An Extension of Lévy's Theorem and Applications to Financial Models Based on Futures Prices

Diego Jara
Department of Mathematical Sciences
Carnegie Mellon University

Ph. D. Thesis
Thesis Advisor: Professor David Heath

April, 2000

Abstract

In this work we introduce financial models based on the evolution of prices of futures contracts. We explore conditions under which these models are free of arbitrage and complete, and therefore are useful for pricing contingent claims with payoffs that are measurable with respect to the information provided by the future contracts. In cases where the contracts are futures on interest rates, the models provide an alternative way of studying the evolution of the term structure of interest rates, in a setup which is similar to the HJM framework. One difference, however, is the possibility of defining models where the state of the futures curve is determined by a low-dimensional vector-valued process.

We study the theoretical feasibility of using future models for financial modeling. In particular, we explore whether information about the distribution of the quadratic variation of the future prices is enough to determine the distribution of the future prices. This is equivalent to studying the possibility of extending Lévy's theorem of characterization of Brownian Motion. In particular, we pose the question of whether two martingales with different laws may have quadratic variations with equal laws. We give answers to this question for two classes of continuous martingales: martingales with divergent and absolutely continuous quadratic variation, and martingales which are weak solutions to driftless Stochastic Differential Equations in which the volatility depends only on the martingale itself. In the first case, we conclude that martingales with different laws may have quadratic variations with equal laws. In the second case, we find that two elements of this class of martingales that have quadratic variations with equal distributions must have the same law, modulo reflection.

ACKNOWLEDGEMENTS

I want to thank my advisor, David Heath, for providing ideas and helping me explore different areas that gave birth to this work. I also want to thank Steven Shreve for his useful comments and discussions, Pierre Collin-Dufresne and Dev Joneja for serving in my committee, and Marc Yor for his helpful suggestions.

I dedicate this work to my wife, Bala, and to my parents, Chacho and Magola.

Contents

Chapter 1

Introduction

The objective of this work is twofold. First, to establish a financial model which allows continuous trading on futures contracts on an arbitrary spot process. Second, to study the possibility of extending Lévy's Theorem of characterization of Brownian Motion to a more general class of martingales.

These two objectives are not independent. It will be seen that for pricing purposes, the determination of a model based on futures prices depends only on the specification of a "volatility" process. In practice, however, it is reasonable to assume that we are given information about the quadratic (cross) variation of future prices. Therefore, from a modeling perspective it is important to decide whether this information is enough to determine the distribution of futures prices. We recall that Lévy's Theorem of characterization of Brownian Motion states that if a continuous martingale $\mathcal{N}$ yields the same quadratic variation as Brownian Motion, then $\mathcal{N}$ is a Brownian Motion. Therefore, the previous lines suggest studying the possibility of extending Lévy's Theorem to classes of martingales larger than the class of Brownian Motions.

The definition of futures-based models is motivated by the increasing demand experienced in the financial world during the past decades for probability models that provide a balance between their capability to:

1. Produce prices for liquid assets that are close to the market prices.

2. Provide simple implementations that allow efficient pricing of contracts.

Following the ideas introduced by Black and Scholes (1973) and Merton (1973), these models produce prices under the assumption that there are no arbitrage opportunities. Under completeness, prices of assets coincide with the prices of their replicating portfolios.

Mathematically, the futures models introduced here are very similar to standard continuous-time financial models found in the literature. However, futures models allow term-structure modeling; i.e., at every time we are given information of a whole curve, as opposed to a (spot) price. This will aid replicating market prices (item 1 above), since current futures prices are used as a starting point, and not as a calibrating target. The tradable securities are interpreted as future contracts with different maturities on an underlying spot process such as the price of a share of stock, a bond, a commodity, or the value of an interest rate. Initially the underlying spot process is left unspecified. In specific situations, we may define a low dimensional process that contains the information necessary to derive the state of the futures curve. This will allow PDE or recombining tree methods to be used for pricing purposes. Briefly, these particular models preserve the theoretical advantages of the HJM framework, and allow the use of standard pricing techniques (item 2 above). Musiela and Rutkowski (1998) have studied the evolution of finitely many futures contracts that trade in discrete time. To the best of our knowledge, ours is the first effort to provide a financial model in continuous time where the tradable assets are interpreted as general future contracts.

The implementation of futures models raises the question of whether the distribution of the quadratic variation of futures prices is enough to determine the model that produced those prices. Mathematically, let $\mathcal{M}$ be a class of continuous martingales starting at 0. Let $M, N \in \mathcal{M}$ be given, and assume that $\langle M \rangle$ and $\langle N \rangle$ have the same law. We are interested in comparing the laws of M and N. In particular, we search for classes $\mathcal{M}$ for which we can conclude that M and N have the same law. Lévy's theorem states that Brownian Motion is characterized by its quadratic variation. This is extended to the case in which $\mathcal{M}$ is the class of Gaussian continuous martingales. To study this question is to analyze the possibility of extending Lévy's theorem in a particular direction. In this work we give answers for two classes of martingales:

1. $\mathcal{M}_1 = \{M : M$ is a continuous martingale, such that $M_0 = 0$,

   $\langle M \rangle_\infty = \infty, \langle M \rangle$ is a.s. absolutely continuous and $\frac{d}{dt}\langle M \rangle_t > 0 \ a.s.\}$

2. $\mathcal{M}_2 = \{M : M$ is a continuous martingale, such that $M_t = \int_0^t g(M_s)dW_s$

   for a Brownian Motion W and a measurable function g whose zeroes

   coincide with the set of points where $\frac{1}{g}$ is not locally square integrable.$\}$

It will be proved that Lévy's theorem does not accept an extension in the case of $\mathcal{M}_1$:

Theorem 1. *For every martingale $M \in \mathcal{M}_1$ that is not Gaussian, there exists $N \in \mathcal{M}_1$ with different law than M, such that $\langle M \rangle \stackrel{d}{=} \langle N \rangle$.*

This result suggests restricting the class of martingales. Unfortunately, it does not provide information on how different the laws of M and N are. In $\mathcal{M}_2$, Lévy's theorem has an extension, modulo reflection of the martingale. More precisely,

Theorem 2. *Let $M, N \in \mathcal{M}_2$ be given. If $\langle M \rangle \stackrel{d}{=} \langle N \rangle$, then $M \stackrel{d}{=} N$ or $M \stackrel{d}{=} -N$.*

In Chapter 2 we introduce futures models in a continuous time setup where the basic securities are future contracts and a money market account. We define trading strategies for these models, and give definitions of arbitrage opportunities and completeness of the markets. We give conditions under which markets are free of arbitrage and complete.

In Chapter 3, examples of future models are exhibited. First, we study futures models with unspecified underlying spot process and volatilities that are linear functions of futures prices. Second, models of future interest rates will be studied. Two examples of such models are given: futures contracts that settle to the spot interest rate, and to 3-month LIBOR. We discuss various practical issues of these models, such as consistency and pricing of fixed income contracts.

In Chapter 4 we prove Theorems 1 and 2, and provide specific applications to financial modeling.

Chapter 2

Futures Models

In this chapter we exhibit a continuous-time financial model with finite horizon, where basic securities are future contracts on an underlying spot process. We consider such contracts for infinitely many maturity dates, but allow trading strategies to take positions on finitely many of these contracts. This setup is very similar to most widely used models in mathematical finance. Different specifications of such economies have led to different versions of the First and Second Fundamental Theorems of Arbitrage Pricing, which aim at characterizing absence of arbitrage and completeness of the models, respectively. We exhibit versions of these theorems for futures models. More precisely, we investigate relationships between existence of an equivalent martingale measure and absence of arbitrage, and the uniqueness of an equivalent martingale measure and completeness of the model in sections 2.2 and 2.3.

In the first section we describe securities and define trading strategies in the model. In the second section we define arbitrage in the economy, and relate the absence of arbitrage in a futures model to the existence of a measure equivalent to the original measure under which basic securities are (local) martingales. In the third section we define completeness of futures models, and link completeness of a model with uniqueness of an equivalent (local) martingale measure. In the fourth section we give comments about the use of futures models. First, we talk about the use of futures models for pricing and hedging of financial contracts. Second, we pose questions about sufficient information to establish a futures model. These questions are the subject of study of the fourth chapter.

2.1 Description of the Futures Based Economy

Let $\tau > 0$ be a given fixed finite horizon. We consider an economy with continuous-time trading on a fixed interval $[0, \tau]$. We assume the existence of an underlying spot process (which can be interpreted as the price of a financial asset) on which standard futures contracts are traded for every maturity prior to τ. We also assume the existence of a spot interest rate process. The basic securities of the economy are a money market account, and "futures portfolios"; a futures portfolio consists of one futures contract, for which the continuous stream of cash flows originated by the marking to market activities are immediately invested in the money market account. The initial value of these portfolios is 0, since entering a futures contract requires no initial exchange of money.

More precisely, let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space on which is given an $n-$dimensional standard Brownian Motion $\{\mathbf{W}(t) = (W_1(t), \ldots, W_n(t)); 0 \leq t \leq \tau\}$. Let $\{\mathcal{F}(t); 0 \leq t \leq \tau\}$ be the augmented, right-continuous complete filtration generated by $\mathbf{W}$. That is, for $t \in [0, \tau]$,

$$\mathcal{F}(t) = \sigma\{\mathcal{F}^{\mathbf{W}}(t) \cup \mathcal{N}\},$$

where $\mathcal{N}$ is the set of $\mathbf{P}$-null subsets of $\mathcal{F}^{\mathbf{W}}(\tau)$. We assume $\mathcal{F} = \mathcal{F}_\tau$.

For $T \in [0, \tau]$, $t \in [0, T]$, let $F(t, T)$ denote the futures price at date t of a contract that settles to the value of the spot process at date T. It is convenient to specify $F(t, T) = F(T, T)$, for $0 \leq T \leq t \leq \tau$. Denote by $r(t)$ the spot interest rate at date $t \in [0, \tau]$. Assume that futures prices and the spot interest rate satisfy

M1. For $0 \leq t \leq T \leq \tau$,

$$F(t, T) = F(0, T) + \int_0^t \alpha(s, T) ds + \int_0^t \boldsymbol{\sigma}(s, T) d\mathbf{W}(s)', \tag{2.1}$$

where

1. $F(0, \cdot)$ is a measurable function from $([0, \tau], \mathcal{B}([0, \tau]))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\int_0^\tau |F(0, s)| ds < \infty.$$

2. For $T \in [0, \tau]$, $\alpha(\cdot, T)$ is an adapted process such that

$$\alpha(t, T) = 0, \; 0 \leq T \leq t \leq \tau, \text{ and}$$
$$\int_0^\tau |\alpha(t, T)| dt < \infty \text{ a.s.}, \; 0 \leq T \leq \tau.$$

3. For $T \in [0, \tau]$, $\boldsymbol{\sigma}(\cdot, T) = (\sigma_1(\cdot, T), \ldots, \sigma_n(\cdot, T))$ is an adapted process such that

$$\boldsymbol{\sigma}(t, T) = (0, \ldots, 0), \; 0 \leq T \leq t \leq \tau, \text{ and}$$
$$\int_0^\tau \|\boldsymbol{\sigma}(t, T)\|^2 dt < \infty \text{ a.s.}, \; 0 \leq T \leq \tau.$$

M2. For $t \in [0, \tau]$,

$$r(t) = r_0 + \int_0^t a(s) ds + \int_0^t \mathbf{b}(s) d\mathbf{W}(s)',$$

where

1. $r(0) = r_0$ is a given initial spot rate.
2. For $T \in [0, \tau]$, a is an adapted process such that

$$\int_0^\tau |a(t)|dt < \infty \ a.s. \tag{2.2}$$

3. For $T \in [0, \tau]$, $\mathbf{b} = (b_1, \ldots, b_n)$ is an adapted process such that

$$\int_0^\tau \|\mathbf{b}(t)\|^2 dt < \infty \ a.s.$$

Definition 1. A *Futures Model* $\mathcal{M}$ is a collection of

1. Future curves F that satisfy [M1].
2. A spot rate process r that satisfies [M2].
3. $\mathcal{T} \subset [0, \tau]$, the maturities of tradable futures contracts. We assume $(0, \tau) \cap \mathcal{T} \neq \varnothing$.
4. A *Final Trading Date*, $T < \tau$. If $\mathcal{T}$ is finite, $T \leq \min(\mathcal{T} \cap (0, \tau))$. Otherwise, we require that $\mathcal{T} \cap (T, \tau]$ contains at least n elements.

Elements of a futures model $\mathcal{M} = \{F, r, \mathcal{T}, T\}$ will be identified by $F^{\mathcal{M}}, r^{\mathcal{M}}, \mathcal{T}^{\mathcal{M}}, T^{\mathcal{M}}$.

Remark 2.1.1. Item 3 is used to distinguish models that allow trading on different futures contracts. Generally, financial models specify a finite set of tradable assets. Futures models may specify an infinite number of such assets. However, trading will be allowed in only finitely many assets (Definition 2 below).

Remark 2.1.2. The underlying process of the futures prices may not be independent of the spot interest rate. In such an event, further restrictions should be imposed on these processes to ensure consistency of the model. This aspect will be considered in particular examples of futures models based on interest rates, which will be studied in Chapter 3. For the remainder of this chapter, we assume that futures models contain spot rate and futures prices consistent with each other.

Trading in Futures Models. We assume no transaction costs, and no constraints in liquidity or short-selling of securities. Trading is allowed in two basic securities.

MM. *Money Market Account.* The value at time $t \in [0, \tau]$ of a unit of the money market account is

$$B(t) = \exp\Big(\int_0^t r(s)ds\Big). \tag{2.3}$$

By (2.2), $0 < B(T) < \infty$ a.s. for all $T \in [0, \tau]$.

FP. *Futures Portfolios.* For a given $T \in [0, \tau]$, the futures portfolio with maturity T contains one futures contract with maturity T, and a money market account on which gains (losses) from the futures contract will automatically be deposited. Briefly, the margin account of the futures contract is a money market account. After the expiration of the futures contract, the only activity of this portfolio is derived from the money market account. The price $Q(t, T)$ at date $t \in [0, \tau]$ of one unit of this portfolio satisfies

$$\begin{aligned} Q(t,T) &= \int_0^t dF(s,T) + \int_0^t \frac{Q(s,T)}{B(s)} dB(s) \\ &= F(t,T) - F(0,T) + \int_0^t Q(s,T) r(s) ds. \end{aligned} \tag{2.4}$$

Using Itô's rule, this is written

$$\frac{Q(t,T)}{B(t)} = \int_0^t \frac{\alpha(s,T)}{B(s)} ds + \int_0^t \frac{\boldsymbol{\sigma}(s,T)}{B(s)} d\mathbf{W}(s)'.$$

Definition 2. A *Trading Strategy* in a futures model $\mathcal{M}$ is a pair $\pi = (\mathcal{R}, \Psi)$ of maturities and positions such that

1. $\mathcal{R} = (T_0, T_1, \dots, T_M) \in (\mathcal{T}^{\mathcal{M}})^{M+1}$, where $0 = T_0 < T_1 < \dots < T_M \le T^{\mathcal{M}}$, for a given $M \in I\!N$.
2. $\Psi = \{\Psi(t) = (\Psi_0(t), \Psi_1(t), \dots, \Psi_M(t)); 0 \le t \le \tau\}$ is an adapted process such that

$$\int_0^\tau \left| \sum_{i=1}^M \Psi_i(t) \alpha(t, T_i) \mathbf{1} \right| dt < \infty \text{ a.s., where } \mathbf{1} = (1, 1, \dots, 1) \in I\!R^M.$$

$$\int_0^\tau \left\| \sum_{i=1}^M \Psi_i(t) \boldsymbol{\sigma}(t, T_i) \right\|^2 dt < \infty \text{ a.s.}$$

The *Wealth Process* of a Trading Strategy $\pi = (\mathcal{R}, \Psi)$ on a futures model $\mathcal{M}$ is denoted $\{V^\pi(t); 0 \le t \le T^{\mathcal{M}}\}$ and satisfies

$$V^\pi(t) = \Psi_0(t) B(t) + \sum_{i=1}^M \Psi_i(t) Q(t, T_i).$$

A trading strategy π on a futures model $\mathcal{M}$ is *Admissible* if there exists $a \in I\!R$ such that $V^\pi(t) > a$ a.s. for every $t \in [0, T^{\mathcal{M}}]$.
A trading strategy $\pi = (\mathcal{R}, \Psi)$ is *Self-Financing* if

$$V^\pi(t) = V^\pi(0) + \sum_{i=1}^M \int_0^t \Psi_i(s) dF(s, T_i) + \int_0^t \frac{V^\pi(s)}{B(s)} dB(s) \ \forall t \in [0, T^{\mathcal{M}}].$$

Remark 2.1.3. A self-financing trading strategy satisfies (for $t \in [0, T^{\mathcal{M}}]$)

$$V^\pi(t) = V^\pi(0) + \sum_{i=1}^{M} \int_0^t \Psi_i(s) dQ(s, T_i) + \int_0^t \Psi_0(s) dB(s), \text{ and}$$
$$\frac{V^\pi(t)}{B(t)} = V^\pi(0) + \sum_{i=1}^{M} \int_0^t \Psi_i(s) \frac{\alpha(s, T_i)}{B(s)} ds + \sum_{i=1}^{M} \int_0^t \Psi_i(s) \frac{\boldsymbol{\sigma}(s, T_i)}{B(s)} d\mathbf{W}(s)'.$$

Remark 2.1.4. The concept of admissible strategy was introduced in Harrison and Pliska (1981). It is common practice to restrict trading strategies in this way to exclude situations such as suicide or doubling strategies.* From a practical perspective, this restriction is sensible, since unbounded temporary losses that eventually yield sure profits are not realistically sustainable.

Remark 2.1.5. The previous definitions allow entering a futures contract at any time. In fact, fix $T \in \mathcal{T}^{\mathcal{M}}$, and $t < T$. Define the self-financing trading strategy $\pi = (\mathcal{R}, \Psi)$ with $M = 1$, $\mathcal{R} = (T)$ and

$$\Psi_0(s) = \begin{cases} 0, & s < t \\ -\frac{Q(t,T)}{B(t)}, & s > t \end{cases}$$
$$\Psi_1(s) = \begin{cases} 0, & s < t \\ 1, & s > t \end{cases}$$

There is no activity for this portfolio before date t. At date t, it costs nothing to enter the strategy. For $s \in [t, T]$, the wealth process satisfies

$$V^\pi(s) = -Q(t, T) \frac{B(s)}{B(t)} + Q(s, T) = F(s, T) - F(t, T) + \int_t^s V^\pi(u) r(u) du.$$

With this portfolio one effectively buys at date t the futures contract with maturity T.

2.2 No Arbitrage in Futures Models

In this section we define arbitrage in a futures models, and exhibit conditions under which a futures model will be free of arbitrage.

Definition 3. A futures model $\mathcal{M}$ admits *Arbitrage* if there exists an admissible self-financing trading strategy $\pi = (\mathcal{R}, \Psi)$ on $\mathcal{M}$ that satisfies

$$V^\pi(0) = 0$$
$$\exists t \in [0, T^{\mathcal{M}}] \; V^\pi(t) \geq 0 \text{ a.s.}, \; \mathbf{P}[V^\pi(t) > 0] > 0.$$

*In suicide strategies, the wealth process starts at 1 dollar, and ends at 0 for sure. However, unlimited gains must be sustained in between. Reversing signs, these strategies bring sure profits upon sustaining unbounded losses. See, for example, Harrison and Pliska (1981).

We are interested in working with models that are free of arbitrage. To this end we give conditions that are necessary and sufficient for the absence of arbitrage.

Definition 4. An *Equivalent (Local) Martingale Measure* (henceforth EMM (ELMM)) for a futures model $\mathcal{M}$ is a measure $\tilde{\mathbf{P}}$ on $(\Omega, \mathcal{F}_{T^{\mathcal{M}}})$ equivalent to $\mathbf{P}$ (i.e., for $A \in \mathcal{F}_{T^{\mathcal{M}}}$, $\mathbf{P}(A) = 0 \leftrightarrow \tilde{\mathbf{P}}(A) = 0$) such that

$$\forall T \in \mathcal{T}^{\mathcal{M}} \left\{ \frac{Q(t,T)}{B(t)}, \mathcal{F}_t : 0 \le t \le T^{\mathcal{M}} \right\} \text{ is a } \tilde{\mathbf{P}}\text{-(local) martingale .}$$

The name *Risk-Neutral Measure* will be used to refer to an equivalent martingale measure.

Lemma 2.2.1. *Let $\mathcal{M}$ be a given futures model. The existence of an ELMM for $\mathcal{M}$ is equivalent to the existence of an adapted process $\boldsymbol{\gamma} = \{(\gamma_1(t), \dots, \gamma_n(t)); 0 \le t \le T^{\mathcal{M}}\}$, such that*

1. *For $T \in \mathcal{T}^{\mathcal{M}}$, $\alpha(t,T) + \boldsymbol{\sigma}(t,T)\boldsymbol{\gamma}(t)' = 0$ a.s. on $([0,T^{\mathcal{M}}] \times \Omega, \mathcal{B}([0,T^{\mathcal{M}}]) \times \mathcal{F}, \lambda \times \mathbf{P})$.*
2. *$\mathbf{P}[\int_0^{T^{\mathcal{M}}} \|\boldsymbol{\gamma}(t)\|^2 dt < \infty] = 1$.*
3. *$\mathbb{E}[\exp\{\int_0^{T^{\mathcal{M}}} \boldsymbol{\gamma}(t) d\mathbf{W}(t)' - \frac{1}{2}\int_0^{T^{\mathcal{M}}} \|\boldsymbol{\gamma}(t)\|^2 dt\}] = 1$.*

Proof. Necessity. Assume $\tilde{\mathbf{P}}$ is an ELMM for $\mathcal{M}$. For $t \le T^{\mathcal{M}}$, call $Z_t = \frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}|_{F_t}$ the Radon-Nykodym derivative of $\tilde{\mathbf{P}}$ with respect to $\mathbf{P}$ with underlying sigma-algebra $\mathcal{F}_t$. Then $\{Z_t, \mathcal{F}_t; 0 \le t \le T^{\mathcal{M}}\}$ is a strictly positive martingale under $\mathbf{P}$,[†] and hence it is continuous.

Therefore, there exists a continuous $\mathbf{P}-$local martingale L such that

$$Z(t) = \exp\{L_t - \frac{1}{2}\langle L \rangle_t\} \; \forall t \in [0, T^{\mathcal{M}}],$$

and $\{\tilde{W}_t = W_t - \langle W, L\rangle_t; 0 \le t \le T^{\mathcal{M}}\}$ is a $\tilde{\mathbf{P}}-$Brownian Motion (this is a form of Girsanov's theorem; Revuz and Yor (1999), proposition 8.1.7 and theorem 8.1.12). The Martingale Representation Theorem (Revuz and Yor (1999), theorem 5.3.5) implies the existence of an adapted, locally square integrable process $\{\boldsymbol{\gamma}(t); 0 \le t \le T^{\mathcal{M}}\}$ taking values in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that

$$L(t) = \int_0^t \boldsymbol{\gamma}(s) d\mathbf{W}(s)'.$$

We need only show that $\boldsymbol{\gamma}$ satisfies condition 1. For this, fix $T \in \mathcal{T}^{\mathcal{M}}$ and consider a sequence of stopping times $\{T_k; k \in \mathbb{N}\}$ such that $T_k \uparrow T^{\mathcal{M}}$ a.s., and $\frac{Q(t \wedge T_k, T)}{B(t \wedge T_k)}$ is a $\tilde{\mathbf{P}}-$martingale. From the expression

$$\frac{Q(t,T)}{B(t)} = \int_0^t \frac{1}{B(s)} \boldsymbol{\sigma}(s,T) d\tilde{\mathbf{W}}(s)' + \int_0^t \frac{1}{B(s)} (\alpha(s,T) + \boldsymbol{\sigma}(s,T)\boldsymbol{\gamma}(s)') ds,$$

[†]By equivalence of $\tilde{\mathbf{P}}$ and $\mathbf{P}$, Z is a.s. positive. We may assume Z to be in fact a strictly positive version.

we get $\alpha(t \wedge T_k, T) + \sum_{i=1}^n \sigma_i(t \wedge T_k, T)\gamma_i(t \wedge T_k) = 0$ a.s. for $k \in I\!N$. By allowing $k \to \infty$ in the previous expression, we get the desired conclusion.

Sufficiency. For $t \in [0, T^{\mathcal{M}}]$ define

$$Z(t) = \exp\Big\{\int_0^t \boldsymbol{\gamma}(s) d\mathbf{W}(s)' - \frac{1}{2}\int_0^t \|\boldsymbol{\gamma}(s)\|^2 ds\Big\}.$$

By condition 3, the positive, continuous local martingale $\{Z_t, \mathcal{F}_t; 0 \le t \le T^{\mathcal{M}}\}$ is a martingale under $\mathbf{P}$. Define $\tilde{\mathbf{P}}$ by

$$\tilde{\mathbf{P}}[A] = I\!E[1_A Z_T] \ \forall A \in \mathcal{F}.$$

By Girsanov's Theorem,

$$\{\tilde{\mathbf{W}}(t) = \mathbf{W}(t) - \int_0^t \boldsymbol{\gamma}(s) ds, \mathcal{F}_t; 0 \le t \le T^{\mathcal{M}}\}$$

is an $n-$dimensional Brownian Motion under $\tilde{\mathbf{P}}$. We rewrite (2.4) using $\tilde{\mathbf{W}}$ and Itô's rule, and obtain, for $(t, T) \in [0, T^{\mathcal{M}}] \times \mathcal{T}^{\mathcal{M}}$,

$$\frac{Q(t,T)}{B(t)} = \int_0^t \frac{1}{B(s)} dF(s,T) = \int_0^t \frac{1}{B(s)} \boldsymbol{\sigma}(s,T) d\tilde{\mathbf{W}}(s)'.$$

Therefore, $\frac{Q(t,T)}{B(t)}$ is a local martingale under $\tilde{\mathbf{P}}$; i.e., $\tilde{\mathbf{P}}$ is an ELMM for $\mathcal{M}$. □

Theorem 2.2.1. *Consider a futures model $\mathcal{M}$. If there exists an ELMM for $\mathcal{M}$, then the model does not admit arbitrage.*

Proof. Let $\tilde{\mathbf{P}}$ be an ELMM for $\mathcal{M}$. Consider an arbitrary admissible, self-financing trading strategy $\pi = (\mathcal{R}, \Psi)$. From the proof of Theorem 2.2.1,

$$\frac{V^\pi(t)}{B(t)} = \int_0^t \frac{1}{B(s)} \sum_{i=1}^M \Psi_i(s) \boldsymbol{\sigma}(s, T_i) d\tilde{\mathbf{W}}(s)',$$

where $\tilde{\mathbf{W}}$ is a Brownian Motion under $\tilde{\mathbf{P}}$. By admissibility, and [M1], $\{\frac{V^\pi(t)}{B(t)}; t \in [0, T^{\mathcal{M}}]\}$ is a $\tilde{\mathbf{P}}$-local martingale bounded below, and hence it is a supermartingale. Therefore

$$I\!E\Big[\frac{V^\pi(t)}{B(t)}\Big] \le V^\pi(0) = 0 \ \forall t \in [0, T^{\mathcal{M}}].$$

Therefore, there are no arbitrage opportunities in $\mathcal{M}$. □

Remark 2.2.1. In general arbitrage-free financial models, futures prices are martingales under equivalent martingale measures. In analogy with the HJM framework, we see that the no arbitrage condition in the futures model is necessary for the absence of drift of (2.1) under a risk neutral measure.

A converse of lemma 2.2.1 requires stronger arguments, and different definitions of arbitrage. Different versions of such a converse are simultaneously called the First Fundamental Theorem of Asset Pricing (FFTAP). They usually state that there are no arbitrage opportunities if and only if there is an equivalent (local) martingale measure. Different assumptions and definitions give rise to different statements. Dalang, Morton, and Willinger (1990) proved the FFTAP in the discrete-time case for general stochastic processes, without a requirement of admissibility in trading strategies; they use EMM instead of ELMM. Delbaen and Schachermayer (1994) proved that in the case of a finite number of bounded semimartingales, no arbitrage implies the existence of an EMM. In their setup, an arbitrage is produced by admissible processes that produce bounded terminal payoffs. More generally they proved that in the case of locally bounded semimartingales, the assumption of NFLVR (No Free Lunch With Vanishing Risk) implies the existence of an ELMM. NFLVR is defined as the absence of a sequence of admissible self-financing trading strategies with terminal payoffs converging to a bounded, nonnegative, not identically zero terminal payoff. Finally, Delbaen and Schachermayer (1997) showed that in the case of general semimartingales, there is NFLVR if and only if there exists an equivalent *sigma-martingale* measure. They define a sigma-martingale to be a general martingale transform; i.e., a stochastic martingale integral, where the integrand is required to be predictable and integrable with respect to the martingale.

In the present setup there are potentially infinitely many continuous semimartingales, which widens the possibility of creating arbitrage strategies. The no arbitrage condition defined here is weaker than the NFLVR condition. Therefore, we cannot obtain an exact converse for the FFTAP. However, we give a necessary condition for no arbitrage to hold in the case of an infinite denumerable set of tradable securities. This condition is very close to the condition equivalent to the existence of an ELMM.

Theorem 2.2.2. *Let $\mathcal{M}$ be a futures model which does not admit arbitrage, with $\mathcal{T}^{\mathcal{M}}$ at most an countably infinite. Then there exists an adapted process $\boldsymbol{\gamma} = \{(\gamma_1(t), \dots, \gamma_n(t)), \mathcal{F}_t; 0 \le t \le T^{\mathcal{M}}\}$, such that for $T \in \mathcal{T}^{\mathcal{M}}$,*

$$\alpha(t,T) + \boldsymbol{\sigma}(t,T)\boldsymbol{\gamma}(t)' = 0 \tag{2.5}$$

a.s. on $([0,T^{\mathcal{M}}] \times \Omega, \mathcal{B}([0,T^{\mathcal{M}}]) \times \mathcal{F}, \lambda \times \mathbf{P})$.

We need the following lemma.

Lemma 2.2.2. *Let $\mathcal{M}$ be a futures model with $\mathcal{T}^{\mathcal{M}}$ an infinite denumerable set. Assume that for every collection $\mathcal{R} = (T_1, \dots, T_{n+1})$ of $n+1$ elements of $\mathcal{T}^{\mathcal{M}}$ there exists an adapted process $\gamma = \{(\gamma_1(t), \dots, \gamma_n(t)), \mathcal{F}_t; 0 \le t \le \tau\}$, that satisfies (2.5) for $T \in \mathcal{R}$. Then there exists an adapted process $\gamma = \{(\gamma_1(t), \dots, \gamma_n(t)), \mathcal{F}_t; 0 \le t \le T^{\mathcal{M}}\}$, that satisfies (2.5) for $T \in \mathcal{T}^{\mathcal{M}}$.*

Proof. We assume $n = 1$, the other cases being similar. Assume the conclusion is false.

Case I. There exist $T_0 \in \mathcal{T}^{\mathcal{M}}$ and $A \in \mathcal{B}([0,T_1]) \times \mathcal{F}$ with positive $\lambda \times \mathbf{P}$ measure on which $\sigma(t,T_0) = 0$ and $\alpha(t,T_0) \neq 0$. In this case there is no process γ for which (2.5) is satisfied for T_0.

Case II. Case I does not hold, and there exist $T_1, T_2 \in \mathcal{T}^{\mathcal{M}}$ with $0 \leq T_1 \leq T_2$ and a set $A \in \mathcal{B}([0,T_1]) \times \mathcal{F}$ with positive $\lambda \times \mathbf{P}$ measure on which $\alpha(t,T_2)\sigma(t,T_1) \neq \alpha(t,T_1)\sigma(t,T_2)$. Then there is no process γ for which (2.5) is satisfied for T_1 and T_2.

Case III. Cases I and II do not hold. For $T \in \mathcal{T}^{\mathcal{M}}$ define

$$A_T = \{(t,\omega) \in [0,T^{\mathcal{M}}] \times \Omega : \sigma(t,T) \neq 0\}.$$

By ruling out cases I and II, the following process is well defined on $[0,T^{\mathcal{M}}] \times \Omega$:

$$\gamma(t,\omega) = \begin{cases} -\frac{\alpha(t,T)}{\sigma(t,T)}, & if\ (t,\omega) \in A_T \text{ for some } T \in \mathcal{T}^{\mathcal{M}} \\ 0, & if\ (t,w) \notin \cup_{T \in \mathcal{T}^{\mathcal{M}}} A_T. \end{cases}$$

By denumerability of $\mathcal{T}^{\mathcal{M}}$, γ is adapted, and it satisfies (2.5). This process contradicts our initial assumption. Therefore, this case is not possible.

It follows that case I or II holds. In either case, there exist $T_1, T_2 \in \mathcal{T}^{\mathcal{M}}$ for which (2.5) does not hold. □

Proof of Theorem 2.2.2. Assume there is no process γ which satisfies the conditions of the lemma. By Lemma 2.2.2, there exist $T_1 < T_2 < \ldots < T_{n+1}$ in $\mathcal{T}^{\mathcal{M}}$ for which there is no process that satisfies (2.5) for $T_k,\ k = 1, \ldots, n+1$. Define $\boldsymbol{\rho}_i(s) = (\sigma_i(s,T_1), \ldots \sigma_i(s,T_{n+1}))$, $i = 1, \ldots, n$. Call G the linear subspace in $\mathbb{R}^{n+1}$ generated by the vectors $\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_n$. For $s \in [0,T^{\mathcal{M}}]$ call $\boldsymbol{\alpha}(s) = (\alpha(s,T_1), \ldots \alpha(s,T_{n+1}))$, and define

$$\boldsymbol{\Phi}(s) = \begin{cases} \mathrm{Pr}_{G^\perp} \frac{\boldsymbol{\alpha}(s)}{\|\boldsymbol{\alpha}(t)\|}, & if\ \boldsymbol{\alpha}(t) \neq 0 \\ 0, & if\ \boldsymbol{\alpha}(t) = 0 \end{cases} \tag{2.6}$$

The initial assumption implies the existence of $A \in \mathcal{B}([0,T^{\mathcal{M}}]) \times \mathcal{F}_{T^{\mathcal{M}}}$ with positive $\lambda \times \mathbf{P}$ measure such that $\boldsymbol{\Phi}(s)\boldsymbol{\alpha}(s)' > 0$ on A. Set

$$\Psi_0(t) = -\int_0^t (\boldsymbol{\Phi}(t) - \boldsymbol{\Phi}(s))\frac{\boldsymbol{\alpha}(s)'}{B(s)}ds - \sum_{i=1}^{n+1}\int_0^t (\Phi_i(t) - \boldsymbol{\Phi}_i(s))\frac{\boldsymbol{\sigma}(s,T_i)}{B(s)}d\mathbf{W}(s)'.$$

By construction, $\pi = ((T_1, \ldots, T_{n+1}), \Psi = (\Psi_0, \boldsymbol{\Phi}))$ is an admissible, self-financing trading strategy. In fact, $V^\pi(s) \geq 0\ \forall s \in [0,T^{\mathcal{M}}]$. Now, $\{\omega \in \Omega : \lambda\{t : (t,\omega) \in A\} > 0\}$ has positive $\mathbf{P}-$measure. Therefore, $\mathbf{P}[V^\pi(T^{\mathcal{M}})] > 0$. This is an arbitrage opportunity. □

2.3 Completeness of Futures Models

We call a futures model $\mathcal{M}$ *standard* if it does not admit arbitrage, $\mathcal{T}^{\mathcal{M}}$ is at most infinitely denumerable, and the process $\boldsymbol{\gamma}$ of lemma 2.2.2 satisfies conditions 2 and 3 of theorem 2.2.1. By Theorem 2.2.1, there exists an ELMM for every standard futures model $\mathcal{M}$. We choose arbitrarily one such measure, and denote it by $\tilde{\mathbf{P}}^{\mathcal{M}}$. We denote by $\tilde{\mathbf{W}}^{\mathcal{M}}$ the Brownian Motion under $\tilde{\mathbf{P}}^{\mathcal{M}}$ obtained in the proof of necessity of Theorem 2.2.1.

Definition 5. A standard futures model $\mathcal{M}$ is *Complete* if for every $T \in [0, T^{\mathcal{M}}]$ and $\mathcal{F}_T-$measurable random variable X that is bounded below and satisfies

$$x_0 = \tilde{\mathbb{E}}^{\mathcal{M}}\Big[\frac{X}{B(T)}\Big] < \infty,$$

there is an admissible, self-financing trading strategy $\pi = (\mathcal{R}, \Psi)$ on $\mathcal{M}$ such that $V^{\pi}(0) = x_0$ and $V^{\pi}(T) = X$ a.s.

Theorem 2.3.1. *Let $\mathcal{M}$ be a given standard futures model, with $\mathcal{T}^{\mathcal{M}}$ finite. The following statements are equivalent:*

1. *$\mathcal{M}$ is complete.*
2. *There exists a unique Equivalent Local Martingale Measure for $\mathcal{M}$.*
3. *There exists a unique (in L^2) adapted process $\gamma = \{(\gamma_1(t), \dots, \gamma_n(t)), \mathcal{F}_t; 0 \le t \le T^{\mathcal{M}}\}$ that satisfies conditions 1,2 and 3 of Theorem 2.2.1.*

Proof. The equivalence of 2 and 3 is a consequence of Girsanov's theorem and the proof of Theorem 2.2.1.

$1 \Rightarrow 2$. Assume $\mathbf{P}_1$ is an ELMM for $\mathcal{M}$. Let $A \in \mathcal{F}_{T^{\mathcal{M}}}$ be given. Consider the $F_{T^{\mathcal{M}}}$-measurable random variables $X_1 = 1_A B(T^{\mathcal{M}})$ and $X_2 = 1_{A^c} B(T^{\mathcal{M}})$. For $j = 1, 2$, let $\pi^{(j)} = (\mathcal{R}^{(j)}, \Psi^{(j)})$ be an admissible, self-financing trading strategy for $\mathcal{M}$ that replicates X_j. Then

$$1_A - \tilde{\mathbf{P}}^{\mathcal{M}}[A] = \int_0^{T^{\mathcal{M}}} \frac{1}{B(s)} \sum_{i=1}^{M^{(1)}} \Psi_i^{(1)}(s)\boldsymbol{\sigma}(s, T_i) d\mathbf{W}_1(s)'$$

$$1_{A^c} - \tilde{\mathbf{P}}^{\mathcal{M}}[A^c] = \int_0^{T^{\mathcal{M}}} \frac{1}{B(s)} \sum_{i=1}^{M^{(2)}} \Psi_i^{(2)}(s)\boldsymbol{\sigma}(s, T_i) d\mathbf{W}_1(s)',$$

where $\mathbf{W}_1$ is a Brownian Motion under $\mathbf{P}_1$ (by Theorem 2.2.1). The right hand sides of the previous equations are $\mathbf{P}_1$-local martingales bounded below (by admissibility of $\pi^{(j)}$), and are therefore supermartingales that start at 0. It follows that

$$\mathbb{E}_1[1_A - \tilde{\mathbf{P}}^{\mathcal{M}}[A]] \le 0$$

$$\mathbb{E}_1[1_{A^c} - 1 + \tilde{\mathbf{P}}^{\mathcal{M}}[A]] \le 0.$$

Therefore, $\mathbf{P}_1(A) = \tilde{\mathbf{P}}^{\mathcal{M}}(A)$. Since $A \in \mathcal{F}_{T^{\mathcal{M}}}$ was arbitrary, we conclude that $\mathbf{P}_1 = \tilde{\mathbf{P}}^{\mathcal{M}}$. That is, $\tilde{\mathbf{P}}^{\mathcal{M}}$ is the unique equivalent local martingale measure for $\mathcal{M}$.

$3 \Rightarrow 1$. We write $\mathcal{T}^{\mathcal{M}} = \{\tau_1, \ldots, \tau_J\}$. Let X be an $\mathcal{F}_T$-measurable random variable bounded below, for $T \le T^{\mathcal{M}}$, such that $\tilde{I\!\!E}^{\mathcal{M}}[\frac{X}{B(T)}] = x < \infty$.

Consider the $\tilde{\mathbf{P}}^{\mathcal{M}}$-martingale $N_t = \tilde{I\!\!E}^{\mathcal{M}}[\frac{X}{B(T)}|\mathcal{F}_t]$. This martingale is bounded below. By a martingale representation theorem (Karatzas and Shreve (1998), Lemma 6.7), there exists a square-integrable, adapted process $\mathbf{\Phi}$ such that

$$N_t = x + \int_0^t \mathbf{\Phi}(s) d\tilde{\mathbf{W}}^{\mathcal{M}}(s)',\ t \in [0,T].$$

Since $\mathcal{T}^{\mathcal{M}}$ is finite and $\{\boldsymbol{\sigma}(s,\tau_1), \ldots, \boldsymbol{\sigma}(s,\tau_J)\}$ generates $I\!\!R^n$ a.s., there exists an adapted process $\mathbf{\Psi} = (\Psi_1(s), \ldots, \Psi_J(s))$ such that

$$\sum_{i=1}^{J} \Psi_i(s)\boldsymbol{\sigma}(s,\tau_i) = B(s)\mathbf{\Phi}(s) \text{ a.s.},\ s \in [0,T],\ \tau_i \in \mathcal{T}^{\mathcal{M}}.$$

Set

$$\Psi_0(t) = x - \int_0^t (\mathbf{\Psi}(t) - \mathbf{\Psi}(s))\frac{\boldsymbol{\alpha}(s)'}{B(s)}ds - \sum_{i=1}^{J}\int_0^t (\Psi_i(t) - \Psi_i(s))\frac{\boldsymbol{\sigma}(s,\tau_i)}{B(s)}d\mathbf{W}(s)'.$$

Then $\pi = (\mathcal{T}^{\mathcal{M}}, \mathbf{\Psi})$ is an admissible, self-financing, replicating strategy for X. □

Theorem 2.3.2. *Let $\mathcal{M} = \{F^{\mathcal{M}},\ r^{\mathcal{M}},\ \mathcal{T}^{\mathcal{M}},\ T^{\mathcal{M}}\}$ be a given standard futures model. Then $\mathcal{M}$ is complete if and only if there exists $\mathcal{R} \subseteq \mathcal{T}^{\mathcal{M}}$ such that $\mathcal{R}$ is finite,* $\min \mathcal{R} \ge T^{\mathcal{M}}$*, and $\mathcal{M}^*_{\mathcal{R}} = \{F^{\mathcal{M}},\ r^{\mathcal{M}},\ \mathcal{R},\ T^{\mathcal{M}}\}$ is a complete futures model.*

Proof. We use a contradiction argument to prove the nontrivial direction. Assume that $\mathcal{M}$ is complete, and that for every $\mathcal{R}$ which is a finite subset of $\mathcal{T}^{\mathcal{M}}$ and for which $\min \mathcal{R} > T^{\mathcal{M}}$, $\mathcal{M}^*_{\mathcal{R}}$ is not complete. Then for every such $\mathcal{R}$, there exists $A_{\mathcal{R}} \in \mathcal{B}([0,T^{\mathcal{M}}])$ with positive Lebesgue measure, such that $\{\boldsymbol{\sigma}(t,R) : R \in \mathcal{R}\}$ does not generate $I\!\!R^n$, for $t \in A_{\mathcal{R}}$.

Fix arbitrarily one such $\mathcal{R}_0 = (R_1, \ldots, R_K)$. Let $\{\mathbf{v}(s); 0 \le s \le T^{\mathcal{M}}\}$ be an adapted n-dimensional process such that for $t \in A_{\mathcal{R}_0}$, $\mathbf{v}(t)$ is not in the linear span of $\{\boldsymbol{\sigma}(t,R) : R \in \mathcal{R}\}$. Without loss of generality, assume $\|\mathbf{v}(t)\| = 1$ for $t \le T^{\mathcal{M}}$.

Let f be a smooth function in $\mathcal{R}$ such that f and f'' are bounded below, and f' never vanishes (for example, $f(x) = \arctan(x)$). Consider the processes defined at time t by

$$X(t) = \int_0^t \mathbf{v}(s) d\tilde{\mathbf{W}}(s)$$

$$Y(t) = f(X(t)) - \frac{1}{2}\int_0^t f''(X(s))ds = f(0) + \int_0^t f'(X(s))\mathbf{v}(s)d\tilde{\mathbf{W}}(s).$$

Then the $\mathcal{F}_{T^{\mathcal{M}}}$-measurable variable $Y(T^{\mathcal{M}})$ is bounded below and has finite expected value (since Y is a martingale). By completeness of the model, we may find an admissible, self-financing strategy $\pi = (\Psi, \mathcal{S} = (S_1, \dots, S_m))$ such that $V^\pi(T^{\mathcal{M}}) = Y(T^{\mathcal{M}})$. Therefore, we have $\frac{\boldsymbol{\Psi}(t)\boldsymbol{\sigma}(t)}{B(t)} = f'(X(t))\mathbf{v}(t)$ for $t \le T^{\mathcal{M}}$, where $\boldsymbol{\sigma}(t)$ is an $m \times n$ matrix with i-th row $\boldsymbol{\sigma}(t, S_i)$, and $\boldsymbol{\Psi}(t) = (\Psi_1(t), \dots, \Psi_m(t))$. However, this is impossible for $t \in A_{\mathcal{R}\cup\mathcal{S}}$, by construction of $\mathbf{v}$, and by observing that $A_{\mathcal{R}\cup\mathcal{S}} \subseteq A_{\mathcal{R}}$. The result follows. □

Remark 2.3.1. From the proof of Theorem 2.3.1, a model is complete if and only if all claims with integrable terminal payoff bounded below can be replicated with a *martingale-generating* trading strategy on a fixed finite set of future contracts. If a futures model $\mathcal{M}$ satisfies the conditions of Theorem 2.3.1, then pricing of claims that pay X at date $T < T^{\mathcal{M}}$ is done through arbitrage arguments. The fair price of the claim at date $t < T$ is the price of its replicating portfolio, which is

$$V^\pi(t) = \tilde{\mathbb{E}}[B(t)\frac{X}{B(T)}|\mathcal{F}_t], \tag{2.7}$$

where $\tilde{\mathbb{E}}[\cdot]$ denotes expectation under *the* equivalent local martingale measure.

Remark 2.3.2. A corollary of Theorems 2.3.1 and 2.3.2 is that completeness implies uniqueness of equivalent local martingale measures for standard futures models for which $\mathcal{T}^{\mathcal{M}}$ is countably infinite. For the converse of this corollary, we need uniqueness of the ELMM for a finite subset of $\mathcal{T}^{\mathcal{M}}$. In fact, consider a futures model $\mathcal{M}$ with $n = 1$, $\mathcal{T}^{\mathcal{M}} = (T_1, T_2, \dots)$, and $T^{\mathcal{M}} < T_1$. Assume $r \equiv 0$ and

$$\begin{aligned} &\alpha(t,T) = 0,\ t \le T^{\mathcal{M}},\ T \in \mathcal{T}^{\mathcal{M}} \\ &\sigma(t,T_i) = 1_{[(1-2^{-i+1})T^{\mathcal{M}},(1-2^{-i})T^{\mathcal{M}}]}(t)\ t \in [0,T^{\mathcal{M}}],\ i \in \mathbb{N}. \end{aligned}$$

Then the wealth process of a trading strategy $\pi = (\mathcal{R}, \Psi)$ is

$$V^\pi(t) = B(t)\sum_{T_i \in \mathcal{R}} \int_{(1-2^{-i+1})T^{\mathcal{M}}}^{(1-2^{-i})T^{\mathcal{M}}} \frac{\Psi_i(s)}{B(s)} dW(s).$$

Given the finiteness of assets allowed in trading strategies, it becomes clear that not every terminal payoff is replicable. For example, there is no replicating portfolio for the variable $W(T^{\mathcal{M}})$. Therefore, $\mathcal{M}$ is a standard model for which there is a unique ELMM, but $\mathcal{M}$ is not complete.

2.4 Comments on the Use of Futures Models

Futures models offer a method to value and hedge contingent claims that are adapted to the given filtration, using no arbitrage arguments. For this purpose, we may assume that $\alpha(t,T) = 0$ for $(t,T) \in [0,\tau]^2$. Lemma 2.2.1 and Theorem 2.2.1 guarantee that the resulting model is free of arbitrage. This offers no loss of generality. In fact, from the sensible assumption that a futures model is standard, we establish the existence of an ELMM under which futures prices will be driftless. The assumption on α implicitly states that we originally start from this measure.

Assume given a complete futures model. Consider a claim with integrable payoff X at date T which is bounded below. Its price at date $t < T$ is given by (2.7). Completeness of the model guarantees the existence of a replicating portfolio for X, from where hedging of the claim follows. This is extended to claims that have cash flows prior to expiration, or early exercise features.

If the model is not complete, minimal super-hedging techniques may be used to derive minimal fair values for claims. There are several works in this direction. See for example, El Karoui and Quenez (1995), and Cvitanić and Karatzas (1993).

Under the assumption that futures prices are driftless under $\mathbf{P}$, we need only specify $\boldsymbol{\sigma}$ to determine the stochastic process followed by them. In practice, volatilities are usually derived statistically from market movements, or obtained to fit current market prices of derivatives. The information used in the former is a realized path of the quadratic cross variation of the process. In the latter the information used involves partial information of the distribution of the quadratic (cross) variation process. It is natural to ask whether this is enough information to determine the distribution of the futures prices processes. Historical market movements are taken under the physical measure. Current prices provide information under the risk-neutral measure. Both give information about quadratic (cross) variations. However, quadratic (cross) variations remain unchanged when going from one measure to the other. Therefore, to determine the futures prices processes from market data for pricing and hedging purposes, the measure used to obtain the information is irrelevant.

Motivated by the previous remarks, we pose the following questions. Consider a continuous martingale M. Assume given the distribution of the quadratic (cross)

variation of the martingale, $\langle M \rangle$. Is it possible to uniquely determine the distribution of the martingale?

Now assume that we restrict to the class of one dimensional martingales that satisfy the equation

$$dM_t = g(M_t)dW_t, \quad M_0 = 0,$$

where W is Brownian Motion and g is a Borel measurable function for which weak existence and uniqueness in the sense of probability law hold for the previous equation. Given the distribution of the quadratic variation of one such martingale, is it possible to uniquely determine the distribution of the martingale? In other words, is it possible to find two martingales within this class with different law that produce quadratic variations with the same law?

These questions are addressed in Chapter 4.

Chapter 3

Term Structure Models Based on Future Rates

Futures models are determined by futures prices processes, a spot rate process, a set of tradable maturities and a final trading date. If we isolate the spot rate process, we derive a term structure model from the spot rate perspective. In this chapter we take the alternative approach. We assume that futures contracts settle to spot rates. The effect is that the spot rate process is not independent of the futures prices. If we assume futures models to be consistent, then we derive term-structure models based on the evolution of future rate curves. Theoretically, this viewpoint is equivalent to the HJM framework, in which the evolution of the forward curve is studied. However, in some particular cases, this perspective gives practical advantages. More precisely, for particular examples of volatilities, the state of the model is given by a vector-valued process with equal dimension as the underlying random shocks.

In the first section we give examples of Gaussian and Lognormal futures models. We study the particular case in which volatilities are sums of exponential functions. This case was studied in Heath and Jara (1999), and it is shown that an n-dimensional process contains the information necessary to derive the state of the futures curve, where n is the dimension of the underlying Brownian Motion. In the second section we study futures models where the futures contracts settle to the continuously compounded spot rate. We give examples of widely known spot rate models written in terms of future rate models. We study the case of exponential volatilities, for which we show the feasibility of using PDE methods or recombining trees for pricing fixed income contingent claims. In the third section we exhibit a model for futures prices that settle to 3-month LIBOR. We explore requirements that need to be satisfied by the spot rate process to obtain a consistent futures model. We study the case of exponential volatilities, and propose methods for determining the parameters of the model from historical observations of Eurodollar Future contracts.

3.1 Linear Volatilities for Futures Models

In this section we exhibit special ways of specifying futures models. In particular, we will consider futures prices volatilities that are linear functions of the futures prices. Following the discussion of section 2.4, we assume that drifts are null, and therefore we stand initially in a risk-neutral measure. We also study a particular form of deterministic $\boldsymbol{\sigma}$, where each component is written as a sum of functions that decay exponentially with time to maturity. These types of volatilities were studied in the HJM framework by Cheyette (1993) and by Ritchken and Sankarasubramanian (1993).

A futures model $\mathcal{M}$ will be called *linear* if $\mathcal{T}^{\mathcal{M}}$ contains at least n elements and is at most countable, and the futures curve satisfies

$$d_t F(t,T) = (K_1 F(t,T) + K_2)\boldsymbol{\rho}(t,T)d\mathbf{W}(t)', \tag{3.1}$$

where $\boldsymbol{\rho}$ is a deterministic process that satisfies condition 3 of [M1], and $(K_1, K_2) \in (\{1\} \times I\!R) \cup \{(0,1)\}$.* That is, a linear futures model specifies that $\alpha = 0$ and $\boldsymbol{\sigma}(t,T) = (K_1 F(t,T) + K_2(T))\boldsymbol{\rho}(t,T)$. A linear futures model is standard.

If $K_1 = 0$, future curves are driftless Gaussian processes, with covariance and cross-covariance (for $t \le T_1 \le T_2 \le \tau$, $t_1 \le t_2 \le T \le \tau$)

$$\begin{aligned}\mathrm{Cov}[F(t,T_1), F(t,T_2)] &= \int_0^t \boldsymbol{\rho}(s,T_1)\boldsymbol{\rho}(s,T_2)'ds \\ \mathrm{Cov}[F(t_1,T), F(t_2,T)] &= \int_0^{t_1} \|\boldsymbol{\rho}(s,T)\|^2 ds\end{aligned}$$

If $K_1 = 1$,

$$F(t,T) = (F(0,T) + K_2)\exp\left(\int_0^t \boldsymbol{\rho}(s,T)d\mathbf{W}(s)' - \frac{1}{2}\int_0^t \|\boldsymbol{\rho}(s,T)\|^2 ds\right) - K_2. \tag{3.2}$$

For a linear futures model $\mathcal{M}$ and $t \in [0, T^{\mathcal{M}}]$, let $V_t^{\mathcal{M}}$ be the linear span of $\{\boldsymbol{\rho}(t,T) : T \in \mathcal{T}^{\mathcal{M}}\}$, if $K_1 = 0$, and the linear span of $\{\boldsymbol{\rho}(t,T) : T \in \mathcal{T}^{\mathcal{M}},\ F(0,T) \neq -K_2\}$ if $K_1 = 1$.

Proposition 3.1.1. *A linear futures model $\mathcal{M}$ is complete if and only if there exists $\mathcal{R}$, a finite subset of $\mathcal{T}^{\mathcal{M}}$, for which the futures model $\mathcal{M}^* = \{F^{\mathcal{M}}, r^{\mathcal{M}}, \mathcal{R}, T^{\mathcal{M}}\}$ satisfies $V_t^{\mathcal{M}^*} = I\!R^n$ a.e on $[0, T^{\mathcal{M}}]$.*

Proof. It is an immediate consequence of Theorem 2.3.1, considering that $\alpha = 0$, and

$$\boldsymbol{\sigma}(t,T) = \begin{cases} \boldsymbol{\rho}(t,T), & K_1 = 0 \\ \Big((F(0,T) + K_2)\times & \\ \quad \exp\left(\int_0^t \boldsymbol{\rho}(s,T)d\mathbf{W}(s)' - \frac{1}{2}\int_0^t \|\boldsymbol{\rho}(s,T)\|^2 ds\right)\Big)\boldsymbol{\rho}(t,T), & K_1 = 1. \end{cases}$$

□

*One can specify K_2 to be a function of T without adding considerable changes to the present analysis.

Given a linear futures model, we can obtain different deterministic processes $\hat{\boldsymbol{\rho}}$ that yield the same model, from a distributional perspective.

Proposition 3.1.2. *Let $\boldsymbol{\rho}, \hat{\boldsymbol{\rho}}$ be deterministic processes that satisfy condition 3 of [M1]. Let $F, \hat{F}$ be the respective solutions of* (3.1) *with the same initial futures curve. For $T_1 \leq \ldots \leq T_m \in [0, \tau]$, the m-dimensional processes*

$$\{(F(t,T_1),\ldots,F(t,T_m)); 0 \leq t \leq T_1\} \text{ and } \{(\hat{F}(t,T_1),\ldots,\hat{F}(t,T_m)); 0 \leq t \leq T_1\}$$

have the same law if and only if

$$\int_0^t \boldsymbol{\rho}(s,T_i)\boldsymbol{\rho}(s,T_j)'ds = \int_0^t \hat{\boldsymbol{\rho}}(s,T_i)\hat{\boldsymbol{\rho}}(s,T_j)'ds \ \forall t \in [0,T_1],\ i,j = 1,\ldots,m.$$

Proof. Let $(T_1,\ldots,T_m) \in [0,\tau]^m$ be given, with $T_1 \leq T_2 \leq \ldots \leq T_m$. Then

$$\{(F(t,T_1),\ldots,F(t,T_m)) : 0 \leq t \leq T_1\} \text{ and } \{(\hat{F}(t,T_1),\ldots,\hat{F}(t,T_m)) : 0 \leq t \leq T_1\}$$

have the same law if and only if the m-dimensional Gaussian processes

$$\left\{\Big(\int_0^t \boldsymbol{\rho}(s,T_1)d\mathbf{W}(s)',\ldots,\int_0^t \boldsymbol{\rho}(s,T_m)d\mathbf{W}(s)'\Big); 0 \leq t \leq T_1\right\}$$
$$\left\{\Big(\int_0^t \hat{\boldsymbol{\rho}}(s,T_1)d\mathbf{W}(s)',\ldots,\int_0^t \hat{\boldsymbol{\rho}}(s,T_m)d\mathbf{W}(s)'\Big); 0 \leq t \leq T_1\right\}$$

have the same law. This is equivalent to

$$\int_0^t \boldsymbol{\rho}(s,T_i)\boldsymbol{\rho}(s,T_j)'ds = \int_0^t \hat{\boldsymbol{\rho}}(s,T_i)\hat{\boldsymbol{\rho}}(s,T_j)'ds \ \forall t \in [0,T_1],\ i,j = 1,\ldots,m.$$

□

Remark 3.1.1. The previous proposition suggests defining the equivalence class of a deterministic process $\boldsymbol{\rho}$,

$$[\boldsymbol{\rho}] = \left\{\hat{\boldsymbol{\rho}} : \int_0^t \boldsymbol{\rho}(s,T_1)\boldsymbol{\rho}(s,T_2)'ds = \int_0^t \hat{\boldsymbol{\rho}}(s,T_1)\hat{\boldsymbol{\rho}}(s,T_2)'ds;\ t \leq T_1 \leq T_2 \leq \tau\right\}.$$

We exhibit a method for obtaining representatives of these classes in terms of the Principal Components of a matrix of inner products. For this, let $\boldsymbol{\rho}(t,T) = (\rho_1(t,T),\ldots,\rho_n(t,T))$ be a given deterministic function that satisfies condition 3 of [M1.]. For $t \in [0,\tau]^2$ define the $n \times n$ symmetric and positive semi-definite matrix of "cross covariances"

$$X(t) = \int_t^\tau \boldsymbol{\rho}(t,s)'\boldsymbol{\rho}(t,s)ds.$$

We obtain diagonalizing processes for $t \in [0, \tau]$. That is, we write

$$X(t) = A(t)' D(t) A(t),$$

where $A(t)$ is orthonormal ($A(t)A(t)' = A(t)'A(t) = I_n$), and $D(t) = \text{diag}(\lambda_1(t), \ldots, \lambda_n(t))$ is the diagonal matrix of (positive real) eigenvalues of $X(t)$, where $\lambda_1(t) \geq \ldots \geq \lambda_n(t)$. The rows of $A(t)$, which we denote by $\mathbf{v}_i(t),\ i = 1, \ldots, n$ are the corresponding eigenvectors. Consider the deterministic process $\mathbf{\Psi}$ defined by

$$\mathbf{\Psi}(t, T) = \boldsymbol{\rho}(t, T) A(t)'.$$

For $t, T_1, T_2 \leq \tau$, $\mathbf{\Psi}(t, T_1)\mathbf{\Psi}(t, T_2)' = \boldsymbol{\rho}(t, T_1)\boldsymbol{\rho}(t, T_2)'$. Therefore, $\mathbf{\Psi} \in [\boldsymbol{\rho}]$. In other words, we should be indifferent between establishing a futures model with $\boldsymbol{\rho}$ or with $\mathbf{\Psi}$. The Principal Component interpretation remains valid. First, for $i, j = 1 \ldots, n,\ i \neq j$ and $t \leq \tau$, $\Psi_i(t, \cdot)$ and $\Psi_j(t, \cdot)$ are orthogonal:

$$\begin{aligned}\int_t^\tau \Psi_i(t,s)\Psi_j(t,s)ds &= \int_t^\tau \boldsymbol{\rho}(t,s)\mathbf{v}_i(t)'\boldsymbol{\rho}(t,s)\mathbf{v}_j(t)'ds \\ &= \mathbf{v}_i(t)X(t)\mathbf{v}_j(t)' = \lambda_j(t)\mathbf{v}_i(t)\mathbf{v}_j(t)' = 0.\end{aligned}$$

Second, for $i = 1, \ldots, n$ and $t \leq \tau$, $\Psi_i(t, \cdot)$ maximizes the "residual variance":

$$\int_t^\tau \Psi_i(t,s)^2 ds = \frac{\mathbf{v}_i(t)X(t)\mathbf{v}_i(t)'}{\|\mathbf{v}_i(t)\|^2} = \lambda_i(t) = \max_{\mathbf{a}\perp\mathbf{v}_1(t),\ldots,\mathbf{v}_{i-1}(t)} \frac{\mathbf{a}X(t)\mathbf{a}'}{\|\mathbf{a}\|^2}.$$

Third, for $i = 1, \ldots, n$, $t \leq \tau$ and $T \geq t$, $\Psi_i(t, T)$ is an eigenvector of the "matrix" function in $[t, \tau]^2$ defined by $Y(t, T_1, T_2) = \boldsymbol{\rho}(t, T_1)\boldsymbol{\rho}(t, T_2)'$:

$$\int_t^\tau Y(t,T,s)\Psi_i(t,s)ds = \boldsymbol{\rho}(t,T)X(t)\mathbf{v}_i(t)' = \lambda_i(t)\Psi_i(t,T).$$

This "representative" depends on $\boldsymbol{\rho}$. However, any representative chosen this way satisfies the last three conditions.

Exponential Volatilities We now exhibit a particular form of $\boldsymbol{\rho}$ that leads to the possibility of using low dimensional PDEs to price derivatives for multifactor future models (Heath and Jara (1999)). Consider linear future models where ρ_i $(i = 1, \ldots, n)$ are sums of exponential functions. This allows describing the state of the futures curve in terms of artificial state variables. Volatilities are parametrized as follows:

$$\rho_i(t, T) = \sum_{j=1}^m \beta_{ij} \exp(-\lambda_j(T - t)), \tag{3.3}$$

where $\beta_{ij} \in I\!R,\ i = 1, \ldots, n\ j = 1, \ldots, m$ and $\lambda_j \in I\!R_+,\ j = 1, \ldots, m$. Assume $\lambda_i \neq \lambda_j$ for $i \neq j$. It is convenient to define $\gamma_{jk} = \sum_{i=1}^n \beta_{ij}\beta_{ik}$ for $j, k = 1, \ldots, m$.

Remark 3.1.2. By the Stone-Weierstrass theorem, $\text{span}\{e^{-\lambda x} : \lambda \geq 0\}$ is sup-norm dense in $\mathcal{C}([0, \tau])$. Therefore, we can approximate any deterministic continuous $\boldsymbol{\rho}$ as closely as required, by using as many terms in (3.3) as necessary. This remains true if the same exponents, λ_j, are used in all the factors.

Remark 3.1.3. A representative $\boldsymbol{\Psi}$ of $[\boldsymbol{\rho}]$ constructed in remark 3.1.1 satisfies

$$\Psi_i(t, \cdot) \in \text{ span}\{\rho_1(t, \cdot), \dots, \rho_n(t, \cdot)\}.$$

Therefore, the futures model defined with $\boldsymbol{\Psi}$ instead of $\boldsymbol{\rho}$ is also an exponential model, and moreover, it preserves the exponents (λ_j's).

Define the following O-U processes:

$$Z_j(t) = \int_0^t \sum_{i=1}^n \beta_{ij} \exp(-\lambda_j (t-s)) dW_i(s), \ j = 1, \dots, m. \tag{3.4}$$

Example 1 (Gaussian Models.). In the case of Gaussian Futures models ($K_1 = 0$), (3.1), (3.3) and (3.4) imply

$$\begin{aligned} F(t,T) &= F(0,T) + \int_0^t \sum_{i=1}^n \sum_{j=1}^m \beta_{ij} e^{-\lambda_j (T-s)} dW_i(s) \\ &= F(0,T) + \sum_{j=1}^m e^{-\lambda_j (T-t)} Z_j(t). \end{aligned}$$

Example 2 (Shifted Lognormal Models.). If $K_1 = 1$, (3.2), (3.3) and (3.4) yield

$$F(t,T) = (F(0,T) + K_2) \exp\Big\{ \sum_{j=1}^m e^{-\lambda_j (T-t)} Z_j(t) - \\ \frac{1}{2} \sum_{j=1}^m \sum_{k=1}^m \frac{\gamma_{jk}}{\lambda_j + \lambda_k} (e^{-(\lambda_j+\lambda_k)(T-t)} - e^{-(\lambda_j+\lambda_k)T}) \Big\} - K_2.$$

The last equality holds if and only if $\lambda_j + \lambda_k \neq 0$ for all $j, \ k \leq m$. If for some $j, \ k$, $\lambda_j + \lambda_k = 0$, then the corresponding terms in the double sum must be replaced by $\gamma_{jk} t$.

Remark 3.1.4. The m variables defined in (3.4) determine the state of the futures curve. Therefore, it is enough to keep track of these variables in the case of linear volatilities expressed as sums of exponential functions. Moreover, for practical purposes we may assume $m = n$, without loss of generality.

In fact, if $m > n$, then the number of factors could be increased without increasing the computational complexity of the implementation of the model. In other words, more random shocks could be used to model the evolution of the futures curve, without keeping track of more than m variables (by using the initial set of λ's).

On the other hand, if $m < n$, a rotation of the Brownian Motions could be carried out such that the resulting model would have effectively fewer factors. In fact, consider (3.1) and (3.3) with $m < n$. Consider the $n \times n$ matrix B with entries

$$(B)_{ij} = \begin{cases} \beta_{ij}, & \text{if } j \leq m; \\ 0, & \text{if } j > m. \end{cases}$$

Let A be an orthogonal matrix in $I\!R_{n\times n}$ such that the last $n-m$ columns of BA are 0. Define $L(s) = (e^{-\lambda_1(s)}, \dots, e^{-\lambda_m(s)}, 0, \dots, 0) \in I\!R^n$. Then

$$dF(t,T) = (K_1F(t,T) + K_2)L(T-t)Bd\mathbf{W}(t)' = (K_1F(t,T) + K_2)L(T-t)BAA'd\mathbf{W}(t)'.$$

Now, $A'd\mathbf{W}'$ is an n-dimensional Brownian Motion in $(\Omega, \mathcal{F})$, of which only the first m components are used in the expression for $dF(t,T)$. Therefore, the futures curve is effectively being model-led with m random shocks.

Proposition 4.3.2 generalizes this analysis. See remark 4.3.2.

We define *Exponential Futures Models* to be linear futures models with exponential volatilities given by (3.3), with $m = n$. If $K_1 = 1$, we further impose that there exist at least n different elements in $\{T \in \mathcal{T}^{\mathcal{M}} : T \geq T^{\mathcal{M}},\ F(0,T) \neq -K_2\}$.

Proposition 3.1.3. *An exponential futures model is complete if and only if* $B = (\beta_{ij})_{i,j=1,\dots,n}$ *is non-singular.*

Proof. From Proposition 3.1.1, it is enough to prove that $V_t^{\mathcal{M}^*} = I\!R^n$ a.e on $[0, T^{\mathcal{M}}]$ for a submodel $\mathcal{M}^* = \{F^{\mathcal{M}}, r^{\mathcal{M}}, \mathcal{R}, T^{\mathcal{M}}\}$, with $\mathcal{R}$ finite, if and only if $B = (\beta_{ij})_{i,j=1,\dots,n}$ is invertible.

Choose $T_1 < T_2 < \dots < T_n \in \mathcal{T}^{\mathcal{M}} \cap [T^{\mathcal{M}}, \tau]$ such that $F(0,T_i) \neq -K_2$, $i = 1, \dots, n$, if $K_1 = 1$. For $j,k = 1, \dots, n$ and $t \leq T_1$ define $S_{jk}(t) = \rho_j(t, T_k)$. Consider the matrix $S(t) = (S_{jk}(t))_{j,k=1\dots,n}$. Then

$$\begin{aligned} \exists \boldsymbol{\nu}(t) =& (\nu_1(t), \dots, \nu_n(t)) \neq 0,\ \boldsymbol{\nu}(t)S(t) = 0 \\ &\Leftrightarrow \exists \boldsymbol{\nu}(t) \neq 0\ \sum_{j=1}^{n} e^{-\lambda_j(T_i - t)} (\sum_{k=1}^{n} \nu_k(t)\beta_{kj}) = 0,\ i = 1, \dots, n \\ &\Leftrightarrow \exists \boldsymbol{\nu}(t) \neq 0\ B\boldsymbol{\nu}(t)' = 0. \end{aligned}$$

□

3.2 Spot Rate Futures

In this section we introduce futures models that settle to the spot rate. We compare this framework with the HJM framework and the spot rate viewpoint. Throughout the remainder of the chapter we work with futures models where futures prices are driftless ($\alpha = 0$). We may do this without loss of generality as was discussed in section 2.4.

3.2.1 Pure Discount Bonds, Forward Rates and Future Rates

We introduce zero-coupon bonds denoted by $P(t,T)$, which represent the value at date t of a default-free contract that pays 1 dollar at date T, for $T \in [0,\tau]$ and $t \in [0,T]$. We assume that

$$P(t,T) > 0, \ \ 0 \le t \le T \le \tau,$$
$$\frac{\partial}{\partial T} P(t,T) \text{ exists for every } T \in [0,\tau], t \in [0,T].$$

The forward rate at date t, with maturity T is defined by

$$f(t,T) = -\frac{\partial}{\partial T} \log(P(t,T)). \tag{3.5}$$

The future rate at date t with maturity T, is the futures price of a contract that settles to the spot rate. Consider a futures model $\mathcal{M}$ where futures prices are interpreted as future rates. Assume that the model is complete. Then every claim bounded below can be hedged with a martingale generating strategy (see Section 2.4). Therefore, **P** is the (unique) risk-neutral measure, which is used for pricing purposes. In particular, we obtain

$$P(t,T) = \mathbb{E}[\exp\{-\int_t^T r(s)ds\}|\mathcal{F}_t].$$

With these considerations, we write the relations among the spot rate, future rates, and forward rates. Fix $0 \le t \le T \le \tau$. Then

$$r(t) = F(t,t) = f(t,t)$$
$$F(t,T) = \mathbb{E}[r(T)|\mathcal{F}_t] \tag{3.6}$$
$$f(t,T) = \frac{\mathbb{E}[r(T)\exp\{-\int_t^T r(s)ds\}|\mathcal{F}_t]}{\mathbb{E}[\exp\{-\int_t^T r(s)ds\}|\mathcal{F}_t]}. \tag{3.7}$$

These equations show that from one among the spot rate process, future rates processes or forward rates processes, we may derive the other two. There is a long list of models specified under the spot rate perspective (see the next section for examples of widely known spot rate processes). Heath, Jarrow and Morton (1992), introduced a framework in which they model the evolution of the forward rates in a similar fashion to the futures prices evolution given in [M1]. In our present setup, we model the evolution of future rates through futures models.

As the previous remarks imply, in futures models of future rates, these ([M1]) and the spot rate ([M2]) should not be given independently. If we specify a spot rate process, then the future rates are completely specified. That is, given the spot rate process under **P**, we establish a futures model by (3.6). In fact, by the martingale representation theorem (assuming the future rates process is square integrable) we obtain adapted, locally square-integrable processes $\boldsymbol{\sigma}$ that satisfy

$$F(t,T) = F(0,T) + \int_0^t \boldsymbol{\sigma}(s,T)d\mathbf{W}(s)', \tag{3.8}$$
$$\mathbb{E}[\int_0^\tau \|\boldsymbol{\sigma}(t,T)\|^2 dt] < \infty,\ 0 \leq T \leq \tau.$$

This is a spot rate perspective. In the next section we show through examples the future rates models implied by several commonly used spot rate models.

We are interested in taking the alternative point of view in which the future rates are specified by [M1], and the spot rate process is derived.

Therefore, assume we are given futures curves that satisfy (3.8). Then the spot rate process is given by

$$\begin{aligned} r(t) &= F(t,t) \\ &= F(0,t) + \int_0^t \boldsymbol{\sigma}(u,t)d\mathbf{W}(u)', \end{aligned}$$

and if $F(0,\cdot) \in \mathcal{C}^1((0,\tau))$, $\boldsymbol{\sigma}(s,\cdot) \in \mathcal{C}^1((s,\tau))$, we obtain the differential representation

$$dr(t) = \frac{d}{dt}F(0,t)dt + (\int_0^t \frac{d}{dt}\boldsymbol{\sigma}(u,t)d\mathbf{W}(u)')dt + \boldsymbol{\sigma}(t,t)d\mathbf{W}(t)'.$$

This expression determines a and b of [M2].

Remark 3.2.1. In this analysis we start from the martingale measure to define the model. From a practical perspective, this is not very realistic. It is more sensible to start from the physical measure, and then change to the equivalent martingale measure.

If we start from the future rates, then we may observe their drift and volatility under the physical measure. To change to the EMM, we need only set the drift equal to 0, and the volatilities remain equal. This observation makes useless the knowledge of the drift, but essential the knowledge of the volatilities, for pricing purposes. The situation is similar under the HJM framework.

On the other hand, when starting from a spot rate process, we may observe its drift and volatility under the physical measure. Volatilities remain equal under the EMM, but drifts change. However, from this perspective it is unspecified how the drift changes when viewed under the EMM. Other methods must be used to determine this, such as estimating the market price of risk. This is not an easy task. See, for example, Fournié and Talay (1991), who exhibit the difficulties in accurately estimating the market price of risk for a CIR movement of the spot rate.

Remark 3.2.2 (Convexity Correction.). One advantage brought upon by HJM models is that at every time, knowledge of the forward curve allows construction of the discount curve. That is, if at date t we know the forward curve $f(t,\cdot)$, then the pure discount bonds are obtained by $P(t,T) = e^{-\int_t^T f(t,u)du}$. This allows immediate pricing of future cash flows. In the spot rate framework, pure discount bonds usually need to be priced by extending the numerical analysis to the expiration of the bond. This is also the case for general future rate models. However, in some cases approximations of the future-forward spread (or convexity correction) may give accurate differences between the futures curve and the forward curve, therefore allowing the construction of a forward curve. From (3.7) we derive the future-forward spread

$$F(t,T) - f(t,T) = \mathbb{E}[r(T)|\mathcal{F}_t] - \frac{\mathbb{E}[\frac{r(T)B(t)}{B(T)}|\mathcal{F}_t]}{P(t,T)}.$$

We study the convexity correction of linear future rate models in section 3.2.3.

3.2.2 Examples of Spot Rate Models Written as Future Rates Models

In the following examples we show how to obtain future rate models from spot rate models.

Example 3 (Hull-White Model). For some models, like the Hull-White model, there is an HJM representation that uses deterministic $\boldsymbol{\sigma}$. In Heath and Jara (1999) it is shown that the same $\boldsymbol{\sigma}$ should be used in the corresponding futures models. The Hull-White model gives the following SDE for the evolution of the spot rate under the risk-neutral measure:

$$dr_t = (\alpha_t - \beta_t r_t)dt + \sigma_t dW_t, \tag{3.9}$$

where α, β, σ are integrable (on $[0, \tau]$) deterministic functions. We explicitly solve this equation and get

$$r_t = e^{-\int_0^t \beta_s ds}(r_0 + \int_0^t e^{\int_0^s \beta_u du}\alpha_s ds + \int_0^t e^{\int_0^s \beta_u du}\sigma_s dW_s),$$

which yields

$$\begin{aligned} F(t,T) = \mathbb{E}[r(T)|\mathcal{F}_t] &= e^{-\int_0^T \beta_s ds}(r_0 + \int_0^T e^{\int_0^s \beta_u du}\alpha_s ds + \int_0^t e^{\int_0^s \beta_u du}\sigma_s dW_s) \\ &= F(0,T) + \int_0^t e^{-\int_s^T \beta_u du}\sigma_s dW_s. \end{aligned}$$

The initial futures curve is given in terms of the model parameters:

$$F(0,T) = e^{-\int_0^T \beta_s ds}(r_0 + \int_0^T e^{\int_0^s \beta_u du}\alpha_s ds).$$

Given the initial curve, calibration of the model (of α and β) follows. The SDE for the futures curve is

$$dF(t,T) = e^{-\int_t^T \beta_s ds}\sigma_t dW_t,$$

and the corresponding HJM model has the differential representation

$$df(t,T) = \theta(t,T)dt + e^{-\int_t^T \beta_s ds}\sigma_t dW_t,$$

where θ satisfies the HJM no arbitrage restriction (see Heath, Jarrow and Morton (1992))

$$\theta(t,T) = e^{-\int_t^T \beta_s ds}\sigma(t)\int_0^t e^{-\int_t^s \beta_v dv}\sigma(s)ds.$$

Example 4 (Cox-Ingersoll-Ross Model). The SDE specified by this model for the spot rate under the risk neutral measure is

$$dr_t = (\alpha_t - \beta_t r_t)dt + \sigma_t\sqrt{r_t}dW_t, \tag{3.10}$$

where α, β, σ are integrable (on $[0,\tau]$), strictly positive, deterministic functions. Set $K_t = \int_0^t \beta_s ds$. Then

$$\begin{aligned} F(t,T) = \mathbb{E}[r_T|\mathcal{F}_t] &= r_0 + \int_0^T (\alpha_s - \beta_s \mathbb{E}[r_s|\mathcal{F}_t])ds + \int_0^t \sigma_s\sqrt{r_s}dW_s \\ &= F(t,t) + \int_t^T (\alpha_s - \beta_s F(t,s))ds \end{aligned}$$

Fix $t \geq 0$. For $s \geq t$ define $y(s) = F(t,s)$. Assume that the futures curve is smooth ($y \in C^1(t,\infty)$). Solving the differential equation we find

$$\begin{aligned} y(s) &= e^{-(K_s - K_t)}(y(t) + \int_t^s e(K_v - K_t)\alpha_v dv), \text{ or} \\ F(t,T) &= e^{-(K_T - K_t)}(F(t,t) + \int_t^T e^{(K_v - K_t)}\alpha_v dv). \end{aligned}$$

The differential Futures Model version of (3.10) is

$$\begin{aligned} d_t F(t,T) &= e^{-\int_t^T \beta_s ds}\sigma_t\sqrt{F(t,t)}dW_t, \\ F(0,T) &= e^{-K_T}F(0,0) + \int_t^T e^{-(K_T - K_v)}\alpha_v dv. \end{aligned} \tag{3.11}$$

Therefore, in the CIR model the evolution of the whole futures curve depends only on the level of the spot rate. The initial futures curve implied by the CIR model depends on the parameters α and β. From a futures perspective, given a smooth initial futures curve, and β and σ, we recover $\alpha_T = \frac{dF(0,T)}{dT} + \beta_T F(0,T)$. Now, for a steeply inverted initial futures curve (high short end, low long end), it is possible that the evolution given in (3.11) implies that negative futures rates will exist with positive probability. Since $F(t,T) = \mathbb{E}[r(T)|\mathcal{F}_t]$, then negative spot rates occur with positive probability, and (3.10) and (3.11) would not be sensible evolutions. However, such situations require initial curves that would imply negative values of α, which are excluded in the definition of the CIR model. Therefore, when specifying a futures CIR model, the initial curve cannot be arbitrarily specified. Extra care should be taken to ensure the strict positivity of α.

Example 5 (Single Factor Black-Karasinski Model). This model specifies the logarithm of the spot rate to follow the mean-reverting process given in (3.9):

$$dX_t = (\alpha_t - \beta_t X_t)dt + \sigma_t dW_t$$
$$r_t = \exp(X_t).$$

If we set $K_t = \int_0^t \beta_s ds$, then the spot rate is

$$r_t = \exp(e^{-K_t}[X_0 + \int_0^t e^{K_s}\alpha_s ds + \int_0^t e^{K_s}\sigma_s dW_s]),$$

from which we derive

$$\begin{aligned} F(t,T) &= I\!E[r_T|\mathcal{F}_t] \\ &= \exp(e^{-K_T}[X_0 + \int_0^T e^{K_s}\alpha_s ds + \int_0^t e^{K_s}\sigma_s dW_s])\exp(\frac{1}{2}\int_t^T e^{-2(K_T-K_s)}\sigma_s^2 ds) \\ &= F(0,T)\exp\Big(\int_0^t e^{-(K_T-K_s)}\sigma_s dW_s - \frac{1}{2}\int_t^T e^{-2(K_T-K_s)}\sigma_s^2 ds\Big). \end{aligned}$$

The differential version of this representation is

$$d_t F(t,T) = e^{-\int_t^T \beta_s ds}\sigma_t F(t,T)dW_t.$$

The Black, Derman, Toy model is a special case, in which $\beta_t = \frac{\sigma_t\prime}{\sigma_t}$. In this case, the differential version of the futures model representation is

$$d_t F(t,T) = \sigma_T F(t,T)dW_t.$$

The volatility depends only on the maturity date. This might be considered an undesirable feature of the model (volatilities that depend on the time to maturity are more sensible from a practical perspective).

3.2.3 Linear Future Rate Models

Consider linear future rate models. For Gaussian models we derive simplified expressions for pure discount bonds, given the Gaussian nature of the spot rate process,

$$r(t) = F(0,t) + \int_0^t \boldsymbol{\rho}(s,t)d\mathbf{W}(s)'.$$

In fact, we obtain

$$P(t,T) = I\!E[e^{-\int_t^T r(u)du}|\mathcal{F}_t] = \exp\Big\{-\int_t^T F(t,u)du + \int_t^T\int_t^u\int_t^s \boldsymbol{\rho}(v,u)\boldsymbol{\rho}(v,s)'dvdsdu\Big\}. \tag{3.12}$$

For shifted lognormal models, which satisfy (3.2), the spot rate process is given by

$$r(t) = (F(0,t) + K_2)\exp\left(\int_0^t \boldsymbol{\rho}(s,t)d\mathbf{W}(s)' - \frac{1}{2}\int_0^t \|\boldsymbol{\rho}(s,t)\|^2 ds\right) - K_2. \tag{3.13}$$

If $K_2 = 0$, we get a lognormal spot rate process.

Remark 3.2.3. Hogan and Weintraub (1993) have shown that spot rate processes with lognormal distributions imply negative infinite Eurodollar futures prices (equivalently, infinite future LIBOR). More precisely, they show that a spot rate that satisfies one of the equations

$$\begin{aligned} dr(t) &= \alpha r(t)dt + \sigma r(t)dW(t) \\ d\log r(t) &= \kappa(\theta - \log r(t))dt + \sigma dW(t) \end{aligned}$$

implies $\mathbb{E}[\frac{1}{P(t,T)}] = \infty$, $0 < t < T \leq \tau$. The claim follows by observing that the Eurodollar futures price at date t for settlement at date $T > t$ is

$$E(t,T) = \mathbb{E}[1 - r^L(T)|\mathcal{F}_t] = \mathbb{E}\Big[1 - \frac{1 - P(T,T+\delta)}{\delta P(T,T+\delta)}\Big],$$

where $\delta = 3$ months, and r^L is 3-month LIBOR (see section 3.3.1).

Exponential Volatilities

We now consider exponential future rate models. The Gaussian case yields

$$r(t) = F(t,t) = F(0,t) + \sum_{j=1}^{m} Z_j(t), \tag{3.14}$$

where $Z_j(t)$ is given by (3.4). In the shifted lognormal case, the corresponding spot rate process is given by

$$r(t) = F(t,t) = (F(0,t) + K_2)\exp\Big\{\sum_{j=1}^{m} Z_j(t) - \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} \frac{\gamma_{jk}}{\lambda_j + \lambda_k}(1 - e^{-(\lambda_j+\lambda_k)t})\Big\} - K_2 \tag{3.15}$$

Remark 3.2.4. Assume $F(0,\cdot) \in \mathcal{C}^1((0,\tau))$ and $F(0,t) > 0$. The corresponding spot rate model implied by the lognormal ($K_2 = 0$) exponential future rate model can be seen as a multifactor extension of the Black-Karasinski model. In fact,

$$\begin{aligned} d\ln(r(t)) =&\Big(\frac{\frac{d}{dt}F(0,t)}{F(0,t)} - \int_0^t \sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_j \beta_{ij} e^{-\lambda_j(t-s)} dW_i(s) + \\ &\frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} \gamma_{jk}(1 - e^{-(\lambda_j+\lambda_k)t})\Big)dt + \sum_{i=1}^{n}\sum_{j=1}^{m} \beta_{ij} dW_i(t) \end{aligned}$$

If $n = m = 1$, then

$$d\ln(r(t)) = (a(t) - b(t)\ln r(t))dt + \sigma(t)dW(t),$$

where

$$\begin{aligned} a(t) &= \frac{\frac{d}{dt}F(0,t)}{F(0,t)} + \lambda \ln F(0,t) + \frac{1}{4}\beta^2(1 - e^{-2\lambda t}) \\ b(t) &= \lambda \\ \sigma(t) &= \beta. \end{aligned}$$

This is a one-factor Black-Karasinski model.

Discount Factors for Exponential Models

Following remark 3.2.2, we investigate prices of pure discount bonds for exponential future rate models, for the Gaussian and Lognormal cases. As expected, for the former case we give closed-form formulas, and for the latter case we suggest an approximation.

Gaussian case. In Heath and Jara (1999) it is shown that Gaussian future rates and the corresponding Gaussian forward rates share common volatilities. Therefore, the corresponding Gaussian HJM model satisfies

$$f(t,T) = f(0,T) + \int_0^t \boldsymbol{\rho}(s,T)d\mathbf{W}(s)' + \int_0^t \boldsymbol{\rho}(s,T)\int_s^T \boldsymbol{\rho}(s,u)'duds.$$

We obtain

$$F(t,T) - f(t,T) = \mathbb{E}[f(T,T)|\mathcal{F}_t] - f(t,T) = \int_t^T \boldsymbol{\rho}(s,T)\int_s^T \boldsymbol{\rho}(s,u)'duds.$$

In the case of exponential volatilities, given by (3.3),

$$F(t,T) - f(t,T) = \sum_{j,k=1}^{m} \gamma_{jk}\Big[\frac{1}{\lambda_j(\lambda_j+\lambda_k)} - \frac{e^{-\lambda_j(T-t)}}{\lambda_j\lambda_k} + \frac{e^{-(\lambda_j+\lambda_k)(T-t)}}{\lambda_k(\lambda_j+\lambda_k)}\Big].$$

For these models, (3.12) is written

$$\begin{aligned} P(t,T) = \exp\Big\{ -\int_t^T F(t,u)du + \\ \sum_{j,k=1}^{m} \gamma_{jk}\Big[\frac{T-t}{\lambda_j(\lambda_j+\lambda_k)} - \frac{1-e^{-\lambda_j(T-t)}}{\lambda_j^2\lambda_k} + \frac{1-e^{-(\lambda_j+\lambda_k)(T-t)}}{\lambda_k(\lambda_j+\lambda_k)^2}\Big]\Big\} \end{aligned} \tag{3.16}$$

Lognormal case. Deriving values for discount bonds and forward rates in the lognormal case is a harder task. We propose numerical approximations. Recall

$$P(0,T) = I\!E[\exp\{-\int_0^T r(u)du\}] = I\!E[\exp\{-\int_0^T F(0,u)e^{\phi(u)-\psi(u)}du\}], \tag{3.17}$$

where

$$\phi(t) = \int_0^t \boldsymbol{\rho}(u,t)d\mathbf{W}(u)'$$
$$\psi(t) = \frac{1}{2}\int_0^t \|\boldsymbol{\rho}(u,t)\|^2 du.$$

We suggest approximating one of the exponential functions in (3.17) with a truncated Taylor series. We approximate the outer exponential function, which gives us more tractability of formulas. For $m \in I\!N$, $t \in [0, T^{\mathcal{M}}]$, define $I_m(t) = \frac{1}{m!}(\int_0^t r(s)ds)^m$. By Fubini's theorem and symmetry arguments,

$$I\!E[I_m(t)] = \int_0^t \int_0^{t_m} \dots \int_0^{t_2} I\!E[r(t_1)r(t_2)\dots r(t_m)]dt_1 \dots dt_m. \tag{3.18}$$

From (3.15) we find (for $t_1 \le t_2 \dots \le t_m$)

$$I\!E[r(t_1)r(t_2)\dots r(t_m)] = \Big(\prod_{i=1}^m F(0,t_i)\Big)\exp\Big\{-\frac{1}{2}\sum_{i=1}^m\sum_{j,k=1}^n \gamma_{j,k}(1-e^{-(\lambda_j+\lambda_k)t_i})\Big\} \times$$
$$I\!E\Big[\exp\{\sum_{i=1}^m\sum_{j=1}^n Z_j(t_i)\}\Big].$$

From (3.4) we get the explicit formula

$$I\!E[r(t_1)r(t_2)\dots r(t_m)] = \Big(\prod_{i=1}^m F(0,t_i)\Big)\exp\Big\{\frac{1}{2}\sum_{i=1}^m\sum_{j,k=1}^n \gamma_{j,k}\Big(-1+e^{-(\lambda_j+\lambda_k)t_i} +$$
$$[e^{(\lambda_j+\lambda_k)t_i} - e^{(\lambda_j+\lambda_k)t_{i-1}}][\sum_{p,q=i}^m e^{-\lambda_j t_p - \lambda_k t_q}]\Big)\Big\}. \tag{3.19}$$

Going back to the pure discount bonds, we write an N-order approximation of (3.17),

$$P(0,T) \approx I\!E[\sum_{m=0}^N (-1)^m I_m(T)].$$

Considering (3.18) and (3.19), the previous expression may be written as a sum of iterated integrals of deterministic functions. For example, for $N = 3$, we obtain

$$
\begin{aligned}
P(0,T) \approx & 1 - \int_0^T F(0,t_1)dt_1 + \int_0^T \int_0^{t_2} F(0,t_1)F(0,t_2)\exp\{\sum_{j,k=1}^n \gamma_{jk} e^{-\lambda_k t_2}(e^{\lambda_k t_1} - e^{-\lambda_j t_1})\}dt_1 dt_2 \\
& - \int_0^T \int_0^{t_3} \int_0^{t_2} F(0,t_1)F(0,t_2)F(0,t_3)\exp\Big\{\sum_{j,k=1}^n \gamma_{jk}\Big(e^{-\lambda_k t_2}(e^{\lambda_k t_1} - e^{-\lambda_j t_1}) + \\
& \qquad e^{-\lambda_k t_3}(e^{\lambda_k t_1} - e^{-\lambda_j t_1} + e^{\lambda_k t_2} - e^{-\lambda_j t_2})\Big)\Big\}dt_1 dt_2 dt_3.
\end{aligned}
$$

A similar study is done in Hansen and Jorgensen (1998) in the case of spot rates that follow geometric Brownian Motions with constant coefficients. This allows them to solve the integral in (3.18) in a recursive way. Their numerical examples suggest that for reasonable values of the parameters, and small times (e.g., five years), third order approximations are accurate to four decimal places.

3.2.4 Pricing of Contingent Claims for Exponential Future Rate Models

Let $\mathcal{M}$ be a given complete exponential future rates model. By proposition 3.1.3, this is equivalent to the nonsingularity of $B = (\beta_{jk})_{j,k=1,\dots,n}$. Then every terminal payoff bounded below may be replicated with a martingale-generating, self-financed, admissible trading strategy (see Section 2.4). For now, consider only European-type payoffs. For definiteness, let $T \leq T^{\mathcal{M}}$ be an expiration date, and X a random variable on $(\Omega, \mathcal{F}_T, \mathbf{P})$ bounded below. This variable is interpreted to be the terminal price of a contingent claim with value derived from the futures curve. By the Markovian nature of the term-structure of future rates with respect to the state variables defined in (3.4), X can be expressed as a function of these:

$$X = \mathbb{E}[X|\mathcal{F}_T] = \mathbb{E}[X|Z_1(T),\dots,Z_n(T)] = X(Z_1(T),\dots,Z_n(T)).$$

Call $V(t)$ the no-arbitrage price of this contract at date $t \leq T$. This price depends on time and on the state variables; i.e., $V(t) = V(t, Z_1(t),\dots,Z_n(t))$, and

$$V(T) = X. \tag{3.20}$$

The no-arbitrage price of this contract at date t for the state vector $(z_1,\dots,z_n)$, is

$$
\begin{aligned}
V(t,z_1,\dots,z_n) &= \mathbb{E}[\exp\{-\int_t^T r(u)du\}X|\mathcal{F}_t] \\
&= \mathbb{E}[\exp\{-\int_t^T F(u,u)du\}X|Z_1(t) = z_1,\dots Z_n(t) = z_n].
\end{aligned} \tag{3.21}
$$

The previous equation may be solved numerically by different methods, such as Monte Carlo simulation or recombining trees. The latter method exploits directly the existence of a low-dimensional state vector. This fact also allows for PDE methods to be used. We will focus on this methodology to give specific pricing procedures. From (3.4) we obtain

$$dZ_j(t) = -\lambda_j Z_j(t)dt + \boldsymbol{\beta}_j d\mathbf{W}(t)',\ j = 1, \ldots, n,$$

where $\boldsymbol{\beta}_j = (\beta_{1j}, \ldots, \beta_{nj})$. The Feynmann-Kac Theorem allows us to express (3.21) as the solution to the PDE[†]

$$\frac{\partial V}{\partial t} = rV + \sum_{j=1}^{m} \lambda_j z_j \frac{\partial V}{\partial z_j} - \frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} \gamma_{jk} \frac{\partial^2 V}{\partial z_j \partial z_k}, \tag{3.22}$$

with domain $[0^+, T) \times I\!R^n$, and terminal condition $V(T) = X$. This PDE is well posed; i.e., it has a unique solution which depends continuously on the terminal condition (Gustaffson, Kreiss and Oliger (1995), theorem 7.1.1). The price of the contract is $V(0, 0, \ldots, 0)$. In this equation, $r(t) = F(t,t)$ is a function of the state variables, given by (3.14) or (3.15).

Rotation of Variables

Mixed derivatives in (3.22) are avoided by rotating coordinates.

Notation. The $m \times m$ diagonal matrix with entries $x_1, \ldots, x_m$ will be denoted $\mathrm{diag}(x_1, \ldots, x_m)$.
I_m will denote the $m \times m$ identity matrix.
For a function f, f_z will denote the partial derivative of f with respect to z; i.e., $f_z = \frac{\partial f}{\partial z}$.

Consider the $n \times n$ matrices B and Γ with entries $(\Gamma)_{ij} = \gamma_{ij}$, $(B)_{ij} = \beta_{ij}$ $i, j = 1, \ldots, n$. Then $\Gamma = B'B$. Let A be an orthonormal matrix ($A'A = AA' = I_n$) such that

$$A\Gamma A' = \mathrm{diag}(\mu_1, \ldots, \mu_n),$$

where $\mu_1 \geq \ldots \geq \mu_n \geq 0$ are the (necessarily real) eigenvalues of Γ. The columns of A are the eigenvectors of Γ, the $i-$th column corresponding to the $i-$th eigenvalue.

For $\mathbf{z} = (z_1, \ldots, z_n) \in I\!R^n$, consider the rotated vector $\mathbf{y} = \mathbf{z}A'$. Assume $V(t, \mathbf{z})$ is a solution to (3.22). Define $\tilde{V}$ on $I\!R_+ \times I\!R^n$ as follows:

$$\tilde{V}(t, \mathbf{y}) = V(T - t, \mathbf{y}A) = V(T - t, \mathbf{z}).$$

Then $V(t, \mathbf{z}) = \tilde{V}(T - t, \mathbf{z}A') = \tilde{V}(T - t, \mathbf{y})$. Since V satisfies (3.22), then $\tilde{V}$ satisfies

$$\frac{\partial \tilde{V}}{\partial t}(t, \mathbf{y}) = -r(T - t, \mathbf{y}A)\tilde{V}(t, \mathbf{y}) - \sum_{j=1}^{m} \lambda_j (\mathbf{y}A)_j \Big(\sum_{i=1}^{n} a_{ij} \tilde{V}_{y_i}(t, \mathbf{y})\Big) + \frac{1}{2}\sum_{j=1}^{m} \mu_j \tilde{V}_{y_j y_j}(t, \mathbf{y}) \tag{3.23}$$

on $(0, T^-] \times I\!R^n$, with initial condition $\tilde{V}(0, \mathbf{y}) = X(\mathbf{y}A)$. The value of the contract is $\tilde{V}(T, 0, \ldots, 0) = V(0, 0, \ldots, 0)$.

[†]The Feynmann-Kac Theorem restates the use of Itô's rule for V, together with the no arbitrage condition that V must grow on average at the spot (risk-free) rate.

Artificial Boundary Conditions

In order to solve this equation using implicit finite difference methods, one must define a spatial boundary and boundary conditions. One possibility is to look at a particular contract with payoff X, and determine a region outside of which the contract may be approximated in simple terms (for example, for European options deep in the money, we might require that on a sufficiently distant boundary the option's value be equal to its discounted intrinsic value). However, this requires considering each contract separately. Therefore, we propose a uniform condition at the boundary (to be used for all contracts). This specification is arbitrary, and we follow a heuristic argument to justify it. In the following discussion we propose boundaries and boundary conditions.

The covariances of the Gaussian processes $\mathbf{Z}(t) = (Z_1(t), \ldots, Z_n(t))$ and $\mathbf{Y}(t) = (Y_1(t), \ldots, Y_n(t)) = (Z_1(t), \ldots, Z_n(t))A'$, at date t (see (3.4)) are given by

$$I\!E[Z_j(t)Z_k(t)] = \int_0^t (\sum_{i=1}^n \beta_{ij} e^{-\lambda_j(t-s)})(\sum_{i=1}^n \beta_{ik} e^{-\lambda_k(t-s)})ds = \frac{\gamma_{jk}}{\lambda_j + \lambda_k}(1 - e^{-(\lambda_j+\lambda_k)t})$$

$$I\!E[Y_j(t)Y_k(t)] = I\!E[(\sum_{i=1}^n A_{ij}Z_i(t))(\sum_{i=1}^n A_{ik}Z_i(t))] = \sum_{i=1}^n \sum_{l=1}^n A_{ij}A_{lk} \frac{\gamma_{il}}{\lambda_i + \lambda_l}(1 - e^{-(\lambda_i+\lambda_l)t}).$$

In particular,

$$\mathrm{Var}[Y_j(t)] = \sum_{i=1}^n \sum_{l=1}^n A_{ij}A_{lj} \frac{\gamma_{il}}{\lambda_i + \lambda_l}(1 - e^{-(\lambda_i+\lambda_l)t}). \tag{3.24}$$

$\mathbf{Y}(t)$ is an $n-$dimensional centered Gaussian process. It is natural to specify as boundary a contour of constant density for the distribution of the process at a given time. These are surfaces of ellipsoids centered at the origin, with orientation and eccentricity that depend on the covariance of the process. Given a small number $\alpha > 0$, there is such an ellipse in which $\mathbf{Y}(T)$ will be with probability $1 - \alpha$. In the spirit of the Neyman-Pearson lemma, this ellipse is the smallest region inside which $\mathbf{Y}(T)$ will be with probability $1 - \alpha$. If α is small enough, it seems reasonable to assume that the process will remain within this boundary until date T, reducing the effect of the boundary on the solution of the PDE.

It is also natural to look for rectangular boundaries, which allow smoother implementations of finite difference schemes. Combining both ideas, we consider rectangular boundaries centered at the origin, where each side has a length of $2K$ standard deviations of the variable corresponding to the axis parallel to the side, for a large K. To determine what we mean by large, fix a probability level $\alpha > 0$, and a standard normal random variable U. Choose K such that $\mathbf{P}[|U| \leq K] = 1 - \frac{\alpha}{n}$. Then $\mathbf{P}[Y_j(T)^2 \leq K^2 \mathrm{Var}[Y_j(T)]] \geq 1 - \alpha \; \forall j = 1, \ldots, n$. For $\alpha = 10^{-5}$, the corresponding K is less than 5, for $n = 1, 2, 3$. Therefore, for $K \geq 5$, we define the boundary $\mathcal{B}_K$ to be an $n-$dimensional rectangle with vertices

$$\{(\pm K\sqrt{\mathrm{Var}[Y_1(t)]}, \ldots, \pm K\sqrt{\mathrm{Var}[Y_n(t)]})\}.$$

Call $\mathcal{D}_K$ the interior of $\mathcal{B}_K$.

The issue of boundary conditions is not trivial, since one should define well-posed boundary equations that approximate the original one. One possibility is to impose null convexity towards the outward boundary. That is, at a point $\mathbf{y}$ on a side of $\mathcal{B}_K$ perpendicular to the y_j axis, we require that $\tilde{V}_{y_j y_j}(t, \mathbf{y}) = 0$ for $t \in (0, T)$. This is not a standard boundary condition. Further work should be done towards studying the well-posedness of the resulting equation. Kangro (1997) proves well-posedness of the Black-Scholes equation with these type of boundary conditions, and provides error estimates for them. The equations presented here are more general, but of similar nature.

Numerical Examples We provide examples of numerical solutions to the PDE with the suggested boundary conditions. For $n = 1$, a Crank-Nicolson scheme is implemented. For $n = 2, 3$, Alternating Directions Implicit (ADI) methods are used. To assess the accuracy of this pricing tool, we value contracts for which there are closed-form prices. In the following table we list prices of a pure discount bond and options on pure discount bonds in Exponential Gaussian models and "futures caplet" contracts in Exponential Lognormal models; i.e., future contracts that settle to $\max(0, r(t) - K)$ at date t, for a given strike price K. The closed-form price of a pure discount bond is given by (3.16). The t-time price of a call option with strike price K and expiration T_1 on a pure discount bond that expires at date $T_2 > T_1$ is given by (Musiela and Rutkowski (1998), Proposition 15.1.1)

$$C(t) = P(t, T_2)\mathcal{N}\Big(\frac{\log(\frac{P(t,T_2)}{KP(t,T_1)}) + 0.5\nu(t)^2}{\nu(t)}\Big) - KP(t, T_1)\mathcal{N}\Big(\frac{\log(\frac{P(t,T_2)}{KP(t,T_1)}) - 0.5\nu(t)^2}{\nu(t)}\Big),$$

where

$$\nu(t)^2 = \int_t^{T_1} \Big\| \int_{T_1}^{T_2} \boldsymbol{\sigma}(u, s) ds \Big\|^2 du = \sum_{j,k=1}^{n} \frac{\sum_{i=1}^{n} \beta_{ij}\beta_{ik}}{\lambda_j \lambda_k (\lambda_j + \lambda_k)} \Big[(e^{-\lambda_j (T_2 - t)} - 1)(e^{-\lambda_k (T_1 - t)} - 1) - (e^{-\lambda_j (T_2 - t)} - e^{-\lambda_j (T_1 - t)})(e^{-\lambda_k (T_2 - t)} - e^{-\lambda_k (T_1 - t)}) \Big].$$

Future caplet prices are derived in Heath and Jara (1999), and shown to be

$$V(0, t) = F(0, t)\mathcal{N}\left(\frac{\ln(\frac{F(0,t)}{K}) + \frac{\tilde{\nu}(t)^2}{2}}{\tilde{\nu}(t)}\right) - K\mathcal{N}\left(\frac{\ln(\frac{F(0,t)}{K}) - \frac{\tilde{\nu}(t)^2}{2}}{\tilde{\nu}(t)}\right), \text{ where}$$

$$\nu^2(t) = \sum_{j=1}^{m}\sum_{k=1}^{m} \frac{\sum_{i=1}^{n} \beta_{ij}\beta_{ik}}{\lambda_j + \lambda_k} (1 - e^{-(\lambda_j + \lambda_k)t}).$$

We use a constant initial futures curve of 5%. For the lognormal case we use the estimated parameters of section 3.3.2; for the Gaussian case, we multiply these by 5%.[‡] In the

[‡]These parameters are used just to compare prices, and we ignore for now their origin.

following table, N_t gives the number of gridpoints used in the time discretization, and N_x gives the number of gridpoints used in (each of) the space directions. n denotes the dimension of the Brownian Motion of the model. The numbers shown in table 3.1 are numerically derived prices; the numbers in parentheses are relative errors with respect to the model prices. The strike price for the caplets is 5%. The strike prices, and maturity dates for the bond options are shown in the table.

Contract	$n=1$, $N_t=60$, $N_x=121$	$n=2$, $N_t=50$, $N_x=41$	$n=2$, $N_t=50$, $N_x=61$	$n=3$, $N_t=50$, $N_x=21$	$n=3$, $N_t=50$, $N_x=31$	$n=3$, $N_t=50$, $N_x=39$
PDB, $T=10$	0.616069 (-0.0005%)	0.616207 (-0.0001%)	0.616207 (-0.0001%)	0.615812 (-0.0519%)	0.615766 (-0.0594%)	0.615769 (-0.0589%)
PDBOption, $T_1=1$ $T_2=2$, $K=0.95$	0.00472657 (0.0052%)	0.00495051 (-0.0605%)	0.00495061 (-0.0583%)	0.00537784 (-0.4972%)	0.00538776 (-0.3137%)	0.00538928 (-0.2856%)
PDBOption, $T_1=3$ $T_2=5$, $K=0.90$	0.0157547 (-0.0198%)	0.0143503 (-0.0752%)	0.0143569 (-0.0295%)	0.0148542 (-0.0207%)	0.0148716 (-0.0964%)	0.0148748 (0.1178%)
PDBOption, $T_1=5$ $T_2=10$, $K=0.75$	0.0424312 (0.0008%)	0.0428356 (-0.0250%)	0.0428432 (-0.0073%)	0.0424176 (-0.2429%)	0.424293 (-0.2155%)	0.0424367 (-0.1981%)
Fut. Caplet, $T=1$	0.0045861 (0.0435%)	0.0049469 (0.129%)	0.0049401 (-0.010%)	0.0048080 (-1.03%)	0.0048488 (-0.188%)	0.0048526 (-0.109%)
Fut. Caplet, $T=5$	0.0091522 (0.0477%)	0.0094442 (-0.090%)	0.0094504 (-0.0248%)	0.0097473 (1.09%)	0.0096458 (0.037%)	0.00961787 (-0.253%)
Fut. Caplet, $T=10$	0.011395 (0.0332%)	0.011798 (-0.436%)	.0.011869 (0.164%)	0.012077 (1.07%)	0.011893 (-0.474%)	0.0119426 (-0.0592%)

Table 3.1: Prices and errors of finite difference (ADI) solutions of the PDE.

Remark 3.2.5 (Contracts with Early Exercise). Within this framework we may also consider derivatives with American or Bermudan types of exercise. Preserving the notation introduced earlier, we consider a function X that represents the intrinsic value of the contract and depends on time and the state variables Z. That is, (3.21) becomes

$$V(t,z_1,\dots,z_n)) = \max_{\substack{t\le\tau\le T\\ \tau\{\mathcal{F}_s\}\text{-stopping time}}} \mathbb{E}[\exp\{-\int_t^\tau r(u)du\}X(\tau)|\mathcal{F}_t]$$

$$= \max_{\substack{t\le\tau\le T\\ \tau\{\mathcal{F}_s\}-\text{stopping time}}} \mathbb{E}[\exp\{-\int_t^\tau F(u,u)du\}X(\tau,Z_1(\tau),\dots Z_n(\tau))|Z_1(t),\dots Z_n(t)].$$

Once again, we need only restrict to stopping times that depend on the state variables. If the derivative is Bermudan, then the stopping may only occur at given exercise dates. This is a Free Boundary problem, and the exercise boundary comes as part of the solution to the PDE. At each point in $(0, T^-] \times \mathcal{D}_K$, $\tilde{V}$ satisfies one of the following:

1. $\dfrac{\partial \tilde{V}}{\partial t}(t, \mathbf{y}) = -r(T-t, \mathbf{y}A)\tilde{V}(t, \mathbf{y}) - \sum_{j=1}^{m} \lambda_j(\mathbf{y}A)_j(\sum_{i=1}^{n} a_{ij}\tilde{V}_{y_i}(t, \mathbf{y})) + \frac{1}{2}\sum_{j=1}^{m} \mu_j \tilde{V}_{y_j y_j}(t, \mathbf{y})$,
   $\tilde{V}(t, \mathbf{y}) > X(t, \mathbf{y}A)$
2. $\tilde{V}(t, \mathbf{y}) = X(t, \mathbf{y}A)$

As before, numerical solutions require artificial boundaries and boundary conditions.

3.2.5 Comparing Spot Rate, HJM and Future Rate Models

We give a brief comparison from theoretical and practical perspectives of the spot rate, HJM, and future rate frameworks.

1. The three frameworks are theoretically equivalent in the sense that given either a spot rate process, a forward rate curve process or a future rate curve process, the other two are uniquely determined.

2. The spot rate methodology does not impose restrictions for the spot rate process under the risk-neutral measure. From HJM and Futures perspectives, no arbitrage assumptions impose restrictions on drifts under the risk-neutral measure. Moreover, these drifts depend exclusively on the volatility of the curves (in the futures framework, the drift is 0 under the risk-neutral measure). Therefore, to determine an HJM or futures model under the risk-neutral measure, only the volatility is needed. This is useful to determine a model from historical information. This helps establishing the model under the physical measure. In the HJM and futures framework, changing to the risk-neutral measure is immediate. For the spot rate framework, this requires knowledge of the market price of risk (which will be multidimensional for multifactor models), which is not easily obtained.

3. Immediate determination of discount factors (pure discount bonds) from the present state of rates is desirable, since this simplifies valuing future cash flows. In the HJM setup, the forward curve gives immediately a discount curve. In general, the spot rate and future rate viewpoints require numerical determination of the discount curve. This may be avoided occasionally, as in the case of Gaussian models.

4. Usually, spot rate models are established with low-dimensional Markovian structures for the spot rate process. Recombining trees and PDE methods are available for pricing purposes. Implementation of HJM models requires non-recombining tree methods which become timely expensive very quickly. Some efforts have been made towards expressing forward curves in terms of a finite number of state variables: Cheyette (1993) and Ritchken and Sankarasubramanian (1995) have shown that HJM models with exponential volatilities may be described with a state vector.

In fact, Ritchken and Sankarasubramanian prove for one-factor models, that the forward curve is Markov with respect to two state variables if and only if the forward curve volatility is a product of a function of the spot rate and a deterministic function of time and maturity. However, two-factor models require using four state variables to determine the forward curve. In general, futures models behave like HJM models, because of the presence of the no-arbitrage conditions. Nevertheless, in the case of exponential models, n-factor models are described with n state variables. Alternative pricing methods become available for a multifactor setup. Moreover, volatilities which are proportional to rates are readily available, in contrast with the HJM framework, where proportional volatilities imply that the forward curve explodes with positive probability in finite time.

5. Spot rate models require one number as initial condition (the spot rate). More than one rate is needed for multifactor models. The current yield curve is then used to calibrate the model. HJM and Futures models start from the initial yield curve, and calibrate to other contracts, giving more potential to accurately imply current market prices.

3.3 Futures LIBOR models

We now assume that futures contracts settle to 3-month LIBOR. We denote the value of this rate at date t by $r^L(t)$, and the futures price of 3-month LIBOR by $F^L(t,T)$. In practice, these contracts are traded with fixed maturities. They settle at specific dates in March, June, September and December and extend to about 10 years of maturity. Therefore, a sensible futures model for the evolution of Eurodollar Futures prices has a set of about forty tradable maturities. In this section we build future models for LIBOR, and study conditions for these models to be consistent; i.e., for the spot rate process to agree with the LIBOR process (these two are not independent processes). We propose two methods to determine exponential future models for LIBOR from historical observations. We give a numerical example derived from observations of Eurodollar Future contracts.

3.3.1 Consistent Future LIBOR Models

To establish consistent future models for LIBOR, we study the dependence between the spot rate and 3-month LIBOR. Roughly, if we are given the spot rate process under the equivalent martingale measure, then the 3-month LIBOR process is fully determined (see equation (3.25) below). However, from the latter, the spot rate is only recovered partially. As before, information of the LIBOR process is equivalent to information of the process for future LIBOR. In fact, the following equations establish this equivalence.

$$\begin{aligned} F^L(t,T) &= \mathbb{E}[F^L(T,T)|\mathcal{F}_t] = \mathbb{E}[r^L(T)|\mathcal{F}_t] \\ r^L(t) &= F^L(t,t). \end{aligned}$$

Now, 3-month LIBOR is the simple rate that provides a fair payoff in three months. We simplify the discussion by uniformly approximating 3 months with $\delta = 0.25$ years. Then

$$1 + \delta r^L(t) = \frac{1}{P(t, t+\delta)}. \tag{3.25}$$

The spot rate process produces the processes of pure discount bond prices. Therefore, we may derive the LIBOR process (and hence the futures LIBOR processes) from the spot rate process. The next example gives the evolution for futures LIBOR derived from a particular class of spot rate processes.

Example 6. Let a, b_i, γ_i be deterministic Borel measurable functions from $([0,\tau], \mathcal{B}([0,\tau]))$ to $(I\!R, \mathcal{B}(I\!R))$, $i = 1, \dots, n$. Assume the spot rate process is given by

$$r(t) = r(0) + \int_0^t (a(s) + \mathbf{b}(s)\mathbf{W}(s)')ds + \int_0^t \boldsymbol{\gamma}(s)d\mathbf{W}(s)'. \tag{3.26}$$

Then

$$\int_T^{T+\delta} r(u)du|_{\mathcal{F}_T} \sim \mathcal{N}\Big(\delta r(T) + \delta\int_T^{T+\delta}(T+\delta-s)(a(s) + \mathbf{b}(s)\mathbf{W}(T))ds,$$
$$2\int_T^{T+\delta}\int_T^u (T+\delta-u)(T+\delta-s)\mathbf{b}(u)((s-T)\mathbf{b}(s)' + \boldsymbol{\gamma}(s)')dsdu + \int_T^{T+\delta}\|\boldsymbol{\gamma}(s)\|^2 ds\Big).$$

Using (3.25) and the formula for the MGF of a normally distributed random variable X with mean μ and variance ν, $I\!E[e^{tX}] = e^{t\mu + \frac{1}{2}t^2\nu}$, $t \in I\!R$, we derive a formula for

$$r^L(t) = \frac{1}{\delta}(I\!E[e^{-\int_t^{t+\delta} r(u)du}|\mathcal{F}_t]^{-1} - 1),$$

from which we obtain $F^L(t,T) = I\!E[r^L(T)|\mathcal{F}_t]$. After simplifying, we get

$$dF^L(t,T) = (1 + \delta F^L(t,T))\Big(\boldsymbol{\gamma}(t) + \int_t^T \mathbf{b}(s)ds + \frac{1}{\delta}\int_T^{T+\delta}(T+\delta-s)\mathbf{b}(s)ds\Big)d\mathbf{W}(t)'. \tag{3.27}$$

For the spot rate defined in (3.26), the corresponding futures LIBOR processes have volatilities

$$\boldsymbol{\sigma}(t,T) = (1 + \delta F^L(t,T))\boldsymbol{\rho}(t,T), \text{ where}$$
$$\boldsymbol{\rho}(t,T) = \boldsymbol{\gamma}(t) + \int_t^T \mathbf{b}(s)ds + \frac{1}{\delta}\int_T^{T+\delta}(T+\delta-s)\mathbf{b}(s)ds.$$

Futures LIBOR may be written explicitly as follows:

$$F^L(t,T) = (F^L(0,T) + \frac{1}{\delta})\exp\Big(\int_0^t \boldsymbol{\rho}(s,T)'d\mathbf{W}(s) - \frac{1}{2}\int_0^t \|\boldsymbol{\rho}(s,T)\|^2 ds\Big) - \frac{1}{\delta}.$$

These are linear futures models. (In fact, if we start by assuming (3.27), then a spot rate process given by (3.26) yields a consistent futures model).

The converse direction does not follow completely. If we are given a spot LIBOR process, then any spot rate process that satisfies [M2] and

$$\mathbb{E}[e^{-\int_t^{t+\delta} r(u)du}|\mathcal{F}_t] = \frac{1}{1+\delta r^L(t)}, \ \forall t \in [0,\tau]$$

will be consistent. In the previous example we observed a specific case of a linear future LIBOR model for which we could explicitly exhibit a consistent spot rate process.

Summarizing, for LIBOR future models it is not enough to specify the evolution of future LIBOR; we need additional requirements on the spot rate process. For practical purposes in using the model for pricing, we may approximate the spot rate at date t with 3-month LIBOR.

Remark 3.3.1. The "market model" for forward LIBOR introduced by Brace, Gatarek and Musiela (1997) and the futures LIBOR models introduced here are analogous, the first one being treated from a forward (HJM) perspective, and the second one from a futures perspective. BGM derive no-arbitrage conditions for the forward LIBOR model similar to the HJM conditions. That condition finds an analogue of no drift in the present framework. In their context, the forward LIBOR specification gives partial information for recovering the forward rate process. In fact, by specifying arbitrary volatilities for pure discount bonds on $[0,\delta)$, the corresponding HJM model may be readily derived. Therefore, given a BGM model, one may readily find a consistent spot rate process.

In specific examples, it is possible to find explicitly a spot rate process consistent with the spot LIBOR process. We now give an example of a spot rate process that is consistent with a lognormal linear future LIBOR model. The idea extends to general linear future LIBOR models.

Example 7. Let $\mathcal{M}$ be a linear future LIBOR model with $K_2 = 0$ (i.e., satisfying (3.2)). Assume that $F^L(0,t), \rho_i(t,t) \in \mathcal{C}^1(0,\tau)$, and $\rho_i(s,t) \in \mathcal{C}^1((s,\tau))$, $i = 1, \ldots, n$, and that $F(0,\cdot)$ is strictly positive. Then LIBOR is

$$r^L(t) = F^L(0,t)\exp\left(\int_0^t \boldsymbol{\rho}(s,t)d\mathbf{W}(s)' - \frac{1}{2}\int_0^t \|\boldsymbol{\rho}(s,t)\|^2 ds\right).$$

Using Itô's rule,

$$\begin{aligned} dr^L(t) = r^L(t)\Big[\Big(\frac{\frac{d}{dt}F^L(0,t)}{F^L(0,t)} + \int_0^t \frac{\partial}{\partial t}\boldsymbol{\rho}(s,t)d\mathbf{W}(s)' - \int_0^t \boldsymbol{\rho}(s,t)\frac{\partial}{\partial t}\boldsymbol{\rho}(s,t)'ds\Big)dt \\ + \boldsymbol{\rho}(t,t)d\mathbf{W}(t)'\Big] \end{aligned} \tag{3.28}$$

We aim at deriving a consistent spot rate process, or equivalently, a consistent HJM model. Therefore, we search for an initial differentiable forward curve $f(0,\cdot)$ and a progressively measurable process $\boldsymbol{\sigma}(t,T)$ with values in $\mathbb{R}^n$ for which forward rates

$$f(t,T) = f(0,T) + \int_0^t \boldsymbol{\sigma}(s,T)\boldsymbol{\sigma}^*(s,T)'ds + \int_0^t \boldsymbol{\sigma}(s,T)d\mathbf{W}(s)', \tag{3.29}$$

where $\boldsymbol{\sigma}^*(s,T) = \int_s^T \boldsymbol{\sigma}(s,u)du$, satisfy

$$r^L(t) = \frac{e^{\int_t^{t+\delta} f(t,u)du} - 1}{\delta}. \tag{3.30}$$

From (3.29), (3.30) and Itô's rule, we derive

$$dr^L(t) = \frac{1}{\delta}(1+\delta r^L(t))\Big[(f(t,t+\delta) - f(t,t) + \|\boldsymbol{\sigma}^*(t,t+\delta)\|^2)dt + \boldsymbol{\sigma}^*(t,t+\delta)d\mathbf{W}(t)'\Big].$$

Comparing with (3.28), we obtain conditions for consistency of the model,

$$\boldsymbol{\sigma}^*(t,t+\delta) = \frac{r^L(t)}{r^L(t)+\frac{1}{\delta}}\boldsymbol{\rho}(t,t) \equiv G(t),\ \ 0 \le t \le T^{\mathcal{M}} \tag{3.31}$$

$$\begin{aligned} f(t,t+\delta) - f(t,t) = \frac{r^L(t)}{r^L(t)+\frac{1}{\delta}}\Big(\frac{\frac{d}{dt}F(0,t)}{F(0,t)} + \int_0^t \frac{\partial}{\partial t}\boldsymbol{\rho}(s,t)d\mathbf{W}(s)' - \int_0^t \boldsymbol{\rho}(s,t)\frac{\partial}{\partial t}\boldsymbol{\rho}(s,t)'ds \\ - \frac{r^L(t)\|\boldsymbol{\rho}(t,t)\|^2}{r^L(t)+\frac{1}{\delta}}\Big) \equiv H(t) \end{aligned}$$

Using (3.29), the last equation may be rewritten as

$$\begin{aligned} f(0,t+\delta) - f(0,t) + \int_0^t \frac{1}{2}\frac{\partial}{\partial t}(\|\boldsymbol{\sigma}^*(s,t+\delta)\|^2 - \\ \|\boldsymbol{\sigma}^*(s,t)\|^2)ds + \int_0^t (\boldsymbol{\sigma}(s,t+\delta) - \boldsymbol{\sigma}(s,t))d\mathbf{W}(s)' = H(t) \end{aligned} \tag{3.32}$$

Equations (3.31) and (3.32) are necessary and sufficient conditions on $\boldsymbol{\sigma}$ for consistency of the model. Assume $\boldsymbol{\sigma}(s,\cdot)$ is differentiable, and $\boldsymbol{\rho}(s,\cdot)$, $\boldsymbol{\rho}(\cdot,\cdot)$ and $F(0,\cdot)$ are twice differentiable. Using Itô's rule, we get

$$dH(t) = \alpha(t)dt + \boldsymbol{\gamma}(t)d\mathbf{W}(t)'$$

for processes α and $\boldsymbol{\gamma}$ which may be expressed in terms of $\boldsymbol{\rho}$ and r^L. Assume for now that

$$\frac{\partial^2}{\partial t^2}(\|\boldsymbol{\sigma}^*(s,t+\delta)\|^2 - \|\boldsymbol{\sigma}^*(s,t)\|^2) = 0. \tag{3.33}$$

From (3.32), we get

$$\begin{aligned} \frac{\partial}{\partial t}(f(0,t+\delta) - f(0,t)) + \frac{1}{2}\frac{\partial}{\partial t}\|\boldsymbol{\sigma}^*(s,t+\delta)\|^2 + \\ \int_0^t \frac{\partial}{\partial t}(\boldsymbol{\sigma}(s,t+\delta) - \boldsymbol{\sigma}(s,t))d\mathbf{W}(s)' = \alpha(t) \end{aligned} \tag{3.34}$$

$$\boldsymbol{\sigma}(t,t+\delta) - \boldsymbol{\sigma}(t,t) = \boldsymbol{\gamma}(t) \tag{3.35}$$

Assume $\boldsymbol{\nu} = (\nu_1, \dots, \nu_n)$ satisfies[§]

$$\nu_i(t) = 0, \;\; d\nu_i(t) = d\Big(\frac{\alpha(t)}{n} - \frac{1}{n}\sum_{i=1}^{n} 2(\gamma_i(t)G_i(t) + \frac{G_i(t)}{\delta})\Big) - \frac{6\delta\nu_i(t)}{G_i(t)}dW_i(t), \;\; i = 1, \dots, n.$$

We take any differentiable initial forward curve $f(0,\cdot)$. For $i = 1, \dots, n$, $t \in [0, T^{\mathcal{M}}]$, define

$$\begin{aligned} a_i(t) &= \frac{\nu_i(t)}{G_i(t)} \\ b_i(t) &= \frac{\gamma_i(t)}{\delta} - 3a_i(t)\delta \\ c_i(t) &= \frac{G_i(t)}{\delta} - \gamma_i(t) - 2a_i(t)\delta^2. \end{aligned}$$

For $t \in [0, T^{\mathcal{M}}]$, $i = 1 \dots, n$, define periodic functions g_i on $\mathbb{R}_+$ (period δ), such that

$$g_i(x) = 2[(a_i(t)x^3 + b_i(t)x^2 + c_i(t)x)(6a_i(t)x + 2b_i(t)) + (3a_i(t)x^2 + 2b_i(t)x + c_i(t))^2]$$

on $[0, \delta)$. For $0 \le s \le t \le T^{\mathcal{M}}$, define

$$\sigma_i^*(s,t) = \sqrt{\int_s^t \int_s^u g_i(x)dxdu.}$$

Then for $t - s \in [0, \delta]$ we have

$$\begin{aligned} \sigma_i^*(s,t) &= a_i(s)(t-s)^3 + b_i(s)(t-s)^2 + c_i(s)(t-s) \\ \sigma_i(s,t) &= 3a_i(s)(t-s)^2 + 2b_i(s)(t-s) + c_i(s). \end{aligned}$$

It is straightforward to check that by construction, $\boldsymbol{\sigma}(t,T) = (\sigma_1(t,T), \dots, \sigma_n(t,T))$ and $f(0,t)$ satisfy (3.31), (3.33), (3.34), and (3.35), and therefore produce a consistent lognormal futures model.

3.3.2 Estimation of Parameters for Exponential Future LIBOR Models

Consider the class $\mathcal{M}$ of exponential future models with n factors for 3-month LIBOR. We represent this set with

$$\mathcal{B} = M(\mathbb{R})_{n\times n} \times \mathbb{R}_+^n.$$

An element of $\mathcal{M}$ is an $n \times n$ real matrix, which we interpret as the matrix $B = (\beta_{ij})_{i,j=1,\dots,n}$, and an n-dimensional positive vector, to be interpreted as $(\lambda_1, \dots, \lambda_n)$. Our goal is to estimate the element(s) $\mathcal{P}_0$ in $\mathcal{B}$ that imply a futures model that best agrees

[§]Here we boldly assume that this equation has a solution.

with given observations of future LIBOR (for example, from Eurodollar Future contracts). Many elements of $\mathcal{B}$ imply the same futures model, in a distributional sense (Proposition 3.1.2. Therefore, we restrict our attention to "Principal Component" representatives of classes of parameters that yield the same model (see remark 3.1.1). We propose two estimation methods; an MLE technique and a minimization of a squared error. We show results in the latter case for 2 years of data with daily observations. For definiteness, we assume the exponential futures model is lognormal. However, the following analysis is readily extended to more general linear future models.

Statement of the Statistical Problem.

The quadratic (cross) variation of a futures model is the same under the physical and the risk-neutral measures. Observations are done under the physical measure and pricing under the risk-neutral measure. Our aim is to estimate the quadratic (cross) variation of the evolution of future LIBOR. From this, we determine the model under the risk-neutral measure for pricing and hedging purposes.

For lognormal exponential models we have the following differential representation of the quadratic (cross) variation for $t \leq T_1 \leq T_2$:

$$dF(t,T_1)dF(t,T_2) = V(T_1 - t, T_2 - t)F(t,T_1)F(t,T_2)dt,$$

where

$$V(\tau_1, \tau_2) = \sum_{k=1}^{n} \left(\sum_{l=1}^{n} \beta_{k,l} \exp(-\lambda_l \tau_1) \right) \left(\sum_{l=1}^{n} \beta_{k,l} \exp(-\lambda_l \tau_2) \right). \tag{3.36}$$

Based on this expression, we will define normally distributed processes which will be used to determine the estimation procedures.

Fix Δt, which is assumed to be small compared to one year. For $0 \leq t \leq T - \Delta t \leq \tau - \Delta t$, define the $\mathcal{F}_{t+\Delta t}$-measurable

$$Y(t,T) = \frac{F(t+\Delta t, T) - F(t,T)}{F(t,T)\sqrt{\Delta t}}. \tag{3.37}$$

Under the physical measure $\hat{\mathbf{P}}$, where future rates have drifts, we obtain

$$Y(t,T) = \frac{\int_t^{t+\Delta t} F(s,T)\boldsymbol{\rho}(s,T)d\hat{\mathbf{W}}(s)' + \int_t^{t+\Delta t} \alpha(s,T)ds}{F(t,T)\sqrt{\Delta t}}.$$

$Y(t,T)$ is a consistent discrete approximation of the normalized jumps of the futures rates on a period of time of length Δt (see Kloeden and Platen (1995)):

$$Y(t,T) = \boldsymbol{\rho}(t,T)\frac{\Delta \mathbf{W}(t)'}{\sqrt{\Delta t}} + O(\sqrt{\Delta t}), \tag{3.38}$$

where $\Delta\hat{\mathbf{W}}(t) = \hat{\mathbf{W}}(t+\Delta t) - \hat{\mathbf{W}}(t)$. Within order of $\sqrt{\Delta t}$, $Y(t_1, T_1)$ is uncorrelated with $Y(t_2, T_2)$ if $t_2 - t_1 \geq \Delta t$. In fact we have the following equations for $0 \leq t \leq T_1 \leq T_2$, $t_1 < t_2 - \Delta t$, where V satisfies (3.36).

$$\begin{aligned}\mathbb{E}[Y(t,T_1)Y(t,T_2)] &= V(T_1 - t, T_2 - t) + O(\sqrt{\Delta t}) \\ \mathbb{E}[Y(t_1,T_1)Y(t_2,T_2)] &= O(\sqrt{\Delta t})\end{aligned}$$

We now define the observable data. Call $E(t,T)$ the price at date t of the Eurodollar Futures contract that settles to the spot rate at date T. Then the futures LIBOR is $F^L(t,T) = 1 - \frac{E(t,T)}{100}$. Assume information is available at trading dates t_i, $i = 1, \dots, M$. Take T_i to be the set of relative maturities observed at trading date t_i. Then $\tau \in T_i$ if and only if there is a futures contract traded at date t_i that expires at date $t_i + \tau$. Let $\{\tau_j\}_{j=1}^N$ be an enumeration of $\bigcup_{i=1}^M T_i$. Define $J_i = \{j \in \{1, \dots, N\} : \tau_j \in T_i\}$. For every $i = 1, \dots, M$ such that $t_i + \Delta t$ is a trading date, and every $j \in J_i$, define

$$X_i^j = \frac{\Delta F^L(t_i, t_i + \tau_j)}{F^L(t_i, t_i + \tau_j)\sqrt{\Delta t}}, \tag{3.39}$$

where

$$\Delta F^L(t_i, t_i + \tau_j) = F^L(t_i + \Delta t, t_i + \tau_j) - F^L(t_i, t_i + \tau_j).$$

That is, $(X_i^j)_{j \in J_i}$ is an observation of $(Y(t_i, t_i + \tau_j))_{j \in J_i}$.

There are four Eurodollar future contracts that mature in a given year, and the maturities extend as far as ten years. Then there are approximately 40 observable relative maturities. Therefore, we will not be able to observe directly the covariation of all different pairs of relative maturities. Consider the "incomplete" matrix X with entries given by (3.39):

$$(X)_{ij} = X_i^j, \quad i = 1, \dots, M, \ j \in J_i.$$

$(X)_{ij}$ is defined only if the change in the rate for relative maturity τ_j is observed at date t_i; this happens if and only if relative maturity τ_j is observed at date t_i, and relative maturity $\tau_j - \Delta t$ is observed at date $t_i + \Delta t$. An "incomplete" covariance structure is obtained from X in the following way: for $j, k \in \{1, \dots, N\}$ define C_{jk} as the trading dates shared by the relative maturities τ_j and τ_k. Call N_{jk} the number of elements in C_{jk}. Mathematically,

$$\begin{aligned}C_{jk} &= \{i \in \{1, \dots, M\} : j, k \in J_i\} \\ N_{jk} &= |C_{jk}|.\end{aligned} \tag{3.40}$$

Then, for $(j,k) \in \{(l,m) \in \{1, \dots, N\}^2 : N_{l,m} > 0\}$, we define

$$\gamma_{jk} = \frac{1}{N_{jk}} \sum_{i \in C_{jk}} X_i^j X_i^k \tag{3.41}$$

This defines an asymptotically unbiased estimate of the covariances. In fact,

$$
\begin{aligned}
\mathbb{E}[\gamma_{jk}] &= \frac{1}{N_{jk}} \sum_{i\in C_{jk}} \mathbb{E}[X_i^j X_i^k] \\
&= \frac{1}{N_{jk}} \sum_{i\in C_{jk}} \mathbb{E}[Y(t_i, t_i+\tau_j)Y(t_i, t_i+\tau_k)] \\
&= V_{jk} + O(\sqrt{\Delta t}) \\
\mathbb{E}[(\gamma_{jk} - V_{jk})^2] &= \frac{\mathrm{Var}[X_i^j X_i^k]}{N_{jk}} + O(\sqrt{\Delta t}) \\
&= \frac{2V_{jk}^2 + V_{jj}V_{kk}}{N_{jk}} + O(\sqrt{\Delta t}),
\end{aligned}
$$

where $V_{ij} = V(\tau_i, \tau_j)$.

Minimization of a Squared Error.

Different volatilities may imply the same variance-covariance structure. In such case, the corresponding future processes are identically distributed. In particular, securities prices predicted by each one coincide. We therefore want to find parameters for the model that imply a theoretical covariance structure that mimics the estimated counterpart as closely as possible. We use a weighted least squared error method.¶

The following function measures the distance between the theoretical and historical variance-covariance structures. It takes as inputs the $n(n+1)$ parameters that define the volatilities (3.3). Define

$$
\begin{aligned}
h(\{\beta_{ij}\}_{i,j}, \{\lambda_j\}_j) &= \sum_{i=1}^{N}\sum_{j=1}^{N} N_{ij}\,(V_{ij} - \gamma_{ij})^2 \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} N_{ij} \left(\sum_{k=1} \sigma_k(\tau_i)\sigma_k(\tau_j) - \gamma_{ij} \right)^2
\end{aligned}
$$

If γ_{ij} is not defined, then $N_{ij} = 0$, and the corresponding term is 0. We aim at solving

$$
\min_{\{\beta_{ij}\}_{i,j}, \{\lambda_j\}_j} h(\{\beta_{ij}\}_{i,j}, \{\lambda_j\}_j)
$$

More explicitly, we need to solve the following problem:

$$
\min_{\{\beta_{ij}\}_{i,j}, \{\lambda_j\}_j} \sum_{i=1}^{N}\sum_{j=1}^{N} N_{ij} \left(\sum_{k=1}^{n} \Big(\sum_{l=1}^{n} \beta_{kl} \exp(-\lambda_l \tau_i) \Big) \Big(\sum_{l=1}^{n} \beta_{kl} \exp(-\lambda_l \tau_j) \Big) - \gamma_{ij} \right)^2 ,
$$

¶Through this approach we avoid the problem of filling the covariance matrix by interpolation or other methods. This matrix should be positive semi-definite, and this may be violated when estimating the missing data.

where N_{ij} and γ_{ij} are given by (3.40) and (3.41).

Remark 3.3.2. An alternative proxy for the error can be obtained by evaluating the squared error between the covariance structure implied by the estimated parameters, and a smoothed historical covariance matrix. Explicitly, we write

$$\tilde{\gamma}_{ij} = \sum_{k,l} M_{kl}^{ij} \gamma_{kl},$$

where $\sum_{k,l} M_{kl}^{ij} = 1$,[‖] and approximate the error with

$$\sum_{i=1}^{N} \sum_{j=1}^{N} N_{ij} \left(V_{ij} - \tilde{\gamma}_{ij} \right)^2 \tag{3.42}$$

This may help to evaluate the improvement achieved by using more factors.

Numerical Estimation. We now give a numerical example of the MSE procedure. We take daily observations of Eurodollar future contracts for the years 1994 and 1995. That is, we fix $\Delta t = 1$ day. To avoid the inclusion of noise, weekly observations might also be a sensible choice.

A Nelder-Mead algorithm was used to find the optimal parameters. From the comments in remark 3.1.1, the estimated parameters are given in terms of the principal components of the covariance matrix of the exponential functions. That is, the principal components may be taken as the representative of the class of volatilities that imply the same variance covariance structure on $(0, T^-] \times I\!R^n$, $\tilde{V}$. The results are shown in tables 3.2, 3.3, 3.4, 3.5. Figure 3.1 shows the graph of the volatilities implied by these parameters, and the corresponding covariance structures. The proxy error shown corresponds to $\sigma_x = 40, \sigma_y = 20$ (see remark 3.3.2).

Maximum Likelihood Estimates.

We show an alternative way of estimating the model parameters. For every element $\boldsymbol{B}$ of $\mathcal{B}$ we may infer the approximate distribution (discarding the $O(\sqrt{\Delta t})$ terms) of the vector $\mathbf{X}_i = (X_i^j)_{j \in J_i}$. Let n_i be the size of J_i. From (3.38) we define a linear model for X_i^j:

$$X_i^j \sim \sigma_1(t_i, t_i + \tau_j) Z_1 + \ldots + \sigma_n(t_i, t_i + \tau_j) Z_n + \epsilon_{ij},$$

where $(Z_1, \ldots, Z_n)$ is a standard n-dimensional Gaussian vector (mean 0, covariance I_n), and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_0^2)$ are independent, normally distributed random errors. For the density

[‖]For example, we can take a Gaussian kernel: $M_{kl}^{ij} \propto \exp\{-\frac{1}{2}(\frac{(\tau_i - \tau_k)^2}{\sigma_x^2} + \frac{(\tau_j - \tau_l)^2}{\sigma_y^2})\}$. σ_x and σ_y should be chosen to smooth the covariance structure, while retaining its general structure.

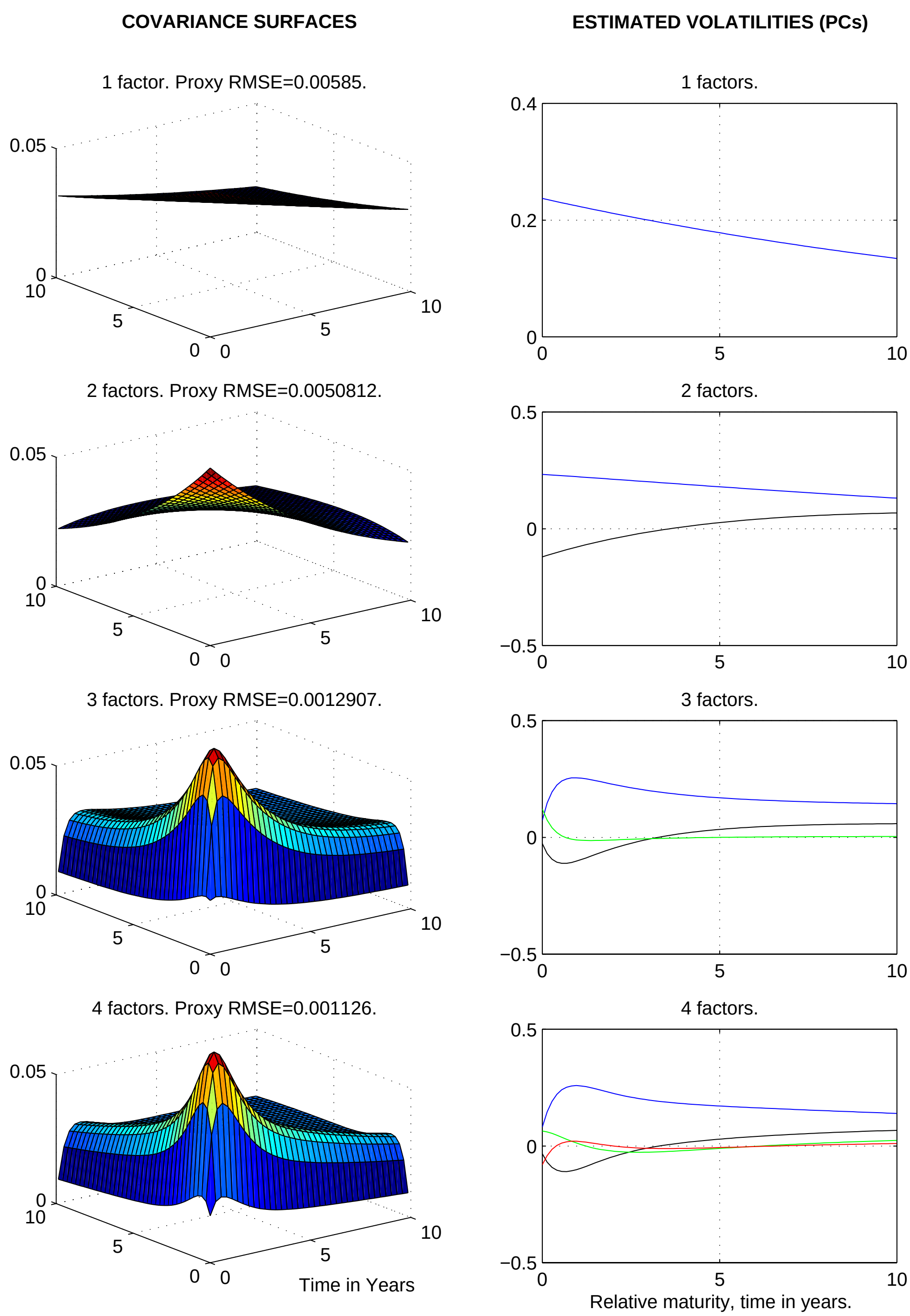

Figure 3.1: Covariance Surfaces and Estimated Volatilities (Principal Components) for one, two, three and four factors.

β	0.2371
λ	0.05691

Table 3.2: Estimated parameters for a one factor exponential lognormal future LIBOR model.

i	1	1	2	2
j	1	2	1	2
β_{ij}	-0.1301	0.3641	-0.5623	0.4438
λ_j	0.1498	0.08255	0.1498	0.08255

Table 3.3: Estimated parameters for a two factor exponential lognormal future LIBOR model.

i	1	1	1	2	2	2	3	3	3
j	1	2	3	1	2	3	1	2	3
β_{ij}	-0.2607	0.1762	0.1574	0.1791	-0.2752	0.0690	0.1561	-0.03943	0.004606
λ_j	2.872	0.4356	0.01042	2.872	0.4356	0.01042	2.872	0.4356	0.01042

Table 3.4: Estimated parameters for a three factor exponential lognormal future LIBOR model.

i	1	1	1	1	2	2	2	2
j	1	2	3	4	1	2	3	4
β_{ij}	-0.9510	0.8245	0.2064	0.001074	0.7208	-0.6975	0.2167	-0.2741
λ_j	1.702	1.265	0.03958	0.1246	1.702	1.265	0.03958	0.1246
i	3	3	3	3	4	4	4	4
j	1	2	3	4	1	2	3	4
β_{ij}	-0.2263	0.3833	0.1161	-0.1954	-0.7236	0.7206	0.08591	-0.1494
λ_j	1.702	1.265	0.03958	0.1246	1.702	1.265	0.03958	0.1246

Table 3.5: Estimated parameters for a four factor exponential lognormal future LIBOR model.

f of this distribution we find

$$f(\mathbf{X}_i|\boldsymbol{B},\sigma_0) = (2\pi)^{-\frac{n_i}{2}}(\det\Sigma_i)^{-\frac{1}{2}}\exp\{-\frac{1}{2}\mathbf{X}_i\Sigma_i^{-1}\mathbf{X}_i'\}, \text{where}$$
$$(\Sigma_i)_{jk} = V(\tau_j,\tau_k) + \sigma_0^2\delta_{j,k}.$$

The likelihood of N of these observations is

$$L(B,\sigma_0) = \prod_{i-1}^{N} f(\mathbf{X}_i|\boldsymbol{B},\sigma_0).$$

To get the maximum likelihood estimate, we find $\boldsymbol{B}, \sigma_0$ that maximize the previous expression. Equivalently, we need to solve

$$\min_{\boldsymbol{B},\sigma_0} \sum_{i=1}^{N} \log(\det \Sigma_i) + \sum_{i=1}^{N} \mathbf{X}_i \Sigma_i^{-1} \mathbf{X}_i.$$

Chapter 4

Martingales Characterized by the Distribution of the Quadratic Variation

In this chapter we study the possibility of extending Lévy's theorem of characterization of Brownian Motion to a wider class of martingales. The financial motivation for studying this question arises from the implementation of models such as HJM, or Futures models, where it is assumed that to determine a model, we need only know the volatility structure. In practical terms, we are given information about the quadratic (cross) variation of a process. However, we do not know in general that the distribution of the quadratic variation of a martingale determines uniquely the distribution of the martingale. The main objective of this chapter is to prove the following two results:

We call a continuous local martingale *divergent* if its quadratic variation diverges almost surely; i.e., if $\langle M\rangle_\infty = \infty$ a.s.

Theorem 1. *Let M be a divergent continuous local martingale with a.s. absolutely continuous quadratic variation such that $\frac{d\langle M\rangle_t}{dt} > 0$ a.s. Assume that for every divergent continuous local martingale N, $\{\langle M\rangle_t\ t \geq 0\} \stackrel{d}{=} \{\langle N\rangle_t; t \geq 0\}$, implies $\{M_t; t \geq 0\} \stackrel{d}{=} \{N_t; t \geq 0\}$. Then M is a Gaussian martingale.*

For a real measurable function f, denote by $Z(f)$ the set of zeroes of f, and by $I(f)$ the set of points where the algebraic inverse of f is not locally square integrable.

Theorem 2. *Let g_1, g_2 be Borel measurable functions on $(I\!R, \mathcal{B}(I\!R))$ such that $I(g_i) = Z(g_i)$, $i = 1, 2$. Consider the stochastic differential equations*

$$X_0 = 0, \quad dX_t = g_1(X_t)dW_t \tag{4.1}$$

$$X_0 = 0, \quad dX_t = g_2(X_t)dW_t. \tag{4.2}$$

Let

$$(X^{(i)}, W^{(i)}),\ (\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbf{P}^{(i)}),\ \{\mathcal{F}_t^{(i)}\},\ (i = 1, 2)$$

be weak solutions of (4.1) and (4.2), respectively. Assume $\{\langle X^{(1)}\rangle_t; t \geq 0\} \stackrel{d}{=} \{\langle X^{(2)}\rangle_t; t \geq 0\}$. *Then* $X^{(1)} \stackrel{d}{=} X^{(2)}$, *or* $X^{(1)} \stackrel{d}{=} -X^{(2)}$.

In the first section we give further motivations for studying these problems. In the second section we prove Theorem 1. We also include a brief discussion of properties of some interesting classes of martingales; in particular, we study Ocone martingales, which are used in the proof of the theorem. In the third section we show the irrelevance of the sign of the volatility when looking for weak solutions of Stochastic Differential Equations. This result will be used in subsequent sections. We extend this result to the multidimensional case (Proposition 4.3.2). This will not be used in the proofs of Theorems 1 and 2, but it will provide interesting financial applications of its own. In the fourth section we characterize one and two-dimensional Borel sets that yield the same finite-dimensional probabilities for Stopped Brownian Motion. That is, we give conditions under which two Borel sets A, B in $\mathbb{R}^n$ satisfy $\mathbf{P}[(V(t_1), \ldots, V(t_n)) \in A] = \mathbf{P}[(V(t_1), \ldots, V(t_n)) \in B]$ for every $t_1, \ldots, t_n$, where V is a (possibly stopped) Brownian Motion. These results will be used in the proof of Theorem 2. In section 5 we show the proof of Theorem 2. The sixth section will conclude the chapter by showing applications of these results to financial modeling. In particular, we show applications to futures models. In section 7 we give proofs of technical results stated throughout the chapter.

4.1 Motivation and Preliminary Discussion

Recall Lévy's Theorem:

Theorem 4.1.1. *(Lévy, 1948) Let M be a continuous square-integrable local martingale such that $\langle M\rangle_t = t$ a.s. Then M is a Brownian Motion.*

This theorem states that there is a unique (in a weak sense) continuous square-integrable martingale with the same quadratic variation as Brownian Motion. We seek conditions under which the law of the quadratic variation of a continuous martingale characterizes the law of the martingale. We further motivate this problem with the following example.

Example 8. Consider a Brownian Motion $\{W_t, \mathcal{F}_t; t \geq 0\}$ on a space $(\Omega, \mathcal{F}, \mathbf{P})$. Set $X_t = \int_0^t 2W_s dW_s = W_t^2 - t$, and $Y_t = -X_t$. We have

$$\langle X\rangle_t = \langle Y\rangle_t = \int_0^t 4W_s^2 ds$$

This is an a.s. equality. However, $\{X_t\}$ and $\{Y_t\}$ do not have the same distribution. In fact, for any $t \in \mathbb{R}^+$,

$$\mathbf{P}(X_t > t) > 0 = \mathbf{P}(Y_t > t).$$

Thus, square-integrable semimartingales may have the same quadratic variation and different finite dimensional distributions. Therefore, Lévy's theorem may not always be extended to more general situations. In fact, consider a nonsymmetric martingale X. Then X and $-X$ are martingales with different distributions, but with equal quadratic variations.

Before proceeding, we recall the definition of the distribution of a stochastic process. Consider a stochastic process $X(t,\omega)$ on a space $(\Omega, \mathcal{F}, \mathbf{P})$, taking values on $(I\!R, \mathcal{B}(I\!R))$. For definiteness, assume the time variable is indexed by $I\!R_+ = [0,\infty)$. On the set $I\!R^{[0,\infty)}$ of functions from $I\!R_+$ to $I\!R$, consider the product sigma-algebra $\mathcal{B}(I\!R^{[0,\infty)})$. Then we may view X as a function from $(\Omega, \mathcal{F}, \mathbf{P})$ to $(I\!R^{[0,\infty)}, \mathcal{B}(I\!R^{[0,\infty)}))$, which is measurable by definition of the product sigma-algebra. The *law* $\mathbf{P}_X$ of X is the induced probability on the image space. That is, for $A \in \mathcal{B}(I\!R^{[0,\infty)})$,

$$\mathbf{P}_X(A) = \mathbf{P}\{\omega : X(\omega) \in A\}.$$

By definition of the product sigma-algebra, this law is determined by the finite-dimensional distributions

$$\mathbf{P}\{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in A\},$$

for $n \in N$, $A \in \mathcal{B}(I\!R^n)$. If all the paths of X lie in a subset $\mathcal{S}$ of $R^{[0,\infty)}$, we may work with the sigma-algebra generated by the finite-dimensional cylinder sets

$$C = \{f \in \mathcal{S} : (f(t_1), \dots, f(t_n)) \in A\},\ n \in I\!N,\ t_1, \dots, t_n \in [0,\infty), A \in \mathcal{B}(I\!R^n).$$

Consider two processes, X and Y, defined on spaces $(\Omega_X, \mathcal{F}_X, \mathbf{P}_X)$, $(\Omega_Y, \mathcal{F}_Y, \mathbf{P}_Y)$, respectively. If X and Y have the same law, we write $(X, \mathbf{P}_X) \stackrel{d}{=} (Y, \mathbf{P}_Y)$, or $X \stackrel{d}{=} Y$ if the meaning is clear.

We explore conditions under which the law of the quadratic variation of a martingale determines uniquely the law of the martingale. The main results were stated at the beginning of the Chapter. Theorem 1 gives an answer to the proposed question: On a fairly general class of continuous divergent martingales, Lévy's theorem is valid only within the subclass of Gaussian martingales. This gives a partial converse to Lévy's Theorem in the following sense. From Lévy's Theorem, if a continuous martingale is Gaussian, then its law is characterized by the law of its quadratic variation, which is deterministic (see Proposition 4.2.1). Theorem 1 gives a converse of this statement.

Theorem 1 suggests studying a different (more restricted) class of martingales within which we may get the desired characterization. This is done in Theorem 2, where we restrict our study to martingales with volatilities that are stochastic only through dependence on the martingale itself. Theorem 2 roughly states that within such a class of martingales, the distribution of the quadratic variation determines almost uniquely the distribution of the martingale. In fact, if two martingales with different law in this class have the same quadratic variation distribution, then one is the reflection of the other.

4.2 Proof of Theorem 1 and Unique Martingales

The statement of Theorem 1 suggests defining classes of martingales with distributions that are uniquely determined by the distribution of the quadratic variation. To do this, we list other classes of martingales and find relations among these classes. We focus our attention on one-dimensional martingales defined on general probability spaces. Using relations between *Ocone* martingales and *Unique* martingales, defined below, we prove Theorem 1. Recall that a continuous local martingale M is called *divergent* if $\langle M \rangle_\infty = \infty$ a.s.

Theorem 4.2.1. *(Dambis(1965), Dubins and Schwarz (1965)) Let $\{M_t, \mathcal{F}_t; t \geq 0\}$ be a divergent continuous local martingale. For each $s \geq 0$ define*

$$T(s) = \inf\{t \geq 0 : \langle M \rangle_t > s\}.$$

Then

$$B_s = M_{T(s)}, \quad \mathcal{G}_s = \mathcal{F}_{T(s)}, \quad s \geq 0$$

is a standard one-dimensional Brownian Motion. We have a.s.

$$M_t = B_{\langle M \rangle_t}.$$

For a divergent continuous local martingale M, the Brownian Motion defined in Theorem 4.2.1 is called its DDS Brownian Motion.

We now list some classes of local martingales. For simplicity, we denote by $\mathcal{D}$ the class of divergent continuous local martingales.

1. $\mathcal{G}$, Gaussian Martingales. These are the continuous martingales with Gaussian finite dimensional distributions.

2. $\mathcal{E}$, Extremal Martingales. A continuous local martingale X is extremal if its law is extremal in $\mathcal{K}$, the class of measures under under which X is a martingale. That is, if it cannot be written as a strict convex combination of other elements of $\mathcal{K}$.

3. $\mathcal{P}$, Pure Martingales. $X \in \mathcal{D}$ is pure if $\mathcal{F}^X_\infty = \mathcal{F}^B_\infty$ where B is the DDS Brownian Motion of X.

4. $\mathcal{O}$, Ocone Martingales. $X \in \mathcal{D}$ is an Ocone martingale if its DDS Brownian Motion is independent of $\langle X \rangle$. This name was used by Vostrikova and Yor (1999) to denote this class of martingales, which was studied and characterized by Ocone (1993).

5. $\mathcal{U}$, Unique Martingales. A continuous local martingale X is *Unique* if for every martingale Y, $\langle X \rangle \stackrel{d}{=} \langle Y \rangle$ implies $X \stackrel{d}{=} Y$.

We first exhibit results that will be used for the proof of Theorem 1. After its proof, we give characterizations and relations between some of these classes. We start by recalling a helpful characterization of $\mathcal{G}$ (see Revuz and Yor (1999)).

Proposition 4.2.1. *Let X be a continuous martingale. Then $X \in \mathcal{G}$ if and only if $\{\langle X\rangle_t;\ t \geq 0\}$ is deterministic.*

We now state Ocone's theorem, which is the starting point of the proof of Theorem 1.

Theorem 4.2.2. *(Ocone, 1993) Let $X \in \mathcal{D}$ be given. Then $X \in \mathcal{O}$ if and only if $\{\int_0^t \epsilon_s dX_s,\ t \geq 0\} \stackrel{d}{=} X$ for every predictable process ϵ such that $|\epsilon| = 1$.*

Ocone showed that the condition of the theorem may be restricted to deterministic processes ϵ of the form $1_{[0,a)} - 1_{[a,\infty)}$ for $a \in I\!\!R_+$. The following result is an immediate consequence of Ocone's Theorem.

Corollary 4.2.1. *Let $X \in \mathcal{D}$ be given. If $X \in \mathcal{U}$ then $X \in \mathcal{O}$.*

We also need the following two results. The first one is given in Vostrikova and Yor (1999) in a slightly more general version, where μ is only required never to vanish.

Theorem 4.2.3. *(Vostrikova and Yor, 1999) Let $\{B_t, \mathcal{F}_t\}$ be a Brownian Motion, and $\{\mu_t\}$ a strictly positive $\{\mathcal{F}_t\}$-adapted process such that $\int_0^T \mu_s^2 ds < \infty\ \forall T$ a.s. and $\int_0^\infty \mu_s^2 ds = \infty$ a.s. Then $\{M_t = \int_0^t \mu_s dB_s; 0 \leq t\} \in \mathcal{O}$ if and only if $\{B_t; t \geq 0\}$ is independent from $\mathcal{N} = \sigma\{\mu_s, s \geq 0\}$.*

Lemma 4.2.2. *Let $\{X_t; t \geq 0\}$ be a nondeterministic process. There exist a Brownian Motion W and a process Y on some probability space, such that $Y \stackrel{d}{=} X$ and W and Y are not independent.*

Proof. Let $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ be probability spaces on which we have, respectively, a Brownian Motion W_1 and a process Y_2 such that $Y_2 \stackrel{d}{=} X$. Consider the product space

$$(\Omega, \mathcal{F}, \mathbf{P}) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mathbf{P}_1 \times \mathbf{P}_2).$$

On this space consider the extensions W and Y of W_1 and Y_2, respectively.

Since Y is nondeterministic, there exist $t_0 \geq 0$ and $a_0 \in I\!\!R$ such that $\alpha = \mathbf{P}_2(A_0) \in (0,1)$, where $A_0 = \{\omega \in \Omega_2 : Y_{t_0}(\omega) \geq a_0\}$. Call b_0 the real number for which $\mathbf{P}_1(B_0) = \alpha$, where $B_0 = \{\omega \in \Omega_1 : W_{t_0}(\omega) \geq b_0\}$. For $B_1 \in \mathcal{F}_1$ and $A_1 \in \mathcal{F}_2$, set

$$\mathcal{Q}(B_1 \times A_1) = \frac{\mathbf{P}((B_1 \cap B_0) \times (A_1 \cap A_0))}{\alpha} + \frac{\mathbf{P}((B_1 \cap B_0^c) \times (A_1 \cap A_0^c))}{1 - \alpha}.$$

We extend this measure to all $\mathcal{F}$ and observe that

$$\begin{aligned}
& \mathcal{Q}(\Omega) = 1 \\
\forall C \in \mathcal{B}(I\!R^{[0,\infty)}) \quad & \mathcal{Q}(Y \in C) = \mathcal{Q}(\Omega_1 \times \{Y_2 \in C\}) \\
& = \mathbf{P}_1(B_0)\frac{\mathbf{P}_2(A_0 \cap \{Y_2 \in C\})}{\alpha} + \mathbf{P}_1(B_0^c)\frac{\mathbf{P}_2(A_0^c \cap \{Y_2 \in C\})}{1-\alpha} \\
& = \mathbf{P}_2(Y_2 \in C) \\
\forall C \in \mathcal{B}(\mathcal{C}([0,\infty))) \quad & \mathcal{Q}(W \in C) = \mathcal{Q}(\{W_1 \in C\} \times \Omega_2) \\
& = \frac{\mathbf{P}_1(B_0 \cap \{W_1 \in C\})}{\alpha}\mathbf{P}_2(A_0) + \frac{\mathbf{P}_2(A_0^c \cap \{Y_2 \in C\})}{1-\alpha}\mathbf{P}_1(B_0^c) \\
& = \mathbf{P}_1(W_1 \in C).
\end{aligned}$$

Therefore, $\mathcal{Q}$ is a measure on $(\Omega, \mathcal{F})$ under which W is a Brownian Motion and $Y \stackrel{d}{=} X$. However,

$$\mathcal{Q}(Y_{t_0} \geq a_0, W_{t_0} \geq b_0) = \alpha \neq \alpha^2 = \mathcal{Q}(Y_{t_0} \geq a_0)\mathcal{Q}(W_{t_0} \geq b_0).$$

It follows that Y and W are not independent. □

Theorem 1. *Let M be a divergent continuous local martingale with a.s. absolutely continuous quadratic variation paths. Assume $\frac{d\langle M\rangle_t}{dt} > 0$ a.s. If $M \in \mathcal{U}$ then $M \in \mathcal{G}$.*

Proof. Assume $M \notin \mathcal{G}$. Then $\{X_t = \frac{d\langle M\rangle_t}{dt}; t \geq 0\}$ is not deterministic. Let W, Y be processes defined in some filtered space $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ such that W is a Brownian Motion, $Y \stackrel{d}{=} X$ and W and Y are not independent (their existence is given by the previous lemma).

Consider the martingale $\{N_t = \int_0^t Y_s dW_s; 0 \leq t\}$. By Theorem 4.2.3, N is not an Ocone martingale. By Proposition 4.2.1, $N \notin \mathcal{U}$. By construction, $\{\langle M\rangle_t; t \geq 0\} \stackrel{d}{=} \{\langle N\rangle_t; t \geq 0\}$. Therefore, $M \notin \mathcal{U}$. □

We now give further results on the relation and characterizations of other classes of martingales previously listed. The first result compares the classes $\mathcal{E}$, $\mathcal{P}$ and $\mathcal{G}$. In the second one, an alternative definition to pure martingales is given. These are standard results. See for example, Revuz and Yor (1999).

Theorem 4.2.4. $\mathcal{G} \cap \mathcal{D} \subset \mathcal{P} \subset \mathcal{E}$.

Proposition 4.2.3. *Let M be a divergent continuous martingale with DDS Brownian Motion W. The following three statements are equivalent.*

1. $M \in \mathcal{P}$.
2. $T_t \equiv \inf\{s \geq 0; \langle M\rangle_s > t\}$ *is* $\mathcal{F}_\infty^W$*-measurable for every* t.

3. $\langle M\rangle_t$ is $\mathcal{F}_\infty^W$-measurable for every t.

Vostrikova and Yor (1999) give another interesting characterization of Ocone martingales, Theorem 4.2.5. They also characterize Ocone martingales that are extremal, in Proposition 4.2.4.

Theorem 4.2.5. *(Vostrikova and Yor, 1999) Let X be a divergent continuous martingale. $X \in \mathcal{O}$ is equivalent to the following condition: If $\{\phi_t, t \geq 0\}$ is a predictable process such that*

$$D_t^\phi = \exp\Big(\int_0^t \phi_s dX_s - \frac{1}{2}\int_0^t \phi_s^2 d\langle X\rangle_s\Big)$$

is a martingale and $\mathcal{Q}_{|\mathcal{F}_t} = D_t^\phi P_{|\mathcal{F}_t}$ defines a new probability, then

$$\tilde{X}_t^\phi = X_t - \int_0^t \phi_s d\langle X\rangle_s$$

satisfies $\{\tilde{X}^\phi, \mathcal{Q}\} \stackrel{d}{=} \{X, \mathcal{P}\}$.

Proposition 4.2.4. *(Vostrikova and Yor, 1999) $\mathcal{O} \cap \mathcal{E} = \mathcal{G} \cap \mathcal{D}$.*

Theorems 4.2.2 and 4.2.5 show that Ocone martingales share important properties with Gaussian martingales. However, there are Ocone martingales which are not Gaussian. Several interesting examples can be found in Vostrikova and Yor (1999). We give a generic idea to show that these two classes are not equal.

Example 9 (Ocone Martingales that are not Gaussian.). Consider a nonnegative continuous process $\{X_t; t \geq 0\}$ on a space $(\Omega, \mathcal{F}, \mathbf{P})$, that is not deterministic and satisfies $\int_0^\infty X_t^2 dt = \infty$ a.s. Assume that on this space we have a Brownian Motion $\{W_t; t \geq 0\}$ independent of X. By definition, the process $\{W_{\int_0^t X_s^2 ds}; t \geq 0\}$ is an Ocone martingale. However, it is not a Gaussian martingale because its quadratic variation is not deterministic.

Finally, we relate the class of pure martingales with a class of solutions to stochastic differential equations which will be studied in section 4.5.

Proposition 4.2.5. *Let (X, W), $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ be a weak solution of*

$$X_0 = 0, \quad dX_t = g(X_t)dW_t,$$

where g is a measurable function that never vanishes. Assume $X \in \mathcal{D}$. Then $X \in \mathcal{P}$.

Proof. This is a corollary of Proposition 4.2.3, recalling that

$$T_s = \int_0^t \frac{du}{g_1(B_u)^2},$$

where B is the DDS Brownian Motion of M. □

Remark 4.2.1. Consider a continuous martingale M. In general, M is not divergent, so it does not necessarily have a DDS Brownian Motion. In fact, for any continuous martingale M, if we define $T_t = \inf\{s \geq 0; \langle M\rangle_s > t\}$, then $W_t = M_{T_t} - M_0$ is a $\mathcal{F}_{T_t}-$Brownian Motion stopped at $\langle M\rangle_\infty$.* We may write $M_t = W_{\langle M\rangle_t} + M_0$ and $W_{t\wedge\langle M\rangle_\infty} = M_{T_t} - M_0$. This suggests defining a class $\mathcal{P}'$ as the set of continuous local martingales M such that $\langle M\rangle_t$ is $\mathcal{F}^W_\infty-$measurable for every t, where $W_t = M_{\inf\{s\geq 0;\langle M\rangle_s > t\}}$ is its "DDS stopped Brownian Motion". A similar extension may be done for $\mathcal{O}$.

4.3 Equivalent Stochastic Differential Equations

In this section we prove that changing the sign of the volatility of stochastic differential equations does not change the distribution of its solutions. The proof of Theorem 2 uses this result. We also prove the extension of this result to the multidimensional case, where absolute values of matrices are interpreted in terms of their Principal Components. This proof is rather technical, but it is an extension of the proof of the one-dimensional case. We will give a financial application of this extension in section 6.

The following multidimensional equations will be studied:

$$\forall\Gamma \in \mathcal{B}(\mathbb{R}^d),\ \mathbf{P}(\mathbf{M}_0 \in \Gamma) = \mu(\Gamma), \quad d\mathbf{M}_t = b(t, \mathbf{M}_t)dt + g(t, \mathbf{M}_t)d\mathbf{W}'_t,$$

where $\mathbf{M} = (M_1, \ldots, M_m)$, and $\mathbf{W} = (W_1, \ldots, W_n)$ is an n-dimensional standard Brownian Motion. Consider two functions g_1, g_2 such that the Principal Components of the matrices $g_i(t,x)g_i(t,x)'$ $(i = 1, 2)$ coincide. We will show that from a weak solution to one of the equations we can construct a weak solution to the other equation by redefining the Brownian Motion of the solution. This implies that when studying this type of SDE's, we need to pay attention only to orthogonal matrices g, since these will "generate" all possible positive semidefinite symmetric matrices.

We start with the following simple result for the one-dimensional case, which will show the idea to follow in several dimensions.

Proposition 4.3.1. *Let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{B}(\mathbb{R})-$measurable function. Consider the following one-dimensional SDE's*

$$M_0 = 0, \quad dM_t = |g(M_t)|dW_t \tag{4.3}$$

$$M_0 = 0, \quad dM_t = g(M_t)dW_t \tag{4.4}$$

Assume that uniqueness in the sense of probability law holds for (4.3)*. Then uniqueness in the sense of probability law holds for* (4.4)*.*

*This is Theorem V.1.7 in Revuz and Yor (1999), and is an extension of the Dambis, Dubins, Schwarz Theorem.

Proof. Let $(X^{(i)}, W^{(i)})$, $(\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbf{P}^{(i)})$, $\{\mathcal{F}_t^{(i)}\}$, $(i = 1, 2)$ be weak solutions of (4.4). Define

$$N_t^{(i)} \equiv \int_0^t \operatorname{sgn}(g(X_s^{(i)}))dW_s^{(i)}, \ (i = 1, 2),$$

where we define sgn(0)=1. Then $N_t^{(i)}$ is a continuous square-integrable martingale, and $\langle N^{(i)}\rangle_t = t$. By Lévy's Theorem, $\{N_t^{(i)}, \mathcal{F}_t^{(i)}\}$ is a Brownian Motion.

We have (see Karatzas and Shreve (1991), corollary 3.2.20)

$$X_t^{(i)} = \int_0^t g(X_s^{(i)})dW_s^{(i)} = \int_0^t |g(X_s^{(i)})|\operatorname{sgn}(g(X_s^{(i)}))dW_s^{(i)} = \int_0^t |g(X_s^{(i)})|dN_s^{(i)}.$$

It follows that $(X^{(i)}, N^{(i)})$, $(\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbf{P}^{(i)})$, $\{\mathcal{F}_t^{(i)}\}$, $(i = 1, 2)$ are weak solutions of (4.3). By weak uniqueness, $X^{(1)}$ and $X^{(2)}$ have the same law, and uniqueness in the sense of probability law for (4.4) follows. □

The proposition provides something stronger; from a solution to (4.4) we construct a solution to (4.3), using the same process, and changing the Brownian Motion. Similarly, from a solution to (4.3) we may construct a solution to (4.4). This will be stated more precisely in the multidimensional case, where the proposition works similarly, provided we set an appropriate analogue of the absolute value (or of the *signum* function). This is done using the principal components of the cross variation process. We also aim at generalizing this result in various directions, by including time dependence, drift, a converse of the proposition, initial conditions and addressing the issue of existence.

Notation. The $n \times n$ identity matrix will be denoted I_n.
If $A_1, A_2, \ldots, A_k$ are matrices of orders $m\times a_1, m\times a_2, \ldots, m\times a_k$, we write $[A_1\ A_2\ \ldots\ A_k]$ to denote the $m \times (a_1 + a_2 + \ldots + a_k)$ matrix made up of concatenating the columns of $A_1, A_2, \ldots, A_k$.

Let $g_{i,j} : I\!R^{d+1} \to I\!R$ be $\mathcal{B}(I\!R)-$measurable functions $(1 \le i \le d,\ 1 \le j \le n)$. Call g the matrix composed of these functions as entries. For $\mathbf{x} \in I\!R^d$ and $t \in I\!R_+$ diagonalize the (symmetric and positive semidefinite) matrix $g(t, \mathbf{x})g(t, \mathbf{x})'$:

$$g(t, \mathbf{x})g(t, \mathbf{x})' = P(t, \mathbf{x})\Lambda(t, \mathbf{x})P(t, \mathbf{x})', \qquad \text{where} \tag{4.5}$$

$\Lambda(t, \mathbf{x}) = \operatorname{diag}(\lambda_1(t, \mathbf{x}), \ldots, \lambda_d(t, \mathbf{x}))$, and $\lambda_1(t, \mathbf{x}) \ge \ldots \ge \lambda_d(t, \mathbf{x}) \ge 0$ are the d (real) eigenvalues of $g(t, \mathbf{x})g(t, \mathbf{x})'$. The matrix $P(t, \mathbf{x})$ is a d-dimensional orthogonal matrix, whose columns are the eigenvectors of $g(t, \mathbf{x})g(t, \mathbf{x})'$ and the columns of $P(t, \mathbf{x})\Lambda^{0.5}(t, \mathbf{x})$ are its principal components. This representation may not be unique. For every $\mathbf{x} \in I\!R^d$, $t \in I\!R^+$ choose one such diagonalization, so that $P(t, \mathbf{x})$ is measurable. Let μ be a given distribution on $I\!R^d$, and $\mathbf{b} : I\!R^{d+1} \mapsto I\!R^d$ a measurable function. Consider the following d-dimensional SDE's

$$\forall\Gamma \in \mathcal{B}(I\!R^d),\ \mathbf{P}(\mathbf{M}_0 \in \Gamma) = \mu(\Gamma), \quad d\mathbf{M}_t = \mathbf{b}(t, \mathbf{M}_t)dt + P(t, \mathbf{M}_t)\Lambda(t, \mathbf{M}_t)^{0.5}d\mathbf{W}_t' \tag{4.6}$$

$$\forall\Gamma \in \mathcal{B}(I\!R^d),\ \mathbf{P}(\mathbf{M}_0 \in \Gamma) = \mu(\Gamma), \quad d\mathbf{M}_t = \mathbf{b}(t, \mathbf{M}_t)dt + g(t, \mathbf{M}_t)d\mathbf{W}_t' \tag{4.7}$$

Note that in (4.6) $\mathbf{W}$ is d-dimensional, and in (4.7) it is n-dimensional.

Proposition 4.3.2. *Let P, Λ be given by (4.5). Assume $(\mathbf{X}, \mathbf{W}^{(1)}), (\Omega, \mathcal{F}, \mathbf{P}), \{\mathcal{F}_t\}$ is a weak solution of (4.7). Then there exists a weak solution of (4.6) on an extended space, say $(\tilde{\mathbf{X}}, \tilde{\mathbf{W}}^{(2)}), (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}}), \{\tilde{\mathcal{F}}_t\}$, such that $\tilde{\mathbf{X}}$ and $\mathbf{X}$ have the same law. Conversely, if $(\mathbf{X}, \mathbf{W}^{(2)}), (\Omega, \mathcal{F}, \mathbf{P}), \{\mathcal{F}_t\}$ is a weak solution of (4.6), then there is a weak solution of (4.7) on an extended space, say $(\tilde{\mathbf{X}}, \tilde{\mathbf{W}}^{(1)}), (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}}), \{\tilde{\mathcal{F}}_t\}$, such that $\tilde{\mathbf{X}}$ and $\mathbf{X}$ have the same law.*

Proof. Let $(\mathbf{X}, \mathbf{W}^{(1)}), (\Omega, \mathcal{F}, \mathbf{P}), \{\mathcal{F}_t\}$ be a weak solution of (4.7). We can extend this space to allow it to support a d-dimensional Brownian Motion independent of the given sigma-algebra (Karatzas and Shreve(1991), remark 3.4.1). We will use a tilde to denote objects in the extended space.

Then $(\tilde{\mathbf{X}}, \tilde{\mathbf{W}}^{(1)})$, $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$, $\{\tilde{\mathcal{F}}_t\}$ is a weak solution of (4.7), and $\{\tilde{\mathbf{B}}_t, \tilde{\mathcal{F}}_t\}$ is a d-dimensional Brownian Motion independent of $(\tilde{\mathbf{X}}, \tilde{\mathbf{W}}^{(1)})$. Define

$$\begin{aligned}
D(s, \tilde{\mathbf{X}}_s) =& \operatorname{diag}(\mathbf{1}_{\{\lambda_1(s,\tilde{\mathbf{X}}_s)=0\}}, \cdots, \mathbf{1}_{\{\lambda_d(s,\tilde{\mathbf{X}}_s)=0\}}) \\
E(s, \tilde{\mathbf{X}}_s) =& \operatorname{diag}(\mathbf{1}_{\{\lambda_1(s,\tilde{\mathbf{X}}_s)>0\}}\lambda_1(s, \tilde{\mathbf{X}}_s)^{-0.5}, \cdots, \mathbf{1}_{\{\lambda_d(s,\tilde{\mathbf{X}}_s)>0\}}\lambda_d(s, \tilde{\mathbf{X}}_s)^{-0.5}) \\
\tilde{\mathbf{W}}_t^{(2)} =& \int_0^t E(s, \tilde{\mathbf{X}}_s)P(s, \tilde{\mathbf{X}}_s)'g(s, \tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(1)\prime} + \int_0^t D(s, \tilde{\mathbf{X}}_s)d\tilde{\mathbf{B}}_s'.
\end{aligned}$$

Then $\tilde{\mathbf{W}}_t^{(2)}$ is a continuous, square-integrable martingale, and $\langle \tilde{\mathbf{W}}^{(2)} \rangle_t = tI_{d\times d}$, by (4.5). By Lévy's Theorem, $\{\tilde{\mathbf{W}}_t^{(2)}, \mathcal{F}_t\}$ is a d-dimensional Brownian Motion.

Observe that

$$\begin{aligned}
&\left\langle \int_0^t P(s, \tilde{\mathbf{X}}_s)D(s, \tilde{\mathbf{X}}_s)P(s, \tilde{\mathbf{X}}_s)'g(s, \tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(2)\prime} \right\rangle_t \\
&\quad = \int_0^t P(s, \tilde{\mathbf{X}}_s)D(s, \tilde{\mathbf{X}}_s)P(s, \tilde{\mathbf{X}}_s)'g(s, \tilde{\mathbf{X}}_s)g(s, \tilde{\mathbf{X}}_s)'P(s, \tilde{\mathbf{X}}_s)D(s, \tilde{\mathbf{X}}_s)'P(s, \tilde{\mathbf{X}}_s)'ds \\
&\quad = \int_0^t P(s, \tilde{\mathbf{X}}_s)D(s, \tilde{\mathbf{X}}_s)\Lambda(s, \tilde{\mathbf{X}}_s)D(s, \tilde{\mathbf{X}}_s)'P(s, \tilde{\mathbf{X}}_s)'ds \\
&\quad = \int_0^t P(s, \tilde{\mathbf{X}}_s)\operatorname{diag}(\mathbf{1}_{\{\lambda_1(s,\tilde{\mathbf{X}}_s)=0\}}\lambda_1(s, \tilde{\mathbf{X}}_s), \cdots, \mathbf{1}_{\{\lambda_d(s,\tilde{\mathbf{X}}_s)=0\}}\lambda_d(s, \tilde{\mathbf{X}}_s))P(s, \tilde{\mathbf{X}}_s)'ds \\
&\quad = 0.
\end{aligned}$$

Therefore, $\int_0^t P(s, \tilde{\mathbf{X}}_s)D(s, \tilde{\mathbf{X}}_s)P(s, \tilde{\mathbf{X}}_s)'g(s, \tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(2)\prime} = 0$, being a martingale with zero quadratic variation starting at 0. Since $\Lambda(s, \tilde{\mathbf{X}}_s)^{0.5}D(s, \tilde{\mathbf{X}}_s) = 0$,

$$\int_0^t P(s, \tilde{\mathbf{X}}_s)\Lambda(s, \tilde{\mathbf{X}}_s)^{0.5}D(s, \tilde{\mathbf{X}}_s)d\tilde{\mathbf{B}}_s' = 0.$$

By the previous remarks,

$$\begin{aligned}\int_0^t P(s,\tilde{\mathbf{X}}_s)\Lambda(s,\tilde{\mathbf{X}}_s)^{0.5}d\tilde{\mathbf{W}}_s^{(2)\prime} &= \int_0^t P(s,\tilde{\mathbf{X}}_s)(I_d - D(s,\tilde{\mathbf{X}}_s))P(s,\tilde{\mathbf{X}}_s)'g(s,\tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(1)\prime} \\ &= \int_0^t g(s,\tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(1)\prime} - \int_0^t P(s,\tilde{\mathbf{X}}_s)D(s,\tilde{\mathbf{X}}_s)P(s,\tilde{\mathbf{X}}_s)'g(s,\tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(1)\prime} \\ &= \int_0^t g(s,\tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(1)\prime} \\ &= \tilde{\mathbf{X}}_t - \int_0^t \mathbf{b}(s,\tilde{\mathbf{X}}_s)ds\end{aligned}$$

It follows that $(\tilde{\mathbf{X}},\tilde{\mathbf{W}}^{(2)}),(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbf{P}}),\{\tilde{\mathcal{F}}_t\}$ is a weak solution of (4.6). By construction, $\tilde{\mathbf{X}}$ and $\mathbf{X}$ have the same law.

The converse direction works similarly, but not identically.

Let $(\mathbf{X},\mathbf{W}^{(2)}),(\Omega,\mathcal{F},\mathbf{P}),\{\mathcal{F}_t\}$ be a weak solution of (4.6). As before, we extend this space to allow it to support an n-dimensional Brownian Motion independent of the given sigma-algebra and we use a tilde to denote objects in the extended space.

Then $(\tilde{\mathbf{X}},\tilde{\mathbf{W}}^{(2)})$, $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbf{P}})$, $\{\tilde{\mathcal{F}}_t\}$ is a weak solution of (4.6), and $\{\tilde{\mathbf{B}}_t,\tilde{\mathcal{F}}_t\}$ is an n-dimensional Brownian Motion independent of $(\tilde{\mathbf{X}},\tilde{\mathbf{W}}^{(2)})$. Define

$$\begin{aligned}D(s,\tilde{\mathbf{X}}_s) &= \mathrm{diag}(\mathbf{1}_{\{\lambda_1(s,\tilde{\mathbf{X}}_s)=0\}},\dots,\mathbf{1}_{\{\lambda_d(s,\tilde{\mathbf{X}}_s)=0\}}) \\ E(s,\tilde{\mathbf{X}}_s) &= \mathrm{diag}(\mathbf{1}_{\{\lambda_1(s,\tilde{\mathbf{X}}_s)>0\}}\lambda_1(s,\tilde{\mathbf{X}}_s)^{-0.5},\dots,\mathbf{1}_{\{\lambda_d(s,\tilde{\mathbf{X}}_s)>0\}}\lambda_d(s,\tilde{\mathbf{X}}_s)^{-0.5}) \\ C(s,\tilde{\mathbf{X}}_s) &= g(s,\tilde{\mathbf{X}}_s)'P(s,\tilde{\mathbf{X}}_s)E(s,\tilde{\mathbf{X}}_s)\end{aligned}$$

Note that $C(s,\tilde{\mathbf{X}}_s)'C(s,\tilde{\mathbf{X}}_s) = I_n - D(s,\tilde{\mathbf{X}}_s)$. Therefore the non-zero columns of $C(s,\tilde{\mathbf{X}}_s)$ form an orthonormal system. In fact, there are as many of these columns as there are non-zero eigenvalues in $\Lambda(s,\tilde{\mathbf{X}}_s)$. Construct (in a measurable way) a matrix $V(s,\tilde{\mathbf{X}}_s)$ such that the non-zero columns of $C(s,\tilde{\mathbf{X}}_s)$ and the columns of $V(s,\tilde{\mathbf{X}}_s)$ form an orthonormal basis of $\mathbb{R}^n$. Define the $n\times(n+d)$ matrix

$$A(s,\tilde{\mathbf{X}}_s) = [C(s,\tilde{\mathbf{X}}_s)\ V(s,\tilde{\mathbf{X}}_s)\ 0].$$

The last $n-k$ columns of A are 0, so d columns of $A(s,\tilde{\mathbf{X}}_s)$ are 0, and the other n columns form an orthonormal system. Therefore, the rows of A form an orthonormal system, and we have the identity

$$A(s,\tilde{\mathbf{X}}_s)A(s,\tilde{\mathbf{X}}_s)' = I_n. \tag{4.8}$$

Moreover, we have

$$g(s,\tilde{\mathbf{X}}_s)A(s,\tilde{\mathbf{X}}_s) = [P(s,\tilde{\mathbf{X}}_s)\ \Lambda(s,\tilde{\mathbf{X}}_s)^{0.5}\ 0], \tag{4.9}$$

where the last n columns are 0.

Now we are ready to define the Brownian Motion that will be part of the weak solution to (4.7). Call $\tilde{\mathbf{U}}$ the $(d+n)$-dimensional Brownian Motion $[\tilde{\mathbf{W}}_t^{(2)}\tilde{B}_t]$. Define $\tilde{\mathbf{W}}_t^{(1)} = \int_0^t A(s,\tilde{\mathbf{X}}_s)d\tilde{\mathbf{U}}'_s$. Then, by (4.8) and Lévy's theorem, $\tilde{\mathbf{W}}_t^{(1)}$ is an n-dimensional Brownian Motion, and by (4.9),

$$\int_0^t g(s,\tilde{\mathbf{X}}_s)d\tilde{\mathbf{W}}_s^{(1)\prime} = \int_0^t P(s,\tilde{\mathbf{X}}_s)\Lambda(s,\tilde{\mathbf{X}}_s)^{0.5}d\tilde{\mathbf{W}}_s^{(2)\prime} = \tilde{\mathbf{X}}_t - \int_0^t \mathbf{b}(s,\tilde{\mathbf{X}}_s)ds.$$

Therefore, $(\tilde{\mathbf{X}},\tilde{\mathbf{W}}^{(1)}),(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbf{P}}),\{\tilde{\mathcal{F}}_t\}$ is a weak solution of (4.7). By construction, $\tilde{\mathbf{X}}$ and $\mathbf{X}$ have the same law. □

Corollary 4.3.1. *Weak existence holds for* (4.6) *if and only if it holds for* (4.7)*. Uniqueness in the sense of probability law holds for* (4.6) *if and only if it holds for* (4.7)*.*

Remark 4.3.1. The proof of Proposition 4.3.2 follows very closely the proof of the Martingale Representation Theorem (see Karatzas and Shreve (1991), Theorem 3.4.2). In fact, as a corollary of the proof of that theorem, we obtain the following result. Let $\{\mathbf{M}_t,\mathcal{F}_t; t\geq 0\}$ be a d-dimensional continuous local martingale with absolutely continuous cross-variations, on a space $(\Omega,\mathcal{F},\mathbf{P})$. Then on an extended space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbf{P}})$ there is a d-dimensional Brownian Motion $\mathbf{W}$ and a $d\times d$ measurable, adapted, and locally square-integrable matrix X such that

$$M_t = M_0 + \int_0^t X_s d\mathbf{W}'_s \text{ a.s.}$$

and $X_t = P_t\Lambda_t^{0.5}$ for $t\geq 0$, where $\{P_t; t\geq 0\}$ is a progressively measurable orthonormal matrix, and $\{\Lambda_t; t\geq 0\}$ is a progressively measurable diagonal matrix with nonnegative entries.

Roughly, this states that every martingale of the form $\int_0^t \boldsymbol{\sigma}_s d\mathbf{W}'_s$ may be written as $\int_0^t P_s\Lambda_s^{0.5}d\tilde{\mathbf{W}}'_s$ for a Brownian Motion $\tilde{\mathbf{W}}$ on a possibly extended space, and P, Λ keep the previous meanings.

Remark 4.3.2. A rather obvious consequence of this analysis is that to represent a continuous martingale as a Brownian integral, the Brownian Motion need not be of a higher dimension than the martingale. See remark 3.1.4.

4.4 Sets with Equal Distributions for Stopped Brownian Motion

Definition 6. Let $\{X(t); 0\leq t<\infty\}$ be a continuous stochastic process on $(\Omega,\mathcal{F},\mathbf{P})$. Fix $k\in I\!N$. We will say that $A,B\in\mathcal{B}(I\!R^k)$ are (X,k)*-equivalent* if

$$\forall t_1,\dots,t_k\geq 0,\ \mathbf{P}((X_{t_1},\dots,X_{t_k})\in A) = \mathbf{P}((X_{t_1},\dots,X_{t_k})\in B).$$

This relation will be denoted by $A \overset{(X,k)}{\sim} B$.

Consider a standard one-dimensional Brownian Motion $\{W_t, \mathcal{F}_t; t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathbf{P})$. Let a, b be given numbers in $I\!R$ with $a < 0 < b$. We are interested in characterizing (X, k)-equivalence for the following three cases:

1. $X_t = W_t$.
2. $X_t = W_{t \wedge T_b}$, where $T_b = \inf\{t \geq 0 : W_t = b\}$ $(b > 0)$.
3. $X_t = W_{t \wedge T_{a,b}}$, where $T_{a,b} = \inf\{t \geq 0 : W_t = a \text{ or } W_t = b\}$ $(a < 0 < b)$.

These results will be used in the proof of Theorem 2. It will be seen that it is enough to study the cases $n = 1$ and $n = 2$. The proofs of the following three lemmas will be given in section 4.7. In the following lines, λ denotes Lebesgue measure.

Lemma 4.4.1. *Consider the case $X_t = W_t$. Then*

$$\forall A, B \in \mathcal{B}(I\!R^2) \ (A \overset{(X,2)}{\sim} B \Leftrightarrow A \cap B^c = -(B \cap A^c) \ \lambda\text{−a.s.})$$

Lemma 4.4.2. *Consider $X_t = W_{t \wedge T_b}$. The following holds true for every A, B in $\mathcal{B}((-\infty, b)^2)$:*

$$A \overset{(X,2)}{\sim} B \Leftrightarrow A = B \ \lambda\text{-a.s.}$$

Lemma 4.4.3. *Consider $X = W_{t \wedge T_{a,b}}$.*

1. *If $\frac{a}{b} = -\frac{m}{n} \in \mathbf{Q}$, where $m, n \in I\!N$, $(m, n) = 1$ then*
   (a) *For every A, B in $\mathcal{B}((a, b))$,*

   $$A \overset{(X,1)}{\sim} B \Leftrightarrow A \cap B^c = -(B \cap A^c) \ \lambda\text{-a.s. on } \left(-\frac{b}{n}, \frac{b}{n}\right)$$
   $$\text{and } A \cap B^c, B \cap A^c \text{ are periodic on } (a, b) \text{ with period } \frac{2b}{n}.$$

   (b) *If $-a \neq b$, then for every A, B in $\mathcal{B}((a, b)^2)$,*

   $$A \overset{(X,2)}{\sim} B \Leftrightarrow A = B \ \lambda\text{-a.s.}$$

   If $-a = b$, then for every A, B in $\mathcal{B}((a, b)^2)$,

   $$A \overset{(X,2)}{\sim} B \Leftrightarrow A \cap B^c = -(B \cap A^c) \ \lambda\text{-a.s.}$$

2. *If $\frac{a}{b} \notin \mathbf{Q}$, then for every A, B in $\mathcal{B}((a, b))$,*

   $$A \overset{(X,1)}{\sim} B \Leftrightarrow A = B \ \lambda\text{-a.s.}$$

4.5 Proof of Theorem 2

Let g_1, g_2 be measurable functions on $I\!R$. Consider the following stochastic differential equations:

$$dX_t = g_1(X_t)dW_t \tag{4.10}$$

$$dX_t = g_2(X_t)dW_t \tag{4.11}$$

We will focus our attention on distributions of solutions to these equations. Therefore, by Proposition 4.3.2, we can assume without loss of generality that $g_1(x), g_2(x) \geq 0$ for every $x \in I\!R$. For a measurable function f define

$$I(f) = \{x \in I\!R : \int_{-\epsilon}^{\epsilon} \frac{dy}{f^2(x+y)} = \infty \ \forall \epsilon > 0\}$$
$$Z(f) = \{x \in I\!R : f(x) = 0\}.$$

Theorem 4.5.1. *(Engelbert and Schmidt (1984)) For every initial distribution μ,* (4.10) *has a solution which is unique in the sense of probability law if and only if $I(g_1) = Z(g_1)$.*

Lemma 4.5.1. *Let $\{W_t, \mathcal{F}_t^W\}$ be a standard one-dimensional Brownian Motion on a space $(\Omega, \mathcal{F}, \mathbf{P})$. Fix an initial distribution μ. Let $\{\mathcal{G}_t\}$ be an extended filtration satisfying the usual conditions on which W remains a Brownian Motion. Let ξ be a $\mathcal{G}_0$-measurable random variable with distribution μ, independent of $\mathcal{F}_t^W$ for $t \geq 0$.*

Assume $I(g_1) = Z(g_1)$, $i = 1, 2$. Assume the quadratic variations of weak solutions to (4.10) *and* (4.11) *with initial distribution μ have the same law. Then*

$$\{g_1(\xi + W_{t \wedge \inf\{s>0: \xi + W_s \in I(g_1)\}}); t \geq 0\} \stackrel{d}{=} \{g_2(\xi + W_{t \wedge \inf\{s>0: \xi + W_s \in I(g_2)\}}); t \geq 0\}.$$

Proof. For $i = 1, 2$, define

$$T_s^{(i)} = \int_0^s \frac{du}{g_i(\xi + W_u)^2}$$
$$A_s^{(i)} = \inf\{t > 0; T_t^{(i)} > s\},$$

and set

$$X_t^{(i)} = \xi + W_{A_t^{(i)}},$$
$$\mathcal{F}_t^{(i)} = \mathcal{G}_{A_t^{(i)}}.$$

Then there exist Brownian Motions $\{B_t^{(i)}, \mathcal{F}_t^{(i)}; t \geq 0\}$ on a possibly extended space such that $(X^{(i)}, B^{(i)})$ $(i = 1, 2)$ are weak solutions of (4.10) and (4.11), respectively, with initial distribution μ and quadratic variations

$$\langle X^{(i)} \rangle_t = \int_0^t g_i^2(X_s^{(i)})ds = A_t^{(i)}$$

(see the proof of Theorem 5.4 in Karatzas and Shreve (1991) †).

†This Theorem is due to Engelbert and Schmidt, and precedes Theorem 4.5.1

By the conditions on g_1 and g_2 and Theorem 4.5.1, $\{A_t^{(1)}\}$ and $\{A_t^{(2)}\}$ have the same law (or equivalently, the same finite dimensional distributions). For $n \in N,\ t_1, \ldots, t_n, s_1, \ldots, s_n \in I\!R_+$, we have

$$\{\omega : T^{(i)}(s_1) < t_1, \ldots, T^{(i)}(s_n) < t_n\} = \{\omega : A^{(i)}(t_1) > s_1, \ldots, A^{(i)}(t_n) > s_n\},\ i = 1, 2.$$

Therefore, $\{T_t^{(1)}\} \stackrel{d}{=} \{T_t^{(2)}\}$. Now, if we write

$$T^{(i)}(s) = \begin{cases} \int_0^s \frac{du}{g_i(\xi + W_u)^2}, & s < A_\infty^{(i)} \\ \infty, & s \geq A_\infty^{(i)} \end{cases}$$

then the processes

$$\Big\{\frac{1}{G_i(\xi + W_s)^2} 1_{s < A_\infty^{(i)}} + \infty 1_{s \geq A^{(i)}\infty}; s \geq 0\Big\},\ i = 1, 2$$

have the same law, for some Borel measurable function G_i which is a.s equal to g_i, $i = 1, 2$. Since $g_i \geq 0$, then the processes

$$\{g_i(\xi + W_s) 1_{s < A_\infty^{(i)}}; s \geq 0\},\ i = 1, 2$$

have the same law. The result follows by observing that

$$R^{(i)} \equiv \inf\{s > 0 : T_s^{(i)} = \infty\} = \inf\{s > 0 : \xi + W_s \in I(g_i)\} \text{ a.s.}$$

(Lemma 5.5.2 in Karatzas and Shreve (1991)). □

This lemma gives an important relation between the functions g_1 and g_2 which can be further exploited using the characterization of sets with equal distributions for stopped Brownian Motion derived in section 4.4. This will lead to Theorem 2. Before proceeding, we include some necessary observations.

Remark 4.5.1. $I(g)$ is closed. To prove this, let z be a limit point of $I(g)$. Let $\epsilon > 0$ be given. Choose w in $I(g) \cap (z - \epsilon, z + \epsilon)$. Then

$$\int_{-2\epsilon}^{2\epsilon} \frac{dy}{g^2(z + y)} \geq \int_{-\epsilon}^{\epsilon} \frac{dy}{g^2(w + y)} = \infty.$$

Therefore $z \in I(g)$.

Notation. We will use μ_a to denote the unitary mass distribution on $I\!R$ with support a, where $a \in I\!R$. That is, μ_a satisfies $\mu_a(\{a\}) = 1$.

Remark 4.5.2. Take g_1, g_2, W to satisfy the conditions of Lemma 4.5.1. Let a be an arbitrary real number, and let μ_a be an initial distribution. From the proof of Lemma 4.5.1, $\{T_t^{(i)}\}$, $i = 1, 2$, have the same law, so $R^{(1)} \stackrel{d}{=} R^{(2)}$. In other words,

$$\inf\{s > 0 : a + W_s \notin I(g_1)\} \stackrel{d}{=} \inf\{s > 0 : a + W_s \notin I(g_2)\}. \quad (4.12)$$

Therefore, $a \in I(g_1) \Leftrightarrow a \in I(g_2)$. Now, if $a \notin I(g_1)$, then for $i = 1, 2$ we define

$$\begin{aligned} a_i &= \sup(I(g_i) \cap (-\infty, a)) \\ b_i &= \inf(I(g_i) \cap (a, \infty)). \end{aligned} \tag{4.13}$$

These numbers may be $\infty, -\infty$. By the previous remark $a_i < a < b_i$, $i = 1, 2$. We rewrite equation (4.12),

$$\inf\{s > 0 : a + W_s \notin (a_1, b_1)\} \stackrel{d}{=} \inf\{s > 0 : a + W_s \notin (a_2, b_2)\}.$$

This implies $a_1 = a_2$ and $b_1 = b_2$, or $a_1 = 2a - b_2$ and $b_1 = 2a - a_2$.[‡]

Now assume that the equality of quadratic variations holds for μ_a, for every $a \in I\!R$. Then $a_1 = a_2,\ b_1 = b_2$.

Lemma 4.5.2. *Let g_1, g_2 be a.s. equal measurable functions. Assume $I(g_1) = Z(g_1) = Z(g_2) = I(g_2)$. Let μ be an arbitrary initial distribution. Solutions to* (4.10) *and* (4.11) *with initial distribution μ have the same law.*

Proof. Let $W, \mathcal{G}, \xi, T_s^{(i)}, A_s^{(i)}$ be defined as before. Then $A_\infty^{(1)} = \inf\{s \geq 0; \xi + W_s \in I(g_1)\} = A_\infty^{(2)}$. For $\omega \in \Omega$, if $\xi(\omega) \in I(g_1)$, then $T_s^{(1)}(\omega) = T_s^{(2)}(\omega) = \infty\ \forall s \geq 0$. Otherwise, if $s < A_\infty^{(1)}$, then $T_s^{(1)}(\omega) = T_s^{(2)}(\omega)$ because $\frac{1}{g_i(\xi + W_t)^2}$, $i = 1, 2$ are a.s. equal integrable functions in $[0, s]$. If $s \geq A_\infty^{(1)}$, then $T_s^{(1)}(\omega) = T_s^{(2)}(\omega) = \infty$.

In any case, $T_s^{(1)}(\omega) = T_s^{(2)}(\omega)$ for $\omega \in \Omega$. Therefore $A_s^{(1)} = A_s^{(2)}$, and $W_{A_t^{(1)}} = W_{A_t^{(2)}}$. Since $I(g_1) = Z(g_1)$ and $Z(g_2) = I(g_2)$, then (4.10) and (4.11) satisfy weak uniqueness, by Theorem 4.5.1. Therefore solutions to these equations have the same distribution as $W_{A_t^{(1)}}$ and $W_{A_t^{(2)}}$. □

The following proposition explores further the relationship between g_1 and g_2 given an equality of distributions of the quadratic variation of solutions to (4.10) and (4.11) starting at 0. It starts from the conclusion of Lemma 4.5.1 and uses Remark 4.5.2. The proof will be given in section 4.7.

Proposition 4.5.3. *Let g_1, g_2 be Borel measurable functions with $g_1(0), g_2(0) \neq 0$, and $Z(g_i) = I(g_i)$, $i = 1, 2$. Assume the quadratic variations of solutions to* (4.10) *and* (4.11) *starting at 0 have the same law. Fix $a = 0$, and let a_i, b_i, $i = 1, 2$ be given by* (4.13) *(possibly equal to ∞).*

1. *If $-a_1 = b_1 = b$, then $-a_2 = b_2 = b$, and $g_1(x) = g_2(x)$ a.s. in $[-b, b]$, or $g_1(x) = g_2(-x)$ a.s. in $[-b, b]$.*
2. *If $-a_1 \neq b_1$ and $a_1 = a_2 = a$, then $b_1 = b_2$, and $g_1(x) = g_2(x)$ a.s. in $[a, b_1]$.*

[‡]This follows from the distribution of exit times for Brownian Motion (Exercise 2.8.11 in Karatzas and Shreve (1991)).

3. If $-a_1 \neq b_1$ and $-a_1 = b_2 = b$, then $-a_2 = b_1$, and $g_1(x) = g_2(-x)$ a.s. in $[-b, b_1]$.

Theorem 2. *Let g_1, g_2 be Borel measurable functions on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $I(g_i) = Z(g_i)$, $i = 1, 2$. Consider the stochastic differential equations*

$$X_0 = 0, \quad dX_t = g_1(X_t)dW_t \tag{4.14}$$

$$X_0 = 0, \quad dX_t = g_2(X_t)dW_t. \tag{4.15}$$

Let

$$(X^{(i)}, W^{(i)}),\ (\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbf{P}^{(i)}),\ \{\mathcal{F}_t^{(i)}\},\ (i = 1, 2)$$

be weak solutions of (4.14) and (4.15), respectively. Assume $\{\langle X^{(1)}\rangle_t; t \geq 0\} \stackrel{d}{=} \{\langle X^{(2)}\rangle_t; t \geq 0\}$. Then $X^{(1)} \stackrel{d}{=} X^{(2)}$, or $X^{(1)} \stackrel{d}{=} -X^{(2)}$.

Proof. From the previous proposition, and keeping the same notation, we have one of the following cases (a_i, b_i may be $\pm\infty$):

$$a_1 = a_2 = a,\ b_1 = b_2 = b \text{ and } g_1(x) = g_2(x) \text{ a.s. on } (a, b)$$

$$a_1 = -b_2 = a,\ b_1 = -a_2 = b \text{ and } g_1(x) = g_2(-x) \text{ a.s. on } (a, b)$$

Since the initial distribution is μ_0, then $X_t^{(1)} \in (a_1, b_1)$, and $X_t^{(2)} \in (a_2, b_2)$ a.s. From the former case we conclude that $X^{(1)} \stackrel{d}{=} X^{(2)}$, by Lemma 4.5.2. The latter case requires an extra step. Define h_2 on $\mathbb{R}$ by $h_2(x) = g_2(-x)$. Then $Y = -X^{(2)}$ satisfies

$$Y_0 = 0, \quad dY_t = -h_2(Y_t)dW_t$$

By Proposition 4.3.1, Y has the same distribution as the solution to

$$Z_0 = 0, \quad dZ_t = h_2(Z_t)dW_t = g_1(Z_t)dW_t.$$

We conclude that $X^{(1)} \stackrel{d}{=} -X^{(2)}$. □

This result has a trivial converse. Under the conditions there, if $X^{(1)} \stackrel{d}{=} X^{(2)}$, or $X^{(1)} \stackrel{d}{=} -X^{(2)}$, then $\{\langle X^{(1)}\rangle_t\} \stackrel{d}{=} \{\langle X^{(2)}\rangle_t\}$. We notice that this equivalent condition was derived using only up to two-dimensional distributions from the result in Lemma 4.5.1.

We relate Theorems 1 and 2 through the following example.

Example 10. Let

$$(X, W),\ (\Omega, \mathcal{F}, \mathbf{P}),\ \{\mathcal{F}_t\}$$

be a weak solution of (4.14), where $g_1(x) = |x| + 1$. Since g_1 is even, $\{X_t; t \geq 0\} \stackrel{d}{=} \{-X_t; t \geq 0\}$. Theorem 2 tells us that no martingale (other than X) that is a solution to a stochastic differential equation of type (4.14) gives a quadratic variation with equal law as that of $\langle X\rangle$. But X is a divergent martingale that is not Gaussian. By Theorem 1, there is a martingale Y with different law than X such that $\langle Y\rangle \overset{d}{=} \langle X\rangle$.

To find one such martingale, we start by observing (Proposition 4.2.5), that X is a pure martingale. Therefore, $X \notin \mathcal{O}$. By Ocone's Theorem, for some $a \geq 0$, we have

$$X^a = \{\int_0^t (1_{[0,a]} - 1_{(a,\infty)})dX_s, t \geq 0\} = \{\int_0^t (1_{[0,a](s)} - 1_{(a,\infty)}(s))(1 + |X_s|)dW_s, t \geq 0\} \overset{d}{\neq} X.$$

In fact,

$$X_t^a = \begin{cases} X_t, & \text{if } t \leq a \\ 2X_a - X_t, & \text{if } t \geq a, \end{cases}$$

so any $a > 0$ will cause the difference in distributions, given that volatilities are increasing in $|X|$.

The following result restates Theorem 2 under more restrictive conditions, for which we get equality of distributions for the martingales. First we start with a helpful Proposition.

Proposition 4.5.4. *Let g_1, g_2 be Borel measurable functions on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $I(g_i) = Z(g_i)$, $i = 1, 2$. Let*

$$(X^{(i),\mu}, W^{(i)}),\ (\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbf{P}^{(i)}),\ \{\mathcal{F}_t^{(i)}\},\ (i = 1, 2)$$

be weak solutions of (4.10) *and* (4.11), *respectively, with initial distribution μ. Assume $\{\langle X^{(1),\mu_a}\rangle_t\} \overset{d}{=} \{\langle X^{(2),\mu_a}\rangle_t\}$ for all $a \in \mathbb{R}$. Then $I(g_1) = I(g_2)$ and $g_1(x) = g_2(x)$ a.s. in $\mathbb{R}$.*

Proof. From Remark 4.5.2 we notice that $I(g_1) = I(g_2)$. Consider an arbitrary real number $a \notin I(g_1)$. Keeping the notation used in that remark, define $I \equiv (a_1, b_1) = (a_2, b_2)$. Define functions h_1, h_2 on $\mathbb{R}$ by $h_i(x) = g_i(x + a)1_I(x)$, $i = 1, 2$. Then

$$(X^{(i),\mu_a} - a, W^{(i)}),\ (\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbf{P}^{(i)}),\ \{\mathcal{F}_t^{(i)}\},\ (i = 1, 2)$$

are weak solutions of (4.10) and (4.11), where the h_i substitutes the g_i. By Proposition 4.5.3, $h_1 = h_2$ a.s. (observing that a can be taken to be any number in I.) Since $a \in \mathbb{R}$ was arbitrary, we conclude that $g_1 = g_2$ a.s.. ☐

Theorem 4.5.2. *Let g_1, g_2 be Borel measurable functions. Assume $I(g_i) = Z(g_i)$, $i = 1, 2$. Let*

$$(X^{(i),\mu}, W^{(i)}),\ (\Omega^{(i)}, \mathcal{F}^{(i)}, \mathbf{P}^{(i)}),\ \{\mathcal{F}_t^{(i)}\},\ (i = 1, 2)$$

be weak solutions of (4.10) *and* (4.11), *respectively, with initial distribution μ. Assume $\{\langle X^{(1),\mu_a}\rangle_t\} \overset{d}{=} \{\langle X^{(2),\mu_a}\rangle_t\}$ for all $a \in \mathbb{R}$. Then $X^{(1),\mu} \overset{d}{=} X^{(2),\mu}$ for every initial distribution μ.*

Proof. The conclusion follows from Proposition 4.5.4 and Lemma 4.5.2. ☐

4.6 Applications to Financial Modeling

In this section we show applications of the results obtained in the past sections, in particular to futures models. They will remain valid for general financial modeling, such as the HJM framework.

Example 11 (Determining the Model from a Sample Path.). Consider the following situation. A practitioner wishes to implement a futures-based model (or an HJM model). As it was suggested before, for pricing purposes, the models are completely determined with a specification of the volatility ($\boldsymbol{\sigma}$). To this end, she restricts $\boldsymbol{\sigma}$ to lie within a predetermined set. She is able to observe the evolution of the prices of Interest Rate futures contracts. Her objective is to choose the model from the restricted set that is most in agreement with the observed paths. For simplicity, assume that observations are done continuously. We also assume that the "real" process lies within the set determined by the practitioner. Is it possible to determine uniquely the corresponding model?

Let d be a given natural number. Let Σ be a set of Borel-measurable functions mapping $I\!R_+ \times I\!R^d$ to $I\!R_{d\times n}$ $(n = 1, 2, \dots)$, such that $\boldsymbol{\rho}(t,x) = f(t)\boldsymbol{\sigma}(x)$ for $\boldsymbol{\rho} \in \Sigma,\ t \in [0,\infty),\ x \in I\!R^d$, where $f : I\!R_+ \to I\!R - \{0\}$, $\boldsymbol{\sigma} : I\!R^d \to I\!R_{d\times n}$ are measurable functions. Consider the following d-dimensional stochastic differential equation:

$$d\mathbf{X}_t = \boldsymbol{\rho}(t, X_t)d\mathbf{W}_t', \tag{4.16}$$

where $\boldsymbol{\rho} = f\boldsymbol{\sigma} \in \Sigma$ and $\mathbf{W}$ stands for n-dimensional Brownian Motion. Assume weak existence holds for this equation, and let $(\mathbf{X}, \mathbf{W})$, $(\Omega, \mathcal{F}, \mathbf{P})$, $\{\mathcal{F}_t\}$ be a weak solution. The practitioner is given a realized path of $\mathbf{X}$, say $\{\mathbf{X}_t(\omega_0); t \geq 0\}$, $\omega_0 \in \Omega$. We assume that this path is *informative* in the following sense:

$$\mathbf{P}(\mathbf{X}_t(\omega) \in \text{ range of } \mathbf{X}(\omega_0)\ \forall t) = 1.$$

In other words, if we choose another path, its image will be contained in the image of $\mathbf{X}(\omega_0)$ almost surely. That is, the path provides as much information as possible. We may obtain the corresponding path of the cross variation of the process, i.e.

$$A_t(\omega_0) = \int_0^t \boldsymbol{\rho}(s, X_s)\boldsymbol{\rho}'(s, X_s)ds,\ t \geq 0. \tag{4.17}$$

We have assumed that the limiting cross variation of the path coincides with (4.17).[§] Therefore, we have knowledge of $f(t)^2\boldsymbol{\sigma}(\mathbf{X}_t(\omega_0))\boldsymbol{\sigma}(\mathbf{X}_t(\omega_0))'$, for $t \geq 0$.

[§]This is a strong assumption; however, the quadratic variation over a finite partition converges (in probability) to the quadratic variation of the process as the partition becomes finer. In fact, fix a date T, and positive numbers η and ϵ. For a finite partition $\Pi = \{0 = t_0 \leq t_1 \leq \dots \leq t_n = T\}$ of $[0,T]$ define the quadratic variation until date T, $V_T^{(2)}$ in the usual way,

$$V_T^{(2)}(\omega) = \sum_{k=1}^n |\mathbf{X}_{t_k}(\omega_0) - \mathbf{X}_{t_{k-1}}(\omega_0)|^2.$$

If the function $f \neq 0$ is known, we may determine the function $\boldsymbol{\sigma}(x)\boldsymbol{\sigma}(x)'$ for x in the image of $X(\omega_0)$. By Proposition 4.3.2, this is equivalent to knowing the Principal Components of $\boldsymbol{\sigma}(x)\boldsymbol{\sigma}(x)'$ for every "relevant" x. That is, two elements of Σ that produce the same Principal Components will generate equally distributed processes. Alternatively, if the function $\boldsymbol{\sigma}(x)$ is known, and $\boldsymbol{\sigma}(x) \neq 0 \; \forall x$, then we may determine the function f.

For definiteness, we restrict this analysis to exponential futures models introduced in Chapter 3 for the case of lognormal distributions. Fix a set of maturity dates $0 < T_1 < T_2, \ldots < T_m$, and a natural number n. Define

$$\Sigma = \{\boldsymbol{\rho} = (\rho_{ij})_{\substack{i=1,\ldots,m \\ j=1\ldots,n}} | \rho_{ij}(t,x) = \sigma_{ij}(t)x = \sum_{k=1}^{m} \beta_{jk} \exp(-\lambda_k(T_i - t))x, \; \lambda_k \geq 0, \; \beta_{jk} \in I\!R\}$$

As before, we make the assumption that given a continuous observation of the evolution of a set of futures rates, we are able to obtain its (theoretical) cross variation; i.e., we have knowledge of the matrix-valued process

$$A_t(\omega_0) = \int_0^t \boldsymbol{\rho}(s)\boldsymbol{\rho}(s)' ds, \; 0 \leq t \leq T_1,$$

and therefore we observe $\{\sum_{k=1}^{n} \sigma_{ik}(t)\sigma_{jk}(t)F(t,T_i)F(t,T_j); 0 \leq t \leq T_1\}$ for $i,j = 1,\ldots,m$. Since we observe the paths $\{F(t,T_i); t \geq 0\}$, $i = 1,\ldots,m$, by assuming future rates never vanish we determine the function $\boldsymbol{\sigma}(t)\boldsymbol{\sigma}(t)'$, $0 \leq t \leq T_1$. The SDE that models the evolution of the futures rates ((3.1) and (3.3)) satisfies weak existence and uniqueness in the sense of probability law, provided $\boldsymbol{\rho} \in \Sigma$. From Proposition (4.3.2), we conclude that the elements of Σ that could have implied this cross variation will produce equally distributed martingales. Therefore, upon knowledge of a sample path of the (theoretical) increasing process, we are able to determine uniquely the distribution of the original m-dimensional martingale (until date T_1). ¶

Example 12 (Determining the Model from Distributions of Quadratic Variations.). Let $\mathcal{N}$ be a given class of one-dimensional martingales. A practitioner wishes to establish an arbitrage-free futures model $\mathcal{M}$ with one tradable maturity, T. Therefore, we may assume that the futures price is model-led under $\mathbf{P}$, the risk-neutral measure (see section 2.4). Assume the futures price of the T-maturity contract is an element N of $\mathcal{N}$. The

For an appropriate fineness of the partition, $|V_T^{(2)}(\omega) - A_T(\omega)| \leq \epsilon$, with probability greater than $1 - \eta$. Extending this, for $M \in I\!N$ and for a partition size small enough, we have

$$\mathbf{P}[|V_\tau^{(2)} - A_\tau| \geq \epsilon] \leq \epsilon, \quad \tau = \frac{T}{2^M}, \frac{2T}{2^M}, \ldots, T.$$

Therefore, relying on the continuity of the quadratic variation, we can get arbitrarily close to $A_t(\omega_0)$, unless ω_0 lies in the probability zero exceptional set, which we assume is not the case. Hence, it is justifiable to assume that we may obtain the corresponding path for the quadratic variation.

¶We observe that there is not a unique set of parameters that produces this martingale; any set that produces the same cross variation will produce "the same" model (distributionally). See remark 3.1.1.

practitioner is given the distribution of the quadratic variation (increasing process) of N; i.e., she knows the law of $\langle N \rangle$. We might interpret this as "weakly" knowing the volatility of the futures price. It is natural to ask whether the practitioner can determine uniquely (the law of) N.

First consider the case

$$\mathcal{N} = \{M : M \text{ is a continuous divergent martingale, such that } M_0 = 0,$$
$$\langle M \rangle \text{ is a.s. absolutely continuous and } \frac{d}{dt}\langle M \rangle_t > 0 \ a.s.\}$$

Theorem 1 tells us that the information given is not enough to uniquely determine the distribution of the martingale. In other words, there is more than one potential futures model arising from $\mathcal{N}$. Moreover, there is no more information that may help us distinguish between these models. Unfortunately, this theorem does not specify how different are the distributions of the potential martingales.

Now consider

$$\mathcal{N} = \{M : M \text{ is a continuous martingale, such that } M_t = M_0 + \int_0^t g(M_s) dW_s,$$
$$\text{for a Brownian Motion } W \text{ and a measurable function } g \text{ such that } I(g) = Z(g)\}.$$

By Theorem 2, the practitioner has at most two potential candidates for the model. In this case, a model may be uniquely determined, modulo a reflection of the futures price process on the initial futures price. In many cases, one of the potential candidates may be ruled out by assuming nonnegativity of futures prices.

The situation proposed in this example is relevant when calibrating a model to current derivative prices. In many cases, the information given by these prices is partial information of the distribution of the quadratic variation of a process. Here we assume complete information of this distribution.

4.7 Proofs

Notation. We denote by ϕ_ν the cdf of the centered normal distribution with variance ν. We call μ_ν the probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ determined by this distribution, and f_ν the corresponding density on $\mathbb{R}$. This notation will also be used in two dimensions. For $0 \le \nu_1 \le \nu_2$, we will denote by μ_{ν_1,ν_2} the probability measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ determined by the centered bivariate normal distribution with covariance matrix

$$\begin{pmatrix} \nu_1 & \nu_1 \\ \nu_1 & \nu_2 \end{pmatrix}$$

and f_{ν_1,ν_2} will denote the corresponding density on $\mathbb{R}^2$. These will correspond to the two-dimensional distributions of Brownian Motion.

Remark 4.7.1.

$$\left(\forall t_1, t_2\ f_{t_1,t_2}(x_1, y_1) = f_{t_1,t_2}(x_2, y_2)\right) \Leftrightarrow (x_1, y_1) = (x_2, y_2) \text{ or } (x_1, y_1) = -(x_2, y_2)$$

Remark 4.7.2. Given nonnegative numbers t_1, t_2, t_3, t_4, there exist nonnegative numbers α, t_5, t_6 which may depend on the first four numbers, such that

$$\forall x, y\ f_{t_1,t_2}(x, y) f_{t_3,t_4}(x, y) = \alpha f_{t_5,t_6}(x, y)$$

Lemma 4.7.1. *Consider the sets*

$$G = \{g : \mathbb{R}^2 \to \mathbb{R}; g \text{ continuous, even and } \lim_{|x|\to\infty} g(x) = 0\}$$

$$G_\infty^\perp = \{h \in L^\infty(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_{1,2}); \iint hg d\mu_{1,2} = 0\ \forall g \in G\}$$

Then $h \in G_\infty^\perp \Rightarrow h(x, y) = -h(-x, -y)$ for almost all $(x, y) \in \mathbb{R}^2$.

Proof. Without loss of generality, assume we have $h \in G_\infty^\perp$ such that $h(x, y) \geq -h(-x, -y) + \epsilon$ on a bounded set $A \in \mathcal{B}(\mathbb{R} \times [0, \infty))$ of positive measure, for some $\epsilon > 0$. Define $g_0 = 1_{A \cup -A}$. Since the continuous functions in $\mathbb{R}^2$ are dense in $L^1(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_{1,2})$, we may find $g_1 \in G$ such that $\|g_0 - g_1\|_{L^1} \leq \frac{\epsilon\mu_{1,2}(A)}{2\|h\|_\infty}$. Then we have

$$\begin{aligned}\iint hg_1 d\mu_{1,2} \geq \iint hg_0 d\mu_{1,2} - \left|\iint h(g_0 - g_1) d\mu_{1,2}\right| &\geq \epsilon\mu_{1,2}(A) - \|h\|_\infty \|g_0 - g_1\|_{L^1} \\ &\geq \frac{\epsilon\mu_{1,2}(A)}{2} > 0.\end{aligned}$$

□

Lemma 4.7.2. *Let H be a measurable bounded real function on $\mathbb{R}^2$ such that*

$$\iint_{\mathbb{R}^2} H f_{t_1,t_2} = 0\ \forall t_2 \geq t_1 \geq 0.$$

Then $H(x, y) = -H(-x, -y)$ for almost all $(x, y) \in \mathbb{R}^2$.

Proof. Consider the sets G and $G_\infty^\perp$ defined in Lemma 4.7.1, and define

$$F = \operatorname{span}\{f_{t_1,t_2} : 0 \leq t_1 \leq t_2 \in \mathbb{R}\}$$

We consider the one point compactification of $\mathbb{R}^2$. Call $\mathcal{R}^2 = \mathbb{R}^2 \cup \{\infty\}$. We extend the domains of elements of F and G so that they vanish at $\{\infty\}$, and call the new sets F^* and G^*. We will need a space of finite measure later, so we fix arbitrarily the distribution $\mu_{1,2}$ on $\mathbb{R}^2$.

By Remark 4.7.2, F^* (and F) is closed under multiplication. Therefore, F^* is a subalgebra of G^*. By Remark 4.7.1, F^* separates non-antipode points. The Stone-Weierstrass

Theorem implies that F^* is dense in G^*, with the L^∞ norm. We conclude that F is dense in G in the space $L^\infty(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_{1,2})$.

Let $g \in G$ be given, and let $\{h_n; n = 1, 2, \dots\}$ be a sequence in F that converges to g (in $L^\infty(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_{1,2})$). Then

$$\begin{aligned} Hh_n &\overset{L^\infty}{\to} Hg \\ Hh_n &\to Hg \text{ a.s.} \end{aligned}$$

Since g is bounded, then $\{Hh_n; n \in \mathbb{N}\}$ is uniformly bounded. Using that $\mu_{1,2}$ is finite, the bounded convergence theorem implies

$$\iint Hh_n d\mu_{1,2} \to \iint Hg d\mu_{1,2}.$$

By Remark 4.7.2, for every $n \in \mathbb{N}$ there exist $k_n \in \mathbb{N}$ and nonnegative α_i, t_{i1}, t_{i2}, $i = 1, \dots, k_n$ such that

$$\iint Hh_n d\mu_{1,2} = \iint \sum_{i=1}^{k_n} \alpha_i H d\mu_{t_{i1}, t_{i2}} = 0.$$

We conclude that $\iint Hg d\mu_{1,2} = 0$. Since $g \in G$ was arbitrary, we have

$$H \in G_\infty^\perp = \{h \in L^\infty(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mu_{1,2}); \iint hg d\mu_{1,2} = 0 \ \forall g \in G\}.$$

The conclusion follows from Lemma 4.7.1. □

Proof of Lemma 4.4.1. Let A and B be given sets in $\mathcal{B}(\mathbb{R}^2)$ and look at the nontrivial direction of the lemma. Assume that $A \overset{(X,2)}{\sim} B$. This can be stated equivalently as

$$\forall t_2 \geq t_1 \geq 0 \ \iint_{\mathbb{R}^2} (1_A - 1_B) f_{t_1, t_2} = 0.$$

By Lemma 4.7.2, $1_A - 1_B$ is a.s. odd. This is equivalent to $A \cap B^c = -(B \cap A^c)$ a.s. □

Notation. For $A \in \mathbb{R}^2$ and real functions f and g with domain $\mathbb{R}$, we define the transformed set $A_{f,g}$ as follows:

$$A_{f,g} = \{(f(x), g(y)); (x, y) \in A\}.$$

We denote the identity function in $\mathbb{R}$ by e.

Lemma 4.7.3. *Let $\{W_t, \mathcal{F}_t; t \geq 0\}$ be a Brownian Motion on a space $(\Omega, \mathcal{F}, \mathbf{P})$, and $b > 0, t_2 > t_1 > 0$. For $A \in (-\infty, b)^2$ we have*

$$\mathbf{P}((W_{t_1 \wedge T_b}, W_{t_2 \wedge T_b}) \in A) = \iint (1_A + 1_{A_{2b-e,e}} - 1_{A_{e,2b-e}} - 1_{A_{2b-e,2b-e}}) d\mu_{t_1, t_2}(x, y).$$

Proof. Applying the reflection principle, for $y, z < b$,

$$\begin{aligned}
&\mathbf{P}(W_{t_1} \le y, W_{t_2} \le z, T_b > t_2) \\
&=\mathbf{P}(W_{t_1} \le y, W_{t_2} \le z) - \mathbf{P}(W_{t_1} \le y, W_{t_2} \le z, t_1 < T_b \le t_2) - \\
&\quad \mathbf{P}(W_{t_1} \le y, W_{t_2} \le z, 0 \le T_b < t_1) \\
&=\mathbf{P}(W_{t_1} \le y, W_{t_2} \le z) - \mathbf{P}(W_{t_1} \le y, W_{t_2} \ge 2b - z, t_1 < T_b) - \mathbf{P}(W_{t_1} \ge 2b - y, W_{t_2} \ge 2b - z) \\
&=\mathbf{P}(W_{t_1} \le y, W_{t_2} \le z) - \mathbf{P}(W_{t_1} \le y, W_{t_2} \ge 2b - z) + \mathbf{P}(W_{t_1} \le y, W_{t_2} \ge 2b - z, t_1 \le T_b) - \\
&\quad \mathbf{P}(W_{t_1} \ge 2b - y, W_{t_2} \ge 2b - z) \\
&=\mathbf{P}(W_{t_1} \le y, W_{t_2} \le z) - \mathbf{P}(W_{t_1} \le y, W_{t_2} \ge 2b - z) + \mathbf{P}(W_{t_1} \ge 2b - y, W_{t_2} \le z) - \\
&\quad \mathbf{P}(W_{t_1} \ge 2b - y, W_{t_2} \ge 2b - z)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\partial}{\partial y}\frac{\partial}{\partial z}\mathbf{P}(W_{t_1} \le y, W_{t_2} \le z, T_b > t_2) =& f_{t_1,t_2}(y,z) + f_{t_1,t_2}(2b-y,z) \\
&- f_{t_1,t_2}(y, 2b-z) - f_{t_1,t_2}(2b-y, 2b-z),
\end{aligned}$$

and the result follows. □

Proof of Lemma 4.4.2. For $C \in \mathcal{B}(I\!R^2)$, set $\zeta_C = 1_C + 1_{C_{2b-e,e}} - 1_{C_{e,2b-e}} - 1_{C_{2b-e,2b-e}}$. Without loss of generality, let A and B be given disjoint sets in $\mathcal{B}((-\infty, b)^2)$ such that $A \overset{(X,2)}{\sim} B$. From Lemmas 4.7.2 and 4.7.3, we conclude that $\zeta_A - \zeta_B$ is an a.s. odd function in $I\!R^2$.

In most of the following discussion we will omit writing a.s. equalities and a.s. belongings. Consider the following four scenarios for $(x, y) \in (-\infty, b)^2$:

1. $(x, y) \in (-b, b)^2$. Then $\zeta_A(x, y) - \zeta_B(x, y) = 1_A(x, y) - 1_B(x, y)$. Therefore,

   $$1_A(x, y) + 1_A(-x, -y) = 1_B(x, y) + 1_B(-x, -y)$$

   From this and the disjointness of A and B we conclude that on $(-b, b)^2$ we have $A = -B$ a.s.

2. $(x, y) \in (-b, b) \times (-\infty, -b)$. By oddness of $\zeta_A - \zeta_B$ we obtain

   $$1_A(x, y) - 1_A(-x, 2b + y) = 1_B(x, y) - 1_B(-x, 2b + y)$$

   From this and the disjointness of A and B we conclude that on $(-b, b) \times (-\infty, -b)$ we have

   (a) $(x, y) \notin A \cup B \Rightarrow (-x, 2b + y) \notin A \cup B$.
   (b) $(x, y) \in A \Rightarrow (-x, 2b + y) \in A$.
   (c) $(x, y) \in B \Rightarrow (-x, 2b + y) \in B$.

3. $(x,y) \in (-\infty,-b) \times (-b,b)$. Similarly to the previous case, we obtain

   (a) $(x,y) \notin A \cup B \Rightarrow (2b+x,-y) \notin A \cup B$.

   (b) $(x,y) \in A \Rightarrow (2b+x,-y) \in A$.

   (c) $(x,y) \in B \Rightarrow (2b+x,-y) \in B$.

4. $(x,y) \in (-\infty,-b) \times (-\infty,-b)$. Using similar arguments, we obtain

   (a) $(x,y) \notin A \cup B \Rightarrow (2b+x,2b+y) \notin A \cup B$.

   (b) $(x,y) \in A \Rightarrow (2b+x,2b+y) \in A$.

   (c) $(x,y) \in B \Rightarrow (2b+x,2b+y) \in B$.

Let $(x,y) \in (-b,b)^2 \cap A$ be given. By the case 3 above, $(-2b+x,-y) \in A$. By the case 4, $(-2b+x,-2b-y) \in A$, and so $(x,-2b-y) \in A$. By the case 2, $(-x,-y) \in A$. Therefore, $A = -A$ a.s. on $(-b,b)^2$. From the disjointness of A and B, and the four cases above, we finally conclude that $A = B = \varnothing$. $\square$

Lemma 4.7.4. *Let $\{W_t, \mathcal{F}_t; t \geq 0\}$ be a Brownian Motion on a space $(\Omega, \mathcal{F}, \mathbf{P})$, and $b > 0 > a, t_2 > t_1 > 0$. For $A \in \mathcal{B}((a,b))$, and $B \in \mathcal{B}((a,b)^2)$ we have*

$$\mathbf{P}(W_{t_1 \wedge T_{a,b}} \in A) = \int \sum_{n \in \mathcal{Z}} (1_{A+2n(b-a)} - 1_{A+2j(b-a)-2a}) d\mu_{t_1}$$

$$\begin{aligned}\mathbf{P}((W_{t_1 \wedge T_{a,b}}, W_{t_2 \wedge T_{a,b}}) \in B) = \iint \sum_{(m,n) \in \mathcal{Z}^2} (&1_{B_{2n(b-a)+e,2m(b-a)+e}} + 1_{B_{2a+2n(b-a)-e,2m(b-a)+e}} \\ &- 1_{B_{2n(b-a)+e,2a+2m(b-a)-e}} - 1_{B_{2a+2n(b-a)-e,2a+2m(b-a)-e}}) d\mu_{t_1,t_2}(x,y)\end{aligned}$$

Proof. The first equality is Proposition 2.8.10 in Karatzas and Shreve (1991). The second one is derived using the proof of this proposition, together with the proof of Lemma 4.7.3. $\square$

Proof of Lemma 4.4.3. We deal first with the one-dimensional cases (1 (a), and 2), and deal later with the two dimensional case of 1 (b) .

[Cases $A \overset{(X,1)}{\sim} B$.] Without loss of generality we assume that A and B are disjoint sets in $\mathcal{B}((a,b))$, and that $-a < b$. From Lemmas 4.7.2 and 4.7.4,

$$\zeta = \sum_{j=-\infty}^{\infty} (1_{A+2j(b-a)} - 1_{A+2j(b-a)-2a} - (1_{B+2j(b-a)} - 1_{B+2j(b-a)-2a}))$$

is an a.s. odd function. As before, and for clearness, we will omit writing a.s. equalities and a.s. belongings.

On $(a,-a)$, $\zeta = 1_A - 1_B$. Then for $x \in (0,-a)$, $1_A(x) + 1_A(-x) = 1_B(x) + 1_B(-x)$. Since A and B are disjoint, then

$$A = -B \text{ on } (-a,a). \tag{4.18}$$

Call $[x] = \{x + 2(jb + ka) : j,k \in \mathcal{Z}\} \cap (a,b)$. Later we will show that for $x \in (a,b)$,

$$x \in A \Rightarrow [x] \subset A,\ [-x] \subset B. \tag{4.19}$$

This will be enough to complete the proofs for the one-dimensional case. In fact, let's consider both cases separately.

i) Assume $\frac{a}{b} = \frac{-m}{n} \in \mathbf{Q}$, where $m,n \in I\!N$, $(m,n) = 1$. Then $[x] = \{x + \frac{2kb}{n} : k \in \mathcal{Z}\}$. We know that $A = -B$ on $(-a,a)$ a.s., so in particular this is true on $(\frac{-b}{n}, \frac{b}{n})$. Now, by (4.19), if $x,\ x + \frac{2b}{n} \in (a,b)$, then $x \in A \Leftrightarrow x + \frac{2b}{n} \in A$. The same is true for B. Therefore A and B are periodic on (a,b) with period $\frac{2b}{n}$.

ii) Assume $\frac{a}{b} \notin \mathbf{Q}$. We will show now that for $C \in \mathcal{B}((a,b))$ we have

$$\lambda(A \cap C) = \lambda(A)\frac{\lambda(C)}{b-a}. \tag{4.20}$$

From this,

$$0 = \lambda(A \cap B) = \lambda(A)\frac{\lambda(B)}{b-a},$$

which means that $\lambda(A) = 0$ or $\lambda(B) = 0$. But clearly one implies the other, since $A \overset{\Psi_X}{\sim} B$, and the claim follows.

Now, let $c,d \in [a,b]$ be given, with $c \le d$. For $x \in (a,b)$, $[x]$ is dense in (a,b). Let $\epsilon > 0$ be given. Choose $j_0, k_0 \in \mathcal{Z}$ such that $0 < j_0 a + k_0 b < \frac{\min(\epsilon, b-a)}{2}$. Set $\theta = \frac{1}{2(j_0 a + k_0 b)}$. For $x \ge a$ let j_x be such that $a + \frac{j_x}{\theta} \le x < a + \frac{j_x+1}{\theta}$. By (4.19), for $j \in \{0,1,\ldots,j_b - 1\}$ we have

$$A \cap (a + \frac{j}{\theta}, a + \frac{j+1}{\theta}) = \frac{1}{\theta} + A \cap (a + \frac{j-1}{\theta}, a + \frac{j}{\theta})$$

Therefore,

$$\lambda(A \cap (a + \frac{j}{\theta}, a + \frac{j+1}{\theta})) = \frac{\lambda(A \cap (a, a + \frac{j_b}{\theta}))}{j_b},$$

and it follows that

$$\frac{\lambda(A) - \frac{1}{\theta}}{j_b} \le \lambda(A \cap (a + \frac{j}{\theta}, a + \frac{j+1}{\theta})) \le \frac{\lambda(A)}{j_b}$$

Now, since

$$\lambda(A \cap (a + \frac{j_c + 1}{\theta}, a + \frac{j_d}{\theta})) \leq \lambda(A \cap (c, d)) \leq \lambda(A \cap (a + \frac{j_c}{\theta}, a + \frac{j_d + 1}{0}\theta)),$$

then

$$(\lambda(A) - \frac{1}{\theta})\frac{j_d - 1 - j_c}{j_b} \leq \lambda(A \cap (c, d)) \leq \lambda(A)\frac{j_d + 1 - j_c}{j_b}.$$

Finally, we use

$$\frac{(j_d - 1 - j_c)/\theta}{(j_b + 1)/\theta} \leq \frac{d - c}{b - a} \leq \frac{(j_d + 1 - j_c)/\theta}{j_b/\theta} \text{ , } and$$
$$\frac{j_d + 1 - j_c}{j_b} - \frac{j_d - 1 - j_c}{j_b + 1} \leq \frac{4\epsilon}{b - a}$$

to reach

$$\lambda(A)(\frac{d - c}{b - a}) - \frac{4\epsilon(d - c)}{b - a} \leq \lambda(A \cap (c, d)) \leq \lambda(A)(\frac{d - c}{b - a} + \frac{4\epsilon}{b - a})$$

Since $\epsilon > 0$ was arbitrary, we conclude that

$$\lambda(A \cap (c, d)) = \lambda(A)\frac{d - c}{b - a},$$

which proves (4.20) for open intervals in (a, b). The result for general Borel sets in (a, b) follows upon using the $\pi - \lambda$ theorem.

We need only prove (4.19). We start by observing that on $(-b, a)$ we have $\zeta = 0$. Therefore, by oddness of ζ we have for $x \in (-a, b)$,

$$1_A(x) - 1_A(x + 2a) = 1_B(x) - 1_B(x + 2a).$$

Since A and B are disjoint, we conclude that

$$x \in A \Leftrightarrow x + 2a \in A$$
$$x \in B \Leftrightarrow x + 2a \in B$$

Therefore,

$$x \in A \Rightarrow \{x + 2ka : k \in \mathcal{Z}\} \cap (a, b) \subset A. \tag{4.21}$$

The same holds true for B.

Before proceeding, observe that on $(b, b-2a)$, $\zeta = -1_{A-2a} + 1_{B-2a}$. On the other hand, from $1_{C-y}(x) = 1_C(x + y)$ it follows that ζ is periodic, with period $2(b - a)$. Therefore, from oddness of ζ we conclude that for $x \in (b, b - 2a)$,

$$1_A(x + 2a) + 1_A(2b - x) = 1_B(x + 2a) + 1_B(2b - x).$$

Using that A and B are disjoint, we arrive to

$$2b - x \in A \Leftrightarrow x + 2a \in B. \tag{4.22}$$

Using (4.21), (4.18) and (4.22) we obtain

$$\begin{aligned} x \in A &\Leftrightarrow x + 2k_0 a \in A \cap (a, -a) \text{ for some } k_0 \in \mathcal{Z} \\ &\Leftrightarrow -x + 2k_0 a \in B \text{ for some } k_0 \in \mathcal{Z} \\ &\Leftrightarrow -x + 2k_1 a \in B \cap (b + 2a, b) \text{ for some } k_1 \in \mathcal{Z} \qquad (4.23) \\ &\Leftrightarrow 2b + 2a + x - 2k_1 a \in A \text{ for some } k_1 \in \mathcal{Z} \\ &\Leftrightarrow \{2b + 2ka + x : k_1 \in \mathcal{Z}\} \cap (a, b) \subset A. \qquad (4.24) \end{aligned}$$

(4.23) and (4.24) give us (4.19).

[Case $A \overset{(X,2)}{\sim} B$.] Without loss of generality we assume that A and B are disjoint sets in $\mathcal{B}((a,b)^2)$, and that $-a \leq b$, where $\frac{a}{b} = \frac{-m}{n} \in \mathbf{Q}$, for some $m, n \in I\!N$, $(m,n) = 1$. From Lemmas 4.7.2 and 4.7.4, $\zeta_A - \zeta_B$ is an a.s. odd function in $I\!R^2$, where

$$\zeta_C = \sum_{(m,n)\in\mathcal{Z}^2} (1_{C_{2n(b-a)+e,2m(b-a)+e}} + 1_{C_{2a+2n(b-a)-e,2m(b-a)+e}} - 1_{C_{2n(b-a)+e,2a+2m(b-a)-e}} - 1_{C_{2a+2n(b-a)-e,2a+2m(b-a)-e}}).$$

As before, and for clearness, we will omit writing a.s. equalities and a.s. belongings. Using the oddness of $\zeta_A - \zeta_B$, and considering that $A, B \subset (a,b)^2$, we derive rules for different (not necessarily excluding) cases, following the proof of Lemma 4.4.2. The results are valid upon replacing A with B, by symmetry.

1. $(x,y) \in (a,-a)^2$. In this set we have $(x,y) \in A \Leftrightarrow (-x,-y) \in B$.
2. $(x,y) \in (a,-a) \times (-a,b)$. Here we obtain $(x,y) \in A \Leftrightarrow (-x, y-2a) \in A$. Therefore, we have
$$(x,y) \in A \Leftrightarrow \{((-1)^k x, 2ka + y) : k \in \mathcal{Z}\} \cap (a,b)^2 \subset A.$$
3. $(x,y) \in (a,-a) \times (b+2a, b)$. In this set we get $(x,y) \in A \Leftrightarrow (-x, 2(b+a) - y) \in B$.
4. $(x,y) \in (-a,b) \times (a,-a)$. This case gives $(x,y) \in A \Leftrightarrow (2a + x, -y) \in B$.
5. $(x,y) \in (b+2a, b) \times (a,-a)$. Then $(x,y) \in A \Leftrightarrow (2(a+b) - x, -y) \in B$.
6. $(x,y) \in (-a,b) \times (-a,b)$. Here we have
$$(x,y) \in A \Leftrightarrow \{(2k + x, 2ka + y) : k \in \mathcal{Z}\} \cap (a,b)^2 \subset A.$$
7. $(x,y) \in (-a,b) \times (b+2a, b)$. Then $(x,y) \in A \Leftrightarrow (2a + x, 2(a+b) - y) \in A$.

8. $(x,y) \in (b+2a, b) \times (-a, b)$. Then $(x,y) \in A \Leftrightarrow (2(a+b)-x, 2a+y) \in A$.

9. $(x,y) \in (b+2a, b) \times (b+2a, b)$. In this case we get $(x,y) \in A \Leftrightarrow (2(a+b)-x, 2(a+b)-y) \in B$.

Rule 1 above gives the result for the case $-a = b$. Now assume $-a < b$. As in the proof of Lemma 4.4.2, successive applications of these rules let us conclude that for $(x,y) \in (-\frac{b}{n}, \frac{b}{n})^2$, $(x,y) \in A \Leftrightarrow (-x,-y) \in A$. To see this, for $(x,y) \in (-\frac{b}{n}, \frac{b}{n})^2$, define $[(x,y)]$ as follows:

$$[(x,y)] = \begin{cases} \{((-1)^{k_1+k_2+k_3}x + 2k_1\frac{b}{n}, (-1)^{k_3}y + 2k_2\frac{b}{n}) : k_1, k_2 \in \mathcal{Z},\ k_3 \in \{0,1\}\}, & n+m \text{ even.} \\ \{((-1)^{k_1}x + 2k_2\frac{b}{n}, (-1)^{k_3}y + 2k_4\frac{b}{n}) : k_2, k_4 \in \mathcal{Z},\ k_1, k_3 \in \{0,1\}\}, & n+m \text{ odd.} \end{cases}$$

We claim that for $(x,y) \in (-\frac{b}{n}, \frac{b}{n})^2$, $(x,y) \in A \Leftrightarrow [(x,y)] \subset A$. Let's take the case when $m+n$ is even, the other one being similar.

Let $(x,y) \in (-\frac{b}{n}, \frac{b}{n})^2$ be given. Then

$$\begin{aligned} (x,y) \in A &\Leftrightarrow (-x,-y) \in B \quad \text{(rule 1)} \\ &\Leftrightarrow ((-1)^{k+1}x, 2ka - y) \in B \cap (a,b) \times (b+2a, b) \text{ for some } k \in \mathcal{Z} \quad \text{(rule 2)} \\ &\Leftrightarrow ((-1)^k x, 2(b+a) - 2ka + y) \in A \text{ for some } k \in \mathcal{Z} \quad \text{(rule 3)} \\ &\Leftrightarrow \{((-1)^k x, 2(b+a) - 2ka + y) : k \in \mathcal{Z}\} \cap (a,b)^2 \subset A \quad \text{(rule 2)} \end{aligned}$$

Repeating this argument, we obtain

$$(x,y) \in A \Leftrightarrow \{((-1)^{k_2}x, 2k_1(b+a) - 2k_2 a + y) : k_1, k_2 \in \mathcal{Z}\} \cap (a,b)^2 \subset A.$$

Since $m+n$ is even, then m and n are odd. This gives

$$(x,y) \in A \Leftrightarrow \{((-1)^{k_2}x, y + 2k_2\frac{b}{n}) : k_2 \in \mathcal{Z}\} \cap (a,b)^2 \subset A.$$

Similar arguments give $(x,y) \in A \Leftrightarrow [(x,y)] \subset A$. In particular, $A = -A$ a.s. on $(-\frac{b}{n}, \frac{b}{n})^2$. Therefore, $A = B$ a.s. on $(a,b)^2$. □

Notation. Let f be a given real function on $I\!R$, and A a Borel set in $I\!R^2$. By the inverse of A through f we mean the coordinate inverses; i.e.,

$$f^{-1}(A) = \{(x,y) \in I\!R^2 : (f(x), f(y)) \in A\}.$$

Recall that $I\!R_+$ denotes the set of (strictly) positive real numbers.

Proof of Proposition 4.5.3. Without loss of generality, we take g_1, g_2 to be 0 outside of $I(g_1)$ and $I(g_2)$, respectively. Remark 4.5.2 gives the proposed exhaustive structure of $a_i, b_i,\ i = 1, 2$. These numbers cannot be 0, given that $g_1(0), g_2(0) \neq 0$.

Let $\{W_t, \mathcal{F}_t^W\}$ be a standard one-dimensional Brownian Motion on a space $(\Omega, \mathcal{F}, \mathbf{P})$. By Lemma 4.5.1,

$$\{g_1(W_{t\wedge T_{a_1,b_1}}); t \geq 0\} \stackrel{d}{=} \{g_2(W_{t\wedge T_{a_2,b_2}}); t \geq 0\},$$

where $T_{-\infty,b} = T_b$, $T_{a,\infty} = T_a$ and $T_{-\infty,\infty} = \infty$. Therefore, for $A \in \mathcal{B}(I\!R^2)$, and $t_2 \geq t_1 \geq 0$,

$$\mathbf{P}((W_{t_1\wedge T_{a_1,b_1}}, W_{t_2\wedge T_{a_1,b_1}}) \in g_1^{-1}(A)) = \mathbf{P}((W_{t_1\wedge T_{a_2,b_2}}, W_{t_2\wedge T_{a_2,b_2}}) \in g_2^{-1}(A)). \tag{4.25}$$

Case 1. $a_1 = a_2 = a$.
In this case, from Remark 4.5.2, $b_1 = b_2 = b$.

1. Assume $-a \neq b$. Then by (4.25) and Lemmas 4.4.2 and 4.4.3, ‖

$$\forall A \in \mathcal{B}((I\!R_+)^2)\ g_1^{-1}(A) = g_2^{-1}(A))\text{ a.s.} \tag{4.26}$$

Assume there exists $C \in \mathcal{B}((a,b))$ with positive measure such that $g_1(x) \neq g_2(x)$ on C. Then we may find $C_0 \in \mathcal{B}((a,b))$ with positive measure, together with $n, k \in I\!N$ such that for $x \in C_0$,

$$|g_1(x) - g_2(x)| > \frac{1}{n} \text{ and } g_1(x) \in (\frac{k}{2n}, \frac{k+1}{2n}].$$

Consider $A = I\!R_+ \times (\frac{k}{2n}, \frac{k+1}{2n}]$. Then $(a,b) \times C_0 \subset g_1^{-1}(A)$, but $(a,b) \times C_0 \not\subseteq g_2^{-1}(A)$. This contradicts (4.26), and therefore $g_1 = g_2$ a.s. on (a,b).

2. Now assume $-a = b$. By (4.25) and Lemmas 4.4.1, and 4.4.3,

$$\forall A \in \mathcal{B}((I\!R_+)^2)\ g_1^{-1}(A) \cap (g_2^{-1}(A))^c = -(g_2^{-1}(A) \cap (g_1^{-1}(A))^c)\text{ a.s.} \tag{4.27}$$

First we will prove that for almost every x in (a,b), $\{g_1(x), g_1(-x)\} = \{g_2(x), g_2(-x)\}$. For $B \in \mathcal{B}(I\!R_+)$, we use $A = B \times I\!R_+$ in (4.27) to obtain

$$g_1^{-1}(B) \cap (g_2^{-1}(B))^c = -(g_2^{-1}(B) \cap (g_1^{-1}(B))^c)\text{ a.s.}$$

Assume the claim is false. Then we may find $C \in \mathcal{B}((a,b))$ and $n, k \in I\!N$ such that for $x \in C$,

$$|g_1(x) - g_2(x)| > \frac{1}{n},\ |g_1(x) - g_2(-x)| > \frac{1}{n} \text{ and } g_1(x) \in (\frac{k}{2n}, \frac{k+1}{2n}].$$

Take $B = (\frac{k}{2n}, \frac{k+1}{2n}]$. Then $D \subset g_1^{-1}(B) \cap (g_2^{-1}(B))^c$, but $-D \not\subseteq g_2^{-1}(B) \cap (g_1^{-1}(B))^c$. This contradiction proves the claim.

‖Recall that we are taking $g_1, g_2 \geq 0$ without loss of generality, due to proposition 4.3.1.

Now assume there exist $C_1, C_2 \in \mathcal{B}((a,b))$ with positive measure such that $g_1(x) \neq g_2(x)$ on C_1 and $g_1(x) \neq g_2(-x)$ on C_2. Then we may find $D_1, D_2 \in \mathcal{B}((a,b))$ with positive measure, together with $n, k_1, k_2 \in I\!N$ such that for $x \in D_1,\ y \in D_2$ we have

$$|g_1(x) - g_2(x)| > \frac{1}{n} \quad \text{and} \quad g_1(x) \in (\frac{k_1}{2n}, \frac{k_1+1}{2n}]$$
$$|g_1(y) - g_2(-y)| > \frac{1}{n} \quad \text{and} \quad g_1(y) \in (\frac{k_2}{2n}, \frac{k_2+1}{2n}].$$

Consider $A = (\frac{k_1}{2n}, \frac{k_1+1}{2n}] \times (\frac{k_2}{2n}, \frac{k_2+1}{2n}]$. Then $D_1 \times D_2 \subset g_1^{-1}(A) \cap (g_2^{-1}(A))^c$, but $-D_1 \times -D_2 \not\subseteq g_1^{-1}(A) \cap (g_2^{-1}(A))^c$. This contradicts (4.27). Therefore $g_1(x) = g_2(x)$ or $g_1(x) = g_2(-x)$ a.s. on (a,b).

Case 2. $-a_1 = b_2 = -a$ and $-a_1 \neq b_1 = b$.

From Remark 4.5.2, $-a_2 = b_1$. Consider the function h defined on $I\!R$ by $h(x) = g_2(-x)$. If

$$(X, W),\ (\Omega, \mathcal{F}, \mathbf{P}),\ \{\mathcal{F}_t\}$$

is a weak solution of (4.11), then $(-X, -W),\ (\Omega, \mathcal{F}, \mathbf{P}),\ \{\mathcal{F}_t\}$ is a weak solution of

$$Y_0 = 0, \quad dY_t = h(Y_t) dW_t.$$

Therefore, the hypotheses of the Proposition are satisfied with h instead of g_2. Since $I(h) = -I(g_2)$, then by the case 1, $g_1(x) = h(x)$ a.s. That is, $g_1(x) = g_2(-x)$ a.s. on (a_1, b_1). □

Bibliography

[1] P. Artzner and F. Delbaen (1989), "Term Structure of Interest Rates: the Martingale Approach," *Advances in Applied Mathematics* **10**: 95-129.

[2] R. Ash (1972), *Real analysis and Probability*, Academic Press.

[3] P. Billingsley (1989), *Probability and Measure*, Wiley.

[4] F. Black and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," *J. Political Econom* **81**: 637-654.

[5] F. Black, E. Derman and W. Toy (1990), "A One-Factor Model of Interest Rates and Its Applications to Treasury Bond Options," *Financial Analysts Journal*, **46**(1): 33-39.

[6] F. Black and P. Karasinski (1991): "Bond and Option Pricing when Short Rates are Lognormal," *Financial Analysts Journal*, **47**(4): 52-59.

[7] A. Brace, D. Gatarek and M. Musiela (1997), "The Market Model of Interest Rate Dynamics," *Mathematical Finance* **7**: 127-154.

[8] O. Cheyette (1996), "Markov Representation of the Heath-Jarrow-Morton Model," Working Paper, Barra Inc.

[9] J. C. Cox, J. E. Ingersoll and S. A. Ross (1985), "A Theory of the Term Structure of Interest Rates," *Econometrica* **53**: 385-407.

[10] J. Cvitanić and I. Karatzas (1993), "Hedging Contingent Claims with Constrained Portfolios," *Ann. Appl. Probab.* **3**: 652-681.

[11] R. C. Dalang, A. Morton and W. Willinger (1990), "Equivalent Martingale Measures and No-Arbitrage in Stochastic Security Market Models," *Stochastics* **29**: 185-201.

[12] F. Delbaen and W. Schachermayer (1994), "A General Version of the Fundamental Theorem of Asset Pricing," *Mathematische Annalen* **300**: 463-520.

[13] F. Delbaen and W. Schachermayer (1997), "The Fundamental Theorem of Asset Pricing for Unbounded Stochastic Processes," Working Paper, ETH Centrum and Institut für Statistik der Universität Wien.

[14] L. E. Dubins, M. Émery and M. Yor (1993), "On the Lévy Transformation of Brownian Motions and Continuous Martingales," *Séminaire de Probabilités* **1557**: 122-132.

[15] N. El Karoui and M. C. Quenez (1995), "Dynamic Programming and Pricing of Contingent Claims in an Incomplete Market," *SIAM J. Control and Optimization* **33**: 29-66.

[16] E. Fournié and D. Talay (1991), "Application de la Statistique des Diffusions á un Modéle de Taux D'intérêt," *Finance* **12**:79-111.

[17] B. Gustafsson, H.-O. Kreiss and J. Oliger (1995), *Time Dependent Problems and Difference Methods*, Wiley.

[18] J. M. Harrison and S. R. Pliska (1981), "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Process. Appl.* **11**(3): 215-260.

[19] J. M. Harrison and S. R. Pliska (1983), "A Stochastic Calculus Model of Continuous Trading: Complete Markets," *Stochastic Process. Appl.* **15**(3): 313-316.

[20] D. Heath and D. Jara (1999), "Term Structure Models Based on Futures Prices," Working Paper, Carnegie Mellon University.

[21] D. Heath, R. Jarrow and A. Morton (1992), "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica* **60**(1): 77-105.

[22] M. Hogan and K. Weintraub (1993), "The Lognormal Interest Rate Model and Eurodollar Futures," Working Paper, Citibank.

[23] F. Jamshidian (1993), "Option and futures evaluation with deterministic volatilities," *Mathematical Finance* **3**: 149-159.

[24] R. Kangro (1997), *Analysis of Artificial Boundary Conditions for Black-Scholes Equations*, Ph.D. thesis, Carnegie Mellon University.

[25] I. Karatzas and S. E. Shreve (1991), *Brownian Motion and Stochastic Calculus*, Springer-Verlag.

[26] I. Karatzas and S. E. Shreve (1998), *Methods of Mathematical Finance*, Springer-Verlag.

[27] P.E. Kloeden and E. Platen (1995), *Numerical Solutions of Stochastic Differential Equations*, Springer-Verlag.

[28] R. C. Merton (1973), "Theory of Rational Option Pricing," *Bell J. Econom. Manag. Sci.* **4**: 141-183.

[29] M. Musiela and M. Rutkowski (1998), *Martingale Methods in Financial Modelling*, Springer-Verlag.

[30] D. Ocone (1993), "A Symmetry Characterization of Conditionally Independent Increment Martingales," *Proceedings of the San Felice Workshop on Stochastic Analysis*, 147-167.

[31] D. Revuz and M. Yor (1999), *Continuous Martingales and Brownian Motion,* Springer-Verlag.

[32] P. Ritchken and L. Sankarasubramanian (1995): "Volatility Structure of Forward Rates and the Dynamics of the Term Structure," *Mathematical Finance* **1**: 55-72.

[33] L. Vostrikova and M. Yor (1999), "Some Invariance Properties (of the Laws) of Ocone's Martingales," Working Paper, Université d'Angers and Université Pierre et Marie Curie.