A BOUNDED AND PERIODIC INTEREST RATE MODEL

by

CHEN CAI

PU WANG, COMMITTEE CHAIR
JUNSOO LEE
TSUN-ZEE MAI
TAVAN TRENT
ZHIJIAN WU

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Mathematics
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2013

ABSTRACT

In financial market, interest rate is crucially important. Its changes and moves have a great impact on consumer's products, inflation rate, bond and stock market, and almost all the aspects in the financial world.

An ideal stochastic model describing the volatility of the short-term interest rate would possess the following nice properties. First it has to have the periodic behavior; this is different from stock price model in which it has an increasing or decreasing trends. Second, it should maintain in a positive range and be bounded. Third, its differential equation should be simple and have an analytical solution so that its density function as well as any moments can be readily derived.

In this dissertation, we propose and investigate such a stochastic differential equation. Its solution involves sine/cosine wave functions of Brownian motion that has all these properties. Their statistical properties such as mean, variance and covariance structure of this interest rate at any time are derived; their relation with martingale is established; both analytical and numerical solutions are obtained. From this interest rate model, the term-structure and the yield curves will also be demonstrated for various settings.

DEDICATION

To my parents.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Pu Wang, for giving me the opportunity to conduct research in the exciting and interesting field of financial mathematics. I truly appreciate and am grateful for his guidance and support on every aspect throughout my study. Without his assistance and inspiration, the work in this dissertation would not be possible.

Lastly, I would also like to thank my wonderful parents, for their constant support throughout my life, academically and emotionally.

CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

1.1 Background

In modern finance and economics, interest rates and their dynamics play an important role as a fundamental part of pricing complicated financial derivatives. Their mathematical research always involves with a challenge. Stochastic interest rate is a foundation to innovatively model term structures, yet the classic models are still convincing and effective. The methods of pricing interest rate derivatives have been addressed by many mathematical and financial researchers. Basically there are two groups of interest rate models: equilibrium models and no-arbitrage models. In an equilibrium model, the initial term structure is an output from the model; in a no-arbitrage model it is an input to the model. The disadvantage of the equilibrium models is that they do not automatically fit today's term structure, while the no-arbitrage model is designed to be exactly consistent with today's term structure. (Hull, 2005) [8]

A list of one-factor short rate models include Merton (1973) [11], Rendleman-Bartter (1980) [13], Ho-Lee (1986) [6], Black-Karasinski (1991) [2], Hull-White (Extended Vasicek) (1993) [7], Ait-Sahalia (1996) [1], Mercurio-Moraleda (2000) [10], etc.. The most famous equilibrium models are the one-factor Vasicek model (1977) [16] and the Cox-Ingersoll-Ross (CIR) model (1985) [4]. These two models have been the subject of considerable research, because they have closed form solutions for the term structures. The Vasicek model (1977) [16] was the first model to capture the mean reversion, an essential characteristic that makes interest rate models different from other financial derivatives. However the interest rate determined by the Vasicek model (1977) [16] could fall below zero, which is not practical in the financial world. The CIR model (1985) [4] not only guarantees that the interest rate is non-negative but also keeps the mean-reverting property. Most subsequent equilibrium models were developed closely following these two models. Usually they have the same drift term, but the difference relies in the diffusion term which is much more important in describing the uncertainty of the future evolution of the interest rate. In Vasicek model (1977) [16] , the diffusion term is a positive constant σ, thus not proportional to the interest rate r. In CIR(1985) [4] , the diffusion is $\sigma\sqrt{r}$, and in Duffie-Kan (1996) [5] it is $\sqrt{\sigma_0 + \sigma_1 r}$. In the Black-Karasinski (1991) [2] model, the state variable is identified as $\ln(r)$ to avoid negative interest rates, but its constant diffusion does not make the Wiener process change dependent on r itself. (Copeland *et al*, 2005) [3] The authors Ait-Sahalia (1996) [1] pointed out that the linearity of the drift term $a(b - r)$ is actually the main source of mis-specification and not the second term as people thought before. To overcome the problem of linearity it is possible to add discontinuous jumps to the process, which means unfortunately that closed form solutions of the bond prices become unattainable. (Copeland *et al*, 2005) [3]

In this dissertation, we will present and investigate a new model that preserves the desirable characteristic, mean-reversion. In addition, this model allow the interest rate to move randomly between an upper and a lower boundary.

1.2 Stochastic Process

A good reference for this section is Sheldon M. Ross's *Stochastic Processes* [14].

A *stochastic process* $\mathbf{X} = \{X(t), t \in T\}$ is a collection of random variables. That is, for each t in the *index set* T, $X(t)$ is a random variable. We often interpret t as time and call $X(t)$ the state of the process at time t. If the index set T is a countable set, we call $\mathbf{X}$ a discrete-time stochastic process, and if T is a continuum, we call it a continuous-time process.

A continuous-time stochastic process$\{X(t), t \in T\}$ is said to have independent increments if for all $t_0 < t_1 < t_2 < \cdots < t_n$, the random variables

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$$

are independent. It is said to possess *stationary* increments if $X(t+s) - X(t)$ has the same distribution for all t. That is, it possesses independent increments if the changes in the processes' value over non-overlapping time intervals are independent; it possesses stationary increments if the distribution of the change in value between any two points depends only on the time difference between those points.

1.3 Markov Processes

Some good references for this section are Sheldon M. Ross's *Stochastic Processes.* [14] and Toshio Nakagawa's *Stochastic processes - with applications to reliability theory* [12].

1.3.1 Markov Chains

Markov process is a very general class of processes. Many stochastic processes can be classified as Markov processes. Among them are Markov chains. Consider a discrete-time stochastic process $\{X_n, n = 0, 1, 2, \ldots\}$ with a finite discrete-state set $S = \{0, 1, 2, \cdots, m\}$. Note that the event $\{X_n = j\}$ represents that the process is in State j $(j = 0, 1, 2, \cdots, m)$ at time n $(n = 0, 1, 2, \cdots)$. That is, the process moves among $(m+1)$ states at some unit of time, time unit, or jump according to a given probability law. Then we want to know mainly in which state the process is at time n and converges as n becomes larger.

If we suppose that

$$P\{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0\} = P\{X_{n+1} = j | X_n = i\} \quad (1.3.1)$$

for all $i_0, i_1, \ldots, i_n$ and all $n \geq 0$, then the process $\{X_n, n = 0, 1, 2, \ldots\}$ is said to be a *Markov chain* (Markov 1806). This property indicates that, given the value of X_n, the future value of X_{n+1} does not depend on the value of X_k for $0 \leq k \leq n-1$. If the probability of X_{n+1} being in state j, given that X_n is in state i, is independent of n, i.e.,

$$P\{x_{n+1} = j | X_n = i\} \equiv P_{ij}$$

then the process has a *stationary* or *homogeneous (one-step) transition probability*, i.e., the process transits from state i to state j with probability P_{ij} and remains in state i with probability P_{ij} regardless of n. In other words, the transition probability from state i to state j does not change even if time varies. We restrict ourselves only to discrete-time Markov chains with stationary transition probabilities and a finite state space, because most reliability systems form such processes.

We introduce with the following definition.

Definition 1.3.1. *A matrix $P \equiv \{P_{ij}\}$ is called a stochastic matrix (also termed probability matrix, transition matrix, substitution matrix, or Markov matrix) , if P_{ij} satisfies*

$$P_{ij} \geq 0 \text{ and } \sum_{i=0}^{m} P_{ij} = 1, \text{ for any } i, j$$

Let P denote the transition matrix given by

$$P = \begin{pmatrix} P_{00} & P_{01} & \cdots & P_{0m} \\ P_{10} & P_{11} & \cdots & P_{1m} \\ \vdots & \vdots & & \vdots \\ P_{i0} & P_{i1} & \cdots & P_{im} \\ \vdots & \vdots & & \vdots \\ P_{m0} & P_{m1} & \cdots & P_{mm} \end{pmatrix}$$

We can call P doubly stochastic matrix and the Markov chain doubly stochastic if, in addition,

$$\sum_{j=0}^{m} P_{ij} = 1, \text{ for any } i, j$$

We have already defined the one-step transition probabilities P_{ij}. We now define the n-step transition probabilities P_{ij}^n to be the probability that a process in state i will be in state j after n additional transitions. That is,

$$P_{ij}^n = P\{X_n = j | X_0 = i\}.$$

Also let $P^{(n)}$ be the matrix of n-step transition probabilities P_{ij}^n, then

$$P^{(n+m)} = P^{(n)} \cdot P^{(m)}, \quad n, \ m \geq 0,$$

where the dot represents matrix multiplication.

1.3.2 Limiting Probability

Two states that communicate are said to be in the same *class*; and any two classes are either disjoint or identical. We say that the Markov chain is *irreducible* if there is only one class - That is, if all states communicate with each other.

State j is said to have period d if $P_{ii}^n = 0$ whenever n is not divisible by d and d is the greatest integer with this property. (If $P_{ii}^n = 0$ for all $n > 0$, then define the period of i to be infinite.) A state with period 1 is said to be *aperiodic*.

Let τ_{ii} denote the *return time* to state i given $X_0 = i$:

$$\tau_{ii} = \min\{n \geq 1 : X_n = i | X_0 = i\}.$$

It represents the amount of time (number of steps) until the chain returns to state i given that it started in state i. We define "never return" by $\tau_{ii} = \infty$ and "return occurs" by $\tau_{ii} < \infty$. If $P(\tau_{ii} < \infty) = 1$, state i is called *recurrent*; if $P(\tau_{ii} < \infty) < 1$, it is called *transient*.

A recurrent state j is called *positive recurrent*, if the expected amount of time to return to state j given that the chain started in state j has finite first moment: $E(\tau_{jj}) < \infty$.
A reccurrent state j for which $E(\tau_{jj}) = \infty$ is called *null recurrent*.

Definition 1.3.2. *If for each $j \in S$, π_j exists as defined by $\pi_j = \lim_{n\to\infty} P_{ij}^n$ and is independent of the initial state i, and $\sum_{j \in S} \pi_j = 1$, then the probability distribution $\pi = (\pi_0, \pi_1, \ldots)$ on the state space S is called the limiting or stationary or steady-state distribution of the Markov chain.*

Theorem 1.3.1. *Suppose $\{X_n\}$ is an irreducible Markov chain with transition matrix P. Then $\{X_n\}$ is positive recurrent if and only if there exists a (non-negative, summing to 1)* *solution, $\pi = (\pi_0, \pi_1, \ldots)$, to the set of linear equations $\pi = \pi P$, in which case π is precisely the unique stationary distribution for the Markov chain.*

1.3.3 Brownian Motion

To create a Brownian motion, we first consider a symmetric random walk that in each time unit is equally likely to take a unit step either to the left or to the right. Now suppose that we speed up this process by taking smaller and smaller steps in smaller and smaller time intervals. If we now go to the limit in the correct manner, what we obtain is Brownian motion.

More precisely suppose that each Δt time units we take a step of size Δx either to the left or to the right with equal probabilities. If we let $X(t)$ denote the position at time t, then

$$X(t) = \Delta x(X_1 + \cdots + X_{[t/\Delta t]}), \tag{1.3.2}$$

where

$$X_i = \begin{cases} +1 & \text{if the } i\text{th step of length } \Delta x \text{ is to the right,} \\ -1 & \text{if it is to the left,} \end{cases}$$

and where the X_i is assumed independent with

$$P\{X_i = 1\} = P\{X_i = -1\} = \frac{1}{2}.$$

Since $E[X_i] = 0, \mathrm{Var}(X_i) = E[X_i^2] = 1$, we see from (1.3.2) that

$$E[X(t)] = 0, \tag{1.3.3}$$

$$\mathrm{Var}(X(t)) = (\Delta x)^2 \left[\frac{t}{\Delta t}\right]. \tag{1.3.4}$$

We shall now let Δx and Δt go to 0. However, we must do it in a way to keep the resulting limiting process nontrivial (for instance, if we let $\Delta x = \Delta t$ and then let $\Delta t \to 0$, then from the above we see that $E[X(t)]$ and $\text{Var}(X(t))$ would both converge to 0 and thus $X(t)$ would equal to 0 with probability 1). If we let $\Delta x = c\sqrt{\Delta t}$ for some positive constant c, then from (1.3.3) and (1.3.4), we see that as $\Delta t \to 0$

$$E[X(t)] = 0,$$
$$\text{Var}(X(t)) \to c^2 t.$$

We now list some intuitive properties of this limiting process obtained by taking $\Delta x = c\sqrt{\Delta t}$ and then letting $\Delta t \to 0$. From (1.3.2) and the central limit theorem we see that:

(i) $X(t)$ is normal with mean 0 and variance $c^2 t$;

(ii) $\{X(t), t \geq 0\}$ has independent increments;

(iii) $\{X(t), t \geq 0\}$ has stationary increments.

Finally, as the distribution of the change in position of the random walk over any time interval depends only on the length of that interval, it would appear that:

Definition 1.3.3. *A stochastic process* $[W_t, t \geq 0]$ *is said to be a Brownian motion process if:*

(i) $W_0 = 0$*;*

(ii) $\{W_t, t \geq 0\}$ *has stationary independent increments;*

(iii) for every $t > 0, W_t$ *is normally distributed with mean 0 and variance* $c^2 t$*.*

When $c = 1$, the Brownian motion is called *standard Brownian motion*. Throughout this dissertation, the Brownian motion we will always use standard Brownian motion. Using the independent increments assumption, we can determine the covariance function of the standard Brownian motion as follows. Let $0 \leq s < t$,

$$\begin{aligned}
\mathrm{Cov}\big[W_s, W_t\big] =& \mathbb{E}\big[W_s W_t\big] - \mathbb{E}\big[W_s\big]\mathbb{E}\big[W_t\big] \\
=& \mathbb{E}\big[W_s(W_t - W_s + W_s)\big] - 0 \\
=& \mathbb{E}\big[W_s\big]\mathbb{E}\big[W_t - W_s\big] + \mathbb{E}\big[W_s\big]^2 \\
=& 0 + \mathbb{E}\big[W_s\big]^2 \\
=& \mathrm{Var}\big[W_s\big] \\
=& s
\end{aligned}$$

In other words, for any values of s and t.

$$\mathrm{Cov}\big[W_s, W_t\big] = \min(s, t) \tag{1.3.5}$$

1.4 Itô Integrals

A good reference for this section is Steven E. Shreve's *Stochastic Calculus for Finance II Continuous-Time Models* [15].

We first construct the integrand $\delta(t)$, $t \geq 0$, where

(i) $\delta(t)$ is adapted to filtration $\mathcal{F}(t)$, $t \geq 0$ (i.e., δ is $\mathcal{F}(t)-$measurable);

(ii) $\delta(t)$ is square-integrable:

$$\mathbb{E}\int_0^T \delta^2(t)dt < \infty, \text{ for all } T. \tag{1.4.1}$$

Then we want to define the Itô Integral

$$I(t) = \int_0^t \delta(u)dW_u,$$

where W_u is one-dimensional Brownian motion with associated filtration $\mathcal{F}(t)$ and it is $\mathcal{F}(t)-$measurable.

Definition 1.4.1. *The Itô integral for the continuously varying integrand $\delta(t)$ is defined by the formula*

$$\int_0^t \delta(u)dW_u = \lim_{n\to\infty} \int_0^t \delta_n(u)dW_u, \quad 0 \le t \le T, \tag{1.4.2}$$

where $\{\delta_n\}$ is a sequence of simple processes such that

$$\lim_{n\to\infty} \mathbb{E} \int_0^T (\delta(t) - \delta_n(t))^2 dt = 0$$

Theorem 1.4.1. *Let T be a positive constant and let $\delta(t)$, $0 \le t \le T$, be an adapted stochastic process that satisfies (1.4.1). Then $I(t) = \int_0^t \delta(u)dW_u$ defined by (1.4.2) has the following properties.*

(i) ***(Continuity)****As a function of the upper limit of integration t, the paths of $I(t)$ are continuous.*

(ii) ***(Adaptivity)****For each t, $I(t)$ is $\mathcal{F}(t)-$measurable.*

(iii) ***(Linearity)****If $I(t) = \int_0^t \delta(u)dW_u$ and $J(t) = \int_0^t \phi(t)dW_u$, then $I(t) \pm J(t) = \int_0^t \big(\delta(t) \pm \phi(t)\big)dW_u$; furthermore, for every constant c, $cI(t) = \int_0^t c\delta(t)dW_u$.*

(iv) ***(Martingale)****$I(t)$ is a martingale.*

(v) ***(Itô isometry)*** *$\mathbb{E}I^2(t) = \mathbb{E}\int_0^t \delta^2 du$.*

(vi) ***(Quadratic variation)*** $[I, I](t) = \mathbb{E}\int_0^t \delta^2 du$.

1.5 Itô-Doeblin formula

A good reference for this section is Steven E. Shreve's *Stochastic Calculus for Finance II Continuous-Time Models* [15].

Theorem 1.5.1. *Let $f(t,x)$ be a function for which the partial derivatives $f_t(t,x)$, $f_x(t,x)$, and $f_{xx}(t,x)$ are defined and continuous, and let W_t be a Brownian motion. Then for every $T \geq 0$,*

$$f\big(T, W_T\big) = f\big(0, W_0\big) + \int_0^T f_t\big(t, W_t\big)dt + \int_0^T f_x\big(t, W_t\big)dW_t + \frac{1}{2}\int_0^T f_{xx}\big(t, W_t\big)dt. \quad (1.5.1)$$

The formula in (1.5.1) can be written in differential form:

$$df\big(t, W_t\big) = f_t\big(t, W_t\big)dt + f_x\big(t, W_t\big)dW(t) + \frac{1}{2}f_{xx}\big(t, W_t\big)dt. \quad (1.5.2)$$

1.6 Risk-Neutral Measure

Some good references for this section are Steven E. Shreve's *Stochastic Calculus for Finance II Continuous-Time Models* [15] and Damien Lamberton and Bernard Lapeyre's *Introduction to Stochastic Calculus Applied to Finance* [9].

1.6.1 Basic Theorem

Theorem 1.6.1. ***(Girsanov's Theorem)***. *Let W_t, $0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{F}(t)$, $0 \leq t \leq T$, be a filtration for this Brownian* *morion. Let $\Theta(t)$, $0 \le t \le T$, be an adapted process. Define*

$$Z(t) = \exp\left\{ -\int_0^t \Theta(u)dW_u - \frac{1}{2}\int_0^t \Theta^2(u)du \right\},$$

$$\widetilde{W}_t = W_t + \int_0^t \Theta(u)du,$$

and assume that

$$\mathbb{E}\int_0^T \Theta^2(u)Z^2(u)du < \infty.$$

Set $Z = Z(T)$. Then $\mathbb{E}Z = 1$ and under the probability measure $\widetilde{\mathbb{P}}$ given by $\widetilde{\mathbb{P}} = \int_A Z(\omega)dP(\omega)$ for all $A \in \mathcal{F}$, the process $\widetilde{W}_t$, $0 \le t \le T$, is a Brownian motion with

$$\widetilde{\mathbb{E}}\big(\widetilde{W}_t\big) = 0, \tag{1.6.1}$$

$$\widetilde{Var}\big(\widetilde{W}_t\big) = t. \tag{1.6.2}$$

1.6.2 Asset Pricing under Risk-Neutral Measure

Definition 1.6.1. *A market is viable if there is no arbitrage opportunity.*

The following result is sometimes referred to as the *Fundamental Theorem of Asset Pricing.*

Theorem 1.6.2. *The market is viable if and ony if there exists a probability measure $\widetilde{\mathbb{P}}$ equivalent* [1] *to $\mathbb{P}$ such that the discounted prices of assets are $\widetilde{\mathbb{P}}-$martingales.*

A risk-neutral measure $\widetilde{\mathbb{P}}$, also called an equivalent martingale measure, is used in the pricing of financial derivatives due to the fundamental theorem of asset pricing, which implies a complete market, in which an asset's price is the expected value of the discounted future

[1] Recall the two probability measure $\mathbb{P}_1$ and $\mathbb{P}_2$ are equivalent if and only if for any event A, $\mathbb{P}_1(A) = 0 \Leftrightarrow \mathbb{P}_2(A) = 0$. Here, $\widetilde{\mathbb{P}}$ equivalent to $\mathbb{P}$ means that, for any $\omega \in \Omega$, $\widetilde{\mathbb{P}_1}(\{\omega\}) > 0$.

payoff under the unique risk-neutral measure $\widetilde{\mathbb{P}}$. We denote the price process of an asset at time t by $V(t)$, $0 \leq t \leq T$. In addition, suppose we have an adapted interest rate process $r(t)$. We define the discount process

$$D(t) = e^{-\int_0^t r(s)ds}. \tag{1.6.3}$$

By Theorem (1.6.2), we have the assumption

$$D(t)V(t) = \widetilde{\mathbb{E}}\Big(D(T)V(T)|\mathcal{F}(t)\Big), \quad 0 \leq t \leq T. \tag{1.6.4}$$

Dividing (1.6.4) by $D(t)$, which is $\mathcal{F}$-measurable and hence can be moved inside the conditional expectation on the right-hand side of (1.6.4), we may write (1.6.4) as

$$V(t) = \widetilde{\mathbb{E}}\Big(e^{-\int_t^T r(s)ds}V(T)|\mathcal{F}(t)\Big), \quad 0 \leq t \leq T. \tag{1.6.5}$$

Relied on Girsanov's theorem, under the risk-neutral probability $\mathbb{P}$, there exists a process $\Theta(u)$, which is called *market price of risk*, such that the process $\widetilde{W_t} = W_t + \int_0^t \Theta(u)du$ is a Brownian motion. Risk-neutral measurement guarantee that the prices of assets do not admit arbitrage and that every derivative security in the model can be hedged.

Chapter 2

Bounded and Periodic functions of Brownian Motion

In the real financial world, It seems that interest rate does not behave like stock prices and randomly moves between a relatively narrow bounded range. We seek an interest rate model that possesses this property. One such candidate function would be sine or cosine function. The short-term risk-free interest rate $r(t)$ (for simplicity, sometimes denoted as r) proposed in this dissertation is described by a process of the form

$$dr = \beta\big(\alpha - r\big)dt + \boldsymbol{\sigma}\sqrt{1-\mathcal{L}^2(r)}\; dW_t. \tag{2.0.1}$$

Here α and β are positive constants, $\mathcal{L}$ is a nonrandom function of r and $\boldsymbol{\sigma}$ is a nonrandom function of a standard Brownian motion W_t. This model is said to have one-factor because the interest rate is determined by only one stochastic differential equation (sde). The interest rate modeled by equation (2.0.1) possesses the property of mean-reversion due to its drift term. The diffusion term implies that

$$-1 \leq \mathcal{L}(r) \leq 1.$$

Assuming the inverse function of $\mathcal{L}$ exists, denoted as $\mathcal{L}^{-1}$, there must be a upper and lower bounds for r. Equation (2.0.1) has the properties we expect for the interest rate. One analytic solutions to the equation (2.0.1) possibly involves with sine/cosine function. Under the assumption of $r = \mathcal{L}^{-1}\Big(\cos(\theta W_t + \gamma)\Big)$, where θ, γ are constants, and a specific function $\boldsymbol{\sigma}$, equation (2.0.1) can be written as the following

$$dr = \beta\big(\alpha - r\big)dt + \sigma\sin(\theta W_t + \gamma)dW_t. \tag{2.0.2}$$

Before we could better understand the model (2.0.1) and (2.0.2), we focus on some interesting properties of $\cos(\theta W_t)$ and $\sin(\theta W_t)$ in this chapter. Then in Chapter 3, we will further discuss this interest rate model and solve the *sde* (2.0.2).

2.1 Statistical Properties of $\cos W_t$ and $\sin W_t$

We first compute the mean and variance of $\cos(\theta W_t)$, where W_t is a standard Brownian motion. To do this, theorem (2.1.1) will be helpful.

Theorem 2.1.1. *For any normally distributed random variable X with mean zero and variance $\sigma^2 < \infty$,*

$$\mathbb{E}\big[\cos(\theta X)\big] = e^{-\frac{\theta^2\sigma^2}{2}}$$

and

$$Var\big[\cos(\theta X)\big] = \frac{1}{2}\big(1 - e^{-\theta^2\sigma^2}\big)^2.$$

Proof. We will use moment generating function (*mgf*) to prove this theorem. The taylor series for $\cos(\theta X)$ is

$$\cos(\theta X) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}(\theta X)^{2n}.$$

Then,

$$\begin{aligned}\mathbb{E}\big[\cos(\theta X)\big] &= \mathbb{E}\Big[\sum_{n=0}^{\infty} \frac{(-1)^n}{2n!}(\theta X)^{2n}\Big] \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!}\mathbb{E}\big[(\theta X)^{2n}\big] \\ &= \sum_{n=0}^{\infty} \frac{(-\theta^2)^n}{2n!}\mathbb{E}\big[(X)^{2n}\big] \qquad (2.1.1)\end{aligned}$$

X is normally distributed and the *mgf* is $M(s) = e^{\frac{1}{2}s^2\sigma^2}$. The *Taylor series* of the $e^{\frac{1}{2}s^2\sigma^2}$ is

$$M(s) = \sum_{m=0}^{\infty} \frac{s^{2m}\sigma^2 m}{2^m m!} = 1 + \frac{s^2\sigma^2}{2} + \frac{s^4\sigma^4}{4 \cdot 2!} + \frac{s^6\sigma^6}{8 \cdot 3!} + \cdots\cdots,$$

then,

$$M'(s) = \sum_{m=0}^{\infty} \frac{2m \cdot s^{2m-1}\sigma^{2m}}{2^m m!} = \sum_{m=1}^{\infty} \frac{2m \cdot s^{2m-1}\sigma^{2m}}{2^m m!}$$

$$M''(s) = \sum_{m=1}^{\infty} \frac{2m \cdot (2m-1) \cdot s^{2m-2}\sigma^{2m}}{2^m m!}$$

$$\vdots$$

and $M'(0) = 0$, since the only constant term "1" in $M(s)$ is gone after taking the first derivative, and all terms in $M'(s)$ contains s. If we differentiate the second time, we get an only constant term again and $M''(0)$ will equal to this constant term. So it is easy to think in this way: if p is a positive integer and $M^{(p)}(s)$ is the pth derivative of $M^{(s)}$, the degrees of s in each term of $M^{(p)}(s)$ only have two possible patterns: "1, 3, 5, 7, 9..." or "0, 2, 4, 6, 8...". Only in pattern "0, 2, 4, 6, 8..." the derivative function of $M(t)$ has a constant term because of the degree "0" on s and this is the pattern when p is an even number. $M^{(p)}(0)$ equals to the constant term. For p is odd, we have all the terms containing s, then $M^{(p)}(0)$ equals to zero. Now we can write down all the moments,

$$M''(0) = \frac{2 \cdot \sigma^2}{2 \cdot 1},$$

$$M^{(3)}(0) = 0,$$

$$\vdots$$

$$M^{(2n-1)}(0) = 0,$$

$$M^{(2n)}(0) = \frac{2n! \cdot \sigma^{2n}}{2^n \cdot n!}.$$

Therefore,

$$\mathbb{E}\big[W_t^{2n}\big] = M^{(2n)}(0) = \frac{2n! \cdot \sigma^{2n}}{2^n \cdot n!}, \tag{2.1.2}$$

and we write (2.1.1) as

$$\begin{aligned}\mathbb{E}\big[\cos(\theta W_t)\big] &= \sum_{n=0}^{\infty} \frac{(-\theta^2)^n}{2n!} \cdot \frac{2n!\sigma^{2n}}{2^n n!} \\ &= \sum_{n=0}^{\infty} \frac{(-\theta\sigma)^2 n}{2^n n!} \\ &= e^{-\frac{\theta^2\sigma^2}{2}}. \end{aligned} \tag{2.1.3}$$

To compute the variance of $\cos(\theta X)$, we use trigonometrical double-angle identity to calculate

$$\mathbb{E}\big[\cos^2(\theta X)\big] = \mathbb{E}\left[\frac{1+\cos(2\theta X)}{2}\right] = \frac{1+e^{-2\theta^2\sigma^2}}{2},$$

then the variance is

$$\begin{aligned}\mathrm{Var}\big[\cos(\theta X)\big] =&\mathbb{E}\big[\cos^2(\theta X)\big] - \mathbb{E}^2\big[\cos(\theta X)\big] \\ =&\frac{1+e^{-2\theta^2\sigma^2}}{2} - e^{-\theta^2\sigma^2} \\ =&\frac{1}{2}\big(1-e^{-\theta^2\sigma^2}\big)^2. \end{aligned}$$

□

Theorem (2.1.1) is used to compute $\mathbb{E}\big[\cos(\theta W_t)\big]$ and $\mathrm{Var}\big[\cos(\theta W_t)\big]$, where W_t is a standard Brownian motion, following distribution $N(0,t)$, then

$$\mathbb{E}\big[\cos(\theta W_t)\big] = e^{-\frac{\theta^2 t}{2}}. \tag{2.1.4}$$

and

$$\mathrm{Var}\big[\cos(\theta W_t)\big] = \frac{1}{2}\big(1 - e^{-\theta^2 t}\big)^2.$$

Now let us discuss about the mean and variance of $\sin(\theta W_t)$. The expected value of $\sin(\theta W_t)$, with respect to the normal probability density function (*pdf*) of W_t, $\phi(x)$, is given by $\mathbb{E}\big(\sin(\theta W_t)\big) = \int_{\infty}^{\infty} \phi(x)\sin(\theta x)dx$. Since $\phi(x)$ is an even function and $\sin(\theta x)$ is an odd one, this Lebesgue integral over symmetric interval is zero. That is

$$\mathbb{E}\big[\sin(\theta W_t)\big] = 0, \tag{2.1.5}$$

and

$$\mathrm{Var}\big[\sin(\theta W_t)\big] = \mathbb{E}\big[\sin^2(\theta W_t)\big] = \mathbb{E}\left[\frac{1-\cos(2\theta W_t)}{2}\right] = \frac{1}{2}\big(1 - e^{-2\theta^2 t}\big).$$

2.2 Probability Distribution of $\cos W_t$

The *probability density function* (*pdf*) of $\cos W_t$ is another quantity of interest. Let $\arccos z = A$, $A \in [0, \pi]$ due to $z \in [-1, 1]$. We start with the *cumulative distribution function* (*cdf*) of $\cos W_t$ which can be calculated as

$$\begin{aligned}
&P(\cos W_t \le z) \\
=&2\sum_{n=0}^{\infty} P\big(2n\pi + A \le W_t \le 2(n+1)\pi - A\big) \\
=&2\sum_{n=0}^{\infty} \Big[P\big(W_t \le 2(n+1)\pi - A\big) - P\big(W_t \le 2n\pi + A\big)\Big] \\
=&2\sum_{n=0}^{\infty} \Big[\int_{-\infty}^{2(n+1)\pi - A} \frac{\exp(-\frac{x^2}{2t})}{\sqrt{2\pi t}}dx - \int_{-\infty}^{2n\pi + A} \frac{\exp(-\frac{x^2}{2t})}{\sqrt{2\pi t}}dx\Big] \\
=&\frac{\sqrt{2}}{\sqrt{\pi t}}\sum_{n=0}^{\infty} \Big[\int_{0}^{2(n+1)\pi - A} e^{-\frac{x^2}{2t}}dx - \int_{0}^{2n\pi + A} e^{-\frac{x^2}{2t}}dx\Big]
\end{aligned}$$

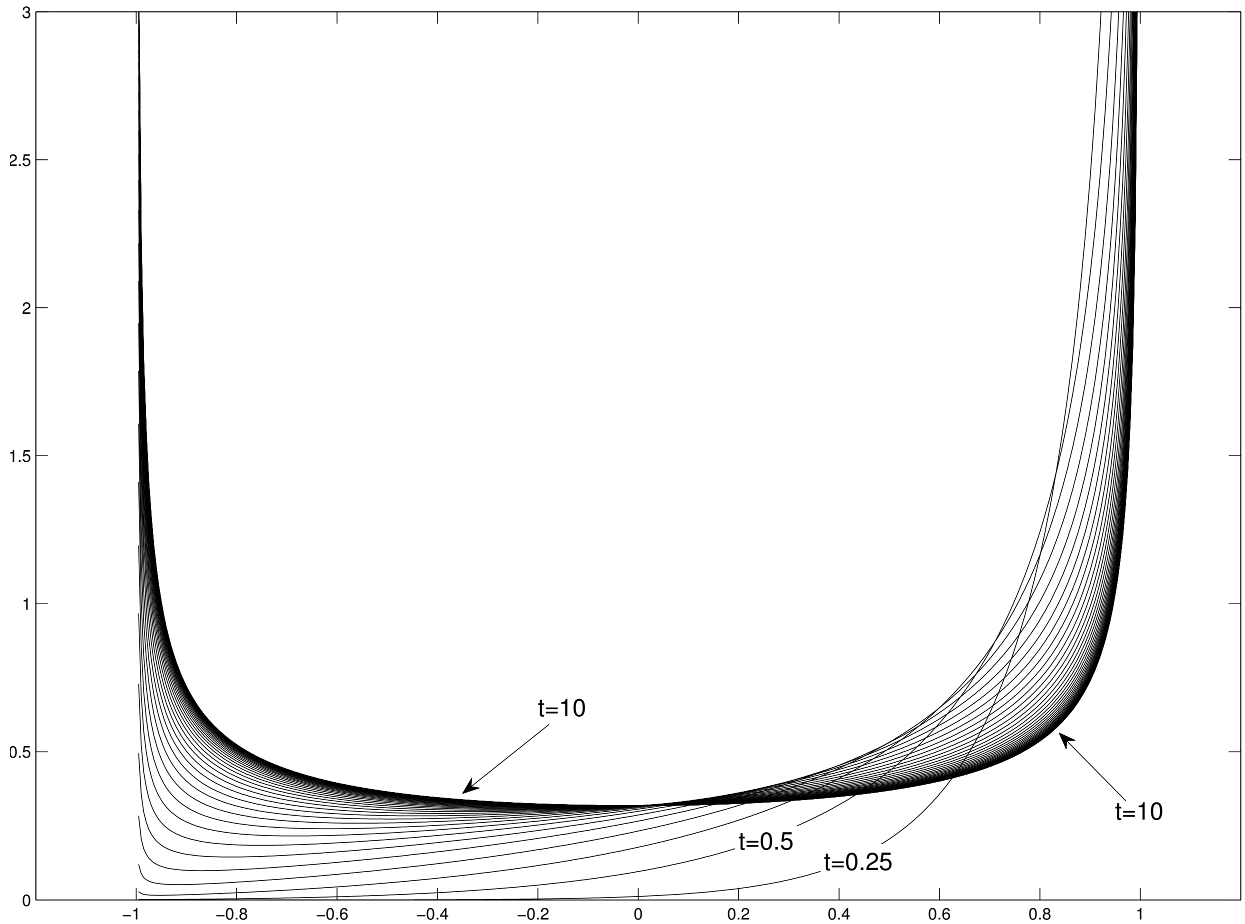


Figure 2.2.1: *Pdf* $f_t(z)$ converges with t increasing

Differentiating this *cdf* function, we get the *pdf* of $\cos W_t$:

$$f_t(z) = \frac{\sqrt{2}}{\sqrt{\pi t}} \frac{1}{\sqrt{1-z^2}} \sum_{n=0}^{\infty} \left[e^{-\frac{(2n\pi + \arccos z)^2}{2t}} + e^{-\frac{(2(n+1)\pi - \arccos z)^2}{2t}} \right]. \tag{2.2.1}$$

The *pdf*s of $\cos W_t$ (2.2.1) at different ts are shown in Figure (2.2.1). As t increases gradually from 0.25 to 10 with increment of 0.25, there is a clear trend for (2.2.1) to converge. Although we have an exact *pdf* of $\cos W_t$, where W_t is the standard Brownian motion, computing its limiting *pdf* denoted by

$$f_\infty(z) = \lim_{t \to \infty} f_t(z)$$

is a much difficult task (if it is not impossible).

We will next derive the limiting *pdf* $f_\infty(z)$. Recalling derivation in section(1.3.3), a Brownian motion can be obtained by taking the limit of random walk as both time increment Δt and distance increment Δx approach zero. In the same manner, we take a detour to finite Markov chain first and then come back to compute the limiting *pdf* of $\cos W_t$.

We know that in $\cos W_t$, W_t can be represented as an angle. On a unit circle, the angle's measurement in radians numerically equal the length of its corresponding arc. We think in this way, instead of moving vertically up or down on a x-y coordinates, our Brownian motion will move along a unit circle, clockwise or counterclockwise, with the initial position W_0 on the positive x-axis (see Figure 2.2.2). For convenience, we define another process W'_t

$$P(\cos W'_t < z) = P(\cos W_t < z), \quad W'_t \in [0, 2\pi), \quad z \in [-1, 1], \tag{2.2.2}$$

where W'_t moves with the same probabilities on the unit circle as a Brownian motion W_t, which makes $\cos W'_t$ has the same probability distribution as $\cos W_t$. W'_t will 'wrap around' upon reaching 2π or 0. That's to say W'_t only takes the value from 0 to 2π. By how we define in equation (2.2.2), we see further that

$$P(\cos W'_\infty < z) = P(\cos W_\infty < z), \quad W'_\infty \in [0, 2\pi), \quad z \in [-1, 1], \tag{2.2.3}$$

where $P(\cos W'_\infty < z)$ is the *cdf* of $\cos W'_t$ as t approaches infinity (limiting *cdf*), so is $P(\cos W_\infty < z)$.

After we define the process for W'_t, we discretize this continuous case into discrete counterparts. Let us first consider a symmetric random walk that moves along a set of m nodes, labelled 0, 1, ..., $m-1$, which are arranged around a circle(see Figure 2.2.3). At each step and each state it is equally likely to move a unit step in either the clockwise or

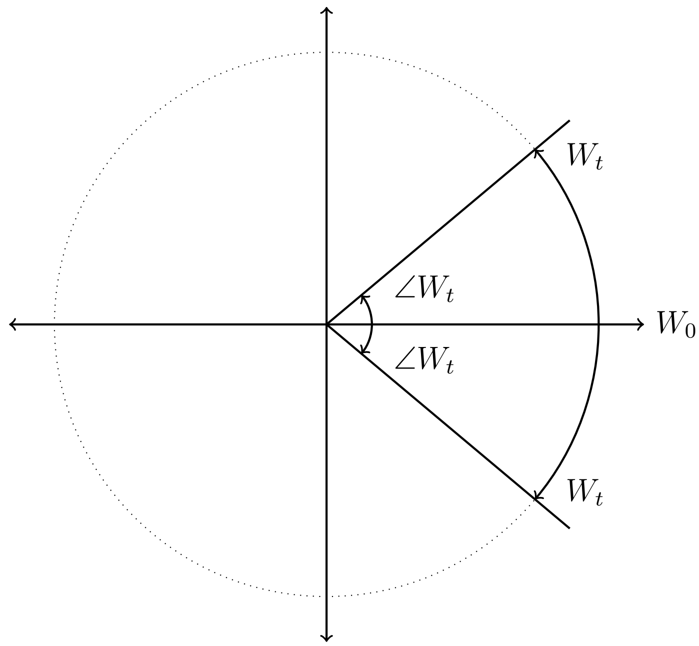


Figure 2.2.2: Brownian motion moving around a unit circle.

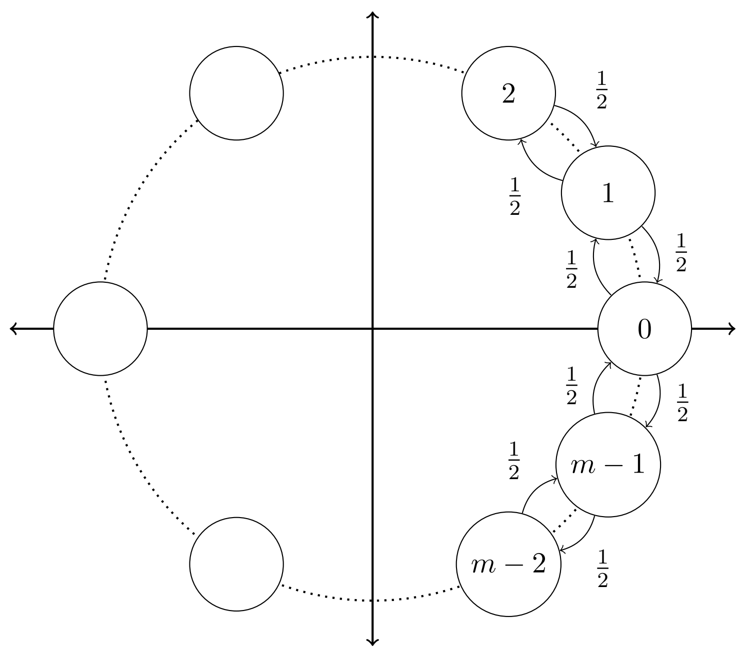


Figure 2.2.3: Particle moving around a circle.

counterclockwise direction. That is, if X_n is the position of the particle after its n-th step then

$$X_n = \begin{cases} i+1 & \text{if the } n\text{-th step is counterclockwise, given that } X_{n-1} = i \\ i-1 & \text{if it is clockwise,} \end{cases}$$

and the probability distribution of each independent step is

$$P\{X_n = i+1 | X_{n-1} = i\} = P\{X_n = i-1 | X_{n-1} = i\} = \frac{1}{2},$$

where $i + 1 \equiv 0$ when $i = m - 1$, and $i - 1 \equiv m - 1$ when $i = 0$. Then the transition probability matrix of this finite Markov chain is

$$P = \begin{pmatrix} 0 & \frac{1}{2} & & & & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & & & \\ & \frac{1}{2} & 0 & \frac{1}{2} & & \\ & & \ddots & \ddots & \ddots & \\ & & & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & & & & \frac{1}{2} & 0 \end{pmatrix} \tag{2.2.4}$$

Lemma 2.2.1. *If a transition matrix for an irreducible Markov chain with a finite state space S is doubly stochastic, its (unique) limiting distribution is uniform over S.*

Proof. Assume that $S = 0, 1, 2, \ldots, m-1$. A vector $(1, 1, \ldots, 1)$ with m entries of 1 is exactly the vector of column sums. This can also be written as $(1, 1, \ldots, 1)P = (1, 1, \ldots, 1)$.

This equation holds if multiplied by a number $\frac{1}{m}$ on both sides, which yields

$$\left(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}\right) P = \left(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}\right).$$

This specifies that $\boldsymbol{\pi} = \left(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}\right)$ is the unique solution to $\boldsymbol{\pi} P = \boldsymbol{\pi}$ with $||\boldsymbol{\pi}|| = 1$. The limiting distribution π is uniform over S and the lemma is proven. □

In equation (2.2.4), each column of the transition matrix P only have two non-zero entries of $\frac{1}{2}$, so each column sum is 1. From Definition (1.3.1) and lemma (2.2.1), we can say that the transition matrix of this symmetric random walk on a circle is a doubly stochastic and let $\boldsymbol{\pi}$ be its limiting distribution, which is

$$\boldsymbol{\pi} = (\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}).$$

In order to obtain $P(\cos W'_\infty < z)$, we next make the finite Markov chain discussed above continuous by obtaining its limit. We consider to find limit by letting the number of states in S, $m+1$, go to infinity. It is supposed that the states in S are evenly distributed on the unit circle. The arc between any two consecutive states is Δx and $m = \frac{2\pi}{\Delta x}$. When $m \to \infty$, $\Delta x \to 0$. Hence, the limiting *cdf* of W'_t is

$$P\big(W'_\infty < \theta\big) = \lim_{\Delta x \to 0} \sum_{\theta} \frac{\Delta x}{2\pi} = \int_0^\theta \frac{1}{2\pi} dx = \frac{\theta}{2\pi}, \quad 0 < \theta \leq 2\pi.$$

The notation "Σ" in the second side of the equality above represents the summation of all the states on the arc from 0 to θ. This is a *cdf* of a uniform distribution from 0 to 2π. From equation (2.2.3), we obtain

$$\begin{aligned} P\Big(\cos W_\infty < z\Big) =& P\Big(\cos W'_\infty < z\Big) \\ =& P\Big(\cos^{-1} z < W'_\infty < 2\pi - \cos^{-1} z\Big) \\ =& \frac{\pi - \cos^{-1} z}{\pi}. \end{aligned}$$

We conclude this discussion in the next theorem.

Theorem 2.2.1. *Let W_t be a standard Brownian motion, the limiting pdf of $cosW_t$, noted as $f_\infty(z)$, is*

$$f_\infty(z) = \frac{1}{\pi\sqrt{1-z^2}}, \quad z \in (-1, 1). \tag{2.2.5}$$

Proof. By differentiation of $P\Big(\cos W_\infty < z\Big) = \frac{\pi - \cos^{-1} z}{\pi}$ with respect to z, the theorem is proven. □

Figure (2.2.4) shows the comparison between a series of $f_t(z)$ at some different t's and the limiting *pdf* $f_\infty(z)$. $f_t(z)$ converges fast to $f_\infty(z)$. $f_{10}(z)$ and $f_\infty(z)$ are very closed to each other. The distribution of $\cos W_\infty$ is symmetric about zero and its mean

$$\mathbb{E}(\cos W_\infty) = \lim_{t\to\infty} e^{-\frac{t}{2}} = 0.$$

2.3 More Statistical Properties

In section, we study the martingale property and derive the covariance structure of the process $\cos W_t$.

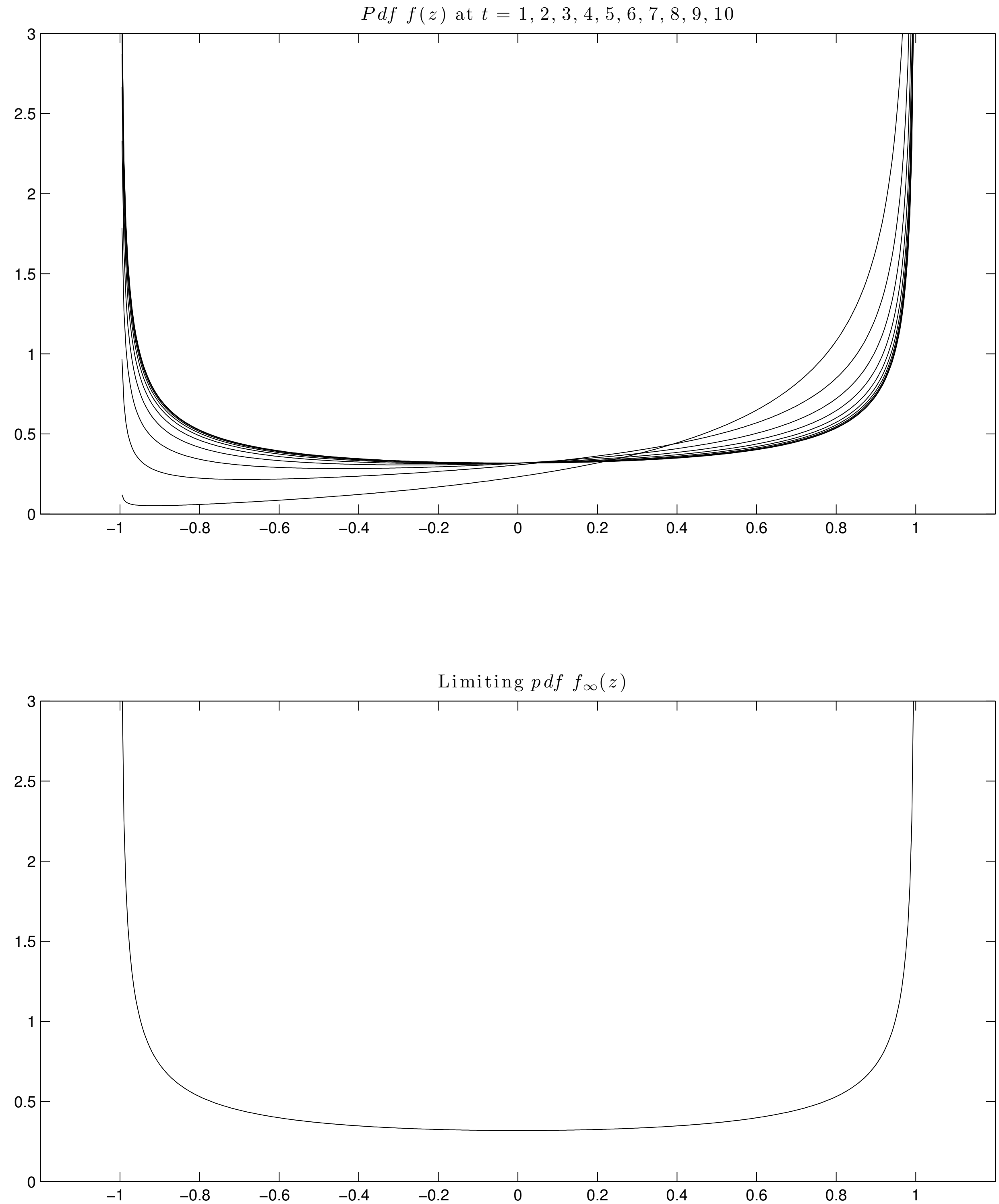


Figure 2.2.4: *Pdf* Comparison between equation (2.2.1) and (2.2.5)

2.3.1 Independent Time Increments

Let $0 < s < t$ be given, and the random variables W_t and $W_t - W_s$ are independent increments. We decompose $\mathbb{E}\big[\cos W_t\big]$ as

$$\begin{aligned}
\mathbb{E}\big[\cos W_t\big] &= \mathbb{E}\big[\cos(W_t - W_s + W_s)\big] \\
&= \mathbb{E}\big[\cos W_s \cos(W_t - W_s) - \sin W_s \sin(W_t - W_s)\big] \\
&= \mathbb{E}\big[\cos W_s \cos(W_t - W_s)\big] - \mathbb{E}\big[\sin W_s\big]\mathbb{E}\big[\sin(W_t - W_s)\big] \\
&= \mathbb{E}\big[\cos W_s \cos(W_t - W_s)\big] \qquad (2.3.1) \\
&= \mathbb{E}\big[\cos W_s\big]\mathbb{E}\big[\cos(W_t - W_s)\big]. \qquad (2.3.2)
\end{aligned}$$

One can derive (2.3.1) and (2.3.2) by using angle-sum identity in the second line and the result $\mathbb{E}\big[\sin W_t\big] = 0$ from equation (2.1.5) in the third line. This result is stated more generally in Theorem (2.3.1)

Theorem 2.3.1. *Let W_t be a standard Brownian motion. For $0 < t_0 < t_1 < t_2 < \cdots < t_n$,*

$$\mathbb{E}\big[\cos W_{t_n}\big] = \mathbb{E}\left[\cos W_{t_0} \prod_{i=1}^{n} \cos W_{t_i - t_{i-1}}\right],$$

or

$$\mathbb{E}\big[\cos W_{t_n}\big] = \mathbb{E}\big[\cos W_{t_0}\big] \prod_{i=1}^{n} \mathbb{E}\big[\cos W_{t_i - t_{i-1}}\big].$$

Proof. By (2.3.2) we have

$$\mathbb{E}\big[\cos W_{t_n}\big] = \mathbb{E}\big[\cos W_{t_{n-1}}\big]\mathbb{E}\big[\cos(W_{t_n} - W_{t_{n-1}})\big].$$

If we repeatedly decompose t_{n-1}, t_{n-3} $\cdots$ and t_1, like the above process, the result is going to be

$$\mathbb{E}\big[\cos W_{t_n}\big] = \mathbb{E}\big[\cos W_{t_0}\big]\prod_{i=1}^{n}\mathbb{E}\big[\cos(W_{t_i} - W_{t_{i-1}})\big].$$

All these consecutive non-overlapping increment are mutually independent. Since the increment $W_{t_i} - W_{t_{i-1}}$ has the same distribution as $W_{t_i - t_{i-1}}$, it is equivalent to write $\mathbb{E}\big[\cos(W_{t_i} - W_{t_{i-1}})\big]$ as $\mathbb{E}\big[W_{t_i - t_{i-1}}\big]$. Then the theorem is proven.

$\square$

2.3.2 Covariance of $\cos W_t$

For $0 < s < t$, by using theorem (2.3.1) the covariance of $\cos W_t$ and $\cos W_s$ can be calculated as

$$\begin{aligned}
\mathrm{Cov}\big[\cos W_t, \cos W_s\big] &= \mathbb{E}\big[\cos W_t \cdot \cos W_s\big] - \mathbb{E}\big[\cos W_t\big]\mathbb{E}\big[\cos W_s\big] \\
&= \mathbb{E}\big[\cos(W_t - W_s)\big] \cdot \mathbb{E}\big[\cos^2 W_s\big] - e^{-\frac{t+s}{2}} \\
&= e^{-\frac{t+s}{2}}\left(\frac{1+e^{-2s}}{2}\right) - e^{-\frac{t+s}{2}} \\
&= \frac{1}{2}e^{-\frac{t+3s}{2}}(e^{2s} + 1 - 2e^s) \\
&= \frac{1}{2}e^{-\frac{t+3s}{2}}(e^s - 1)^2.
\end{aligned}$$

The movement at time t and s is always positively correlated. That is to say $\cos W_s$ and $\cos W_t$ have the same trend to move up or down. This is similar to a standard Brownian motion W_t.

2.3.3 Martingale Property

If we compute the conditional expectation $\mathbb{E}\big[\cos W_t \big| \mathcal{F}(s)\big]$, by theorem (2.3.1) and for $s < t$,

$$\mathbb{E}\big[\cos W_t \big| \mathcal{F}(s)\big] = \mathbb{E}\big[\cos W_s \big| \mathcal{F}(s)\big]\mathbb{E}\big[\cos W_{t-s} \big| \mathcal{F}(s)\big] = \cos W_s \cdot \mathbb{E}\big[\cos W_{t-s}\big] = e^{-\frac{t-s}{2}} \cos W_s.$$

The processes of $\cos W_t$ and $\sin W_t$ themselves are not martingales. However the result above helps us find martingales.

Theorem 2.3.2. *Let $W(t)$, $t \geq 0$, be a Brownian motion with a filtration $\mathcal{F}(t)$, $t \geq 0$. Then,*

(i) the process $e^{\frac{t}{2}} \cos W_t$ is a martingale;

(ii) the process $e^{\frac{t}{2}} \sin W_t$ is also a martingale.

Proof. For $0 \leq s \leq t$, by using theorem (2.3.1) we have

$$\begin{aligned}
\mathbb{E}\big[e^{\frac{t}{2}} \cos W_t \big| \mathcal{F}_s\big] &= e^{\frac{t}{2}} \mathbb{E}\big[\cos(W_t - W_s) \cos W_s \big| \mathcal{F}(s)\big] \\
&= e^{\frac{t}{2}} \cdot e^{-\frac{t-s}{2}} \cdot \cos W_s \\
&= e^{\frac{s}{2}} \cos W_s.
\end{aligned} \tag{2.3.3}$$

In a similar manner we can show that $e^{\frac{t}{2}} \sin W_t$ is also a martingale:

$$\begin{aligned}
\mathbb{E}\big[e^{\frac{t}{2}} \sin W_t \big| \mathcal{F}(s)\big] &= e^{\frac{t}{2}} \mathbb{E}\big[\cos(W_t - W_s) \sin W_s + \sin(W_t - W_s) \cos W_s \big| \mathcal{F}(s)\big] \\
&= e^{\frac{t}{2}} \mathbb{E}\big[\cos(W_t - W_s) \sin W_s \big| \mathcal{F}(s)\big] + \mathbb{E}\big[\sin(W_t - W_s) \cos W_s \big| \mathcal{F}(s)\big] \\
&= e^{\frac{t}{2}} \cdot e^{-\frac{t-s}{2}} \cdot \sin W_s + 0 \\
&= e^{\frac{s}{2}} \sin W_s.
\end{aligned}$$

□

Chapter 3

The Interest Rate Model

3.1 Introduction to One-factor Interest Rate Model

Some good references for this section are John C. Hull's *Options, Futures, and Other Derivatives.* [8] and Steven E. Shreve's *Stochastic Calculus for Finance II Continuous-Time Models* [15].

Equilibrium models usually start with assumptions about economic variables and derive a process for the short-term risk-free rate, r. They then explore what the process implies for bond prices and option prices. The short rate, r, at time t is the rate that applies to an infinitesimally short period of time at time t. It is sometimes referred to as the instantaneous short rate. It is important to emphasize that it is not the process for r in the real world that matters. Bond prices, option prices, and other derivative prices depend only on the process followed by r in a risk-neutral world.

In a one-factor model, the process for r involves only one source of uncertainty. Usually the short rate is described by an Itô process of the form

$$dr = m(r)dt + s(r)dW_t.$$

The instantaneous drift, m, and instantaneous standard deviation (diffusion), s, are assumed to be functions of r, but independent of time. The assumption of a single factor is not as restrictive as it might appear. A one-factor model implies that all rates move in the same direction over any short time interval, but not that they all move by the same amount. It does not, as is sometimes supposed, imply that the term structure always has the same shape. A fairly rich pattern of term structures can occur under a one factor model.

Here are several classic equilibrium models:

$$m(r) = \mu r, \quad s(r) = \sigma r \quad \text{(Rendleman and Bartter model);} \tag{3.1.1}$$

$$m(r) = \beta(\alpha - r), \quad s(r) = \sigma \quad \text{(Vasicek model);} \tag{3.1.2}$$

$$m(r) = \beta(\alpha - r), \quad s(r) = \sigma\sqrt{r} \quad \text{(Cox, Ingersoll, and Ross model).} \tag{3.1.3}$$

Mean Reversion

Rendleman and Bartter's assumption that the short-term interest rate behaves like a stock price is less than ideal. One important difference between interest rates and stock prices is that interest rates appear to be pulled back to some long-run average level over time. This phenomenon, known as *mean reversion*, is not captured by the Rendleman-Bartter model. When r is high, mean reversion tends to cause it to have a negative drift; when r is low, mean reversion tends to cause it to have a positive drift. Mean reversion is illustrated in Figure (3.1.1).

The Vasicek model is the first interest rate model that introduced the desirable mean-reverting property. The mathematical explanation of this property is: when $r(t) = \alpha$, the drift term in (3.1.2) is zero; when $r(t) > \alpha$, this term is negative, which pushes $r(t)$ back toward α; when $r(t) < \alpha$, this term is positive, which again pushes $r(t)$ back toward α. If $r(0) = \alpha$, then $\mathbb{E}\big(r(t)\big) = \alpha$ for all $t \geq 0$. If $r(0) \neq \alpha$, then $\lim_{t\to\infty} \mathbb{E}\big(r(t)\big) = \alpha$.

The Cox, Ingersoll, and Ross model (CIR model) has the same drift term as Vasicek model, so it captures the mean-reverting property as Vasicek model.

There are compelling economic arguments in favor of mean reversion. When rates are high, the economy tends to slow down and there is less requirement for funds on the part of borrowers. As a result, rates decline. When rates are low, there tends to be a high demand for funds on the part of borrowers. As a result, rates tend to rise.

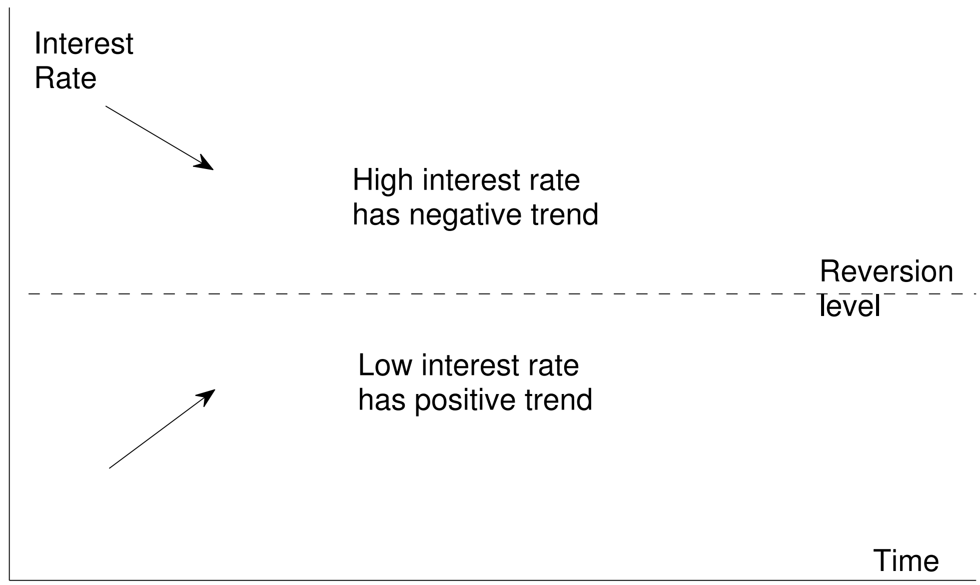


Figure 3.1.1: Mean reversion.

Non-negativity

If interest rate is negative, rather than paying us interest on deposits, the bank would charge us for storing cash. This sounds impossible, however, interest rates are normally positive, but not always. In real financial world, negative interest rate is more complicated involving government policies. For convenience and simplicity, we consider non-negativity a theoretical assumption and ideal property for the mathematical model.

The advantage of CIR over the Vasicek model is that the interest rate in the CIR model does not become negative. In CIR, If $r(t)$ reaches zero, the diffusion term $\sigma\sqrt{r}$ multiplying dW_t vanishes and the positive drift term $\alpha\beta dt$ in (3.1.3) drives the interest rate back into positive territory. The Vasicek model can not guarantee that the interest rate is always non-negative. However, we would like to point out that none of these models has an upper bound.

3.2 Basic Assumptions

Let W_t, $t \geq 0$, be a standard Brownian motion. The model we propose to solve is in the form of

$$dr = \beta\big(\alpha - r\big)dt + \boldsymbol{\sigma}(W_t)\sqrt{1 - \mathcal{L}^2(r)}\; dW_t, \tag{3.2.1}$$

where α and β are positive constants, $\mathcal{L}$ is a given non-random function and $\boldsymbol{\sigma}(W_t)$ is a given function of a standard Brownian motion W_t. Like the Vasicek equation (3.1.2) and the CIR equation (3.1.3), (3.2.1) has the same drift term $\beta\big(\alpha - r\big)$, which preserves mean-reverting property for this model. The significant difference comes from the diffusion term, the diffusion in (3.2.1) has a square-root which is similar to the CIR. The square-root makes $0 \leq r$ in the CIR. In the same way, the square-root in (3.2.1) gives a restriction on r, that is

$$-1 \leq \mathcal{L}(r) \leq 1.$$

For an existing $\mathcal{L}^{-1}$, there must exist a lower bound l and an upper bound u such that

$$l \leq r \leq u.$$

A simple case is that $\mathbb{L}$ is a linear function. Then l and u are two constant determined by the constant coefficients from $\mathbb{L}$. By choosing the coefficient of the linear function $\mathbb{L}$, the interest rate r is controlled to move randomly between a lower level l and a upper lever u. The l can be chosen as zero to guarantee the interest rate non-negative. The u can prevent the interest rate from reaching an extremely high level. The mean-reverting in neither Vasicek nor CIR model can eliminate the possibility that the interest rate might go suddenly very high. Not like stock prices, we expect that interest rate moves within a relative narrow range and our proposed model (3.2.1) guarantees this desirable property.

In order to solve this stochastic differential equation for a possible solution, we first assume that

$$r = \mathcal{L}^{-1}\big(\cos(\theta W_t + \gamma)\big). \tag{3.2.2}$$

Here θ and γ are positive constants. Then based on (3.2.2), we may assume function $\boldsymbol{\sigma}(W_t) = \sigma \cdot \text{sgn}\Big(\sin(\theta W_t + \gamma)\Big)$, where σ is a positive constant and

$$\text{sgn}\Big(\sin(\theta W_t + \gamma)\Big) = \begin{cases} +1 & \text{if } \sin(\theta W_t + \gamma) \geq 0, \\ -1 & \text{if } \sin(\theta W_t + \gamma) < 0. \end{cases}$$

With the above assumptions, *sde* (3.2.1) can be simplified to

$$dr = \beta(\alpha - r)dt + \sigma \sin(\theta W_t + \gamma)dW_t. \tag{3.2.3}$$

3.3 Solving the Stochastic Differential Equation

To solve *sde* (3.3.7), we first establish a function

$$f(t, r) = e^{\beta t} r.$$

For the Itô-Doeblin formula $df = f_t dt + f_r dr + \frac{1}{2} f_{rr}(dr)^2$, we shall need the following partial derivatives of $f(t, r)$:

$$f_t = \beta e^{\beta t} r, \quad f_r = e^{\beta t} \quad \text{and} \quad f_{rr} = 0.$$

Since $f_{rr} = 0$, we shall not need $(dr)^2$. The Itô-Doeblin formula states that

$$\begin{aligned}
&df(t, r) \\
=&d(e^{\beta t} r) \\
=&f_t dt + f_r dr + \frac{1}{2} f_{rr}(dr)^2 \\
=&\beta e^{\beta t} r dt + \beta e^{\beta t}(\alpha - r)dt + \sigma e^{\beta t} \sin(\theta W_t + \gamma)dW_t \\
=&\alpha\beta e^{\beta t} dt + \sigma e^{\beta t} \sin(\theta W_t + \gamma)dW_t.
\end{aligned} \tag{3.3.1}$$

Integration of both sides from 0 to t and using the initial condition $r(0) = r_0$, we obtain the formula

$$\begin{aligned}
&f(t, r(t)) - f(0, r(0)) \\
=&e^{\beta t} r(t) - r_0 \\
=&\alpha\beta \int_0^t e^{\beta s} ds + \sigma \int_0^t e^{\beta s} \sin(\theta W_s + \gamma)dW_s \\
=&\alpha(e^{\beta t} - 1) + \sigma \int_0^t e^{\beta s} \sin(\theta W_s + \gamma)dW_s.
\end{aligned}$$

Then,

$$r(t) = e^{-\beta t} r_0 + \alpha(1 - e^{-\beta t}) + \sigma e^{-\beta t} \int_0^t e^{\beta s} \sin(\theta W_s + \gamma)dW_s. \tag{3.3.2}$$

In order to solve the random integral $\int_0^t e^{\beta s}\sin(\theta W_s+\gamma)dW_s$ in the third term of (3.3.2), we first use Itô-doeblin formula to compute

$$\begin{aligned}
&de^{\beta s}\cos(\theta W_s+\gamma)\\
=&-\theta\sin(\theta W_s+\gamma)dW_s+\beta e^{\beta s}\cos(\theta W_s+\gamma)ds-\frac{\theta^2}{2}e^{\beta s}\cos(\theta W_s+\gamma)(dW_s)^2\\
=&-\theta e^{\beta s}\sin(\theta W_s+\gamma)dW_s+(\beta-\frac{\theta^2}{2})e^{\beta s}\cos(\theta W_s+\gamma)ds.
\end{aligned}$$

We can see that if $\beta=\frac{\theta^2}{2}$, the second term on the right-hand side of the above result disappears, which leads to

$$de^{\frac{\theta^2 s}{2}}\cos(\theta W_s+\gamma)=-\theta e^{\frac{\theta^2 s}{2}}\sin(\theta W_s+\gamma)dW_s,$$

or

$$e^{\frac{\theta^2 s}{2}}\sin(\theta W_s+\gamma)dW_s=-\frac{1}{\theta}de^{\frac{\theta^2 s}{2}}\cos(\theta W_s+\gamma). \tag{3.3.3}$$

Integration of both sides of (3.3.3) and the condition $W_0=0$ result in

$$\begin{aligned}
\int_0^t e^{\frac{\theta^2 s}{2}}\sin(\theta W_s+\gamma)dW_s=&-\frac{1}{\theta}\int_0^t de^{\frac{\theta^2 s}{2}}\cos(\theta W_s+\gamma)\\
=&\frac{1}{\theta}\left(1-e^{\frac{\theta^2 t}{2}}\cos(\theta W_t+\gamma)\right).
\end{aligned} \tag{3.3.4}$$

A more generic integration from t to T is

$$\int_t^T e^{\frac{\theta^2 s}{2}}\sin(\theta W_s+\gamma)dW_s=\frac{1}{\theta}\left(e^{\frac{\theta^2 t}{2}}\cos(\theta W_t+\gamma)-e^{\frac{\theta^2 T}{2}}\cos(\theta W_T+\gamma)\right). \tag{3.3.5}$$

Substitute the (3.3.4) for the third term in (3.3.2), we obtain

$$r(t) = e^{-\frac{\theta^2 t}{2}} r_0 + \alpha(1 - e^{-\frac{\theta^2 t}{2}}) + \frac{\sigma}{\theta}\Big(e^{-\frac{\theta^2 t}{2}} - \cos(\theta W_t + \gamma)\Big) \tag{3.3.6}$$

which is the solution to

$$dr = \frac{\theta^2}{2}\big(\alpha - r\big)dt + \boldsymbol{\sigma}\sqrt{1 - \mathcal{L}^2(r)}\; dW_t, \tag{3.3.7}$$

Recalling in Chapter 2 that

$$\mathbb{E}\big[\cos(\theta W_t)\big] = e^{-\frac{\theta^2 t}{2}},$$

$$\mathrm{Var}\big[\cos(\theta W_t)\big] = \frac{1}{2}\big(1 - e^{-\theta^2 t}\big)^2,$$

and

$$\mathrm{Var}\big[\sin(\theta W_t)\big] = \frac{1}{2}\big(1 - e^{-2\theta^2 t}\big),$$

we calculate the expectation and variance of $\cos(\theta W_t + \gamma)$ by using angle-sum identity as followings

$$\begin{aligned}\mathbb{E}\big[\cos(\theta W_t + \gamma)\big] =& \mathbb{E}\big[\cos(\theta W_t)\cos\gamma - \sin(\theta W_t)\sin\gamma\big]\\ =& \cos\gamma\mathbb{E}\big[\cos(\theta W_t)\big]\\ =& e^{-\frac{\theta^2 t}{2}}\cos\gamma,\end{aligned} \tag{3.3.8}$$

$$\begin{aligned}\mathrm{Var}\big[\cos(\theta W_t + \gamma)\big] =& \mathrm{Var}\big[\cos(\theta W_t)\cos\gamma - \sin(\theta W_t)\sin\gamma\big]\\ =& \cos^2\gamma\mathrm{Var}\big[\cos(\theta W_t)\big] + \sin^2\gamma\mathrm{Var}\big[\sin(\theta W_t)\big]\\ =& \frac{1}{2}\Big(\cos^2\gamma\big(1 - e^{-\theta^2 t}\big)^2 + \sin^2\gamma\big(1 - e^{-2\theta^2 t}\big)\Big).\end{aligned} \tag{3.3.9}$$

Then the expectation and variance of $r(t)$ are calculated as

$$\mathbb{E}\big[r(t)\big] = e^{-\frac{\theta^2 t}{2}} r_0 + \alpha(1 - e^{-\frac{\theta^2 t}{2}}) + \frac{\sigma}{\theta} e^{-\frac{\theta^2 t}{2}} \left(1 - \cos\gamma\right), \tag{3.3.10}$$

and

$$\mathrm{Var}\big[r(t)\big] = \frac{\sigma^2}{\theta^2}\left(\cos^2\gamma\big(1 - e^{-\theta^2 t}\big)^2 + \sin^2\gamma\big(1 - e^{-2\theta^2 t}\big)\right). \tag{3.3.11}$$

In particular, as t goes to infinity, all the exponential term in $r(t)$ (3.3.6), $\mathbb{E}\big[r(t)\big]$ (3.3.10) and $\mathrm{Var}\big[r(t)\big]$ (3.3.11) vanish and

$$r(t) = \alpha - \frac{\sigma}{\theta}\cos(\theta W_t + \gamma), \tag{3.3.12}$$

with

$$\mathbb{E}\big[r(t)\big] = \alpha, \quad \text{and} \quad \mathrm{Var}\big[r(t)\big] = \frac{\sigma^2}{2\theta^2}.$$

We discussed in last section that a lower bound l exists for $r(t)$ and we want to know when l is zero. By using equation (3.3.12), we have

$$r(t) = \alpha - \frac{\sigma}{\theta}\cos(\theta W_t + \gamma) \geqslant l,$$

Because $\cos(\theta W_t + \gamma) \in [-1, 1]$, and we want $\alpha - \frac{\sigma}{\theta}\cos(\theta W_t + \gamma) \geqslant l$ be always true, then the following must be satisfied:

$$\alpha - \frac{\sigma}{\theta} = l. \tag{3.3.13}$$

When $l = 0$, we have

$$\alpha - \frac{\sigma}{\theta} = 0.$$

On the contrary, for the upper bound u, we obtain another inequality

$$\alpha - \frac{\sigma}{\theta}\cos(\theta W_t + \gamma) \leqslant u,$$

and it is always true if

$$\alpha + \frac{\sigma}{\theta} = u. \tag{3.3.14}$$

The equation (3.3.12) shows that $r(t)$ is a linear function of $\cos(\theta W_t + \gamma)$, or equivalently $\cos(\theta W_t + \gamma)$ is a linear function of $r(t)$. In the assumption (3.2.2)

$$r = \mathcal{L}^{-1}\big(\cos(\theta W_t + \gamma)\big),$$

$\mathcal{L}$ or $\mathcal{L}^{-1}$ is a linear function. We rewrite the differential equation (3.3.7)

$$dr = \frac{\theta^2}{2}\big(\alpha - r\big)dt + \boldsymbol{\sigma}\sqrt{1 - \mathcal{L}^2(r)}\ dW_t$$

in the form of

$$dr = \frac{\theta^2}{2}\big(a - r\big)dt + \theta\sqrt{(u - r)(r - l)}\ dW_t.$$

In which u and l are identified as boundaries of interest rate r due to $(u - r)(r - l) \geq 0$. We summarize the discussion in this section by presenting the following theorem.

Theorem 3.3.1. *Suppose that the stochastic differential equation governing the short term interest rate is given by*

$$dr(t) = \frac{\theta^2}{2}\big(a - r(t)\big)dt + \theta\sqrt{\big(u - r(t)\big)\big(r(t) - l\big)}\; dW_t \tag{3.3.15}$$

where the constant parameters a, θ, u and l satisfies $a = \frac{u+l}{2}$, $0 < \theta$ and $l < u$. The initial condition is $r(0) = r_0$, $r_0 \in [l, u]$. The solution to this sde is in the form of

$$r(t) = a + \frac{u - l}{2}\cos(\theta W_t + \gamma), \tag{3.3.16}$$

where

$$\gamma = \arccos\big(\frac{2(r_0 - a)}{u - l}\big).$$

Proof. The equation (3.3.16) is equivalent as

$$\cos(\theta W_t + \gamma) = \frac{2}{u - l}(r - a).$$

Using Itô-Doeblin formula, we have

$$dr = (r_t + \frac{1}{2} r_{WW})dt + r_W dW_t$$

where

$$r_t = 0, \quad r_W = -\frac{\theta(u - l)}{2}\sin(\theta W_t + \gamma), \quad r_{WW} = -\frac{\theta^2(u - l)}{2}\cos(\theta W_t + \gamma).$$

The diffusion term is

$$
\begin{aligned}
r_W dW_t = &- \frac{\theta(u-l)}{2}\sin(\theta W_t+\gamma)dW_t \\
= &\pm \frac{(u-l)}{2}\theta\sqrt{\sin^2(\theta W_t+\gamma)}dW_t \\
= &\pm \frac{(u-l)}{2}\theta\sqrt{1-\cos^2(\theta W_t+\gamma)}dW_t \\
= &\pm \frac{(u-l)}{2}\theta\sqrt{1-\left(\frac{2}{u-l}(r-a)\right)^2}dW_t \\
= &\pm \theta\sqrt{\left(\frac{u-l}{2}\right)^2-(r-a)^2dW_t} \\
= &\pm \theta\sqrt{(u-r)(r-l)}\ dW_t.
\end{aligned}
$$

Based on the definition of Itô process, W_t and dW_t are independent, which indicates that the process $\pm\theta\sqrt{(u-r)(r-l)}$ is independent on dW_t. Additionally, because $d(-W_t)$ and dW_t have the same distribution, the diffusion term can be written as

$$
r_W dW_t = \theta\sqrt{(u-r)(r-l)}\ dW_t.
$$

The drift term is

$$
\begin{aligned}
(r_t+\frac{1}{2}r_{WW})dt = &-\frac{\theta^2}{2}\frac{u-l}{2}\cos(\theta W_t+\gamma)dt \\
= &\frac{\theta^2}{2}\left(a-a-\frac{u-l}{2}\cos(\theta W_t+\gamma)\right)dt \\
= &\frac{\theta^2}{2}\big(a-r\big)dt.
\end{aligned}
$$

Next we check the initial condition:

$$
r_0 = a+\frac{u-l}{2}\cos(\theta W_t+\gamma)
$$

which implies

$$\gamma = \arccos\big(\frac{2(r_0 - a)}{u - l}\big).$$

□

For simplicity, let $b = \frac{u-l}{2}$, then the solution in theorem (3.3.1) is

$$r(t) = a + b\cos(\theta W_t + \gamma)$$

where $\gamma = \arccos\big(\frac{r_0 - a}{b}\big)$. The mean-reverting level is a, upper bound $u = a + b$ and lower bound $l = a - b$.

3.4 Monte Carlo Simulation

This section presents the simulation of the interest rate process. Monte Carlo method is a collection of mathematical and statistical techniques to estimate information from a stochastic model, where we face a high degree of uncertainty. With the process of obtaining random number, estimating deterministic parameters, deciding number of trials, generating discretized paths and calculating the expectation, Monte Carlo has led to wide spread applications to a variety of problems where closed form solutions are either impossible to attain or extremely hard to find. This method not only tells us something about the end result of a random system, but also presents the behavior along the time. We are interested in what the interest rate determined by theorem (3.3.1) looks like, so in this section we use the Monte Carlo method to generate some paths of the interest rate model which is studied in this dissertation.

The interest rate equation modeled by theorem (3.3.1) is in the discrete-time form of

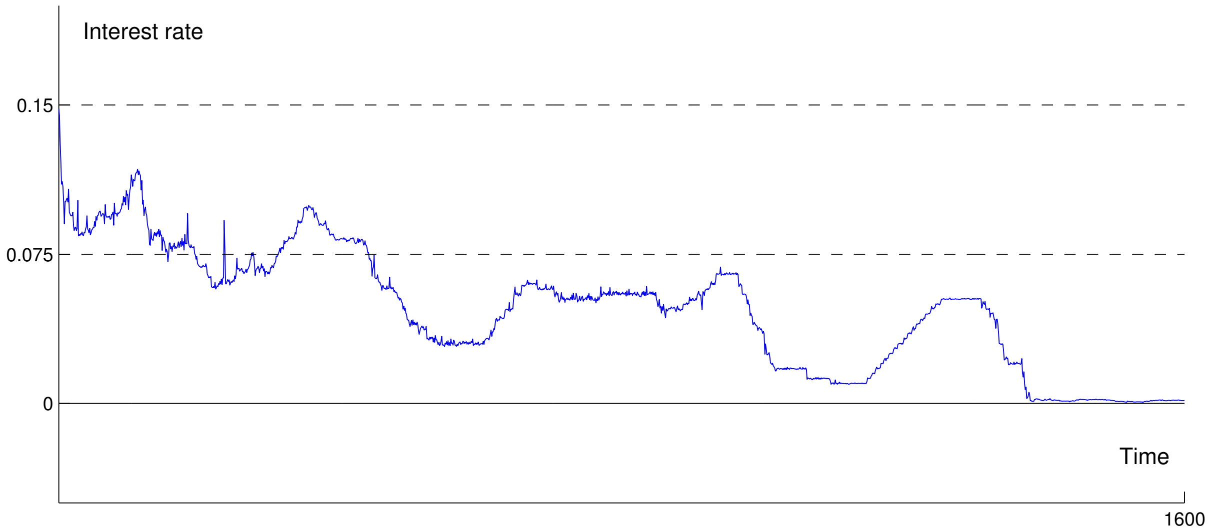
$$r(t_{i+1}) = a + b\cos\left(\theta W_{t_i} + \arccos\left(\frac{r(t_i) - a}{b}\right)\right), \quad i = 1, 2, 3, \cdots, \tag{3.4.1}$$

where W_{t_i} is a standard Brownian motion following a normal distribution $N(0, \Delta t)$. The deterministic parameters in (3.4.1) are a, b and θ. These parameters can be estimated by historical interest rate data. Let $\hat{a}$, $\hat{b}$ and $\hat{\theta}$ be the estimators. Then the sample path simulation for (3.4.1) is performed by using the following discrete-time expression:

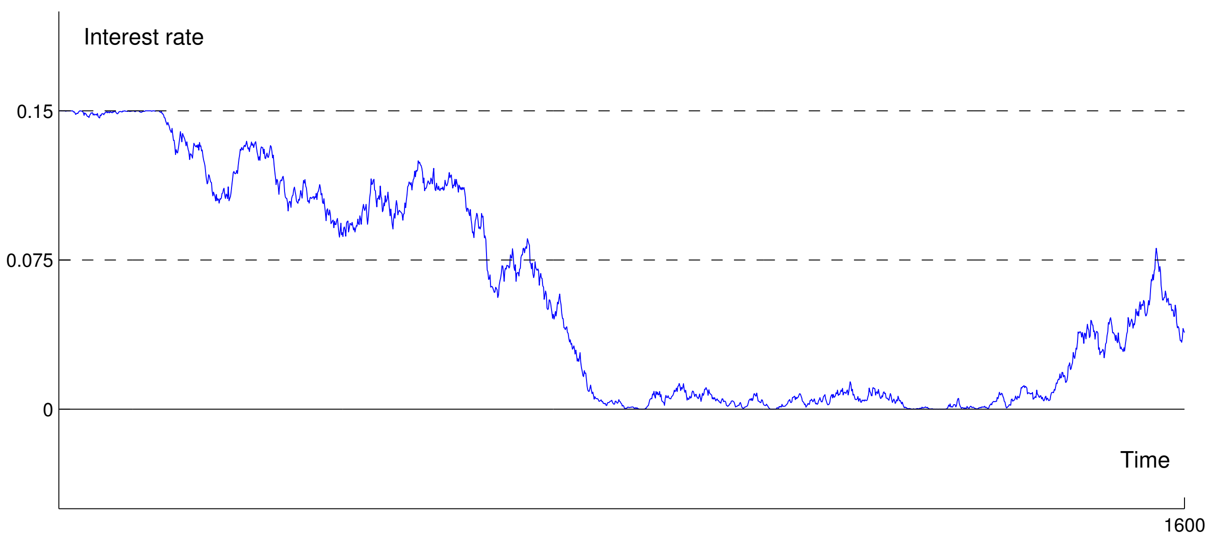
$$r(t_{i+1}) = \hat{a} + \hat{b}\cos\left(\hat{\theta}W_{t_i} + \arccos\left(\frac{r(t_i) - \hat{a}}{\hat{b}}\right)\right), \quad i = 1, 2, 3, \cdots, \tag{3.4.2}$$

For $r(t_i)$ in equation (3.4.2), the mean-reverting level is $\hat{a}$, upper bound $\hat{u} = \hat{a} + \hat{b}$ and lower bound $\hat{l} = \hat{a} - \hat{b}$.

The top figure in (3.4.1) shows US weekly (Wednesday) interest rate (Federal funds) from June 30th, 1982 to Feb 20th, 2013. It is a path of 1600 data moving between 0 and 15%. the bottom figure shows a trial path of 1600 steps with $a = 0.075$, $b = 0.075$, $\theta = 4$ and initial rate $r_0 = 0.15$ (these numbers are not estimators from the real data, since we just want to see how the paths look like). The reverting level is $a = 0.075$. The lower bound is set at $l = 0.075 - 0.075 = 0$ and the upper bound $u = 0.075 + 0.075 = 0.15$.



US weekly (Wednesday) interest rate (Federal funds)
from 6/30/1982 to 2/20/2013



A 1600−step simulated path
with a=0.075, b=0.075, θ=4, Δt=0.01, r_0=0.15

Figure 3.4.1: Historical interest rate data (top) and a simulated path of $r(t)$ (bottom)

Chapter 4

Term structures and Yield Curves

4.1 Introduction

A good reference for this section is Steven E. Shreve's *Stochastic Calculus for Finance II Continuous-Time Models* [15].

Real financial markets do not have a single interest rate. Instead, they have bonds of different maturities, some paying coupons and others not paying coupons. From these bonds, one can ultimately determine prices of zero-coupon bonds for a number of different maturity dates. Each of these bonds has a *yield* specific to its maturity, where yield is defined to be the constant continuously compounding interest rate over the lifetime of the bond that is consistent with its price:

$$\text{price of zero-coupon bond} = \text{face value} \times e^{\text{-yield} \times \text{time to maturity}}.$$

The *face value* of a zero-coupon bond is the amount it promises to pay upon maturity. The formula above implies that capital equal to the price of the bond, invested at a continuously compounded interest rate equal to the yield, would, over the lifetime of the bond, result in a final payment of the face value. In this chapter, we shall assume zero-coupon bonds by taking the face value to be 1.

In summary, instead of having a single interest rate, real markets have a *yield curve*, which one can regard either as a function of finitely many yields plotted versus their corresponding maturities or more often as a function of a nonnegative real variable (time) obtained by interpolation from the finitely many maturity-yield pairs provided by the market. The interest rate (sometimes called the *short rate*) is an idealization corresponding to the shortest maturity yield or perhaps the overnight rate offered by the government, depending on the particular application.

We assume throughout this chapter that the bonds have no risk of default. One generally regards U.S. government bonds to be non-defaultable.

For some asset pricing models, we assume an evolution of the price of a primary asset or the prices of multiple primary assets under the actual measure and then determine the market prices of risk that enable us to switch to a risk neutral measure(relying on Girsanov's theorem). Zero-coupon bond prices are given by the risk neutral pricing formula, which implies that discounted zero-coupon bond prices are martingales under the risk-neutral measure. This implies that no arbitrage can be achieved by trading in the zero-coupon bonds and the money market. The actual probability measure and the market prices of risk never enter the picture.

The *discount process* is as given in (1.6.3),

$$D(t) = e^{-\int_0^t r(s)ds}.$$

A zero-coupon bond is a contract promising to pay a certain 'face' amount, which we take to be 1, at a fixed maturity date T. Prior to that, the bond makes no payments. We denote the price of zero-coupon bond at time t by $B(t,T)$, $0 \le t \le T$. Note that $B(T,T) = 1$. We assume that the price process $(B(t,T))_{0\le t\le T}$ is adapted.

The modeling of the bond price will be based upon the following hypothesis:

(H). *There is a probability* $\widetilde{\mathbb{P}}$ *equivalent to* $\mathbb{P}$*, under which, the process* $B^*(t,T)$ *defined by*

$$B^*(t,T) = D(t)B(t,T)$$

is a martingale.

This hypothesis has some interesting consequences. Indeed, the martingale property under $\widetilde{\mathbb{P}}$ and the equality $B(T,T) = 1$ yield

$$B^*(t,T) = \widetilde{\mathbb{E}}\big[B^*(T,T)|\mathcal{F}(t)\big]$$

which implies

$$D(t)B(t,T) = \widetilde{\mathbb{E}}\big[D(T)B(T,T)|\mathcal{F}(t)\big] = \widetilde{\mathbb{E}}\left[e^{-\int_0^T r(s)ds}\middle|\mathcal{F}(t)\right],$$

Because $D(t)$ is $\mathcal{F}(t)$-measurable, we eliminate it on both sides,

$$B(t,T) = \widetilde{\mathbb{E}}\left[e^{-\int_t^T r(s)ds}\middle|\mathcal{F}(t)\right]. \tag{4.1.1}$$

4.2 Pricing Bond

The interest rate equation determine by theorem (3.3.1) is

$$r(t) = a + b\cos(\theta W_t + \gamma).$$

For solving the bond price, we start directly from the pricing formula (4.1.1) under risk-neutral measure. That is

$$\begin{aligned} B(t,T) =& \widetilde{\mathbb{E}}\left[e^{-\int_t^T r(s)ds}\middle|\mathcal{F}(t)\right] \\ =& \widetilde{\mathbb{E}}\left[e^{-\int_t^T \left(a+b\cos(\theta\widetilde{W}_s+\gamma)\right)ds}\middle|\mathcal{F}(t)\right] \\ =& e^{-a(T-t)}\ \widetilde{\mathbb{E}}\left[e^{-b\int_t^T \cos(\theta\widetilde{W}_s+\gamma)\big)ds}\middle|\mathcal{F}(t)\right]. \end{aligned} \tag{4.2.1}$$

where

$$\widetilde{\mathbb{E}}(\widetilde{W_t}) = 0,$$

and

$$\widetilde{\mathrm{Var}}(\widetilde{W_t}) = t.$$

First we consider to compute $\mathbb{E}\left[e^{\int_t^T \cos \widetilde{W}_s ds}\middle|\mathcal{F}(t)\right]$. Representing $\cos \widetilde{W}_s$ by its Taylor series expansion, we have

$$\begin{aligned}
&\widetilde{\mathbb{E}}\left[e^{\int_t^T \cos \widetilde{W}_s ds}\middle|\mathcal{F}(t)\right] \\
=&\widetilde{\mathbb{E}}\left[\sum_{n=0}^{\infty} \frac{\left(\int_t^T \cos \widetilde{W}_s ds\right)^n}{n!}\middle|\mathcal{F}(t)\right] \\
=&1 + \widetilde{\mathbb{E}}\left[\sum_{n=1}^{\infty} \frac{\int_t^T \int_t^T \cdots \int_t^T \prod_{i=1}^n \cos \widetilde{W}_{s_i} ds_1 ds_2 \cdots ds_n}{n!}\middle|\mathcal{F}(t)\right].
\end{aligned} \tag{4.2.2}$$

The double integral $\int_a^b \int_a^b \cos W_{s_1} \cos W_{s_2} ds_1 ds_2$, with $s_1,\ s_2 \in [t, T]$, can be calculated as a sum of two integrals

$$\int_t^T \int_t^{s_2} \cos W_{s_1} \cos W_{s_2} ds_1 ds_2, \qquad \text{for } t \leq s_1 < s_2 \leq T$$

and

$$\int_t^T \int_t^{x_1} \cos W_{s_1} \cos W_{s_2} ds_2 ds_1, \qquad \text{for } t \leq s_2 < s_1 \leq T.$$

Since this two integrals are identical, then we have

$$\int_t^T \int_t^T \cos W_{s_1} \cos W_{s_2} ds_1 ds_2 = 2 \int_t^T \int_t^{s_2} \cos W_{s_1} \cos W_{s_2} ds_1 ds_2, \qquad \text{for } t \leq s_1 < s_2 \leq T.$$

By induction, the following statement can be proven,

$$\int_t^T \int_t^T \cdots \int_t^T \prod_{i=1}^{n} \cos \widetilde{W}_{s_i} ds_1 ds_2 \cdots ds_n = n! \int_t^T \int_t^{s_n} \cdots \int_t^{s_2} \prod_{i=1}^{n} \cos \widetilde{W}_{s_i} ds_1 ds_2 \cdots ds_n,$$

$$\text{for } t < s_1 < s_2 < \cdots < s_n < T.$$

Applying this result on the multiple integral $\int_t^T \int_t^T \cdots \int_t^T \prod_{i=1}^{n} \cos \widetilde{W}_{s_i} ds_1 ds_2 \cdots ds_n$ in (4.2.2), if $t < s_1 < s_2 < \cdots < s_n < T$ is assumed, $\widetilde{\mathbb{E}}\left[e^{\int_t^T \cos \widetilde{W}_s ds} \middle| \mathcal{F}(t)\right]$ is simplified as

$$\begin{aligned} &\widetilde{\mathbb{E}}\left[e^{\int_t^T \cos W_s ds} \middle| \mathcal{F}(t)\right] \\ =&1 + \widetilde{\mathbb{E}}\Bigg[\sum_{n=1}^{\infty} \frac{n! \int_t^T \int_t^{s_n} \cdots \int_t^{s_2} \prod_{i=1}^{n} \cos \widetilde{W}_{s_i} ds_1 ds_2 \cdots ds_n}{n!} \Bigg| \mathcal{F}(s)\Bigg] \\ =&1 + \sum_{n=1}^{\infty} \int_t^T \int_t^{s_n} \cdots \int_t^{s_2} \widetilde{\mathbb{E}}\Bigg[\prod_{i=1}^{n} \cos W_{s_i} \Bigg| \mathcal{F}(t)\Bigg] ds_1 ds_2 \cdots ds_n. \end{aligned} \tag{4.2.3}$$

Next we work on $\widetilde{\mathbb{E}}\left[\prod_{i=1}^{n} \cos W_{s_i} \middle| \mathcal{F}(t)\right]$, which is the expectation of the product of a sequence of $\cos W_{s_i}$. In order to solve for the expectation, we can repeat angle-product-to-sum identity for cosine, which is

$$\cos A \cos B = \frac{1}{2}\big(\cos(A+B) + \cos(A-B)\big),$$

so that the product of the sequence can be expanded to a sum of cosine functions, the iterative procedure is shown as a binomial tree in Figure (4.2.1) and the formula is expressed as

$$\prod_{i=1}^{n} \cos W_{s_i} = \frac{1}{2^{n-1}} \sum \cos(W_{s_1} \pm W_{s_2} \pm \cdots \pm W_{s_n}). \tag{4.2.4}$$

The sign of the first Brownian motion W_{s_1} in each linear combination is always positive. To be consistent to the other Brownian motion $\pm W_{s_2}, \cdots, \pm W_{s_n}$, considering the symmetrical

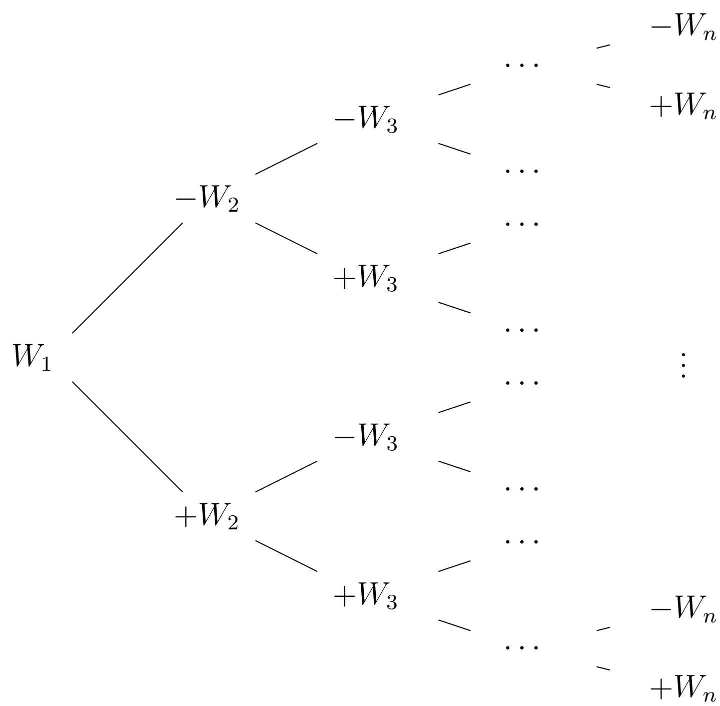


Figure 4.2.1: Binomial tree for linear combination of W_{s_i} in each cosine function

property of cosine function

$$\cos(W_{s_1}) = \cos(-W_{s_1}),$$

equation (4.2.4) is equivalent to

$$\prod_{i=1}^{n} \cos W_{s_i} = \frac{1}{2^n} \sum \cos(\pm W_{s_1} \pm W_{s_2} \pm \cdots \pm W_{s_n}). \tag{4.2.5}$$

Thus equation (4.2.5) shows that the product of n cosine functions is converted into a sum of 2^n cosine function. This identity will help us later to compute their expectations. This result (4.2.5) is stated in the next Theorem.

Theorem 4.2.1. *Let $\sum_{q=1}^{n} a_{p,q}X_q$ be any of 2^n possible combination of a sequence of n variables X_q, where $1 \le q \le n$, $1 \le p \le 2^n$ and $a_{p,q} = \pm 1$, then*

$$\prod_{i=1}^{n} \cos W_{s_i} = \frac{1}{2^n} \sum_{p=1}^{2^n} \cos\left(\sum_{q=1}^{n} a_{p,q} X_q \right),$$

where $\sum_{p=1}^{2^n}$ denotes the sum of all 2^n possible combinations.

By Theorem (4.2.1), the expectation $\widetilde{\mathbb{E}}\left[\prod_{i=1}^{n} \cos W_{s_i} \middle| \mathcal{F}(t) \right]$ in (4.2.3) is written as

$$\widetilde{\mathbb{E}}\left[\prod_{i=1}^{n} \cos W_{s_i} \middle| \mathcal{F}(t) \right] = \frac{1}{2^n} \sum_{p=1}^{2^n} \widetilde{\mathbb{E}}\left[\cos\left(\sum_{q=1}^{n} a_{p,q} W_{s_q} \right) \middle| \mathcal{F}(t) \right]. \tag{4.2.6}$$

To 'take out' what is $\mathcal{F}$-measurable from (4.2.6), let $C_p = \sum_{q=1}^{n} a_{p,q}$, where $1 \le q \le n$ and $1 \le p \le 2^n$, recalling the theorem (2.3.1) derived in chapter 2,

$$\begin{aligned}
&\widetilde{\mathbb{E}}\left[\prod_{i=1}^{n} \cos W_{s_i} \middle| \mathcal{F}(t)) \right] \\
=&\frac{1}{2^n} \sum_{p=1}^{2^2} \widetilde{\mathbb{E}}\left[\cos\left(\sum_{q=1}^{n} a_{p,q} W_{s_q} \right) \middle| \mathcal{F}(t) \right] \\
=&\frac{1}{2^n} \sum_{p} \widetilde{\mathbb{E}}\left[\cos\left(\sum_{q} a_{p,q}(W_{s_q} - W_t) + C_p W_t \right) \middle| \mathcal{F}(t) \right] \\
=&\frac{1}{2^n} \sum_{p} \cos(C_p W_t) \widetilde{\mathbb{E}}\left[\cos\left(\sum_{q} a_{p,q} W_{s_q - t} \right) \right].
\end{aligned} \tag{4.2.7}$$

Now we are focusing on the task of computing $\mathbb{E}\left[\cos\left(\sum_{q=1}^{n} a_{p,q} W_{s_q - t} \right) \right]$. Theorem (2.1.1) says that for any normally distributed variable X,

$$\mathbb{E}\cos(\theta X) = e^{-\frac{\theta^2 \mathrm{Var}(X)}{2}}. \tag{4.2.8}$$

Let $X = \sum_{q=1}^{n} a_{p,q} W_{s_q - t}$. We know that a linear combination of mutually independent random variables following normal distributions also has normal distribution. W_{s_1-t}, W_{s_2-t}, $\cdots$ and W_{s_n-t} are not independent to each other, however each of them can be decomposed into non-overlapping time increments, which are mutually independent. In this case X is the linear combination of these time increments and X is normally distributed. Therefore, theorem (2.1.1) in Chapter 2 can be used.

$$\begin{aligned}
&\widetilde{\text{Var}}(X) \\
=&\widetilde{\text{Var}}\Bigg[\sum_{q=1}^{n}(a_{p,q}W_{s_q-t})\Bigg] \\
=&\sum_{q=1}^{n}\widetilde{\text{Var}}(a_{p,q}\widetilde{W}_{s_q-t}) + \sum_{h=1}^{n}\sum_{k=1}^{n} a_{p,h}a_{p,k}\widetilde{\text{Cov}}(\widetilde{W}_{s_h-t}, \widetilde{W}_{s_k-t}) \\
=&\sum_{q}(s_q - t) + \sum_{h}\sum_{k} a_{p,h}a_{p,k}\min(s_h - t, s_k - t) \\
=&\sum_{q}(s_q - t) + 2\sum_{1\le h<}\sum_{k\le n} a_{p,h}a_{p,k}(s_i - t).
\end{aligned} \tag{4.2.9}$$

We put the results (4.2.7), (4.2.8) and (4.2.9) together and obtain

$$\begin{aligned}
&\widetilde{\mathbb{E}}\Bigg[\prod_{i=1}^{n}\cos\widetilde{W}_{s_i}\Bigg|\mathcal{F}(t))\Bigg] \\
=&\frac{1}{2^n}\sum_{p}\cos(C_p\widetilde{W}_t)\exp\Bigg[-\frac{1}{2}\Bigg(\sum_{q=1}^{n}(s_q - t) + 2\sum_{1\le h<}\sum_{k\le n} a_{p,h}a_{p,k}(s_i - t)\Bigg)\Bigg].
\end{aligned} \tag{4.2.10}$$

Finally we go back to (4.2.3)

$$\begin{aligned}
&\widetilde{\mathbb{E}}\Bigg[e^{\int_t^T \cos\widetilde{W}_s ds}\Bigg|\mathcal{F}(t)\Bigg] \\
=&1 + \sum_{n=1}^{\infty}\int_t^T\int_t^{s_n}\cdots\int_t^{s_2}\widetilde{\mathbb{E}}\Bigg[\prod_{i=1}^{n}\cos\widetilde{W}_{s_i}\Bigg|\mathcal{F}(t)\Bigg]ds_1 ds_2\cdots ds_n,
\end{aligned}$$

and substitute $\widetilde{\mathbb{E}}\left[\prod_{i=1}^n \cos\widetilde{W}_{s_i}\middle|\mathcal{F}(t)\right]$ by (4.2.10). Right now it is easier to see a little further that

$$\begin{aligned}
&\widetilde{\mathbb{E}}\left[e^{-b\int_t^T \cos(\theta\widetilde{W}_s+\gamma)ds}\middle|\mathcal{F}(t)\right]\\
=&1+\sum_{n=1}^{\infty}(-b)^n\int_t^T\int_t^T\cdots\int_t^T \widetilde{\mathbb{E}}\left[\prod_{i=1}^n \cos(\theta\widetilde{W}_{s_i}+\gamma)\middle|\mathcal{F}(t)\right]ds_1ds_2\cdots ds_n,
\end{aligned}\tag{4.2.11}$$

where

$$\begin{aligned}
&\widetilde{\mathbb{E}}\left[\prod_{i=1}^n \cos(\theta\widetilde{W}_{s_i})\middle|\mathcal{F}(t))\right]\\
=&\frac{1}{2^n}\sum_p \cos\big(C_p(\theta\widetilde{W}_t+\gamma)\big)\exp\left[-\frac{\theta^2}{2}\left(\sum_{q=1}^n(s_q-t)+2\sum_{1\le h<}\sum_{k\le n}a_{p,h}a_{p,k}(s_i-t)\right)\right].
\end{aligned}\tag{4.2.12}$$

4.3 Yield Curve

The yield curve is used to show the relationship between yield and maturity, which is know as the *term* of the same type of bond. Once zero-coupon bond prices have been computed, we can define the yield between time t and T to be

$$Y(t,T) = -\frac{1}{T-t}\ln B(t,T). \tag{4.3.1}$$

Recalling equation (4.2.1)

$$B(t,T) = e^{-a(T-t)}\;\widetilde{\mathbb{E}}\left[e^{-b\int_t^T \cos(\theta\widetilde{W}_s+\gamma)\big)ds}\middle|\mathcal{F}(t)\right].$$

By (4.3.1),

$$Y(t,T) = a - \frac{1}{T-t}\ln\widetilde{\mathbb{E}}\left[e^{-b\int_t^T \cos(\theta\widetilde{W}_s+\gamma)\big)ds}\middle|\mathcal{F}(t)\right].$$

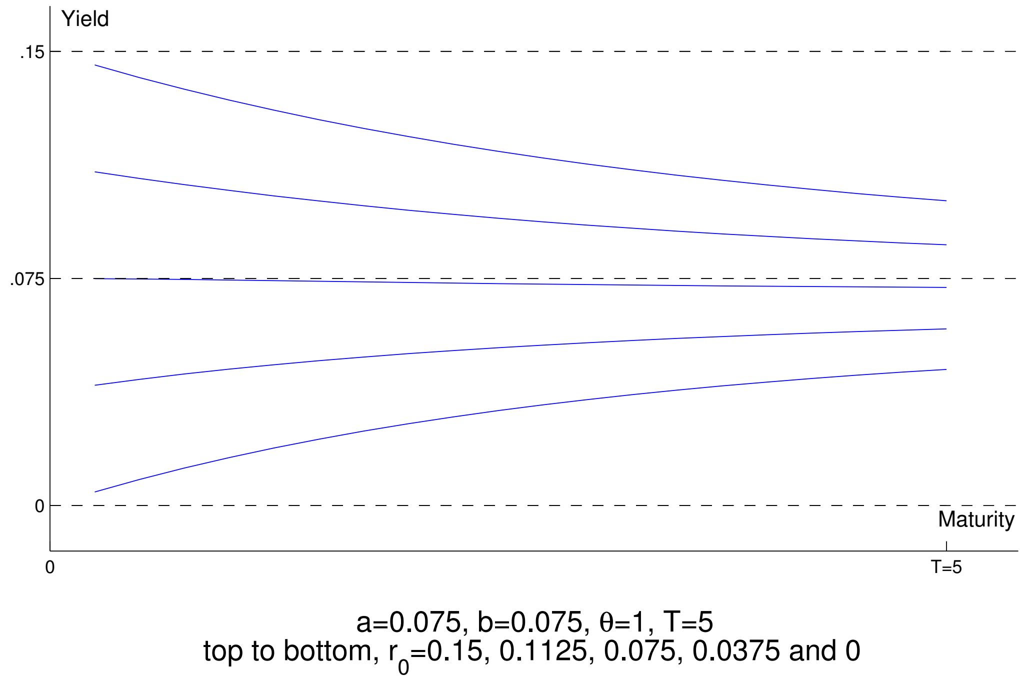


Figure 4.3.1: Five possible yield curves

Here $\widetilde{\mathbb{E}}\left[e^{-b\int_t^T \cos(\theta\widetilde{W}_s+\gamma)\big)\,ds}\middle|\mathcal{F}(t)\right]$ is calculated as in equation (4.2.11) and (4.2.12). The number of exponential terms in equation (4.2.11) and (4.2.12) increases exponentially as n grows. Hence it is difficult (if possible) to compute exact values of yield by using the above formulas. To demonstrate some of the possible yield curves by this model, Figure (4.3.1) is generated in MATLAB through approximation (computation truncates at $n = 6$), with $a = 0.075$, $b = 0.075$, $\theta = 1$ and $T = 5$. In a top-down order, the initial interest rates r_0 are 0.15, 0.1125, 0.075, 0.0375 and 0. The mean-reverting level is $a = 0.075$, lower bound is $l = a - b = 0$ and upper bound is $u = a + b = 0.15$. We notice that yield curves tend to start from initial interest rates and approach the mean-reverting level with maturity. The yield curve (middle) starts from 0.075 is supposed to maintain the same level, however it deviates a little with maturity. We believe this trend is set by the truncation error and all the yield curves should have minute deviations from the real ones.

4.4 Summary and Future Research

The main work of this dissertation is to propose a new process for the short-term risk-free rate, r,

$$dr = \beta(\alpha - r)dt + \boldsymbol{\sigma}\sqrt{1 - \mathcal{L}^2(r)}\, dW_t.$$

This process is based on the very classic equilibrium model of Vasicek, which was subsequently improved by Cox-Ingersoll-Ross (CIR)(1985) [4]. Vasicek model (1977) [16] brings us the key feature of mean-reversion, however it allows the interest rate to fall below zero. Cox-Ingersoll-Ross later introduced a lower-bound of zero on the Vasicek model)(1977) [16]. Even though mean reversion could draw the interest rate to the level defined by the drift term, both models could not provide control on how high the interest rate can potentially grow. The model studies in this dissertation kept both features of Vasicek (1977) [16] and CIR (1985) [4] and essentially introduces an upper limitation to the picture.

To solve this *sde*, we want to leverage $\cos(W_t)$, which leads to a study on properties of $\cos(W_t)$. Monte-Carlo technique was used to simulate the path of interest rate. We then searched for what this process with $\cos(W_t)$ implies for bond prices and the term structures.

A key component in this dissertation is the Brownian motion on a circle, that is $\cos W_t$. One extension of this model is to study the behavior of the Brownian motion on a general closed contour. For instance, we may consider the Brownian motion moving on a square, an ellipse, or any closed simple contour. Those contours are continuous but do not have to be smooth or differentiable. Such non-differentiable curves would enable us to model random rate jumps and/or pauses at certain levels. All the results would hold except for more lengthy algebraic manipulation.

As for numerical implementation of the term-structure, the MATLAB files could be improved to accelerate the computational speed for more accurate solution and longer expiration times.

Finally, the parameters involved in this stochastic interest rate model need to be determined by statistical estimation from historical data, which would be a challenging but exciting research beyond the scope of this dissertation.

REFERENCES

[1] Yacine Ait-Sahalia. Testing continuous-time models of the spot interest rate. *Review of Financial Studies*, 9(2):385–426, 1996.

[2] Fischer Black and Karasinski Piotr. Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, 47(4):52–59, 7 1991.

[3] Thomas E. Copeland, J. Fred Weston, and Kuldeep Shastri. *Financial Theory and Corporate Policy*. Pearson Addison Wesley, fourth edition, 2005.

[4] John C. Cox, Jonathan E. Jr. Ingersoll, and Ross Stephen A. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 3 1985.

[5] Darrell Duffie and Rui Kan. A yield factor model of interest rates. *Mathematical Finance*, 6(4):379–406, 12 1996.

[6] Thomas S. Y. Ho and Sang-Bin Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41(5):1011–1029, 12 1986.

[7] John Hull and Alan White. One-factor interest-rate models and the valuation of interest-rate derivative securities. *Journal of Financial and Quantitative Analysis*, 28(2):235–254, 6 1993.

[8] John C. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, sixth edition, 2005.

[9] Damien Lamberton and Bernard Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. Financial Mathematics. Chapman & Hall/CRC, second edition, 2008.

[10] F. Mercurio and J. M. Moraleda. An analytically tractable interest rate model with humped volatility. *European Journal of Operational Research*, 120(1):205–214, 1 2000.

[11] Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[12] Toshio Nakagawa. *Stochastic Processes - with Applications to Reliability Theory*, chapter 4. Springer, 2011.

[13] Richard J. Jr. Rendleman and Brit J. Bartter. The pricing of options on debt securities. *Journal of Financial and Quantitative Analysis*, 15(1):11–24, 3 1980.

[14] Sheldon M. Ross. *Stochastic Processes*. Probability And Mathematical Statistics. John Wiley & Sons, Inc, second edition, 1996.

[15] Steven E. Shreve. *Stochastic Calculus for Finance II Continuous-Time Models*. Springer, 2004.

[16] Oldrich Vasicek. An equilibrium characterization of the term structure of interest rates. *Journal of Financial Economics*, 5:177–188, 1977.