

STOCHASTIC DIFFERENTIAL EQUATIONS: THEORY AND NUMERICAL SIMULATIONS

BY: ADAM QUINN JAFFE
ADVISOR: LENYA RYZHIK

A THESIS SUBMITTED IN PARTIAL SATISFACTION OF THE REQUIREMENTS FOR
THE DEGREE OF BACHELOR OF SCIENCE IN MATHEMATICS WITH HONORS

ABSTRACT. The present work contains a brief introduction to the theory of stochastic differential equations. First, we precisely define stochastic integration and prove a basic existence and uniqueness result for stochastic ordinary differential equations (SODEs). Then we introduce some general principles for simulating solutions to SODEs and study a class of numerical schemes. We then shift our attention to stochastic partial differential equations (SPDEs), restricting our attention for brevity's sake to the stochastic heat equation in one spatial dimension: We define stochastic integration in this setting, prove a basic existence and uniqueness result, and then explore a numerical schemes for numerically solving the SPDE.

1. INTRODUCTION

From a variety of perspectives, there is great interest in studying stochastic differential equations. In applied mathematics, the number of problems in which such models are relevant is staggering: statistical mechanics, wave propagation in random media, population genetics (and mathematical biology more broadly), and the pricing of financial assets, just to name a few. In pure mathematics, too, such problems arise as quite natural objects of study in stochastic processes and dynamical systems.

There is, additionally, great interest in developing numerical schemes for solving stochastic differential equations, especially for applications to the problems mentioned above. This turns out, however, to be a quite difficult task, since many things can go wrong if one uses naive methods to simulate solutions to stochastic differential equations. In this way, a strong theoretical background in the mathematical theory of stochastic differential equations will help one go a long way in avoiding some major pitfalls in simulations of solutions.

This thesis is intended to serve as an introduction to this very rich theory, but it by no means comprehensive. It represents a small amount of the material I studied during my senior year as an undergraduate at Stanford, which in turns represents a tiny amount of material in the entire discipline. The present work draws on the sources [5], [3], and [8] most directly, although many other sources have been instrumental in my studies; readers will notice that I have changed a great deal of notation and selected material quite tacitly so as to create the most coherent presentation possible.

The remainder of this thesis is structured as follows: In Section 2 we develop the theory of stochastic ordinary differential equations (SODEs), and in Section 3 we

study numerical simulations for SODEs. In Section 4 we move on to the theory of stochastic partial differential equations (SPDEs), and in Section 5 we study numerical simulations for SPDEs. The appendices contain some bounds and inequalities that we will need, as well as the MATLAB code used to generate our numerical simulations and plots.

2. THEORY OF SODE

The goal of this section is to introduce a basic theory of stochastic ordinary differential equations (SODEs). That is, we aim to study ordinary differential equations (ODEs) which have an additional driving force of a white noise. For the sake of simplicity, we limit our scope to 1-dimensional SODEs.

As an informal motivation, let us consider an arbitrary 1-dimensional ODE whose dynamics are perturbed slightly by adding a “white noise” term W_t , which has scaling that may depend on the value of the function and on time:

$$(1) \quad \frac{dU_t}{dt} = \mu(t, U_t) + \sigma(t, U_t)W_t, \quad U_0 = u_0$$

There are many questions to be asked about even such a system: existence and uniqueness of solutions, continuous dependence on initial data, stability of solutions, model-specific specifications, and more. However, before we can ask any of these questions, we must decide how to make sense of the equation above, and, in particular, how to make sense of the white noise process $\{W_t\}$.

What are some properties that we hope that white noise should have? At the very least, we need the following: For any t , we need W_t to be a $N(0, 1)$ random variable, and, for any distinct t_1, t_2 , we need W_{t_1} and W_{t_2} to be independent. In a discrete-time system, this makes perfect sense; in a continuous-time system, it turns out that no such $\{W_t\}_t$ exists. In many disciplines, the existence of this process is simply assumed. In a more mathematical treatment like this, we will not avoid this subtlety.

Not all hope is lost in making (1) into a mathematically precise statement. To do this, observe that we can write it equivalently as the following integral equation:

$$(2) \quad U_t = \int_0^t \mu(s, U_s) ds + \int_0^t \sigma(s, U_s) W_s ds$$

Hence, the only difficulty is in defining the second integral on the right side.

Although we cannot make sense of the white noise as a true stochastic process, we can recall that there is a stochastic process, which is said to be, heuristically speaking, the antiderivative of the white noise. This process is the Brownian motion $\{B_t\}_t$ for which we can, formally, write $dB_t/dt = W_t$. Integration by parts then suggests that we can write the mysterious integral as:

$$(3) \quad \int_0^t \sigma(s, U_s) W_s ds = \int_0^t \sigma(s, U_s) dB_s$$

From this perspective, we have an idea of how to define the stochastic integral: We will integrate the coefficient function $\{\sigma(t, U_t)\}_t$ “against” a Brownian motion $\{B_t\}_t$. Once we make this precise, we will be able to define the SODE and explore some of its consequences.

2.1. Brownian Motion. Our first goal is to construct and study the Brownian motion. While there are many ways to do this construction, we will focus on one method that will be easy to generalize to higher dimensions, as this will be needed when we eventually undertake the study of stochastic partial differential equations in Section 4.

We begin by defining the white noise, not as time-indexed stochastic process, but as a stochastic process indexed by Borel subsets of \mathbb{R}^d .

Definition 2.1. A *white noise* on \mathbb{R}^d on $(\Omega, \mathcal{F}, \mathbb{P})$ is a function $W : \mathcal{B}(\mathbb{R}^d) \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P})$ such that the following properties hold:

- (2.1.1) For all $A \in \mathcal{B}(\mathbb{R}^d)$, we have $W(A) \sim N(0, \lambda(A))$.
- (2.1.2) For all $A_1, A_2 \in \mathcal{B}(\mathbb{R}^d)$ with $A_1 \cap A_2 = \emptyset$, the random variables $W(A_1)$ and $W(A_2)$ are independent, and $W(A_1 \cup A_2) = W(A_1) + W(A_2)$ holds \mathbb{P} -almost surely.

For the present work, we take it as granted that we can construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a white noise W on \mathbb{R}^d on this space. The details of this construction are highly nontrivial, but not worth discussing here. For the remainder of this subsection, let us fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which W , a white noise on \mathbb{R} is defined. Let us also restrict attention to a subset $[0, T] \subseteq \mathbb{R}$ for some fixed $T > 0$.

We now define a Brownian motion from a white noise as follows:

Definition 2.2. Given any white noise W , the *Brownian motion* is a stochastic process $\{B_t\}_{t \in [0, T]}$ defined from the white noise as $B_t = W((0, t])$.

The definition above does not give us a very intuitive understanding of the distributional properties of the Brownian motion. To do this, we prove the following easy result:

Lemma 2.3. The Brownian motion $\{B_t\}_t$ enjoys the following properties, for any $s, t \in [0, T]$ with $s < t$:

- (2.3.1) $\mathbb{P}(B_0 = 0) = 1$
- (2.3.2) $B_t - B_s$ is distributed as $N(0, t - s)$.
- (2.3.3) $B_t - B_s$ is independent of B_s .

Proof. For (2.3.1), note that $B_0 = W(\emptyset) \sim N(0, 0)$, hence $B_0 = 0$ holds \mathbb{P} -almost surely. For (2.3.2) and (2.3.3), note that we have $W([0, t]) = W([0, s] \cup [s, t]) = W([0, s]) + W([s, t])$, hence $B_t - B_s = W([s, t])$. This implies (2.3.2) because $W([s, t]) \sim N(0, t - s)$, and (2.3.3) because $W([s, t])$ is independent of $B_s = W([0, s])$. \square

One desirable property of the Brownian motion still eludes us: Are the sample paths of $\{B_t\}_t$ continuous? It turns out that the answer, in general, is no. However, we can construct a stochastic process with continuous sample paths \mathbb{P} -almost surely whose finite-dimensional distributions match those of the Brownian motion. To do this, we appeal to the following extremely important result, whose proof can be found in any standard textbook on stochastic processes, as in, say, [2]:

Theorem 2.4 (Kolmogorov-Chentsov). Let $\{X_t\}_{t \in [0, T]}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. If there exist positive constants α, β , and K , such that

$$(4) \quad \mathbb{E}[|X_t - X_s|^\alpha] \leq K|t - s|^{1+\beta}$$

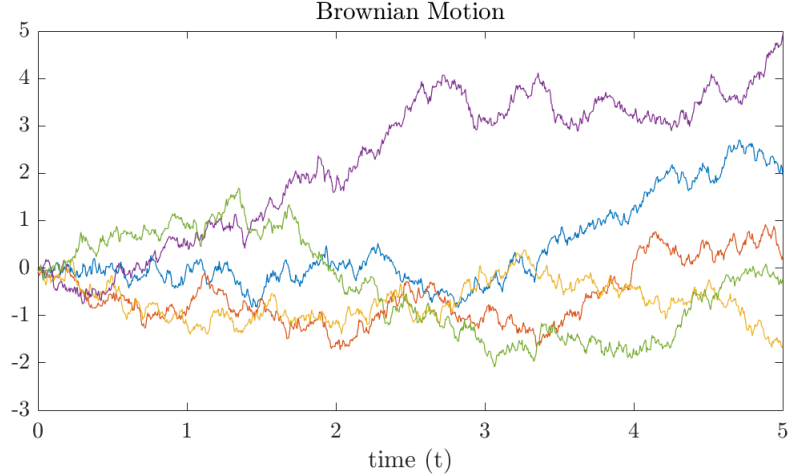
holds for all $s, t \in [0, T]$, then there exists a stochastic process $\{\tilde{X}_t\}_{t \in [0, T]}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(X_t = \tilde{X}_t) = 1$ holds for all $t \in [0, T]$ and whose sample paths are \mathbb{P} -almost surely Hölder continuous with any exponent $\gamma \in (0, \beta/\alpha)$.

Let us apply this result to the Brownian motion. Since for any $s, t \in [0, T]$, the random variable $B_t - B_s$ is distributed as $N(0, |t - s|)$, we know that its fourth moment satisfies

$$(5) \quad \mathbb{E} \left[|B_t - B_s|^4 \right] = 3|t - s|^2$$

Therefore, applying Theorem 2.4 with $\alpha = 4, \beta = 1$, and $K = 3$ gives that there is a stochastic process $\{\tilde{B}_t\}_t$ such that $\mathbb{P}(B_t = \tilde{B}_t) = 1$ holds for all $t \in [0, T]$ and whose sample paths are, in particular, continuous \mathbb{P} -almost surely. For the remainder of this paper, we will always assume that $\{B_t\}_t$ is a continuous version of a Brownian motion, constructed in this way.

To visualize the Brownian motion, we consult the following image which plots 5 sample paths of a Brownian motion on the time interval $[0, T]$ for $T = 5$: We leave the details of how exactly this simulation works to Section 3.



The next property of Brownian motion that we will need is its relationship with the random walk. In particular, the following result makes precise the idea that the Brownian can be characterized as the limit of a sequence of suitably scaled random walks.

Theorem 2.5 (Donsker). Let $\{X_n\}_{n=1}^\infty$ be a sequence of independent, identically distributed random variables with $\mathbb{E}[X_1] = 0$ and $\text{Var}(X_1) = 1$, and, for any $N > 0$, set

$$(6) \quad \hat{S}_t^N = \frac{1}{\sqrt{N}} \sum_{i=1}^n X_i$$

whenever $t = nT/N$ for $n = 0, 1, \dots, N$, and by linear interpolation otherwise. Then, we have

$$(7) \quad \{\hat{S}_t^N\}_t \rightarrow \{B_t\}_t$$

in law as $C([0, T])$ -valued random variables.

The importance of Theorem 2.5 is that the Brownian motion is the limiting object of a suitably scaled random walk, no matter the particular distribution of the increments. Of particular importance to us will be the case of $X_1 \sim N(0, 1)$, but occasionally other choices may be useful.

The final property of the Brownian motion that we investigate is the fundamental reason that stochastic calculus differs from ordinary calculus.

Lemma 2.6. The Brownian motion $\{B_t\}_t$ up to time $t \in [0, T]$ has quadratic variation equal to t and total variation equal to ∞ , both holding \mathbb{P} -almost surely.

Proof. First we prove that, \mathbb{P} -almost surely, the quadratic variation equals t . To do this, fix $t \in [0, T]$, and for each positive integer k , set Π_k to be a partition of $[0, t]$. Let n_k be the number of intervals in this partition and let $|\Pi_k| = \max_{j=0,1,\dots,k-1} (t_{j+1} - t_j)$ denote the largest size of interval in this partition. Assume also that $\{\Pi_k\}_k$ are such that $n_k \rightarrow \infty$ and $|\Pi_k| \rightarrow 0$ with $n_k |\Pi_k|^2$ summable.

For arbitrary $\varepsilon > 0$, we will use Chebyshev's inequality to bound the probability:

$$(8) \quad \mathbb{P} \left(\left| \sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|^2 - t \right| \geq \varepsilon \right).$$

In order to do this, we will have to compute the variance of the random variable on the left side of the inequality. Note that we can rewrite this equivalently as

$$(9) \quad \sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|^2 - t = \sum_{j=0}^{n_k-1} \left(|B_{t_{j+1}} - B_{t_j}|^2 - (t_{j+1} - t_j) \right).$$

Now observe that $|B_{t_{j+1}} - B_{t_j}|^2 - (t_{j+1} - t_j)$ for $j = 0, 1, \dots, k-1$ are independent, mean zero random variables. Hence, the variance of their sum is the sum of their variances:

$$(10) \quad \text{Var} \left(\sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|^2 - t \right) = \sum_{j=0}^{n_k-1} \text{Var} \left(|B_{t_{j+1}} - B_{t_j}|^2 - (t_{j+1} - t_j) \right),$$

and we can compute the variances individually as follows:

$$(11) \quad \text{Var} \left(|B_{t_{j+1}} - B_{t_j}|^2 - (t_{j+1} - t_j) \right)$$

$$(12) \quad = \mathbb{E} \left[\left| |B_{t_{j+1}} - B_{t_j}|^2 - (t_{j+1} - t_j) \right|^2 \right]$$

$$(13) \quad = \mathbb{E} \left[|B_{t_{j+1}} - B_{t_j}|^4 \right] + (t_{j+1} - t_j)^2 - 2\mathbb{E} \left[|B_{t_{j+1}} - B_{t_j}|^2 \right] (t_{j+1} - t_j)$$

$$(14) \quad = 3(t_{j+1} - t_j)^2 + (t_{j+1} - t_j)^2 - 2(t_{j+1} - t_j)^2$$

$$(15) \quad = 2(t_{j+1} - t_j)^2$$

Plugging this into (10), we have concluded:

$$(16) \quad \text{Var} \left(\sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|^2 - t \right) = 2 \sum_{j=0}^{n_k-1} (t_{j+1} - t_j)^2$$

Combining this all, we get the bound:

$$(17) \quad \sum_{k=1}^{\infty} \mathbb{P} \left(\left| \sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|^2 - t \right| \geq \varepsilon \right) \leq \frac{2}{\varepsilon^2} \sum_{k=1}^{\infty} \sum_{j=0}^{n_k-1} (t_{j+1} - t_j)^2$$

$$(18) \quad \leq \frac{2}{\varepsilon^2} \sum_{k=1}^{\infty} n_k |\Pi_k|^2,$$

which is finite by assumption.

Hence, by Borel-Cantelli, we have

$$(19) \quad \mathbb{P} \left(\left| \sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|^2 - t \right| \geq \varepsilon \text{ finitely often} \right) = 1.$$

In other words, for almost all ω we have

$$(20) \quad \sum_{j=0}^{n_k-1} |B_{t_{j+1}}(\omega) - B_{t_j}(\omega)|^2 \rightarrow t$$

This proves that, \mathbb{P} -almost surely, the quadratic variation equals t .

Next we prove that, \mathbb{P} -almost surely, the total variation is infinite. Towards a contradiction, suppose that there were a set of positive probability on which the total variation were finite. Then let $\{\Pi_k\}_k$ be partitions satisfying the same conditions as above, and consider the following inequality:

$$(21) \quad \sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|^2 \leq \max_{j=0,1,\dots,n_k-1} |B_{t_{j+1}} - B_{t_j}| \sum_{j=0}^{n_k-1} |B_{t_{j+1}} - B_{t_j}|$$

Take the lim sup as $k \rightarrow \infty$ of the above. We know that, \mathbb{P} -almost surely, the left side goes to t . Also \mathbb{P} -almost surely, the first factor on the right side goes to 0. But, by assumption, the second factor on the right remains bounded with positive probability. Hence, with positive probability, the right side goes to zero while the left side goes to t . This is a contradiction and completes the proof. \square

The importance of Lemma 2.6 deserves some elaboration. If we were to be cavalier for a moment, we might write the sum above as an integral and come to the following strange conclusion:

$$(22) \quad \int_0^t (dB_s)^2 = t$$

Differentiating this, we conclude $(dB_t)^2 = dt$. This equation has no precise meaning mathematically, but it turns out to be an extremely useful heuristic for stochastic calculus.

Generally in calculus, one is interested with finding polynomial approximations to smooth functions. That is, as some small step size h approaches zero, we deem terms important if they go to zero like h^p for $p \leq d$, and we deem them negligible if they go to zero like h^p for $p > d$, where d is some fixed degree of interest. In ordinary calculus, this process is easy since p is integer-valued. In stochastic calculus, however, there are terms which take on half-integer values as well; consider our most recent discovery of $dB_t = (dt)^{1/2}$.

In this sense, the nonzero quadratic variation of the Brownian motion forces us to take on a totally different perspective when deciding which terms are important and which terms are negligible. Generally one likes to think of dt as being on the order of $O(h)$ and d as being fixed at 1. We then deduce the following identities:

$$(23) \quad (dt)^2 = 0 \quad dt dB_t = 0 \quad (dB_t)^2 = dt$$

While these identities are not precise as written, they give rise to useful heuristics that will almost never lead one astray. In many disciplines, these identities are simply taken as fact and no more is ever said about them.

Now that we have undertaken a brief introduction to the Brownian motion, we are ready to define the stochastic integral of (3).

2.2. Stochastic Integration. As we saw in the introduction, our goal is to make sense of integrals of the form:

$$(24) \quad \int_0^T f dB_t$$

where $f : [0, T] \times \Omega \rightarrow \mathbb{R}$ is some random function and $\{B_t\}_t$ is a Brownian motion.

However, the sample paths of a Brownian motion have unbounded variation almost surely, so this integral cannot be computed in a Riemann-Stieltjes sense, despite what the notation suggests. We must find a more clever construction if one is to define a reasonable meaning for stochastic integration.

Our approach will be to leverage tools from functional analysis. The effect of this treatment is not only to simplify the exposition, but also highlight the fundamental structure of a construction that is often presented as rather “magical”. To be specific, we demonstrate that the canonical construction of the Ito integral on its usual domain is nothing more than the continuous extension of a linear transformation which is initially defined on one of its dense subsets. We will spend the remainder of the subsection making this idea precise and exploring some of its basic consequences. To get started, let us fix some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which a Brownian motion $\{B_t\}_t$ is defined, and let us use $\mathbb{F} = \{\mathcal{F}_t\}_t$ to denote the natural filtration of $\{B_t\}_t$.

It makes sense to begin by defining the Ito integral for a small and easy-to-work-with class of functions. Recalling the construction of the Lebesgue integral, then, it is natural to define the following:

Definition 2.7. A stochastic process $\phi : [0, T] \times \Omega \rightarrow \mathbb{R}$ is called *elementary* with respect to \mathbb{F} if it can be written as

$$(25) \quad \phi(t, \omega) = \sum_{j=1}^k E_j(\omega) \mathbf{1}_{(t_j, t_{j+1}]}(t)$$

for some positive integer k , discretization $0 = t_0 < t_1 < \dots < t_k = T$, and random variables $\{E_j\}_j \subseteq L^2(\Omega)$ such that E_j is \mathcal{F}_{t_j} -measurable for all j . We use $\mathcal{E}_{\mathbb{F}}$ to denote the collection of all elementary processes with respect to \mathbb{F} .

The assumption that each E_j be \mathcal{F}_{t_j} -measurable should not be glossed over. In fact, this assumption is rather important as a modeling concern in many problems where the stochastic integral should never be able to “look into the future”.

Definition 2.8. The *Ito integral* is the map $I_T : \mathcal{E}_{\mathbb{F}} \rightarrow L^2(\Omega)$ defined as follows: For $\phi \in \mathcal{E}_{\mathbb{F}}$ with discretization $\{t_j\}_j$ and random variables $\{E_j\}_j$, set

$$(26) \quad \int_0^T \phi dB_t(\omega) = I_T(\phi)(\omega) = \sum_{j=0}^{k-1} E_j(\omega) (B_{t_{j+1}}(\omega) - B_{t_j}(\omega))$$

We will use the notation $\int_0^T \phi dB_t$ and $I_T(\phi)$ interchangeably throughout the remainder of our work. The former is a fairly intuitive representation of the Ito integral, but it can be misleading at times. The latter is preferred in many settings, since it emphasizes the operator structure of the Ito integral. Next, we will explore some basic properties of this map.

Lemma 2.9. Let $\phi, \psi \in \mathcal{E}_{\mathbb{F}}$ be any elementary processes and $\alpha, \beta \in \mathbb{R}$ be any reals. The map $I_T : \mathcal{E}_{\mathbb{F}} \rightarrow L^2(\Omega)$ enjoys the following properties:

- (2.9.1) (Linearity) $I_T(\alpha\phi + \beta\psi) = \alpha I_T(\phi) + \beta I_T(\psi)$
- (2.9.2) (Mean Zero) $\mathbb{E}[I_T(\phi)] = 0$
- (2.9.3) (Ito Isometry) $\|I_T(\phi)\|_{L^2(\Omega)} = \|\phi\|_{L^2([0,T] \times \Omega)}$
- (2.9.4) (Martingale) The stochastic process $\{I_t(\phi)\}_t$ is a martingale with respect to the filtration $\{\mathcal{F}_t^B\}_t$.
- (2.9.5) (Continuity) The stochastic process $\{I_t(\phi)\}_t$ is \mathbb{P} -almost surely continuous in t .

Proof. To prove (2.9.1), we first show that the process $\alpha\phi + \beta\psi$ is elementary. Take $\{t_j\}_j$ to be the (ordered) union of the discretizations of ϕ and ψ , and take $\{E_j\}_j$ and $\{F_j\}_j$ be the obvious alignments of the random variables of ϕ and ψ along the new discretization $\{t_j\}_j$. Then we can easily compute the the Ito integral of the linear combination as follows:

$$(27) \quad \int_0^T (\alpha\phi + \beta\psi) dB_t = \sum_{j=0}^{k-1} (\alpha E_j + \beta F_j) (B_{t_{j+1}} - B_{t_j})$$

$$(28) \quad = \alpha \sum_{j=0}^{k-1} E_j (B_{t_{j+1}} - B_{t_j}) + \beta \sum_{j=0}^{k-1} F_j (B_{t_{j+1}} - B_{t_j})$$

$$(29) \quad = \alpha \int_0^T \phi dB_t + \beta \int_0^T \psi dB_t.$$

This proves linearity.

To prove (2.9.2), we easily compute

$$(30) \quad \mathbb{E} \left[\int_0^T \phi dB_t \right] = \mathbb{E} \left[\sum_{j=0}^{k-1} E_j (B_{t_{j+1}} - B_{t_j}) \right]$$

$$(31) \quad = \sum_{j=0}^{k-1} \mathbb{E}[E_j] \mathbb{E}[B_{t_{j+1}} - B_{t_j}] = 0.$$

Likewise for (2.9.3):

$$(32) \quad \mathbb{E} \left[\left| \int_0^T \phi dB_t \right|^2 \right] = \sum_{j=0}^{k-1} \mathbb{E} [|E_j (B_{t_{j+1}} - B_{t_j})|^2]$$

$$(33) \quad = \sum_{j=0}^{k-1} \mathbb{E} [|E_j|^2] (t_{j+1} - t_j)$$

$$(34) \quad = \int_0^T \mathbb{E} [|\phi(t, \cdot)|^2] dt$$

Next, we prove (2.9.4). If $S = 0$, the claim is immediate. Otherwise, choose a nonnegative integer m such that $t_m < S \leq t_{m+1}$ holds. Then note that the discretization $\{t_j\}_j$ of ϕ can be refined by inserting the value S , so that the elementary function ϕ can be written as:

$$(35) \quad \phi(t, \omega) = \sum_{j=0}^{m-1} E_j(\omega) \mathbf{1}_{(t_j, t_{j+1}]}(t)$$

$$(36) \quad + E_m(\omega) \mathbf{1}_{(t_m, S]}(t) + E_m(\omega) \mathbf{1}_{(S, t_{m+1}]}(t) + \sum_{j=m+1}^{k-1} E_j(\omega) \mathbf{1}_{(t_j, t_{j+1}]}(t)$$

Using this form of ϕ , we compute that the conditional expectation of the Ito integral must be:

$$(37) \quad \mathbb{E} \left[\int_0^T \phi dB_t \middle| \mathcal{F}_S^B \right] = \int_0^S \phi dB_t + \mathbb{E}[E_m(B_{t_{m+1}} - B_S) | \mathcal{F}_S]$$

$$(38) \quad + \sum_{j=m+1}^k \mathbb{E} [E_j (B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_S]$$

$$(39) \quad = \int_0^S \phi dB_t + E_m \mathbb{E}[B_{t_{m+1}} - B_S]$$

$$(40) \quad + \sum_{j=m+1}^k \mathbb{E}[B_{t_{j+1}} - B_{t_j}] \mathbb{E}[E_j | \mathcal{F}_S]$$

$$(41) \quad = \int_0^S \phi dB_t,$$

This proves that the Ito integral is a martingale.

Finally, we prove (2.9.5). Note that if ϕ is elementary on $[0, T]$, with representation

$$(42) \quad \phi(t, \omega) = \sum_{j=1}^k E_j(\omega) \mathbf{1}_{(t_j, t_{j+1}]}(t),$$

then it is also elementary on $[0, t]$, with representation

$$(43) \quad \phi(s, \omega) = \sum_{j=1}^{j(t)-1} E_j(\omega) \mathbf{1}_{(t_j, t_{j+1}]}(s) + E_{j(t)}(\omega) \mathbf{1}_{(t_{j(t)}, t]}(s),$$

where $j(t) = \max\{j : t_j \leq t\}$. But the Ito integral of the latter is just

$$(44) \quad I_t(\phi) = \sum_{j=0}^{j(t)-2} E_j(\omega) (B_{t_{j+1}}(\omega) - B_{t_j}(\omega)) + E_{j(t)}(\omega) (B_t(\omega) - B_{t_{j(t)}}(\omega)),$$

and this is continuous \mathbb{P} -almost surely since $B_t(\omega) - B_{t_{j(t)}}(\omega)$ is continuous \mathbb{P} -almost surely. This completes the proof. \square

This result essentially closes the book on Ito integral for elementary processes, so we are now left to extend the definition to a larger collection of stochastic processes. What exactly should this larger collection be? Our functional analysis perspective will lead us to a very natural choice: The results of Lemma 2.9 tells us that $I_T : \mathcal{E}_{\mathbb{F}} \rightarrow L^2(\Omega)$ is a bounded linear operator, when $\mathcal{E}_{\mathbb{F}}$ is viewed as a subset of $L^2([0, T] \times \Omega)$. Hence, we can use the bounded linear transformation theorem to extend our definition to all processes in the closure, $\overline{\mathcal{E}_{\mathbb{F}}}$. This is exactly the “magic” which we can use to bypass the obstacles of a Riemann-Stieltjes-type construction!

It still remains to characterize the stochastic processes $f \in \overline{\mathcal{E}_{\mathbb{F}}}$. Since we need $\mathcal{E}_{\mathbb{F}} \subseteq L^2([0, T] \times \Omega)$, it is clear that any $f \in \overline{\mathcal{E}_{\mathbb{F}}}$ must satisfy, at the very least, $\|f\|_{L^2([0, T] \times \Omega)} < \infty$. Of course, we should also impose some measurability constraints on the stochastic processes we are able to Ito integrate. This point actually demands some caution: Oksendal [5] proceeds by merely taking f to be adapted, but this is not enough because we need to be able to integrate f on the space $([0, T] \times \Omega, \mathcal{B}([0, T]) \otimes \mathcal{F}, \lambda \otimes \mathbb{P})$. Lalley [4] identifies this mistake and corrects it, but then moves to a very general collection of processes that the technical details become a bit cumbersome. In the present work, we compromise with the following:

Definition 2.10. Call a stochastic process $f : [0, T] \times \Omega \rightarrow \mathbb{R}$ *predictable* with respect to \mathbb{F} if it is $\sigma(\mathcal{E}_{\mathbb{F}})$ -measurable. Also define the set $\mathcal{P}_{\mathbb{F}}^2 = L^2([0, T] \times \Omega, \sigma(\mathcal{E}_{\mathbb{F}}), \lambda \otimes \mathbb{P})$ of all processes which are predictable with respect to \mathbb{F} and square integrable.

The next few results will help us get a handle on which stochastic processes lie in $\mathcal{P}_{\mathbb{F}}^2$. In particular, we will generate useful sufficient conditions to do the heavy lifting for us.

Lemma 2.11. For any probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and filtration $\mathbb{F} = \{\mathcal{F}_t\}_t$ of \mathcal{F} , we have that $\mathcal{P}_{\mathbb{F}}^2$ is a Hilbert space.

Proof. Since $\mathcal{E}_{\mathbb{F}}$ is a sub- σ -algebra of $\mathcal{B}([0, T]) \otimes \mathcal{F}$, we get that $L^2([0, T] \times \Omega, \sigma(\mathcal{E}_{\mathbb{F}}), \lambda \otimes \mathbb{P})$ is a closed linear subspace of $L^2([0, T] \times \Omega, \mathcal{B}([0, T]) \otimes \mathcal{F}, \lambda \otimes \mathbb{P})$. The latter is complete, and a closed subset of a complete space is complete. \square

Lemma 2.12. With respect to any filtration $\mathbb{F} = \{\mathcal{F}_t\}_t$, an \mathbb{F} -adapted and continuous stochastic process is \mathbb{F} -predictable.

Proof. Let $f : [0, T] \times \Omega \rightarrow \mathbb{R}$ be any \mathbb{F} -adapted and continuous process, and for each positive integer n and $j \in \{0, 1, \dots, 2^n\}$, set $t_j = jT/2^n$. Now define the following approximations:

$$(45) \quad \phi_n(t, \omega) = \sum_{j=0}^{2^n-1} f(t_j, \omega) \mathbf{1}_{(t_j, t_{j+1}]}(t)$$

Note that $f(t_j, \omega)$ is \mathcal{F}_{t_j} -measurable, so each ϕ_n is elementary with respect to \mathbb{F} . Moreover, by continuity, we have $\phi_n(t, \omega) \rightarrow f(t, \omega)$ for all (t, ω) . Hence, for any $\alpha \in \mathbb{R}$ we have

$$(46) \quad \{(t, \omega) : f(t, \omega) < \alpha\} = \bigcup_{k=1}^{\infty} \bigcap_{n \geq k} \{(t, \omega) : \phi_n(t, \omega) < \alpha\}$$

The right side of the above is in $\sigma(\mathcal{E}_{\mathbb{F}})$, and the intervals $(-\infty, \alpha)$ generate $\mathcal{B}(\mathbb{R})$, so f is $\sigma(\mathcal{E}_{\mathbb{F}})$ -measurable, hence \mathbb{F} -predictable. \square

Lemma 2.13. The space $\mathcal{E}_{\mathbb{F}} \subseteq L^2([0, T] \times \Omega)$ is dense in $\mathcal{P}_{\mathbb{F}}^2$. That is, for any $f \in \mathcal{P}_{\mathbb{F}}^2$, there exists a sequence $\{\phi_n\}_n$ of elementary processes with respect to \mathbb{F} with $\phi_n \rightarrow f$ in $L^2([0, T] \times \Omega)$.

Proof. The proof proceeds in three steps, which should remind the reader of the so-called “standard machine” in measure theory. That is, each step will prove the desired result for an iteratively larger collection of processes, and we will eventually come to the desired conclusion for all of $\mathcal{P}_{\mathbb{F}}^2$.

Step 1: If $f \in \mathcal{P}_{\mathbb{F}}^2$ has $f(t, \omega)$ continuous in t for all ω and is bounded uniformly by M across all (t, ω) , then there exists a sequence of $\{\phi_n\}_n \subseteq \mathcal{E}_{\mathbb{F}}$ with $\phi_n \rightarrow f$ in $L^2([0, T] \times \Omega)$.

For each positive integer n let $\{t_{n,j}\}_j$ be the endpoints of a partition of $[0, T]$ into 2^n intervals, such that $\max_{j=0,1,\dots,2^n-1} (t_{n,j+1} - t_{n,j}) \rightarrow 0$ as $n \rightarrow \infty$. Then define

$$(47) \quad \phi_n(t, \omega) = \sum_{j=0}^{2^n-1} f(t_{n,j}, \omega) \mathbf{1}_{(t_{n,j}, t_{n,j+1}]}(t)$$

Since f is, in particular, \mathbb{F} -adapted, we have that $f(t_{n,j}, \omega)$ is $\mathcal{F}_{t_{n,j}}$ -measurable, and hence ϕ_n is elementary with respect to \mathbb{F} . Now let ω be fixed. By the continuity of and uniform bound on f , we have the following for all t : $|f(t, \omega) - \phi_n(t, \omega)|^2 \rightarrow 0$ and $|f(t, \omega) - \phi_n(t, \omega)|^2 \leq 4M^2$. Hence, the dominated convergence theorem gives $\phi_n(\cdot, \omega) \rightarrow f(\cdot, \omega)$ in $L^2([0, T]) \rightarrow 0$ for all ω . But f is uniformly bounded across ω as well, so we have $\|f(\cdot, \omega) - \phi_n(\cdot, \omega)\|_{L^2([0, T])}^2 \leq 4M^2T$. Therefore, another application of the dominated convergence theorem, along with Fubini-Tonelli, gives $\phi_n \rightarrow f$ in $L^2([0, T] \times \Omega)$, as needed.

Step 2: If $f \in \mathcal{P}_{\mathbb{F}}^2$ is bounded uniformly by M across all (t, ω) , then there exists a sequence of $\{\phi_n\}_n \subseteq \mathcal{E}_{\mathbb{F}}$ with $\phi_n \rightarrow f$ in $L^2([0, T] \times \Omega)$.

Let $K : \mathbb{R} \rightarrow \mathbb{R}$ be any nonnegative continuous function with $\text{supp}(K) \subseteq [0, T]$ and $\int_{\mathbb{R}} K(t)dt = 1$, for example a suitably scaled triangular pulse between 0 and T . In particular, this implies that K is bounded. Then we can define the functions $\{K_n\}_n$ by $K_n(t) = nK(nt)$, and

$$(48) \quad g_n(t, \omega) = (K_n * f(\cdot, \omega))(t) = \int_0^T K_n(s) f(t-s, \omega) ds$$

$$(49) \quad = \int_0^T K_n(t-s) f(s, \omega) ds$$

where we have extended $f(t, \omega)$ to take on the value zero whenever $t \leq 0$. Note that g_n is continuous, since it is a convolution of a continuous function K_n with a bounded function f . Moreover, we have

$$(50) \quad |g_n(t, \omega)| \leq TM \sup_s K(s)$$

so g_n is bounded uniformly across all (t, ω) .

Now we prove that g_n is \mathbb{F} -predictable for each n . By Lemma 2.12 and the fact that g_n is continuous, it suffices to prove that g_n is \mathbb{F} -adapted. Note that we can write

$$(51) \quad g_n(t, \omega) = \int_0^T K_n(t-s) f(s, \omega) ds = \int_0^t K_n(t-s) f(s, \omega) ds,$$

and the last line holds since $s \geq t$ implies $K_n(t-s) = 0$. But $f(s, \omega)$ is \mathcal{F}_t -measurable whenever $s \leq t$, so its integral up to time t is also \mathcal{F}_t -measurable. This proves that each g_n is \mathbb{F} -predictable.

Next, we prove as an intermediate step that g_n converges in $L^2([0, T] \times \Omega)$ to f . First let ω be fixed, and compute the following bound, using Minkowski's integral inequality (Lemma A.7):

$$(52) \quad \|g_n(\cdot, \omega) - f(\cdot, \omega)\|_{L^2([0, T])}$$

$$(53) \quad = \left(\int_0^T \left| \int_0^T K(s) \left(f\left(t - \frac{s}{n}, \omega\right) - f(t, \omega) \right) ds \right|^2 dt \right)^{\frac{1}{2}}$$

$$(54) \quad \leq \int_0^T K(s) \left\| f\left(\cdot - \frac{s}{n}, \omega\right) - f(\cdot, \omega) \right\|_{L^2([0, T])} ds.$$

The integrand is bounded by $2M \sup_s K(s)$, so, if we can prove that it goes to zero pointwise in s as $n \rightarrow \infty$, then dominated convergence will give $g_n(\cdot, \omega) \rightarrow f(\cdot, \omega)$ in $L^2([0, T])$.

To do this, it suffices to prove that the operators $\tau_a : L^2([0, T]) \rightarrow L^2([0, T])$ defined by $(\tau_a f)(t) = f(t-a)$ for $a \geq 0$ satisfy $\tau_a f \rightarrow f$ in $L^2([0, T])$ as $a \rightarrow 0$. Obviously τ_a is bounded and continuous for all a , and $\|\tau_a\|_{\text{op}} \leq 1$. Then we can bound:

$$(55) \quad \|\tau_a f - f\|_{L^2([0, T])} \leq \|\tau_a g - g\|_{L^2([0, T])} + 2\|f - g\|_{L^2([0, T])}$$

Since $C([0, T])$ is dense in $L^2([0, T])$ and since the claim is true for $g \in C([0, T])$, this proves the claim for $f \in L^2([0, T])$. Now we know that $g_n(\cdot, \omega) \rightarrow f(\cdot, \omega)$ holds in $L^2([0, T])$, and also that $\|g_n(\cdot, \omega) - g_n(\cdot, \omega)\|_{L^2([0, T])} \leq 2MT \sup_s K(s)$ holds across ω . Hence, another application of dominated convergence, and Fubini-Tonelli, gives $g_n \rightarrow f$ in $L^2([0, T] \times \Omega)$, as desired. This also implies $\|g_n\|_{L^2([0, T] \times \Omega)} \rightarrow \|f\|_{L^2([0, T] \times \Omega)}$, so each g_n is in $\mathcal{P}_{\mathbb{F}}^2$.

Since each g_n satisfies the hypotheses of Step 1, we can construct a sequence of elementary processes $\{\phi_{k,n}\}_k$ for each n with $\phi_{k,n} \rightarrow g_n$ in $L^2([0, T] \times \Omega)$ as $k \rightarrow \infty$. By diagonalization, then, $\{\phi_{n,n}\}_n$ is a sequence of elementary processes with $\phi_{n,n} \rightarrow f$ in $L^2([0, T] \times \Omega)$ as $n \rightarrow \infty$. This completes the proof of Step 2.

Step 3: If f is in $\mathcal{P}_{\mathbb{F}}^2$, then there exists a sequence of $\{\phi_n\}_n \subseteq \mathcal{E}_{\mathbb{F}}$ with $\phi_n \rightarrow f$ in $L^2([0, T] \times \Omega)$.

For each n define

$$(56) \quad g_n(t, \omega) = \begin{cases} -n & \text{if } f(t, \omega) \leq -n \\ f(t, \omega) & \text{if } -n \leq f(t, \omega) \leq n \\ n & \text{if } f(t, \omega) \geq n \end{cases}.$$

For all (t, ω) , we have $|g_n(t, \omega) - f(t, \omega)|^2 \rightarrow 0$ and $|g_n(t, \omega) - f(t, \omega)|^2 \leq 4|f(t, \omega)|^2$, so dominated convergence convergence gives $g_n \rightarrow f$ in $L^2([0, T] \times \Omega)$. Again we use diagonalization: Since each processes $\{g_n\}_n$ is in $\mathcal{P}_{\mathbb{F}}^2$ and is bounded uniformly across all (t, ω) , we can use Step 2 to construct a sequence of elementary processes $\{\phi_{k,n}\}_k$ that converges in $L^2([0, T] \times \Omega)$ to g_n for each n , and therefore we have that $\{\phi_{n,n}\}_n$ converges $L^2([0, T] \times \Omega)$ to f . This completes Step 3 and hence the proof of the desired result. \square

Definition 2.14. The *Ito integral* is the unique extension of $I_T : \mathcal{E}_{\mathbb{F}} \rightarrow L^2(\Omega)$ to $\mathcal{P}_{\mathbb{F}}^2 \subseteq \mathcal{E}_{\mathbb{F}}$. We denote this operator (by abuse) by

$$(57) \quad I_T(f) = \int_0^T f dB_t$$

All of our hard work at the beginning of the section has paid off by making our definition of the Ito integral rather painless. Even more so, we inherit all of the following properties from the result of Lemma 2.9:

Lemma 2.15. Let $f, g \in \mathcal{P}_{\mathbb{F}}^2$ be any predictable, square integrable processes, and let $\alpha, \beta \in \mathbb{R}$ be any reals. The map $I_T : \mathcal{P}_{\mathbb{F}}^2 \rightarrow L^2(\Omega)$ enjoys the following properties:

- (2.15.1) (Linearity) $I_T(\alpha f + \beta g) = \alpha I_T(f) + \beta I_T(g)$
- (2.15.2) (Mean Zero) $\mathbb{E}[I_T(f)] = 0$
- (2.15.3) (Ito Isometry) $\|I_T(f)\|_{L^2(\Omega)} = \|f\|_{L^2([0, T] \times \Omega)}$
- (2.15.4) (Martingale) The stochastic process $\{I_t(f)\}_t$ is a martingale with respect to the filtration \mathbb{F} .

Proof. First observe that (2.15.1) follows directly from the definition of the extension. Moreover, note that the remaining properties all follow from the fact that I_T , \mathbb{E} , $\mathbb{E}[\cdot | \mathcal{F}_s]$, and the norms $\|\cdot\|_{L^2(\Omega)}$ and $\|\cdot\|_{L^2([0, T] \times \Omega)}$ are continuous as maps to and from the obvious spaces. Indeed, take sequence $\{\phi_n\}_n$ in $\mathcal{E}_{\mathbb{F}}$ converging to $f \in \mathcal{P}_{\mathbb{F}}^2$ with respect to the norm of $L^2([0, T] \times \Omega)$. For (2.15.2):

$$(58) \quad \mathbb{E}[I_T(f)] = \mathbb{E}\left[I_T\left(\lim_{n \rightarrow \infty} \phi_n\right)\right] = \lim_{n \rightarrow \infty} \mathbb{E}[I_T(\phi_n)] = 0.$$

For (2.15.3):

$$(59) \quad \|I_T(f)\|_{L^2(\Omega)} = \left\| I_T\left(\lim_{n \rightarrow \infty} \phi_n\right) \right\|_{L^2(\Omega)}$$

$$(60) \quad = \lim_{n \rightarrow \infty} \|I_T(\phi_n)\|_{L^2(\Omega)}$$

$$(61) \quad = \lim_{n \rightarrow \infty} \|\phi_n\|_{L^2([0,T] \times \Omega)}$$

$$(62) \quad = \left\| \lim_{n \rightarrow \infty} \phi_n \right\|_{L^2([0,T] \times \Omega)} = \|f\|_{L^2([0,T] \times \Omega)}$$

For (2.15.4):

$$(63) \quad \mathbb{E}[I_t(f)|\mathcal{F}_s] = \mathbb{E}\left[I_t\left(\lim_{n \rightarrow \infty} \phi_n\right)|\mathcal{F}_s\right]$$

$$(64) \quad = \lim_{n \rightarrow \infty} \mathbb{E}[I_t(\phi_n)|\mathcal{F}_s]$$

$$(65) \quad = \lim_{n \rightarrow \infty} I_s(\phi_n)$$

$$(66) \quad = I_s\left(\lim_{n \rightarrow \infty} \phi_n\right) = I_s(f)$$

This completes the proof. \square

A drawback to this approach is that there is no reason that an $L^2(\Omega)$ limit should preserve pathwise properties at all. In particular, we do not inherit that the process $t \mapsto I_t(f)$ is continuous for arbitrary $f \in \mathcal{P}_{\mathbb{F}}^2$. In some sense, this is not surprising, since the Brownian motion itself failed to have pathwise continuity before transitioning to a continuous version. Fortunately, we can always make a continuous version of the Ito integral:

Lemma 2.16. For any $f \in \mathcal{P}_{\mathbb{F}}^2$, there exists a stochastic process with continuous sample paths \mathbb{P} -almost surely whose finite-dimensional distributions agree with those of $\int_0^t f dB_s$.

Proof. Take arbitrary $f \in \mathcal{P}_{\mathbb{F}}^2$ and choose a sequence $\{\phi_n\}_n$ in $\mathcal{E}_{\mathbb{F}}$ which converge to f in the sense of $L^2([0, T] \times \Omega)$. Then use Doob's martingale inequality (Lemma A.5) and the Ito isometry to get:

$$(67) \quad \mathbb{P}\left(\sup_{0 \leq t \leq T} |I_t(\phi_n) - I_t(\phi_m)| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} \mathbb{E}\left[|I_T(\phi_n) - I_T(\phi_m)|^2\right]$$

$$(68) \quad \leq \frac{1}{\varepsilon^2} \int_0^T \mathbb{E}\left[|\phi_n - \phi_m|^2\right] ds$$

for all $\varepsilon > 0$. Since the right side goes to zero as $m, n \rightarrow \infty$, we can choose a subsequence $\{\phi_{n_k}\}_k$ such that

$$(69) \quad \mathbb{P}\left(\sup_{0 \leq t \leq T} |I_t(\phi_{n_{k+1}}) - I_t(\phi_{n_k})| > \frac{1}{2^k}\right) < \frac{1}{2^k}$$

holds for all k .

Now we have

$$(70) \quad \sum_{k=1}^{\infty} \mathbb{P} \left(\sup_{0 \leq t \leq T} |I_t(\phi_{n_{k+1}}) - I_t(\phi_{n_k})| > \frac{1}{2^k} \right) \leq \sum_{k=1}^{\infty} \frac{1}{2^k} < \infty,$$

so Borel-Cantelli implies that

$$(71) \quad \mathbb{P} \left(\sup_{0 \leq t \leq T} |I_t(\phi_{n_{k+1}}) - I_t(\phi_{n_k})| > \frac{1}{2^k} \text{ finitely often} \right) = 1.$$

In other words, we have

$$(72) \quad \sum_{k=1}^{\infty} \|I_t(\phi_{n_{k+1}}) - I_t(\phi_{n_k})\|_{\text{sup}} < \infty$$

holding \mathbb{P} -almost surely. But $C([0, T])$ is a Banach space, so we get that $\{I_t(\phi_{n_k})\}_k$ converges uniformly to a continuous limit, \mathbb{P} -almost surely. Then, since the limit of $\{I_t(\phi_{n_k})\}_k$ and the limit of $\{I_t(\phi_n)\}_n$ must have the same finite-dimensional distributions, the proof is complete. \square

2.3. Integral Equations. At the beginning of the section we anticipated that it would be easiest to define an SODE as an integral equation which could, in turn, be interpreted as a sort of differential equation. We are now ready to make this idea precise, and then explore its consequences.

Definition 2.17. An *Ito stochastic differential equation* on the interval $[0, T]$ is an integral equation of the form

$$(73) \quad U_t = u_0 + \int_0^t \mu(s, U_s) ds + \int_0^t \sigma(s, U_s) dB_s.$$

Here, μ is called the *drift*, σ is called the *diffusion*, and $u_0 \in \mathbb{R}$ is called the *initial condition*.

We will often abbreviate (73) by simply using the shorthand:

$$(74) \quad dU_t = \mu(t, U_t) dt + \sigma(t, U_t) dB_t, \quad U_0 = u_0$$

This shorthand, of course, makes no sense on its own; it should really be viewed as a formal symbol which means exactly the same as (73).

We would hope to find conditions on μ, σ , and u_0 for which the SODE (74) has solutions, or perhaps, has a unique solution. However, it is not yet clear what it even means for a stochastic process $\{U_t\}_t$ to be a solution to an SODE. The definition which we will use for the remainder of the paper is as follows:

Definition 2.18. If for any probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which a Brownian motion $\{B_t\}_t$ with canonical filtration $\mathbb{F} = \{F_t^B\}_t$ is defined, there exists a stochastic process $\{U_t\}$ in $\mathcal{P}_{\mathbb{F}}^2$ such that (74) holds \mathbb{P} -almost surely per fixed t and that $\{U_t\}_t$ has continuous sample paths \mathbb{P} -almost surely, then $\{U_t\}_t$ is called a *strong solution* of the SODE.

It is important to remark that Definition 2.18 is not the only meaningful definition of a solution to an SODE (as may already be clear from the terminology). One may alternatively be interested in a weaker notion of solution, which an algebraic perspective can help to disambiguate: A *weak solution* consists of the tuple of a Brownian motion and a stochastic process, and a *strong solution* consists of a map from Brownian motions to stochastic processes. For some SODEs with strong solutions, there is an explicit map that gives the solution process in terms of the Brownian motion, but in general this is not possible. In either case, of course, we demand that the stochastic process satisfies some conditions related to the SODE itself. We will focus only on strong solutions for the remainder of this paper, but there is a great deal of literature studying weak solutions to SODEs in a manner similar to what we are doing here.

Our next task is to come up with criteria under which we have existence and uniqueness of strong solutions to a given SODE. The most classical result towards this end is the following:

Theorem 2.19. Suppose that $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ and $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ satisfy

$$(75) \quad |\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq C_1 |x - y|$$

$$(76) \quad |\mu(t, x)| + |\sigma(t, x)| \leq C_2(1 + |x|),$$

and that the initial condition is deterministic $u_0 \in \mathbb{R}$. Then, the SODE (74) admits a strong solution on $[0, T]$. Moreover, the solution is unique in the sense that any two solutions defined on the same probability space have the same finite-dimensional distributions.

Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which a Brownian motion $\{B_t\}_t$ is defined. The plan is to construct a candidate stochastic process, show that it satisfies the SODE, and then show that it has continuous sample paths \mathbb{P} -almost surely. To construct our candidate, we use Picard's iteration method: Define a sequence of random functions $\{\{U_t^k\}_t\}_k$ by

$$(77) \quad U_t^0 = u_0, \quad U_t^{k+1} = u_0 + \int_0^t \mu(s, U_s^k) ds + \int_0^t \sigma(s, U_s^k) dB_s,$$

Now use Lemma A.1 to bound the $L^2(\Omega)$ -distance between adjacent terms as follows:

$$(78) \quad \mathbb{E} [|U_t^{k+1} - U_t^k|^2] \leq 2\mathbb{E} \left[\left| \int_0^t (\mu(s, U_s^k) - \mu(s, U_s^{k-1})) ds \right|^2 \right] \\ + 2\mathbb{E} \left[\left| \int_0^t (\sigma(s, U_s^k) - \sigma(s, U_s^{k-1})) dB_s \right|^2 \right],$$

We can now bound the two terms on the right side separately. For the first, use Cauchy-Schwarz, the Lipschitz bound (75), and Fubini-Tonelli:

$$(79) \quad \mathbb{E} \left[\left| \int_0^t (\mu(s, U_s^k) - \mu(s, U_s^{k-1})) ds \right|^2 \right] \leq C \int_0^t \mathbb{E} [|U_s^k - U_s^{k-1}|^2] ds$$

For the second, use the Ito isometry and the Lipschitz bound (75):

$$(80) \quad \mathbb{E} \left[\left| \int_0^t (\sigma(s, U_s^k) - \sigma(s, U_s^{k-1})) dB_s \right|^2 \right] \leq C \int_0^t \mathbb{E} [|U_s^k - U_s^{k-1}|^2] ds$$

Combining these, we have shown:

$$(81) \quad \mathbb{E} [|U_t^{k+1} - U_t^k|^2] \leq C \int_0^t \mathbb{E} [|U_s^k - U_s^{k-1}|^2] ds$$

A similar argument using the bound (76) shows:

$$(82) \quad \mathbb{E} [|U_t^1 - U_t^0|^2] = \mathbb{E} \left[\left| \int_0^t \mu(s, u_0) ds + \int_0^t \sigma(s, u_0) dB_s \right|^2 \right]$$

$$(83) \quad \leq C \int_0^t \mathbb{E} [|1 + |u_0| |^2] ds < \infty$$

It follows that the sequence of functions $\{\{\mathbb{E} [|U_t^{k+1} - U_t^k|^2]\}_t\}_k$ satisfies the hypotheses of Gronwall's inequality (Lemma A.3), and hence we obtain the bound:

$$(84) \quad \mathbb{E} [|U_t^{k+1} - U_t^k|^2] \leq C_1 \frac{(C_2 T)^k}{k!}$$

for all $t \in [0, T]$. Therefore, for each t , the random variables $\{U_t^k\}_k$ have an $L^2(\Omega)$ limit which we denote by U_t . Combining these together for all t , we have constructed out candidate stochastic process $\{U_t\}_t$.

Next we show that the stochastic process $\{U_t\}_t$ has continuous sample paths \mathbb{P} -almost surely. By Lemma 2.16, the function $\{U_t^k\}_t$ is continuous for each k . Moreover, by Lemma 2.15, the function $\{U_t^k\}_t$ is a martingale for each k , hence $\{|U_s^{k+1} - U_s^k|^2\}_t$ is a submartingale for each k . Applying Doob's martingale inequality (Lemma A.5) and the bound (84), we get

$$(85) \quad \mathbb{P} \left(\sup_{0 \leq t \leq T} |U_t^{k+1} - U_t^k|^2 \geq \varepsilon \right) \leq \frac{\mathbb{E} [|U_T^{k+1} - U_T^k|^2]}{\varepsilon} \leq C_1 \frac{(C_2 T)^k}{\varepsilon \cdot k!}$$

for all $\varepsilon > 0$. Note in particular that we get

$$(86) \quad \sum_{k=1}^{\infty} \mathbb{P} \left(\sup_{0 \leq t \leq T} |U_t^{k+1} - U_t^k|^2 \geq 2^{-k} \right) \leq C_1 \sum_{k=1}^{\infty} \frac{(2C_2 T)^k}{k!} < \infty$$

Therefore Borel-Cantelli gives

$$(87) \quad \mathbb{P} \left(\sup_{0 \leq t \leq T} |U_t^{k+1} - U_t^k|^2 \geq 2^{-k} \text{ infinitely often} \right) = 0$$

In other words, for almost all ω , we have some $K(\omega)$ such that $k \geq K(\omega)$ implies

$$(88) \quad \|U_t^{k+1} - U_t^k\|_{\sup} < 2^{-k/2}$$

This proves that we have, \mathbb{P} -almost almost surely:

$$(89) \quad \sum_{k=1}^{\infty} \|U_t^{k+1} - U_t^k\|_{\sup} \leq \sum_{k=1}^{\infty} 2^{-k/2} < \infty$$

Since $C([0, T])$ is a Banach space, this implies that the sequence $\{\{U_t^k\}_t\}_k$ converges uniformly to its limit, which proves that $\{U_t\}_t$ is continuous \mathbb{P} -almost surely.

Next we verify that the stochastic process $\{U_t\}_t$ satisfies the SODE \mathbb{P} -almost surely per fixed t . We already know that the left side of (77) converges \mathbb{P} -almost surely to the left side of (74), hence it suffices to show that the right side of (77) converges in $L^2(\Omega)$ to the right side of (74). Indeed, we can prove this for each term individually: For the first, use Cauchy-Schwarz, Fubini-Tonelli, and the Lipschitz bound (75) to get

$$(90) \quad \mathbb{E} \left[\left| \int_0^t (\mu(s, U_s^k) - \mu(s, U_s)) ds \right|^2 \right] \leq C \int_0^t \mathbb{E} [|U_s^k - U_s|^2] ds.$$

Going back to (84), we see that $\{U_t^k\}_k \rightarrow \{U_t\}_t$ in $L^2(\Omega)$ holds uniformly in t , hence the right side of the above goes to zero as $k \rightarrow \infty$. For the second term, we use the Ito isometry and the Lipschitz bound (75) and then apply the same argument. This proves that $\{U_t\}$ satisfies the SODE.

Finally, it remains to check uniqueness. For $\{U_t^1\}_t$ and $\{U_t^2\}_t$ both solutions to the SODE, we can argue as above to get:

$$(91) \quad z(t) = \sup_{0 \leq s \leq t} \mathbb{E} [|U_s^1 - U_s^2|^2]$$

$$(92) \quad \leq C \sup_{0 \leq s \leq t} \int_0^t \mathbb{E} [|\mu(s, U_s^1) - \mu(s, U_s^2)|^2] ds$$

$$(93) \quad \leq C \int_0^t z(s) ds$$

By Gronwall's inequality (Lemma A.4), we get

$$(94) \quad z(T) = \sup_{0 \leq t \leq T} \mathbb{E} [|U_t^1 - U_t^2|^2] = 0$$

This proves that, for each t , we have $\mathbb{P}(U_t^1 = U_t^2) = 1$, which completes the proof. \square

The result of Theorem 2.19 will be powerful enough to guarantee the existence and uniqueness of strong solutions in many cases of interest. However, it does nothing to help one find an explicit strong solution of an SDE even if it exists. Our main computational tool, in this regard, is the generalization of the fundamental theorem of calculus for the Ito calculus:

Lemma 2.20 (Ito's Lemma). Take any function $g \in C^2([0, T] \times \mathbb{R})$, and let $\{U_t\}_t$ be a stochastic process given by

$$(95) \quad U_t = U_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s$$

with $\mu \in \mathcal{P}_{\mathbb{F}}^1$ and $\sigma \in \mathcal{P}_{\mathbb{F}}^2$. Then the stochastic process $\{g(t, U_t)\}_t$ can be written in integral form as:

$$(96) \quad g(t, U_t) = g(0, U_0) + \int_0^t \left(\frac{\partial g}{\partial t}(s, U_s) + \mu_s \frac{\partial g}{\partial x}(s, U_s) + \frac{|\sigma_s|^2}{2} \frac{\partial^2 g}{\partial x^2}(s, U_s) \right) ds$$

$$(97) \quad + \int_0^t \frac{\partial g}{\partial x}(s, U_s) \sigma_s dB_s$$

Proof. We provide a heuristic (non-rigorous) proof of the claim in differential form. That is, suppose that $\{U_t\}_t$ is a stochastic process given by

$$(98) \quad dU_t = \mu_t dt + \sigma_t dB_t$$

Now write out the Taylor expansion for $g(t, x)$ near $(t, x) = (0, 0)$:

$$(99) \quad g(t, x) = g(0, 0) + \frac{\partial g}{\partial t}(t, x) \Delta t$$

$$(100) \quad + \frac{\partial g}{\partial x}(t, x) \Delta x + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, x) (\Delta x)^2 + R(t)$$

where the remainder term R is on the order of

$$(101) \quad R(t) = O((\Delta t)(\Delta x)) + O((\Delta t)(\Delta x)^2) + \sum_{i \geq 2} \sum_{j \geq 3} O((\Delta t)^i (\Delta x)^j).$$

Now we substitute $x = U_t$ and take the limit $\Delta t \rightarrow dt$. We show that R becomes negligible according to the rules of (23): First, compute

$$(102) \quad (dU_t)^2 = (\mu_t dt)^2 + 2\mu_t \sigma_t dt dB_t + (\sigma_t dB_t)^2 = \sigma_t^2 dt,$$

which, in other words, means $dU_t = \sigma_t(dt)^{1/2}$. Then plug this into (101) to get

$$(103) \quad R(t) \rightarrow O((dt)^{3/2}) + O((dt)^2) + \sum_{i \geq 2} \sum_{j \geq 3} O((dt)^{i+\frac{j}{2}}) = 0.$$

Now we see that, in the limit, (99) becomes

$$(104) \quad dg(t, U_t) = \frac{\partial g}{\partial t}(t, U_t) dt$$

$$(105) \quad + \frac{\partial g}{\partial x}(t, U_t) dU_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, U_t) (dU_t)^2.$$

Plugging in the formulas for dU_t and $(dU_t)^2$, we get:

$$(106) \quad dg(t, U_t) = \left(\frac{\partial g}{\partial t}(t, U_t) + \mu_t \frac{\partial g}{\partial x}(t, U_t) + \frac{|\sigma_t|^2}{2} \frac{\partial^2 g}{\partial x^2}(t, U_t) \right) dt$$

$$(107) \quad + \frac{\partial g}{\partial x}(t, U_t) \sigma_t dB_t.$$

Rewriting this in integral form gives the desired result. \square

2.4. Examples. At last we have the tools to briefly study a few interesting examples of SODEs.

Example 2.21 (Geometric Brownian Motion). Consider the SODE

$$(108) \quad dU_t = \mu U_t dt + \sigma U_t dB_t, \quad U_0 = u_0.$$

on the time interval $[0, T]$. Since the drift $\mu(t, x) = \mu x$ and the diffusion $\sigma(t, x) = \sigma x$ satisfy the hypotheses of Theorem 2.19, we see that the SODE admits strong solutions. We can actually go one step further and compute the solution explicitly: Consider the function $g(t, x) = \log(x)$ and use Ito's Lemma to compute:

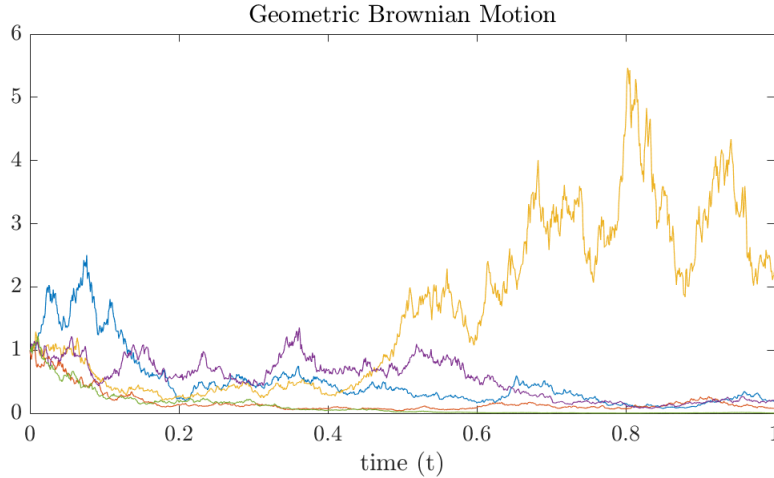
$$(109) \quad \log(U_t) = \log(u_0) + \int_0^t \left(\mu - \frac{|\sigma|^2}{2} \right) ds + \int_0^t \sigma dB_s$$

$$(110) \quad = \log(u_0) + \left(\mu - \frac{|\sigma|^2}{2} \right) t + \sigma B_s.$$

Therefore, the solution to the SODE is given explicitly by

$$(111) \quad U_t = u_0 \exp \left(\left(\mu - \frac{|\sigma|^2}{2} \right) t + \sigma B_s \right).$$

We can plot a few sample paths of the solution to this SODE, say, using the parameters $\mu = -1, \sigma = 2, u_0 = 1$. The plot is as follows:



Example 2.22 (Brownian Bridge). Consider the SODE

$$(112) \quad dU_t = -\frac{U_t}{1-t} dt + dB_t, \quad U_0 = 0,$$

on the time interval $[0, 1]$. Theorem 2.19 guarantees the existence of strong solutions, since the coefficient functions satisfy the hypotheses of the theorem. However, this result does not give us a good qualitative understanding of the solutions to the SODE.

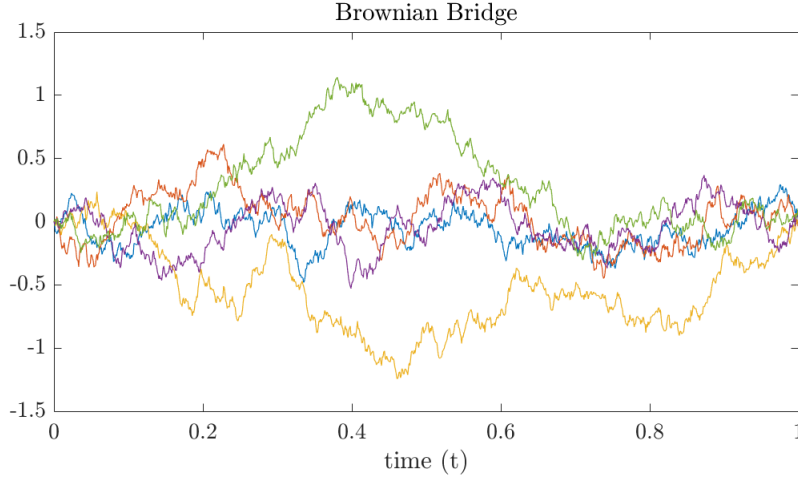
To accomplish this, observe that the scaling of the noise does not depend on the value of U_t . Hence, we expect that we may be able to eliminate the drift term by instead considering an equivalent SODE in terms of the stochastic process $Y_t = f(t)U_t$ for some deterministic function f . To find the right choice of f , we use Ito's Lemma and compute:

$$(113) \quad Y_t = \int_0^t U_s \left(f'(s) - \frac{f(s)}{1-s} \right) ds + \int_0^t f(s) dB_s$$

We are successful if we can choose f such that $f'(s) - f(s)/(1-s) = 0$ holds. This is nothing more than a linear ODE, so we can find its solution to be $f(s) = C/(1-s)$ for any constant C . Now modify (113) by substituting in $Y_t = f(t)U_t$ and $f(t) = C/(1-t)$ for $C \neq 0$, resulting in:

$$(114) \quad U_t = (1-t) \int_0^t \frac{1}{1-s} dB_s.$$

This is a very useful characterization of the solution to the SODE for the Brownian bridge, and we can easily read off some of its qualitative properties: We have $U_0 = U_1 = 0$, and U_t has the same smoothness as the Brownian motion. This properties can be seen in the following plots, which shows a few sample paths of the Brownian bridge:



As the plot suggests, it can be proven that the law of the Brownian bridge is nothing more than the law of a Brownian motion $\{B_t\}_t$ conditioned on the event $\{\omega \in \Omega : B_0(\omega) = B_1(\omega) = 0\}$.

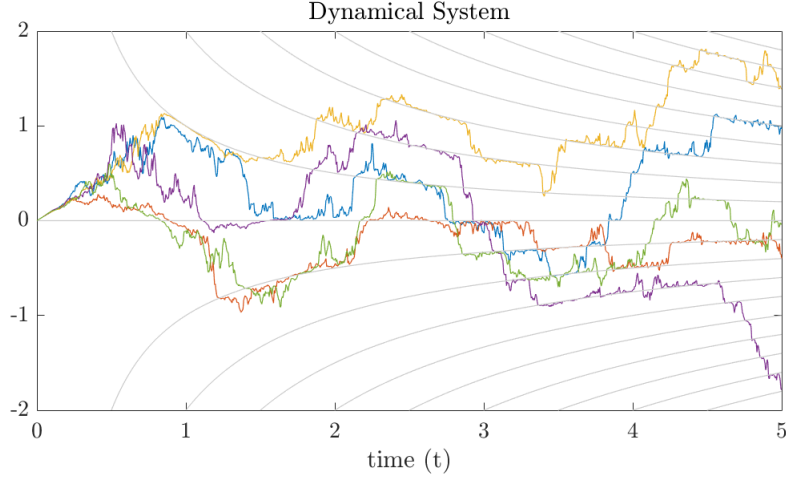
Our next example illustrates the interplay between SODEs and dynamical systems, and provides an example where we do not have an analytical solution to the SODE in terms of the driving Brownian motion.

Example 2.23 (Dynamical System). Consider the SODE

$$(115) \quad dU_t = \cos(\pi t)dt + \sin(\pi t U_t)dB_t, \quad U_0 = 0.$$

The coefficients satisfy all of the hypotheses of Theorem 2.19, so this SODE admits strong solutions. However, it is not clear whether or not we can write them explicitly in as a function of the driving Brownian motion at all. Instead, we can hope to get a qualitative understanding of the sample paths of the solution to the SODE.

Note that we have $\sin(\pi t U_t) \approx 0$ if and only if $U_t \approx k/t$ for some integer k , i.e. the noise has a weak effect on the system when U_t is near one of these hyperbolas. In other words, we expect that the solutions stay roughly constant along these hyperbolas for at least a short amount of time. The following plot of a few sample paths corroborates this guess:



All of the plots presented in this section were generated by simulating solutions to these SODEs, even when an analytical solution was known. For the exact scripts used to generate these plots, we refer the reader to Appendix B, but for a mathematical understanding of how such a task is even possible, we proceed to the next section.

3. NUMERICAL SODE

The goal of this section is to introduce the main ideas of numerical analysis of SODEs. That is, how can one simulate the solution to a given SODE in a computer and guarantee that the results are close to the true solution? For the sake of simplicity, we restrict our attention to time-discrete, one-step approximations with a fixed time step for 1-dimensional SODEs.

For the remainder of this section, let us consider the arbitrary SDE (74) on the time interval $[0, T]$. Suppose that μ and σ satisfy the conditions to admit strong solutions, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be any probability space on which a Brownian motion $\{B_t\}_t$ is defined. Whenever a constant N called the *resolution* is fixed, we set $\Delta t = T/N$ and $t_j = j\Delta t$ for $j = 0, \dots, N$. Also let $j(s) = \inf\{j \leq t : j = 0, 1, \dots, N\} = \lfloor s/\Delta t \rfloor$ for any $s \in [0, T]$. Additionally, we will abbreviate $\Delta B_j(\omega) = B_{t_{j+1}}(\omega) - B_{t_j}(\omega)$.

3.1. General Principles. Some terminology will be useful throughout the remainder of this section.

Definition 3.1. A *numerical scheme* is an algorithm, which, for each positive integer N , returns a discrete-time stochastic process $\{V_j^N\}_{j=0}^N$ satisfying $V_0^N = u_0$ defined on the same probability space as the SODE.

We will often abbreviate this notation by writing $\{V_j^N\}_j$ to make the parallelism with $\{U_t\}_t$ more clear. While we really only need to know the values of $\{V_j^N\}_j$ at the discrete time values, one can imagine extending to some continuous-time function of interest: As such, we may write $\{\bar{V}_t^N\}_t$ and $\{\hat{V}_t^N\}_t$ to denote the piecewise-constant and piecewise-linear interpolations of $\{V_j^N\}_j$, respectively. As a sanity check, observe that we have

$$(116) \quad V_j^N = \bar{V}_{t_j}^N = \hat{V}_{t_j}^N$$

for any $N > 0$ and any $j = 0, 1, \dots, N$. Of primary interest to us will be the piecewise linear case.

In general, an extremely wide class of algorithms are considered to be numerical schemes. However, we will focus our attention on the following collection of numerical schemes:

Definition 3.2. A *one-step numerical scheme* is a numerical scheme completely determined by some function $\Psi : [0, T] \times \mathbb{R} \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$. Specifically, each $\omega \in \Omega$ gives rise to $\{V_j^N(\omega)\}_j$ via the recursive structure:

$$(117) \quad V_0^N(\omega) = u_0$$

$$(118) \quad V_{j+1}^N(\omega) = V_j^N(\omega) + \Psi(t_j, V_j^N(\omega), \Delta t, \Delta B_j(\omega))$$

The appeal of one-step schemes is their simplicity: Since they only depend on the Brownian motion through its increments (which are normally distributed) we can use a random walk to simulate the path of the Brownian motion.

It still remains to precisely define what is meant by a numerical scheme for an SODE being “close” to the true solution. Indeed, many such notions exist, but we will focus our attention on the following:

Definition 3.3. A numerical scheme $\{V_j^N\}_j$ is *strongly quadratically convergent of order $\gamma > 0$* if

$$(119) \quad \mathbb{E} \left[\sup_{0 \leq t \leq T} |U_t - \hat{V}_t^N|^2 \right] = O(N^{-2\gamma})$$

3.2. Choosing the Resolution. The interpretation of a numerical scheme with strong order of convergence equal to γ is the following: Given a numerical approximation of the solution at an initial resolution, increasing the resolution by a factor of K increases the precision of the approximation by a factor of K^γ . However, the order of convergence tells us nothing about how to actually choose the resolution. In problems where the scale of the drift and diffusion are not known, there may be no a priori reason to set the resolution at a certain value.

What the strong order of convergence does well is that it gives us control on how well the scheme improves as we improve the resolution. But this task demands some caution: We must always make sure that the initial random walk and the increased-resolution random walk are generated according to the same ω . Otherwise, there is

no reason at all that the output of the numerical scheme for one should be anywhere close to the output of the numerical scheme for the other.

Therefore, we need a way to simulate a random walk of high resolution conditional on the event that the random walk of a lower resolution passed through some specified points—The rest of this subsection is dedicated to fleshing out this idea. To get started, let $(\Omega, \mathcal{F}, \mathbb{P})$ be any probability space on which a Brownian motion $\{B_t\}_t$ is defined, and consider the following as our basic tool:

Lemma 3.4. Let $t_1 < t_2$ be fixed times and $b_1, b_2 \in \mathbb{R}$ any reals. For $\bar{t} = (t_1 + t_2)/2$, it follows that the conditional distribution of $B_{\bar{t}}$ with respect to $\{B_{t_1} = b_1, B_{t_2} = b_2\}$ is normal with mean $(b_1 + b_2)/2$ and variance $(t_2 - t_1)/2$.

Proof. Define the random variable

$$(120) \quad X = B_{\bar{t}} - \frac{B_{t_1} + B_{t_2}}{2}.$$

We can easily see that X is normal with mean zero and variance $(t_1 - t_2)/2$. Moreover, we can compute $\text{Cov}(X, B_{t_1}) = 0$ and $\text{Cov}(X, B_{t_2}) = 0$, so it follows that X is independent of both B_{t_1} and B_{t_2} .

In particular, X is independent of the event

$$(121) \quad A = \{\omega \in \Omega : B_{t_1}(\omega) = b_1, B_{t_2}(\omega) = b_2\}$$

since A is $\sigma(B_{t_1}, B_{t_2})$ -measurable.

Now compute the characteristic function of $B_{\bar{t}}$ conditional on $\sigma(A)$: For any $s \in \mathbb{R}$, we have

$$(122) \quad \mathbb{E} \left[\exp(isB_{\bar{t}}) \middle| \sigma(A) \right]$$

$$(123) \quad = \mathbb{E} [\exp(isX)] \mathbb{E} \left[\exp \left(is \frac{B_{t_1} + B_{t_2}}{2} \right) \middle| \sigma(A) \right].$$

But, of course, the conditional distribution of $(B_{t_1} + B_{t_2})/2$ given A is just $(b_1 + b_2)/2$, so we conclude

$$(124) \quad B_{\bar{t}} | \sigma(A) \stackrel{d}{=} X + \frac{b_1 + b_2}{2},$$

which proves the claim. \square

Definition 3.5. Construct a sequence of stochastic processes $\{\{R_j^n\}_{j=0}^{2^n}\}_{n=0}^\infty$ as follows: Set $R_0^0 = 0$ and draw $R_0^1 \sim N(0, T)$. Then recursively set

$$(125) \quad R_j^{n+1} = \begin{cases} R_{j/2}^n & \text{if } j \text{ even} \\ (R_{(j+1)/2}^n + R_{(j-1)/2}^n)/2 + X_{(j+1)/2}^n & \text{if } j \text{ odd} \end{cases}$$

where for each n , the collection $\{X_j^n\}_{j=1}^{2^n}$ are drawn from $N(0, T/2^{n+2})$, independent from each other and independent from $\{\{R_j^m\}_{j=0}^{2^m}\}_{m=0}^n$.

Corollary 3.6. For each n , the discrete-time stochastic process $\{R_j^n\}_{j=0}^{2^n}$ is equal in law to the sampled Brownian motion $\{B_{jT/2^n}\}_{j=0}^{2^n}$.

Proof. The sampled Brownian motion is characterized as the unique mean-zero Gaussian process with $\text{Cov}(B_{iT/2^n}, B_{jT/2^n}) = \min\{i, j\}T/2^n$. Hence, we prove by induction on n that $\{R_j^n\}_{j=0}^{2^n}$ satisfies these conditions. That this is a mean-zero Gaussian process is immediate, so we proceed to the inductive step.

Take any $i, j \in \{0, 1, \dots, 2^{n+1}\}$ and assume without loss of generality that $i \leq j$. If both i and j are even, then we can compute:

$$(126) \quad \text{Cov}(R_i^{n+1}, R_j^{n+1}) = \text{Cov}\left(R_{i/2}^n, R_{j/2}^n\right) = i \frac{T}{2^{n+1}}$$

as needed.

If i is even and j is odd, then in particular we have $i \leq j-1$ and we can compute:

$$(127) \quad \text{Cov}(R_i^{n+1}, R_j^{n+1}) = \text{Cov}\left(R_{i/2}^n, \frac{R_{(j-1)/2}^n + R_{(j+1)/2}^n}{2} + X_{(j+1)/2}^n\right)$$

$$(128) \quad = \frac{1}{2} \text{Cov}\left(R_{i/2}^n, R_{(j-1)/2}^n\right) + \frac{1}{2} \text{Cov}\left(R_{i/2}^n, R_{(j+1)/2}^n\right)$$

$$(129) \quad = \min\left\{\frac{i}{2}, \frac{j-1}{2}\right\} \frac{T}{2^{n+1}} + \min\left\{\frac{i}{2}, \frac{j+1}{2}\right\} \frac{T}{2^{n+1}}$$

$$(130) \quad = i \frac{T}{2^{n+1}}.$$

If i odd and j even, then we also have $i \leq j-1$, so an analogous proof goes through.

The only remaining case is both i and j odd. Let us compute:

$$(131) \quad \text{Cov}(R_i^{n+1}, R_j^{n+1})$$

$$(132) \quad = \text{Cov}\left(\frac{R_{(i-1)/2}^n + R_{(i+1)/2}^n}{2} + X_{(i+1)/2}^n, \frac{R_{(j-1)/2}^n + R_{(j+1)/2}^n}{2} + X_{(j+1)/2}^n\right)$$

$$(133) \quad = \text{Cov}\left(\frac{R_{(i-1)/2}^n + R_{(i+1)/2}^n}{2}, \frac{R_{(j-1)/2}^n + R_{(j+1)/2}^n}{2}\right)$$

$$(134) \quad + \text{Cov}\left(X_{(i+1)/2}^n, X_{(j+1)/2}^n\right)$$

Now we split into further cases. If $i = j$, then the first covariance term equals

$$(135) \quad \text{Var}\left(\frac{R_{(i-1)/2}^n + R_{(i+1)/2}^n}{2}\right)$$

$$(136) \quad = \frac{1}{4} \text{Var}\left(R_{(i-1)/2}^n\right) + \frac{1}{4} \text{Var}\left(R_{(i+1)/2}^n\right) + \frac{1}{2} \text{Cov}(R_{(i-1)/2}^n, R_{(i+1)/2}^n)$$

$$(137) \quad = i \frac{T}{2^{n+1}} - \frac{T}{2^{n+2}}$$

and the second equals

$$(138) \quad \text{Var}\left(X_{(i-1)/2}^n\right) = \frac{T}{2^{n+2}},$$

so summing these together gives the desired result.

Otherwise we have $i < j$, but, both being odd, this implies $i+1 \leq j-1$. Then the second covariance term vanishes and we are left with:

$$(139) \quad \text{Cov} \left(\frac{R_{(i-1)/2}^n + R_{(i+1)/2}^n}{2}, \frac{R_{(j-1)/2}^n + R_{(j+1)/2}^n}{2} \right)$$

$$(140) \quad = \frac{1}{4} (\text{Cov}(R_{(i-1)/2}^n, R_{(j-1)/2}^n) + \text{Cov}(R_{(i-1)/2}^n, R_{(j+1)/2}^n)$$

$$(141) \quad + \text{Cov}(R_{(i+1)/2}^n, R_{(j-1)/2}^n) + \text{Cov}(R_{(i+1)/2}^n, R_{(j+1)/2}^n))$$

$$(142) \quad = i \frac{T}{2^{n+1}}$$

This completes the induction, and finishes the proof. \square

We now linearly interpolate each $\{R_j^n\}_{j=0}^{2^n}$ to get $\{\hat{R}_t^n\}_t$, a $C([0, T])$ -valued random variable. The advantage of this construction is stated below:

Theorem 3.7. The stochastic process $\{\{\hat{R}_t^n\}_t\}_{n=0}^\infty$ is an \mathbb{N} -indexed, $C([0, T])$ -valued Markov chain such that, \mathbb{P} -almost surely, we have $\{\hat{R}_t^n\}_t$ converging uniformly to $\{B_t\}_t$ as $n \rightarrow \infty$.

Proof. That $\{\{\hat{R}_t^n\}_t\}_{n=0}^\infty$ is a Markov chain defined on $(\Omega, \mathcal{F}, \mathbb{P})$ follows immediately from Lemma 3.4 and Definition 3.5. To see that the desired convergence holds, define the following random variable for each n :

$$(143) \quad Z_n = \|\hat{R}_t^{n+1} - \hat{R}_t^n\|_{\text{sup}} = \max_{j=1,1,\dots,2^n} |X_j^n|$$

By Lemma A.2, we have $\mathbb{E}[Z_n] \leq C 2^{-n} (n^{1/2} + n^{-1/2})$ for some constant C . Hence, the random variable $Z = \sum_{n=0}^\infty Z_n$ satisfies

$$(144) \quad \mathbb{E}[Z] \leq C \sum_{n=1}^\infty \frac{1}{2^n} \left(\sqrt{n} + \frac{1}{\sqrt{n}} \right) < \infty$$

But since Z is nonnegative, its finite expectation implies that it is finite \mathbb{P} -almost surely. Hence, we have shown

$$(145) \quad \sum_{n=1}^\infty \|\hat{R}_t^{n+1}(\omega) - \hat{R}_t^n(\omega)\|_{\text{sup}} < \infty,$$

so $\{\hat{R}_t^n(\omega)\}_t$ possesses a limit in the Banach space $C([0, T])$ for almost all ω . In other words, the sequence $\{\hat{R}_t^n(\omega)\}_t$ of $C([0, T])$ -valued random variables converges to a limit \mathbb{P} -almost surely, and, in particular, in law. But Theorem 2.5 and Lemma 3.6 together imply that $\{\hat{R}_t^n\}_t \rightarrow \{B_t\}_t$ holds in law in $C([0, T])$. Since the limit in law of a sequence of random variables is unique, this implies that $\{\hat{R}_t^n(\omega)\}_t \rightarrow \{B_t\}_t$ holds \mathbb{P} -almost surely in $C([0, T])$. \square

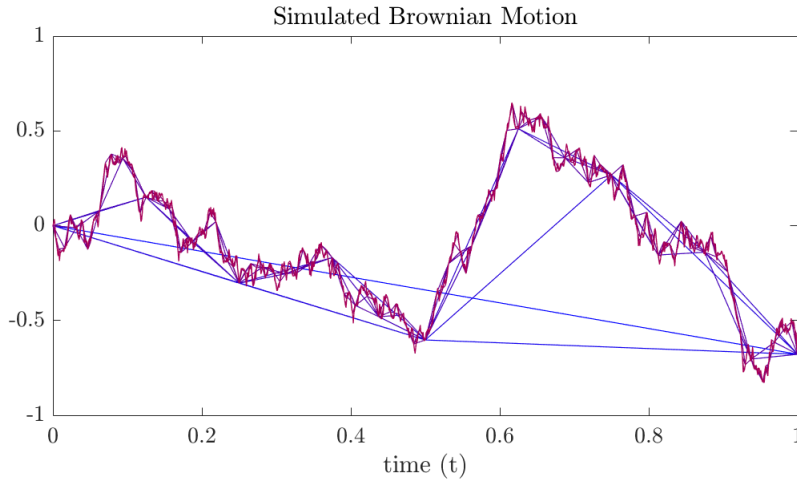
Theorem 3.7 is useful because it guarantees that we can run this refinement procedure and the resulting random variables converge to a Brownian motion. The Markov property is especially helpful, since it means we can generate the next approximation to the Brownian motion given only the current approximation. From this perspective we propose a quite general method for “choosing” the resolution $N = 2^n$ probabilistically: Use any \mathbb{N} -valued stopping time which is adapted to the canonical filtration of $\{\{\hat{R}_t^n\}_t\}_{n=0}^\infty$. This quite flexible method allows us to come

up with many rules for determining when a random walk is “close enough” to a Brownian motion.

As an example, considering the following stopping time, for some tolerance parameter $\varepsilon > 0$:

$$(146) \quad \tau^\varepsilon = \min\{n > 0 : \|\hat{R}_t^n - \hat{R}_t^{n-1}\|_{\text{sup}} < \varepsilon\}.$$

Let’s run a computational example, taking $T = 1$ and $\varepsilon = 2^{-4}$. We then generate the following plot, which displays the successive approximations to the Brownian motion colored by their resolution (blue being the lowest and red being the highest):



This exhibits the behavior we expected, and it seems to have stopped at a fairly reasonable resolution.

A slightly more useful stopping time would be as one that depends on the convergence, not of the Brownian motion, but of the output of a numerical scheme. For example, letting \hat{V}_t^N be any one-step numerical scheme, we might set:

$$(147) \quad \tau^\varepsilon = \min\{n > 0 : \|\hat{V}_t^{2^n}(\hat{R}_t^n) - \hat{V}_t^{2^{n-1}}(\hat{R}_t^{n-1})\|_{\text{sup}} < \varepsilon\}.$$

where $\hat{V}_t^{2^n}(\hat{R}_t^n)$ represents the output of $\hat{V}_t^{2^n}$ when the increments to Φ are taken from \hat{R}_t^n . To further explore this idea, we dedicate the next subsection towards studying a few one-step numerical schemes.

3.3. Numerical Schemes. Let’s now explore some basic numerical schemes for solving a given SODE. Throughout the remainder of this subsection, suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space on which a Brownian motion $\{B_t\}_t$ is defined, and consider the SODE (74).

Definition 3.8. The *Euler-Maruyama scheme* is the one-step numerical scheme determined by

$$(148) \quad \Psi(t, x, \Delta t, \Delta x) = \mu(t, x)\Delta t + \sigma(t, x)\Delta x.$$

Theorem 3.9. If for $f = \mu$ and σ we have

$$(149) \quad |f(t, x) - f(t, y)| \leq C_1 |x - y|$$

$$(150) \quad |f(t, x)| \leq C_2(1 + |x|)$$

$$(151) \quad |f(t, x) - f(s, x)| \leq C_3(1 + |x|)\sqrt{|t - s|},$$

then the Euler-Maruyama scheme is strongly quadratically convergent of order $\gamma = 1/2$.

The proof relies on a two intermediate results which we state and prove below:

Lemma 3.10. Under the hypotheses of Theorem 3.9, we have

$$(152) \quad \mathbb{E} \left[(1 + |U_t|)^2 \right] \leq C$$

for some constant C not depending on $t \in [0, T]$.

Proof. Methods similar to the proof of Theorem 2.19 show the following:

$$(153) \quad \mathbb{E} \left[|U_t|^2 \right] = \mathbb{E} \left[\left| \int_0^t \mu(s, U_s) ds + \int_0^t \sigma(s, U_s) dB_s \right|^2 \right]$$

$$(154) \quad \leq C \int_0^t \mathbb{E} \left[(1 + |U_s|)^2 \right] ds$$

$$(155) \quad = C_1 + C_2 \int_0^t \mathbb{E} \left[|U_s|^2 \right] ds$$

Therefore, we can apply Gronwall (Lemma A.4) to the function

$$(156) \quad z(t) = \sup_{0 \leq s \leq t} \mathbb{E} \left[|U_s|^2 \right]$$

to get

$$(157) \quad z(T) \leq \sup_{0 \leq s \leq T} \mathbb{E} \left[|U_s|^2 \right] \leq C_1 \exp(C_2 T)$$

as desired. □

Lemma 3.11. Under the hypotheses of Theorem 3.9, we have

$$(158) \quad \int_0^T \mathbb{E} \left[|U_r - U_{t_j(r)}|^2 \right] dr \leq C \Delta t$$

for some constant C .

Proof. Observe that we can write

$$(159) \quad U_r - U_{t_j(r)} = \int_{t_j(r)}^r \mu(s, U_s) ds + \int_{t_j(r)}^r \sigma(s, U_s) dB_s$$

As always, we can bound this as:

$$(160) \quad \mathbb{E} \left[|U_r - U_{t_{j(r)}}|^2 \right] \leq C \int_{t_{j(r)}}^r \mathbb{E} \left[(1 + |U_s|)^2 \right] ds$$

By Lemma 3.10, the integrand is finite, and the limits of integration are bounded like $r - t_{j(r)} \leq \Delta t$. Hence, the result is proved. \square

Proof of Theorem 3.9. Define the function

$$(161) \quad z_N(t) = \mathbb{E} \left[\sup_{0 \leq s \leq t} |U_s - \hat{V}_s^N|^2 \right].$$

Our goal is find a bound on z_N which goes to zero like $O(N^{-1})$ uniformly in $t \in [0, T]$, and we will prove this, of course, via the Gronwall inequalities.

Building towards that point, note that we can write:

$$(162) \quad |U_s - \hat{V}_s^N|^2 = \left| \int_0^s \left(\mu(r, U_r) - \mu(t_{j(r)}, \hat{V}_{t_{j(r)}}^N) \right) dr \right.$$

$$(163) \quad \left. + \int_0^s \left(\sigma(r, U_r) - \sigma(t_{j(r)}, \hat{V}_{t_{j(r)}}^N) \right) dB_r \right|^2.$$

The terms on the right side are not quite in the form of our hypotheses, but we can massage the into the right quite easily. Note that we can write the integrand as the sum of three differences

$$(164) \quad f(r, U_r) - f(t_{j(r)}, \hat{V}_{t_{j(r)}}^N) = g_1^f(t, \omega) + g_2^f(t, \omega) + g_3^f(t, \omega)$$

where we have defined

$$(165) \quad g_1^f(t, \omega) = f(r, U_r) - f(t_{j(r)}, U_r)$$

$$(166) \quad g_2^f(t, \omega) = f(t_{j(r)}, U_r) - f(t_{j(r)}, U_{t_{j(r)}})$$

$$(167) \quad g_3^f(t, \omega) = f(t_{j(r)}, U_{t_{j(r)}}) - f(t_{j(r)}, \hat{V}_{t_{j(r)}}^N)$$

for either $f = \mu$ or σ . The benefit of this notation is that the hypotheses (149), (150), and (151) are exactly of the form $|g_i^f(t, \omega)| \leq h_i(t, \omega)$ for h_i some specified function for $i = 1, 2, 3$. Specifically, we have:

$$(168) \quad h_1(t, \omega) = C_3(1 + |U_r|)\sqrt{r - t_{j(r)}}$$

$$(169) \quad h_2(t, \omega) = C_1|U_r - U_{t_{j(r)}}|$$

$$(170) \quad h_3(t, \omega) = C_1|U_{t_{j(r)}} - \hat{V}_{t_{j(r)}}^N|.$$

Now we replace the integrands of (162) with the sum of g_i^f for $i = 1, 2, 3$ for both $f = \mu, \sigma$. Using Lemma A.1, we can bound the square of the sum by the sum of squares, and this leaves us with the bound

$$(171) \quad z_N(t) \leq 6 \sum_{i=1}^3 (R_i^\mu(t) + R_i^\sigma(t)),$$

where $R_i^f(t)$ is defined by the following for $f = \mu$:

$$(172) \quad R_i^\mu(t) = \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| \int_0^s g_i^\mu(r, \omega) dr \right|^2 \right],$$

and the following for $f = \sigma$:

$$(173) \quad R_i^\sigma(t) = \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| \int_0^s g_i^\sigma(r, \omega) dB_r \right|^2 \right].$$

It now remains to bound the R_i^f terms individually. While each of these will be handled differently, their first few steps are all the same so we will perform them at once. For any $i = 1, 2, 3$, we can bound the $f = \mu$ terms as follows: First use Cauchy-Schwarz, then use $s \leq T$, then use that the integrand is nonnegative to form an upper bound by increasing the upper limit of integration to T . Finally, use Fubini-Tonelli to push the expectation inside the integral, and use the bound (168) to get:

$$(174) \quad R_i^\mu(t) = \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| \int_0^s g_i^\mu(r, \omega) dr \right|^2 \right]$$

$$(175) \quad \leq \mathbb{E} \left[\sup_{0 \leq s \leq t} s \int_0^s |g_i^\mu(r, \omega)|^2 dr \right]$$

$$(176) \quad \leq \mathbb{E} \left[T \int_0^T |g_i^\mu(r, \omega)|^2 dr \right]$$

$$(177) \quad \leq T \int_0^T \mathbb{E} \left[|g_i^\mu(r, \omega)|^2 \right] dr.$$

Next note that the $f = \sigma$ terms can be bounded as follows: Use Doob's martingale inequality (Lemma A.6), the Ito isometry, and then increase the upper limit of integration to T . Now the sup disappears, and we are left with:

$$(178) \quad R_i^\sigma(t) = \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| \int_0^s g_i^\sigma(r, \omega) dB_r \right|^2 \right]$$

$$(179) \quad = 4 \sup_{0 \leq s \leq t} \mathbb{E} \left[\left| \int_0^s g_i^\sigma(r, \omega) dB_r \right|^2 \right]$$

$$(180) \quad = 4 \int_0^T \mathbb{E} \left[|g_i^\sigma(r, \omega)|^2 \right] dr$$

Combining these, we have proven:

$$(181) \quad R_i^\mu(t) + R_i^\sigma(t) \leq C \int_0^T \mathbb{E} \left[|g_i^\mu(r, \omega)|^2 + |g_i^\sigma(r, \omega)|^2 \right] dr,$$

and this will be our starting point for the three terms to be bounded.

For $i = 1$, we use the bounds (168) and $r - t_{j(r)} \leq \Delta t$ to get:

$$(182) \quad R_1^\mu(t) + R_1^\sigma(t) \leq C \int_0^T \mathbb{E} \left[|g_1^\mu(r, \omega)|^2 + |g_1^\sigma(r, \omega)|^2 \right] dr,$$

$$(183) \quad \leq 2CC_3^2 \int_0^T \mathbb{E} \left[(1 + |U_r|)^2 \right] (r - t_{j(r)}) dr$$

$$(184) \quad \leq 2CC_3^2 \Delta t \int_0^T \mathbb{E} \left[(1 + |U_r|)^2 \right] dr$$

By Lemma 3.10, the leftover integral factor is finite, so we conclude

$$(185) \quad R_1^\mu(t) + R_1^\sigma(t) \leq C \Delta t$$

For $i = 2$, use (169) to get

$$(186) \quad R_2^\mu(t) + R_2^\sigma(t) \leq C \int_0^T \mathbb{E} \left[|g_2^\mu(r, \omega)|^2 + |g_2^\sigma(r, \omega)|^2 \right] dr$$

$$(187) \quad \leq 2CC_1^2 \int_0^T \mathbb{E} \left[|U_r - U_{t_{j(r)}}|^2 \right] dr$$

By Lemma 3.11, we have

$$(188) \quad R_2^\mu(t) + R_2^\sigma(t) \leq C \Delta t.$$

Finally, consider $i = 3$, using the bound (170) to get:

$$(189) \quad R_3^\mu(t) + R_3^\sigma(t) \leq C \int_0^T \mathbb{E} \left[|g_3^\mu(r, \omega)|^2 + |g_3^\sigma(r, \omega)|^2 \right] dr$$

$$(190) \quad \leq 2CC_1^2 \int_0^T \mathbb{E} \left[|U_{t_{j(r)}} - \hat{V}_{t_{j(r)}}^N|^2 \right] dr$$

$$(191) \quad \leq 2CC_1^2 \int_0^T \mathbb{E} \left[\sup_{0 \leq r \leq s} |U_{t_{j(r)}} - \hat{V}_{t_{j(r)}}^N|^2 \right] dr.$$

In other words, we have proven:

$$(192) \quad R_3^\mu(t) + R_3^\sigma(t) \leq C \int_0^t z_N(r) dr$$

Combining these bounds all together, we have proven

$$(193) \quad z_N(t) \leq C_1 \Delta t + C_2 \int_0^t z_N(r) dr,$$

so Gronwall (Lemma A.4) gives

$$(194) \quad z_N(t) \leq C_1 \Delta t \exp(C_2 t)$$

In other words, we have

$$(195) \quad \mathbb{E} \left[\sup_{0 \leq s \leq t} |U_s - \hat{V}_s^N|^2 \right] \leq \frac{C}{N}$$

which completes the proof. \square

It may be surprising that the order of convergence of the Euler-Maruyama method does not match the order of the Euler method for numerical solutions to ODE—This discrepancy is due to the quadratic variation of the Brownian motion. We can achieve the desired order of convergence of one if we add more terms to the function determining the one-step method, but, of course, this comes at a cost of assuming extra smoothness conditions on μ and σ .

This insight leads us to the next numerical scheme of interest, whose proof of convergence we omit for the sake of brevity:

Definition 3.12. The *Milstein scheme* is a one-step numerical scheme determined by

$$(196) \quad \Psi(t, x, \Delta t, \Delta x) = \mu(t, x)\Delta t + \sigma(t, x)\Delta x$$

$$(197) \quad + \frac{1}{2}\sigma(t, x)\frac{\partial \sigma}{\partial x}(t, x)((\Delta x)^2 - \Delta t)$$

Theorem 3.13. If for all $f = \mu, \sigma$, and $\sigma \frac{\partial \sigma}{\partial x}$ we have

$$(198) \quad |f(t, x) - f(t, y)| \leq C_1|x - y|$$

$$(199) \quad |f(t, x)| \leq C_2(1 + |x|)$$

$$(200) \quad |f(t, x) - f(s, x)| \leq C_3(1 + |x|)\sqrt{|t - s|},$$

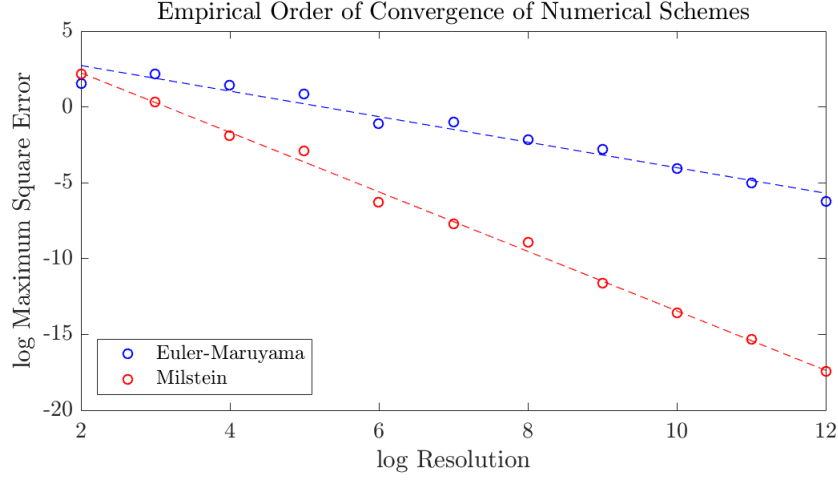
then the Milstein scheme is strongly quadratically convergent of order $\gamma = 1$.

Let us illustrate this point by revisiting the SODE (108) describing geometric Brownian motion and using Monte Carlo to empirically confirm that the orders of convergence are as claimed. To do this, observe that we have

$$(201) \quad \log_2(\text{error}) = a \log_2(N) + b \quad \Leftrightarrow \quad \text{error} = 2^b N^a.$$

This guarantees that the strong quadratic order of convergence of a numerical scheme can be determined empirically by halving the slope of the line of best fit of a log-log plot of the results of an experiment.

To perform this experiment, take the parameters $\mu = -1, \sigma = 2$, and $u_0 = 1$ as before. Then, for 100 independent trials, compute the sup norm of the square difference between the numerical solution and the true solution, and perform this calculation as the resolution ranges over $N = 2^2, 2^3, \dots, 2^{12}$. See Appendix B for the exact code used to create the following plot:



From here we can observe the expected behavior: The slope of -1 on the line for the Euler-Maruyama scheme corresponds to its strong quadratic order of convergence of $1/2$, and the slope of -2 on the line for the Milstein scheme corresponds to its strong quadratic order of convergence of 1 .

It turns out that the Euler-Maruyama scheme and the Milstein scheme can be combined into one general framework: The infinite Taylor series of an SDE, where Ito's formula is used for every derivative, takes on a useful form: The m th term is just a scalar multiple of some function of the Brownian- and time-increments, and the coefficient is either μ or $L^m \sigma$ where L is the operator $L = \sigma \frac{\partial}{\partial x}$. Then, the truncated series up to the M th term is a numerical scheme which converges strongly quadratically of order equal to $\gamma = M/2$, provided that all coefficient functions satisfy the bounds (198), (199), and (200). For the present we do not develop this further, but the proof of Theorem 3.13 and more on this general theory of Ito-Taylor approximation can be found in Chapter 10 of [3].

4. THEORY OF SPDE

The goal of this section is to study partial differential equations (PDEs) which are driven by white noise. Since a general theory of PDEs can be rather difficult to establish, it makes sense to restrict our attention to a small class of PDEs: For the present, we choose to study parabolic PDE in $(1 + 1)$ dimensions. It is well known that all such PDEs reduce, by a suitable change of coordinates, to the heat equation, so that will be our main object of study.

Arguing as before, let us consider an informal motivation. We will consider the following stochastic heat equation as our general object of study:

$$(202) \quad \frac{\partial U}{\partial t} = \alpha \frac{\partial^2 U}{\partial x^2} + f(t, x, U(t, x))W(t, x), \quad U(0, x) = u_0(x).$$

where (t, x) ranges over $[0, T] \times I$ for some connected set $I \subseteq \mathbb{R}$, possibly with boundary conditions. Here we assume that $f : [0, T] \times I \times \mathbb{R} \rightarrow \mathbb{R}$ is some sufficiently smooth function, $u_0 : I \rightarrow \mathbb{R}$ is some compactly supported initial data, and $W(t, x)$ represents a space-time white noise.

Recall from Section 2 the lesson that it is much easier to define integral equations than differential equations. Duhamel's principle gives us a way to do this for parabolic PDE; We can write (202) equivalently as

$$(203) \quad U(t, x) = \int_I u_0(y) G_\alpha(t, x; y) dy$$

$$(204) \quad + \int_0^t \int_I f(s, y, U(s, y)) G_\alpha(t - s, x; y) W(s, y) dy ds$$

where G_α is the Green's function for the usual heat equation on $[0, T] \times I$ with the same boundary conditions. Of course, our main work will be defining the stochastic integral in two dimensionals with respect to a space-time white noise.

However, the present setting introduces a new consideration which was not salient in the case of SODEs: The specification of $I \subseteq \mathbb{R}$ and the boundary conditions can, in principle, be very complicated. Fortunately, the move to integral equations means that we can encode all of the boundary conditions information into the Green's function, at least for stochastic heat equations that correspond to well-posed deterministic heat equations.

Several examples will be important to us. For the *unbounded* case of $I = \mathbb{R}$, no boundary conditions are necessary and the Green's function is the usual heat kernel:

$$(205) \quad G_\alpha(t, x; y) = \frac{1}{\sqrt{4\pi\alpha t}} \exp\left(-\frac{|x - y|^2}{4\alpha t}\right).$$

We are also interested in the case of *periodic boundary conditions*. That is, for some $L > 0$, set $I = [-L/2, L/2]$ and impose boundary conditions $U(t, -L/2) = U(t, L/2)$ and $\partial_x U(t, -L/2) = \partial_x U(t, L/2)$. In this case the Green's function is given by the infinite sum

$$(206) \quad G_\alpha^P(t, x; y) = \frac{1}{\sqrt{4\pi\alpha t}} \sum_{n=-\infty}^{\infty} \exp\left(-\frac{|x - y - nL|^2}{4\alpha t}\right).$$

Finally, we consider the case of *Dirichlet boundary conditions*. That is, for some $L > 0$, set $I = [0, L]$ and impose $U(t, 0) = U(t, L) = 0$. Here the Green's function takes on the form

$$(207) \quad G_\alpha^D(t, x; y) = \frac{1}{\sqrt{4\pi\alpha t}} \sum_{n=-\infty}^{\infty} \left(\exp\left(-\frac{|x - y - 2nL|^2}{4\alpha t}\right) + \exp\left(-\frac{|x + y - 2nL|^2}{4\alpha t}\right) \right).$$

The remainder of this section is dedicated to making these ideas precise and exploring the basic properties of the stochastic heat equation. To simplify the exposition, the structure of this section is meant to parallel that of Section 2 as much as possible.

4.1. White Noise. In order to make precise the theory of stochastic partial differential equations (SPDEs), we will need to define a suitable concept of a rough driving force like a Brownian motion. By Definition 2.1, let W be a white noise on $[0, T] \times \mathbb{R}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 4.1. The *Brownian martingale measure* is the stochastic process $B : [0, T] \times \mathcal{B}(\mathbb{R}) \rightarrow L^2(\Omega)$ is defined as

$$(208) \quad B(t, A) = W((0, t] \times A).$$

The order induced on the index set $[0, T] \times \mathcal{B}(\mathbb{R})$ is the total ordering on $[0, T]$. Hence, the natural filtration of $\{B(t, A)\}_{t,A}$ is

$$(209) \quad \mathcal{F}_t = \sigma(\{B(s, A) : s \leq t, A \in \mathcal{B}(\mathbb{R})\})$$

We denote this filtration by $\mathbb{F} = \{\mathcal{F}_t\}_t$.

It should be made explicit that we are treating the variables t and x in fundamentally different ways. On the one hand, the t variable is taken to range over the compact set $[0, T]$ which we have endowed with a “direction” so as to make sense of “not seeing into the future”. On the other hand, the x variable is taken to range over a noncompact space and it does not make sense to endow its domain with any kind of direction. This is why our construction of the Brownian object must be a measure in the x -variable and not a function; it does not make sense to use the naive construction of $B(t, x) = W([0, t], [0, x])$ as our integrator.

Lemma 4.2. The Brownian martingale measure $\{B(t, A)\}_{t,A}$ enjoys the following properties:

- (4.2.1) For all $A \in \mathcal{B}(\mathbb{R})$, we have $B(0, A) = 0$ holds \mathbb{P} -almost surely.
- (4.2.2) For each $t \in [0, T]$ and $A \in \mathcal{B}(\mathbb{R})$, the law of $B(t, A)$ is σ -finite on \mathbb{R} .
- (4.2.3) For all $A \in \mathcal{B}(\mathbb{R})$, the stochastic process $\{B(t, A)\}_t$ is a martingale with respect to \mathbb{F} .

Proof. We get (4.2.1) immediately, since $B(0, A)$ is distributed as $N(0, 0)$. Moreover, (4.2.2) is immediate since the law of each $B(t, A)$ is a probability measure on \mathbb{R} . To verify (4.2.3), we only need to check, for arbitrary $s \leq t$, the following: $\mathbb{E}[B(t, A)|\mathcal{F}_s] = \mathbb{E}[W((0, s] \times A) + W((s, t] \times A)|\mathcal{F}_s] = W((0, s] \times A) + \mathbb{E}[W((s, t] \times A)] = B(s, A)$. \square

It turns out that we can define stochastic integration for any stochastic process $M : [0, T] \times \mathcal{B}(\mathbb{R}) \rightarrow L^2(\Omega)$ satisfying the conclusions of Lemma 4.2; such processes are called *martingale measures*. While we will not study martingale measures in generality, observe that most of the work of the next subsection carries through without any changes whenever $\{B(t, A)\}_{t,A}$ is replaced by a general martingale measure $\{M(t, A)\}_{t,A}$.

4.2. Stochastic Integration. In analogy with Section 2, our next goal is to make sense of

$$(210) \quad \int_0^T \int_A f dB(t, x)$$

where $f : [0, T] \times \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ is some random function, $\{B(t, A)\}_{t,A}$ is the Brownian martingale measure and A is an arbitrary Borel set in \mathbb{R} . If we are successful, this integral should be a martingale measure when viewed as a stochastic process in the arguments $(t, A) \in [0, T] \times \mathcal{B}(\mathbb{R})$.

As before, we start by defining the integral for elementary processes:

Definition 4.3. A stochastic process $\phi : [0, T] \times \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ is called *elementary* with respect to $\mathbb{F} = \{\mathcal{F}_t\}_t$ if it can be written as

$$(211) \quad \phi(t, x, \omega) = \sum_{j=0}^{k-1} E_j(\omega) \mathbf{1}_{[t_j, t_{j+1})}(t) \mathbf{1}_{A_j}(x)$$

for some positive integer k , discretization $0 = t_0 < t_1 < \dots < t_k = T$, disjoint Borel sets $\{A_j\}_j$ in \mathbb{R} , and random variables $\{E_j\}_j \subseteq L^2(\Omega)$ such that E_j is \mathcal{F}_{t_j} -measurable for all j . We use $\mathcal{E}_{\mathbb{F}}$ to denote the collection of all elementary processes with respect to \mathbb{F} .

Definition 4.4. The *Ito integral* for fixed T, A is the map $I_{T,A} : \mathcal{E}_{\mathbb{F}} \rightarrow L^2(\Omega)$ defined as follows: For $\phi \in \mathcal{E}_{\mathbb{F}}$ with discretization $\{t_j\}_j$, disjoint Borel sets $\{A_j\}_j$, and random variables $\{E_j\}_j$, set

$$(212) \quad I_{T,A}(\phi)(\omega) = \int_0^T \int_A \phi dB(t, x)(\omega)$$

$$(213) \quad = \sum_{j=0}^{k-1} E_j(\omega) (B(t_{j+1}, A_j \cap A)(\omega) - B(t_j, A_j \cap A)(\omega))$$

Lemma 4.5. Let $\phi, \psi \in \mathcal{E}_{\mathbb{F}}$ be any elementary processes and $\alpha, \beta \in \mathbb{R}$ be any reals. The map $I_{T,A} : \mathcal{E}_{\mathbb{F}} \rightarrow L^2(\Omega)$ enjoys the following properties:

- (4.5.1) (Linearity) $I_{T,A}(\alpha\phi + \beta\psi) = \alpha I_{T,A}(\phi) + \beta I_{T,A}(\psi)$
- (4.5.2) (Mean Zero) $\mathbb{E}[I_{T,A}(\phi)] = 0$
- (4.5.3) (Ito Isometry) $\|I_{T,A}(\phi)\|_{L^2(\Omega)} = \|\phi\|_{L^2([0,T] \times A \times \Omega)}$
- (4.5.4) (Martingale Measure) The stochastic process $\{I_{t,A}(\phi)\}_{t,A}$ is a martingale measure with respect to the filtration \mathbb{F} .
- (4.5.5) (Continuity) For all $A \in \mathcal{B}(\mathbb{R})$, the stochastic process $\{I_{t,A}(\phi)\}_t$ is \mathbb{P} -almost surely continuous in t .

Proof. The proofs of (4.5.1), (4.5.2), (4.5.3), (4.5.5) are identical to the corresponding proofs in Lemma 2.9. To prove (4.5.4), we must check all three conditions (4.2.1), (4.2.2), and (4.2.3): the first two of these are immediate, and the third is verified by the proof of (2.9.4). \square

Definition 4.6. Call a stochastic process $f : [0, T] \times \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ *predictable* with respect to \mathbb{F} if it is $\sigma(\mathcal{E}_{\mathbb{F}})$ -measurable. Also define the set $\mathcal{P}_{\mathbb{F}}^2 = L^2([0, T] \times \mathbb{R} \times \Omega, \sigma(\mathcal{E}_{\mathbb{F}}), \lambda \otimes \lambda \otimes \mathbb{P})$ of all processes which are predictable with respect to \mathbb{F} and square integrable.

Lemma 4.7. For any probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and filtration $\mathbb{F} = \{\mathcal{F}_t\}_t$ of \mathcal{F} , we have that $\mathcal{P}_{\mathbb{F}}^2$ is a Hilbert space.

Proof. Exactly as before: $\mathcal{P}_{\mathbb{F}}^2$ is a closed linear subspace of a Hilbert space, so it is a Hilbert space. \square

The main obstruction in adapting the proofs of Section 2 to the present setting is that x is now allowed to range over a noncompact space. This necessitates a few extra steps but does not change the nature of the proofs too much.

Lemma 4.8. With respect to any filtration $\mathbb{F} = \{\mathcal{F}_t\}_t$, an \mathbb{F} -adapted and continuous stochastic process is \mathbb{F} -predictable.

Proof. Suppose that f is \mathbb{F} -adapted and continuous, and, for each n , set $f_n(t, x, \omega) = f(t, x, \omega)\mathbf{1}_{[-n, n]}(x)$. Now each f_n is continuous and compactly supported, so a straightforward adaptation of the proof of Lemma 2.12 goes through. That is, f_n is in $\mathcal{P}_{\mathbb{F}}^2$. But pointwise limit of measurable functions are measurable and we have we have $f_n(t, x, \omega) \rightarrow f(t, x, \omega)$ for all (t, x, ω) , so f is also in $\mathcal{P}_{\mathbb{F}}^2$. \square

Lemma 4.9. The space $\mathcal{E}_{\mathbb{F}} \subseteq L^2([0, T] \times \mathbb{R} \times \Omega)$ is dense in $\mathcal{P}_{\mathbb{F}}^2$. That is, for any $f \in \mathcal{P}_{\mathbb{F}}^2$, there exists a sequence $\{\phi_n\}_n$ of elementary processes with respect to \mathbb{F} with $\phi_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$.

Proof. Proof is similar to the proof of Lemma 2.13. That is, we proceed in four steps:

Step 1: If $f \in \mathcal{P}_{\mathbb{F}}^2$ has $f(t, x, \omega)$ continuous in (t, x) for all ω , is supported on $[0, T] \times \Gamma$ with respect to (t, x) where the compact set $\Gamma \subseteq \mathbb{R}$ does not depend on ω , and is bounded uniformly by M across all (t, x, ω) , then there exists a sequence of $\{\phi_n\}_n \subseteq \mathcal{E}_{\mathbb{F}}$ with $\phi_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$.

For each positive integer n set partitions $\{t_j\}_j$ of $[0, T]$ and $\{x_j\}_j$ of Γ , each consisting of 2^n intervals, such that $\max_{j=0,1,\dots,2^n-1} (t_{j+1} - t_j)(x_{j+1} - x_j) \rightarrow 0$ as $n \rightarrow \infty$. Then define

$$(214) \quad \phi_n(t, x, \omega) = \sum_{j=0}^{2^n-1} f(t_j, x_j, \omega) \mathbf{1}_{(t_j, t_{j+1}]}(t) \mathbf{1}_{(x_j, x_{j+1}]}(x)$$

Since f is, in particular, \mathbb{F} -adapted, we see that $h(t_j, x_j, \omega)$ is \mathcal{F}_{t_j} -measurable. Hence, ϕ_n is elementary with respect to \mathbb{F} for all n .

Now let ω be fixed. By the continuity of $f(\cdot, \cdot, \omega)$, we have $|\phi_n(t, x, \omega) - f(t, x, \omega)|^2 \rightarrow 0$ and $|\phi_n(t, x, \omega) - f(t, x, \omega)|^2 \leq M^2$ for all (t, x) . Hence, the dominated convergence theorem gives $\phi_n(\cdot, \cdot, \omega) \rightarrow f(\cdot, \cdot, \omega)$ in $L^2([0, T] \times \mathbb{R})$ for each ω . Since we also have the bound $\|\phi_n(\cdot, \cdot, \omega) - f(\cdot, \cdot, \omega)\|_{L^2([0, T] \times \mathbb{R})}^2 \leq 4M^2 T \lambda(\Gamma)$ holding uniformly across ω , another application of the dominated convergence theorem along with Fubini-Tonelli gives $\phi_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$. Therefore, $\{\phi_n\}_n$ is as claimed.

Step 2: If $f \in \mathcal{P}_{\mathbb{F}}^2$ is supported on $[0, T] \times \Gamma$ with respect to (t, x) where the compact set $\Gamma \subseteq \mathbb{R}$ does not depend on ω and if f is bounded uniformly by M across all (t, x, ω) , then there exists a sequence of $\{\phi_n\}_n \subseteq \mathcal{E}_{\mathbb{F}}$ with $\phi_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$.

Let $K : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be any nonnegative continuous function with compact support contained in $[0, T] \times \mathbb{R}$ and $\int_{\mathbb{R}^2} K d\lambda = 1$, for example a suitably scaled cone with center at $(t, x) = (T/2, 0)$ and with base radius equal to $T/2$. In particular, this implies that K is bounded. Then we can define the functions $\{K_n\}_n$ by $K_n(t, x) = n^2 K(nt, nx)$, and

$$(215) \quad g_n(t, x, \omega) = (K_n * f(\cdot, \cdot, \omega))(t, x) = \int_0^T \int_{\mathbb{R}} K_n(s, y) f(t-s, x-y, \omega) dy ds$$

$$(216) \quad = \int_0^T \int_{\mathbb{R}} K_n(t-s, x-y) f(s, y, \omega) dy ds$$

where $f(t, x, \omega)$ is extended to be equal to zero whenever $t \leq 0$.

Let's prove some basic properties about every g_n . First, note that $g_n(\cdot, \cdot, \omega)$ is continuous in (t, x) for all ω , since it is the convolution of a compactly supported continuous function and a bounded function. Also, the support of $g_n(\cdot, \cdot, \omega)$ satisfies

$$(217) \quad \text{supp}(g_n(\cdot, \cdot, \omega)) = \overline{\text{supp}(K) + \text{supp}(f(\cdot, \cdot, \omega))} \subseteq \overline{\text{supp}(K) + ([0, T] \times \Gamma)},$$

and, by assumption, the set on the right side, which we denote by S , does not depend on ω . Finally, we can bound

$$(218) \quad |g_n(t, x, \omega)| \leq M\lambda(\text{supp}(K)) \sup_{s, y} K(s, y)$$

so each g_n is bounded uniformly across (t, x, ω) ; Let us denote this uniform bound by M_n .

Now let us prove that g_n is \mathbb{F} -predictable for each n . As before, Lemma 4.8 and the fact that g_n is continuous imply that it is sufficient to prove that g_n is \mathbb{F} -adapted. Note that we can write

$$(219) \quad g_n(t, x, \omega) = \int_0^T \int_{\mathbb{R}} K_n(t-s, x-y) f(s, y, \omega) dy ds$$

$$(220) \quad = \int_0^t \int_{\mathbb{R}} K_n(t-s, x-y) f(s, y, \omega) dy ds,$$

and the last line holds since $s \geq t$ implies $K_n(t-s, x-y) = 0$. But $f(s, y, \omega)$ is \mathcal{F}_t -measurable for all $s \leq t$, so the double integral is also \mathcal{F}_t -measurable. This proves that each g_n is \mathbb{F} -predictable.

Next, we prove as an intermediate step that g_n converges in $L^2([0, T] \times \mathbb{R} \times \Omega)$ to f . First let ω be fixed, and compute the following bound, using Minkowski's integral inequality (Lemma A.7):

$$\begin{aligned} & \|g_n(\cdot, \cdot, \omega) - f(\cdot, \cdot, \omega)\|_{L^2([0, T] \times \mathbb{R})} \\ &= \left(\int_{[0, T] \times \mathbb{R}} \left| \int_{[0, T] \times \mathbb{R}} K(s, y) \left(f\left(t - \frac{s}{n}, x - \frac{y}{n}, \omega\right) - f(t, x, \omega) \right) d\lambda(s, y) \right|^2 d\lambda(t, x) \right)^{\frac{1}{2}} \\ &\leq \int_{[0, T] \times \mathbb{R}} K(s, y) \left\| f\left(\cdot - \frac{s}{n}, \cdot - \frac{y}{n}, \omega\right) - f(\cdot, \cdot, \omega) \right\|_{L^2([0, T] \times \mathbb{R})} d\lambda(s, y). \end{aligned}$$

The integrand is bounded by $2M \sup_{s, y} K(s, y)$, and the same argument from before shows that $f(\cdot - \frac{s}{n}, \cdot - \frac{y}{n}, \omega) \rightarrow f(\cdot, \cdot, \omega)$ holds in $L^2([0, T] \times \mathbb{R})$. Hence, dominated convergence gives $g_n(\cdot, \cdot, \omega) \rightarrow f(\cdot, \cdot, \omega)$ in $L^2([0, T] \times \mathbb{R})$. But also we have the uniform bound

$$(221) \quad \|g_n(\cdot, \cdot, \omega) - f(\cdot, \cdot, \omega)\|_{L^2([0, T] \times \mathbb{R})}$$

$$(222) \quad \leq \|g_n(\cdot, \cdot, \omega)\|_{L^2([0, T] \times \mathbb{R})} + \|f(\cdot, \cdot, \omega)\|_{L^2([0, T] \times \mathbb{R})}$$

$$(223) \quad \leq M_n \lambda(S) + MT \lambda(\Gamma)$$

where M_n and S were defined during the earlier part of Step 2. Since this holds across all ω , dominated convergence and Fubini-Tonelli give $g_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$. In particular, this implies $\|g_n\|_{L^2([0, T] \times \mathbb{R} \times \Omega)} < \infty$, so $g_n \in \mathcal{P}_{\mathbb{F}}^2$.

Finally, we have proven that each g_n satisfies the hypotheses of Step 1. Hence for each n we construct $\{\phi_{k,n}\}_k$ to be a sequence of processes in $\mathcal{E}_{\mathbb{F}}$ with $\phi_{k,n} \rightarrow g_n$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$ as $k \rightarrow \infty$. Then by diagonalization we have $\{\phi_{n,n}\}_n$, a sequence of processes in $\mathcal{E}_{\mathbb{F}}$ with $\phi_{n,n} \rightarrow g$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$ as $n \rightarrow \infty$. This completes the proof of Step 2.

Step 3: If $f \in \mathcal{P}_{\mathbb{F}}^2$ is supported on $[0, T] \times \Gamma$ with respect to (t, x) where the compact set $\Gamma \subseteq \mathbb{R}$ does not depend on ω , then there exists a sequence of $\{\phi_n\}_n \subseteq \mathcal{E}_{\mathbb{F}}$ with $\phi_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$.

For each n let us set

$$(224) \quad g_n(t, x, \omega) = \begin{cases} -n & \text{if } f(t, x, \omega) \leq -n \\ f(t, x, \omega) & \text{if } -n \leq f(t, x, \omega) \leq n \\ n & \text{if } f(t, x, \omega) \geq n \end{cases}.$$

For all (t, x, ω) , we have $|g_n(t, x, \omega) - f(t, x, \omega)|^2 \rightarrow 0$ and $|g_n(t, x, \omega) - f(t, x, \omega)|^2 \leq 4|f(t, x, \omega)|^2$, so dominated convergence gives $g_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$. Now the following properties take no work verify: g_n is bounded by n uniformly across all (t, x, ω) , we have $g_n \in \mathcal{P}_{\mathbb{F}}^2$, and $\text{supp}(g_n(\cdot, \cdot, \omega)) \subseteq [0, T] \times \Gamma$. As such, we can use Step 2 to get a sequence $\{\phi_{k,n}\}_k$ with $\phi_{k,n} \rightarrow g_n$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$, and, by diagonalization, this implies $\phi_{n,n} \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$.

Step 4: For any $f \in \mathcal{P}_{\mathbb{F}}^2$, there exists a sequence of $\{\phi_n\}_n \subseteq \mathcal{E}_{\mathbb{F}}$ with $\phi_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$.

For each n define $g_n(t, x, \omega) = f(t, x, \omega) \mathbf{1}_{[-n, n]}(x)$, and note that the monotone convergence theorem gives $g_n \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$. But each g_n satisfies the hypotheses of Step 3, so we can get a sequence $\{\phi_{k,n}\}_k$ in $\mathcal{E}_{\mathbb{F}}$ converging to g_n in $L^2([0, T] \times \mathbb{R} \times \Omega)$. Diagonalizing as always, $\{\phi_{n,n}\}_n$ is a sequence of elementary processes with respect to \mathbb{F} with $\phi_{n,n} \rightarrow f$ in $L^2([0, T] \times \mathbb{R} \times \Omega)$. This completes Step 4, and hence proves the result. \square

Definition 4.10. The *Ito integral* is the unique extension of $I_{T,A} : \mathcal{E}_{\mathbb{F}} \rightarrow L^2(\Omega)$ to $\mathcal{P}_{\mathbb{F}}^2 \subseteq \overline{\mathcal{E}_{\mathbb{F}}}$. We denote this operator by

$$(225) \quad I_{T,A}(f) = \int_0^T \int_A f dB(t, x)$$

Lemma 4.11. Let $f, g \in \mathcal{P}_{\mathbb{F}}^2$ be any predictable, square integrable processes, and let $\alpha, \beta \in \mathbb{R}$ be any reals. The map $I_{T,A} : \mathcal{P}_{\mathbb{F}}^2 \rightarrow L^2(\Omega)$ enjoys the following properties:

- (4.11.1) (Linearity) $I_T(\alpha f + \beta g) = \alpha I_{T,A}(f) + \beta I_{T,A}(g)$
- (4.11.2) (Mean Zero) $\mathbb{E}[I_{T,A}(f)] = 0$
- (4.11.3) (Ito Isometry) $\|I_{T,A}(f)\|_{L^2(\Omega)} = \|f\|_{L^2([0, T] \times \Omega)}$
- (4.11.4) (Martingale Measure) The stochastic process $\{I_{t,A}(f)\}_{t,A}$ is a martingale measure with respect to the filtration \mathbb{F} .

Proof. As in the proof of Lemma 2.15, these properties are either immediate or follow immediately from the continuity of various operators. \square

4.3. Integral Equations. We now study the properties of solutions to the stochastic heat equation defined on the domain $[0, T] \times I$ with Green's function $G_\alpha(t, x; y)$.

Definition 4.12. An *Ito stochastic heat equation with nonlinear inhomogeneity* on $[0, T] \times I$ is an integral equation of the form

$$(226) \quad U(t, x) = \int_I u_0(y) G_\alpha(t, x - y) dy$$

$$(227) \quad + \int_0^t \int_I f(s, y, U(s, y)) G_\alpha(t - s, x; y) dB(s, y),$$

$$(228) \quad U(0, x) = u_0(x)$$

where $\{U(t, x)\}_{t,x}$ is a stochastic process defined on $[0, T] \times I$.

Definition 4.13. If for any probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which a Brownian martingale measure $\{B_{t,A}\}_{t,A}$ with canonical filtration $\mathbb{F} = \{\mathcal{F}_t\}_t$ is defined, there exists a stochastic process $\{U(t, x)\}_{t,x}$ in $\mathcal{P}_{\mathbb{F}}^2$ such that (226) holds \mathbb{P} -almost surely per fixed (t, x) , then $\{U(t, x)\}_{t,x}$ is called a *strong solution* of the SPDE.

Compare the above to Definition 2.18, in which it was demanded that the \mathbb{P} -almost all sample paths of the solution were continuous. This condition is not necessary in the SPDE case, since continuity is guaranteed for any strong solution:

Theorem 4.14. Any strong solution of the SPDE (226) is automatically continuous in t and in x .

For the sake of brevity, we omit the proof of this result from the current work, and instead direct readers to [8] in which a stronger statement about the Hölder continuity of strong solutions is proved. For now, we move on to our main existence and uniqueness result:

Theorem 4.15. Suppose that $u_0 : I \rightarrow \mathbb{R}$ is compactly supported and $f : [0, T] \times I \times \mathbb{R} \rightarrow \mathbb{R}$ satisfies the Lipschitz bound

$$(229) \quad |f(t, x, u) - f(t, x, v)| \leq C|u - v|$$

$$(230) \quad |f(t, x, u)| \leq C(1 + |u|)$$

Suppose further tha the Green's function satisfies the uniform bound

$$(231) \quad \int_I G_\alpha^2(t, x; y) dy \leq g(t)$$

with $g \in L^q([0, T])$ for some $q > 1$. Then, the SPDE (226) admits strong solutions.

Proof. The proof is similar to the proof of Theorem 2.19. To begin, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which a Brownian martingale measure $\{B_{t,A}\}_{t,A}$ is defined. Then use Picard's iteration method to define a sequence of random functions $\{\{U^k(t, x)\}_{t,x}\}_k$ by

$$(232) \quad U^0(t, x) = u_0(x)$$

$$(233) \quad U^{k+1}(t, x) = \int_I u_0(y)G(t, x; y)dy$$

$$(234) \quad + \int_0^t \int_I f(s, y, U^k(s, y))G(t-s, x; y)dB(s, y),$$

We now bound the $L^2(\Omega)$ -distance between adjacent terms using the Ito isometry and the Lipschitz bound (229):

$$(235) \quad \begin{aligned} & \mathbb{E} [|U^{k+1}(t, x) - U^k(t, x)|^2] \\ &= \mathbb{E} \left[\left| \int_0^t \int_I (f(s, y, U^k(s, y)) - f(s, y, U^{k-1}(s, y)))G(t-s, x; y)dB(s, y) \right|^2 \right] \\ &\leq \int_0^t \int_I \mathbb{E} [|f(s, y, U^k(s, y)) - f(s, y, U^{k-1}(s, y))|^2] G^2(t-s, x; y)dyds \\ &\leq C \int_0^t \int_I \mathbb{E} [|U^k(s, y) - U^{k-1}(s, y)|^2] G^2(t-s, x; y)dyds \end{aligned}$$

Now define

$$(236) \quad z_k(t) = \sup_{x \in I} \mathbb{E} [|U^k(t, x) - U^{k-1}(t, x)|^2]$$

so that (235) can be further bounded using (231) as

$$(237) \quad z_{k+1}(t) \leq C \sup_{x \in I} \int_0^t z_k(s) \int_I G^2(t-s, x; y)dyds$$

$$(238) \quad \leq C \int_0^t z_k(s)g(t-s)ds.$$

Now use Hölder's inequality for conjugate exponents p, q to get

$$(239) \quad z_{k+1}(t) \leq C \left(\int_0^t z_k^p(s)ds \right)^{1/p} \left(\int_0^t |g(t-s)|^q ds \right)^{1/q}$$

The second integral is finite by assumption, so we have proven

$$(240) \quad z_{k+1}^p(t) \leq C \int_0^t z_k^p(s)ds.$$

A similar argument but with the bound (230) proves

$$(241) \quad z_1(t) = \sup_{x \in I} \mathbb{E} \left[\left| \int_0^t \int_I f(s, y, u_0(y))G_\alpha(t-s, x; y)dB(s, y) \right|^2 \right]$$

$$(242) \quad \leq \sup_{x \in I} \int_0^t \int_I (1 + |u_0(y)|)^2 G_\alpha^2(t-s, x; y)dyds.$$

The integral is of course uniformly bounded in x , so the right side is finite.

Applying Gronwall's inequality (Lemma A.3) to the sequence $\{\{z_k^q(t)\}_t\}_k$, we obtain the following bound:

$$(243) \quad \mathbb{E} \left[|U^k(t, x) - U^{k-1}(t, x)|^2 \right] \leq C_1 \left(\frac{(C_2 T)^k}{k!} \right)^{\frac{1}{q}}$$

which holds uniformly across all (t, x) . Since this goes to zero as $k \rightarrow \infty$, we have constructed a candidate solution $\{U(t, x)\}_{t,x}$.

To see that this is truly a solution, it suffices to prove

$$(244) \quad \int_0^t \int_I f(s, y, U^k(t, x)) G_\alpha(t - s, x; y) dB(s, y)$$

$$(245) \quad \rightarrow \int_0^t \int_I f(s, y, U(t, x)) G_\alpha(t - s, x; y) dB(s, y)$$

in $L^2(\Omega)$. This can be done easily exactly as before, simply by applying Gronwall (Lemma A.3) to the bound

$$(246) \quad \begin{aligned} & \mathbb{E} [|U^k(t, x) - U(t, x)|^2] \\ & \leq C \int_0^t \int_I \mathbb{E} [|U^{k-1}(s, y) - U(s, y)|^2] G^2(t - s, x; y) dy ds \end{aligned}$$

Similarly, uniqueness is proven with the same method. If $\{U^1(t, x)\}_{t,x}$ and $\{U^2(t, x)\}_{t,x}$ are both solutions, then we have

$$(247) \quad \begin{aligned} & \mathbb{E} [|U^1(t, x) - U^2(t, x)|^2] \\ & \leq C \int_0^t \int_I \mathbb{E} [|U^1(s, y) - U^2(s, y)|^2] G^2(t - s, x; y) dy ds \end{aligned}$$

which implies

$$(248) \quad \mathbb{E} [|U^1(t, x) - U^2(t, x)|^2] = 0$$

so $U^1(t, x) = U^2(t, x)$ holds \mathbb{P} -almost surely. \square

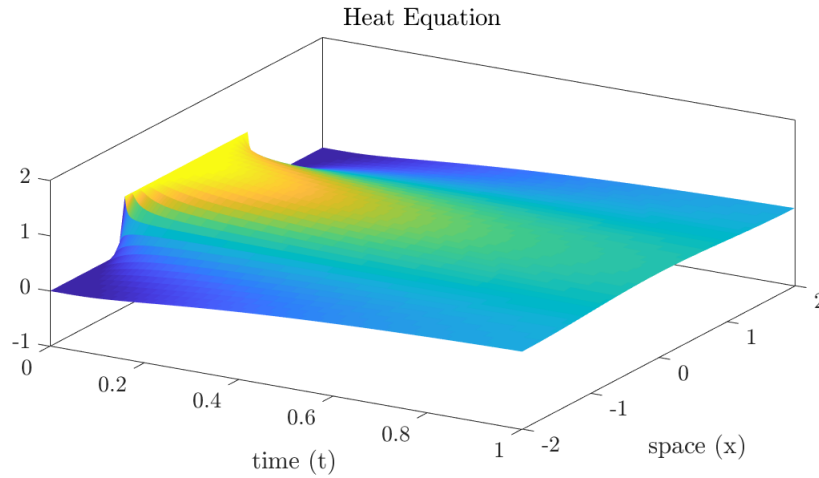
As a consequence of Theorem 4.15 we can obtain strong solutions to the SPDE (226) on all of the domains that we considered at the beginning of the section. Indeed, the following can all be verified by computing some simple bounds on Green's functions:

Corollary 4.16. If f and u_0 satisfy the hypotheses of Theorem 4.15, then the SPDE (202) admits strong solutions in all of the following settings:

- (4.16.1) Domain $[0, T] \times \mathbb{R}$ and no boundary conditions
- (4.16.2) Domain $[0, T] \times [-L/2, L/2]$ and periodic boundary conditions
- (4.16.3) Domain $[0, T] \times [0, L]$ and Dirichlet boundary conditions

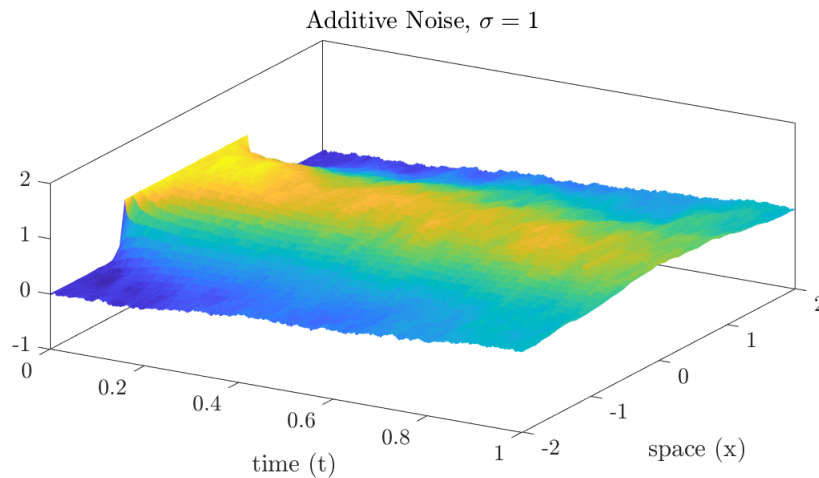
4.4. Examples. We can now explore some examples of the stochastic heat equation. For the sake of implementation in a computer, we restrict our attention to the bounded domain with periodic boundary conditions.

We will fix all but the inhomogeneity f from now on. Take our domain to be $[0, T] \times [-L/2, L/2]$ for $T = 1$ and $L = 4$, fix $\alpha = 1$, and let our initial condition be $u_0(x) = \mathbf{1}_{[-1,1]}(x)$. As a benchmark, we note that the deterministic heat equation ($f = 0$) with these specifications has the following solution:



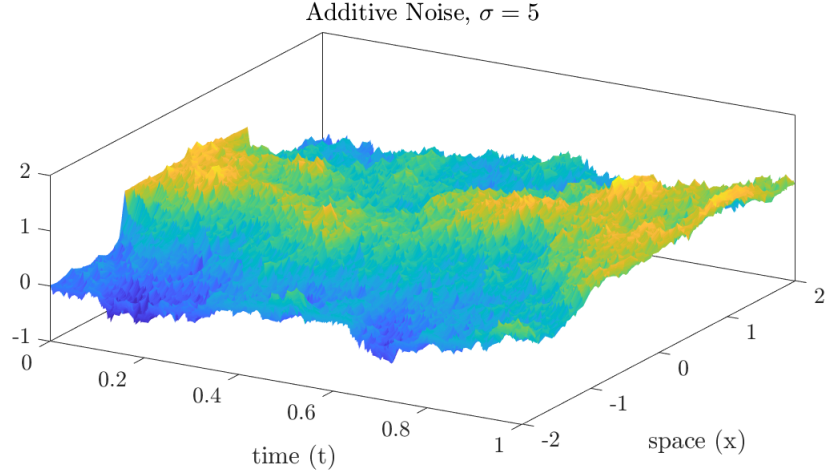
An important qualitative feature is that the solution “stabilizes” near the constant value $U(t, x) = 1/2$, essentially due to the fact that periodic boundary conditions force heat to distribute itself throughout the domain. Another feature of note is that the singular initial conditions become smooth instantaneously.

Example 4.17 (Additive Noise). Consider the simplest nontrivial inhomogeneity, $f(t, x, u) = \sigma$ for some fixed constant $\sigma > 0$. As a first example, take $\sigma = 1$ and see the following plot:



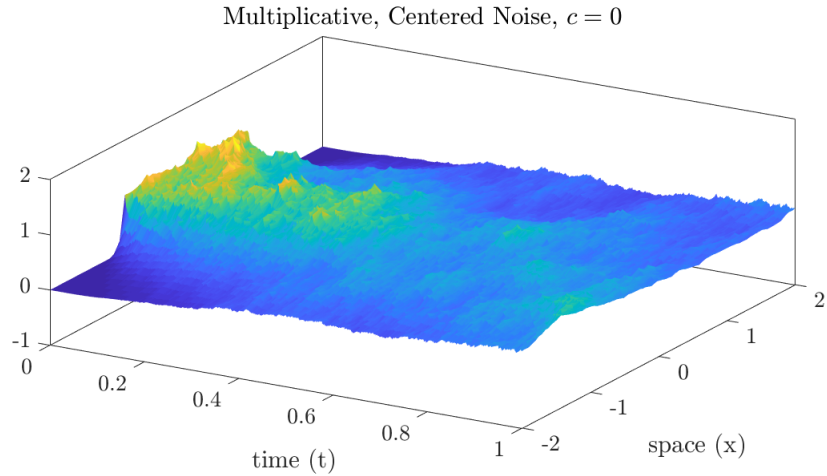
It appears that the solution retains the stabilization property of the deterministic heat equation, but the rate of convergence, which is of course probabilistic, appears to have changed. Additionally, it appears that the smoothness of the solution has changed, although this can be hard to see in the plot.

We now increase the strength of the noise to $\sigma = 5$ and check the resulting plot for these same properties:



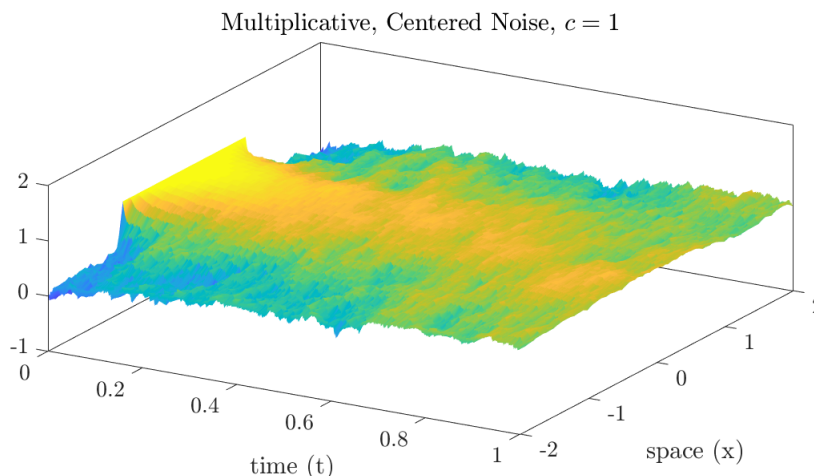
From this plot, we can see that we have totally lost both the stabilization and the smoothness of the solution. That the stability fails should be rather intuitive: The force of heat distribution is not fast enough to spread out the large random peaks that are appearing. That is, even on a longer time-scale, we do not expect these solutions to stabilize.

Example 4.18 (Multiplicative, Centered Noise). We now consider nonlinear inhomogeneities of the form $f(t, x, u) = \sigma|u - c|$ for some $\sigma > 0$ and $c \in \mathbb{R}$. Let us fix $\sigma = 5$ and vary c across a few different values. As a start, consider $c = 0$:



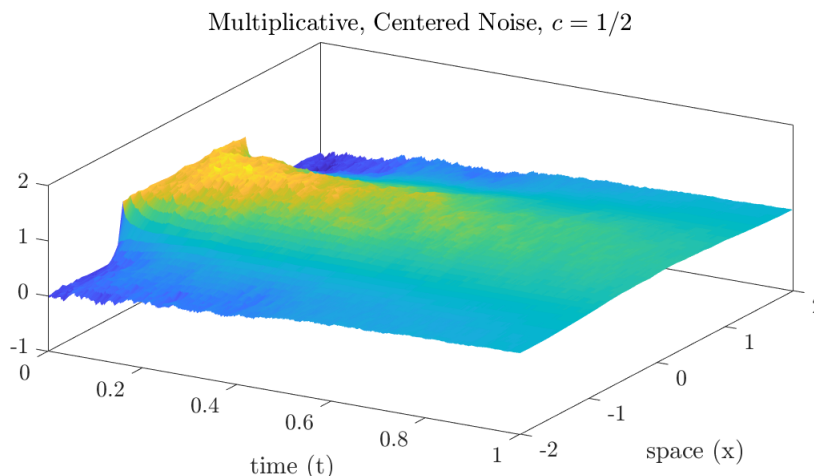
Observe that at time $t = 0$, the strength of the noise changes the total heat in the system slightly. Since the noise is relatively weak for $t > 0$, this change propagates all the way to time $t = 1$. This is evidenced by the fact we can see the system stabilize at a value strictly smaller than $1/2$.

Now let's investigate the effect of moving the center to $c = 1$. Consider the following plot:



As expected, this noise has the effect of “pulling” the solutions up to the constant value of $U(t, x) = 1$.

Finally, let us consider taking the center to be $c = 1/2$. Since we know that the deterministic heat equation already stabilizes at this value, it is not what effect the noise will have, if any at all. So, we look to the following plot:



It is clear the noise has some effect at time $t = 0$, but this solution looks remarkably similar to the solution of the deterministic heat equation we saw at the beginning of this subsection.

For details on the how these simulations we computed, we move on to the final section of this thesis.

5. NUMERICAL SPDE

This brief section is dedicated to introducing the theory of numerical analysis for SPDE. Most of the work contained in this section is rather heuristic, as the development of precise results can become quickly quite complicated.

Our main goal is to describe the FTCS scheme for numerically solving the stochastic heat equation. As a reminder, the SPDE we aim to solve is the following:

$$(249) \quad \frac{\partial U}{\partial t} = \alpha \frac{\partial^2 U}{\partial x^2} + f(t, x, U(t, x))W(t, x), \quad U(0, x) = u_0(x).$$

For simplicity, we will fix a variety of parameters: Assume that the domain of the problem is $[0, T] \times [-L/2, L/2]$ for $T > 0$ and $L > 0$, and impose periodic boundary conditions. Also take $\alpha = 1$ and assume that f and u_0 satisfy the hypotheses for the existence of strong solutions to the SPDE. Now let $(\Omega, \mathcal{F}, \mathbb{P})$ be any probability space on which a Brownian martingale measure $\{B(t, A)\}_{t, A}$ is defined.

We will use the following notation: Whenever we have fixed a time resolution N and space resolution M , we set $\Delta t = T/N$ and $\Delta x = L/M$ along with $t_i = i\Delta t$ for $i = 0, 1, \dots, N$ and $x_j = j\Delta x - L/2$ for $j = 0, 1, \dots, M$. It will also be convenient to reference the constant $r = \alpha\Delta t/(\Delta x)^2$ and the Brownian martingale measure increment $\Delta B_{i, A}(\omega) = B(t_{i+1}, A)(\omega) - B(t_i, A)(\omega)$.

The numerical scheme that we present in this section is discretized in both time and space, and is based on the forward-time, centered-space scheme for deterministic parabolic PDE. That is, it draws on the idea of one-step methods that we explored in Section 3 by approximating the derivatives in the SPDE with finite differences.

Definition 5.1. The *stochastic FTCS scheme* is a numerical scheme for the stochastic heat equation on $[0, T] \times [-L/2, L/2]$ given by the following recursive structure:

$$(250) \quad V_{0,j}(\omega) = u_0(x_j)$$

$$(251) \quad V_{i+1,j}(\omega) = r(V_{i,j-1}(\omega) + V_{i,j+1}(\omega)) + (1 - 2r)V_{i,j}(\omega)$$

$$(252) \quad + f(t_i, x_j, V_{i,j}(\omega))\Delta B_{i, (x_{j-1/2}, x_{j+1/2})}(\omega).$$

We may include a superscript (N, M) to denote the time- and space-resolutions if they are not clear from context.

Some remarks are necessary at this point: Since the values of $V_{i+1,j}(\omega)$ and $V_{i,j}(\omega)$ should be relatively close to each other, we need to have $r < 1/2$ so that the coefficient of the latter will be positive. Indeed, if $r \geq 1/2$ then this numerical scheme is unstable, meaning its errors propagate in an unbounded way across time. This condition, that $r < 1/2$, can be satisfied by maintining the asymptotic orders $N_x^2 \sim N_t$, which is often done in practice.

Additionally, we highlight the set that is included in the Brownian martingale measure difference: $(x_{j-1/2}, x_{j+1/2}]$. These are nothing more than intervals of width Δx which are centered at $\{x_j\}_j$, and this is quite natural since we have chosen to treat the variable x in a centered way with respect to its derivatives.

Finally, we point out that the value of a point in the stochastic FTCS scheme is just a linear combination of a few values at the previous time, plus an error depending on the value of the previous term. In particular, it can be written as a matrix multiplication so as to emphasize its structure as a one-step scheme in time. Specifically, write V_i as shorthand for $\{V_{i,j}\}_{j=0}^{N_x}$ and then we have

$$(253) \quad V_{i+1}(\omega) = AV_i(\omega) + f_i(\omega)^T \Delta B_i(\omega).$$

Here we have used the shorthand

$$(254) \quad f_i(\omega) = \begin{pmatrix} f(t_i, x_0, V_{i0}(\omega)) \\ f(t_i, x_1, V_{i1}(\omega)) \\ \vdots \\ f(t_i, x_{N_x-1}, V_{i,N_x-1}(\omega)) \end{pmatrix}, \quad \Delta B_i(\omega) = \begin{pmatrix} \Delta B_{i,(x_{0-1/2}, x_{0+1/2})}(\omega) \\ \Delta B_{i,(x_{1-1/2}, x_{1+1/2})}(\omega) \\ \vdots \\ \Delta B_{i,(x_{N_x-1/2}, x_{N_x+1/2})}(\omega) \end{pmatrix}$$

as well as:

$$(255) \quad A = \begin{pmatrix} 1-2r & r & 0 & 0 & \cdots & 0 & 0 & 0 & r \\ r & 1-2r & r & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & r & 1-2r & r & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & r & 1-2r & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1-2r & r & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & r & 1-2r & r & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & r & 1-2r & r \\ r & 0 & 0 & 0 & \cdots & 0 & 0 & r & 1-2r \end{pmatrix}$$

These formulas are very easy to implement in a program, and can be run very efficiently. See AppendixB for an implementation of the FTCS scheme in vectorized form, as well as the scripts used to simulate the plots of the previous section.

It does remain, however, to prove a precise statement about the (strong quadratic) order of convergence of the stochastic FTCS scheme. It is known that the FTCS scheme for deterministic parabolic PDE achieves an order of convergence equal to 1 in the time resolution and equal to 2 in the space resolution, so in the Ito calculus we expect that its (strong quadratic) orders of convergence will not exceed these. Perhaps Monte Carlo simulations can lead one to a more precise conjecture.

ACKNOWLEDGEMENTS

This thesis would not have been possible without plenty of support from my friends, family, and mentors. My math friends (Eric L., Eric K., Bora) helped me survive many long nights working on homework assignments over the last four years, and my thesis friends (Brad, Sienna, Atticus, Josh) helped me power through the research and writing process for this thesis during our last year. Additionally, my parents and sister showed me great support and encouragement as I wrote this thesis and applied to graduate schools. Finally, my advisor Lenya deserves a huge acknowledgement for his expertise in teaching, his genuine curiosity in my project, and his patience throughout the last year.

REFERENCES

- [1] D.L. Cohn. *Measure Theory: Second Edition*. Birkhäuser Advanced Texts Basler Lehrbücher. Springer New York, 2013.
- [2] R. Durrett. *Probability: Theory and Examples*. Cambridge series on statistical and probabilistic mathematics. Cambridge University Press, 2010.
- [3] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013.
- [4] Steve Lalley. Notes on the Ito Calculus, May 2012.
- [5] B. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer Berlin Heidelberg, 2013.
- [6] M. Reed and B. Simon. *Methods of Modern Mathematical Physics: Functional Analysis*. Methods of Modern Mathematical Physics. Academic Press, 1980.
- [7] A. Vasy. *Partial Differential Equations*. Graduate Studies in Mathematics. American Mathematical Society, 2015.
- [8] John B. Walsh. An introduction to stochastic partial differential equations. In P. L. Hennequin, editor, *École d'Été de Probabilités de Saint Flour XIV - 1984*, pages 265–439, Berlin, Heidelberg, 1986. Springer Berlin Heidelberg.

APPENDIX A. INEQUALITIES

Lemma A.1. If I is any finite index set and $\{c_\alpha\}_{\alpha \in I}$ is a collection of reals, then

$$(256) \quad \left| \sum_{\alpha \in I} c_\alpha \right|^2 \leq |I| \sum_{\alpha \in I} |c_\alpha|^2$$

Lemma A.2. If $\{Y_i\}_{i=1}^k$ are independent $N(0, \sigma^2)$ random variables, then

$$(257) \quad \mathbb{E} \left[\max_{j=1, \dots, k} |Y_j| \right] \leq C\sigma \left(\sqrt{\log k} + \frac{1}{\sqrt{\log k}} \right)$$

for some universal constant $C > 0$.

Lemma A.3 (Gronwall I). Let $\{u_n(t)\}$ be a sequence of nonnegative functions with u_0 constant. If there is some nonnegative constant C such that $\{u_n(t)\}$ satisfies the implicit inequality

$$(258) \quad u_{n+1}(t) \leq C \int_0^t u_n(s) ds,$$

then it satisfies the explicit inequality

$$(259) \quad u_n(t) \leq u_0 \frac{(Ct)^n}{n!}$$

for all $t \in [0, T]$.

Lemma A.4 (Gronwall II). Let $u(t)$ be any nonnegative function. If there is some nonnegative constants A, C such that $u(t)$ satisfies the implicit inequality

$$(260) \quad u(t) \leq A + C \int_0^t u(s) ds,$$

then it satisfies the explicit inequality:

$$(261) \quad u(t) \leq A \exp(Ct)$$

for all $t \in [0, T]$.

Lemma A.5 (Doob I). If $\{M_t\}$ is a submartingale on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then for all $\varepsilon > 0$ we have

$$(262) \quad \mathbb{P} \left(\sup_{0 \leq t \leq T} M_t \geq \varepsilon \right) \leq \frac{\mathbb{E}[M_T]}{\varepsilon}$$

Lemma A.6 (Doob II). If $\{M_t\}$ is a nonnegative submartingale on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then for all $\varepsilon > 0$ and $p > 1$ we have

$$(263) \quad \mathbb{E} \left[\sup_{0 \leq t \leq T} |M_t|^p \right] \leq \left(\frac{p}{p-1} \right)^p \mathbb{E}[M_T]$$

Lemma A.7 (Minkowski). For (X, μ_X) and (Y, μ_Y) any σ -finite measure spaces, any measurable function $f : X \times Y \rightarrow \mathbb{R}$, and any $p \in (1, \infty)$, we have:

$$(264) \quad \left(\int_Y \left| \int_X f(x, y) d\mu_X(x) \right|^p d\mu_Y(y) \right)^{\frac{1}{p}} \leq \int_X \left(\int_Y |f(x, y)|^p d\mu_Y(y) \right)^{\frac{1}{p}} d\mu_X(x)$$

When we take $X = \{1, 2\}$ with the counting measure, we recover the “usual” Minkowski inequality.

APPENDIX B. MATLAB CODE

This appendix contains the functions and scripts used to generate the plots in the preceding sections. For the sake of exposition, the code snippets contained in this appendix are presented in an ordering that deviates slightly from the chronological presentation in the main body of the text. All files were written to be executed with MATLAB 2017 under no additional licenses.

B.1. Brownian Motion. First we present the script used to generate the sample paths of the Brownian motion in Subsection 2.1. As expected, the simulation relies on simulating normal random variables to represent the increments of the Brownian motion, and then forming the cumulative sum of these.

% Simulating sample paths of the Brownian motion
% Plot in Subsection 2.1

% Parameters

```
T = 1;
N = 2^10;
ntrials = 5;
```

% The simulation

```
t = linspace(0, T, N+1);
Dt = T/N;
figure;
set(gcf, 'Position', [0, 0, 600, 300]);
hold on; box on;
for trial = 1:ntrials
    bm = cumsum([0 normrnd(0, sqrt(Dt), [1, N])]);
    plot(t, bm);
end
```

```
title('Brownian Motion');
xlabel('time (t)')
```

Next we describe the script used to generate the plot at the end of Subsection 3.2, in which we choose the resolution of a numerical scheme probabilistically according to some stopping rule. In particular, the following implementation is straightforward:

```
% Probabilistically choosing the resolution of
% a Brownian motion
% Plot in Subsection 3.2

% Paramters
T = 1;
Nmin = 2^0;
tol = 2^-4;

% The simulation
N = Nmin;
Dt = T/N;
t = linspace(0,T,N+1);
t_old = t;
bm_approx = [0 cumsum(normrnd(0,sqrt(Dt),[1,N]))];
bm_approx_old = bm_approx;
refine = Inf(1,N+1);

figure;
set(gcf, 'Position', [0, 0, 600, 300])
hold on; box on;
while max(abs(refine)) >= tol
    plot(t,bm_approx,'Color',[log2(N)/(15),0,1-log2(N)/(15)]);
    t_old = t;
    bm_approx_old = bm_approx;
    Dt = T/N;
    t = linspace(0,T,2*N+1);
    bm_approx = interp1(t_old,bm_approx_old,t);
    refine = normrnd(0,sqrt(Dt/4),[1,N]);
    bm_approx(2:2:end) = bm_approx(2:2:end) + refine;
    N = 2*N;
end
title('Simulated Brownian Motion');
xlabel('time (t)')
```

B.2. SODEs. The next collection of code snippets was used to simulate solutions to SODEs throughout the text. First, we present the function used to implement the one-step schemes outlined in Subsection 3.3; For an appropriate choice of Ψ , the following can be used to implement the Euler-Maruyama scheme and the Milstein scheme.

```
function num_sol = one_step_scheme(Psi,N,t,bm,u0)
    % Computes the numerical solution to an SDE
```

```

% using a one-step scheme.
% INPUTS
% Psi - The function determining the one-step scheme
% N - The number of time steps
% t - The discretized time values
% bm - The driving Brownian motion
% u0 - The initial condition
% OUTPUTS
% num_sol - the numerical solution
num_sol = zeros(1,N+1);
num_sol(1)= u0;
for ii = 2:N+1
    delt_bm = bm(ii) - bm(ii-1);
    delt_t = t(ii) - t(ii-1);
    num_sol(ii) = num_sol(ii-1) ...
        + Psi(t(ii-1),num_sol(ii-1),delt_t,delt_bm);
end
end

```

Now we move on to the scripts used to simulate the solutions to example SODEs in Subsection 2.4. The following script can be used to generate all of Examples 2.21, 2.22, and 2.23:

```

% Simulating the solutions to a few SODEs
% Plots in Subsection 2.4

% SODE parameters
SODE = 'Geometric Brownian Motion';
if strcmp(SODE,'Geometric Brownian Motion') % Example 2.21
    T = 1;
    mu = @(t,x) -1*x;
    sigma = @(t,x) 2*x;
    u0 = 1;
elseif strcmp(SODE,'Brownian Bridge') % Example 2.22
    T = 1;
    mu = @(t,x) -x./(1-t);
    sigma = @(t,x) 1;
    u0 = 0;
elseif strcmp(SODE,'Dynamical System') % Example 2.23
    T = 5;
    mu = @(t,x) cos(pi*t);
    sigma = @(t,x) sin(pi*t.*x);
    u0 = 0;
    ngridlines = 10;
else
    throw(MException('Unknown SODE'))
end

% Numerical parameters
N = 2^10;

```

```

EM = @(t,x,dx) mu(t,x)*dt + sigma(t,x)*dx;

% Display paramters
ntrials = 5;

% The simulation
t = linspace(0,T,N+1);
Dt = T/N;
figure;
set(gcf, 'Position', [0, 0, 600, 300])
hold on; box on;
for trial = 1:ntrials
    bm = cumsum([0 normrnd(0,sqrt(Dt),[1,N]))]);
    sol = one_step_scheme(EM,N,t,bm,u0);
    plot(t,sol);
end

if strcmp(SODE,'Dynamical System')
    axis manual;
    for k = -ngridlines:ngridlines
        gridline = k./t;
        plot(t,gridline,'Color',[0.8,0.8,0.8]);
    end
end

title(SODE);
xlabel('time (t)')

```

The last script we present in this section is the one used to compare the empirical order of convergence of the . Euler-Maruyama and Milstein schemes at the end of Subsection 3.3:

```

% Comparing the empirical order of convergence for
% the Euler-Maruyama and Milstein schemes, using
% the Geometric Brownian motion as case study.
% Plot in Subsection 3.3

% SODE paramters
T = 1;
m = -1;
s = 2;
mu = @(t,x) m*x;
sigma = @(t,x) s*x;
sigma_deriv = @(t,x) s;
u0 = 1;

% Experiment parameters
Nmax = 2^12;
Ns = 2.^linspace(2,12,11)';
ntrials = 100;

```

```

% Numerical schemes
EM = @(t,x,dt,dx) mu(t,x)*dt + sigma(t,x)*dx;
Milstein = @(t,x,dt,dx) EM(t,x,dt,dx) ...
    + 0.5*sigma(t,x).*sigma_deriv(t,x)*((dx)^2-dt);

% The simulation
nNs = length(Ns);
EM_errors = zeros(nNs,ntrials);
Milstein_errors = zeros(nNs,ntrials);
for j = 1:ntrials
    Dt = T/Nmax;
    bm_full = cumsum([0 normrnd(0,sqrt(Dt),[1,Nmax])]);
    t_full = linspace(0,T,Nmax+1);
    for i = 1:nNs
        N = round(Ns(i));
        t = downsample(t_full,Nmax/N);
        bm = downsample(bm_full,Nmax/N);

        sol_EM = one_step_scheme(EM,N,t,bm,u0);
        sol_Milstein = one_step_scheme(Milstein,N,t,bm,u0);
        sol_explicit = u0*exp((m-s^2/2)*t+s*bm);

        EM_errors(i,j) = max((sol_EM - sol_explicit).^2);
        Milstein_errors(i,j) = max((sol_Milstein - sol_explicit).^2);
    end
end

avg_EM_errors = mean(EM_errors,2);
avg_Milstein_errors = mean(Milstein_errors,2);

% Plot the results
figure;
set(gcf, 'Position', [0, 0, 600, 300])
hold on; box on;

log_avg_EM_errors = log2(avg_EM_errors);
log_avg_Milstein_errors = log2(avg_Milstein_errors);
log_Ns = log2(Ns);
X = [ones(nNs,1) log_Ns];

p_EM = scatter(log_Ns,log_avg_EM_errors,'blue');
p_Milstein = scatter(log_Ns,log_avg_Milstein_errors,'red');

LR_EM = X\log_avg_EM_errors; % linear regression
plot(log_Ns,X*LR_EM,'b--');

```

```

LR_Milstein = X\log_avg_Milstein_errors; % linear regression
plot(log_Ns,X*LR_Milstein,'r--');

xlabel('log Resolution');
ylabel('log Maximum Square Error');

legend([p_EM,p_Milstein],{'Euler-Maruyama','Milstein'}, ...
       'Location','southwest')
title('Empirical Order of Convergence of Numerical Schemes');

```

B.3. SPDEs. Finally, we present the code used to simulate solutions to SPDEs. Our first function is the algorithm for the stochastic FTCS scheme:

```

function num_sol = stoch_FTCS(f,u0,T,X,W,boundary)
    % The stochastic FTCS scheme for computing the numerical
    % solution to the stochastic heat equation.
    % INPUTS
    % f      - the multiplicative inhomogeneity function
    % u0     - the initial condition
    % X      - the time domain
    % T      - the space domain
    % W      - the driving Wiener sheet
    % boundary - the boundary conditions
    % OUTPUTS
    % num_sol - the numerical solution
    t = T(1,:);
    x = X(:,1);
    Nt = length(t);
    Nx = length(x);
    Dt = t(2)-t(1);
    Dx = x(2)-x(1);
    num_sol = zeros(Nx,Nt);

    r = Dt/(Dx^2);
    if r >= 0.5
        warning('Discretization ratio (%.3f) may lead to numerical instability.', r);
    end

    A = diag((1-2*r)*ones(Nx,1)) + diag(r*ones(Nx-1,1),1) +diag(r*ones(Nx-1,1),-1);
    if boundary == 'periodic'
        A(1,end) = r;
        A(end,1) = r;
    elseif boundary == 'dirichlet'
        A(1,:) = zeros(1,Nx);
        A(1,1) = 1;
        A(end,:) = zeros(1,Nx);
        A(end,end) = 1;
    else
        throw(MException('Unknown Boundary conditions'))
    end
end

```

```

num_sol(:,1) = u0;
for ii=2:Nt
    num_sol(:,ii) = A*num_sol(:,ii-1) + f(num_sol(:,ii-1)).*W(:,ii-1);
end
end

```

Using this, we can use the following script to simulate our solutions to the stochastic heat equation:

```

% Simulating the solutions to the stochastic heat equation
% on the bounded domain with periodic boundary conditions
% Plots in Subsection 4.4

```

```

% SPDE Parameters

```

```

T = 1;
L = 4;
u0 = @(x) (x > -L/4) & (x < L/4);
SPDE = 'Multiplicative Noise';
if strcmp(SPDE,'Heat Equation')
    f = @(u) 0;
elseif strcmp(SPDE,'Additive Noise') % Example 4.17
    sigma = 1;
    f = @(u) sigma;
elseif strcmp(SPDE,'Multiplicative Noise') % Example 4.18
    sigma = 5;
    c = 0;
    f = @(u) sigma*abs(u-c);
else
    throw(MException('Unknown SPDE'))
end

```

```

% Numerical Parameters

```

```

Nt = 16001;
Nx = 40;

```

```

% The simulation

```

```

Dt = T/Nt;
Dx = L/Nx;
t = linspace(0,T,Nt+1);
x = linspace(-L/2,L/2,Nx+1);
[tvals,xvals] = meshgrid(t,x);
W = normrnd(0,sqrt(Dt*Dx),[Nx+1,Nt+1]);
sol = stoch_FTCS(f,u0(x),tvals,xvals,W,'periodic');

```

```

figure;
set(gcf, 'Position', [0, 0, 600, 300])
hold on; box on;
surf(tvals,xvals,sol,'EdgeColor','none');
xlabel('time (t)')

```

```
ylabel('space (x)')

if strcmp(SPDE,'Heat Equation')
    title('Heat Equation');
elseif strcmp(SPDE,'Additive Noise')
    title(sprintf('Additive Noise,  $\sigma = %0.2f$ ',sigma));
elseif strcmp(SPDE,'Multiplicative Noise')
    title(sprintf('Multiplicative Noise,  $c = %0.2f$ ',c));
end
view([30,45])
axis([0 T -L/2 L/2 -1 2])
```