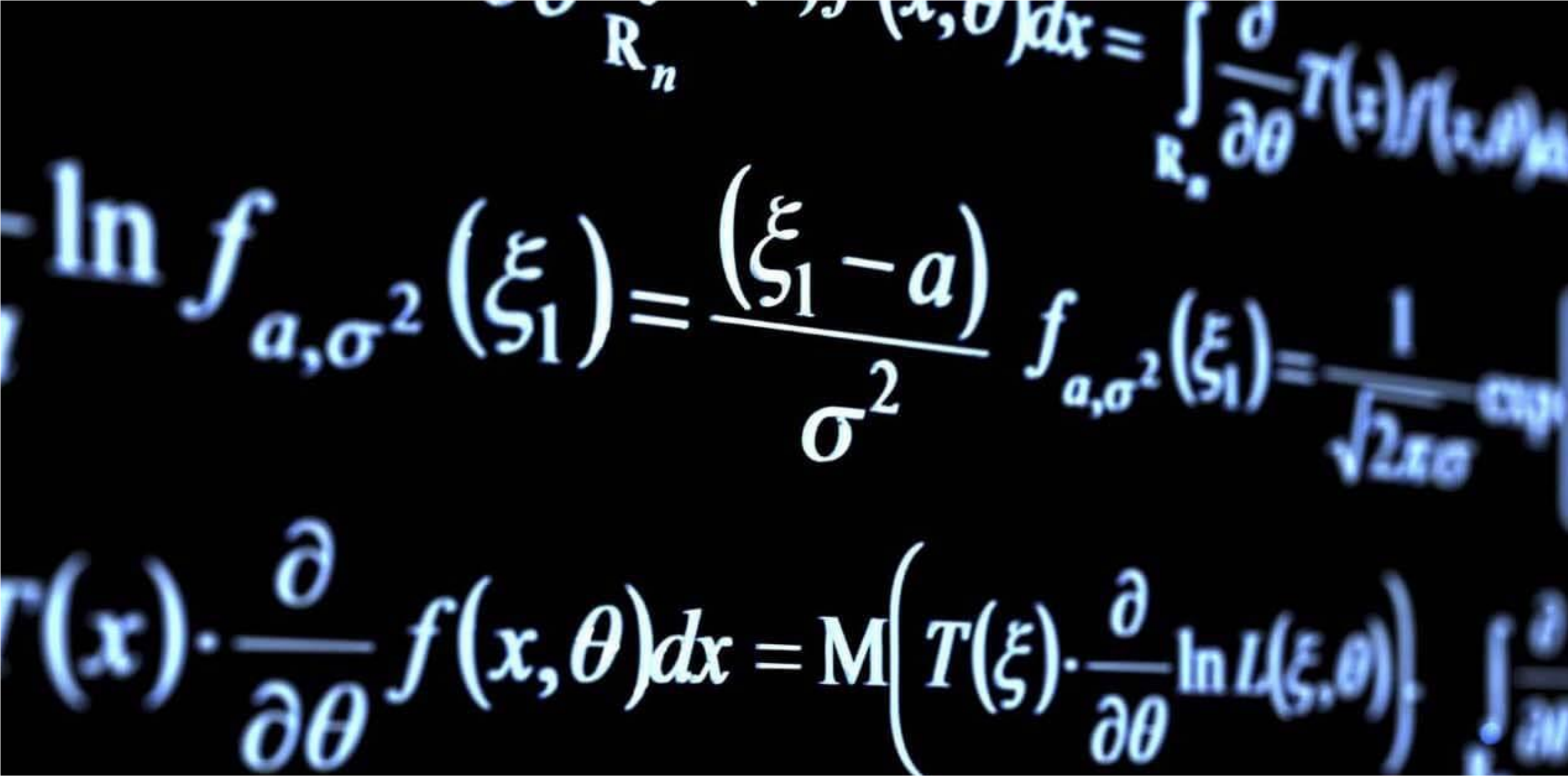
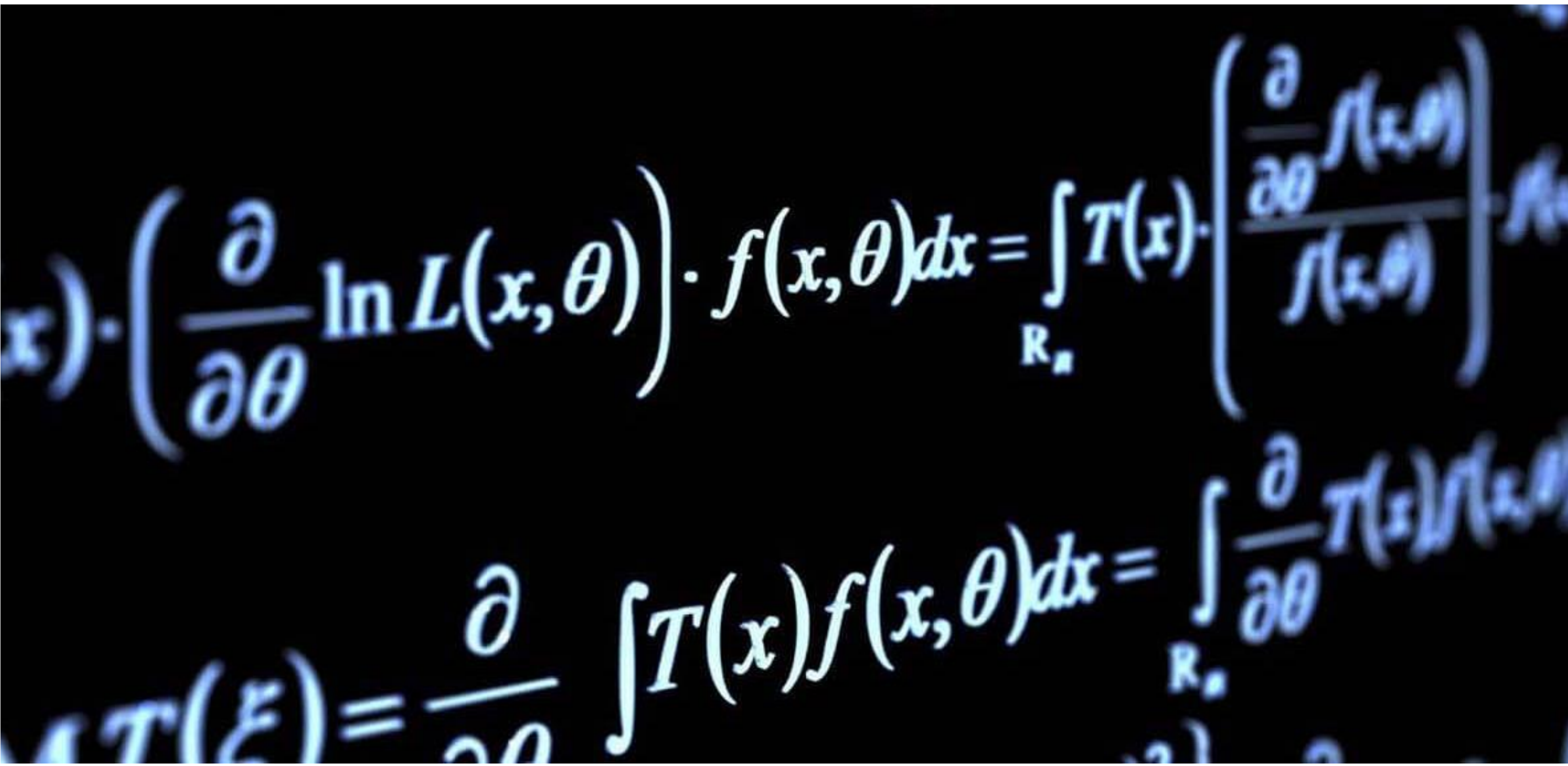


ACTSC 972 Course Notes

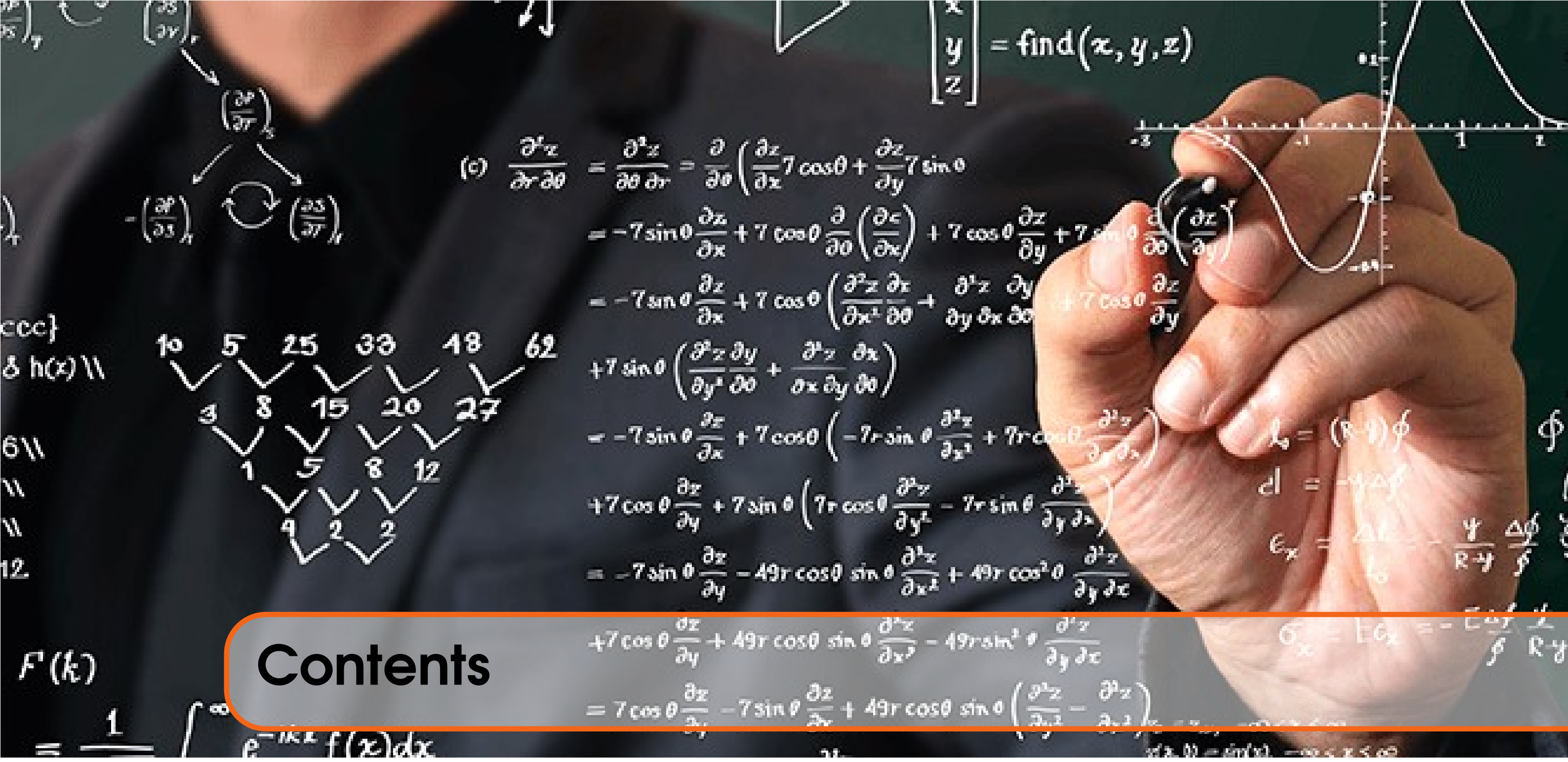
University of Waterloo

The One And Only
Waterloo 76er
Bill Zhuo



This note is for non-commercial use. Since this is not really used by a broad audience (probably just me), I will only include the necessary key points and figures. Lots of examples might be omitted. Additional commentaries from an industry perspective might be added. Use this at one's own discretion. #NotFinancialAdvise

September 2021



Contents

1	**Preliminaries**	**5**
1.1	**Prologue**	**5**
1.1.1	Motivations	5
1.1.2	Digression: Functions of Bounded Variation	6
1.2	**Ito Calculus**	**11**
1.2.1	Partitions	11
1.2.2	Quadratic Variation	11
1.4.1	The Basic Ito Formula	13
1.4.2	Covariation	14
1.6	**The Multidimensional Ito Formula**	**17**
2	**Volatility**	**21**
2.1	**Options**	**21**
2.1.1	Payoff Replication	21
2.1.2	Implied Volatility vs. Realized Volatility	23
2.2	**Variance Swaps and the VIX**	**24**
2.2.1	Quadratic Variation Approximation on VS	24
2.3	**Local Volatility and Dupire's Formula**	**26**
2.4.1	The Dupire Formula	28
2.5	**The Heston Model**	**29**
2.5.1	The CIR Process	29
2.5.2	The Hestion Model as an Affine Diffusion	32
3	**Dynamic Portfolio Strategies**	**33**
3.1	**Volatility Drag**	**33**

3.2 Leveraged ETFs 36
3.3 CPPI 37
3.4 Equal-weighted Portfolios 37
3.5 Universal Portfolios 37
3.6 Stochastic Portfolio Theory 37

4 Price Impact and Execution 39
4.1 Optimal Trade Execution 39
4.2 Pegged Currency Markets 39
4.3 Gamma Squeezes 39
4.4 Predatory Trading 39
4.5 Dark Pools 39

5 Term Structure Modelling 41
5.1 Volatility Term Structure Modelling 41
5.2 Interest Rate Term Structure Modelling 41
5.3 Libor Market Models 41

1. Preliminaries

1.1 Prologue

1.1.1 Motivations

Why continuous-time modelling?

Asset prices are discrete quantities by definition, so why do we bother modelling them as continuous-time models? For example, future markets trade at various discrete tick sizes. The main reasons are:

1. Continuity is important for the use of calculus to find underlying relationships.
2. Moreover, discrete-time binomial asset pricing model (like Cox-Ross-Rubinstein model) become computationally extremely infeasible. You are looking at $\mathscr{O}(2^T)$ complexity here where T grows with the time horizon.

What price to pay?

Reality is indeed discrete. This idealization might be exploitable by arbitrage theory.

Is it possible to model underlying by continuous path?

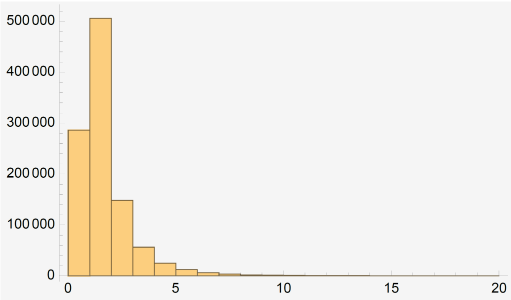
Sure if you zoom out, not so sure if you zoom in. Moreover, different underlyings have different trading time. Overnight trading and extended trading hour might lead to occasional **gap risk**. Simply put, it is the risk that you were happy when you left the office and got fired the next morning.

R So who are those people doing extended trading and overnight trading?

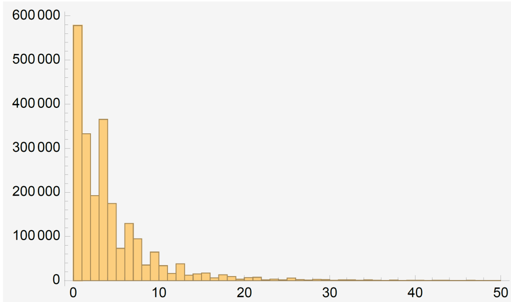
Other things to consider? Sure, **liquidity** in the market might also be a problem. In the most liquid forex market, the tick band is very tight when it is calm, thus, more like a continuous process. But certain event (press release) might move it drastically instantly, which acts more like a discrete process.

How big are the gap risks?

Generally quite small if you look at historical record and plot them on histograms but illiquidity needs to be bared in mind.



(a) Histogram of price difference ticks (tick size \$0.25) in one-minute ES futures data from Nov 2010 through Jan 2015. 13



(b) Histogram of price difference ticks (tick size \$0.01) in one-minute IBM stock price data from Feb 1998 to Nov 2019. (Some large gaps occur during after-hours trading and are probably due to illiquidity).

So let's do continuous-time modelling

Based on our discussion here. This seems viable.

1.1.2 Digression: Functions of Bounded Variation

Simple question to ask: why Brownian motion is used often to model asset prices? But not using general functions of bounded variation, or BV-function.

BV-Functions

Proposition 1.1.1 Suppose that $G : [0,T] \to [0,\infty)$ is right-continuous and non-decreasing. Then there exists a positive finite Borel measure μ on $[0,T]$ such that $G(t) = \mu([0,t])$.

R Such a function need not be differentiable. For example, every analyst's favorite staircase, the Cantor function, is a valid G as mentioned above but it is not differentiable.

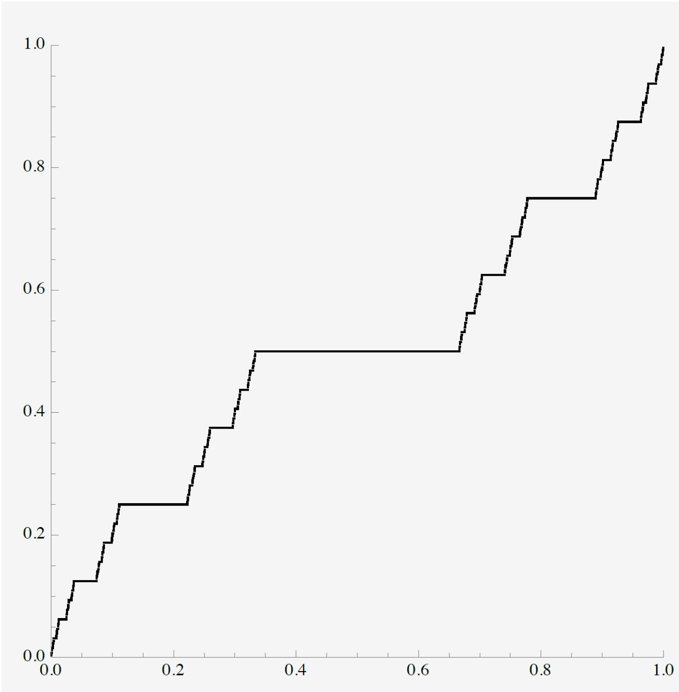


Figure 1.1.2: Devil's staircase

Definition 1.1.1 — BV-function. A right-continuous function $G:[0,T]\to\mathbb{R}$ is called a function of bounded variation or BV-function, if it can be represented as the difference

$$G(t)=G^{+}(t)-G^{-}(t)$$

of two non-decreasing right-continuous function G^{+} and G^{-}.

Let $G(t)=G^{+}(t)-G^{-}(t)$ be as above and let $\mu^{\pm}$ be non-negative Borel measures on $[0,T]$ such that $\mu^{\pm}([0,t])=G^{\pm}(t)$. Then we can define the **signed measure**

$$\mu:\mu^{+}-\mu^{-}$$

and get

$$G(t)=\mu([0,t])$$

Conversely, any function that canb e represented in this form by means of a finite signed Borel measure μ is of bounded variation.

R How to prove the finite case here? Can we do countable? Uncountable?

In particular, every continuously differentiable function G is of bounded variation, because it can be represented as $G=G^{+}-G^{-}$ for

$$G^{\pm}(t):=\int_{0}^{t}(G'(s))^{\pm}ds$$

The following theorem explains the notion "bounded variation".

Theorem 1 A right-continuous function $G:[0,T]\to\mathbb{R}$ is of bounded variation, if and only if

$$\mathsf{variation}(G) := \sup_n \sum_{i=1}^{N_n} \left|G(s_i^n)-G(s_{i-1}^n)\right| < \infty$$

where $0=s_0^n<s_1^n<\cdots<s_{N_n}^n=T$ and $\max_i \left|s_i^n-s_{i-1}^n\right|\to 0$. The quantity $\mathsf{variation}(G)$ is called the **total** variation of G.

Proof. 1. ($\Longrightarrow$): if G is non-decreasing, then

$$\sum_{i=1}^{N_n}\left|G(s_i^n)-G_{i-1}^n\right| = \sum_{i=1}^{N_n}\left(G(s_i^n)-G_{i-1}^n\right) = G(s_{N_n}^n)-G(s_0^n) = G(T)-G(0)$$

hence, the $\mathsf{variation}(G)=G(T)-G(0)$. Hence, if $G=G^+-G^-$ with $G^\pm$ being non-decreasing, then

$$\begin{aligned}\mathsf{variation}(G) &\le \mathsf{variation}(G^+)+\mathsf{variation}(G^-)\\ &= G^+(T)-G^+(0)+G^-(T)-G^-(0)<\infty\end{aligned}$$

2. ($\Longleftarrow$): left as an exercise

■

If we want to integrate only over a subinternal $[0,t)$ or $[0,t]\subseteq[0,T]$, then we write

$$\int_{[0,t]} f(s)dG(s) = \int f\mathbf{1}_{[0,t]}d\mu \text{ or } \int_{[0,t)} f(s)dG(s) = \int f\mathbf{1}_{[0,t)}d\mu$$

The function G is continuous if and only if $\mu(\{x\})=0,\forall x$ (good exercise). In this case, the two integrals above do coincide and we can write

$$\int_0^t f(s)dG(s)$$

From measure theory, we have the following nice result.

Theorem 2 If $f:[0,T]\to\mathbb{R}$ is continuous, then $\int_{[0,t]} f(s)dG(s)$ exists as a limit of Riemann sums:

$$\int_{[0,t]} f(s)dG(s) = \lim_{n\uparrow\infty}\sum_{s_i^n\le t} f(\tau_i^n)\left(G(s_i^n)-G(s_{i-1}^n)\right)$$

where $0=s_0^n<s_1^n<\cdots<s_{N_n}^n=T$, $\max_i|s_i^n-s_{i-1}^n|\to 0$, and $s_{i-1}^n\le\tau_i^n\le s_i^n$. That is, for continuous integrands, the Lebesgue-Stieltjes coincides with the Riemann-Stieltjes integral.

A good way to identify BV-function is given below.

Theorem 3 Suppose that $G:[0,T]\to\mathbb{R}$ is any function for which the Riemann-Stieltjes integral exists for all continuous integrands. Then G is a BV-function.

Theorem 4 — FTC for Stieltjes Integrals. Suppose that $f:\mathbb{R}\to\mathbb{R}$ is continuously differentiable and G is a continuous BV-function. Then,

$$f(G(t))-f(G(0)) = \int_0^t f'(G(s))dG(s)$$

Proof. For $0 = s_0 < s_0 < \cdots < S_N = T$, we use the mean value theorem to write

$$f(G(t)) - f(G(0)) = \sum_{i=1}^{N} (f(G(s_i)) - f(G(s_{i-1}))) \\ = \sum_{i=1}^{N} f'(\xi_i)(G(s_i) - G(s_{i-1}))$$

where ξ_i is between $G(s_i) - G(s_{i-1})$. Representing ξ_i as $G(\tau_i)$ for some $\tau_i \in [s_{i-1}, s_i]$, we get

$$f(G(t)) - f(G(0)) = \sum_{i=1}^{N} f'(G(\tau_i))(G(s_i) - G(s_{i-1}))$$

now take $N \to \infty$ and shrink the mesh to zero, we have the result. ∎

How should one model asset prices in continuous time?

We consider a simple setup:

1. A risk less bond $B_t = 1, \forall t$
2. A risky asset modelled by continuous path $S_t, \forall t \geq 0$

The trader can at time

$$0 = t_0 < t_1 < \cdots < t_N = T$$

trading is described by a dynamic trading strategy

$$\{\xi_{t_i}, \eta_{t_i}\}_{i=1}^{N-1}$$

one for risky and one for bond during trading period from t_i to t_{i+1}. The portfolio value is given by

$$V_{t_i} := \xi_{t_i} S_{t_i} + \eta_{t_i}$$

we call a trading strategy **self-financing** if

$$\xi_{t_i} S_{t_{i+1}} - \eta_{t_i} = \xi_{t_{i+1}} S_{t_{i+1}} + \eta_{t_{i+1}} = V_{t_{i+1}},\ i = 0, \cdots, N-1$$

intuitively, this requirement means tat the portfolio s always rearranged in such a way that its present value is preserved. Another more precise way to look at it is the following.

$$V_{t_i} - V_0 = \sum_{k=1}^{i} \xi_{t_{k-1}} (S_{t_k} - S_{t_{k-1}})$$

This really resembles a martingale transform. Now, we don't want any indice but continuous-time trading. The self-financing condition can be derived as

$$V_{t_i} - V_0 = \int_0^t \xi_s dS_s,\ 0 \leq t \leq T$$

R What is the technicality here for the convergence?

We assume $t \mapsto S_t$ is a BV-function. If f is differentiable, by FTC, we have

$$f(S_t) - f(S_0) = \int_0^t f'(S_s) dS_s$$

this gives us a simple way to derive (ξ_t, η_t) and to check whether there exists arbitrage opportunity. Unfortunately, they do exist using this approach.

■ **Example 1.1** Consider $f(x) = (x - S_0)^2$. Then we obtain

$$(S_t - S_0)^2 - 0 = \int_0^t 2(S_s - S_0)dS_s, \forall t \geq 0$$

the value process of a strategy (ξ, η) is given by $V_t = \xi_t S_t + \eta_t$, and the strategy is self-financing if and only if $V_t = V_0 + \int_0^t \xi_s dS_s$ for $0 \leq 0 \leq T$. An obvious choice here is

$$\begin{cases} \xi_t = 2(S_t - S_0) \\ \eta_t = \int_0^t 2(S_s - S_0)dS_s - 2(S_t - S_0)S_t \end{cases}$$

but after integration, we note that

$$V_t = (S_t - S_0)^2 \geq 0 = V_0, \forall t$$

This is a risk-free investment that makes positive profit as soon as $S_t \neq S_0$, which yields an arbitrage opportunity. ■

R When S is modeled by a BV-function, then our trading strategy yields an arbitrage opportunity. Hence, it follows that it is not possible to choose S as a BV-function. FTC and non-arbitrage theory seem to be contradictory at this point. And we need non-classical integral calculus.

Proposition 1.1.2 — Motivation for Brownian Motion. Let $t \mapsto S_t$ be a continuous function on $[0, T]$ and let

$$\xi_t = 2(S_t - K),\ 0 \leq t \leq T$$

where $K \in \mathbb{R}$ is given. Then $\int_0^t \xi_t dS_t$ exists for all t as the limit of the corresponding Riemann sums, i.e.,

$$\int_0^t \xi_s dS_s = \lim_{n\uparrow\infty} \sum_{t_k^n \leq t} \xi_{t_k^n}\left(S_{t_{k+1}^n} - S_{t_k^n}\right)$$

if and only if the **quadratic variation of** S,

$$\langle S \rangle_t := \lim_{n\uparrow\infty} \sum_{t_k^n \leq t} \left(S_{t_{k+1}^n} - S_{t_k^n}\right)^2$$

exists for all t. In this case, it follows that

$$\int_0^t \xi_s dS_s = (S_t - K)^2 - (S_0 - K)^2 - \langle S \rangle_t$$

Proof. We write

$$\begin{aligned} \xi_{t_k^n}\left(S_{t_{k+1}^n} - S_{t_k^n}\right) &= 2(S_{t_k^n} - K)(S_{t_{k+1}^n} - S_{t_k^n}) \\ &= \left(S_{t_{k+1}^n} - K\right)^2 - (S_{t_k^n} - K)^2 - \left(S_{t_{k+1}^n} - S_{t_k^n}\right)^2 \end{aligned}$$

sum over the mesh yields

$$\sum_{t_k^n \leq t} \xi_{t_k^n}\left(S_{t_{k+1}^n} - S_{t_k^n}\right) = (S_{t_N} - K)^2 - (S_0 - K)^2 - \sum_{t_k^n \leq t} \left(S_{t_{k+1}^n} - S_{t_k^n}\right)^2$$

where $t_N = \max\left\{t_{k+1}^n : t_k^n \leq t\right\} \searrow t$ as $n \uparrow \infty$. Clearly, the claim is true. ■

R It follows that we must impose the existence of the quadratic variation $\langle S\rangle$ so as to be able to define some basic self-financing trading strategies in continuous time, including those of the $\xi_t = 2(S_t - K)$. Moreover, we do have consistency when $A : [0,T] \to \mathbb{R}$ is a continuous BV-function, then its quadratic variation $\langle A\rangle$ vanishes. Since A is a BV-function, $\mathsf{variation}(A) < \infty$, then

$$\begin{aligned}
\langle A\rangle_t &= \lim_{n\uparrow\infty} \sum_{t_k^n \le t} \left(A_{t_{k+1}^n} - A_{t_k^n}\right)^2 \\
&= \lim_{n\uparrow\infty} \sum_{t_k^n \le t} \left|A_{t_{k+1}^n} - A_{t_k^n}\right| \cdot \left|A_{t_{k+1}^n} - A_{t_k^n}\right| \\
&\le \lim_{n\uparrow\infty} \max_l \left|A_{t_{l+1}^n} - A_{t_l^n}\right| \cdot \sum_{t_k^n \le t} \left|A_{t_{k+1}^n} - A_{t_k^n}\right| \\
&\le \lim_{n\uparrow\infty} \underbrace{\max_l \left|A_{t_{l+1}^n} - A_{t_l^n}\right|}_{\to 0} \cdot \underbrace{\mathsf{variation}(A)}_{<\infty} = 0
\end{aligned}$$

1.2 Ito Calculus

1.2.1 Partitions

For $T > 0$, a partition of the interval $[0,T]$ is a finite set

$$\mathbb{T} := \{t_0, t_1, \cdots, t_n\}$$

such that $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_n = T$. If a partition $\mathbb{T}$ of the interval $[0,T]$ has been fixed, we let t' be the successor for any $t \in \mathbb{T}$. The **mesh** of $\mathbb{T}$ is defined as

$$\mathsf{mesh}(\mathbb{T}) := \max_{t\in\mathbb{T}} \left|t' - t\right|$$

a sequence $(\mathbb{T}_n)_{n\in\mathbb{N}}$ of partitions of $[0,T]$ is called a refining sequence of partitions of $[0,T]$ if $\mathbb{T}_1 \subseteq \mathbb{T}_2 \subseteq \cdots$ and $\lim_n \mathsf{mesh}(\mathbb{T}_n) = 0$.

■ **Example 1.2 — Dyadic partitions.** The sequence of dyadic partitions of $[0,1]$ with $\mathbb{T}_n := \{k2^{-n} : k = 0, \cdots, 2^n\}$. ■

1.2.2 Quadratic Variation

Let $T > 0$ and $(\mathbb{T}_n)_n$ be a partition.

Definition 1.2.1 — Quadratic Variation. A continuous function $X \in C[0,T]$ admits a continuous quadratic variation $\langle X\rangle$ along the sequence $(\mathbb{T}_n)_n$, if for each $t \in [0,T]$ the following limit exists,

$$\langle X\rangle_t := \lim_{n\uparrow\infty} \sum_{s\in\mathbb{T}_n, s\le t} (X_{s'} - X_s)^2$$

and if the function $t \mapsto \langle X\rangle_t$ is continuous on $[0,T]$. The class of all $X \in C[0,T]$ that admit a continuous quadratic variation along the sequence $(\mathbb{T}_n)_n$ will be denoted by $QV[0,T]$.

R Important differences between total variation and quadratic variation:
1. The quadratic variation might not exist
2. If the quadratic variation exists, it may depends of the choice of $(\mathbb{T}_n)_n$
3. The space $QV[0,T]$ depends on $(\mathbb{T}_n)_n$ and is typically not a vector space

Lemma 1.3 For $X \in QV[0,T]$, the function $x \mapsto \langle X\rangle_t$ is non-decreasing.

Proof. For $s < t$ we have

$$\begin{aligned}\langle X\rangle_t - \langle X\rangle_s &= \lim_{n\uparrow\infty}\left(\sum_{r\in\mathbb{T}_n, r\le t}(X_{r'}-X_r)^2 - \sum_{r\in\mathbb{T}_n, r\le s}(X_{r'}-X_r)^2\right)\\ &= \lim_{n\uparrow\infty}\sum_{r\in\mathbb{T}_n, s<r\le t}(X_{r'}-X_r)^2 \ge 0\end{aligned}$$

■

This implies that, for every $X \in QV[0,T]$ and each function $f \in C[0,T]$, we can define the Riemann-Stieltjes integral

$$\int_0^t f(s)d\langle X\rangle_s$$

since $\langle X\rangle_s$ is non-decreasing and $d\langle X\rangle_s$ must be of BV. It can be written as

$$\int_0^t f(s)(dX_s)^2$$

why written like this?

Lemma 1.4 Suppose $X \in QV[0,T]$ and $f \in C[0,T]$. Then, for all choices of intermediate time $\tau_s^n \in [s,s'], s \in \mathbb{T}_n$,

$$\sum_{s\in\mathbb{T}_n, s\le t} f(\tau_s^n)\underbrace{(X_{s'}-X_s)^2}_{\text{“}(dX_s)^2\text{”}} \xrightarrow{n\to\infty} \int_0^t f(s)d\langle X\rangle_s$$

Proof. Define a sequence of positive finite measures:

$$\mu_n := \sum_{s\in\mathbb{T}_n}(X_{s'}-X_s)^2\delta_{\tau_s^n}$$

then,

$$\mu_n([0,t]) = \sum_{s\in\mathbb{T}_n}(X_{s'}-X_s)^2\mathbf{1}_{[0,t]}(\tau_s^n) = \sum_{s\in\mathbb{T}_n, \tau_s^n\le t}(X_{s'}-X_s)^2 \xrightarrow{n\to\infty} \langle X\rangle_t$$

Since the CDFs $\mu_n \to \mu = \langle X\rangle$ pointwisely with respect to t. This is a weak convergence. Then,

$$\sum_{s\in\mathbb{T}_n, \tau_s^n\le t} f(\tau_s^n)(X_{s'}-X_s)^2 = \int f\mathbf{1}_{[0,1]}d\mu_n \to \infty f\mathbf{1}_{[0,1]}d\mu = \int_0^t f(s)d\langle X\rangle_s$$

■

Proposition 1.4.1 For $X \in QV[0,T]$ and $f \in C^1(\mathbb{R})$, the function

$$Y_t := f(X_t)$$

belongs to $QV[0,T]$ and has the quadratic variation

$$\langle Y\rangle_t = \int_0^t (f'(X_s))^2 d\langle X\rangle_s,\ 0\le t\le T$$

Proof. For $s \in \mathbb{T}_n$, the mean value theorem of calculus yields a real number ξ between $X_s, X_{s'}$ such that

$$f(X_{s'}) - f(X_s) = f'(\xi)(X_{s'}-X_s)$$

Since $X \in QV[0,T]$, we know that $X \in C[0,T]$ at least. Then, by intermediate value theorem, there exists $\tau_s \in [s,s'] \subseteq [0,T]$ such that $\xi = X_{\tau_s}$. Then,

$$\sum_{s\in\mathbb{T}_n, s\le t} (f(X_{s'}) - f(X_s))^2 = \sum_{s\in\mathbb{T}_n, s\le t} \left(f'(X_{\tau_s})\right)^2 (X_{s'} - X_s)^2$$

By Lemma 1.4, we have that

$$\langle Y\rangle_t = \lim_{n\uparrow\infty} \sum_{s\in\mathbb{T}_n, s\le t} \left(f'(X_{\tau_s})\right)^2 (X_{s'} - X_s)^2 = \int_0^t (f'(X_s))^2 d\langle X\rangle_s$$

■

R If M is a continuous martingale on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), P)$, then $P-a.s$ every trajectory of M admits a continuous quadratic varaition along $(\mathbb{T}_n)_n$. In the special case of Brownian motion W and $\mathbb{T}_n$ being the dyadic partition. This can be seen easily. Consider

$$\sum_{k=1}^{2^n} \left(W_{k2^{-n}} - W_{(k-1)2^{-n}}\right)^2$$

since W has independent increments, the random variables $(W_{2^{-n}} - W_0)^2, \cdots, (W_1, W_{1-2^{-n}})^2$ are independent. Moreover, each increment follows $N(0, 2^{-n})$. Then,

$$\left(W_{k2^{-n}} - W_{(k-1)2^{-n}}\right)^2 = \left(2^{-n/2} Z_k\right)^2$$

for independent standard normal random variables $Z_1, \cdots, Z_{2^n}$. Therefore,

$$\sum_{k=1}^{2^n} \left(W_{k2^{-n}} - W_{(k-1)2^{-n}}\right)^2 = \frac{1}{2^n}\sum_{k=1}^{2^n} Z_k^2$$

this converges to 1 by the law of large numbers. When changing the step size to be t, we shall have $\langle W\rangle_t = t$.

However, the martingale (or semimartingale) assumption is not needed for the existence of a quadratic variation, which is the key assumption for Ito calculus. It works with single trajectories-just as asset price evolutions are also just given as single trajectories. This is an important feature when it comes to model risk stemming from model uncertainty.

1.4.1 The Basic Ito Formula

The Ito formula is the FTC for integrators with continuous quadratic variation.

Theorem 5 — Basic Ito Formula. For $f \in C^2(\mathbb{R})$ and $X \in QV[0,T]$,

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s,\ 0\le t\le T$$

where the **Ito integral**, $\int_0^t f'(X_s)dX_s$ is given by the following limit

$$\int_0^t f'(X_s)dX_s = \lim_{n\uparrow\infty} \sum_{s\in\mathbb{T}_n, s\le t} f'(X_s)(X_{s'} - X_s)$$

Proof. By Taylor's theorem and intermediate value theorem, we know there exists $\tau_s^n \in [s,s']$ for $s \in \mathbb{T}_n$ such that

$$f(X_{s'}) - f(X_s) = f'(X_s)(X_{s'} - X_s) + \frac{1}{2} f''(X_{\tau_s^n})(X_{s'} - X_s)^2$$

if we sum both side, the left-hand telescoping sum converges to $f(X_t) - f(X_0)$ as $n \to \infty$. And Lemma 1.4 implies that

$$\frac{1}{2} \sum_{s \in \mathbb{T}_n, s \leq t} f''(X_{\tau_s^n})(X_{s'} - X_s)^2 \to \frac{1}{2} \int_0^t f''(X+s) d \langle X \rangle_s$$

thus, $\lim_n \sum_{s \in \mathbb{T}_n, s \leq t} f'(X_s)(X_{s'} - X_s)$ must exist as well and we define it to be the Ito integral. ■

R We can think of the quadratic variation integral part as a correction to the classical FTC. If the integrator X is a BV-function, then $\langle X \rangle = 0$, which makes the Ito formula coincide with the classical FTC.

■ **Example 1.3** Let $f(x) = x^2$ and $X \in QV[0,T]$. Then Theorem 5 tells us

$$\begin{aligned} X_t^2 - X_0^2 &= f(X_t) - f(X_0) \\ &= \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d \langle X \rangle_s \\ &= 2 \int_0^t X_s + \langle X \rangle_t \end{aligned}$$

■

R If M is a continuous martingale on a filtered probability space. We know that sample paths of M admit $P-a.s.$ the continuous quadratic variation $\langle M \rangle$. Hence, Ito integrals with respect to M exist, and these are (local) martingales. Moreover, these Ito integrals coincide with the corresponding stochastic integrals, which are often defined as L^2-limit of elementary integrals of simple step functions.

1.4.2 Covariation

If we fix a partition $\mathbb{T}_n$ for some $n \in \mathbb{N}$ and $X \in C[0,T]$. We write

$$\langle X \rangle_t^{(n)} := \sum_{s \in \mathbb{T}_n, s \leq t} (X_{s'} - X_s)^2, \ t \in [0,T], n \in \mathbb{N}$$

then, if $X, Y \in C[0,T]$, we have

$$\langle X+Y \rangle_t^{(n)} = \langle X \rangle_t^{(n)} + \langle Y \rangle_t^{(n)} + 2 \underbrace{\sum_{s \in \mathbb{T}_n, s \leq t} (X_{s'} - X_s)(Y_{s'} - Y_s)}_{\langle X,Y \rangle_t^{(n)}}$$

if both $X, Y \in QV[0,T]$, then $\langle X \rangle_t^{(n)} \to \langle X \rangle_t, \langle Y \rangle_t^{(n)} \to \langle Y \rangle_t$. Thus, the quadratic variation $\langle X+Y \rangle_t$ exists for all $t \in [0,T]$ if and only if the limit of

$$\langle X,Y \rangle_t^{(n)} := \sum_{s \in \mathbb{T}_n, s \leq t} (X_{s'} - X_s)(Y_{s'} - Y_s)$$

exists for all $t \in [0,T]$. This observation motivates the following definition.

Definition 1.4.1 — Covariation. Two functions $X, Y \in C[0,T]$ admit the continuous covariation $\langle X, Y \rangle$ along $(\mathbb{T}_n)_n$ if for each $t \in [0,T]$ the following limit exists

$$\langle X,Y \rangle_t := \lim_n \langle X,Y \rangle_t^{(n)}$$

and if the function $t \mapsto \langle X,Y\rangle_t$ is continuous.

R Note that for $X \in QV[0,T]$, $\langle X,X\rangle = \langle X\rangle$.

Proposition 1.4.2 For $X,Y \in QV[0,T]$, the following two conditions are equivalent

1. $X+Y \in QV[0,T]$
2. X,Y admit the continuous covariation $\langle X,Y\rangle$ along $(\mathbb{T}_n)_n$

Moreover, if these equivalent conditions are satisfied, then the following polarization identity,

$$\langle X,Y\rangle_t = \frac{1}{2}(\langle X+Y\rangle_t - \langle X\rangle_t - \langle Y\rangle_t)$$

and the following Cauchy-Schwarz inequality hold,

$$|\langle X,Y\rangle_t| \leq \sqrt{\langle X\rangle_t \langle Y\rangle_t}$$

Proof. The equivalence of 1. and 2. are immediate based on covariation motivation. Say these two are satisfied, we take $n \to \infty$ to get

$$\langle X+Y\rangle_t = \langle X\rangle_t + \langle Y\rangle_t + 2\langle X,Y\rangle_t$$

thus, we have the result. Now, for the inequality, it is just a simple Cauchy-Schwarz

$$\left|\langle X,Y\rangle_t^{(n)}\right| = \left|\sum_{s\in\mathbb{T}_n, s\leq t}(X_{s'}-X_s)(Y_{s'}-Y_s)\right| \leq \left\{\langle X\rangle_t^{(n)}\langle Y\rangle_t^{(n)}\right\}$$

then take $n \to \infty$, we have the result by comparison theorem. ■

Proposition 1.4.3 Suppose that $X,A \in QV[0,T]$ and that $\langle A\rangle_T = 0$. Then $X+A \in QV[0,T]$ and we have

$$\langle X+A\rangle_t = \langle X\rangle_t \quad \text{and} \quad \langle X,A\rangle_t = 0, \forall t$$

Proof. By CS-inequality, we have

$$\left|\langle X,A\rangle_t^{(n)}\right| \leq \sqrt{\langle X\rangle_t^{(n)}\langle A\rangle_t^{(n)}} \leq \sqrt{\langle X\rangle_T^{(n)}\langle A\rangle_T^{(n)}}$$

note that $\langle X\rangle_t^{(n)}$ is bounded and $\langle A\rangle_T^{(n)}$ converges to 0 as $n \to \infty$. Thus, $\langle X,A\rangle_t = 0, \forall t$. By the polarization identity, we have

$$\langle X+A\rangle_t^{(n)} \xrightarrow{n\to\infty} \langle X\rangle_t$$

■

Proposition 1.4.4 If $X \in QV[0,T]$ and g is continuously differentiable, then the Ito integral $\int_0^t g(X_s)dX_s$ exists. Moreover, as a function of t it belongs if $QV[0,T]$ and

$$\left\langle \int_0^{\cdot} g(X_s)dX_s \right\rangle_t = \int_0^t (g(X_s))^2 d\langle X\rangle_s$$

Proof. Let G be the antiderivative of g. By Ito's formula, we have

$$G(X_t) - G(X_0) = \int_0^t g(X_s)dX_s + \frac{1}{2}\int_0^t g'(X_s)d\langle X\rangle_s$$

since $X \in QV[0,T]$, we let $A_t = \int_0^t g'(X_s) d\langle X \rangle_s$ and it exists. Since both $G(X_t), G(X_0)$ exists. Our Ito integral $Y_t = \int_0^t g(X_s) dX_s$ must also exist. Then, note that A_t is of BV (not hard to show given that g' is continuous). Then, by Proposition 1.4.3, we know that

$$\langle Y_t \rangle = \langle G(X) \rangle_t = \int_0^t (G'(X_s))^2 d\langle X \rangle_s = \int_0^t (g(X_s))^2 d\langle X \rangle_s$$

where the last equation comes from Proposition 1.4.1 that G is clearly at least C^1. ■

In the sequel, we will work with d-dimensional Euclidean space $\mathbb{R}^d$ equipped with the standard Euclidean norm and vector-valued functions. The vector-valued functions will be denoted in the following way:

$$\begin{aligned} \mathbf{X} &: [0,T] \to \mathbb{R}^d \\ t &\mapsto \mathbf{X}_t = \left(X_t^1, \cdots, X_t^d\right)^\top \end{aligned}$$

if each component X^i of $\mathbf{X}$ is continuous, $\mathbf{X} \in C([0,T], \mathbb{R}^d)$.

Definition 1.4.2 For $d \in \mathbb{N}$, we denote $QV^d[0,T]$ the space of all functions $\mathbf{X} = (X^1, \cdots, X^d)^\top \in C([0,T], \mathbb{R}^d)$ for which $X^i + X^j \in QV[0,T]$ for $i, j = 1, \cdots, d$.

It is not hard to see that $QV^d[0,T]$ consists precisely of those functions $\mathbf{X} \in C([0,T], \mathbb{R}^d)$ whose components $X^1, \cdots, X^d$ admit continuous quadratic variations and covariations along $(\mathbb{T}_n)_n$.

Exercise 1.1 Show that the covariation $\langle \cdot, \cdot \rangle$ has the following "inner product" like properties. For $(X,Y,Z)^\top \in QV^3[0,T]$,

1. $\langle X, X \rangle = \langle X \rangle \geq 0$
2. $\langle X, Y \rangle = \langle Y, X \rangle$
3. $\langle X, Y+Z \rangle = \langle X, Z \rangle + \langle X, Z \rangle$
4. For $a, b \in \mathbb{R}$, we have $aX + bY \in QV[0,T]$ and $\langle aX, bY \rangle = ab \langle X, Y \rangle$

■

R Why not a inner product space? Note that a quadratic variation of 0 does not mean the associated process is zero. It could be any process of continuous bounded variation.

Proposition 1.4.5 For $(X,Y)^\top \in QV^2[0,T]$, the functions $\langle X \rangle, \langle Y \rangle$, and $\langle X, Y \rangle$ are continuous BV function.

Proof. This is almost immediate. We know that $\langle X \rangle, \langle Y \rangle, \langle X+Y \rangle$ are non-decreasing and continuous functions and so they are BV functions. Moreover, $\langle X, Y \rangle$ is the difference of two non-decreasing continuous functions by the the polarization identity. Hence, it is also a BV function by the original definition. ■

This proposition implies that $\langle X, Y \rangle$ is also an admissible integrator for Riemann-Stieljes integral. Using the polarization identity, for any $f \in C[0,T]$, we have

$$\int_0^t f(s) d\langle X, Y \rangle_s = \frac{1}{2} \left(\int_0^t f(s) d\langle X+Y \rangle_s - \int_0^t f(s) d\langle X \rangle_s - \int_0^t f(s) d\langle Y \rangle_s \right)$$

Lemma 1.5 Suppose that $(X,Y)^\top \in QV^2[0,T]$ and $g \in C[0,T]$. Then, for all choices of intermediate times $\tau_s^n \in [s, s'], s \in \mathbb{T}_n$,

$$\sum_{s \in \mathbb{T}_n, s \leq t} g(\tau_s^n)(X_{s'} - X_s)(Y_{s'} - Y_s) \xrightarrow{n \to \infty} \int_0^t g(s) d\langle X, Y \rangle_s$$

The preceding lemma motivates the alternative notation

$$\int_0^t g(s)d\langle X,Y\rangle_s = \int_0^t g(s)dX_s dY_s$$

Proposition 1.5.1 Let $\mathbf{X} \in QV^d[0,T]$, U be an open subset of $\mathbb{R}^d$ such that $\mathbf{X}_t \in U$ for all $t \in [0,T]$ and $f \in C^1(U)$. Then the function $Y_t = f(\mathbf{X}_t)$ belongs to $QV[0,T]$ and its quadratic varaition along $(\mathbb{T}_n)_n$ is given by

$$\langle Y\rangle_t = \sum_{k,l=1}^d \int_0^t f_{x^k}(\mathbf{X}_t) f_{x^l}(\mathbf{X}_s) d\left\langle X^k, X^l\right\rangle_s$$

Exercise 1.2 In the situation of the preceding proposition, let $f,g \in C^1(U)$ and define $Y_t := f(\mathbf{X}_t)$ and $Z_t := g(\mathbf{X}_t)$. Show that $(Y,Z)^\top \in QV^2[0,T]$ and compute $\langle Y,Z\rangle_t$. ■

Exercise 1.3 For $(X,Y)^\top \in QV^2[0,T]$ show the following extension of the CS-type inequality: for $f,g \in C[0,T]$ and $t \in [0,T]$,

$$\left|\int_0^t f(s)g(s)d\langle X,Y\rangle_s\right| \le \sqrt{\int_0^t (f(s))^2 d\langle X\rangle_s} \cdot \sqrt{\int_0^t (g(s))^2 d\langle Y\rangle_s}$$

Note: this is a pathwise version of the so-called Kunita-Watanabe inequality. ■

1.6 The Multidimensional Ito Formula

Theorem 6 — Multidimensional Ito Formula. Let $U \subseteq \mathbb{R}^d$ be open and $f \in C^2(U)$, and $\mathbf{X} \in QV^d[0,T]$ be such that $\mathbf{X}_t \in U, \forall t \in [0,T]$. Then, for $t \in [0,T]$,

$$f(\mathbf{X}_t) - f(\mathbf{X}_0) = \int_0^t \nabla_{\mathbf{x}} f(\mathbf{X}_s) d\mathbf{X}_s + \frac{1}{2}\sum_{k,l=1}^d \int_0^t f_{x^k x^l}(\mathbf{X}_s) d\left\langle X^k, X^l\right\rangle_s$$

where $\int_0^t f_{x^k x^l}(\mathbf{X}_s) d\left\langle X^k, X^l\right\rangle_s$ is taken in the sense of Stieltjes integrals and the pathwise Ito integral is given as the following limit of nonanticipating Riemann sums,

$$\int_0^t \nabla_{\mathbf{x}} f(\mathbf{X}_s) d\mathbf{X}_s = \lim_{n\uparrow\infty} \sum_{s\in\mathbb{T}_n, s\le t} \nabla_{\mathbf{x}} f(\mathbf{X}_s)(\mathbf{X}_{s'} - \mathbf{X}_s)$$

whose convergence is uniform in $t \in [0,T]$.

R In the context of the preceding theorem, it may be tempting to write

$$\int_0^t \nabla_{\mathbf{x}} f(\mathbf{X}_s) d\mathbf{X}_s = \sum_{k=1}^d \int_0^t f_{x^k}(\mathbf{X}_s) dX_s^k$$

However, it is not clear whether the integrals on the right-hand side exist individually as pathwise Ito integrals. This is one of the drawbacks of pathwise Ito calculus. On the other hand, for stochastic integrals, as defined for martingale or semimartingale integrators, such a decomposition is always possible.

Corollary 1.6.1 — Ito's Product Rule. For $\begin{bmatrix} Y \\ Z \end{bmatrix} \in QV^2[0,T]$,

$$Y_tZ_t - Y_0Z_0 = \int_0^t [Z_s, Y_s] d \begin{bmatrix} Y_s \\ Z_s \end{bmatrix} + \langle Y, Z \rangle_t$$

Proof. We apply the 2-dimensional Ito formula to

$$\mathbf{X}_t := \begin{bmatrix} Y_t \\ Z_t \end{bmatrix}$$

and the smooth function $f(y,z) = yz$. Then,

$$\nabla f(y,z) = [z,y], \ f_{yy} = f_{zz} = 0, \text{ and } f_{yz} = f_{zy} = 1$$

plug into Theorem 6, we have the result. ■

R In the preceding corollary, it might be tempting to write

$$Y_tZ_t - Y_0Z_0 = \int_0^t Z_s dY_s + \int_0^t Y_s dZ_s + \langle Y, Z \rangle_t$$

this indeed resembles the classical integration-by-parts and which holds in standard stochastic calculus. In pathwise Ito calculus, these two integrals on the right hand side might not exist individually. However, once one component, say Z is a BV function, then $\int_0^t Y_s dZ_s$ is well-defined as Stieltjes integral, then the other component is also well-defined. In this case, $\langle Y, Z \rangle$ vanishes, so the formula simplifies to

$$Y_tZ_t - Y_0Z_0 = \int_0^t Z_s dY_s + \int_0^t Y_s dZ_s$$

Corollary 1.6.2 — Time-dependent Ito Formula. Suppose that $X \in QV[0,T]$, A is a continuous BV function, and $f(t,a,x)$ is a twice continuously differentiable function. Then

$$\begin{aligned} f(t,A_t,X_t) - f(0,A_0,X_0) = & \int_0^t f_x(s,A_s,X_s) dX_s \\ & + \int_0^t f_a(s,A_s,X_s) dA_s + \int_0^t f_t(s,A_s,X_s) ds \\ & + \frac{1}{2} \int_0^t f_{xx}(s,A_s,X_s) d\langle X \rangle_s \end{aligned}$$

■ **Example 1.4** The function

$$f(t,x) = e^{x - \frac{1}{2}t}$$

is smooth and satisfies the backward heat equation

$$f_t + \frac{1}{2} f_{xx} = 0$$

Hence, for $X \in QV[0,T]$,

$$f(\langle X \rangle_t, X_t) - f(0, X_0) = \int_0^t f_x(\langle X \rangle_t, X_s) dX_s = \int_0^t f(\langle X \rangle_s, X_s) dX_s$$

let $Z_t := f(\langle X \rangle_t, X_t)$, it follows that the pathwise Ito integral $\int_0^t Z_s dX_s$ exists and that

$$Z_t - Z_0 = \int_0^t Z_s dX_s$$

This is an example of an Ito integral equation. If $X_t = t$, this determines the standard exponential function: $Z_t = Z_0 e^t$. Therefore, $Z_t = e^{X_t - \frac{1}{2}\langle X \rangle_t}$ can be regarded as an analogue of the exponential function for Ito calculus. Therefore,

$$\mathscr{X}_t := e^{X_t - \frac{1}{2}\langle X \rangle_t},\ t \in [0,T]$$

is called the Doleans-Dade exponential of X. ■

Exercise 1.4 Suppose that $\begin{bmatrix} Y \\ Z \end{bmatrix} \in QV^2[0,T]$ is such that $Z_t > 0, \forall t$. Show that

$$\frac{Y_t}{Z_t} = \frac{Y_0}{Z_0} + \int_0^t \left[\tfrac{1}{Z_s}, -\tfrac{Y_s}{Z_s^2}\right] d\begin{bmatrix} Y_s \\ Z_s \end{bmatrix} + \int_0^t \frac{Y_s}{Z_s^3} d\langle Z \rangle_s - \int_0^t \frac{1}{Z_s^2} d\langle Y, Z \rangle_s$$

Hint: consider $f(y,z) = \frac{y}{z}$ which is C^2 almost everywhere but $(y,0), \forall y$. ■

Let $\mathbf{X} \in QV^d[0,T]$ and suppose that $\xi^{(1)}, \cdots, \xi^{(\nu)}$ are admissible integrands for $\mathbf{X}$. When setting

$$Y_t^l := \int_0^t \xi_s^{(l)} d\mathbf{X}_s,\ l = 1, \cdots, \nu$$

then $\mathbf{Y}_t = \left(Y_t^1, \cdots, Y_t^\nu\right)$ is a continuous trajectory that admits the continuous covariations

$$\left\langle Y^k, Y^l \right\rangle_t = \sum_{i,j=1}^d \int_0^t \xi_s^{(k),i} \xi_s^{(l),j} d\left\langle X^i, X^j \right\rangle_s,\ k,l = 1, \cdots, \nu$$

Therefore, $\mathbf{Y}$ belongs to $QV^\nu[0,T]$ and we can define Ito integrals with respect to $\mathbf{Y}$. The following associativity rule for the pathwise Ito integral shows that one can express a pathwise a pathwise Ito integral with respect to $\mathbf{Y}$ as a pathwise Ito integral with respect to $\mathbf{X}$.

Theorem 7 — Associativity of the pathwise Ito integral. Suppose that $\mathbf{X} \in QV^d[0,T]$, $\xi^{(1)}, \cdots, \xi^{(\nu)}$ are admissible integrands for $\mathbf{X}, \mathbf{Y}$ is defined through preceding part, and $\eta = \left(\eta^1, \cdots, \eta^\nu\right)$ is an admissible integrand for $\mathbf{Y}$. Then $\sum_{l=1}^\nu \eta^l \xi^{(l)}$ is an admissible integrand for $\mathbf{X}$ and

$$\int_0^t \eta_s d\mathbf{Y}_s = \int_0^t \sum_{l=1}^\nu \eta_s^l \xi_s^{(l)} d\mathbf{X}_s$$

2. Volatility

2.1 Options

2.1.1 Payoff Replication

We inherit standard option math notations from any previous mathematical finance course. All notations should be self-explanatory. For a plain vanilla put and call options, we primarily consider American exercise and European exercise (for people looking for Bermudan, the exit door is on the right). For example, the SPX options have a lot of different expiries and strikes. All of these are European style since you cannot directly trade the SPX index. Thus, these are cash settled at maturity. This hints that American options usually require a physical settlement for early exercise. Unless you are trading future as underlying. Most of time, it is more liquid and cost-effective to only consider OTM put and call options. I mean you can always replicate a ITM option given a OTM one and some future. But OTM option market is usually the most liquid. There could be a tons of different trading strategies with only OTM put/call options, future, and cash. In fact for a general payoff (the solid line in the graph below), one can always try to interpolate/approximate this payoff behaviour using these 4 instruments at hand.

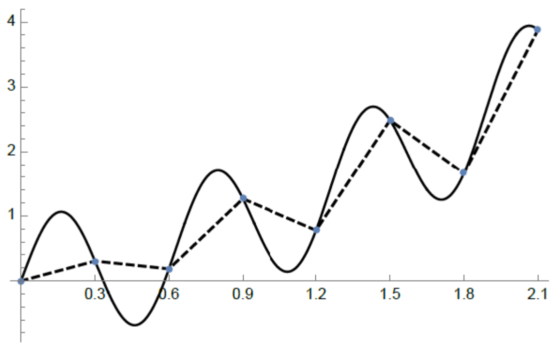


Figure 2.1.1: Target Payoff and Approximation

We aim to do this in a systematic manner. Consider a European derivative with a target payoff $h(S_T)$

at time T. We can always break up h into convex pieces and piece up the finally approximation together. Without loss of generality, suppose that $h : (0,\infty) \to \mathbb{R}$ is a convex function. Then its right-hand derivative exists and given by

$$h'_+(x) := \lim_{\delta \downarrow 0} \frac{h(x+\delta) - h(x)}{\delta}$$

exists for all $x \in (0,\infty)$, and h'_+ is a non-decreasing and right-continuous function on $(0,\infty)$. Hence, it can be represented as follows by means of a nonnegative Borel measure μ in $(0,\infty)$ given by

$$h'_+(y) - h'_+(x) = \mu((x,y]),\ 0 < x < y < \infty$$

Note that $\mu((0,\infty))$ could be infinite. But for any compact set K, we do have $\mu(K)$ being finite. Thus, this is also a Radon measure. More generally, we suppose that $h = f - g$ is the difference of two convex functions. Then, the right-hand derivative h'_+ will still exist as a right-continuous function, but not necessarily non-decreasing, but we could introduce a signed Radon measure to resolve this.

Note that every C^2 function h can be written as the difference of two convex function. Namely, we can pick

$$h^+(x) = \int_1^x (h''(y))^+ dy,\ \ h^-(x) = \int_1^x (h''(y))^- dy$$

where $h' = h^+ - h^-$. Then,

$$h = \int h^+ dy - \int h^- dy$$

are two convex functions. This means h' is of BV. In this case, we have

$$\mu(dx) = h''(x)dx$$

Let S_0 be today's spot price. We first consider the case $S_T > S_0$, then

$$\begin{aligned}
h(S_T) &= h(S_0) + \int_{S_0}^{S_T} h'_+(x)dx \\
&= h(S_0) + \int_{S_0}^{S_T} h'_+(S_0) + \mu((S_0,x])dx \\
&= h(S_0) + (S_T - S_0)h'_+(S_0) + \int_{S_0}^{S_T} \int \mathbf{1}_{S_0 \le K \le x \le S_T} \mu(dK)dx \\
&\overset{\text{Fubini}}{=} h(S_0) + (S_T - S_0)h'_+(S_0) + \int \int \mathbf{1}_{S_0 \le K \le x \le S_T} dx\mu(dK) \\
&= h(S_0) + (S_T - S_0)h'_+(S_0) + \int_{(S_0,S_T]} (S_T - K)\mu(dK) \\
&= h(S_0) + (S_T - S_0)h'_+(S_0) + \int_{(S_0,\infty)} (S_T - K)^+ \mu(dK)
\end{aligned}$$

the last part already looks like a portfolio of calls. When $S_T \le S_0$, we get

$$
\begin{aligned}
h(S_T) &= h(S_0) + \int_{S_0}^{S_T} h'_+(x)dx \\
&= h(S_0) + \int_{S_0}^{S_T} h'_+(S_0) - \mu((x,S_0])dx \\
&= h(S_0) + (S_T - S_0)h'_+(S_0) - \int_{S_0}^{S_T} \int \mathbf{1}_{S_T \le x < \le S_0} \mu(dK)dx \\
&= h(S_0) + (S_T - S_0)h'_+(S_0) + \int_{(S_T,S_0]} \int \mathbf{1}_{S_T \le x < K \le S_0} dx\mu(dK) \\
&= h(S_0) + (S_T - S_0)h'_+(S_0) + \int_{(S_T,S_0]} (K - S_T)\mu(dK) \\
&= h(S_0) + (S_T - S_0)h'_+(S_0) + \int_{(0,S_0]} (K - S_T)^+ \mu(dK)
\end{aligned}
$$

Put both scenarios together, we have the following replication theorem.

Theorem 8 — Replication Theorem. Under the previous conditions, we have

$$
h(S_T) = \underbrace{h(S_0)}_{\text{cash}} + \underbrace{(S_T - S_0)h'_+(S_0)}_{\text{future}} + \underbrace{\int_{(0,S_0]} (K - S_T)^+ \mu(dK)}_{\text{OTM puts}} + \underbrace{\int_{(S_0,\infty)} (S_T - K)^+ \mu(dK)}_{\text{OTM calls}}
$$

R Indeed, these integrals might be hard to be evaluated exactly or use potentially infinite amount of trades to replicate a complicated payoff. One might consider different approximations such as Riemann sums (different quadrature rules) to get some close while being practical.

There is an elementary standard butterfly example that will be left as an exercise here.

Corollary 2.1.1 Under the conditions previously mentioned and assume zero interest rates, the price of $h(S_T)$ is given by

$$
P(h(S_T)) = h(S_0) + \int_{(0,S_0]} P(T,K)\mu(dK) + \int_{(S_0,\infty)} C(T,K)\mu(dK)
$$

where $P(T,K)$ and $C(T,K)$ are the respective prices of put and call options with maturity T and strike K.

If there is non-zero interest rates r, then one needs to replace S_0 with the forward price. When $h \in C^2$, the corollary simplifies to

$$
P(h(S_T)) = h(S_0) + \int_{(0,S_0]} P(T,K)h''(K)dK + \int_{(S_0,\infty)} C(T,K)h''(K)dK
$$

2.1.2 Implied Volatility vs. Realized Volatility

Prices of options are usually measured by their related **implied volatility** on the volatility surfaces, which can be considered as $\sigma_{imp} = \sigma(T,K)$. It is well-known in the equity volatility world that the put skew is a common case (why? 3 possible reasons at least). Meanwhile, the term structure is usually decreasing as the expiry increases.
Unlike implied volatility that is derived from option prices, realized volatility is more statistically interpretive. Suppose we look at $N-$day realized quadratic variation of the log-transformed

underlying

$$\sum_{k=0}^{N-1}\left(\log\left(\frac{S_{t+k+1}}{S_{t+k}}\right)\right)^2$$

thee annualized version is the annualized $N-$day variance of S

$$\frac{252}{N}\sum_{k=0}^{N-1}\left(\log\left(\frac{S_{t+k+1}}{S_{t+k}}\right)\right)^2$$

the square root of this would be the annual σ_{real}.

2.2 Variance Swaps and the VIX

Definition 2.2.1 — Variance Swap. Let $S_t, t \geq 0$ be the spot price of an asset or index. We assume for simplicity that interest rates are zero. A variance swap is a path-dependent derivative whose payoff if (an affine transformation of)

$$VS = \sum_{i=1}^{N}(\log S_{t_i} - \log S_{t_{i-1}})^2$$

where N is teh number of (trading) days until maturity and S_{t_i} is the closing price of the asset or index on the $i-$th trading day.

2.2.1 Quadratic Variation Approximation on VS

We will work with the approximation

$$VS \approx \langle \log S \rangle_T$$

where $T = t_N$ and $t_0 = 0$. From previous work, we know that

$$\langle \log S \rangle_T = \int_0^T \frac{1}{S_t^2} d\langle S \rangle_t$$

by Ito's formula on $f(x) = \log x$, we have

$$\log S_T - \log S_0 = \int_0^T \frac{1}{S_t} dS_t - \frac{1}{2}\int_0^T \frac{1}{S_t^2} d\langle S \rangle_t$$

then, this implies

$$VS = 2\log S_0 - 2\log S_T + \int_0^T \frac{2}{S_t} dS_t$$

we call $\log S_T$ as a **log-contract**. Using the replication theorem with $h(x) = \log x$, we have

$$\log S_T = \log S_0 + (S_T - S_0)\frac{1}{S_0} - \int_0^{S_0}(K - S_T)^+ \frac{1}{K^2} dK - \int_{S_0}^{\infty}(S_T - K)^0 \frac{1}{K^2} dK$$
$$VS = -\frac{2}{S_0}(S_T - S_0) + \int_0^{S_0}(K - S_T)^+ \frac{2}{K^2} dK + \int_{S_0}^{\infty}(S_T - K)^+ \frac{2}{K^2} dK + \int_0^T \frac{2}{S_t} dS_t$$

this is a replicating strategy of a variance swap with the following components:

1. shorting $\frac{2}{S_0}$ futures contracts: $-\frac{2}{S_0}(S_T - S_0)$
2. A portfolio of OTM put options, in which the infinitesimal amount of $\frac{2}{K^2}dK$ puts is held for each strike K

3. A portfolio of OTM call options, in which the infinitesimal amount of $\frac{2}{K^2}dK$ puts is held for each strike K
4. Holding $\frac{2}{S_t}$ shares at each time t

We note that this hedging strategy is model-independent. We don't care about the dynamics of the price process S. Since future contracts and self-financing trading strategy with zero initial investment have zero costs, we can derive the price of a VS as follow.

$$PV(VS) = \int_0^{S_0} P(T,K)\frac{2}{K^2}dK + \int_{S_0}^{\infty} C(T,K)\frac{2}{K^2}dK$$

note that this should be the same as the price of 2 log-contracts. The price of a variance swap is to a large extent determined by the price of far OTM puts, by the left-hand end of the implied vol skew.

Definition 2.2.2 — VIX. CBOE gives the following formula

$$VIX = \sqrt{\frac{2e^{rT}}{T}\sum_{i:K_i\le K_0}\frac{\Delta K_i}{K_i^2}P(T,K_i) + \frac{2e^{rT}}{T}\sum_{i:K_i\ge K_0}\frac{\Delta K_i}{K_i^2}C(T,K_i) - \frac{1}{T}\left(\frac{F}{K_0}-1\right)^2}$$

where:

1. r is the risk-free rate
2. T is the time of expiration of SPX options closet to 30 days
3. F is the options-implied forward price
4. K_0 first strike below F
5. K_i strikes of puts $(i \le 0)$ and calls $(i \ge 0)$
6. $P(T,K_i), C(T,K_i)$ prices of puts and calls
7. ΔK_i half the difference between the strike on either side of K_i:

$$\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$$

8. last term is a correction term for counting the strike K_0 twice.

We can consider the sums in under the square-root as discrete approximation for the sum of the two integrals mentioned before. Of course, this is using the mid-point rule for selecting ΔK_i, one might try other rules. We also observe that

$$VIX^2 \approx \frac{1}{T}PV(VS)$$

Therefore, the VIX is the square root of the current market price of the time-average quadratic variation of the SPX index during the next 30 days.

R It might be expected that VIX being slightly higher than the realized volatility due to market makers collecting premiums on implied volatility.

You cannot trade the VIX directly since you would need to trade all possible SPX options and the far OTM ones have very big bid-ask spread. This means giving out large amount edges to the market makers or other counterparties. Instead, people trade VIX future contracts which are usually higher than the actual VIX. The term structure is indeed a **contango**: the prices are above the underlying and increase with the expiration date. This can be explained by the θ decay of a option long.

■ **Example 2.1 — VXX.** VXX is an ETF that tracks the contango by rolling VIX futures. If it is in a contango situation, this should decay over time. ■

Values of the VIX as options prices changes as function of the underlying spot with implied volatility kept constant. As the spot drops, the VIX increases significantly.

For physical delivery underlyings, such as gold, we might have a lot of American options. We cannot use the same formula as we assume F has a implied future with the same expiry date T. This is not true in this case.

2.3 Local Volatility and Dupire's Formula

Local volatility model can be used to perform perfect model calibration using vast amount of option data.

Definition 2.3.1 — Local Volatility. If $\sigma(t,x)$ is a bounded and Lipschitz in x pointwise and uniformly in t, then $\tilde{\sigma}(t,x) = \sigma(t,x)x$ is local Lipschitz in x and of linear growth. This ensures the uniqueness of solution of the following SDE

$$dS_t = \sigma(t,S_t)S_t dW_t$$

this is a local volatility model.

■ **Example 2.2 — CEV is a special case.** CEV model assumes

$$dS_t = S_t^{\gamma+1} dW_T$$

which is a special case where $\sigma(t,x) = x^\gamma$ that is independent of t. ■

Lemma 2.4 The solution S of the local volatility SDE is a square-integrable martingale with quadratic variation

$$\langle S \rangle_t = \int_0^t (\sigma(s,S_s))^2 S_s^2 ds$$

Proof. The full expression of the SDE is

$$S_t = S_0 + \int_0^t \sigma(s,S_s)S_s dW_s$$

By Proposition 1.4.1, given $\langle W \rangle_t = t$, we have

$$\langle S \rangle_t = \int_0^t (\sigma(s,S_s)S_s)^2 d\langle W \rangle_s = \int_0^t (\sigma(s,S_s))^2 S_s^2 ds$$

To show square-integrable martingale. Let K be an upper bound for σ. By $(a+b)^2 \le 2a^2 + 2b^2$, we get that $u(t) := \mathbb{E}(S_t^2)$ satisfies

$$\begin{aligned}
u(t) &= \mathbb{E}\left[\left(S_0 + \int_0^t \sigma(s,S_s)S_s dW_s\right)^2\right] \\
&\le 2S_0^2 + 2\mathbb{E}\left[\left(\int_0^t \sigma(s,S_s)S_s dW_s\right)^2\right] \\
&= 2S_0^2 + 2\mathbb{E}\left[2\int_0^t X_s dX_s + \langle X\rangle_t\right] \qquad X_t = \int_0^t \sigma(s,S_s)S_s dW_s \\
&= 2S_0^2 + 2\mathbb{E}\left[\int_0^t ((\sigma(s,S_s))^2 S_s^2 ds\right] \\
&\le 2S_0^2 + 2K^2\mathbb{E}\left[\int_0^t S_s^2 ds\right] \\
&= 2S_0^2 + 2K^2 \int_0^t u(s)ds
\end{aligned}$$

By Gronwall's inequality, we have

$$u(t) \le 2S_0^2 e^{2K^2 t} \le 2S_0^2 e^{2K^2 T}, \forall t \in [0,T]$$

Thus, S is square-integrable. It is a local-martingale already, (`https://math.stackexchange.com/questions/38908/criteria-for-being-a-true-martingale`), then we check whether its quadratic variation is finite or not.

$$\mathbb{E}\left[\left\langle \int_0^{\cdot} \sigma(s,S_s)S_s dW_s \right\rangle_t\right] \le \mathbb{E}\left[\left\langle \int_0^{\cdot} \sigma(s,S_s)S_s dW_s \right\rangle_T\right] \le K^2 \int_0^T u(s)ds < \infty$$

■

Now let $f : (0,\infty) \to \mathbb{R}$ be a C^2 function. Then Ito's formula yields that

$$f(S_t) - f(S_0) = \int_0^t f'(S_s)\sigma(s,S_s)S_s dW_s + \int_0^t \frac{1}{2} f''(S_s)(\sigma(s,S_s))^2 S_s^2 ds$$

where the Ito integral itself is a continuous local martingale since f' is bounded. We denote L_t as the infinitesimal generator of S defined as

$$L_t f(x) = \frac{1}{2}(\sigma(t,x)x)^2 f''(x)$$

If we further assume that σ is a smooth and bounded away from 0. Then, the law $\mathbf{P} \circ S_t^{-1}$ of S_t, which is a probability measure of $(0,\infty)$ admits a smooth Lebesgue density $\psi(t,x)$. That is, for $f \ge 0$,

$$\mathbb{E}[f(S_t)] = \int_0^\infty f(x)\psi(t,x)dx$$

Proposition 2.4.1 The density ψ satisfies the following **Fokker-Planck equation**,

$$\frac{\partial}{\partial t}\psi(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2(t,x)x^2\psi(t,x))$$

Proof. Take any f smooth with compact support in $(0,\infty)$ and apply Ito's formula

$$f(S_t) = f(S_0) + \int_0^t f'(S_s)dS_s + \frac{1}{2}\int_0^t f''(S_s)(\sigma(s,S_s))^2 S_s ds$$

The Ito integral is again a martingale since f' is bounded. Taking expectation and use product rule repetitively (along with integral of f vanishes at boundary of the compact support), we get

$$\begin{aligned}\int_0^\infty f(x)\psi(t,x)dx &= \mathbb{E}[f(S_t)] \\ &= f(S_0) + \frac{1}{2}\int_0^t \mathbb{E}[f''(S_s)(\sigma(s,S_s))^2 S_s^2]ds \\ &= f(S_0) + \frac{1}{2}\int_0^t \int_0^\infty f''(x)(\sigma(s,x))^2 x^2 \psi(s,x)dxds \\ &= f(S_0) + \frac{1}{2}\int_0^t \int_0^\infty f(x)\frac{\partial^2}{\partial x^2}(\sigma^2(s,x)x^2\psi(s,x))dxds\end{aligned}$$

taking the derivative with respect to t gives

$$\int_0^\infty f(x)\frac{\partial}{\partial t}\psi(t,x)dx = \int_0^\infty f(x)\frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2(t,x)x^2\psi(t,x))dx$$

since f was arbitrary and all integrands are continuous, we are done. ■

2.4.1 The Dupire Formula

Let $C(t,K)$ be the observed European call options with smooth assumption on $t \in [0,T]$ and $K \geq 0$. For non-arbitrage principle, we assume the following:

1. $t \mapsto C(t,K)$ is nondecreasing for all K. Otherwise, spread arbitrage.
2. $K \mapsto C(t,K)$ is convex for every t. Otherwise, butterfly arbitrage.
3. $C(t,0) = S_0$ for all t if interest rate is zero since this option is essentially a forward contract
4. $C(t,K) \to 0$ as $K \to \infty$ deep OTM
5. $\frac{\partial}{\partial t}C(t,K) \to 0$ as $K \to \infty$, θ decay should be diminishing.

These conditions, in particular, hold in Black-Scholes setting.
Assuming that our local volatility model S perfectly replicates the call price surface, we get

$$\begin{aligned}C(t,K) = \mathbb{E}\left[(S_t - K)^+\right] &= \int_0^\infty (x-K)^+\psi(t,x)dx \\ &= \int_0^\infty \psi(t,x)\int_K^\infty \mathbf{1}_{y\leq x}dydx = \int_K^\infty \int_y^\infty \psi(t,x)dxdy\end{aligned}$$

Hence,

$$\frac{\partial^2}{\partial K^2}C(t,K) = \psi(t,K)$$

This gives us a way to obtain this mysterious ψ function. Plug this into the Fokker-Planck equation and solve to get

$$\frac{1}{2}\sigma^2(t,x)x^2\psi(t,x) = \frac{\partial}{\partial t}C(t,x) + \alpha(t)x + \beta(t)$$

Since S_t is square-integrable, we know

$$\int_0^\infty x^2\psi(t,x)dx < \infty$$

Thus, the left hand side of the equation needs to go to 0 as $x \to \infty$ since σ is also bounded by assumption. This implies that $\alpha = \beta = 0$. Thus, we get a famous formula to compute the $\sigma(t,K)$ function.

Theorem 9 — Dupire Formula. If we let

$$\sigma(t,K) = \frac{1}{K}\sqrt{\frac{2\frac{\partial}{\partial t}C(t,K)}{\frac{\partial^2}{\partial K^2}C(t,K)}}$$

and σ is sufficiently smooth, then the corresponding local volatility model S satisfies

$$C(t,K) = \mathbb{E}[(S_t - K)^+], \forall K \geq 0, t \in [0,T]$$

Usage and Criticisms

1. Somehow practical
2. Exclude static arbitrage on plain vanilla instruments. So, this works for variance swap
3. Model is complete
4. Sensitive to interpolation of $C(t,K)$
5. Rather unrealistic when compared to fitted implied volatility

2.5 The Heston Model

Definition 2.5.1 — The Heston Model. For vanishing interest rates, the risk-neutral evolution is defined through the two-dimensional SDE of two standard Brownian motions W,B with $\langle W,B\rangle_t = \rho t$ and constants $\rho \in [-1,1], \xi > 0, \kappa, \theta \geq 0$.

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t dW_t \\ dv_t &= \xi\sqrt{v_t} dB_t + \kappa(\theta - v_t)dt \end{aligned}$$

Interpretation:

1. S is the the stock price process with stochastic volatility $\sqrt{v}$
2. v is the variance process
3. ρ is the correlation
4. ξ is the vol of vol
5. κ is the mean reversion speed
6. θ is the mean reversion level

2.5.1 The CIR Process

Definition 2.5.2 — Cox-Ingersoll-Ross Process (CIR Process). The solution of the SDE

$$dv_t = \xi\sqrt{v_t} dB_t + \kappa(\theta - v_t)dt$$

is called the CIR process or Bessel process with drift.

Proposition 2.5.1 The following SDE has a unique strong solution v that satisfies $v_t \geq 0$ for all t $\mathbf{P}-$ a.s.

$$dv_t = \xi\sqrt{0 \vee v_t} dB_t + \kappa(\theta - v_t)dt$$

Proof. To show that the solution must be nonnegative, let $\varepsilon > 0$ and define the stopping time

$$\sigma := \inf\{t \geq 0 = v_t = -\varepsilon\}$$

suppose for contradiction that $\mathbf{P}(\sigma < \infty) > 0$. If $\sigma(\omega) < \infty$ then there must be $r < \sigma$ such that $v_t(\omega) < 0$ for $r < t < \sigma(\omega)$. Hence,

$$dv_t(\omega) = \kappa(\theta - v_t(\omega))dt, \;\; t \in (r, \sigma(\omega))$$

Since $v_t(\omega) < 0, r < t < \sigma(\omega)$, the map $t \mapsto v_t(\omega)$ is strictly increasing on $(r, \sigma(\omega))$. But this is impossible. ■

Let $d := \frac{4\kappa\theta}{\xi^2}$.

Proposition 2.5.2 Suppose that $d \in \{2,3,\cdots\}$ and let $X^1,\cdots,X^d$ be d independent Ornstein-Uhlenbeck processes,

$$dX_t^i = \frac{\xi}{2}dW_t^i - \frac{\kappa}{2}X_t^i dt$$

for independent Brownian motions $W^1,\cdots,W^d$ and with $\sum_{i=1}^d (X_0^i)^2 = v_0$. Then, there exists a Brownian motion B such that

$$v_t := \sum_{i=1}^d (X_t^i)^2$$

solves

$$dv_t = \xi\sqrt{v_t}dB_t + \kappa(\theta - v_t)dt$$

Proof. For $i \neq j$, $\left\langle W^i, W^j \right\rangle_t = 0$ and hence also $\left\langle X^i, X^j \right\rangle_t = 0$. Therefore the $d-$dimensional Ito formula yields that $v_t := \sum_{i=1}^d (X_t^i)^2$ satisfies

$$\begin{aligned}
dv_t &= \sum_{i=1}^d 2X_t^i dX_t^i + \frac{1}{2}\sum_{i=1}^d 2d\left\langle X^i \right\rangle_t \\
&= \xi\sum_{i=1}^d X_t^i dW_t^i - \kappa\sum_{i=1}^d X_t^i X_t^i dt + d\frac{\xi^2}{4}dt \\
&= \xi\sum_{i=1}^d X_t^i dW_t^i + \left(\underbrace{\frac{d\xi^2}{4}}_{=\theta\kappa} - \kappa v_t\right)dt
\end{aligned}$$

thus, satisfies the SDE. Next we find such a Brownian motion

$$B_t = \sum_{i=1}^d \int_0^t \frac{X_s^i}{\sqrt{v_s}}dW_s^i$$

Since the integrand is of BV, B is a continuous local martingale with $B_0 = 0$, then

$$\langle B \rangle_t = \sum_{i=0}^d \int_0^t \frac{(X_s^i)^2}{v_s}ds = t$$

thus, by Levy's theorem, B is a Brownian motion and

$$\sum_{i=1}^d X_t^i dW_t^i = \sqrt{v_t}dB_t$$

■

We check whether v may hit zero. Let $\zeta = \inf\{t : v_t = 0\}$ be the stopping time of hitting 0.

Theorem 10 We have the following dichotomy:
1. For $d \geq 2$, we have $\zeta = \infty$ $\mathbf{P}-$a.s.
2. For $d < 2$, we have $\zeta < \infty$ $\mathbf{P}-$a.s.

Proof. We define a scale function of v

$$s(x)=\int_1^x e^{\frac{2\kappa y}{\xi^2}}y^{-\frac{2\kappa\theta}{\xi^2}}dy$$

we have

$$s'(x)=e^{\frac{2\kappa x}{\xi^2}}x^{-\frac{2\kappa\theta}{\xi^2}}$$

and

$$s''(x)=\frac{2\kappa}{\xi^2}s'(x)-\frac{2\kappa\theta}{\xi^2}\frac{1}{x}s'(x)=\frac{2\kappa}{\xi^2}\left(1-\frac{\theta}{x}\right)s'(x)$$

Thus, s satisfies the ODE

$$\frac{\xi^2}{2}xs''(x)+\kappa(\theta-x)s'(x)=0$$

Thus, s is a harmonic function that satisfies $Ls=0$ where

$$L=\frac{\xi^2}{2}x\frac{\partial^2}{\partial x^2}+\kappa(\theta-x)\frac{\partial}{\partial x}$$

which is the infinitesimal generator of v. Thus, by Ito's formula, for $t<\zeta$,

$$ds(v_t)=s'(v_t)\xi\sqrt{v_t}dB_t+Ls(v_t)dt=s'(v_t)\xi\sqrt{v_t}dB_t$$

therefore, $s(v_t)$ is a continuous local martingale. Let $\varepsilon>0$ and $M>0$

$$\tau_\varepsilon=\inf\{t:v_t\le\varepsilon\},\ \tau^M=\inf\{t:v_t\ge M\},\ \tau=\tau_\varepsilon\wedge\tau^M$$

we have

$$\begin{aligned}\max_{y\in[\varepsilon,M]}(s(y))^2&\ge\mathbb{E}\left[(s(v_{t\wedge\tau}))^2\right]\\&=\mathbb{E}\left[(s(v_0)+\int_0^{t\wedge\tau}s'(v_s)\xi\sqrt{v_s}dB_s)^2\right]\\&=s^2(v_0)+\mathbb{E}\left[\int_0^{t\wedge\tau}(s'(v_s))^2\xi^2v_sds\right]\\&\ge s^2(v_0)+\min_{y\in[\varepsilon,M]}(s'(y))^2\xi^2\varepsilon\mathbb{E}[t\wedge\tau]\end{aligned}$$

as $t\to\infty$, $\mathbb{E}[t\wedge\tau]\to\mathbb{E}[\tau]$ that is finite. Thus, $\tau<\infty$ $\mathbf{P}$−a.s. Hence, by optional sampling theorem,

$$s(v_0)=\mathbb{E}[s(v_{t\wedge\tau})],\forall t$$

since $s(v_{t\wedge\tau})$ is bounded, dominated convergence yields

$$s(v_0)=\lim t\to\infty\mathbb{E}[s(v_{t\wedge\tau})]=\mathbb{E}[s(v_\tau)]=s(\varepsilon)\mathbf{P}(\tau_\varepsilon<\tau^M)+s(M)\mathbf{P}(\tau^M<\tau_\varepsilon)$$

Note that

$$e^{\frac{2\kappa y}{\xi^2}}y^{-\frac{2\kappa\theta}{\xi^2}}\overset{y\downarrow 0}{\sim}y^{-\frac{d}{2}}$$

for $d\ge 2$, we have $s(\varepsilon)\to-\infty$ as $\varepsilon\to 0$. Thus,

$$\lim_{\varepsilon\to 0}\mathbf{P}(\tau_\varepsilon<\tau^M)=0,\forall M>v_0$$

this means $\infty=\lim_{\varepsilon\to 0}\tau_\varepsilon=\xi$ $\mathbf{P}$−a.s. The other case is analogous. ■

R In reality, we encounter $d < 2$ a bit more. There exists variance process that does hit zero, say holiday or market close.

2.5.2 The Hestion Model as an Affine Diffusion

We solve the first equation in Heston to get

$$S_t = S_0 \exp\left(\int_0^t \sqrt{v_s} dW_s - \frac{1}{2}\int_0^t v_s ds\right)$$

we assume $S_0 > 0$ and work with log-transformed version,

$$X_t := \log S_t = X_0 + \int_0^t \sqrt{v_s} dW_s - \frac{1}{2}\int_0^t v_s ds$$

then the first SDE is chagned to

$$dX_t = \sqrt{v_t} dW_t - \frac{1}{2} v_t dt$$

let δ be $\frac{1}{2}$ to get a more general form. Then, we have

$$\begin{aligned} d\langle X\rangle_t &= v_t d\langle W\rangle_t = v_t dt \\ d\langle v\rangle_t &= \xi^2 v_t d\langle B\rangle_t = \xi^2 v_t dt \\ d\langle X, v\rangle_t &= \xi v_t d\langle W, B\rangle_t = \rho\xi v_t dt \end{aligned}$$

By 2D-Taylor expansion, we compute the infinitesimal generator of this SDE

$$Lf(x,v) = -\delta v f_x + \kappa(\theta - v) f_v + \frac{\xi^2}{2} v f_{vv} + \rho \xi v f_{xv} + \frac{1}{2} v f_{xx}$$

This is an affine diffusion. The characteristic function of their marginal distributions can be computed in close form. We proceed to derive them.
Recall that the characteristic function of the marginals of the Heston model is

$$\mathbb{E}[e^{i\lambda X_t + i\mu v_t}], i = \sqrt{-1}, \lambda, \mu \in \mathbb{R}$$

we claim that for fixed λ, μ, we can write the characteristic function as

$$e^{\phi(t) + \gamma(t) X_0 + \psi(t) v_0}$$

for certain functions ϕ, γ, ψ. This is common for all affine diffusion (basic affine jump-diffusion models). Suppose these smooth functions do exist. We fix $T > 0$. Then,

$$M_t = \mathbb{E}[e^{i\lambda X_T + i\mu v_T} | \mathscr{F}_t],\ 0 \le t \le T$$

is a martingale. By the Markov property of SDEs

$$M_t(\omega) = \mathbb{E}[e^{i\lambda X_{T-t} + i\mu v_{T-t}} | X_0 = X_t(\omega), v_0 = v_t(\omega)] = e^{\phi(T-t) + \gamma(T-t) X_t(\omega) + \psi(T-t) v_t(\omega)}$$

thus,

$$M_t = e^{\phi(T-t) + \gamma(T-t) X_t + \psi(T-t) v_t} = f(T-t, X_t, v_t)$$

by Ito's formula,

$$dM_t = d(\text{continuous local martingale}) - \left(\frac{\partial f}{\partial t}(T-t, X_t, v_t) - Lf(T-t, X_t, v_t)\right) dt$$

since M is a Martingale, the dt term should be 0.

3. Dynamic Portfolio Strategies

3.1 Volatility Drag

Definition 3.1.1 — CAGR. Let V_t be the portfolio value at time t. The portfolio growth is often measured in **compound annual growth rate (CAGR)** defined by

$$\mathbf{CAGR}(t_0, t_1) = \left(\frac{V_{t_1}}{V_{t_0}}\right)^{\frac{1}{t_1 - t_0}} - 1$$

where t_0, t_1 are meansured in years.

■ **Example 3.1 — Investment in SPX.** Take monthly CGAR of SPX since October, 1950. The average of all monthly CAGRs is **21.3%**. The average of all annual CAGRS is **9.3%**. How come this is significantly lower? What if we take the CAGR of the untire 71 years? It would just **7.9%**. One would have similar observations for other assets. What's going on? ■

Going back to geometric Brownian motion (GBM),

$$dS_t = \sigma S_t dW_t + \mu S_t dt \Longrightarrow S_t = S_0 e^{\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t}$$

Then the expected one-year CAGR is

$$\mathbb{E}\left[\frac{S_1}{S_0} - 1\right] = \mathbb{E}\left[e^{\sigma W_1 + (\mu - \frac{1}{2}\sigma^2)} - 1\right] - 1 = e^{\mu} - 1$$

this is what you see if you average CAGR over many separate one-year periods. Actual realized CAGR per trajectory is

$$\begin{aligned}
\mathbf{CAGR}(T) &= \left(\frac{S_T}{S_0}\right)^{\frac{1}{T}} - 1 \\
&= \exp\left(\frac{\sigma}{T} W_T + \left(\mu - \frac{1}{2}\sigma^2\right)\right) - 1 \\
&\xrightarrow[T \to \infty]{LLN} e^{\mu - \frac{1}{2}\sigma^2} - 1 < e^{\mu} - 1
\end{aligned}$$

It can even be negative if $\frac{1}{2}\sigma^2 > \mu$ even if $\mu > 0$!
We note that
$$\frac{1}{T}\log S_T = \frac{1}{T}\log S_0 + \frac{\sigma}{T}W_T + \mu - \frac{1}{2}\sigma^2 \xrightarrow{\mathbf{P}\text{-a.s.}} \mu - \frac{1}{2}\sigma^2$$
When this is negative, $S_T \sim e^{T(\mu-\frac{1}{2}\sigma^2)} \to 0$ $\mathbf{P}-$a.e. trajectory will go to 0. If this positive, we have exponential growth with rate $\mu - \frac{1}{2}\sigma^2$.

> "Almost certain frustration of investors."

R Suppose $\mu = 0$, then $S_t = S_0 e^{\sigma W_t - \frac{1}{2}\sigma^2 t}$ which is a Martingale greater or equal to 0. By Martinagle convergence theorem, $S_t \to S_\infty$ $\mathbf{P}-$a.s. And $S_\infty = 0$ $\mathbf{P}-$a.s. Moreover, $\mathbb{E}[S_t] = S_0, \forall t$. This means the Maringale S_t is not uniformly integrable.

Instead of investing everything into risky asset, retain some fraction α of total wealth in a store of value such as a bond. Assume the bond value is described as $B_t = e^{rt}$. Then, the self-financing strategy is
$$\eta_t = \frac{\alpha V_t}{B_t},\ \xi_t = \frac{(1-\alpha)V_t}{S_t}$$
the self-financing condition can be derived as follow (assuming S_t follows GBM)
$$\begin{aligned} dV_t &= \frac{(1-\alpha)V_t}{S_t}dS_t + \frac{\alpha V_t}{B_t}dB_t \\ &= (1-\alpha)V_t\sigma dW_t + (1-\alpha)V_t\mu dt + \alpha V_t r dt \end{aligned}$$
This has the following solution
$$V_t = V_0 \exp\left((1-\alpha)\sigma W_t + \left((1-\alpha)\mu + \alpha r - \frac{1}{2}(1-\alpha)^2\sigma^2\right)t\right)$$
The asymptotic growth rate is
$$\lim_{T\to\infty}\frac{1}{T}\log V_T = (1-\alpha)\mu + \alpha r - \frac{1}{2}(1-\alpha)^2\sigma^2$$

Theorem 11 — Kelly Criterion (1956). Find the optimal α.
$$f(\alpha) = (1-\alpha)\mu + \alpha r - \frac{1}{2}(1-\alpha)^2\sigma^2$$
this is a downward sloping parabola and the optimizer is at
$$\alpha = 1 - \frac{\mu - r}{\sigma^2}$$
which yields the asymptotic growth rate
$$r + \frac{(\mu - r)^2}{\sigma^2} > f(0) = \mu - \frac{1}{2}\sigma^2$$

One thing to keep in mind is that this is not a static strategy. If stocks go up, then the percentage of wealth in stock goes up and you need to sell stocks and buy bonds vice versa. In practice, 60/40 portfolio.

First alternative approach to Kelly criterion
Fix $T > 0$,

$$V_T = V_0 \exp\left(\underbrace{(1-\alpha)\sigma W_T}_{\text{median}=0} + \left((1-\alpha)\mu + \alpha r - \frac{1}{2}(1-\alpha)^2\sigma^2\right)T\right)$$

Then, the median of V_0 is

$$V_0 \exp\left(\left((1-\alpha)\mu + \alpha r - \frac{1}{2}(1-\alpha)^2\sigma^2\right)T\right)$$

so, Kelly optimal α also maximizes the median of the portfolio value.

Second alternative approach to Kelly criterion
At each time t, invest fraction π_t of the current wealth into the risky asset and $1-\pi_t$ into the bond. Then,

$$\xi_t = \frac{\pi_t V_t}{S_t}, \quad \eta_t = \frac{(1-\pi_t)V_t}{B_t}$$

the value process

$$V_t^\pi = \xi_t S_t + \eta_t B_t$$

Again, we want a self-financing condition,

$$\begin{aligned}
dV_t^\pi &= \xi_t dS_t + \eta_t dB_t \\
&= \sigma \pi_t V_t^\pi dW_t + \mu \sigma_t V_t^\pi dt + r(1-\pi_t)V_t^\pi dt \\
V_t^\pi &= V_0 \exp\left(\sigma \int_0^t \pi_s dW_s + \int_0^t (\mu \pi_s + r(1-\pi_s) - \frac{1}{2}\sigma^2\pi_s^2)ds\right)
\end{aligned}$$

we need the stochastic integral being well-defined. In particular, $\int_0^T \pi_s^2 ds$ exists is definitely needed. We shall assume that

$$\mathbb{E}\left[\int_0^T \pi_s^2 ds\right] < \infty$$

and π progressive. Our goal is to maximize the expected log-utility

$$\mathbb{E}[\log V_T^\pi]$$

over all π — **Merton problem**. We can compute now,

$$\begin{aligned}
\mathbb{E}[\log V_T^\pi] &= \log V_0 + \underbrace{\mathbb{E}\left[\sigma \int_0^T \pi_s dW_s\right]}_{=0 \text{ by our assumption}} + \mathbb{E}\left[\int_0^T (\mu \pi_s + r(1-\pi_s) - \frac{1}{2}\sigma^2\pi_s^2)ds\right] \\
&= \log V_0 + \mathbb{E}\left[\int_0^T (\mu \pi_s + r(1-\pi_s) - \frac{1}{2}\sigma^2\pi_s^2)ds\right]
\end{aligned}$$

the integrand is maximized for each (t,ω) by $\pi_t^* = \frac{\mu - r}{\sigma^2}$. Thus,

$$\mathbb{E}[\log V_T^\pi] \leq \mathbb{E}\left[\log V_T^{\pi^*}\right]$$

this yields the same optimizer as Kelly-optimal strategy.

3.2 Leveraged ETFs

Typical leverage factors for ETFs are $m \in \{-3, -2, -1, -0.5, 2, 3\}$. What would be the worst case for an inverse ETF with $m = -1$ in one day? Double or more of the underlying index would wipe out all value.

Let $\{S_n\}_{n=0,1,\cdots}$ be the value of QQQ. It has a return at time k,

$$\frac{S_k - S_{k-1}}{S_{k-1}} =: R_k$$

m is a leverage factor as introduced before. Let X_k be the price of $m-$leveraged ETF. Then, the return at time k must be

$$mR_k = m\frac{S_k - S_{k-1}}{S_{k-1}}$$

then $X_k = X_{k-1}(1 + mR_k)$. Then,

$$X_n = X_0 \prod_{k=1}^{n} (1 + mR_k)$$

Over 1 day, S changes by $x\%$ then X changes by $mx\%$, Y changes by $-mx\%$.

■ **Example 3.2 — Volatility Drag.** Over 2 days with $m = 3$, there was a round trip in the index,

$$S : 100 \underset{+10\%}{\rightarrow} 110 \underset{+100/11}{\rightarrow} 100$$

then,

$$X : 100 \underset{+30\%}{\rightarrow} 130 \underset{-300/11}{\rightarrow} 94.55$$

and

$$Y : 100 \underset{-30\%}{\rightarrow} 70 \underset{+300/11\%}{\rightarrow} 89.09$$

both X, Y decrease after the round trip. This is due to the volatility drag. ■

■ **Example 3.3 — Bull!.**

$$S : 100 \rightarrow 110 \rightarrow 121 \implies +21\%$$

$$X : 100 \rightarrow 130 \rightarrow 169 \implies +69\% > 3 \times 21\%$$

$$Y : 100 \rightarrow 70 \rightarrow 49 \implies -57\% > -3 \times 21\%$$

■

How can we better understand the long-term payoff?

Assume $X_0 = 1$,

$$\begin{aligned}
\log X_n &= \log \prod_{k=1}^{n} (1 + mR_k) \\
&= \sum_{k=1}^{n} \log(1 + mR_k) \\
&= \sum_{k=1}^{n} mR_k - \frac{1}{2} m^2 R_k^2 + o(m^2 R^2) \\
&\approx \sum_{k=1}^{n} mR_k - \frac{1}{2} \sum_{k=1}^{n} m^2 R_k^2 \\
&= m \sum_{k=1}^{n} \frac{1}{S_{k-1}} (S_k - S_{k-1}) - \frac{m^2}{2} \sum_{k=1}^{n} \frac{1}{S_{k-1}^2} (S_k - S_{k-1})^2
\end{aligned}$$

3.3 CPPI

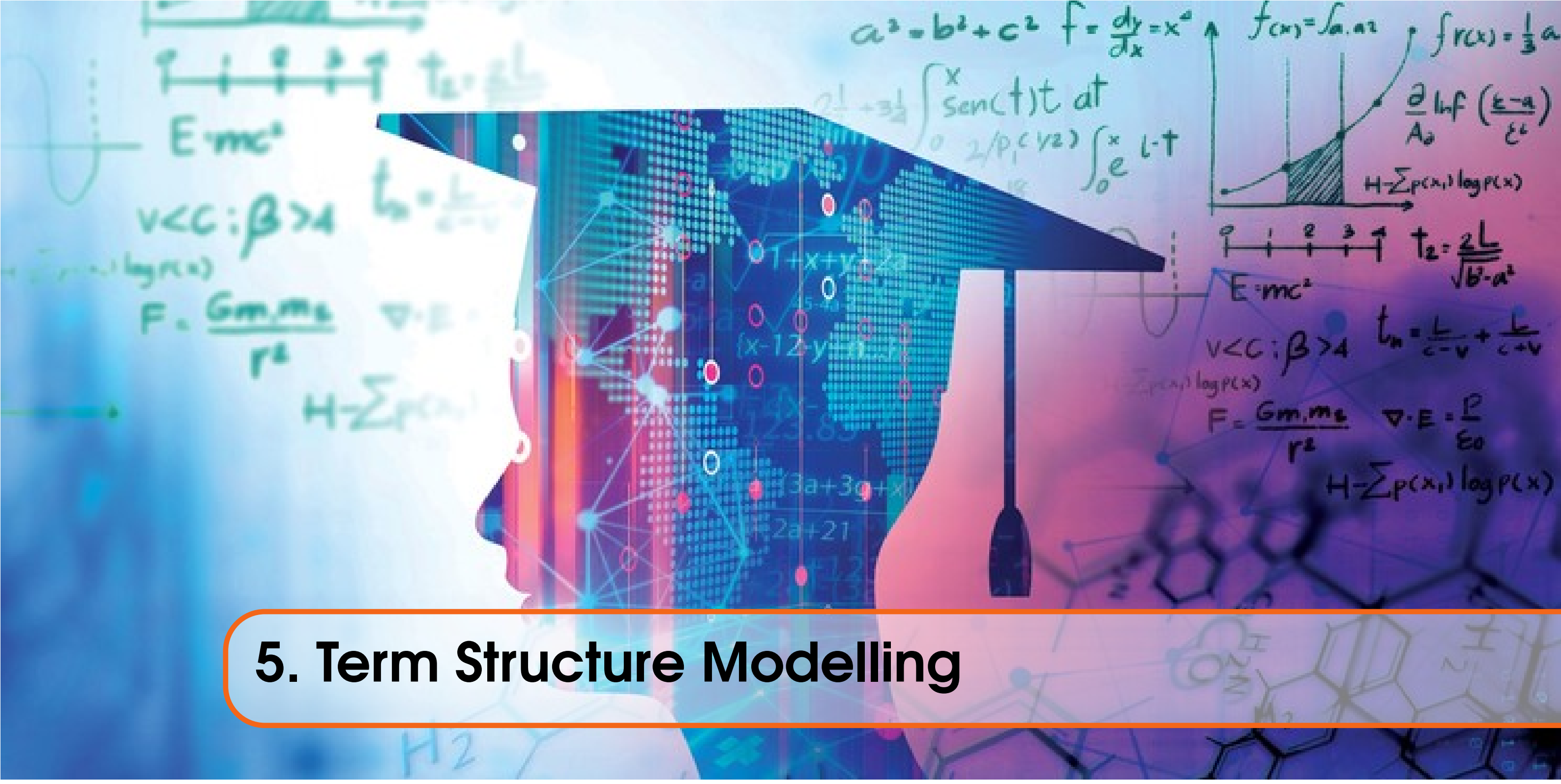
3.4 Equal-weighted Portfolios

3.5 Universal Portfolios

3.6 Stochastic Portfolio Theory

4. Price Impact and Execution

4.1 Optimal Trade Execution
4.2 Pegged Currency Markets
4.3 Gamma Squeezes
4.4 Predatory Trading
4.5 Dark Pools



5. Term Structure Modelling

5.1 Volatility Term Structure Modelling

5.2 Interest Rate Term Structure Modelling

5.3 Libor Market Models