

# Stochastic Processes and Branching Brownian Motion

by

Abhijeet Anand Mulgund  
B.A., Rice University, 2019

THESIS

Submitted in partial fulfillment of the requirements  
for the degree of Masters of Science in Mathematics  
in the Graduate College of the  
University of Illinois at Chicago, 2022

Chicago, Illinois

Defense Committee:  
William Perkins, Chair and Advisor  
Cheng Ouyang  
Marcus Michelen

Copyright by  
Abhijeet Anand Mulgund  
2022

TODO: A brief dedication to someone you care about. For example, “Dedicated to my cats, Neo and Trinity, who are purrfect in every way.”.

An example of “Dedication” can be found on page 15 of the thesis manual<sup>1</sup>.

---

<sup>1</sup>[http://grad.uic.edu/sites/default/files/pdfs/ThesisManual\\_rev\\_06Oct2016.pdf](http://grad.uic.edu/sites/default/files/pdfs/ThesisManual_rev_06Oct2016.pdf)

## ACKNOWLEDGMENT

TODO: A page or two so of shout-outs to people you appreciate. Don't forget your advisor and committee members!

PES

## CONTRIBUTIONS OF AUTHORS

This Masters Thesis is a culmination of my studies on various topics of Stochastic Processes, Martingales, and Branching Brownian Motion. It is by no means a comprehensive study, though it hopefully can serve as a resource to others who wish to learn more about these topics.

## TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
<b>1 STOCHASTIC PROCESSES AND MARTINGALES . . . . .</b>	<b>1</b>
1.1 Conditional Expectation . . . . .	1
1.2 Regular Conditional Probabilities . . . . .	3
1.3 Martingales . . . . .	3
1.4 Decomposition of Discrete-Time martingales . . . . .	7
1.5 Strategies and Discrete-Time martingales . . . . .	14
1.6 Convergence of Discrete-time Martingales . . . . .	19
1.7 Continuous-time martingales, definitions . . . . .	20
1.8 Continuous-time martingales, stopping times . . . . .	20
1.9 Convergence of Continuous-time Martingales . . . . .	20
1.10 Polya's Urn . . . . .	20
1.11 Borel-Cantelli Lemma . . . . .	20
1.12 Differential Equation Method? . . . . .	20
<b>2 MARKOV PROCESSES . . . . .</b>	<b>21</b>
2.1 Basics . . . . .	21
<b>3 BROWNIAN MOTION AND THE HEAT EQUATION . . . . .</b>	<b>22</b>
3.1 Basics . . . . .	22
3.2 Heat Equation . . . . .	22
3.3 Brownian Motion as a Solution . . . . .	22
<b>4 BRANCHING BROWNIAN MOTION . . . . .</b>	<b>23</b>
4.1 Basics . . . . .	23
4.2 F-KPP Equation . . . . .	23
4.3 McKean Representation . . . . .	23
4.4 Kolmogorov's Result . . . . .	23
<b>CITED LITERATURE . . . . .</b>	<b>24</b>
<b>VITA . . . . .</b>	<b>25</b>

## TABLE OF CONTENTS (Continued)

CHAPTER

PAGE

## SUMMARY

We will work to first build up the basic theory of Stochastic Processes and Martingales. We will establish conditional expectation, but take most of the basic results and relevant measure theoretic results for granted. From there, we will turn to martingales and start with studying "strategies" on martingales and how they help us prove convergence. From there we will look into decomposition of processes in terms of martingales and a handful of important examples.

With an understanding of martingales, we will then turn to Markov Processes. [[TODO]]

Brownian Motion is both a martingale and markov process, so our prior work will help elucidate some of its properties. Unlike in the previous sections, however, we will not study brownian motion in isolation. Instead, we will examine its deep relationship with a deterministic object, the heat equation.

Finally, we will put together our work from the previous 3 chapters towards a study of Branching Brownian Motion. Branching Brownian motion involves Brownian motions with lifetimes determined by a standard exponential distribution. Upon death, a Brownian motion will split into two Brownian motions, and so on. Like with regular Brownian motion, we will find an interesting connection with diffusion partial differential equations. With a basic understanding of this connection in hand, we will use it to study the distribution of the maximal point of a standard branching Brownian motion.



## CHAPTER 1

### STOCHASTIC PROCESSES AND MARTINGALES

#### 1.1 Conditional Expectation

Suppose we are given a probability space  $(\Omega, \mathcal{B}, \mathbb{P})$  and a random variable  $X \in L^1(\mathcal{B})$ . We would like to describe the operation of "viewing" this random variable from a sub-sigma algebra  $\mathcal{G} \subset \mathcal{B}$ . We call this operation *conditional expectation* and define it as follows:

**Definition 1.1.1.** *The **conditional expectation** of  $X$  given  $\mathcal{G}$ , denoted  $\mathbb{E}(X|\mathcal{G})$ , is a random variable with the following properties*

1.  $\mathbb{E}(X|\mathcal{G}) \in \mathcal{G}$
2.  $\int_A X d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{G}) d\mathbb{P}$  for all  $A \in \mathcal{G}$ .

The existence and  $\mathbb{P}$ -a.s. uniqueness of a conditional expectation follows from the following measure-theoretic theorem.

**Theorem 1.1.2** (Radon-Nikodym Theorem (1)). *Let  $(X, \mathcal{M})$  be a measure space with sigma-finite nonnegative measure  $\mu$  and sigma-finite signed measure  $\nu$  so that  $\nu \ll \mu$ . Then there exists  $\mu$ -a.e. unique  $f \in L^1(\mathcal{M})$  so that for all  $A \in \mathcal{M}$ ,*

$$\int_A f d\mu = \nu(A).$$

We denote  $\frac{d\nu}{d\mu} := f$ .

A thorough proof of the above result can be found in (1). Note that  $\nu(A) = \int_A X d\mathbb{P}$  is a signed measure on  $(\Omega, \mathcal{B})$  and  $\mathbb{P}$  is a finite measure. Restricting  $\nu$  to  $\mathcal{G}$  does not change this fact, and applying 1.1.2 to  $\nu|_{\mathcal{G}}$  gives us our  $\mathbb{P}$ -a.s. unique conditional expectation

$$\mathbb{E}(X|\mathcal{G}) = \frac{d\nu|_{\mathcal{G}}}{d\mathbb{P}|_{\mathcal{G}}}$$

in  $L^1(\mathcal{G})$ .

Conditional expectation has many useful properties, some of which generalize from expectation:

**Proposition 1.1.3.** *Let  $X, Y \in L^1(\mathcal{B})$ .  $\mathbb{E}(\cdot|\mathcal{G})$  has the following properties*

1. *Linearity: if  $c \in \mathbb{R}$ , we have*

$$\mathbb{E}(X + cY|\mathcal{G}) = \mathbb{E}(X|\mathcal{G}) + c\mathbb{E}(Y|\mathcal{G}).$$

2. *Monotonicity: if  $X \leq Y$ , then we have*

$$\mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G}).$$

3. *Monotone Convergence: if  $0 \leq X_n \in \mathcal{B}$ ,  $X_n \uparrow X$ , we have*

$$\mathbb{E}(X_n|\mathcal{G}) \uparrow \mathbb{E}(X|\mathcal{G}).$$

4. *Jensen's inequality: if  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  is a convex function, then*

$$\psi(\mathbb{E}(X|\mathcal{G})) \leq \mathbb{E}(\psi(X)|\mathcal{G})$$

5. *Smoothing: if  $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{B}$ , then*

$$\mathbb{E}(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1) = \mathbb{E}(X|\mathcal{F}_1) = \mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2)$$

6. *Known Information: if  $X \in \mathcal{G}$ ,  $XY \in L^1(\mathcal{B})$ , then*

$$\mathbb{E}(XY|\mathcal{G}) = X\mathbb{E}(Y|\mathcal{G})$$

For any random variable  $Z \in \mathcal{B}$ , we can define  $\mathbb{E}(X|Y) := \mathbb{E}(X|\sigma(Y))$ . If we want to condition on the event  $[Y = y]$ , we can define  $\mathbb{E}(X|Y = y) = \mathbb{E}(X|Y)(\omega)$  for any  $\omega \in [Y = y]$ . For any event  $A \in \mathcal{B}$  that is not defined in terms of another random variable, we can define  $\mathbb{E}(X|A) = \mathbb{E}(X|1_A = 1)$ .

## 1.2 Regular Conditional Probabilities

TODO: Standard results 1. existence on "nice" spaces 2. relationship to conditional expectation

## 1.3 Martingales

We start our initial study of martingales by first setting up some underlying concepts. Again, we let  $(\Omega, \mathcal{B}, \mathbb{P})$  be our default probability space.

**Definition 1.3.1.** Let  $T$  be a total ordering and let  $S, \Sigma$  be a measure space. A **stochastic process** is a random element  $X : \Omega, \mathcal{B} \rightarrow S^T, \Sigma^T$ , where  $\Sigma^T$  denotes the  $\Sigma$ -product sigma algebra on  $S^T$ . We call  $S, \Sigma$  the **state space**.

Normally, however, we think of a stochastic process as some kind random object that progresses through "time". We could alternately have defined it as an indexed collection of random elements  $\{X_t : \Omega, \mathcal{B} \rightarrow S, \Sigma\}_{t \in T}$ . The definitions are equivalent due to the following proposition

**Proposition 1.3.2 ((1)).** Let  $(Z, \mathcal{M})$ ,  $(Y_\alpha, \mathcal{N}_\alpha)$  be measure spaces for  $\alpha \in A$ . Let  $Y = \prod_{\alpha \in A} Y_\alpha$  and let  $\mathcal{N} = \bigotimes_{\alpha \in A} \mathcal{N}_\alpha$ . Let  $f : Z \rightarrow Y$  and denote  $f_\alpha := \pi_\alpha \circ f$ , where  $\pi_\alpha$  is the projection map for  $\alpha \in A$ . Then  $f$  is  $(\mathcal{M}, \mathcal{N})$ -measureable iff  $f_\alpha$  is  $(\mathcal{M}, \mathcal{N}_\alpha)$ -measureable for every  $\alpha \in A$ .

Therefore, we will use the two definitions interchangeably.

The total ordering in the definition of stochastic process gives us our concept of "time". Usually,  $T = \mathbb{N}$  or  $\mathbb{R}$ . However, we can imagine more exotic stochastic processes by choosing a total ordering that cannot be order embedded into  $\mathbb{R}$ . For example, we could consider  $\mathbb{R}^2, \preceq$ , endowed with  $\preceq$  is the lexicographic ordering of  $\mathbb{R}^2$ . Observe that

$$\{\{t\} \times \mathbb{R} \subset \mathbb{R}^2 \mid t \in \mathbb{R}\}$$

is a partition of  $\mathbb{R}^2$  into uncountably many nondegenerate  $\preceq$ -intervals. Thus, if  $\mathbb{R}^2, \preceq$  could be order embedded into  $\mathbb{R}$ , then  $\mathbb{R}$  would uncountably many disjoint nondegenerate intervals, each of which must contain a rational. But then the rationals would be uncountable, giving us a contradiction (2).

A stochastic process defined with this more exotic total ordering would look like a grid of random variables, with each vertical slice representing a stochastic process over  $\mathbb{R}$ . We could instead choose to think of it as a stochastic process of stochastic processes indexed by  $\mathbb{R}$ .

A stochastic process on its own does not have much more structure than a random element. We would like to incorporate the idea of measurability with respect to some growing information.

**Definition 1.3.3.** A *filtration* is an non-decreasing collection of sub-sigma algebras  $\{\mathcal{F}_t \subset \mathcal{B}\}_{t \in \mathbb{T}}$ , indexed by our total ordering  $\mathbb{T}$ .

**Definition 1.3.4.** An *adapted process* with respect to filtration  $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$  is a stochastic process  $\{X_t\}_{t \in \mathbb{T}}$  so that  $\forall t \in \mathbb{T}, X_t \in \mathcal{F}_t$ .

Finally, we are ready to give our definition of a martingale, our concept of a fair game.

**Definition 1.3.5.** Let  $X : \Omega \rightarrow \mathbb{R}^{\mathbb{T}}$  be an adapted process wrt filtration  $\mathcal{F}_* = (\mathcal{F}_t)_{t \in \mathbb{T}}$ .  $X$  is a *martingale* if it satisfies the following properties:

1.  $X_t \in L^1(\mathcal{B}) \forall t \in \mathbb{T}$ ,
2.  $\mathbb{E}(X_t | \mathcal{F}_r) = X_r$  ( $\mathbb{P}$ -a.s.) for any  $r \leq t$ .

Note that for condition (2) of definition 1.3.5, we need only check that it holds for  $r < t$ , since it already holds for  $r = t$  by the assumption that  $X$  is an adapted process and by the known information property of proposition 1.1.3.

We can also define the related sub(super)-martingale by changing the equality in condition (2) of definition 1.3.5 to a  $\geq$  ( $\leq$ ). A submartingale represents a favorable game and a super-

martingale represents an unfavorable game. We could have alternatively defined submartingale and supermartingale first, then defined a martingale as an adapted process that is both a submartingale and a supermartingale. We will later prove theorems about submartingales, but they will automatically carry over to martingales. To get the corresponding property for supermartingales, we can just look at the supermartingale's negation to turn it into a submartingale.

Note that by applying *smoothing* and the identity  $\mathbb{E}(Y) = \mathbb{E}(Y|\{\emptyset, \Omega\})$ , we get that for any  $r \leq t \in T$ ,

$$\mathbb{E}(X_t) = \mathbb{E}(X_t|\{\emptyset, \Omega\}) = \mathbb{E}(\mathbb{E}(X_t|\mathcal{F}_r)|\{\emptyset, \Omega\}) = \mathbb{E}(X_r|\{\emptyset, \Omega\}) = \mathbb{E}(X_r)$$

When our total ordering  $T = \mathbb{N}$ , we have the following alternative characterization of a martingale

**Proposition 1.3.6.** *Let  $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$  be an adapted process wrt filtration  $\mathcal{F}_* = (\mathcal{F}_n)_{n \in \mathbb{N}}$ .  $X$  is a martingale if and only if it satisfies the following properties:*

1.  $X_n \in L^1(\mathcal{B}) \ \forall n \in \mathbb{N}$ ,
2.  $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$  ( $\mathbb{P}$ -a.s.) for any  $n \in \mathbb{N}$ .

*Proof.* ( $\Rightarrow$ ) Condition (2) of 1.3.6 is just a special case of condition (2) of 1.3.5 when we set  $t = n + 1 \geq n = r$ .

( $\Leftarrow$ ): This follows by induction. First observe that for  $n = 1$ , the only  $m \in \mathbb{N}$  s.t.  $m \leq n$  is  $m = 1$ . Therefore,  $\forall m \leq n$ , we have that  $\mathbb{E}(X_n|\mathcal{F}_m) = \mathbb{E}(X_1|\mathcal{F}_1) = X_1$ . For the inductive step,

suppose this claim holds for  $n \in \mathbb{N}$ . Consider  $m \leq n + 1$ . If  $m = n + 1$ , then  $\mathbb{E}(X_{n+1}|\mathcal{F}_m) = \mathbb{E}(X_{n+1}|\mathcal{F}_{n+1}) = X_{n+1}$ , so we're good. On the other hand, if  $m < n + 1$ , then we have that

$$\mathbb{E}(X_{n+1}|\mathcal{F}_m) = \mathbb{E}(\mathbb{E}(X_{n+1}|\mathcal{F}_n)|\mathcal{F}_m) = \mathbb{E}(X_n|\mathcal{F}_m) = X_m$$

by the *smoothing* property of proposition 1.1.3 and induction.  $\square$

We refer to such martingales as **discrete-time** martingales.

~~TODO: - definition of stochastic process - remark about total ordering - definition of adapted process and filter - definition of martingale (total ordering), super, sub - connection to natural number definition~~

#### 1.4 Decomposition of Discrete-Time martingales

A useful concept for discrete-time martingales is a predictable process

**Definition 1.4.1.** Let  $X : \Omega \rightarrow S^{\mathbb{N}}$  be an adapted process wrt filtration  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  and state space  $S, \Sigma$ .  $X$  is a **predictable process** if for all  $n \in \mathbb{N}$ ,  $X_n \in \mathcal{F}_{n-1}$ , with  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ .

This concept is capturing the idea that we can "predict" the next value of a stochastic process  $X$  using the information we have now. Indeed, for a real-valued predictable process, if at timestep  $n$  we can discern between whether a given event in  $\mathcal{F}_n$  occurred, then for any  $x \in \mathbb{R}$  we can determine whether or not  $X_{n+1} = x$  since  $[X_{n+1} = x] \in \mathcal{F}_n$ . By setting  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ , we enforce that  $X_1$  is a constant.

The following proposition reveals that we can think of any adapted process indexed by  $\mathbb{N}$  as a predictable process with some "fair" noise.

**Proposition 1.4.2.** *Let  $X : \Omega \rightarrow \mathbb{R}^N$  be a discrete-time adapted process wrt filtration  $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$  so that  $X_n \in L^1(\mathcal{B})$  for all  $n \geq 1$ . Then there exists a martingale  $(M_n)_{n \geq 1}$  and a predictable process  $(A_n)_{n \geq 1}$  with  $A_1 = 0$  so that for  $n \geq 1$*

$$X_n = M_n + A_n.$$

*This decomposition is  $\mathbb{P}$ -a.s. unique.*

Before we proceed with the proof of proposition 1.4.2, we first establish a useful concept and lemma.

**Definition 1.4.3.** *Let  $d : \Omega \rightarrow \mathbb{R}^N$  be a discrete-time process adapted to filtration  $(\mathcal{F}_n)_{n \in \mathbb{N}}$ . Then, it is a **martingale difference sequence** if*

1.  $d_j \in L^1(\mathcal{B}) \ \forall j \in \mathbb{N}$
2.  $\mathbb{E}(d_{j+1} | \mathcal{B}_j) = 0 \ \forall j \in \mathbb{N}$

**Lemma 1.4.4.**  *$(M_n)_{n \geq 1}$  is a martingale adapted to filtration  $\mathcal{F}_* = (\mathcal{F}_n)_{n \geq 1}$  iff there exists a martingale difference sequence  $d : \Omega \rightarrow \mathbb{R}^N$  adapted to the same filtration and  $c \in \mathbb{R}$  so that*

$$M_n := c + \sum_{j=1}^n d_j$$

*for all  $n \in \mathbb{N}$ .*

*Proof.*  $(\Rightarrow)$ : Let  $d_n := M_n - M_{n-1}$  for  $n \geq 1$ , where  $M_0 = \mathbb{E}(M_1)$ . Then for  $n \geq 1$ , the sum

$$M_0 + \sum_{j=1}^n d_j = M_0 + \sum_{j=1}^n M_j - M_{j-1} \text{ is telescoping and is equal to } M_n.$$



Now we need simply show that  $(d_n)_{n \geq 1}$  is a martingale difference sequence. First note that  $(d_n)_{n \geq 1}$  is adapted to  $\mathcal{F}_*$  since  $M_n - M_{n-1} \in \mathcal{F}_n$  for all  $n \geq 1$ . Checking condition (1) of 1.4.3, we see for  $n \geq 1$  that  $\mathbb{E}|d_n| = \mathbb{E}|M_n - M_{n-1}| \leq \mathbb{E}|M_n| + \mathbb{E}|M_{n-1}| < \infty$ , so  $d_n \in L^1(\mathcal{B})$ .

Checking condition (2), we see that

$$\mathbb{E}(d_{n+1}|\mathcal{F}_n) = \mathbb{E}(M_{n+1}|\mathcal{F}_n) - \mathbb{E}(M_n|\mathcal{F}_n) = M_n - M_n = 0.$$

Both conditions are satisfied so  $(d_n)_{n \geq 1}$  is a martingale difference sequence.  $\checkmark$

( $\Leftarrow$ ): We are given difference sequence  $(d_n)_{n \geq 1}$  and constant  $c \in \mathbb{R}$ . Note that since  $(d_n)_{n \geq 1}$  is adapted to  $\mathcal{F}_*$ , we know  $M_n = c + \sum_{j=1}^n d_j \in \mathcal{F}_n$  for all  $n \geq 0$ , so  $(M_n)_{n \geq 1}$  is adapted to  $\mathcal{F}_*$ . Now, checking for  $n \in \mathbb{N}$  that  $M_n \in L^1(\mathcal{B})$ , consider

$$\begin{aligned} \mathbb{E}|M_n| &= \mathbb{E} \left| c + \sum_{j=1}^n d_j \right| \\ &\leq |c| + \sum_{j=1}^n \mathbb{E}|d_j| \\ &< \infty \end{aligned}$$

since  $d_j \in L^1(\mathcal{B}) \ \forall j \in \mathbb{N}$ .

Next we check condition (2) of definition 1.3.5. Letting  $n \in \mathbb{N}$

$$\begin{aligned}
 \mathbb{E}(M_{n+1}|\mathcal{F}_n) &= c + \sum_{j=1}^{n+1} \mathbb{E}(d_j|\mathcal{F}_n) \\
 &= \mathbb{E}(d_{n+1}|\mathcal{F}_n) + c + \sum_{j=1}^n d_j \\
 &= c + \sum_{j=1}^n d_j \quad ((d_n)_{n \geq 1} \text{ is a difference sequence}) \\
 &= M_n.
 \end{aligned}$$

Both conditions are satisfied, so  $(M_n)_{n \geq 1}$  is a martingale. ✓

□

Note that our proof above also gives a procedure to construct a martingale difference sequence from a martingale. Now we proceed to the proof of proposition 1.4.2

*Proof of Proposition 1.4.2.* We will deconstruct  $(X_n)_{n \geq 1}$  into

$$\begin{aligned}
 M_n &:= X_0 + \sum_{j=1}^n X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1}) \\
 A_n &:= \sum_{j=1}^n \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1}
 \end{aligned}$$

where we let  $\mathcal{F}_0 := \{\emptyset, \Omega\}$  and  $X_0 = \mathbb{E}(X_1|\mathcal{F}_0) = \mathbb{E}(X_1)$ . Observe that  $A_1 = \mathbb{E}(X_1|\mathcal{F}_0) - X_0 = \mathbb{E}(X_1) - \mathbb{E}(X_1) = 0$ . With this choice of  $(M_n)_{n \geq 1}$  and  $(A_n)_{n \geq 1}$  we have that

$$\begin{aligned} M_n + A_n &= X_0 + \sum_{j=1}^n X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1}) + \sum_{j=1}^n \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \\ &= X_0 + \sum_{j=1}^n X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1}) + \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \\ &= X_0 + \sum_{j=1}^n X_j - X_{j-1} \\ &= X_n. \end{aligned}$$

Next we show that  $(M_n)_{n \geq 1}$  is a martingale and  $(A_n)_{n \geq 1}$  is a predictable sequence. To that end, observe that for  $n \geq 1$ :

1.  $X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1}) \in \mathcal{F}_n$  since  $(X_n)_{n \geq 1}$  is an adapted process and  $\mathcal{F}_{n-1} \subset \mathcal{F}_n$ .
2. We are given that  $X_n \in L^1(\mathcal{B})$  and  $\mathbb{E}(X_n|\mathcal{F}_{n-1}) \in L^1(\mathcal{B})$  by definition, so  $X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1}) \in L^1(\mathcal{B})$ .
3.  $\mathbb{E}(X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1})|\mathcal{F}_{n-1}) = \mathbb{E}(X_n|\mathcal{F}_{n-1}) - \mathbb{E}(X_n|\mathcal{F}_{n-1}) = 0$

so  $X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1})$  is a martingale difference sequence. By lemma 1.4.4, noting that  $X_0 = \mathbb{E}(X_1)$  is a constant, we have that  $M_n = X_0 + \sum_{j=1}^n X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1})$  is a martingale.

On the other hand, for all  $j \geq 1$ , we know that  $\mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \in \mathcal{F}_j$  since  $(X_n)_{n \geq 1}$  is an adapted process. Because  $\mathcal{F}_*$  is a filtration, for all  $m \geq j$ ,  $\mathcal{F}_j \subset \mathcal{F}_m$ . Therefore, for all  $n \geq 1$ ,  $A_n = \sum_{j=1}^n \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \in \mathcal{F}_n$  and it is predictable.

Finally we show that this decomposition is (almost surely) unique. Suppose there existed another decomposition with the same properties:  $X_n = N_n + B_n$ . We will apply induction to show that these decompositions are almost surely the same. For our base case, let  $n = 1$ . We are given that  $A_1 = B_1 = 0$ . Therefore, after subtracting the two decompositions,  $0 = M_1 + A_1 - N_1 + B_1 = M_1 - N_1$ , so  $M_1 = N_1$ . For our inductive step, assume for some  $n \geq 1$  that  $M_n = N_n$  and  $A_n = B_n$  almost surely. We know that  $0 = M_{n+1} + A_{n+1} - N_{n+1} - B_{n+1}$ . Taking conditional expectations of both sides conditioned on  $\mathcal{F}_n$ , we get

$$\begin{aligned}
0 &= \mathbb{E}(M_{n+1} + A_{n+1} - N_{n+1} - B_{n+1} | \mathcal{F}_n) \\
&= M_n - N_n + \mathbb{E}(A_{n+1} | \mathcal{F}_n) - \mathbb{E}(B_{n+1} | \mathcal{F}_n) && ((M_n)_{n \geq 1}, (N_n)_{n \geq 1} \text{ are martingales}) \\
&= M_n - N_n + A_{n+1} - B_{n+1} && ((A_n)_{n \geq 1}, (B_n)_{n \geq 1} \text{ are predictable}) \\
&= A_{n+1} - B_{n+1} && (\text{inductive hypothesis})
\end{aligned}$$

Therefore  $A_{n+1} = B_{n+1}$  almost surely. Then  $0 = M_{n+1} + A_{n+1} - N_{n+1} - B_{n+1} = M_{n+1} - N_{n+1}$ , so  $M_{n+1} = N_{n+1}$  almost surely, proving the inductive hypothesis for  $n + 1$ . Therefore, by induction, we have that  $M_{n+1} + A_{n+1}$  is the almost surely unique decomposition of  $X_n$  for  $n \geq 1$  with the desired properties.  $\square$

We can interpret proposition 1.4.2 as saying any adapted process is some kind of "trend" with "fair" additive noise applied on top. This result naturally gives us some well-known decomposition results.

**Corollary 1.4.5** (Doob's Submartingale Decomposition). *Let  $(Y_n)_{n \geq 1}$  be a discrete-time adapted process wrt filtration  $\mathcal{F}_* := (\mathcal{F}_n)_{n \geq 1}$  so that  $Y_n \in L^1(\mathcal{B})$  for all  $n \geq 1$ . Then it is a submartingale iff there exists a martingale  $(M_n)_{n \geq 1}$  and a non-decreasing predictable process  $(A_n)_{n \geq 1}$  with  $A_1 = 0$  so that for  $n \geq 1$*

$$Y_n = M_n + A_n$$

*This decomposition is almost surely unique.*

*Proof.* ( $\Rightarrow$ ): Apply proposition 1.4.2 to get our almost surely unique martingale  $(M_n)_{n \geq 1}$  and predictable process  $(A_n)_{n \geq 1}$  so that  $Y_n = M_n + A_n$  for  $n \geq 1$ . We must simply show that  $A_n$  is non-decreasing. To that end, observe that for  $n \geq 2$

$$\begin{aligned}
 0 &\leq \mathbb{E}(Y_n | \mathcal{F}_{n-1}) - Y_{n-1} \\
 &= \mathbb{E}(M_n + A_n | \mathcal{F}_{n-1}) - A_{n-1} \\
 &= \mathbb{E}(M_n | \mathcal{F}_{n-1}) - M_{n-1} + \mathbb{E}(A_n | \mathcal{F}_{n-1}) - A_{n-1} \\
 &= M_{n-1} - M_{n-1} + \mathbb{E}(A_n | \mathcal{F}_{n-1}) - A_{n-1} && ((M_n)_{n \geq 1} \text{ is a martingale}) \\
 &= A_n - A_{n-1}, && ((A_n)_{n \geq 1} \text{ is a predictable process})
 \end{aligned}$$

giving us that  $(A_n)_{n \geq 1}$  is non-decreasing.  $\checkmark$

( $\Leftarrow$ ): We must show that  $Y_n = M_n + A_n$  for  $n \geq 1$  is a submartingale. We only need to check condition (2) since we are given  $Y_n \in L^1(\mathcal{B})$ . Checking, for  $n \geq 1$ , we see

$$\begin{aligned}
 \mathbb{E}(Y_{n+1}|\mathcal{F}_n) &= \mathbb{E}(M_{n+1} + A_{n+1}|\mathcal{F}_n) \\
 &= \mathbb{E}(M_{n+1}|\mathcal{F}_n) + A_{n+1} \\
 &= M_n + A_{n+1} && ((M_n)_{n \geq 1} \text{ is a martingale}) \\
 &\geq M_n + A_n && ((A_n)_{n \geq 1} \text{ is non-decreasing}) \\
 &= Y_n,
 \end{aligned}$$

so  $(Y_n)_{n \geq 1}$  is indeed a submartingale.  $\checkmark$

□

TODO: - ~~predictable process definition~~ - ~~decomp of adapted process into predictable process and martingale~~ - ~~Doob's Decomposition~~

## 1.5 Strategies and Discrete-Time martingales

To further analyze martingales, we introduce the concept of a "strategy" for playing the "fair game", where our strategy determines our move at the end of each turn.

**Definition 1.5.1.** Let  $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$  be a discrete-time adapted process wrt filtration  $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$  and let  $H : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$  be a predictable process on the same filtration. Then we call  $H$  a **strategy** for  $X$  with **winnings**

$$(H \cdot X)_n := \sum_{m=1}^n H_m(X_m - X_{m-1})$$

for  $n \geq 1$  (where  $X_0 = 0$ ).

The following proposition will establish that we cannot change the fairness of a game by choosing the "best" (or "worst") strategy.

**Proposition 1.5.2.** *Let  $(Y_n)_{n \geq 1}$  be a discrete-time martingale wrt filtration  $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$  and let  $(H_n)_{n \geq 1}$  be bounded strategy for  $(Y_n)_{n \geq 1}$ . Then  $((H \cdot Y)_n)_{n \geq 1}$  is a martingale.*

*Proof.* First note that since  $Y_0, Y_1, \dots, Y_n, H_1, \dots, H_n \in \mathcal{F}_n$ , we have that  $(H \cdot Y)_n \in \mathcal{F}_n$ , so  $(H \cdot Y)_n$  is adapted to  $\mathcal{F}_*$ .

Next note that for  $n \geq 1$  because  $H_n$  is bounded, say by  $M \in \mathbb{R}_{\geq 0}$

$$\begin{aligned}
 \mathbb{E}|(H \cdot Y)_n| &= \mathbb{E} \left| \sum_{m=2}^n H_m(Y_m - Y_{m-1}) \right| \\
 &\leq M \sum_{m=1}^n \mathbb{E}|Y_m - Y_{m-1}| && \text{(Triangle Inequality)} \\
 &\leq M \sum_{m=1}^n \mathbb{E}|Y_m| + \mathbb{E}|Y_{m-1}| && \text{(Triangle Inequality)} \\
 &< \infty
 \end{aligned}$$

since  $Y_m \in L^1(\mathcal{B}) \forall m \geq 0$ . So  $(H \cdot Y)_n \in L^1(\mathcal{B})$  as well.  $\checkmark$

Observe that for  $n \geq 1$

$$\begin{aligned}
\mathbb{E}((H \cdot Y)_{n+1} | \mathcal{F}_n) &= \mathbb{E}\left(\sum_{m=1}^{n+1} H_m(Y_m - Y_{m-1}) | \mathcal{F}_n\right) \\
&= \sum_{m=1}^{n+1} \mathbb{E}(H_m(Y_m - Y_{m-1}) | \mathcal{F}_n) \\
&= \sum_{m=1}^n H_m(Y_m - Y_{m-1}) + \mathbb{E}(H_{n+1}(Y_{n+1} - Y_n) | \mathcal{F}_n) \\
&= (H \cdot Y)_n
\end{aligned}$$

since  $(Y_n)_{n \geq 1}$  is a martingale, so

$$\begin{aligned}
\mathbb{E}(H_{n+1}(Y_{n+1} - Y_n) | \mathcal{F}_n) &= H_{n+1}(\mathbb{E}(Y_{n+1} | \mathcal{F}_n) - Y_n) \\
&= 0. \checkmark
\end{aligned}$$

Therefore, we have that  $((H \cdot Y)_n)_{n \geq 1}$  is a martingale. □

Using the Doob decomposition, we can easily generalize this result to submartingales (and therefore supermartingales by negation):

**Corollary 1.5.3.** *Let  $(Y_n)_{n \geq 1}$  be a discrete-time submartingale wrt filtration  $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$  and let  $(H_n)_{n \geq 1}$  be non-negative bounded strategy for  $(Y_n)_{n \geq 1}$ . Then  $((H \cdot Y)_n)_{n \geq 1}$  is a submartingale.*



*Proof.* Using corollary 1.4.5, decompose  $(Y_n)_{n \geq 1}$  into  $(M_n + A_n)_{n \geq 1}$ , where  $(M_n)_{n \geq 1}$  is a martingale and  $(A_n)_{n \geq 1}$  is a predictable process with  $A_1 = 0$ . Denote  $M_0 = A_0 = 0$  to make  $Y_0 = M_0 + A_0 = 0$ . Then,

$$\begin{aligned}
 (H \cdot Y)_n &= \sum_{m=1}^n H_m(Y_m - Y_{m-1}) \\
 &= \sum_{m=1}^n H_m(M_m + A_m - M_{m-1} - A_{m-1}) \\
 &= \sum_{m=1}^n H_m(M_m - M_{m-1}) + \sum_{m=1}^n H_m(A_m - A_{m-1}) \\
 &= (H \cdot M)_n + (H \cdot A)_n
 \end{aligned}$$

We know by proposition 1.5.2 that  $((H \cdot M)_n)_{n \geq 1}$  is a martingale. Furthermore, we have that

1.  $((H \cdot A)_n)_{n \geq 1}$  is a predictable process because  $(H_n)_{n \geq 1}$  and  $(A_n)_{n \geq 1}$  are predictable processes.
2.  $(H \cdot A)_1 = H_1 A_1 = 0$ .
3. For all  $n \geq 1$ ,  $(H \cdot A)_{n+1} = (H \cdot A)_n + H_{n+1}(A_{n+1} - A_n) \geq (H \cdot A)_n$  because  $H_{n+1} \geq 0$  and  $A_{n+1} \geq A_n$  (it is non-decreasing). Thus  $((H \cdot A)_n)_{n \geq 1}$  is a non-decreasing sequence.

Therefore, by the converse case of corollary 1.4.5, we have that  $((H \cdot Y)_n)_{n \geq 1} = ((H \cdot M)_n + (H \cdot A)_n)_{n \geq 1}$  is a submartingale. □

Corresponding results hold for submartingales and martingales. If  $(Y_n)_{n \geq 1}$  is a martingale, then bounded  $(H_n)_{n \geq 1}$  will suffice, it does not need to be non-negative.

We can look at playing the game up to a stopping point as a special kind of strategy. The classical definition of a "stopping time" is as follows

**Definition 1.5.4.** Let  $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$  be a filtration. A stopping time on  $\mathcal{F}_*$  is a random variable  $\nu : \Omega \rightarrow \bar{\mathbb{N}}$  so that for each  $n \in \mathbb{N}$ ,  $[\nu = n] \in \mathcal{F}_n$ . Here,  $\bar{\mathbb{N}}$  is the extended natural numbers.

There are a couple equivalent conditions for being a stopping time.

**Proposition 1.5.5.** Suppose  $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$  is a filtration on  $\Omega$ . Let  $\nu : \Omega \rightarrow \bar{\mathbb{N}}$ . The following conditions are equivalent:

1.  $\nu$  is a stopping time.
2.  $[\nu \leq n] \in \mathcal{F}_n, \forall n \in \mathbb{N}$
3.  $[\nu > n] \in \mathcal{F}_n, \forall n \in \mathbb{N}$

*Proof.* (1)  $\Rightarrow$  (2): We have for  $n \in \mathbb{N}$

$$[\nu \leq n] = \bigcup_{m \leq n} [\nu = m] \in \mathcal{F}_n$$

since  $[\nu = m] \in \mathcal{F}_m \subset \mathcal{F}_n \forall m \leq n$ .  $\checkmark$

(2)  $\Rightarrow$  (3): Let  $n \in \mathbb{N}$ . Since  $[\nu \leq n] \in \mathcal{F}_n$ , we know by definition of a sigma-algebra that  $[\nu > n] = [\nu \leq n]^c \in \mathcal{F}_n$ .  $\checkmark$

(3)  $\Rightarrow$  (1): Let  $n \in \mathbb{N}$ . Since  $[\nu > n] \in \mathcal{F}_n$  and because  $\mathcal{F}_n$  forms a filtration, we have that

$$[\nu = n] = [\nu > n - 1] \cap ([\nu > n])^c \in \mathcal{F}_n,$$

where  $[\nu > 0] = \Omega$ . ✓ □

We can alternatively view a stopping time as a special kind of strategy:

**Proposition 1.5.6.** *Suppose  $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$  is a filtration on  $\Omega$ . Let  $\nu : \Omega \rightarrow \bar{\mathbb{N}}$ . If  $\nu$  is a stopping time, then  $(1_{\nu \geq n})_{n \geq 1}$  is a predictable process.*

*Proof.* We need only check that  $1_{\nu \geq n} \in \mathcal{F}_{n-1}$  for  $n \geq 1$  (with  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ ). To that end we have  $1_{\nu \geq 0} = 1$  since  $\nu(\omega) \in \bar{\mathbb{N}}$  for all  $\omega \in \Omega$ , so  $1_{\nu \geq 0} \in \mathcal{F}_0$ . For  $n \geq 1$ , we have that  $1_{\nu \geq n} = 1 - 1_{\nu \leq n-1} \in \mathcal{F}_{n-1}$  since  $[\nu \leq n-1] \in \mathcal{F}_{n-1}$  by proposition 1.5.5. □

For a adapted process  $X : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$ , we can denote the process resulting from stopping the game after  $\nu$  turns as:

$$X_{n \wedge \nu} := (1_{\nu \geq \cdot} \cdot X)_n = \sum_{m=1}^n 1_{\nu \geq m} (X_m - X_{m-1}),$$

which is precisely how much wealth we would have if

Combining with our result about strategies, we get the following result:

**Corollary 1.5.7.**

TODO: - Definition - preservation of fairness - relationship to stopping times

## 1.6 Convergence of Discrete-time Martingales

TODO: - Upcrossing inequality - convergence of discrete-time martingales - pathological examples

### 1.7 CotnINUOUS-time martingales, definitions

- indistinguishable - section 1.1 Karatzas shreve

### 1.8 Continuous-time martingales, stopping times

-section 1.2 Karatzas shreve

### 1.9 Convergence of Continuous-time Martingales

- section 1.3 karatzas shreve - submartingale inequalities (Karatzas thm 3.8) - convergence  
of right-continuous martingales

### 1.10 Polya's Urn

### 1.11 Borel-Cantelli Lemma

### 1.12 Differential Equation Method?

TODO: - Azuma's inequality - 2nd case study with graphs

## CHAPTER 2

### MARKOV PROCESSES

#### 2.1 Basics

## CHAPTER 3

### BROWNIAN MOTION AND THE HEAT EQUATION

#### 3.1 Basics

- Definition - Existence Remark - Martingale - Markov Process

#### 3.2 Heat Equation

#### 3.3 Brownian Motion as a Solution

## CHAPTER 4

### BRANCHING BROWNIAN MOTION

#### 4.1 Basics

TODO: - definition

#### 4.2 F-KPP Equation

- F-KPP equation

#### 4.3 McKean Representation

- McKean Representation

#### 4.4 Kolmogorov's Result

## CITED LITERATURE

1. Folland, G. B.: Real analysis: Modern techniques and their applications. "John Wiley & Sons", 2011.
2. ([https://math.stackexchange.com/users/39362/haskell curry](https://math.stackexchange.com/users/39362/haskell%20curry)), H. C.: Is the set of real numbers the largest possible totally ordered set? Mathematics Stack Exchange. URL:<https://math.stackexchange.com/q/255152> (version: 2013-02-10).



## VITA

<b>NAME</b>	Sheena The Punk Rocker
<b>EDUCATION</b>	B.A., Composition, Rock n' Roll High School, New York City, NYC, 1976
<b>TEACHING</b>	Music lessons (CS342, Summer 1980)
<b>PUBLICATIONS</b>	<p>Author 1, Author 2, and Author 3. "Paper title 1." In Proceedings of the CONFERENCE NAME (CONFERENCE ABBR, YEAR).</p> <p>Author 1, Author 2, and Author 3. "Paper title 2." In Proceedings of the CONFERENCE NAME (CONFERENCE ABBR, YEAR).</p>