**Ornstein-Uhlenbeck Process**

So the Wiener process forms the backbone of the Ornstein-Uhlenbeck process, which seems to adequately describe all the transfer matrix modeling going on.

**Scalar OU Process**

Now let’s say that W(t) drives the evolution of another random variable X(t) such that:



Now a(X,t) is the ‘drift’ term, whereas bw(t) is a random white noise ‘variance’ term. What does this mean? Well this gives an equation for each increment’s probability distribution. What is the probability distribution of dX? It’s linearly related to dW which is Gaussian distributed. Probability distribution of dXj is just the same as probability distribution of a(Xj)dt + b(Xj)dW(tj). So it’s just a normal distribution as well, with a mean of a(X)dt and variance b(X)2Ddt. Note that the p.d.f. of dW(t) is independent of X’s though, at that moment. The p.d.f. of each dXj would be different. We can get the pdf of X(t) via:



which translates to the following discrete version:



Note X is a Markov process because Xn+1 only depends on Xn. But it may not be a Martingale since <dX> = a(X)dt ≠ 0 probably. Even though the pdf of each dX is Gaussian, each increment has a different mean and variance, usually, and so the central limit theorem doesn’t apply and so p.d.f. of X(t) won’t usually be Gaussian. We’ll find it useful to switch from an Ito OU process to a Stratonovich OU process and vice versa. Let’s do Ito → Stratonovich. Note we’ll say dt = ti+1 – ti and dW = W(ti+1) – W(ti) for short.



Note we kept only terms up to O(dt). Filling in the ΔX term,



Now we can replace the b(X(ti+1/4),ti+1/4) with b(X(ti+1/2),ti+1/2) since the error involved in doing so is O(dW), which will ultimately vanish in the limit. So then proceeding:



Now we can as usual replace the dW product with its average, which is Ddt/2. So we get:



Dividing by dt, we get:



So in general, going from one to the other we have:



And,



**Integral functions of OU processes**

Let’s consider a function f(X,t). We want to examine the integral of this quantity, defined as:



To that end, let’s consider the antiderivative F(X,t) for which ∂F/∂X = f. Then,



Now we can relate ΔX to Δt, presumably as we’ve done before:



But we should check:



So yeah this should check out again. Then we can stop at this order in the expansion, and so then we have our result, backing up a little:



And we could solve for the integral too. Denoting ∂F/∂X = f.



**Differential functions of OU process**

Now we’ll reverse the integral formula and consider differentiation. We’d define:



and so,



So dividing by dt, we finally obtain:



Or more concisely,



So if we do the ‘extended calculus’ with Taylor series, then it works out. Now let’s consider the product rule:



and so we have:



Again, let’s consider the Ito heuristic:



So using the extended product rule would deliver the Ito version, if D were symmetric. I’d expect it always is. What of the chain rule?



And so we have:



**Vector OU Process**

Now let’s say that W(t) drives the evolution of another random variable X(t) via:



where tm+λ = tm + λ(tm+1 – tm) is some time within (tm, tm+1). In normal calculus, it would not matter where tm+λ was located within the interval, but we’ll see it does here. Two special cases stand out: Ito (λ = 0), and Stratonovich (λ = ½). The former is what we implicitly analyzed in the previous section. But the non-zero λ cases result in quite different evolution for X(t) as the a,b and dW terms will now be correlated.

**Integral functions of OU processes**

Let’s consider an integral of a function of a OU process. Let’s have function F(**X**,t):



Now we will now perform a Taylor series expansion of the difference, starting at time tk+λ. As before, we’ll find that the result *does* depend on what λ is. We’ll expand out to first order in dt:



Turns out the last term can be written as a purely deterministic function. To start, we’ll substitute the differential dX equation into this term, keeping only O(dt) terms.



One could demur with respect to the time we evaluate the b’s in the last term, but any choice of t within the interval (tk, tk+1) is equivalent, to our desired order. Now two stochastic functions are considered equal if their average squared difference equals zero (this effectively means the probability of a difference between the two functions is zero). This will imply we can replace the [dWdW - dWdW] term with its average: (1-2λ)DkℓΔtm. To prove this, we’ll take the difference between our present expression and the same expression with the [dWdW – dWdW] term replaced by its average, and analyze the expectation of its square. For concision, we’ll use h(**X**,t) to denote the ∂2F/∂Xi∂Xj·bikbjl term:



Now we can replace h(X(tm),tm) by its supremum over the interval (0,t):



And then using the identity, for independent variables Ai:



we can say:



But now, E[ ] = 0, and var[ ] term will go to zero in the large n limit since (Δtm)2 is of order 1/n2. So we have:



Or in other words:



**Differential functions of OU processes**

We basically have our result above, once the derivative of both sides is taken, but let’s do it independently. So as before, let’s start by considering some function F(X(t)), and its evolution, sans averaging at the moment. To proceed, we’ll expand it out to roughly first order in dt.



Now the dX term is problematic because the X’s in F are evaluated tm, while those in dX are evaluated at tm+λ. Let’s massage everything into the latter time. The ∂F/∂t term can be evaluated at tm+λ w/o cost because the error involved would ultimately be > O(dt) vis a vis dF. So



For the second term, we will make a Taylor expansion about tm+λ, keeping terms out to dt.



One could demur with respect to the time we evaluated *b* in the [Xj(tm) – Xj(tm+λ)] term, but any choice of t within the interval (tm, tm+1) would’ve been equivalent to any other, up to our desired order. Now turns out we can replace the product of W’s with their average in the small dt limit. Heuristically, this is because higher moments of that term are O(dt)2 (but really, this must be proved by placing the derivative within the context of integration). Perhaps this only happens for white noise then?



In dF’s third term, we may switch **X**(tm) → **X**(tm+λ) because the difference doesn’t matter in that term up to our desired order. Then expanding the d**X**’s and keeping only the appropriate order terms we have:



Again we can replace the W product with its average:



Putting it all together we come to:



Finally, dividing both sides by dtm, we arive at our result:



Can the Ito version be obtained heuristically?



But this heuristic will apparently only work for λ = 0. Now let’s consider the product rule:



and so we have:



Again, let’s consider the Ito heuristic:



So using the extended product rule would deliver the Ito version, if D were symmetric. I’d expect it always is. What of the chain rule?



And so we have:



As we can see, the Stratonovich version of the process follows the normal rules of Calculus. This makes it convenient for formal manipulations. Pertinently, one may take an Ito-defined SDE, manipulate it into a Stratonovich-defined SDE, and solve it using the normal rules of Calculus. Once **X**(t) is known, its moments and probability distribution function may be worked out, in principle.

**Fokker-Plank Equations**

Now let’s consider how to obtain an evolution equation for the probability distribution function P(X(t)). Formally all we have to do is take the average of F(X(t)), where F(X(t)) = δ(ξ – X(t)). But in its present form, the differential equation for F(X(t)) is inconvenient from the standpoint of evaluating expectations because some of the **X**’s are evaluated right in the middle of a d**W** interval, which makes the two correlated. What we want is for all **X** terms to be evaluated at the beginning of the interval, i.e., tm. So starting from the original equation,



let’s consider the first term in dF. This time we need to do the opposite of what we did last time, shift all the **X**’s in the d**X** term to tm. Doing so out to order O(dt),



Replacing the **W** product with its average, we get:



We may with impunity to O(dt) change the evaluation point of the **X**’s in the last term to tm, and then we may replace the **W** product with its average.



Altogether then, we have:



Now the d**W**’s are decoupled from all the **X**’s. Taking the average over all variables, which is implicitly,



and dividing by dt we get:



where,



We’ll note that the Ito version is similar to the previous section’s result. Even if the equation cannot be solved, a Taylor series expansion of the statistic can be calculated, via repeated application of the H-operator:



With this result in hand, we can write down the evolution equation for the probability distribution function: P(**ξ**,t) = <δ(**ξ** – **X**(t))>. Filling δ(**ξ** – **X**(t)) into F(**X**(t)). Like before, we will get:



and integrating by parts, we find:



where,



Let’s look at this from a different perspective, as is done in the Brownian motion file. Consider some F(X). How would its expectation evolve with time? Let’s look at it from the Ito perspective:



Now,



So filling this in…



I don’t think we could in general separate the average over the F and ΔX terms. For instance, we can’t say <X(t)X(tʹ)> = <X(t)><X(tʹ)> just because tʹ > t. But in an Ito process, the X’s in ΔX are evaluated at time t. So we can pull the X(s)’s out of the integral.



Then doing the averages, we find:



Finally we can divide both sides by τ:



**Example 1**

First the easiest:



We can integrate it to obtain,



Let’s take X(t) = at + bW(t). Then we see that:



And so now we could more or less straightforwardly ascertain the probability distribution of X. For instance,



which is consistent with our other calculations. Can we get a FP equation? So I guess there are two different ways to write P(ξ,t+dt).



I think we need to use the former in this case? Proceeding,



**Example 2**

Now one that commonly shows up in the physics of Brownian motion.



(where X would be velocity) We can formally solve this equation for X(t), assuming some specific input w(t), using GF’s. I think the result is:



Can we evaluate the integral? Well, in this case:



So,



So can’t do this exactly, it would seem? We can see from this solution that we cannot consider X to be X(W,t) exactly. Note that this is a weighted sum of Gaussian random variables, and so is itself a Gaussian distribution? All we’d need to calculate the probability distribution is knowledge of the mean and variance. These can now be obtained.



And,



In the long time limit, this would become,



And at the same time we’d have:



Now consider the differential:



So it does satisfy the ODE. What would a Taylor series expansion of X look like? Well it wouldn’t be in powers of dt alone, but also dW. Can see that to first order in t, we’d have (taking W(0) = 0):



I suppose that even simply follows from the ODE. Let’s look at the probability distribution function. So first, we have:



and so the probability distribution ought to be:



That’s interesting. What if I tried it a different way:



Yeah, don’t know how to do that. But at least we do have a formula for it. Let’s do an FP equation:



**Example 3**

Let’s move to another differential equation:



Let’s do a Taylor series expansion,



Keeping just to order t,



Now let’s consider taking the small t limit. In this limit it would seem that W2 → Ddt, in which case we’d return to the original ODE.

Now to solve the ODE, we would think we could divide by X, but as we know, we don’t necessarily have d(lnX)/dt ≠ X-1dX/dt. So the procedure we used earlier, treating w as a real variable, doesn’t always work. It only worked because the Ito SDE happened to be the same as its Stratonovich version. So let’s consider. If we assume that X = F(W,t), then it follows from the general rule dF/dt = ∂F/∂t + ∂F/∂W∙w + (D/2)∂2F/∂W2 that a(X,t) = ∂X/∂t + (D/2)∂2X/∂W2, and b(X,t) = ∂X/∂W. So then we would propose:



Which leaves us with the same result. More specifically, say we had supposed X = eG(W,t).



Now even the ∂G/∂t term could have a W in it. So there’s no point in equating coefficients. But let’s suppose ∂G/∂W = b nonetheless. Then G = bW + ψ(t). Filling this in, we’d have:



So our solution is:



Imposing initial conditions, we’d arrive at:



Another possibility here is to translate it to the Stratonovich version, for which ‘ordinary’ calculus applies. So,



Then we can do standard manipulations,



which is the same as before. Note the Taylor expansion is:



and this matches. In retrospect, I can see now how treating the preceding linear SDE as an ODE worked out as its Stratonovich version was the same as the Ito version. Now let’s look at the probability distribution:



**Example 4**

Let’s do a matrix example. So our equations are:



To solve, we need to put in Stratonovich form,



and,



So we have:



So for short:



One option is to go to the interaction picture,



Then (and note this product rule would not work unless we were in Stratonovich form)



and, then,



We can write this in matrix form as, trying to remind ourselves that w(t) is a time-dependent variable:



Does **B** commute with itself?



So we can write:



So then,



Now to evaluate this matrix, we need to diagonalize it?



So,



And so,



and so….



So,



**Example 5**

Consider:



(summation implicit) Let’s use GF again. Assume a solution of the form,



Plugging this into the equation we have:



From which it appears that (intepreting as matrices now)



I think we can eliminate the w, because it’s got whatever d.o.f., and then the B by multiplying both sides by its inverse.



So solution is:



**Example 6**

Suppose we wanted simply an expression for the evolution of the average of X, where dX/dt = a(X) + b(X)w. Starting from the probability distribution, we’d have:



and this agrees with the ODE since:



Note we can separate the averages in the last term since b(X(t)) depends on w(t) for all times up to but not including t. We have a similar relationship proceeding from the integral expression:



We can neglect ∫b(X)dW because these are independent random variables in a sense. What about the variance?



And from direct differentiation:



Taking the expectation of both sides we have:



And this matches!

**Series solution of SDE’s**

We’re interested in a Taylor series expansion of a stochastic variable since we often cannot solve the differential equation, or for the probability distribution function, discussed below. Let’s start with a deterministic example first. Consider that we can write (note this can’t be read as simply recursively substituting X(t) into D[ ], ‘cause it doesn’t work out; rather I’m just taking D(X(t)) and using the same kind of identity I used on X(t), that it’s equal to its initial value plus the integral from 0 to t):



This is a kind of nice derivation of the Taylor series expansion + remainder term. So then, suppose we have the stochastic equation:



We start by writing wj(t)dt = dWj(t). And then, we make manipulations similar to what was done before. I’ll stop at ‘second’ order.



(noting ∫dW = W) Next we work out these derivatives:



and,



Then we plug these into the expansion. And we’d do the f(x,t) = f(x(0),0) + ∫df/dt∙dt thing again. Suppose we got the following coupled differential equations. A series expansion of these guys would be a lot easier, since the differential operator is *linear*. We can do it like is done for the propagator in QM.



Integrating:



Filling the result back into the equations:



Simplifying, and keeping only up to order z,



Now suppose, just suppose, we have the following ODE’s (I’m going to make the dots implicit):



Let’s integrate



and then we would need to develop an expansion for the integrands. These are just the u’s themselves, and we already have the expansion right in front of us. So we plug these back into the integrals.



keeping only terms up to order z, keeping in mind that u+ ~ 1, and u- ~ √z



Filling in u+(0) = 1, then we have:



Working this out,



Doing same for u-, we have after filling the u’s back in:



keeping only up to ~ z, and filling in u+(z) ~ 1 as appropriate:



Perhaps we could afford to be less meticulous if we used the Stratonovich version of the differential?