Stratonovich's Theory

18.676 Spring 2024

Rowechen Zhong

Spring 2024

In this paper, we argue on an complete filtered probability space $\left(\Omega, (\mathscr{F}_t)_{t \geq 0}, \mathbb{P}\right)$. As usual, we identify $\mathbb{P}$-indistinguishable processes. We define finite variation processes (FVPs), continuous local martingales (CLMs), and continuous semimartingales (CSMs) as in [2]. All processes are forever $(\mathscr{F}_t)$-measurable and continuous. For any locally bounded progress process (LBPP) H, we define the stochastic integral of H with respect to a CSM X, $H \cdot X$, as in [2].

ℭ1 Stratonovich Integrals

The purpose of this paper is to discuss the various properties and applications of the Stratonovich integral.

> **Definition 1** (Stratonovich Integral) — Given two continuous semimartingales with decompositions $X = M_X + A_X$ and $Y = M_Y + A_Y$, and a C^1 function f, we define the **Stratonovich integral** of $f(Y)$ with respect to X as
>
> $$(f(Y) \circ X)_t \equiv \int_0^t f(Y_s) \circ dX_s = \int_0^t f(Y_s) dX_s + \frac{1}{2} \int_0^t f'(Y_s) d\langle Y, X \rangle_s.$$
>
> This has CLM and FVP components
>
> $$\int_0^t f(Y_s) dA_X + \frac{1}{2} \int_0^t f'(Y_s) d\langle Y, X \rangle_s, \qquad \int_0^t f(Y_s) dM_X.$$
>
> We make the obvious generalization when X takes values in $\mathbb{R}^d$, $f \in C^1\left(\mathbb{R}^d; \mathbb{R}^n\right)$, and Y takes values in $\mathbb{R}^n$.

CSMs are not closed under C^1 functions, hence this definition is not redundant. We immediately observe the possibly surprising fact that the Stratonovich integral with respect to a continuous local martingale is not necessarily itself a CLM. This distinction is important, and we will return to it often. We also observe that compared to the Ito integral, our integrand must satisfy more regularity conditions than a generic locally bounded progressive process. This is less of a concern, as this class is more than enough to discuss practical cases of interest (solutions to stochastic differential equations, etc.).

Throughout the first half of this paper, we will visit ways of obtaining Stratonovich objects as the limit of pathwise-ordinary objects. This means that for any fixed ω, the object in question is a bona fide deterministic integral, or differential equation, or physical process.

ℭ2 Integrals obtained as limits of meshes

Suppose we are given continuous semimartingales X and Y, and make a naive attempt to define the stochastic integral "$\int Y\,dX$" a la Riemann. For any series of partitions Δ of $[0, t]$ whose mesh $|\Delta|$ goes to 0, the limit

$$\lim_{|\Delta|\to 0} \sum_{i=1}^{n} Y_{t_i} \left(X_{t_{i+1}} - X_{t_i} \right)$$

exists in probability, and recovers the familiar Ito integral $\int_0^t Y_s dX_s$. However, with some experience we realize that we were perhaps lucky in this naive definition, because it is not clear a priori that we should evaluate Y at t_i instead of t_{i+1}, and such a change makes a significant difference.

Given CSMs X, Y, $f \in C^1$, a mesh Δ, and a probability measure μ on $[0, 1]$, we define the following two classes of pathwise-Riemann approximations to the stochastic processes:

$$P_\Delta^\mu = \sum_i \left(Y_{t_{i+1}} - Y_{t_i} \right) \int_0^1 f \left(X_{t_i} + s \left(X_{t_{i+1}} - X_{t_i} \right) \right) \mu(ds)$$

and

$$Q_\Delta^\mu = \sum_i \left(Y_{t_{i+1}} - Y_{t_i} \right) \int_0^1 f \left(X_{t_{i+1}+s\left(X_{t_{i+1}} - X_{t_i}\right)} \right) \mu(ds).$$

> **Claim** (Integrals as the Limits of Meshes) — Let $\bar\mu = \int_0^1 s\,d\mu(s)$. In fact,
>
> - In probability,
>
> $$\lim_{|\Delta|\to 0} P_\Delta^\mu = \int_0^t f(X_s)dY_s + \bar\mu \int_0^t f'(X_s)d\langle X, Y\rangle_s.$$
>
> - If $d\langle X, Y\rangle$ is absolutely continuous with respect to Lebesgue measure, then Q_Δ^μ obtains the same limit in probability.

The proof of this point-set result is not interesting: proofs can be found in Strook[4] and Revuz[3].

Thus, large families of such mesh-like approximations to stochastic-integral like objects are in fact of the same form as the Ito and Stratonovich integrals. Notably, the arguably more natural *uniform measure* has mean $\bar\mu = \frac{1}{2}$, and yields the Stratonovich integral, not the Ito integral!

We also thus see that the Ito integral is *unique* in that integrals with respect to CLMs remain CLMs. Heuristically, the reason is that for all μ except δ_0, the integrand has the privilege of "looking into the future" to coordinate its movement with the change $X_{t_{i+1}} - X_{t_i}$.

ℭ2.1 Coordinate Invariance

Recall the celebrated Itô's formula, which is a sort of "chain rule" in the context of stochastic integration. We immediately see that when written with the Stratonovich integral, the formula appears exactly as in the deterministic case.

> **Theorem 1** (Itô's Formula) — Let X be a continuous semimartingale on $\mathbb{R}^n$ and $F \in C^2(\mathbb{R}^n; \mathbb{R})$. We use Einstein summation notation, such that indices are implicitly summed over. Then, for all $t \geq 0$,
>
> $$F(Z(t)) - F(Z(0)) = \int_0^t \partial_i F(Z_s) dZ_s^i + \frac{1}{2} \int_0^t \partial_i \partial_j F(Z_s) d\langle Z^i, Z^j \rangle_s$$
> $$= \int_0^t \partial_i F(Z_s) \circ dZ_s^i.$$

The generalizations to higher dimensions are immediate. Stratonovich's chain rule solves stochastic differential equations faster than Ito's. Here's an illustrative example:

> **Problem 2.1** (Exponential Martingale) — Let X, A be $\text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$-valued continuous semimartingales such that
>
> $$dX_{ij}(t) = X_{ik}(t) dA_{kj}(t)$$
>
> as an *Ito* SDE. Obtain a closed form for X as a function of A and quadratic variations thereof.

Proof. This differential equation is trivial to solve in the ODE case; "$X_t = e^{A_t}$". If the differential equation read

$$dX_{ij}(t) = X_{ik}(t) \circ dY_{kj}(t) \tag{$\star$}$$

we would immediately obtain the same conclusion – $de^{Y_t} = e^{Y_t} \circ dY_t$.

Thus, an efficient way to solve this Ito equation is by transforming the desired into a Stratonovich integral! Consider the substitutions

$$Y_{ij} = A_{ij} - \frac{1}{2} \langle A_{ik}, A_{kj} \rangle \implies \langle Y_{ij}, Y_{k\ell} \rangle = \langle A_{ij}, A_{k\ell} \rangle.$$

We immediately obtain $\star$.

$$X(t) = e^{A(t) - \frac{1}{2} \langle A(t), A(t) \rangle}.$$

(The bracket here is the matrix $\langle A_{ij}(t), A_{jk}(t) \rangle$). $\square$

ℭ2.2 Mollifying the Integrator

Let W be a Brownian motion, and let $\psi(\eta, t)$ have continuous partial derivatives on $\mathbb{R} \times [a, b]$. This time, wish to approximate an integral of the form

$$\int_a^b \psi(W_t, t) \circ dW_t$$

via a limit of **finite variation processes** which approximate W. Recall that finite variation processes are significantly smoother than continuous semimartingales! Specifically, we construct FVPs W_t^n such that for almost all ω,

- $W_t^n(\omega) \to W_t(\omega)$ for all $t \in [a, b]$.

- There exist finite real numbers $n_0(\omega)$ and $k(\omega)$ such that for all $n > n_0$ and all $t \in [a, b]$, $|W_t^n(\omega)| \leq k(\omega)$. If ψ is independent of t, then this condition can be dropped.

3

> **Theorem 2** (Baby Wong-Zakai Theorem) — Almost surely,
>
> $$\lim_{n \to \infty} \int_a^b \psi(W_t^n, t) \circ dW_t^n = \int_a^b \psi(W_t, t) \circ dW_t.$$

It is heuristically clear that the only possible stochastic limit of the (usual, Lebesgue) integrals with respect to W_t^n is the Stratonovich integral, because the left-hand-side obeys the chain rule!

Proof. As a reminder, this Stratonovich integral is

$$\int_a^b \psi(W_t, t) dW_t + \frac{1}{2} \langle \psi(W_t, t), W_t \rangle_a^b = \int_a^b \psi(W_t, t) dW_t + \frac{1}{2} \int_a^b \frac{\partial \psi}{\partial \eta}(W_t, t) dt$$

To prove this, we allow $F(\lambda, t) = \int_a^\lambda \psi(\eta, t) d\eta$. This is a C^2 function in λ which we can hopefully apply Ito's formula to. Then,

$$F(W_b^n, b) - F(W_a^n, a) = \int_a^b \psi(W_t^n, t) \, dW_t^n + \int_a^b \left(\frac{\partial F}{\partial t} \right)(W_t^n, t) dt.$$

This is simply true for almost all ω, by the Fundamental theorem of calculus, as (W_t^n, t) is continuous.

Now, the main reason we did all of this was because it is very hard to compute $\lim_{n \to \infty} \int_a^b \psi(W_t^n, t) \, dW_t^n$ in any sense. Your finite variation process is supposed to converge to a CLM, and neither Portmanteau theorem nor Dominated convergence theorem let you compute this limit in your integrand and integrator simultaneously. The above ω-pointwise equality lets us rewrite this integral in terms of three very smooth well-understood objects.

First off, $F(W_b^n, b) - F(W_a^n, a) \to F(W_b, b) - F(W_a, a)$ everywhere, because $W_n \to W$ and F is continuous.

> **Claim** — For almost all ω,
>
> $$\int_a^b \left(\frac{\partial F}{\partial t} \right)(W_t^n, t) dt \to \int_a^b \left(\frac{\partial F}{\partial t} \right)(W_t, t) dt. \qquad (\star)$$

Proof. F is C^1, thus

$$\frac{\partial F}{\partial t}(\lambda, t) = \int_a^\lambda \frac{\partial \psi(\eta, t)}{\partial t} d\eta$$

Note that if ψ is independent of t, we immediately conclude.

So, fix an ω. Observe that $\frac{\partial \psi(\eta, s)}{\partial s}$ is continuous everywhere, and thus bounded by say K on $[-k(\omega), k(\omega)] \times [a, b]$.

Fix a $a \le t \le b$. We'll show that $\left(\frac{\partial F}{\partial t} \right)(W_t^n, t)$ converges. Look at

$$\int_a^{W_t^n} \frac{\partial \psi(\eta, t)}{\partial t} d\eta = \int_a^{k(\omega)} \mathbb{1}_{\eta \le W_t^n} \frac{\partial \psi(\eta, t)}{\partial t} d\eta$$

These converge to $\left(\frac{\partial F}{\partial t} \right)(W_t, t)$. This is true for each fixed t. Hence, because $\frac{\partial F}{\partial t}(W_t^n, t)$ and $\frac{\partial F}{\partial t}(W_t, t)$ are each continuous functions, they converge pointwise on $[a, b]$ almost surely. Now, for each t, $\frac{\partial F}{\partial t}(W_t^n, t)$ is dominated by

$$\int_a^{k(\omega)} K d\eta = (k(\omega) - a)K < \infty.$$

Hence we conclude. □

Now, by Ito's formula, we know that $F(W_b^n, b) - F(W_a^n, a)$ converges to

$$\int_a^b \psi(W_t, t)dW_t + \int_a^b \left(\frac{\partial F}{\partial t}\right)(W_t, t)dt + \frac{1}{2}\int_a^b \frac{\partial \psi}{\partial \eta}(W_t, t)dt$$

Plugging in $\star$, we conclude.

$\square$

This proof holds when W is replaced by any CSM, so long as the approximating processes are FVPs.

☙3 Wong-Zakai Theorem

Going forward, our intuition is that the ω-pointwise limits of smooth integrals must be Stratonovich. This continues to hold when we turn to look at stochastic differential equations.

Specifically, let us consider the stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t) \circ dW_t, \qquad X_0 = x_0$$

where b is a K-Lipschitz $\mathbb{R}^d$-valued function, σ is a K-Lipschitz $\text{Hom}(\mathbb{R}^d, \mathbb{R}^m)$-valued function, and X takes on values in $\mathbb{R}^d$. Suppose we have a series of continuous piecewise-differentiable(so as to automatically admit Lebesgue integrals) processes W_t^n which approximate Brownian motion; there exists an m-dimensional Wiener process W_t such that

$$\sup_{t \in [0,T]} \|W_t - W_t^n\| \to 0$$

almost surely. Then, we might hope that the usual solution to the differential equation, $\frac{d}{dt}X_t^n = b(X_t^n) + \sigma(X_t^n)\xi_t^n$, where $\xi_t^n = \frac{dW_n^t}{dt}$, coincides with the SDE solution. This is in fact true. We present two versions of this theorem, although there are generalizations with significantly less constraints that are also significantly harder to prove.

> **Theorem 3** (Wong-Zakai with External Noise) — Suppose $\sigma(x) = \sigma$ is constant, which would occur if such a system was driven by an external noise source. Then, $\sup_{t \in [0,T]} \|X_t^n - X_t\| \to 0$ almost surely.

Proof. Indeed,

$$X_t^n - X_t = \int_0^t \left(b(X_s^n) - b(X_s)\right)ds + \sigma(W_t^n - W_t)$$

thus by triangle inequality and Lipschitz,

$$\|X_t^n - X_t\| \le K \int_0^t \|X_s^n - X_s\| \, ds + \|\sigma\| \sup_{t \in [0,T]} (W_t^n - W_t).$$

The supremum goes to 0 almost surely. Then, Gronwall's Lemma concludes immediately.

$\square$

> **Theorem 4** (Wong-Zakai in 1 dimension) — Otherwise, let us restrict ourselves with the following regularity conditions:
>
> 1. X, W, X^n, ξ^n are $\mathbb{R}^1$-valued.

2. b, σ are K-Lipschitz and bounded.

3. σ is C^1 and $\sigma(x)\frac{d\sigma(x)}{dx}$ is Lipschitz continuous.

4. There is some $\beta > 0$ such that $\sigma(x) \geq \beta > 0$ for all x.

Then, $\sup_{t\in[0,T]} |X_t^n - X_t| \to 0$ almost surely.

Proof. We consider the function $\Phi(x) = \int_0^x (\sigma(y))^{-1} \, dy$, which is well-defined and C^2, because $\sigma(y)$ is $C(1)$ and $\sigma(y) \geq \beta > 0$. Then,

$$\frac{d}{dt}\Phi\left(X_t^n\right) = (\sigma(X_t^n))^{-1}\left(b(X_t^n) + \sigma(X_t^n)\xi_t^n\right) = \frac{b(X_t^n)}{\sigma(X_t^n)} + \xi_t^n.$$

Meanwhile,

$$d\Phi(X_t) = \sigma^{-1}(X_t)\left\{b(X_t)dt + \sigma(X_t)dW_t + \frac{1}{2}\sigma'(X_t)\sigma(X_t)dt\right\} - \frac{1}{2}\frac{\sigma'}{\sigma^2}\sigma(X_t)^2 dt$$

$$= \sigma^{-1}\left(X_t\right)\left(b\left(X_t\right)dt + \sigma(X_t)dW_t\right)$$

$$= \frac{b(X_t)}{\sigma(X_t)}dt + dW$$

because the Stratonovich integral obeys the chain rule. We now want to smash with Gronwall.

$$\|\Phi\left(X_t^n\right) - \Phi(X_t)\| \leq \int_0^t \left|\frac{b(X_s^n)}{\sigma(X_s^n)} - \frac{b(X_s)}{\sigma(X_s)}\right| ds + \sup_{t\in[0,T]}|W_t^n - W_t|.$$

Now, because σ is bounded by L, σ^{-1} is bounded by $\frac{1}{L}$ from below, hence $|\Phi(x) - \Phi(z)| \geq \frac{1}{L}|x - z|$. Meanwhile,

$$\left|\frac{b(x)}{\sigma(x)} - \frac{b(z)}{\sigma(z)}\right| \leq \frac{1}{\sigma(x)}|b(x) - b(z)| + \frac{b(z)}{\sigma(x)\sigma(z)}|\sigma(x) - \sigma(z)| \leq \frac{K}{\beta}|x - z| + \frac{LK}{\beta^2}|x - z|.$$

Gronwall kills now. $\square$

These limits of stochastic differential equations provides intuition on how one ought to construct and interpret models of physical processes with stochastic noise. We provide a pedagogical exposition of the *Langevin equation* in an appendix. It has been omitted from the main text, because some results lack the rigor expected in the mathematical community.

☙4 Stochastic Integration on Manifolds

Away from the theory of probability, when discussing theories of differential geometry and algebra, a fundamental consideration is *invariance under natural transformations*. In linear algebra, this places an emphasis on the *eigenvalues and eigenvectors* of a linear map over its explicit presentation as a matrix; in the study of manifolds, this places an emphasis on the C^k/*smooth/analytic/holomorphic structure* of the manifold as opposed to any particular atlas. In our specific case, we will often talk about *invariance under coordinate transformations,* and this is a special case of the general principle.

From this perspective, the Stratonovich integral is much more natural than the Ito integral. The manifest coordinate invariance of this Stratonovich formulation allows us to easily define the most basic objects in stochastic calculus on manifolds.

Our main reference for this section is [1].

☾4.1 Elementary Definitions

Let M be a smooth manifold.

> **Definition 2** (M-valued Semimartingale) — Given a stopping time τ, an M-valued process X defined on $[0, \tau)$ is an M-valued semimartingale if, for all $f \in C^\infty(M)$, $f(X)$ is a real-valued semimartingale.

By Ito's formula, it is clear that this reduces to the usual definition of a semimartingale when $M = \mathbb{R}^d$.

Suppose $V_\alpha : \alpha = 1, \ldots, \ell$ are smooth vector fields on $\mathbb{R}^N$, and consider the Stratonovich SDE

$$X_t = X_0 + \int_0^t V_\alpha(X_s) \circ dZ_s^\alpha.$$

The equivalent Ito SDE is

$$X_t = X_0 + \int_0^t V_\alpha(X_s) dZ_s^\alpha + \frac{1}{2} \int_0^t \nabla_{V_\beta} V_\alpha(X_s) d \langle Z^\alpha, Z^\beta \rangle_s$$

where $\nabla_{V_\beta} V_\alpha$ is the derivative of V_α in the direction of V_β.

Then, if X is a solution to such an SDE and $f \in C^2(\mathbb{R}^d)$, Ito's formula immediately yields that

$$f(X_t) = f(X_0) + \int_0^t V_\alpha f(X_s) \circ dZ_s^\alpha, \qquad t < e(X).$$

Seeing this, we can immediately define solutions to stochastic differential equations on manifolds. Upgrade V_α to be smooth vector fields on M, and let Z be a CSM on $\mathbb{R}^\ell$. Then, we might consider an M-valued random variable $X_0 \in \mathscr{F}_0$, and attempt to define solutions to the SDE

$$dX_t = V_\alpha(X_t) \circ dZ_t^\alpha.$$

> **Definition 3** (Solution to SDE on Manifold) — A solution to the SDE $dX_t = V_\alpha(X_t) \circ dZ_t^\alpha$ is an M-valued semimartingale X and a stopping time τ such that, for all $f \in C^\infty(M)$,
>
> $$f(X_t) = f(X_0) + \int_0^t V_\alpha f(X_s) \circ dZ_s^\alpha, \qquad t < \tau.$$
>
> We say that X is a solution to the SDE $SDE(V_1, \ldots, V_\ell; Z, X_0)$.

This reduces to the usual definition of a solution to an SDE when $M = \mathbb{R}^d$, because notably we can pick the coordinate functions $f = x_i$.

The coordinate-invariance really shines through here:

> **Claim** (Diffeomorphisms preserve solutions to SDEs) — Let $\Phi : M \to N$ be a diffeomorphism between smooth manifolds, and let X be a solution to $SDE(V_1, \ldots, V_\ell; Z, X_0)$. Then, $\Phi(X)$ is a solution of the SDE $SDE(\Phi_* V_1, \ldots, \Phi_* V_\ell; Z, \Phi(X_0))$ on N.

Indeed, let f be any smooth function on N. Then, $f \circ \Phi$ is a smooth function on M, and we have

$$f(Y_t) = f(\Phi(X_0)) + \int_0^t V_\alpha(f \circ \Phi)(X_s) \circ dZ_s^\alpha$$

$$= f(\Phi(X_0)) + \int_0^t (\Phi_* V_\alpha)(f)(\Phi(X_s)) \circ dZ_s^\alpha$$

which is exactly the definition of a solution to the SDE on N.

❧4.2 Existence and Uniqueness of Solutions

It is a bit unsatisfying to end here, as we have not shown how to obtain any solutions to these SDEs. Thus, we present a quick and dirty (extrinsic) way to establish the existence and uniqueness of solutions to SDEs on manifolds.

> **Theorem 5** (Whitney's Embedding Theorem) — Any smooth n-dimensional manifold M can be smoothly embedded in $\mathbb{R}^{2n}$ as a closed submanifold. Let $N = 2n$.

Thus, without loss of generality, we identify M with its image in $\mathbb{R}^N$.

> **Idea** (Remark on Explosions) — Then, ∞_M is mapped to $\infty_{\mathbb{R}^N}$, in the sense that a sequence of points $x_n \in M$ will converge to ∞_M in $\hat{M}$ iff $|x_n| \to \infty$. Indeed, if $x_n \to \infty_M$, then for all R, x_n is eventually outside the compact subset $M \cap B_0(R)$, hence $|x_n| \to \infty$. Conversely, if $|x_n| \to \infty$, it is clear that $|x_n|$ do not converge in $\mathbb{R}^N$ (or M), hence by compactness of $\hat{M}$, $x_n \to \infty_M$.

By Whitney extension theorem, we extend each V_α to a smooth vector field $\tilde{V}_\alpha$ on $\mathbb{R}^n$. Then, we can give (disappointingly coordinate-laden) re-phrasings of semimartingales and solutions to SDEs.

> **Claim** (Semimartingales Characterized by Coordinate functions) — First, X is semimartingale on M iff X is a semimartingale on $\mathbb{R}^N$.
>   Also, consider some $SDE(V_1, \ldots, V_\ell; Z, X_0)$ on M. Then, X is a solution to $SDE(V_1, \ldots, V_\ell; Z, X_0)$ on M iff X is a solution to $SDE(\tilde{V}_1, \ldots, \tilde{V}_\ell; Z, X_0)$ on $\mathbb{R}^N$.

Proof. First, the only-if direction is trivial because each coordinate function is a smooth function. Conversely, any function $f \in C^\infty(M)$ can be extended to a function $\tilde{f} \in C^\infty(\mathbb{R}^N)$ (by the Whitney extension theorem), where the statement is trivial.

For the second claim, the only-if direction is trivial as each coordinate function is a smooth function. Conversely, for any $f \in C^\infty(M)$, we can extend f to a function $\tilde{f} \in C^\infty(\mathbb{R}^N)$, such that $f(X_t) = \tilde{f}(X_1^t, \ldots, X_N^t)$. By Ito's formula,

$$\begin{aligned}
d[f(X_t)] &= \partial_i \tilde{f}\left(X_1^t, X_2^t, \ldots, X_n^t\right) \circ d\left[X_i^t\right] \\
&= \partial_i \tilde{f}\left(X_1^t, X_2^t, \ldots, X_n^t\right) \circ V_\alpha^i \circ d\left[Z_i^\alpha\right] \\
&= \left(\partial_i \tilde{f}\left(X_1^t, X_2^t, \ldots, X_n^t\right) V_\alpha^i\right) \circ d\left[Z_i^\alpha\right] \\
&= V_\alpha f(X_t) \circ dZ_t^\alpha
\end{aligned}$$

$\square$

In order to use our well-understood formalism for SDEs in $\mathbb{R}^N$, we need to show that solutions to the extended equation

$$X_t = X_0 + \int_0^t \tilde{V}_\alpha(X_s) \circ dZ_s^\alpha$$

which begin at $X_0 \in M$ remain on M for all $0 \le t < e(X)$. This is the main non-trivial step in porting our SDE solutions on $\mathbb{R}^N$ to M.

> **Claim** — Let X be the solution of the extended equation up to its explosion time $e(X)$, with constant $X_0 \in M$. Then, $X_t \in M$ for all time.

We omit the details of this proof[1], which is a careful local calculation bounding the distance of X_t from M. Combining these two preceding results, we know that:

- Given any solution to SDE on M, we obtain a solution to SDE on $\mathbb{R}^N$.

- Given any solution to SDE on $\mathbb{R}^N$ which starts at $X_0 \in M$, we obtain a solution to SDE on $M \subset \mathbb{R}^N$ for all time, which is exactly a solution to SDE on M.

By well-known results (see appendix B), we know that if $\tilde{V}_\alpha$ are locally Lipschitz, then the SDE on $\mathbb{R}^N$ has a unique solution. Thus, we obtain the following result:

> **Theorem 6** — If $\{V_i\}$ have a smooth extension $\tilde{V}_i$ to $\mathbb{R}^N$ which is locally Lipschitz, then there exists a unique solution to $SDE(V_1, \ldots, V_\ell; Z, X_0)$ on M up to its explosion time.

✺5 Bibliography

[1] Elton P. Hsu. *Stochastic Analysis on Manifolds.* American Mathematical Society, 2002.

[2] Jean-François Le Gall. *Brownian motion, martingales, and stochastic calculus.* Springer, 2016.

[3] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion.* Grundlehren der mathematischen Wissenschaften. Springer Berlin, Heidelberg, 1999.

[4] DANIEL W. STROOCK. *Elements of stochastic calculus and analysis.* SPRINGER, 2019.

✺A Langevin Equation

Now, we will talk about the Langevin equation, which is a stochastic differential equation that describes a great variety of physical phenomena, to give a real-world example of how a stochastic differential equation arises as the heuristic limit of a sequence of ordinary differential equations. The purpose of this presentation is to elaborate upon the physical motivation for stochastic differential equations as the limit of a series of pathwise-ordinary differential equations.

> **Definition 4** (Langevin Equation) — In its simplest form, the **Langevin equation** is a stochastic differential equation in two variables X_t, P_t taking the form:
>
> $$dX_t = \frac{P_t}{M}dt, \qquad dP_t = -V'(X_t)dt - \frac{\gamma}{M}P_t dt + \sqrt{2\frac{\gamma}{\beta}}dW_t.$$

We will present a rigorous derivation of a different, "ordinary" equation in the context of the movement of a **colloid** (mesoscopic object) in a fluid. Then, we will discuss some well-known results in the physics literature. The only physics knowledge needed for this section is used in establishing the equations of motion, which are then solved as a purely mathematical problem.

☾A.1 Colloid Model

In physics, one usually posits a theory through a **Hamiltonian** or energy function. To this end, consider the following Hamiltonian describing a single colloid particle with mass M, connected to n particles of mass $\{m_i\}_{i=1}^N$ through springs with varying spring constants $\{\omega_i\}_{i=1}^N$:

$$H = \frac{P^2}{2M} + V(X) + \sum_i \frac{p_i^2}{2m_i} + \sum_i \frac{m_i \omega_i^2}{2}(X - q_i)^2$$

Here, X, q_i are the positions of the colloid and microscopic particles, and P, p_i are their momenta. X, q_i, P, p_i are a collection of $2n + 2$ functions from $\mathbb{R}_+$ (time) to vectors in $\mathbb{R}^d$ space. We will take $d = 1$ for simplicity (but generalizations are immediate). The potential V is some smooth positive function $\mathbb{R}^d \to \mathbb{R}$.

  Any physicist can immediately re-write this as a system of $2n + 2$ first-order differential equations:

$$\dot{q}_i = \frac{p_i}{m_i}, \quad \dot{p}_i = m_i \omega_i^2 (X - q_i), \quad \dot{X} = \frac{P}{M}, \quad \dot{P} = -V'(X) + \sum_i m_i \omega_i^2 (q_i - X) \quad (\star)$$

  Such a first-order ODE has smooth solutions. For regularization purposes, we assume that there is some large L, T such that we only consider $|X|, |P| \le L$ for times $t \in [0, T]$.

  In principle, $\star$ is a deterministic ordinary differential equation, which can be exactly solved given precise information about the initial states of the particles. In real experiments however, we do not have such precise knowledge of the microscopic state of the system. Fortuitously, physicists are generally uninterested in the vast majority of the $2n + 2$ degrees of freedom present in the system, and only care about a small subset. In our case, we are solely interested in the evolution of X and P.

  To this end, we model the initial state of the system as a **Gaussian distribution** in the phase space of the particles. In statistical mechanics literature, this is known as the **Boltzmann distribution**:

$$P(q_i(0), p_i(0)) = \frac{1}{Z} \exp\left(-\beta \left[\sum_i \frac{p_i^2}{2m_i} + \frac{m_i \omega_i^2}{2}(q_i^0 - X(0))^2\right]\right)$$

where β is some fixed constant called the **inverse temperature**, and Z is a normalization factor.

☾A.2 ODE Solution

We can now apply general ODE theory to solve this equation path-wise. Through Laplace transformations one immediately obtains the solution

$$q_i(t) = \underbrace{q_i^0 \cos(\omega_i t) + \frac{p_i^0}{m_i \omega_i} \sin(\omega_i t)}_{\text{homog. sol.}} + \omega_i (\sin(\omega_i \bullet) * X)(t)$$

By integration by parts (note that any solution X, P must be smooth),

$$\omega_i (\sin(\omega_i \bullet) * X)(t) = X(t) - \cos(\omega_i t)X(0) - \left(\cos(\omega_i \bullet) * \frac{P}{M}\right)(t).$$

Returning to $\star$, we obtain

$$\dot{P} = -V'(X) - \sum_i m_i \omega_i^2 \left\{ q_i^0 \cos(\omega_i t) + \frac{p_i^0}{m_i \omega_i} \sin(\omega_i t) - \cos(\omega_i t)X(0) - \left(\cos(\omega_i \bullet) * \frac{P}{M}\right)(t)\right\}$$

where we observe that the $X(t)$ from our integration by parts has nicely canceled. Hence, our original system ($\star$) has reduced to the following two equations:

$$\dot{X} = \frac{P}{M}, \qquad \dot{P} = -V'(X) - \frac{1}{M}(P * K)(t) + \xi(t) \tag{♠}$$

where $K(t) = \sum_i m_i \omega_i^2 \cos(\omega_i t)$ is the **friction kernel**, and

$$\xi(t) = \sum_i m_i \omega_i^2 (q_i^0 - X(0)) \cos(\omega_i t) + \omega_i p_i^0 \sin(\omega_i t)$$

is the **noise** which has absorbed all of the dependence on the (random) microscopic initial conditions.

♠ is a pre-cursor to the Langevin equation, which we will obtain after a suitable limit.

☙A.3 Fluctuations and friction

The friction kernel K can be thought of as an *average force* which controls dissipation, while the noise term ξ can be thought of as a *random force* which injects energy into the system. Observe that ξ is manifestly a centered Gaussian process, whose covariance is given by

$$\mathbb{E}\left[\xi(t)\xi(t')\right] = \frac{1}{\beta} \sum_i m_i \omega_i^2 \cos(\omega_j(t - t'))$$

Note this is exactly $\frac{1}{\beta} K(t - t')$, where K is the friction kernel. This is a manifestation of the **fluctuation-dissipation relation**, which is a general physical phenomenon.

The limit we desire to take is the one in which K becomes a point distribution at 0. Then, ♠ becomes a stochastic differential equation without the convolution, and ξ becomes "white noise," such that the $\xi(t)dt$ term becomes dW_t for some Brownian motion. The fluctuation-dissipation relation is encouraging, because it says that these two limits are compatible.

As an aside, in the physics literature, we distinguish these limiting cases. Observe that K is a function of the composition of the fluid (m_i, ω_i model the distribution of particle types and their interaction strengths).

- **White noise**: $K(t - t') \propto \delta(t - t')$, such that the spectrum is flat. The fluid is referred to as **Newtonian**.

- **Colored noise**: $K(t - t')$ is not a delta function, and the noise is correlated in time. The fluid is referred to as **visco-elastic**.

☙A.4 Taking a limit of the damping term

We will now present a series of results which fall short of establishing the desired limit. We make no claims of rigor in this section, but hope that results such as these lead to a real proof.

Fejer's theorem from elementary Fourier theory states there exists a sequence of m_i^N and ω_i^N as $N \to \infty$ such that $K(t)$ uniformly converges to any smooth function with period $2T$. We are only interested in the region $[0, T]$, so we will drop the periodicity condition. Hence, for every ε, δ we obtain a set of parameters m_i and ω_i such that $K(t)$ is uniformly within ε from the sharp Gaussian function

$$K_\delta(t) = \frac{2\gamma}{\sqrt{2\pi\delta}} \exp\left(-t^2/2\delta\right)$$

where γ is some constant.

Next, we would like to bound

$$|(P * K_\delta)(t) - \gamma P(t)|.$$

This is an intuitively clear but difficult-to-prove result. Heuristically,

$$(P * K_\delta)(t) = \int_0^t P(s)K(t-s)ds \approx \int_0^t P(t)K_\delta(t-s)ds = \gamma P(t)$$

For any particular set of these parameters, remember that we are still solving a smooth first-order ODE. Suppose we knew that with high probability along this sequence of parameters, sample paths of P on $[0, T]$ are α-Holder-continuous with constant C, independent of δ. This is hard to justify rigorously. The Wiener process and the **Ornstein-Uhlenbeck process**, where $V(x) = \frac{1}{2}x^2$, exhibits Holder continuity on $[0, T]$; this follows immediately from the representation as a scaled, time-transformed Wiener process. Then, one could hope to bound the above as the sum over a vanishing quantity around t, and a vanishing quantity over all time.

To conclude the proof, we wish to show that this limit of Gaussian processes $W_t^N = \int_0^t \xi_t^N$, as $K(t)$ becomes more and more peaked, is a Brownian motion, in distribution. We know that W_t^N are zero-mean, with covariance functions $\mathbb{E}\left[W_t^N W_{t'}^N\right]$ converging pointwise to $t \wedge t'$. In the finite-dimensional Brownain motion case, this is already sufficient to establish convergence in distribution to a Brownian motion.

The conclusion of this proof likely requires Gronwall's lemma or a more sophisticated result such as Wong-Zakai to transform the given bounds into bounds on the full solution.

☞B Results on Stochastic Differential Equations

We briefly review some details about SDEs on $\mathbb{R}^n$.

Give me a **diffusion coefficient matrix** $\sigma : \mathbb{R}^n \to \text{Hom}(\mathbb{R}^\ell, \mathbb{R}^n)$ and a **driving semimartingale** Z. σ is always locally Lipschitz, and Z is always a $(\mathscr{F}_t)_{t \geq 0}$-adapted continuous semimartingale with values on $\mathbb{R}^\ell$. If X_0 is $\mathscr{F}_0$-measurable $\mathbb{R}^n$-valued random variable, then a **solution to the SDE** $SDE(\sigma, Z, X_0)$ is a CSM X_t and a stopping time τ such that there exists a sequence of stopping times $\tau_n \uparrow \tau$ such that for all n,

$$X_{t \wedge \tau_n} = X_0 + \int_0^{t \wedge \tau_n} \sigma(X_s)dZ_s, \qquad t \geq 0.$$

Pictorially, we write

$$dX_t = \sigma(X_t)dZ_t, \qquad t < \tau.$$

Theory then states that

> **Theorem 7** (Existence and Uniqueness of Global Solution) — If σ is globally Lipschitz and $X_0 \in L^2$, then $SDE(\sigma, Z, X_0)$ has a unique global solution ($\tau = \infty$).

If σ is not declared to be globally Lipschitz, the possibility of *explosion times* arises. For instance, the ODE $dx_t = x_t^2 dt$ with $x_0 = 1$ has solution $X_t = \frac{1}{1-t}$, which explodes at $t = 1$.

The standard way to deal with this is to consider the one-point compactification of your space. Explicitly, if M is a locally compact metric space, let $\hat{M} = M \cup \{\infty_M\}$ be its one-point compactification. Then, a path on M with explosion time e is a continuous map $[0, \infty) \to \hat{M}$ such that $x_t \in M$ for $t < e$ and $x_t = \infty_M$ for $t \geq e$. The collection of all such paths is called the path space of M, denoted $W(M)$.

Theorem 8 (Existence and Uniqueness of Local Solution) — If σ is locally Lipschitz and X_0 is merely $\mathscr{F}_0$-measurable, then there exists a unique $W(\mathbb{R}^n)$-valued process X which is a solution of $SDE(\sigma, Z, X_0)$ up to its explosion time $\tau = e(X)$.