STOCHASTIC CALCULUS
THEORY AND FORMALISMS

JOHN THICKSTUN

1. Overview

This is a brief tutorial on how to do informal stochastic calculus and why it works. I state all the theoretical results needed to compute, I properly compute solutions to two famous SDEs, and I show you how to get the same answers by manipulating Leibniz notation. Then I shows how the Leibniz notation relates to all those theoretical results (there is a neat sort of "lexical isomorphism"). Finally, I show how the Leibniz notation is really helpful–more helpful than the theory–when it comes to solving problems. In particular, it can help you figure out what function to plug in to Ito's formula.

2. Important Theorems

We begin by quoting several rigorous results from Protter. They are intended for reference in later computations and can be skimmed or skipped entirely. The first two establish key linearity properties of the stochastic integral. The third is the celebrated Bichteler-Dellacherie Theorem, which in conjunction with linearity and the fourth theorem allows us to simplify stochastic integrals by splitting them into separate classically integrable and martingale parts. The utility of the fifth and sixth results may be less obvious, but as we will see they are central to computation and closely tied to formal operations on stochastic differentials. The seventh result is Ito's lemma, which allows us to change variables (quoted here only for continuous semimartingales). The rest of the results have to do with quadratic variations. The eighth is Levy characterization. The ninth and tenth and eleventh will be important in computation.

Theorem 2.1. *Let X be a semimartingale and let $H, J \in L(X)$. Then $\alpha H + \beta J \in L(X)$ and $(\alpha H + \beta J) \cdot X = \alpha H \cdot X + \beta J \cdot X$. That is, $L(X)$ is a linear space.*

Theorem 2.2. *Let X, Y be semimartingales and suppose $H \in L(X)$ and $H \in L(Y)$. Then $H \in L(X + Y)$ and $H \cdot (X + Y) = H \cdot X + H \cdot Y$.*

Theorem 2.3. *A process X is a semimartingale if and only if $X = M + A$ where M is a local martingale and A is an FV process.*

Theorem 2.4. *Let X be a semimartingale with paths of finite variation on compacts. Let $H \in L(X)$ be such that the Stieltjes integral $\int_0^t |H_s||dX_s|$ exists a.s., each $t \geq 0$. Then the stochastic integral $H \cdot X$ agrees with a path-by-path Stieltjes integral.*

Theorem 2.5. *Let X be a semimartingale with $K \in L(X)$. Then $H \in L(K \cdot X)$ if and only if $HK \in L(X)$, in which case $H \cdot (K \cdot X) = (HK) \cdot X$.*

Theorem 2.6. *Let X, Y be semimartingales and let $H \in L(X)$, $K \in L(Y)$. Then*

$$[H \cdot X, K \cdot Y]_t = \int_0^t H_s K_s d[X,Y]_s.$$

Theorem 2.7. *Let X be a continuous n-dimensional semimartingale and let $f : \mathbb{R}^n \to \mathbb{R}$ have continuous second order partial derivatives. Then $f(X)$ is a semimartingale and the following formula holds:*

$$f(X_t) - f(X_0) = \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial X_i}(X_s) dX_s^i + \frac{1}{2} \sum_{1 \le i,j \le n} \int_0^t \frac{\partial^2 f}{\partial X_i \partial X_j}(X_s) d[X^i, X^j]_s.$$

Theorem 2.8. *A stochastic process X is a standard Brownian motion if and only if it is a continuous local martingale with $[X,X]_t = t$.*

Theorem 2.9. *Let X and Y be two semimartingales. Then $\Delta[X,Y] = \Delta X \Delta Y$.*

Theorem 2.10. *Let X be a quadratic pure jump semimartingale. Then for any semimartingale Y we have*

$$[X,Y]_t = X_0 Y_0 + \sum_{0 < s \le t} \Delta X_s \Delta Y_s.$$

Theorem 2.11. *If X is adapted, cadlag, with paths of finite variation on compacts, then X is a quadratic pure jump semimartingale.*

3. The Setup

Let X be a semimartingale and $H \in L(X)$. Then by theorems 3 and 2 respectively,

$$\int_0^t H_t dX_t = \int_0^t H_t d(M_t + A_t) = \int_0^t H_t dM_t + \int_0^t H_t dA_t.$$

Observe that by theorem 4, the latter integral is simply a classical Stieltjes integral. Of course, if A is absolutely continuous in t then we have the classical result

$$\int_0^t H_t dA_t = \int_0^t H_t A_t' dt = \int_0^t \mu(H_t, t) dt.$$

It now seems natural to consider the class of integral equations

$$H_t = \int_0^t \mu(H_t, t) dt + \int_0^t \sigma(H_t, t) dM_t.$$

Such equations are more commonly expressed in differential form. We can transform the above equation into a (formal) differential equation by taking a formal derivative:

$$dH_t = \mu(H_t, t) dt + \sigma(H_t, t) dM_t.$$

Martingales are not very differentiable and so it is theoretically difficult to make sense of this expression except as a notational shorthand for the integral version. Brownian motion in particular is nowhere differentiable. The following theorem gives some insight into the difficulty of talking about martingale differentials.

Theorem 3.1. *Let M be a local martingale with continuous paths. If $[M, M]_t = 0$ for all t then $M_t = 0$ for all t.*

Because martingales don't drift, their variation is of the jittery, non-smooth, quadratic type. Nevertheless, stochastic differential equations are intuitively evocative formulas. The equation under consideration tells us that H evolves locally by a drift determined by μ and a random martingale diffusion determined by σ.

4. Examples

Let's now consider some specific stochastic differential equations. A particularly famous example is geometric brownian motion:

$$dH_t = \mu H_t dt + \sigma H_t dB_t$$

The dynamics of this equation are clear: it drifts and diffuses proportional to its current value. We will proceed to solve the equation in two ways. First, we will convert it to an integral equation and derive our solution using the theorems described at the top of this paper. Second, we will operate directly with the differential equation using Leibniz-style formal manipulations. Following after the theoretically justified approach, it will become clear that these formal manipulations are actually just a veneer on the deeper theory.

Proposition 4.1. *If H is geometric brownian motion then*

$$H_t = H_0 \exp\big((\mu - \sigma^2/2)t + \sigma B_t\big).$$

Proof. We begin by converting the differential characterization of H to integrals:

$$H_t = H_0 + \mu \int_0^t H_t dt + \sigma \int_0^t H_t dB_t.$$

The first integral above is newtonian (theorem 4) so we will focus our attention on the more difficult brownian term. By Ito's lemma (theorem 7) using $f(x) = \log(x)$ we have

$$\log(H_t/H_0) = \int_0^t \frac{dH_t}{H_t} - \frac{1}{2}\int_0^t \frac{d[H, H]_t}{H_t^2}.$$

And by the integral dynamics of H_t, linearity (theorem 2) and associativity (theorem 5) respectively

$$\int_0^t \frac{dH_t}{H_t} = \int_0^t \frac{1}{H_t} d\left(\mu \int_0^t H_t dt + \sigma \int_0^t H_t dB_t\right)$$

$$= \int_0^t \frac{1}{H_t} d\left(\mu \int_0^t H_t dt\right) + \int_0^t \frac{1}{H_t} d\left(\sigma \int_0^t H_t dB_t\right) = \mu \int_0^t dt + \sigma \int_0^t dB_t = \mu t + \sigma B_t.$$

By theorem 6, the above calculation, bilinearity of quadratic variation, and theorem 8 respectively

$$\int_0^t \frac{d[H,H]_t}{H_t^2} = \left[\int_0^t \frac{dH_t}{H_t}, \int_0^t \frac{dH_t}{H_t}\right]$$

$$= [\mu t, \mu t] + 2[\mu t, \sigma B_t] + [\sigma B_t, \sigma B_t] = \mu^2[t,t] + 2[\mu t, \sigma B_t] + \sigma^2[B_t, B_t].$$

Because t is FV, combining theorems 11 and 10 we have $[t,t] = 0$. And combining theorems 11 and 9 gives us $[\mu t, \sigma B_t] = 0$. By these results and Levy characterization (theorem 9) the above reduces to

$$\sigma^2[B_t, B_t] = \sigma^2 t.$$

We therefore have

$$\log(H_t/H_0) = \mu t + \sigma B_t - \frac{\sigma^2}{2}t = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t.$$

And exponentiating gives us our result. □

Let us now repeat this calculation, making indiscriminate use of formal manipulations.

Proof. We recall that if H is geometric brownian motion then

$$dH_t = \mu H_t dt + \sigma H_t dB_t.$$

First, let's write down a formal differential analog to Ito's lemma (theorem 7):

$$df = \frac{\partial f}{\partial H_t} dH_t + \frac{1}{2}\frac{\partial^2 f}{\partial H_t^2} d[H_t, H_t].$$

Using $f(x) = \log(x)$ and replacing the quadratic variation term with a formal "square" this becomes

$$d\log(H_t) = \frac{dH_t}{H_t} - \frac{dH_t^2}{2H_t^2}.$$

Substituting the differential dynamics of H_t we have

$$\frac{dH_t}{H_t} = \frac{\mu H_t dt + \sigma H_t dB_t}{H_t} = \mu dt + \sigma B_t.$$

Turning to the second term, we must now deal with the differential "square." Substituting in our sde,

$$dH_t^2 = (\mu H_t dt + \sigma H_t dB_t)^2 = \mu^2 H_t^2 dt^2 + 2\mu\sigma H_t^2 dt dB + \sigma^2 H_t^2 dB_t^2.$$

We will use the heuristics $dt^2 = 0$, $dtdB = 0$ and $dB = \sqrt{dt}$, which give us

$$dH_t^2 = \sigma^2 H_t^2 \sqrt{dt}^2 = \sigma^2 H_t^2 dt.$$

It follows that

$$\frac{dH_t^2}{H_t^2} = \sigma^2 t.$$

Putting this all together,

$$d\log(H_t) = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dB_t.$$

And formally integrating this expression gives us our result:

$$\log(H_t) - \log(H_0) = \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma B_t$$

□

Let's take a look at a second derivation before stepping back to abstract what we've learned here. This next result is an example of an Ornstein-Uhlenbeck process. As before we will first give a theoretically justified proof.

Proposition 4.2. *Suppose H has differential dynamics $dH_t = -\alpha H_t dt + \sigma dB_t$. Then*

$$H_t = e^{-\alpha t}\left(H_0 + \sigma \int_0^t e^{\alpha s} dB_s\right).$$

Proof. By formal integration, we can rewrite the dynamics of H as

$$H_t = -\alpha \int_0^t H_t dt + \sigma \int_0^t dB_s.$$

By Ito's lemma, using $f(x,t) = xe^{at}$ and suppressing terms with vanishing quadratic variation,

$$H_t e^{at} = H_0 + \int_0^t e^{as} dH_s + \alpha \int_0^t H_s e^{as} ds.$$

Focusing on the second term, from the integral dynamics of H_t, linearity (theorem 2) and associativity (theorem 5) respectively,

$$\int_0^t e^{as} dH_t = \int_0^t e^{as} d\left(-\alpha \int_0^t H_s ds + \sigma \int_0^t dB_s\right)$$

$$= -\alpha \int_0^t H_s e^{as} dt + \sigma \int_0^t e^{as} dB_s$$

And substituting this back in to our earlier expression gives us

$$H_t e^{\alpha t} = H_0 + \sigma \int_0^t e^{\alpha s} dB_s.$$

□

Before giving the informal derivation, let's pause and make a couple observations. First, this computation proceeded almost identically to the computation for geometric brownian motion. Where it varied, it actually went a little easier because all the quadratic variation terms dropped out. We will now proceed with the differential version.

Proof. We recall that the differential dynamics of H are given by

$$dH_t = -\alpha H_t dt + \sigma dB_t.$$

By Ito's lemma we have

$$df = \frac{\partial f}{\partial H_t} dH_t + \frac{\partial f}{\partial t} dt + \frac{1}{2}\frac{\partial^2 f}{\partial H_t^2} dH_t^2.$$

And using $f(x, t) = xe^{\alpha t}$ gives us

$$df(H_t, t) = e^{\alpha t} dH_t + \alpha H_t e^{\alpha t} dt.$$

Substituting in the differential dynamics for H_t,

$$df(H_t, t) = e^{\alpha t}(-\alpha H_t dt + \sigma dB_t) + \alpha H_t e^{\alpha t} dt = \sigma e^{\alpha t} dB_t.$$

And by formal integration we may conclude that

$$H_t e^{\alpha t} = H_0 + \sigma \int_0^t e^{\alpha s} dB_s.$$

□

5. The Informal Formulas

Let's now step back and talk about what we've done. I'm going to go through and, tongue in cheek, reformulate the theorems given at the start of this paper with corresponding propositions about differentials. Perhaps we can call them formal theorems, or maybe just formulas.

Formula 5.1.

$$\int (\alpha H + \beta J) dX = \alpha \int H dX + \beta \int J dX.$$

Formula 5.2.

$$\int H d(X + Y) = \int H dX + \int H dY.$$

Formula 5.3.

$$X = M + A.$$

Formula 5.4. *If* $X = A$ *then*

$$\int X = \int A.$$

Formula 5.5.

$$\int X d\left(\int Y dZ\right) = \int XY dZ.$$

Formula 5.6.

$$dH_t^2 = (dH_t)^2$$

Formula 5.7.

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial X_i} dX_i + \frac{1}{2} \sum_{1 \leq i,j \leq n} \frac{\partial^2 f}{\partial X_i \partial X_j} dX_i X_j.$$

We can encapsulate the final four theorems in a single heuristic about newtonian and brownian differentials.

Formula 5.8.

$$dt^2 = dtdB = dBdt = 0; \quad dB^2 = \sqrt{dt}.$$

The most difficult step in computing stochastic integrals is, perhaps, finding the appropriate function f with which to invoke Ito's lemma. The differential formalism can be very helpful for us here. So far, we have mimicked our integral derivations when working with differentials. We did this deliberately, because it allowed us to see the connections between the theory and the formalisms. But now let us break from theory and see how useful the formalisms can be in helping us compute.

6. Conclusion

Returning to geometric brownian motion, we recall that the single-variable Ito formula is given by

$$df = \frac{\partial f}{\partial H_t} dH_t + \frac{1}{2} \frac{\partial^2 f}{\partial H_t^2} dH_t^2.$$

In our earlier derivation we immediately chose $f(x) = \log(x)$, pulling a function out of thin air that magically resolved the equations. Now we will pretend ignorance, and proceed as if we did not know the correct choice of f. Instead, we will immediately substitute the differential dynamics of H into this formula:

$$df = \frac{\partial f}{\partial H_t} (\mu H_t dt + \sigma H_t dB_t) + \frac{1}{2} \frac{\partial^2 f}{\partial H_t^2} dH_t^2.$$

We can also resolve the differential square to $\sigma^2 H_t^2 dt$. Simplifying algebraically yields

$$\left(\mu H_t \frac{\partial f}{\partial H_t} + \frac{\sigma^2}{2} H_t^2 \frac{\partial^2 f}{\partial H_t^2} \right) dt + \sigma H_t \frac{\partial f}{\partial H_t} dB_t.$$

And now it is clear that substituting $f(x) = \log(x)$ makes all the complications disappear.

Now suppose H follows the Ornstein-Uhlenbeck diffusion from above. Recall that Ito's lemma tells us

$$df = \frac{\partial f}{\partial H_t} dH_t + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial H_t^2} dH_t^2.$$

Substituting the differential dynamics gives us

$$\begin{aligned} df &= \frac{\partial f}{\partial H_t} (-\alpha H_t dt + \sigma dB_t) + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial H_t^2} dH_t^2 \\ &= \left(-\alpha H_t \frac{\partial f}{\partial H_t} + \frac{\partial f}{\partial t} \right) dt + \sigma \frac{\partial f}{\partial H_t} dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial H_t^2} dH_t^2. \end{aligned}$$

From here, it is clear that we would like $\partial^2 f/\partial H_t^2$ to vanish and

$$\frac{\partial f}{\partial t} = \alpha H_t \frac{\partial f}{\partial H_t}.$$

The first condition requires that $f(x,t) = xg(t)$ and substituting this into the second condition gives us

$$H_t g'(t) = \alpha H_t g(t).$$

And now we see that it is very natural to choose $f(x,t) = xe^{\alpha t}$. Of course it still requires a little vision to see these problems through and get around the difficulties, but the differential formalism gives us a rich algebraic playground in which we can experiment and tinker with the solution.