Brief Review of Yamada-Watanabe Theorem

Second Homework of Martingale Theory of Stochastic Integration

Beining Wu*

January 6, 2022

§1   Introduction, Problem Settings

1.1   Notations and Preliminaries

We begin with several notations that would be frequently used in this essay.

Filtered Space   Our effort is generally based on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$. Here $\mathscr{F}_t$ is an increasing σ-algebra, while $\mathbb{P}$ is the probability measure defined on the measurable space $(\Omega, \mathscr{F})$. Often we would pay attention to the complete filtration, where all the null sets are measurable. Definition of measurability and σ-algebra can be found in the classical text [Durrett, 2019].

Wiener Space   Let $W := C(\mathbb{R}^+, \mathbb{R})$ be the space of all continuous functions on $\mathbb{R}^+$. The measurable structure on W is defined by with the canonical process $(X(w))_t := w(t)$. Define the filtration as $\mathscr{B}_t := \sigma\langle X_s, s \leq t\rangle$, and the σ-algebra as $\mathscr{B}(W) := \sigma\langle \mathscr{B}_t, t \geq 0\rangle$. One could check that this σ-algebra coincides with the Borel σ-algebra on W, where the topology is induced form the metric

$$d(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\sup_{[0,k]} |x(t) - y(t)|}{1 + \sup_{[0,k]} |x(t) - y(t)|}.$$

This space is usually used to construct our canonical probability space for the processes. We will use W^d and $\mathscr{B}(W^d)$ for the multidimensional version.

Process   A stochastic process $(X_t)_{t\geq 0}$ is an indexed collection of random variables, or function-valued random elements. A process is **adapted** if, for every $t \geq 0$ we have X_t is $\mathscr{F}_t$ measurable. It's **progressive** if, for every $t \geq 0$, the mapping $X : (\omega, t) \mapsto X_t(\omega)$ is $\mathscr{F}_t \otimes \mathscr{B}([0, t])$. Moreover, we can define the **progressive σ-algebra** $\mathscr{P}$ as the sets of all the sets $A \in \mathscr{F} \otimes \mathscr{B}(\mathbb{R}^+)$ such that the process $1_A(\omega, t)$ is progressive. One can check that a process if progressive iff it's $\mathscr{P}$ measurable.

Brownian Motion and Wiener Measure   A $\mathscr{F}_t$-Brownian motion $(B_t)_{t\geq 0}$ is a $\mathscr{F}_t$ adapted process with independent increments and continuous sample paths. Moreover

$$B_t - B_s \sim N(0, t - s).$$

For the construction and properties of Brownian motion, one can refer to [Gall, 2016] and [Rogers and Williams, 2000a]. The wiener measure is the law of Brownian motion on W, i.e., the push-forward probability measure on W.

1.2   What is Stochastic Differential Equations?

Cited from [Gall, 2016], the theory of stochastic differential equations aims to provide a model for a ordinary differential equation with a random perturbation. An simple ordinary equation can be formulated as

$$y'(t) = b(y(t)),$$

---

*mail: `andrewwu@mail.ustc.edu.cn`, Student ID: PB19151833

or in the differential form,

$$\mathrm{d}y_t = b(y_t)\mathrm{d}t.$$

Now we would involve the random noise term. Clearly a Brownian motion would be an ideal noise for its properties of independent increments. The model can be unrigorously written as

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma \mathrm{d}B_t,$$

where σ is a fixed constant.

Our model would be a bit more complex than the aforementioned one.

Definition 1 (Stochastic Differential Equations)**.** Let $\sigma : \mathbb{R}^+ \times \mathbb{R}^n \to L(\mathbb{R}^d; \mathbb{R}^n)$, which is the matrix valued mapping, and $b : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ be the locally bounded measurable functions. The stochastic differential equation we are going to study is of the form

$$X_t = X_0 + \int_0^t \sigma(s, X_s)\mathrm{d}B_s + \int_0^t b(s, X_s)\mathrm{d}s, \tag{1}$$

or written in coordinate form

$$X_t^i = X_0^i + \int_0^t \sum_{j=1}^d \sigma(t, X_s)_{ij}\mathrm{d}B_s^j + \int_0^t b(s, X_s)_i \mathrm{d}s.$$

Remark (Technical Notations)**.** Here are some assumptions that would be necessary for the definitions to be rational. First is measurability assumptions. Recall that in the definition of the stochastic integration, the integrand processes should be progressive and consequently the integral processes are adapted. The following assumption if from [Rogers and Williams, 2000b], wherein the settings are more complex.

Assumption (Progressive Functional)**.** The mapping σ and b are the progressive functional, which means, the process $\sigma(t, X_t)$ and $b(t, X_t)$ are progressive if X_t is adapted.

Another assumption comes from the rationality of stochastic integration. Note that we have required σ and b to be locally bounded. This could be replaced with a more general assumption.

Assumption (Finiteness)**.** The mapping σ and b should follow

$$\forall t \geqslant 0 \qquad \int_0^t \left(\|\sigma(s, X_s)\|_F^2 + |b(s, X_s)| \right) \mathrm{d}s < \infty \quad \text{a.s.}$$

But as [Rogers and Williams, 2000b] noted, this finiteness requirement is nothing to worry much about, because we can use the stopping time trick to restrict our attention before explosion. In rest of this essay, we assume that this finiteness assumption is always true for the solution.

1.3   What We Want to Study?

From [Gall, 2016] we have known that, if we assume σ and b to be Lipschitz continuous, then the pathwise uniqueness easily follows, and we can construct the solution with Picard approximation series. However there are plenty of cases where Lipschitz continuity fails but pathwise uniqueness still holds. In Yamada and Watanabe's classical paper [Yamada and Watanabe, 1971], they proved that the pathwise uniqueness implies the weak uniqueness, and moreover, there exists a strong solution. In our essay, we will go over the definition of the uniqueness and existence of 1. Then we will go over the proof of the classical Yamada-Watanabe theorem.

§2   Definition of Existence and Uniqueness

In this section, we will quickly go over the definition of uniqueness and existence in different senses. Here and after, we will use the term **set-up** $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geqslant 0}, \mathbb{P}, \xi, B)$, where ξ is a $\mathscr{F}_0$ measurable random variable and $B = (B_t)_{t \geqslant 0}$ is the $\mathscr{F}_t$ Brownian motion. One of the crucial concepts we will encounter frequently is the canonical set-up, which is a generalization of the Wiener space.

Definition 2 (Canonical Space and Set-up)**.** In the following texts, a **canonical set-up** with initial distribution μ means:

- A probability space: $\Omega = \mathbb{R}^n \times W^d$, which stand for the initial value and Brownian motion trajectory space.

- Probability measure: $\mathbb{P} = \mu \times \mathbb{W}$, where $\mathbb{W}$ is the Wiener measure on W^d.

- Canonical initial value and Brownian paths: $\xi(y, w) = y$ and $B(y, w)_t = w(t)$ for $(y, w) \in \Omega$.

- Filtration: $\mathscr{F}_t = \overline{\sigma \langle \xi, B_s, s \leqslant t \rangle}$, the completion with $\mathbb{P}$-null sets.

2.1 Weak Scenario

We begin with the weak scenario, where all the existence and uniqueness are defined in the sense of distribution.

Definition 3 (Weak Solution). We say that the SDE

$$X_t = X_0 + \int_0^t b(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}B_s \tag{2}$$

has a weak solution with initial distribution μ, if there exists

- a $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geqslant 0}, \mathbb{P})$ that satisfy the usual conditions.

- a $\mathscr{F}_t$-Brownian motion B_t

- a semimartingale X_t such that equation 2 and assumption 1.2 hold, and X_0 has law μ.

Definition 4 (Weak Uniqueness). We say that the solution (weak in this sense) of the equation 2 is weakly unique, or unique in distribution, if whenever X_t and X'_t are both the weak solution of 2 with same initial distribution, then X_t and X'_t are identical in distribution.

2.2 Strong Scenario

Definition 5 (Pathwise Uniqueness). We shall say that the **pathwise uniqueness** holds for equation 1, if for fixed setup $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geqslant 0}, \mathbb{P}, \xi, B)$ and any two continuous semimartingales X and X' such that

$$X_t = \xi + \int_0^t b(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}B_s,$$
$$X'_t = \xi + \int_0^t b(s, X'_s)\mathrm{d}s + \int_0^t \sigma(s, X'_s)\mathrm{d}B_s,$$

then we have

$$\text{a.s.} \quad X_t = X'_t, \ \forall t \geqslant 0$$

In this case, we can conclude that given any set-up, there is at most one semimartingale X_t (unique up to indistinguishablity) that solves the equation. We shall define the equation that has exactly one solution as the following.

Definition 6 (Exact SDE). The equation 1 is **exact** if given any setup $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geqslant 0}, \mathbb{P}, \xi, B)$, there is exactly one semimartingale that solves the equation 1.

Clearly the pathwise uniqueness is true for exact SDEs. Note that in the definition of exact SDEs, we require no restriction on the given set-up, which means we could adopt the minimal set-up, generated by ξ and B. This follows the definition of strong solution. Briefly, a strong solution is adapted to the canonical filtration of the driven Brownian motion.

Definition 7. Given set-up $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geqslant 0}, \mathbb{P}, \xi, B)$. A strong solution of the SDE 1 is adapted with respect to the completed canonical filtration generated form (ξ, B).

§3 Main Result: Yamada-Watanabe Theorem

In this section we give the formal statement of the famous Yamada-Watanabe theorem.

Theorem 1 (Yamada-Watanabe [Yamada and Watanabe, 1971]). For progressive path functional σ and b, the stochastic differential equation

$$\mathrm{d}X_t = \sigma(t, X_t)\mathrm{d}B_t + b(t, X_t)\mathrm{d}B_t$$

is exact if the following two conditions hold

- The SDE 1 has at least one weak solution.

- The SDE 1 shares the pathwise uniqueness property.

Moreover, one can conclude weak uniqueness from pathwise uniqueness.

If the 1 is exact, then these conditions easily follow. We shall only consider the other direction.

Before we begin our proof, we shall introduce a few technical lemmas, which would be the stepping stones in our proof. The first one (Theorem 10.4 in [Rogers and Williams, 2000b]) enables us to move from canonical set-up to the general set-up.

Theorem 2. Suppose that on the canonical set-up with initial distribution μ the SDE 1 has a solution

$$X_t = \xi + \int_0^t \sigma(s, X_s)\mathrm{d}B_s + \int_0^t b(s, X_s)\mathrm{d}s.$$

Then there exists a function $F_\mu : \mathbb{R}^n \times W^d \to W^n$, such that for all $t \geqslant 0$,

$$F_\mu^{-1}(\mathcal{B}_t(W^n)) \subset \mathcal{F}_t.$$

Such that a.s. $F_\mu(\xi, B) = X$. Moreover, on ant set-up $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geqslant 0}, \tilde{\xi}, \tilde{B})$, the process $\tilde{X} = F_\mu(\tilde{\xi}, \tilde{B})$ solves

$$\tilde{X} = \tilde{\xi} + \int_0^t \sigma(s, \tilde{X}_s)\mathrm{d}B_s + \int_0^t b(s, \tilde{X}_s)\mathrm{d}s.$$

Proof. See [Rogers and Williams, 2000b, Theorem 10.4]. ∎

This theorem means, if for any initial distribution μ we can find a solution on the canonical set-up with initial distribution μ, then we can find the solution on any set-up with a transition function.

Before beginning our formal proof, a technical lemma would be introduced here. This lemma is somehow critical, as it clarifies the stability of stochastic integral with respect to the Brownian motion under different set-ups.

Lemma 1. On canonical set-up, if $\varphi : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$ is $\mathscr{P}/\mathscr{B}(\mathbb{R}^d)$ measurable, and satisfies the integrability condition

$$\mathbb{P}\text{-a.s.} \quad \int_0^t |\varphi(s; \xi, B)|^2 \mathrm{d}s < \infty, \quad \forall t \geqslant 0.$$

Then there exists a function $\Phi : \mathbb{R}^n \times W^d \to W^1$ which is $\mathscr{B}(\mathbb{R}^n) \times \mathscr{B}(W^d)/\mathscr{B}(W^1)$ measurable such that

$$\mathbb{P}\text{-a.s.} \quad \int_0^\cdot \varphi(s; \xi, B)\mathrm{d}s = \Phi(\xi, B), \quad \forall t \geqslant 0.$$

Moreover, on any set-up $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geqslant 0}, \tilde{\xi}, \tilde{B})$, we have

$$\tilde{\mathbb{P}}\text{-a.s.} \quad \int_0^\cdot \varphi(s; \tilde{\xi}, \tilde{B})\mathrm{d}s = \Phi(\tilde{\xi}, \tilde{B}), \quad \forall t \geqslant 0.$$

Proof. See, [Rogers and Williams, 2000b, Lemma 10.1]. ∎

In the rest of our essay, we would follow [Rogers and Williams, 2000b] to prove this main theorem in two parts. In the first part, we would deduce that the pathwise uniqueness implies weak uniqueness.

§4 From the Pathwise Uniqueness to Weak Uniqueness

The intrinsic problem is that weak uniqueness **does not** require the set-up to be the same, instead the pathwise uniqueness does. This key difference motivates us to construct an associated processes which share the same set-up.

Proof of Part 1. Let μ be the initial distribution, then by the weak existence condition, there exists a weak solution (X, B) (a set-up, actually). Let π be the distribution of (X_0, X, B), i.e., a push forward probability measure on $\mathbb{R}^n \times W^n \times W^d$. And define the regular conditional probability $\mathbb{Q} : (\mathbb{R}^n \times W^d) \times \mathscr{B}(W^n) \to [0, 1]$. This is the conditional distribution of solution process given the initial value and trajectory $x_0 \in \mathbb{R}^n, w \in W^d$. Let (X', B') be another weak solution, and also define the same conditional distribution.

Now we construct the probability space
$$S = \mathbb{R}^n \times W^d \times W^n \times W^n,$$
with probability measure
$$\overline{\mathbb{P}}(dy, d\beta, dx, dx') = \mathbb{W}(d\beta)\mu(dy)\mathbb{Q}((y, \beta), dx)\mathbb{Q}'((y, \beta), dx').$$

Here $\mathbb{W}$ is the Wiener measure on W^d. Then we define the filtration on S to be
$$\mathscr{G}_t := \overline{\sigma \langle \beta_s, x_s, x'_s, s \leqslant t \rangle},$$
which is the canonical filtration completed with $\overline{\mathbb{P}}$-null sets. For $s = (y, \beta, x, x') \in S$, define the canonical process
$$\beta_t(s) := \beta(t), \quad x_t(s) := x(t), \quad x'_t(s) = x'(t), \quad \xi(s) = y$$

Then we have the following lemma, which we will prove later.

Lemma 2. The canonical process β_t is a $\mathscr{G}_t$-Brownian motion.

Proof. See [Rogers and Williams, 2000b, Lemma 17.12]. ∎

Clearly the $\overline{\mathbb{P}}$ marginal distribution of (x_0, x, β) is the same as π, and we can prove that
$$x_t = \xi + \int_0^t \sigma(s, x_s)d\beta_s + \int_0^t b(s, X_s)ds.$$

And same equation for x'. Now x' and x are the solutions of the same set-up, from pathwise uniqueness we have
$$\overline{\mathbb{P}} - \text{a.s.} \quad x_t = x'_t, \ \forall t \geqslant 0.$$

This implies that the $\overline{\mathbb{P}}$-marginal distribution of (y, β, x) and (y, β, x') are identical, and moreover identical to π, π'. Hence $\pi \sim \pi'$, thus the uniqueness in distribution is proved. ∎

A byproduct of this proof is that, we can express x in terms of (y, β).

Corollary 1. There exists a function $F_\mu : \mathbb{R}^n \times W^d \to W^n$ such that
$$F_\mu(y, \beta) = x.$$

Moreover, F_μ is $\overline{\mathscr{B}(\mathbb{R}^n) \times \mathscr{B}(W^d)}/\mathscr{B}(W^n)$-measurable.

Proof of Corollary. From the proof in the last theorem, we know that conditional distribution $\mathbb{Q} \times \mathbb{Q}'$ concentrates on the diagonal $\{x = x'\} =: D \subset W^n \times W^n$. This means, for $\mu \times \mathbb{W}$-a.s. $(y, \beta) \in \mathbb{R}^n \times W^d$, there exists $x = x' \in W^n$ such that $\mathbb{Q}((y, \beta), \cdot) = \delta_x(\cdot)$ and $\mathbb{Q}'((y, \beta), \cdot) = \delta_{x'}(\cdot)$. Otherwise the product would assign positive mass off the diagonal. Define the function $F_\mu : (y, \beta) \mapsto x$, then it's clearly measurable to the completed σ-algebra $\overline{\mathscr{B}(\mathbb{R}^n) \times \mathscr{B}(W^d)}$. ∎

§5 Existence of Strong Solution

Now for any initial distribution μ, we shall find the solution on the canonical set-up with initial distribution μ. Once this is done, theorem 2 implies that there is a strong solution with any set-up.

Proof of the Part 2. Consider the function F from the last corollary. We shall prove that,

$$F_\mu(\xi, B)$$

solves the equation on the canonical set-up. From [Rogers and Williams, 2000b, Lemma 17.9], we know that this function satisfies

$$F_\mu^{-1}(\mathscr{B}_t(W^n)) \subset \mathscr{F}_t.$$

Consider the map

$$\varphi(s; \xi, B) = \sigma(s, F_\mu(\xi, B)).$$

This map satisfies all the conditions of lemma 1. Hence there exists a map $\Phi : \mathbb{R}^n \times W^d \to W^n$, such that on any set-up $(\Omega', \mathscr{F}', (\mathscr{F}'_t), \mathbb{P}', \xi', B')$, we have

$$\mathbb{P}'\text{-a.s.} \qquad \int_0^\cdot \sigma(s, F_\mu(\xi', B'))\mathrm{d}B'_s = \Phi(\xi', B').$$

Consider, on the space $(S, \mathscr{G}, \overline{\mathbb{P}}, y, \beta)$ that was defined in the last part, we have

$$\overline{\mathbb{P}}\text{-a.s.} \quad F_\mu(y, \beta) = y + \Phi(y, \beta) + \int_0^\cdot b(s, F_\mu(y, \beta))\mathrm{d}s.$$

Since $\overline{\mathbb{P}}$ distribution of (y, β) is identical to the $\mathbb{P}$ distribution of (ξ, β), we now have

$$\mathbb{P}\text{-a.s.} \quad F_\mu(\xi, B) = \xi + \Phi(\xi, B) + \int_0^\cdot b(s, F_\mu(\xi, B))\mathrm{d}s.$$

From the definition of Φ we have

$$\mathbb{P}\text{-a.s.} \quad F_\mu(\xi, B) = \xi + \int_0^t \sigma(s, F_\mu(\xi, B))\mathrm{d}B_s + \int_0^\cdot b(s, F_\mu(\xi, B))\mathrm{d}s.$$

Which means $F_\mu(\xi, B)$ solves the equation on the canonical set-up. Now by the virtue of theorem 2, the equation is exact. ∎

References

[Durrett, 2019] Durrett, R. (2019). *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

[Gall, 2016] Gall, J. (2016). *Brownian Motion, Martingales, and Stochastic Calculus*. Graduate Texts in Mathematics. Springer International Publishing.

[Rogers and Williams, 2000a] Rogers, L. and Williams, D. (2000a). *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*. Cambridge Mathematical Library. Cambridge University Press.

[Rogers and Williams, 2000b] Rogers, L. and Williams, D. (2000b). *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*. Cambridge Mathematical Library. Cambridge University Press.

[Yamada and Watanabe, 1971] Yamada, T. and Watanabe, S. (1971). On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1):155 – 167.