1 \int_t^\infty \mathrm{d}t_2\, u(t_1 - t)\, u(t_2 - t)\, k(t_1)\, k(t_2)$$

Changing the order of integration

$$\log\left[\hat{\rho}_(k(t))\right] = -\frac{1}{2} \int_0^\infty \mathrm{d}t_1 \int_0^\infty \mathrm{d}t_2\, k(t_1)\, k(t_2)\, K(t_1, t_2)$$

where

$$K(t_1, t_2) = \sigma^2 \int_0^{\min(t_1,t_2)} u(t_1 - t)\, u(t_2 - t)\,\mathrm{d}t \tag{14}$$

This defines a *Gaussian process* where $K(t_1, t_2)$ is the variance of the Gaussian. The fourier transform, $\hat{\rho}(k(t))$, acts as a generating function for the moments

$$\begin{aligned}
\mathrm{i}\left.\frac{\partial \hat{\rho}}{\partial k(t)}\right|_{\kappa(t)=0} &= \langle x(t) \rangle = 0 \\
-\left.\frac{\partial^2 \hat{\rho}}{\partial k(t) \partial k(t')}\right|_{\kappa(t)=0} &= \langle x(t) x(t') \rangle = K(t, t')
\end{aligned}$$

Putting $k(t) = k_1\,\delta(t_1 - t)$ and taking the inverse fourier transform we find

$$\rho(x, t) = \frac{1}{\sqrt{2\pi K(t,t)}}\,\mathrm{e}^{-x^2/(2K(t,t))}$$

note that $K(0,0) = 0$ so that this describes the diffusion of a particle diffussion from a point $x = 0$ at time $t = 0$, thus it is equal to the Green's function

$\rho(x,t|0,0)$. We want to know the Green's function $\rho(x_2,t_2|x_1,t_1)$ which we can obtain from the identity

$$\rho(x_2,t_2|x_1,t_1) = \frac{\rho(x_2,t_2;x_1,t_1|0,0)}{\rho(x_1,t_1|0,0)}$$

where $\rho(x_2,t_2;x_1,t_1|0,0)$ is the probability for a particle starting at $x=0$ at time $t=0$ to pass through the point x' at time t' and then through a point x at time t. To calculate this probability we put

$$k(t) = k_1\delta(t-t_1) + k_2\delta(t-t_2)$$

into the fourier transform for the probability distribution, $\hat{\rho}(k(t))$. Doing this we obtain

$$\rho(x_2,t_2|x_1,t_1) = \frac{\mathrm{e}^{\dfrac{-\left(x_2 - K^2(t_1,t_2)/K(t_1,t_1)\right)^2}{2\left(K(t_2,t_2) - (K^2(t_1,t_2)/K(t_1,t_1))x_1\right)}}}{\sqrt{2\pi\left(K(t_2,t_2) - K^2(t_1,t_2)/K(t_1,t_1)\right)}}. \tag{15}$$

This is the Green's function for a Gaussian diffusion equation. The corresponding backwards diffusion equation is

$$u_t(x,t) + \left(\alpha(t) + x\beta(t)\right) u_x(x,t) + \gamma(t) u_x x(x,t) = 0. \tag{16}$$

This is the most general form of a backwards diffusion equation satisfied by a Gaussian diffusion process.

Itô Calculus

Itô calculus gives rules for transforming equation involving infintesimal random variables

$$\mathrm{d}w \equiv \sqrt{\mathrm{d}t}N(0,1) \tag{17}$$

where $N(\mu,\sigma)$ means a random normally distributed variable with mean μ and standard deviation σ. We call this infinesimal "d Wiener" since it is the element of a Wiener process. Equations involving $\mathrm{d}w$ are called Stochastic differential equations; for example, *Is this right?*

$$\mathrm{d}x = a\,\mathrm{d}t + b\,\mathrm{d}w \tag{18}$$

the first term defines a drift while the second term defines a diffusion. All diffusion equations can be written in terms of such stochastic differential equations whatever the underlying mechanism causing the diffusion. The normal rules of calculus do not apply to stochastic differential equations since the functions involved are not sufficiently smooth. Nevertheless it is possible to define a consistent calculus. Various different calculuses can be constructed, the most useful is Itô's calculus.

To get a feel for what is involved let us consider integrating equation (18) for a very short time. Taking the expectation we find

$$\langle \mathrm{d}x\rangle = \langle x' - x\rangle = a\,\mathrm{d}t + \langle b\,\mathrm{d}w\rangle$$

for this to be a diffusion process we expect $\langle x'-x\rangle$ to be equal to $a\,\mathrm{d}t + o(\mathrm{d}t)$ by property 4i of the Green's function, thus $\langle b\,\mathrm{d}w\rangle$ should vanish. Similarly if we consider

$$\left\langle (\mathrm{d}x)^2\right\rangle = \left\langle (x'-x)^2\right\rangle = a^2(\mathrm{d}t)^2 + 2\,\langle ab\,\mathrm{d}w\rangle\;\mathrm{d}t + b^2\,\mathrm{d}t$$

where we have use equation (17). By property 4ii of the Green's function we expect this to be proportional to $b^2\,\mathrm{d}t$, the first term vanishes being proportional to $(\mathrm{d}t)^2$, thus we require $\langle ab\,\mathrm{d}w\rangle = 0$. Thus the solution to these stochastic differential equations will also be solution to the diffusion equation provided we can insure that

$$\langle f(w)\,\mathrm{d}w\rangle = 0 \tag{19}$$

that is the expectation of any function of Wiener d Wiener is zero. This will be the case if we can construct the function $f(w)$ so that it is independent of $\mathrm{d}w$ — since then we can average over $\mathrm{d}w$ separately which gives zero by equation (17). This is the condition that b is "non-anticipatory". When this holds we can immediately write down the corresponding backwards diffusion equation

$$u_t + au_x + \frac{b^2}{2}u_{xx} = 0. \tag{20}$$

Using the correspondence between the stochastic differential equation and the backwards diffusion equation we can derive the rules of Itô calculus for the stochastic differential equation. This will then allow us to make coordinate transformation and find how they transform the diffusion equation. We consider a change of variables

$$y = y(x,t)$$

where x satisfies the stochastic differential equation (18). The corresponding solution to the diffusion equation is $u(x,t)$ (note that this is a smooth function which obeys the normal rules of calculus). Writing $u(x,t)$ interms of y

$$u(x,t) = v(y(x,t),t)$$

then

$$\begin{aligned} u_t &= v_y\,y_t \\ u_x &= v_y\,y_x \\ u_{xx} &= v_{yy}\,y_x^2 + v_y\,y_{xx} \end{aligned}$$

using equation (20) we find $v(y,t)$ satisfies the equation

$$v_t + (y_t + ay_y + \frac{b^2}{2}y_{xx})v_y + \frac{1}{2}(by_x)v_{yy} = 0.$$

This has the form of a backwards diffusion equation so we can write down the corresponding stochastic differential equation

$$\mathrm{d}y = (y_t + ay_x + \frac{b^2}{2}y_{xx})\,\mathrm{d}t + by_x\,\mathrm{d}w. \tag{21}$$

But using equation (18) this is just

$$\mathrm{d}y = y_t\,\mathrm{d}t + y_x\,\mathrm{d}x + \frac{b^2}{2}y_{xx}\,\mathrm{d}t. \tag{22}$$

The first two terms are the usual terms you would expect from a coordinate transformation; the last term is a new term coming from the invariance of the diffusion equation — it is called *Itô's lemma.*

As an example let as consider

$$\mathrm{d}x = \mathrm{d}w \qquad \text{and} \qquad y = x^2$$

then by Itô's lemma (using $y_{xx} = 2$)

$$\mathrm{d}y = 2x\,\mathrm{d}x + \mathrm{d}t = 2\sqrt{y}\,\mathrm{d}w + \mathrm{d}t$$

and thus the corresponding backwards diffusion equation is

$$u_t + u_y + 2y\,u_{yy} = 0.$$

(Note that if we are given such a diffusion equation we could perform the inverse coordinate transformation and solve the much easier equation $\mathrm{d}w = \mathrm{d}x$, we then have an answer for y in terms of w. This is the typical form of the solution given by Itô's method.)

We can ask whether we can define a sensible integration procedure for solving stochastic differntial equations. Taking the last example if we assume that $x(0) = 0$ so that $x = w$ then from

$$\mathrm{d}y = \mathrm{d}w^2 = 2x\,\mathrm{d}x + \mathrm{d}t = 2w\,\mathrm{d}w + \mathrm{d}t$$

we deduce

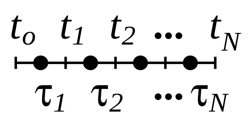
$$w\,\mathrm{d}w = \frac{1}{2}\,\mathrm{d}w^2 - \frac{1}{2}\,\mathrm{d}t$$

the last term arising from Itô's lemma. Integrating we would have

$$\int_{w_0}^{w_t} w\,\mathrm{d}w = \frac{1}{2}w^2(t) - \frac{1}{2}t$$

The expectation is thus equal to zero; this is what we expect from equation (19) — if we did not have the extra Itô term then this would not be true. Let us consider taking integration in the Riemann sense (we could and probably should use measure theory to do this properly but then the subject becomes too dull). The integral is just approximated by a sum

$$\int_0^t w\,\mathrm{d}w \overset{N^{th}\,approx}{\longrightarrow} \sum_{n=1}^{N} w(\tau_n)\left(w(t_n) - w(t_{n-1})\right)$$



For a sufficiently smooth function the result would be independent of where we choose the τ_n's — this will not be true in this case as w is not a smooth function. To evaluate the sum we separate it into three terms $I_1 + I_2 + I_3$ where

$$\begin{aligned}
I_1 &= \sum_{n=1}^{N} w(t_{n-1})\left(w(t_n) - w(t_{n-1})\right) \\
I_2 &= \sum_{n=1}^{N} \left(w(\tau_n) - w(t_{n-1})\right)^2 \\
I_3 &= \sum_{n=1}^{N} \left(w(t_n) - w(\tau_n)\right)\left(w(\tau_n) - w(t_{n-1})\right)
\end{aligned}$$

We consider I_2 first. We know that

$$\left\langle (\Delta w)^2 \right\rangle = \left\langle \left(w(\tau_n) - w(t_{n-1})\right)^2 \right\rangle = \tau$$

but we would like to make the stronger statement $(\mathrm{d}w)^2 = \mathrm{d}\tau$. To show this is indeed the case we consider the expected variation

$$\left\langle \left((\mathrm{d}w)^2 - \mathrm{d}\tau\right)^2 \right\rangle = \left\langle (\mathrm{d}w)^4 \right\rangle - 2\mathrm{d}t \left\langle (\mathrm{d}w)^2 \right\rangle + (\mathrm{d}t)^2 = 3(\mathrm{d}t)^2 = 0$$

Thus

$$I_2 = \sum_{n=1}^{N} (\tau_n - t_{n-1})$$

when $\tau_n - t_{n-1} = \alpha(t_n - t_{n-1})$ then $I_2 = \alpha t$. The term I_3 can be written as

$$I_3 = \frac{1}{2} \sum_{n=1}^{N} (A_n + B_n)^2 + A_n^2 + B_n^2$$

where

$$A_n = \big(w(t_n) - w(\tau_n)\big) \qquad B_n = \big(w(\tau_n) - w(t_{n-1})\big)$$

But using $(\mathrm{d}w)^2 = \mathrm{d}\tau$ we find

$$I_3 = \frac{1}{2} \sum_{n=1}^{N} t_n - t_{n-1} - (t_n - \tau_n) - (\tau_n - t_{n-1}) = 0$$

To evaluate I_1 we use summation by parts. Thus

Summation by parts

$\sum_{n=1}^{N} a_{n-1}(b_n - b_{n-1})$

$= \sum_{n=1}^{N} a_{n-1} b_n$

$\quad - \sum_{n=0}^{N-1} a_n b_n$

$= -\sum_{n=1}^{N} b_n(a_n - a_{n-1})$

$\quad + a_N b_N - a_0 b_0$

$$\begin{aligned} I_1 &= \sum_{n=1}^{N} w(t_{n-1}) \big(w(t_n) - w(t_{n-1})\big) \\ &= -\sum_{n=1}^{N} w(t_n) \big(w(t_n) - w(t_{n-1})\big) + w_N^2 - w_0^2 \end{aligned}$$

Adding these two expressions for I_1 we find

$$\begin{aligned} 2I_1 &= -\sum_{n=1}^{N} \big(w(t_n) - w(t_{n-1})\big)^2 + w_N^2 - w_0^2 \\ &= -t + w_N^2 - w_0^2 \end{aligned}$$

Thus putting together these results we find

$$\int_0^t w \, \mathrm{d}w = \frac{1}{2}(w_N^2 - w_0^2) - \frac{t}{2} + \alpha t$$

where α depends on where we place τ_n in the interval between t_n and t_{n+1}. In the Itô calculus we take $\alpha = 0$ so that τ_n is at the beginning of the interval (thus insuring $f(w)$ and $\mathrm{d}w$ are independent). The choice $\alpha = 1/2$ gives the Stranonvich integral which makes the calculus look like normal calculus but then adds the complication that equation (19) is not satisfied.

We saw above (equation (21)) that making a change of coordinates from x to $y(x,t)$ where x satisfies the equation (18) then y satisfies

$$dy = (y_t + a y_x + \frac{b^2}{2} y_{xx}) \, \mathrm{d}t + b y_x \, \mathrm{d}w.$$

Suppose we want to solve the backwards diffusion equation

$$y_t + a y_x + \frac{b^2}{2} y_{xx} = 0$$

this tells us that

$$\int_t^T \mathrm{d}y = \int_t^T b y_x \,\mathrm{d}w$$

but we know that

$$\left\langle \int b y_x \,\mathrm{d}w \right\rangle = 0$$

by equation (19) thus

$$\langle y(x(T), T) - y(x,t) | x, t \rangle = 0$$

that is starting from x at time t the expectation of y at a later time is the value of y at the time t — a process with this property is known as a *Martingale*. A typical example of such a process is a random walk where the expectation for the position of a particle is just the last measured position. This then provides an interpretation of the backwards solution of the diffusion equation

$$y(x,t) = \langle y(x(T), T) | x, t \rangle$$

it is the expected value of $y(x(T), T)$ starting from a point x at time t.

To use Itô's calculus we need to know how to transform from ones set of coordinates, $y(x,t)$ say, to another, $z(x,t)$. Suppose that

$$\begin{aligned}
\mathrm{d}x &= \mu \,\mathrm{d}t + \sigma \,\mathrm{d}w \\
\mathrm{d}y &= a \,\mathrm{d}t + b \,\mathrm{d}w \\
\mathrm{d}z &= \alpha \,\mathrm{d}t + \beta \,\mathrm{d}w
\end{aligned}$$

then by Itô's lemma

$$\begin{aligned}
\mathrm{d}(yz) &= \mathrm{d}t(y_t z + y z_t) + \mathrm{d}x(y_x z + y z_x) + \frac{\sigma^2}{2}\,\mathrm{d}t(y_{xx} z + y z_{xx} + 2 y_x z_x) \\
&= y \,\mathrm{d}z + z \mathrm{d}\, y + \sigma^2 y_x z_x \,\mathrm{d}t
\end{aligned}$$

This then provides us with the rule to make the coordinate transform $w \to yz$.
Example 1. We consider the equation

$$y_t + \mu y_t + \frac{\sigma^2}{2} y_{xx} = r(x,t) y(t)$$

we know from equation (21) that this is just equivalent to the stochastic differential equation

$$\mathrm{d}y = r y \,\mathrm{d}t + \sigma y_x \mathrm{d}w$$

where x satisfies the stochastic differential equation

$$\mathrm{d}x = \mu \mathrm{d}t + \sigma \mathrm{d}w$$

Now to solve the stochastic differential equation for y we make the coordinate transform $y = uz$ where u satisfies the differential equation

$$\mathrm{d}u = r u \mathrm{d}t$$

(note that this is not a diffusion equation). Solving this differential equation we find

$$u(x,t) = \exp\left(\int_0^t r(x,t')\mathrm{d}t'\right)$$

since u does not have a diffusion term $\mathrm{d}y = \mathrm{d}(uz)$ has the simple form

$$\mathrm{d}y = z\,\mathrm{d}u + u\,\mathrm{d}z = ry\mathrm{d}t + u\mathrm{d}t.$$

Thus

$$\mathrm{d}z = \frac{\sigma y_x}{u}\mathrm{d}w$$

and

$$\left\langle \int \mathrm{d}z \right\rangle = \left\langle \int \frac{\sigma y_x}{u}\mathrm{d}w \right\rangle = 0$$

Thus $z = y/u$ is a Martigale so that

$$y(x,t) = \left\langle y(x(T),T)\,\mathrm{e}^{-\int_t^T r\,\mathrm{d}t'}|x,t \right\rangle$$

This is known as *Black Shaol's Equation.* Is this right?

As a very simple example we consider

Example 2.

$$\mathrm{d}x = \sigma(t)\mathrm{d}w$$

making the change of variables

$$\mathrm{d}w = \sqrt{\frac{\mathrm{d}t}{\mathrm{d}s}}\mathrm{d}w_s$$

and putting

$$\sqrt{\frac{\mathrm{d}t}{\mathrm{d}s}} = \frac{1}{\sigma(t)}$$

then so that

$$s = \int^t \sigma^2(t')\mathrm{d}t'$$

then

$$\int \mathrm{d}x = x(t) - x(0) = N\left(0, \sqrt{\int_0^t \sigma^2(t')\mathrm{d}t'}\right)$$

As a final simple example we consider

Example 3.

$$\mathrm{d}x = \big(a(t) + xb(t)\big) + \sigma(t)\mathrm{d}w$$

introducing the integrating factor

$$\mathrm{e}^{\int^t b(t')\mathrm{d}t'}$$

then

$$\mathrm{d}\left(x\,\mathrm{e}^{\int^t b(t')\mathrm{d}t'}\right) = \mathrm{e}^{\int^t b(t')\mathrm{d}t'}(\mathrm{d}x + bx\mathrm{d}t) = \mathrm{e}^{\int^t b(t')\mathrm{d}t'}(a(t)\mathrm{d}t + \sigma(t)\mathrm{d}w)$$

aborbing the integrating factors into the functions $a(t)$ and $\sigma(t)$ we obtain an equation with the familiar form of a diffusion equation.

There is one class of problems which cannot be solved by path integral techniques because they do not describe Gaussian diffusion — the distribution of x is non-analytic. This is the class

$$\mathrm{d}x = \big(a(t) + x b(t)\big) + \big(A(t) + x B(t)\big)\,\mathrm{d}w$$

If $B(t) = 0$ this would be of the form of Gaussian process. To solve this we first consider the simpler equation (with $a(t) = A(t) = 0$)

$$\frac{\mathrm{d}x}{x} = b\,\mathrm{d}t + B\,\mathrm{d}w$$

But

$$\mathrm{d}\ln(x) = \frac{1}{x}\,\mathrm{d}x + \frac{1}{x^2} + B^2 x^2\,\mathrm{d}t$$

where the second term arises because of Itô's lemma. Thus

$$\mathrm{d}\ln(x) = \left(b - \frac{B^2}{2}\right)\,\mathrm{d}x + B\,\mathrm{d}w$$

To find the full form of the solution we substitute $x = uv$ and using variation of parameter.