

Introduction to Option Pricing with Fourier Transform: Option Pricing with Exponential Lévy Models

Kazuhisa Matsuda

Department of Economics
The Graduate Center, The City University of New York,
365 Fifth Avenue, New York, NY 10016-4309
Email: maxmatsuda@maxmatsuda.com
<http://www.maxmatsuda.com/>

December 2004

Abstract

This sequel is designed as an introduction to Fourier transform option pricing for readers who have zero previous knowledge of Fourier transform. First part of this sequel is devoted for the basic understanding of Fourier transform and discrete Fourier transform using numerous examples and providing important properties. Second part of this sequel applies FT and DFT option pricing approach for three exponential Lévy models: Classic Black-Scholes model which is the only continuous exponential Lévy model, Merton jump-diffusion model (1976) which is an exponential Lévy model with finite arrival rate of jumps, and variance gamma model by Madan, Carr, and Chang (1998) which is an exponential Lévy model with infinite arrival rate of jumps. Some readers may question that what the need for FT option pricing is since all three models above can price options with closed form formulae. The answer is that these three models are special cases of more general exponential Lévy models. Options cannot be priced with general exponential Lévy models using the traditional approach of the use of the risk-neutral density of the terminal stock price because it is not available. Therefore, Carr and Madan (1999) rewrite the option price in terms of a characteristic function of the log terminal stock price by the use of FT. The advantage of FT option pricing is its generality in the sense that the only thing necessary for FT option pricing is a characteristic function of the log terminal stock price. This generality of FT option pricing speeds up the calibration and Monte Carlo simulation with various exponential Lévy models. It is no doubt to us that FT option pricing will be a standard option pricing method from now on.

© 2004 Kazuhisa Matsuda All rights reserved.

JEL classification: G10; G12; G13

Keywords: Fourier transforms; Discrete Fourier transforms; Option pricing; Lévy processes; Random time change; Jump diffusion model

Contents

1	Introduction to the Option Pricing with Fourier Transform	1
2	Prerequisite for Fourier Transform	7
2.1	Radian	7
2.2	Wavelength	7
2.3	Frequency, Angular Frequency, and Period of a Wave	7
2.4	Sine and Cosine	9
2.5	Derivative and Integral of Sine and Cosine Function	10
2.6	Series Definition of Sine and Cosine Function	11
2.7	Euler's Formula	11
2.8	Sine Wave: Sinusoid	12
3	Fourier Transform	15
3.1	Definition of Fourier Transform	15
3.2	Examples of Fourier Transform	17
3.2.1	Double-Sided Exponential Function	17
3.2.2	Rectangular Pulse	19
3.2.3	Dirac's Delta Function	20
3.2.4	Gaussian Function	22
3.2.5	Cosine Wave	24
3.2.6	Sine Wave	26
3.3	Properties of Fourier Transform	29
3.3.1	Dirac's Delta Function	29
3.3.2	Useful Identity: Dirac's Delta Function	31
3.3.3	Linearity of Fourier Transform	34
3.3.4	FT of Even and Odd Functions	35
3.3.5	Symmetry of Fourier Transform	37
3.3.6	Differentiation of Fourier Transform	38
3.3.7	Time Scaling of Fourier Transform	39
3.3.8	Time Shifting of Fourier Transform	40
3.3.9	Convolution: Time Convolution Theorem	41
3.3.10	Frequency-Convolution Theorem	42
3.3.11	Frequency Shifting: Modulation	44
3.3.12	Parseval's Relation	44
3.3.13	Summary of Important Properties of Fourier Transform	46
3.4	Existence of the Fourier Integral	48

4 Characteristic Function	49
4.1 Definition of a Characteristic Function	49
4.2 Properties of a Characteristic Function	50
4.3 Characteristic Exponent: Cumulant-Generating Function	51
4.4 Laplace Transform	52
4.5 Relationship with Moment Generating Function	52
4.6 Summary: How to Calculate Standardized Moments from Characteristic Function and Moment Generating function	55
4.7 Examples of Characteristic Functions	56
5 Discrete Fourier Transform	57
5.1 Intuitive Derivation of DFT: Approximation of FT	57
5.2 Definition of DFT	59
5.2.1 Physicists' Definition of DFT	59
5.2.2 Signal Processing Definition of DFT	59
5.2.3 Requirement of DFT	60
5.2.4 Sampling Theorem and Nyquist Rule: How to Determine the Time Domain Sampling Interval Δt	61
5.3 Examples of DFT	63
5.3.1 Sine Wave	63
5.3.2 Double-Sided Exponential	71
5.3.3 Rectangular Pulse	72
5.3.4 Gaussian Function	73
5.3.5 Cosine Wave	74
5.4 Properties of DFT	75
5.4.1 Linearity of DFT	75
5.4.2 DFT of Even and Odd Functions	75
5.4.3 Symmetry of DFT	77
5.4.4 Time Shifting of DFT	78
5.4.5 Frequency Shifting: Modulation	79
5.4.6 Discrete Convolution: Time Convolution Theorem	80
5.4.7 Discrete Frequency-Convolution Theorem	81
5.4.8 Parseval's Relation	82
5.4.9 Summary of DFT Properties	83
6 Lévy Processes	84
6.1 Definition of Lévy Process	84
6.2 Standard Brownian Motion Process: The Only Continuous Lévy Process Generated by A Normal Distribution	87
6.3 Poisson Process and Compound Poisson Process: Finite Activity Lévy Processes Generated by Poisson Distribution	87
6.4 Lévy-Itô Decomposition and Infinite Activity Lévy Process	90
6.5 Lévy-Khintchin Representation	91

6.6	Stable Processes	92
7	Black-Scholes Model as an Exponential Lévy Model	95
7.1	Standard Brownian Motion Process: A Lévy Process Generated by a Normal Distribution	95
7.2	Black-Scholes' Distributional Assumptions on a Stock Price	96
7.3	Traditional Black-Scholes Option Pricing: PDE Approach by Hedging	99
7.4	Traditional Black-Scholes Option Pricing: Martingale Pricing Approach	101
7.5	Alternative Interpretation of Black-Scholes Formula: A Single Integration Problem	105
7.6	Black-Scholes Model as an Exponential Lévy Model	107
8	Option Pricing with Fourier Transform: Black-Scholes Example	109
8.1	Motivation	109
8.2	Derivation of Call Price with Fourier Transform: Carr and Madan (1999)	111
8.3	How to Choose Decay Rate Parameter α : Carr and Madan (1999)	113
8.4	Black-Scholes Model with Fourier Transform Pricing Method	115
8.5	Derivation of Near-Maturity OTM Vanilla Option Pricing Function with Fourier Transform: Carr and Madan (1999)	119
8.6	Derivation of Call Pricing Function with Discrete Fourier Transform (DFT): Carr and Madan (1999)	126
8.7	Implementation and Performance of DFT Pricing Method with Black-Scholes Model	130
8.8	Summary of Formulae of Option Price with Fourier Transform	137
9	Merton (1976) Jump-Diffusion Model	139
9.1	Model Type	139
9.2	Model Derivation	140
9.3	Log Stock Price Process for Merton Jump-Diffusion Model	149
9.4	Lévy Measure for Merton Jump-Diffusion Model	150
9.5	Option (Contingent Claim) Pricing: PDE Approach by Hedging	152
9.6	Option (Contingent Claim) Pricing: Martingale Approach	155
9.7	Option Pricing Example of Merton Jump-Diffusion Model	158
10	Merton (1976) Jump-Diffusion Model with Fourier Transform Pricing	162
10.1	Merton JD Model with Fourier Transform Pricing Method	162
10.2	Discrete Fourier Transform (DFT) Call Pricing Formula with Merton Jump-Diffusion Model	167
10.3	Implementation and Performance of DFT Pricing Method	

with Merton Jump-Diffusion Model	167
10.4 Summary of Formulae of Option Price with Fourier Transform in Merton Jump-Diffusion Model	170
11 Variance Gamma (VG) Model by Madan, Carr, and Chang (1998)	172
11.1 Model Type	172
11.2 Model Derivation	173
11.2.1 Subordination Theorem of Lévy Processes	173
11.2.2 Tempered α -Stable Subordinator: General Case	174
11.2.3 Gamma Subordinator (Process): Special Case of Tempered α -Stable Subordinator When $\alpha = 0$	176
11.2.4 Re-Parameterizing Tempered α -Stable Subordinator Using Its Scale-Invariance (Self-Similarity) Property	177
11.2.5 Gamma Subordinator (Process): Special Case of Tempered α -Stable Subordinator with Unit Mean Rate When $\alpha = 0$	179
11.2.6 Subordinated Brownian Motion Process with Tempered α -Stable Subordinator with Unit Mean Rate: Normal Tempered α -Stable Process	181
11.2.7 Variance Gamma (VG) Process: Subordinated Brownian Motion Process with Tempered 0-Stable Subordinator with Unit Mean Rate: Normal Tempered 0-Stable Process	182
11.3 Lévy Measure for Variance Gamma (VG) Process	187
11.4 Modeling Stock Price Dynamics with VG Process and VG Log-Return Density	189
11.5 Option Pricing with VG Model	193
11.6 Option Pricing Example of VG Model	195
11.7 Lévy Measure of VG Log-Return z_t	200
12 VG (Variance Gamma) Model with Fourier Transform Pricing	202
12.1 VG Model with Fourier Transform Pricing Method	202
12.2 Discrete Fourier Transform (DFT) Call Pricing Formula with VG Model	210
12.3 Implementation and Performance of DFT Pricing Method with VG Model	210
12.4 Summary of Formulae of Option Price with Fourier Transform in VG Model	215
13 Conclusion	216
Appendix	217
A.1 Set Theory: Notation and Basics	217
A.2 Measure	218
A.2.1 Null Sets	218

A.2.2	Outer Measure	219
A.2.3	Lebesgue Measurable Sets and Lebesgue Measure	220
A.2.4	σ -field	221
A.2.5	Borel σ -field	222
A.2.6	Probability	222
A.3	Stochastic Process	224
A.3.1	Filtration (Information Flow)	224
A.3.2	Nonanticipating (Adapted) Process	224
A.4	Martingales	224
A.4.1	General Concept	224
A.4.2	Martingale Asset Pricing	226
A.4.3	Continuous Martingales, Right Continuous Martingales, Square-Integrable Martingales	226
A.5	Poisson Process	227
A.5.1	Exponential Distribution	227
A.5.2	Poisson Distribution	228
A.5.3	Compensated Poisson Process	229
A.6	Other Distributions Used	230
A.6.1	Gamma Function	230
A.6.2	Incomplete Gamma Function	230
A.6.3	Gamma Distribuiton	231
A.7	Modified Bessel Functions: Modified Bessel Function of the First Kind $I_v(z)$ and Second Kind $K_v(z)$	232
A.8	Itô Formula	234
A.8.1	Itô Formula for Brownian Motion	234
A.8.2	Wimott's (1998) Rule of Thumb of Itô Formula for Brownian Motion	236
A.8.3	Itô Formula for Brownian Motion with Drift	236
A.8.4	Itô Formula for Brownian Motion with Drift in Higher Dimensions	237
A.8.5	Itô Formula for Jump-Diffusion (Finite Activity Lévy) Processes	237
A.8.6	Itô Formula for General (Finite and Infinite Activity) Scalar Lévy Processes	238
	Bibliography	240

[1] Introduction

Many of the option pricing models assume that a stock price process $\{S_t; 0 \leq t \leq T\}$ follows an exponential (geometric) Lévy process:

$$S_t = S_0 e^{L_t},$$

where $\{L_t; 0 \leq t \leq T\}$ is a Lévy process. The reason of the popularity of exponential (geometric) Lévy models is its mathematical tractability which comes from the independent and stationary increments of Lévy processes. Classic Black-Scholes (BS) model chooses a Brownian motion with drift process which is the only continuous Lévy process as their choice of a (risk-neutral) Lévy process:

$$L_t = \left(r - \frac{1}{2} \sigma^2 \right) t + \sigma B_t,$$

where $\{B_t; 0 \leq t \leq T\}$ is a standard Brownian motion process. This specification leads to a normally distributed conditional risk-neutral log return density:

$$\mathbb{Q}(\ln(S_T / S_0) | \mathcal{F}_0) = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp \left[-\frac{\left\{ \ln(S_T / S_0) - \left(r - \frac{1}{2} \sigma^2 \right) T \right\}^2}{2\sigma^2 T} \right].$$

BS call price can be simply calculated as the discounted value of the expected terminal payoff under risk-neutral measure \mathbb{Q} :

$$C(S_0, T) = e^{-rT} \int_K^\infty (S_T - K) \mathbb{Q}(S_T | \mathcal{F}_0) dS_T, \quad (1.1)$$

where $\mathbb{Q}(S_T | \mathcal{F}_0) = \frac{1}{S_T \sqrt{2\pi\sigma^2 T}} \exp \left[-\frac{\left\{ \ln S_T - \left(\ln S_0 + \left(r - \frac{1}{2} \sigma^2 \right) T \right) \right\}^2}{2\sigma^2 T} \right]$ which is a

lognormal density.

But even before BS model was developed, researchers knew that the empirical log return density is not normal, it shows excess kurtosis and skewness. Thus, all the option pricing models after BS (so called beyond BS) try to capture excess kurtosis and negative skewness of the risk-neutral log return density by the use of different techniques.

This sequel deals with Merton jump-diffusion model (we call it Merton JD model) and variance gamma model by Madan, Carr, and Chang (1998) (we call it VG model). These are both exponential Lévy models of different types. Merton's choice of Lévy process is a Brownian motion with drift process plus a compound Poisson jump process which has a continuous path with occasional jumps:

$$L_t = \left(r - \frac{\sigma^2}{2} - \lambda k \right) t + \sigma B_t + \sum_{i=1}^{N_t} Y_i .$$

Merton JD model can be categorized as a finite activity exponential Lévy model because the expected number of jumps per unit of time (i.e. intensity λ) is finite and small. In other words, the Lévy measure $\ell(dx)$ of Merton JD model is finite:

$$\int \ell(dx) < \infty .$$

The only but important difference between the BS and the Merton JD model is the addition of a compound Poisson jump process $\sum_{i=1}^{N_t} Y_i$. Merton introduces three extra parameters λ (intensity of the Poisson process), μ (mean log stock price jump size), and δ (standard deviation of log stock price jump size) to the original BS framework and controls the (negative) skewness and excess kurtosis of the log return density.

Choice of Lévy process by Madan, Carr, and Chang (1998) is a VG process plus a drift:

$$L_t \equiv \left\{ r + \frac{1}{\kappa} \ln \left(1 - \theta \kappa - \frac{\sigma^2 \kappa}{2} \right) \right\} t + VG(x_t; \theta, \sigma, \kappa).$$

A VG process $VG(x_t; \theta, \sigma, \kappa)$ is defined as a stochastic process $\{X_t; 0 \leq t \leq T\}$ created by random time changing (i.e. subordinating) a Brownian motion with drift process $\theta t + \sigma B_t$ by a tempered 0-stable subordinator (i.e. a gamma subordinator) $\{S_t; 0 \leq t \leq T\}$ with unit mean rate:

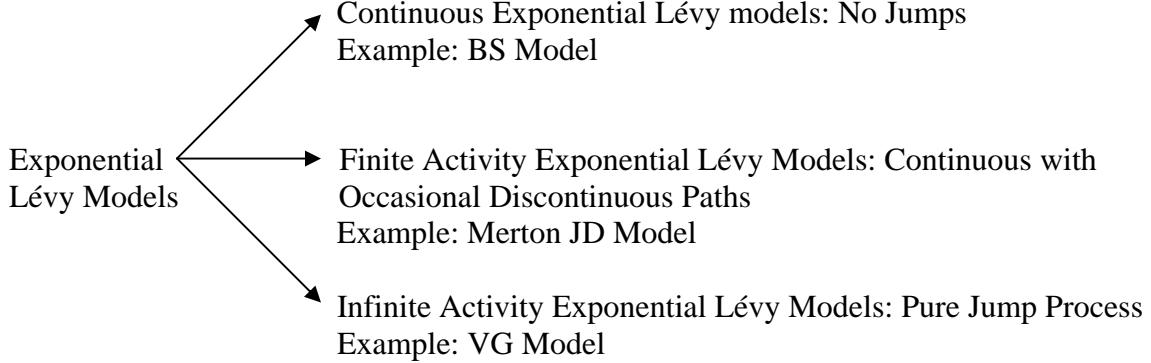
$$X_t \equiv \theta(S_t) + \sigma B_{S_t} .$$

A VG process $VG(x_t; \theta, \sigma, \kappa)$ is characterized as a pure jump Lévy process with infinite arrival rate of jumps. In other words, the Lévy measure of a VG process has an infinite integral:

$$\int_{-\infty}^{\infty} \ell(x) dx = \infty .$$

This means that a VG process has infinitely many small jumps but a finite number of large jumps. VG model introduces two extra parameters: One is variance rate parameter

κ which controls the degree of the randomness of the subordination and the larger κ implies the fatter tails of the log return density. The other is the drift parameter of the subordinated Brownian motion process θ which captures the skewness of the log return density.



Traditionally, both Merton JD call price and VG call price have been expressed as BS type closed form function using the conditional normality of both models. Merton JD call price can be expressed as the weighted average of the BS call price conditional on that the underlying stock price jumps i times to the expiry with weights being the probability that the underlying jumps i times to the expiry. Because of the subordination structure $X_t \equiv \theta(S_t) + \sigma B_{S_t}$ of the VG process $\{X_t; 0 \leq t \leq T\}$, the probability density of VG process can be expressed as the conditionally normal by conditioning on the fact that the realized value of the random time change by a gamma subordinator S_t with unit mean rate is $S_t = g$:

$$VG(x_t | t = S_t = g) = \frac{1}{\sqrt{2\pi\sigma^2}g} \exp\left\{-\frac{(x_t - \theta g)^2}{2\sigma^2 g}\right\}.$$

Using this conditional normality of VG process, the conditional call price can be obtained as a BS type formula after lengthy and tedious process of numerous changes of variables.

The fact that a call price can be expressed as a BS type formula implies that the model has a closed form expression for the log return density $\mathbb{Q}(\ln(S_T / S_0) | \mathcal{F}_0)$. Merton JD model has a log return density of the form:

$$\mathbb{Q}_{Merton}(\ln(S_T / S_0) | \mathcal{F}_0) = \sum_{i=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^i}{i!} N\left(\ln(S_T / S_0); (r - \frac{\sigma^2}{2} - \lambda k)t + i\mu, \sigma^2 t + i\delta^2\right),$$

where $N(x; a, b) \equiv \frac{1}{\sqrt{2\pi}b} \exp\left\{-\frac{(x-a)^2}{2b}\right\}$ which is a normal density. VG model has a log return density of the form:

$$\mathbb{Q}_{VG}(\ln(S_T / S_0) | \mathcal{F}_0) = \frac{\sqrt{2} \exp\left(\frac{\theta}{\sigma^2} x_T\right)}{\kappa^{T/\kappa} \sigma \sqrt{\pi} \Gamma(T/\kappa)} \left(\frac{x_T^2}{\frac{2\sigma^2}{\kappa} + \theta^2} \right)^{\frac{T-1}{2\kappa-4}} K_{\frac{T-1}{\kappa-2}} \left(\frac{\sqrt{x_T^2 \left(\frac{2\sigma^2}{\kappa} + \theta^2 \right)}}{\sigma^2} \right),$$

where $x_T \equiv \ln(S_T / S_0) - \left\{ r + \frac{1}{\kappa} \ln \left(1 - \theta \kappa - \frac{\sigma^2 \kappa}{2} \right) \right\} t$. The existence of the closed form expression for the log return density $\mathbb{Q}(\ln(S_T / S_0) | \mathcal{F}_0)$ enables the use of the equation (1.1) to calculate a call price.

But Merton JD model and VG model are special cases of exponential Lévy models in the sense that more general exponential Lévy models do not have closed form log return densities $\mathbb{Q}(\ln(S_T / S_0) | \mathcal{F}_0)$ or they cannot be expressed using special functions of mathematics. Therefore, we cannot price plain vanilla options using the equation (1.1). How do we price options using general exponential Lévy models? The answer is to use a very interesting fact that characteristic functions (CF) of general exponential Lévy processes are always known in closed-forms or can be expressed in terms of special functions of mathematics although their probability densities are not. There is one-to-one relationship between a probability density and a CF (i.e. CF is just a Fourier transform of a probability density with FT parameters (1,1)) and both of which uniquely determine a probability distribution. If we can somehow rewrite (1.1) in terms of a characteristic function of the conditional terminal stock price $S_T | \mathcal{F}_0$ (i.e. log of $S_T | \mathcal{F}_0$ to be more precise) instead of its probability density $\mathbb{Q}(S_T | \mathcal{F}_0)$, we will be able to price options in general exponential Lévy models.

The purpose of this sequel is to introduce the basics of Fourier transform option pricing approach developed by Carr and Madan (1999) to the readers who have no previous knowledge of Fourier transforms. Carr and Madan (1999)'s contribution was rewriting the equation (1.1) in terms of a CF of the conditional log terminal stock price

$\phi(\ln S_T | \mathcal{F}_0)$:

$$C(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_r(\omega - (\alpha + 1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha + 1)\omega} d\omega. \quad (1.2)$$

And approximate the equation (1.2) using DFT (i.e. simply by taking a sample of size N):

$$C_T(k_n) \approx \frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k} \\ \times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\} \exp(-i2\pi jn/N), \quad (1.3)$$

which improves significant amount of computational time. This simple operation is critically important in the pricing with general exponential Lévy models. Why? Because without Fourier transform pricing approach, we cannot price options in general exponential Lévy models or we have to spend tremendous amount of energy just to come up with closed form solution like the VG model. The excellence of FT option pricing is its simplicity and generality and FT pricing works as long as a CF of the conditional log terminal stock price $S_T | \mathcal{F}_0$ is obtained in closed form.

The structure of this sequel is as follows. Chapter 2 provides readers with minimal necessary knowledge before learning Fourier transform. Chapter 3 defines FT and then, numerous examples are presented in order to intuitively understand the meaning of FT. We also present the important properties of FT. These are not necessary for the beginners, but we believe these will help readers as they proceed to more advanced integral transform pricing methods. In Chapter 4, a characteristic function is defined and its properties are discussed. We show, using several examples, how to obtain moments by using a CF (or characteristic exponent). Moment generating function is also dealt. Chapter 5 gives an introduction to the discrete Fourier transform which is just an approximation of FT. This approximation is done by sampling a finite number of points N of a continuous time domain function $g(t)$ with time domain sampling interval Δt (seconds) and sampling a finite number of points N of a continuous FT $\mathcal{G}(\omega)$ with angular frequency sampling interval $\Delta\omega$ Hz. In other words, both the original continuous time domain function $g(t)$ and the original continuous FT $\mathcal{G}(\omega)$ are approximated by a sample of N points. The use of DFT improves the computation time by a tremendous amount. Following Chapter 3, the examples of DFT and its properties are examined. In Chapter 6, a Lévy process is defined and its properties are discussed. We frequently use Lévy-Khintchin representation to obtain a CF of a Lévy process. Chapter 7 revisits the Black-Scholes model as an exponential Lévy model and its basic properties are reviewed. Chapter 8 gives Carr and Madan (1999)'s general FT call price and our version of general DFT call price. These general FT and DFT call prices are applied in the BS framework which shows that the original BS, BS-FT, and BS-DFT call prices are identical as expected. Chapter 9 illustrates the derivation of Merton JD model, factors which determine the skewness and the excess kurtosis of the log return density, Lévy measure of jump-diffusion process, and the traditional (i.e. PDE and martingale approach) option pricing with Merton JD model. Chapter 10 applies the general FT and DFT call price to the Merton JD model. This is simply done by substituting the CF of Merton JD log terminal stock price. Chapter 11 presents the derivation of VG model, factors which

determine the skewness and the excess kurtosis of the log return density, Lévy measure of VG process, and its closed form and numerical call price. Chapter 12 is an application of the general FT and DFT call price to the VG model which is simply done by substituting the CF of VG log terminal stock price. Chapter 13 gives concluding remarks.

[2] Prerequisite for Fourier Transform

In this section we present prerequisite knowledge before moving to Fourier transform.

[2.1] Radian

Radian is the unit of angle. A complete circle has $2\pi = 6.28319$ radians which is equal to 360° . This in turn means that one radian is equal to:

$$1 \text{ radian} = \frac{360^\circ}{2\pi} = 57.2958^\circ.$$

[2.2] Wavelength

Wavelength λ of a waveform is defined as a distance (d) between peaks or troughs. In other words, wavelength is the distance at which a waveform completes one cycle:

$$\lambda \equiv \frac{\text{distance}}{1 \text{ cycle}}. \quad (2.1)$$

Let v be the speed (distance/second), and f be the frequency (cycles/second, explained soon) of a waveform. These are related by:

$$\lambda(\text{distance / cycle}) \equiv \frac{v(\text{distance/second})}{f(\text{cycles/second})}. \quad (2.2)$$

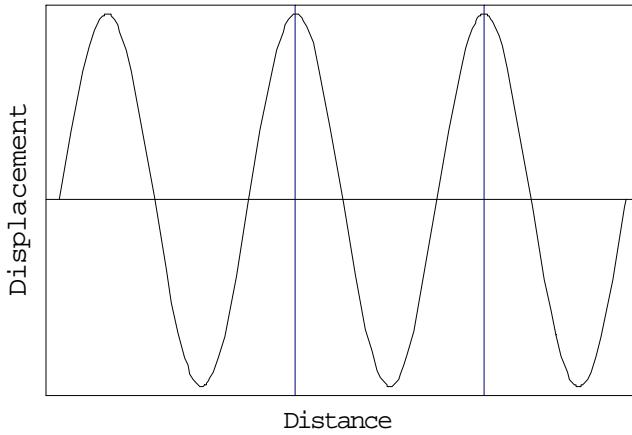


Figure 2.1: Illustration of Wavelength λ .

[2.3] Frequency, Angular Frequency, and Period of a Waveform

Period T of oscillation of a wave is the seconds (time) taken for a waveform to complete one wavelength:

$$T \equiv \frac{\text{seconds}}{1 \text{ wavelength (cycle)}}. \quad (2.3)$$

Period is by definition a reciprocal of a frequency. Let f be the frequency of a wave. Then:

$$T(\text{seconds/cycle}) \equiv \frac{1}{f(\text{cycles/second})}. \quad (2.4)$$

Frequency f of a wave measures the number of times for a wave to complete one wavelength (cycle) per second:

$$f \equiv \frac{\text{number of wavelengths (cycles)}}{1 \text{ second}}. \quad (2.5)$$

By definition, f is calculated as a reciprocal of the period of a wave:

$$f(\text{cycles/second}) \equiv \frac{1}{T(\text{seconds/cycle})}. \quad (2.6)$$

Frequency f is measured in Hertz (Hz). 1 Hz wave is a wave which completes one wavelength (cycle) per second. The frequency of the AC (alternating current) in U.S. is 60 Hz. Human beings can hear frequencies from about 20 to 20,000 Hz (called range of hearing). Alternatively, frequency can be calculated using the speed v and wavelength λ of a wave as:

$$f(\text{cycles/second}) \equiv \frac{v(\text{distance/second})}{\lambda(\text{distance/cycle})}. \quad (2.7)$$

Angular frequency (also called angular speed or radian frequency) ω is a measure of rotation rate (i.e. the speed at which an object rotates). The unit of measurement for ω is radians per 1 second. Since one cycle equals 2π radians, angular frequency ω is calculated as:

$$\omega(\text{radians/second}) = \frac{2\pi(\text{radians})}{T(\text{seconds/cycle})} = 2\pi f(\text{cycles/second}). \quad (2.8)$$

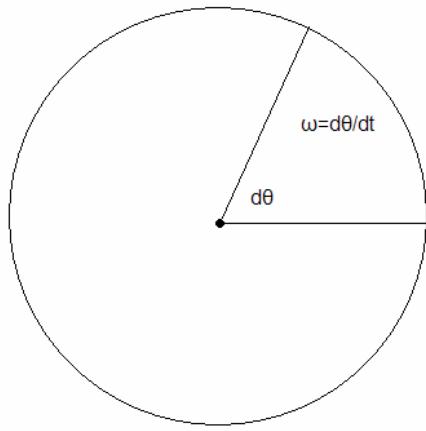


Figure 2.2: Definition of Angular Frequency.

Consider a sine wave $g(t) = \sin(2\pi(5)t)$ which is illustrated in Figure 2.3 for the time between 0 and 2 seconds. This sine wave has frequency $f = 5$ Hz (5 cycles per second) and angular frequency $\omega = 10\pi$ Hz (10 π radians per second). Its period is

$$T \equiv \frac{1}{f} = \frac{1}{5} = 0.2 \text{ (seconds/cycle)}.$$

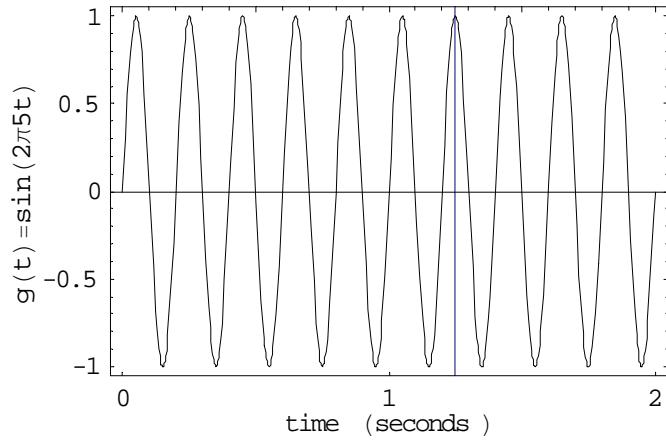


Figure 2.3: Plot of 5 Hz Sine Wave $g(t) = \sin(2\pi(5)t)$.

[2.4] Sine and Cosine

Let θ be an angle which is measured counterclockwise from the x -axis along an arc of a unit circle. Sine function ($\sin \theta$) is defined as a vertical coordinate of the arc endpoint. Cosine function ($\cos \theta$) is defined as a horizontal coordinate of the arc endpoint. Sine and cosine functions $\sin \theta$ and $\cos \theta$ are periodic functions with period 2π as illustrated in Figure 2.5:

$$\sin \theta = \sin(\theta + 2\pi h),$$

$$\cos \theta = \cos(\theta + 2\pi h),$$

where h is any integer.

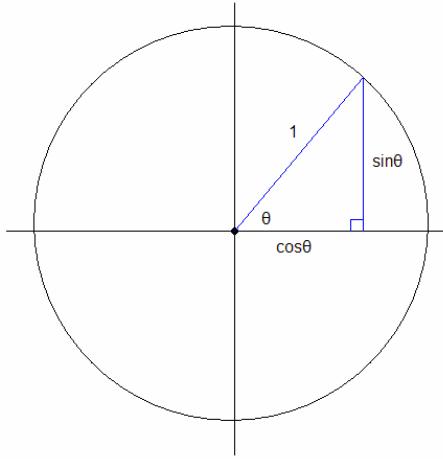


Figure 2.4: The Definition of Sine and Cosine Function with Unit Circle

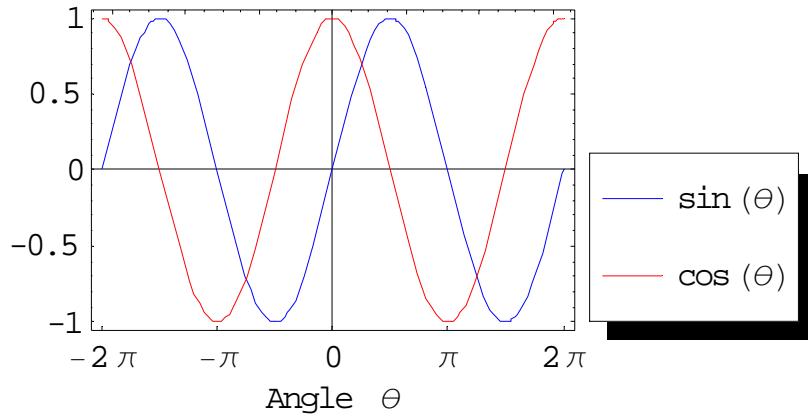


Figure 2.5: Sine and Cosine Function $\sin(\theta)$ and $\cos(\theta)$.

Following Pythagorean theorem, we have the identity:

$$\sin^2 \theta + \cos^2 \theta = 1. \quad (2.9)$$

[2.5] Derivative and Integral of Sine and Cosine Function

Let $\sin(x)$ and $\cos(x)$ be sine and cosine functions for $x \in \mathbb{R}$. The derivative of $\sin(x)$ can be expressed as:

$$\frac{d \sin(x)}{dx} = \cos(x). \quad (2.10)$$

The derivative of $\cos(x)$ can be expressed as:

$$\frac{d \cos(x)}{dx} = -\sin(x). \quad (2.11)$$

The integral of $\sin(x)$ can be expressed as:

$$\int_{-\infty}^{\infty} \sin(x) dx = -\cos(x). \quad (2.12)$$

Refer to any graduate school level trigonometry textbook for proofs.

[2.6] Series Definition of Sine Function and Cosine Function

For any $x \in \mathbb{R}$:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad (2.13)$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}. \quad (2.14)$$

Refer to any graduate school level trigonometry textbook for proofs.

[2.7] Euler's Formula

Euler's formula gives a very important relationship between the complex (imaginary) exponential function and the trigonometric functions. For any $x \in \mathbb{R}$:

$$e^{ix} = \cos(x) + i \sin(x). \quad (2.15)$$

From (2.15), variants of Euler's formula are derived:

$$e^{-ix} = \cos(x) - i \sin(x), \quad (2.16)$$

$$e^{ix} + e^{-ix} = 2 \cos(x), \quad (2.17)$$

$$e^{ix} - e^{-ix} = 2i \sin(x). \quad (2.18)$$

Consider sine and cosine functions with complex arguments z . Then, Euler's formula tells:

$$\sin z = \operatorname{Im}(e^{iz}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} = \frac{e^{iz} - e^{-iz}}{2i}, \quad (2.19)$$

$$\cos z = \operatorname{Re}(e^{iz}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} = \frac{e^{iz} + e^{-iz}}{2}. \quad (2.20)$$

Refer to any graduate school level trigonometry textbook for proofs.

[2.8] Sine Wave: Sinusoid

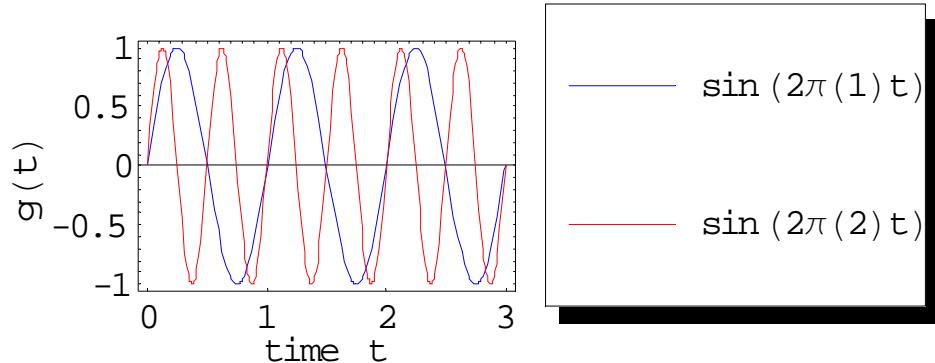
Sine wave is generally defined as a function of time t (seconds):

$$g(t) = a \sin(2\pi f_0 t + b), \quad (2.21)$$

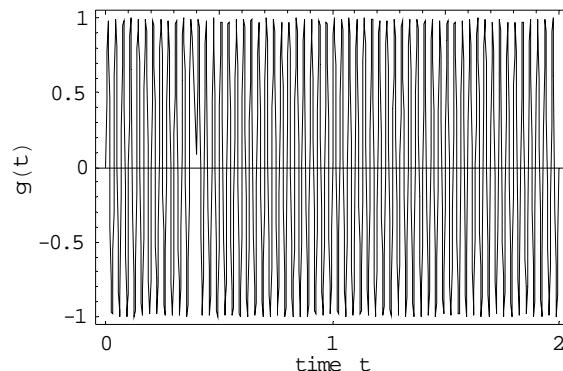
where a is the amplitude, f_0 is the fundamental frequency (cycles/second, Hz), and b changes the phase (angular position) of a sinusoid. In terms of a fundamental angular frequency ω_0 (radians/second), a sine wave is defined as (i.e. $\omega_0 \equiv 2\pi f_0$):

$$g(t) = a \sin(\omega_0 t + b). \quad (2.22)$$

Figure 2.6 illustrates the role of a fundamental frequency f_0 . When a fundamental frequency f_0 doubles from 1 (1 cycle /second) to 2 (cycles/second), its period becomes half from 1 to 1/2 seconds as illustrated in Panel A.



A) 1 Hz sine wave $\sin(2\pi(1)t)$ versus 2 Hz sine wave $\sin(2\pi(2)t)$.



B) 30 Hz sine wave $\sin(2\pi(30)t)$.

Figure 2.6: Plot of a Sine Wave $g(t) = \sin(2\pi ft)$ with Different Fundamental Frequency f_0 .

The role of amplitude a is to increase or decrease the magnitude of an oscillation. Figure 2.7 illustrates how magnitude of an oscillation changes for three different amplitudes $a = 1/2, 1$, and 2 . In audio study amplitude a determines how loud a sound is.

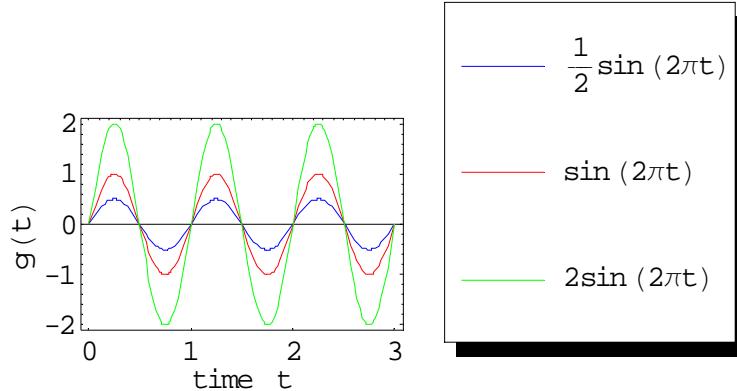


Figure 2.7: Plot of a 1 Hz Sine Wave $g(t) = a \sin(2\pi t)$ with Different Amplitudes $a = 1/2, 1$, and 2 .

Consider three 1 Hz sine waves $\sin(2\pi t)$, $\sin(2\pi t + \frac{\pi}{2})$, and $\sin(2\pi t - \frac{\pi}{2})$. A sine wave $\sin(2\pi t)$ has a phase 0 , $\sin(2\pi t + \frac{\pi}{2}) = \cos(2\pi t)$ has a phase $\pi/2$, and $\sin(2\pi t - \frac{\pi}{2})$ has a phase $-\pi/2$. The role of a parameter b is to change the position of a waveform by an amount b as illustrated in Figure 2.8.

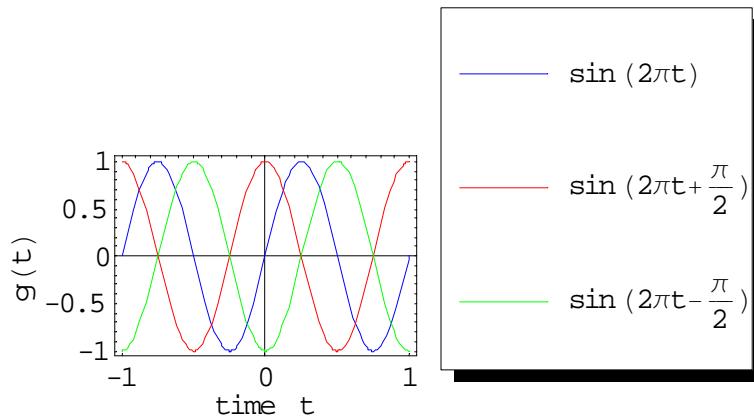


Figure 2.8: Plot of a 1 Hz Sine Wave $g(t) = \sin(2\pi t + b)$ with Different Phase $b = 0, \pi/2$, and $-\pi/2$.

This means that $g(t) = a \sin(2\pi f_0 t)$ is a sinusoid at phase zero and $g(t) = a \cos(2\pi f_0 t)$ is a sinusoid at phase $\pi/2$. For the purpose of defining a sinusoid, it really does not matter whether $\sin(\)$ or $\cos(\)$ is used.

[3] Fourier Transform (FT)

[3.1] Definition of Fourier Transform

We consider Fourier transform of a function $g(t)$ from a time domain t into an angular frequency domain ω (radians/second). This follows the convention in physics. In the field of signal processing which is a major application of FT, frequency f (cycles/second) is used instead of ω . But this difference is not important because ω and f are measuring the same thing (rotation speed) in different units and related by (2.8):

$$\omega = 2\pi f . \quad (3.1)$$

Table 3.1 gives a clear-cut relationship between frequency f and angular frequency ω .

Table 3.1: Relationship between Frequency f and Angular Frequency ω

Frequency f (cycles/second)	Angular Frequency ω (radians/second)
1 Hz	$2\pi = 360^\circ$ Hz
10 Hz	$20\pi = 360^\circ \times 10$ Hz
100 Hz	$200\pi = 360^\circ \times 100$ Hz

We start from the most general definition of FT. FT from a function $g(t)$ to a function $\mathcal{G}(\omega)$ (thus, switching domains from t to ω) is defined using two arbitrary constants a and b called FT parameters as:

$$\mathcal{G}(\omega) \equiv \mathcal{F}[g(t)](\omega) \equiv \sqrt{\frac{|b|}{(2\pi)^{1-a}}} \int_{-\infty}^{\infty} e^{ib\omega t} g(t) dt . \quad (3.2)$$

Inverse Fourier transform from a function $\mathcal{G}(\omega)$ to a function $g(t)$ (thus, switching domains from ω to t) is defined as (i.e. the reverse procedure of (3.2)):

$$g(t) \equiv \mathcal{F}^{-1}[\mathcal{G}(\omega)](t) \equiv \sqrt{\frac{|b|}{(2\pi)^{1+a}}} \int_{-\infty}^{\infty} e^{-ib\omega t} \mathcal{G}(\omega) d\omega . \quad (3.3)$$

For our purpose which is to calculate characteristic functions, FT parameters are set as $(a, b) = (1, 1)$. Thus, (3.2) and (3.3) become:

$$\mathcal{G}(\omega) \equiv \mathcal{F}[g(t)](\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt , \quad (3.4)$$

$$g(t) \equiv \mathcal{F}^{-1}[\mathcal{G}(\omega)](t) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{G}(\omega) d\omega . \quad (3.5)$$

Euler's formula (2.15) is for $t \in \mathbb{R}$:

$$e^{it} = \cos t + i \sin t .$$

Thus, the FT of (3.4) can be rewritten as:

$$\mathcal{G}(\omega) \equiv \mathcal{F}[g(t)](\omega) \equiv \int_{-\infty}^{\infty} \cos(\omega t) g(t) dt + i \int_{-\infty}^{\infty} \sin(\omega t) g(t) dt .$$

Intuitively speaking, FT is a decomposition of a waveform $g(t)$ (i.e. in time domain t) into a sum of sinusoids (i.e. sine and cosine functions) of different frequencies (Hz) which sum to the original waveform. In other words, FT enables any function in time domain to be represented by an infinite number of sinusoidal functions. Therefore, FT is an angular frequency representation (i.e. different look) of a function $g(t)$ and $\mathcal{G}(\omega)$ contains the exact same information as the original function $g(t)$.

We can check if the inverse Fourier transform (3.5) is true:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{G}(\omega) d\omega &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{i\omega\tau} g(\tau) d\tau \right) e^{-i\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\tau) \left(\int_{-\infty}^{\infty} e^{i\omega(\tau-t)} d\omega \right) d\tau \\ &= \int_{-\infty}^{\infty} g(\tau) \delta(\tau - t) d\tau \\ &= g(t), \end{aligned}$$

where we used the identity of Dirac's delta function (3.12) which is proven soon.

Although FT parameters $(a, b) = (1, 1)$ are used for calculating characteristic functions, different pairs of (a, b) are used for other purposes. For example, $(a, b) = (1, -1)$ in pure mathematics:

$$\begin{aligned} \mathcal{G}(\omega) &\equiv \mathcal{F}_t[g(t)](\omega) \equiv \int_{-\infty}^{\infty} e^{-i\omega t} g(t) dt, \\ g(t) &\equiv \mathcal{F}_{\omega}^{-1}[\mathcal{G}(\omega)](t) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \mathcal{G}(\omega) d\omega. \end{aligned}$$

Modern physics uses $(a, b) = (0, 1)$:

$$\begin{aligned} \mathcal{G}(\omega) &\equiv \mathcal{F}_t[g(t)](\omega) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt, \\ g(t) &\equiv \mathcal{F}_{\omega}^{-1}[\mathcal{G}(\omega)](t) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{G}(\omega) d\omega. \end{aligned}$$

In the field of signal processing and most of the standard textbooks of FT, FT parameters $(a, b) = (0, -2\pi)$ are used (i.e. they use frequency f instead of angular frequency ω):

$$\mathcal{G}(f) \equiv \mathcal{F}_t[g(t)](f) \equiv \int_{-\infty}^{\infty} e^{-2\pi i f t} g(t) dt, \quad (3.6)$$

$$g(t) \equiv \mathcal{F}_f^{-1}[\mathcal{G}(f)](t) \equiv \int_{-\infty}^{\infty} e^{2\pi i f t} \mathcal{G}(f) df. \quad (3.7)$$

We use (3.6) and (3.7) frequently because this definition of FT is mathematically simpler to handle for the purpose of proofs.

In general, Fourier transform $\mathcal{G}(\omega)$ is a complex quantity:

$$\mathcal{G}(\omega) = \text{Re}(\omega) + i \text{Im}(\omega) = |\mathcal{G}(\omega)| e^{i\theta(\omega)}, \quad (3.8)$$

where $\text{Re}(\omega)$ is the real part of the FT $\mathcal{G}(\omega)$, $\text{Im}(\omega)$ is the imaginary part of the FT, $|\mathcal{G}(\omega)|$ is the amplitude of a time domain function $g(t)$, and $\theta(\omega)$ is the phase angle of the FT $\mathcal{G}(\omega)$. $|\mathcal{G}(\omega)|$ and $\theta(\omega)$ can be expressed in terms of $\text{Re}(\omega)$ and $\text{Im}(\omega)$ as:

$$|\mathcal{G}(\omega)| = \sqrt{\text{Re}^2(\omega) + \text{Im}^2(\omega)}, \quad (3.9)$$

$$\theta(\omega) = \tan^{-1} \left[\frac{\text{Im}(\omega)}{\text{Re}(\omega)} \right]. \quad (3.10)$$

[3.2] Examples of Fourier Transform

Before discussing important properties of FT, we present representative examples of FT in this section to get the feeling of what FT does.

[3.2.1] Double-Sided Exponential Function

Consider a double-sided exponential function with $A, \alpha \in \mathbb{R}$:

$$g(t) = A e^{-\alpha|t|}.$$

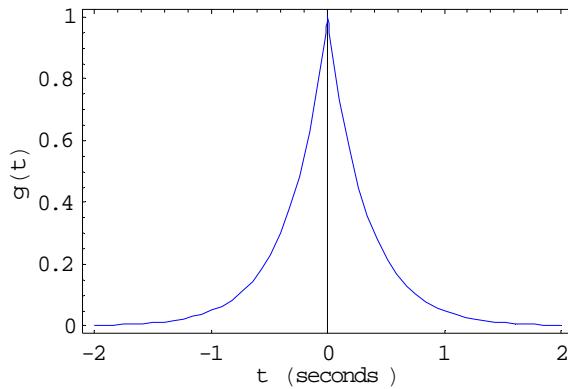
From (3.4):

$$\begin{aligned} \mathcal{G}(\omega) &\equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt = \int_{-\infty}^{\infty} e^{i\omega t} A e^{-\alpha|t|} dt = A \left(\int_{-\infty}^0 e^{i\omega t} e^{\alpha t} dt + \int_0^{\infty} e^{i\omega t} e^{-\alpha t} dt \right) \\ \mathcal{G}(\omega) &= A \left(\frac{1}{\alpha + i\omega} + \frac{1}{\alpha - i\omega} \right) = \frac{2A\alpha}{\alpha^2 + \omega^2}. \end{aligned}$$

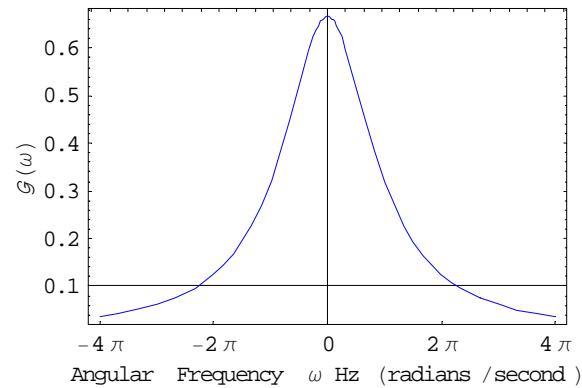
When $A = 1$ and $\alpha = 3$, the time domain function $g(t) = e^{-3|t|}$ and its Fourier transform

$$\mathcal{G}(\omega) = \frac{6}{9 + \omega^2}$$

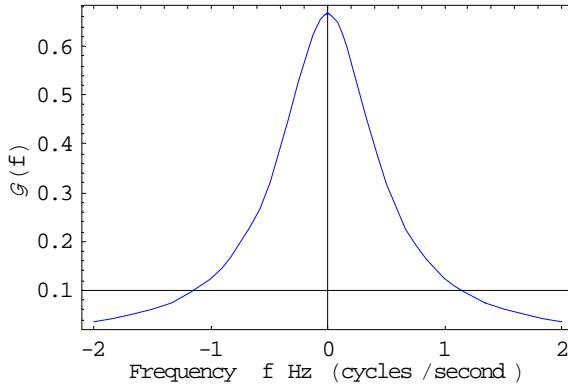
is plotted in Figure 3.1.



A) Plot of a double-sided exponential function $g(t) = e^{-3|t|}$.



B) Plot of FT of $g(t) = e^{-3|t|}$ in Angular Frequency Domain, $\mathcal{G}(\omega)$.



C) Plot of FT of $g(t) = e^{-3|t|}$ in Frequency Domain, $\mathcal{G}(f)$.

Figure 3.1: Plot of Double-Sided Exponential Function $g(t)$ and Its Fourier Transforms $\mathcal{G}(\omega)$ and $\mathcal{G}(f)$.

Using the signal processing definition of $\text{FT}(a,b) = (0, -2\pi)$ of the definition (3.6), FT of $g(t) = Ae^{-\alpha|t|}$ is computed as (which is simply obtained by substituting $\omega = 2\pi f$ into $\mathcal{G}(\omega)$):

$$\mathcal{G}(f) = \frac{2A\alpha}{\alpha^2 + 4\pi^2 f^2}.$$

When $A = 1$ and $\alpha = 3$, $\mathcal{G}(f) = \frac{6}{9 + 4\pi^2 f^2}$ which is plotted in Panel C of Figure 3.1.

[3.2.2] Rectangular Pulse

Consider a rectangular pulse with $A, T_0 \in \mathbb{R}$:

$$g(t) = \begin{cases} A & -T_0 \leq t \leq T_0 \\ 0 & |t| > T_0 \end{cases},$$

which is an even function of t (symmetric with respect to t).

From (3.4) and use Euler's formula (2.18) $e^{ix} - e^{-ix} = 2i \sin(x)$:

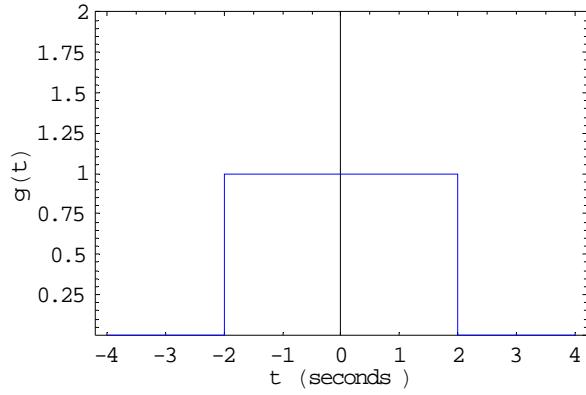
$$\begin{aligned} \mathcal{G}(\omega) &\equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt = A \int_{-T_0}^{T_0} e^{i\omega t} dt = A \left\{ \frac{-\left(e^{-i\omega T_0} - e^{i\omega T_0}\right)}{i\omega} \right\}, \\ &= A \left\{ \frac{2i \sin(\omega T_0)}{i\omega} \right\} = \frac{2A \sin(\omega T_0)}{\omega}. \end{aligned}$$

Using the signal processing definition of $\text{FT}(a, b) = (0, -2\pi)$ of the definition (3.6), FT of a rectangular pulse is computed as:

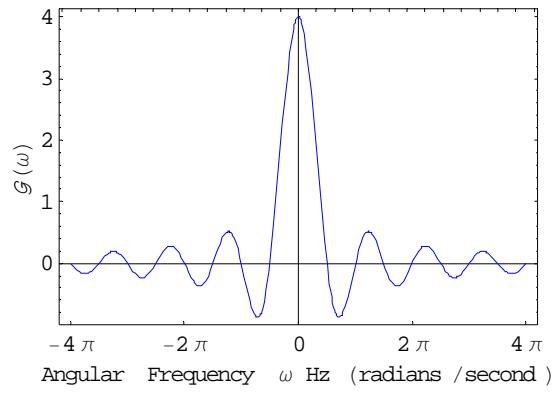
$$\mathcal{G}(f) \equiv \int_{-\infty}^{\infty} e^{-2\pi ift} g(t) dt = A \int_{-T_0}^{T_0} e^{-2\pi ift} dt = \frac{A \sin(2\pi f T_0)}{\pi f}.$$

When $A = 1$ and $T_0 = 2$, the time domain function $g(t)$, Fourier transform

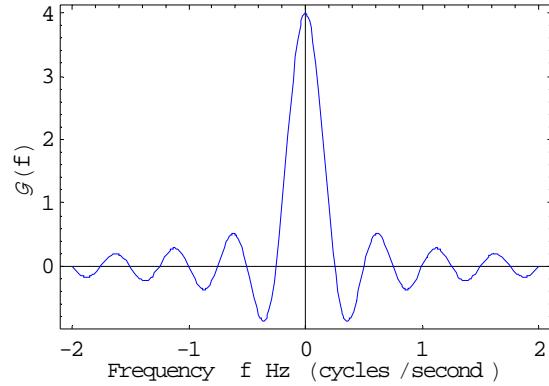
$\mathcal{G}(\omega) = \frac{2 \sin(2\omega)}{\omega}$ in angular frequency, and Fourier transform $\mathcal{G}(f) = \frac{\sin(4\pi f)}{\pi f}$ in frequency domain are plotted in Figure 3.2.



A) Plot of a rectangular pulse $g(t)$.



B) Plot of the FT of $g(t)$ in Angular Frequency Domain, $\mathcal{G}(\omega)$.



C) Plot of the FT of $g(t)$ in Frequency Domain, $\mathcal{G}(f)$.

Figure 3.2: Plot of Rectangular Pulse $g(t)$ and Its Fourier Transforms $\mathcal{G}(\omega)$ and $\mathcal{G}(f)$.

[3.2.3] Dirac's Delta Function (Impulse Function)

Consider Dirac's delta function scaled by $a \in \mathbb{R}^+$ (Dirac's delta function is discussed in detail in section 3.3.1.):

$$g(t) = a\delta(t).$$

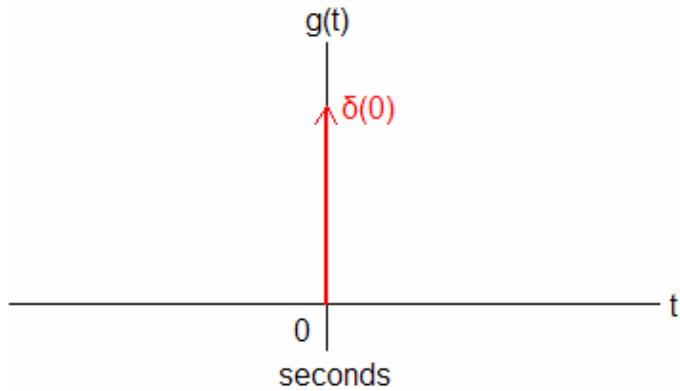
From (3.4):

$$\mathcal{G}(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt = \int_{-\infty}^{\infty} e^{i\omega t} a\delta(t) dt = ae^{i\omega 0} = a.$$

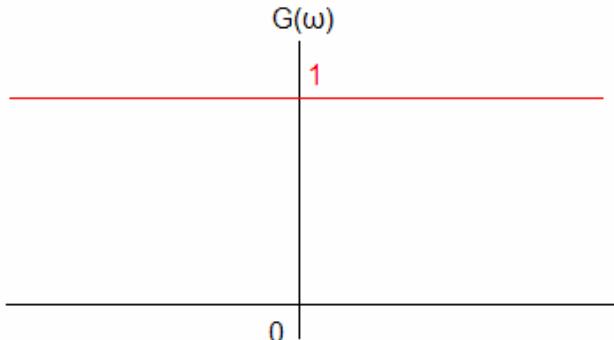
Using the signal processing definition of $\text{FT}(a, b) = (0, -2\pi)$ of the definition (3.6), FT of a scaled Dirac's delta is computed as:

$$\mathcal{G}(f) \equiv \int_{-\infty}^{\infty} e^{-2\pi ift} g(t) dt = \int_{-\infty}^{\infty} e^{-2\pi ift} a\delta(t) dt = ae^{-2\pi if 0} = a.$$

When $a = 1$ (i.e. pure Dirac's delta), the time domain function $g(t) = \delta(t)$ and Fourier transforms $\mathcal{G}(\omega) = 1$ and $\mathcal{G}(f) = 1$ are plotted in Figure 3.3.

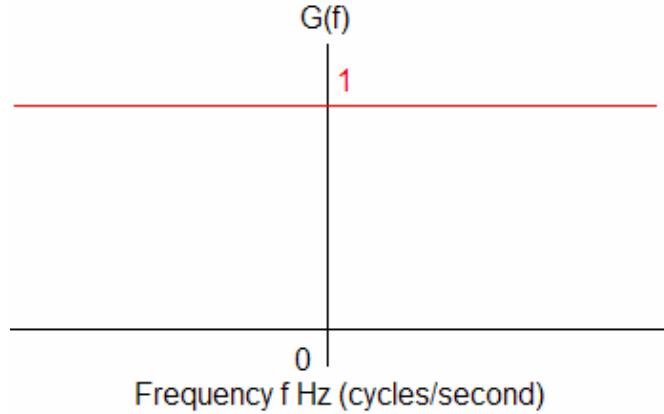


A) Plot of $g(t) = \delta(t)$.



Angular Frequency ω Hz (radians/second)

B) Plot of FT of $g(t)$ in Angular Frequency Domain, $\mathcal{G}(\omega) = 1$.



C) Plot of FT of $g(t)$ in Frequency Domain, $\mathcal{G}(f) = 1$.

Figure 3.3: Plot of Dirac's Delta Function $g(t) = \delta(t)$ and Its Fourier Transforms $\mathcal{G}(\omega)$ and $\mathcal{G}(f)$.

[3.2.4] Gaussian Function

Consider a Gaussian function with $A \in \mathbb{R}^+$:

$$g(t) = e^{-At^2}.$$

From (3.4):

$$\mathcal{G}(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt = \int_{-\infty}^{\infty} e^{i\omega t} e^{-At^2} dt.$$

Use Euler's formula (2.15):

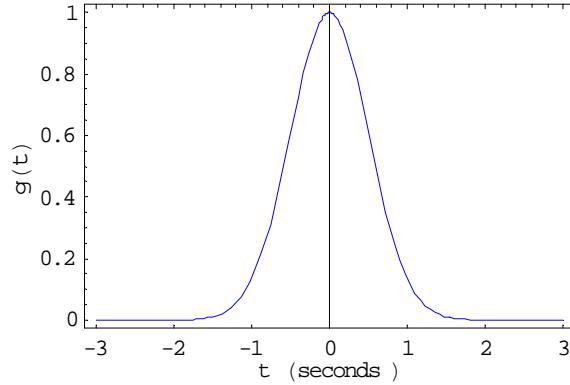
$$\begin{aligned} \mathcal{G}(\omega) &= \int_{-\infty}^{\infty} e^{-At^2} \{ \cos(\omega t) + i \sin(\omega t) \} dt \\ &= \int_{-\infty}^{\infty} e^{-At^2} \cos(\omega t) dt + i \int_{-\infty}^{\infty} e^{-At^2} \sin(\omega t) dt \\ &= \sqrt{\frac{\pi}{A}} e^{-\omega^2/4A} + i0 = \sqrt{\frac{\pi}{A}} e^{-\omega^2/4A}. \end{aligned}$$

Using the signal processing definition of $\text{FT}(a, b) = (0, -2\pi)$ of the definition (3.6), FT of a Gaussian function is computed as:

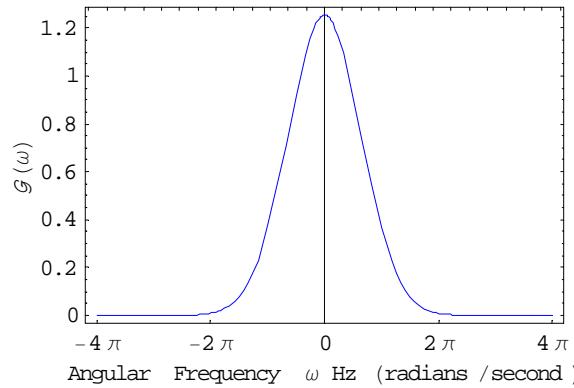
$$\mathcal{G}(f) \equiv \int_{-\infty}^{\infty} e^{-2\pi ift} g(t) dt = \int_{-\infty}^{\infty} e^{-2\pi ift} e^{-At^2} dt = \sqrt{\frac{\pi}{A}} e^{-\pi^2 f^2 / A}.$$

When $A = 2$, the time domain function $g(t) = e^{-2t^2}$ and its Fourier transforms

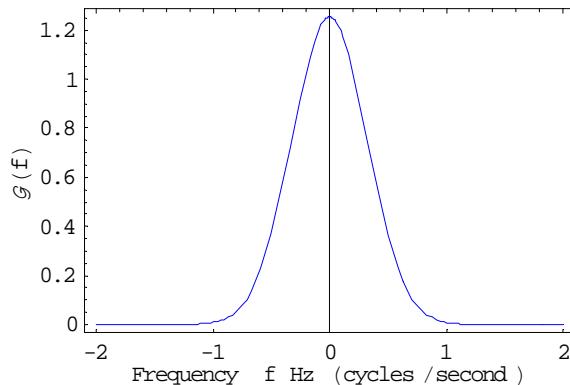
$\mathcal{G}(\omega) = \sqrt{\frac{\pi}{2}} e^{-\omega^2/8}$ and $\mathcal{G}(f) = \sqrt{\frac{\pi}{2}} e^{-\pi^2 f^2/2}$ are plotted in Figure 3.4. Note that FT of a Gaussian Function is another Gaussian function.



A) Plot of Gaussian function $g(t) = e^{-2t^2}$.



B) Plot of FT of $g(t)$ in Angular Frequency Domain, $\mathcal{G}(\omega)$.



C) Plot of FT of $g(t)$ in Frequency Domain, $\mathcal{G}(f)$.

Figure 3.4: Plot of Gaussian Function $g(t) = e^{-2t^2}$ and Fourier Transforms $\mathcal{G}(\omega)$ and $\mathcal{G}(f)$.

[3.2.5] Cosine Wave $g(t) = A \cos(2\pi f_0 t) = A \cos(\omega_0 t)$

Consider a sinusoid $g(t) = A \cos(2\pi f_0 t) = A \cos(\omega_0 t)$. Using the signal processing definition of $\text{FT}(a, b) = (0, -2\pi)$ of the definition (3.6), FT of a cosine wave is computed as:

$$\mathcal{G}(f) \equiv \int_{-\infty}^{\infty} e^{-2\pi i ft} g(t) dt = \int_{-\infty}^{\infty} e^{-2\pi i ft} A \cos(2\pi f_0 t) dt.$$

From Euler's formula (2.17):

$$\begin{aligned} \mathcal{G}(f) &= A \int_{-\infty}^{\infty} e^{-2\pi i ft} \frac{1}{2} (e^{i2\pi f_0 t} + e^{-i2\pi f_0 t}) dt \\ &= \frac{1}{2} A \left\{ \int_{-\infty}^{\infty} e^{-2\pi i ft} e^{i2\pi f_0 t} dt + \int_{-\infty}^{\infty} e^{-2\pi i ft} e^{-i2\pi f_0 t} dt \right\} \\ &= \frac{1}{2} A \left\{ \int_{-\infty}^{\infty} e^{-2\pi i(f-f_0)t} dt + \int_{-\infty}^{\infty} e^{-2\pi i(f+f_0)t} dt \right\}. \end{aligned}$$

Use the identity (3.12) of Dirac's delta function:

$$\delta(x-a) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-a)} d\omega.$$

Therefore, we obtain:

$$\begin{aligned} \mathcal{G}(f) &= \frac{1}{2} A \{ \delta(f-f_0) + \delta(f+f_0) \} \\ \mathcal{G}(f) &= \frac{A}{2} \delta(f-f_0) + \frac{A}{2} \delta(f+f_0), \end{aligned}$$

which is two impulse functions at $f = \pm f_0$. Thus, FT of a cosine wave (which is an even function) is a real valued even function which means $\mathcal{G}(f)$ is symmetric about $f = 0$.

Next, in terms of angular frequency ω from (3.4):

$$\mathcal{G}(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt = \int_{-\infty}^{\infty} e^{i\omega t} A \cos(\omega_0 t) dt.$$

From Euler's formula (2.17):

$$e^{ix} + e^{-ix} = 2 \cos(x).$$

Therefore:

$$\begin{aligned}
\mathcal{G}(\omega) &= A \int_{-\infty}^{\infty} e^{i\omega t} \cos(\omega_0 t) dt = A \int_{-\infty}^{\infty} e^{i\omega t} \left\{ \frac{1}{2} (e^{i\omega_0 t} + e^{-i\omega_0 t}) \right\} dt \\
&= \frac{A}{2} \int_{-\infty}^{\infty} e^{i\omega t} e^{i\omega_0 t} dt + \frac{A}{2} \int_{-\infty}^{\infty} e^{i\omega t} e^{-i\omega_0 t} dt = \frac{A}{2} \int_{-\infty}^{\infty} e^{i(\omega+\omega_0)t} dt + \frac{A}{2} \int_{-\infty}^{\infty} e^{i(\omega-\omega_0)t} dt.
\end{aligned}$$

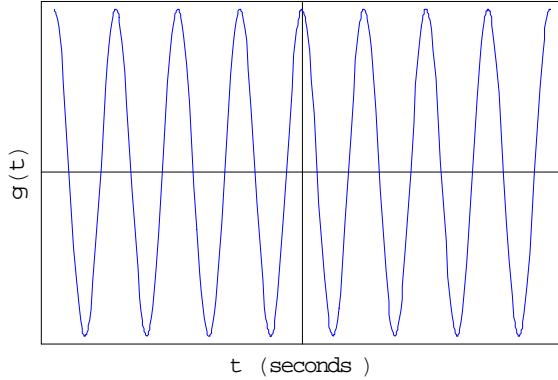
Use the identity of Dirac's delta function (3.12):

$$\delta(x-a) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-a)} d\omega.$$

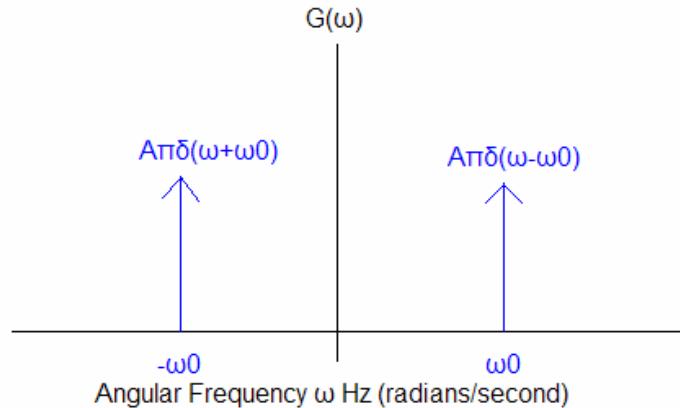
Thus:

$$\begin{aligned}
\mathcal{G}(\omega) &= \frac{A}{2} 2\pi \delta(\omega + \omega_0) + \frac{A}{2} 2\pi \delta(\omega - \omega_0) \\
\mathcal{G}(\omega) &= A\pi \delta(\omega + \omega_0) + A\pi \delta(\omega - \omega_0),
\end{aligned}$$

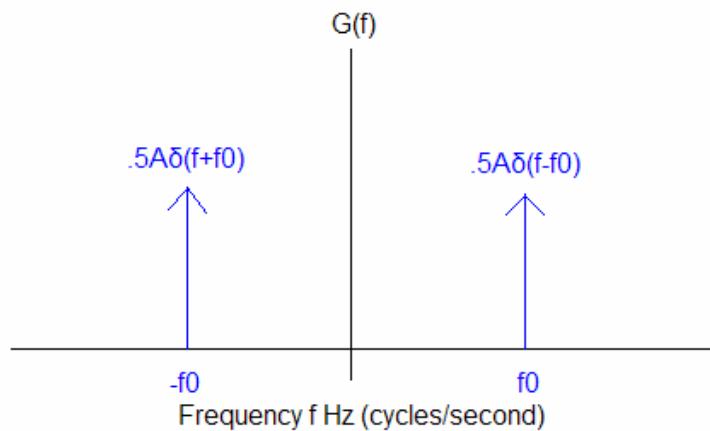
which is two impulse functions at $\omega = \pm\omega_0$. Figure 3.5 illustrates a cosine wave $g(t) = A \cos(2\pi f_0 t) = A \cos(\omega_0 t)$ and FTs $\mathcal{G}(\omega) = A\pi \delta(\omega + \omega_0) + A\pi \delta(\omega - \omega_0)$ and $\mathcal{G}(f) = \frac{A}{2} \delta(f - f_0) + \frac{A}{2} \delta(f + f_0)$.



A) Plot of a Cosine Wave $g(t) = A \cos(2\pi f_0 t) = A \cos(\omega_0 t)$. Amplitude of the wave is given by A .



B) Plot of FT of $g(t)$ in Angular Frequency Domain, $\mathcal{G}(\omega)$.



C) Plot of FT of $g(t)$ in Frequency Domain, $\mathcal{G}(f)$.

Figure 3.5: Plot of a Cosine Wave $g(t) = A \cos(2\pi f_0 t) = A \cos(\omega_0 t)$ and Fourier Transforms $\mathcal{G}(\omega)$ and $\mathcal{G}(f)$.

[3.2.6] Sine Wave $g(t) = A \sin(2\pi f_0 t) = A \sin(\omega_0 t)$

Consider a sinusoid $g(t) = A \sin(2\pi f_0 t) = A \sin(\omega_0 t)$. Using the signal processing definition of $\text{FT}(a, b) = (0, -2\pi)$ of the equation (3.6), FT of a sine wave is computed as:

$$\mathcal{G}(f) \equiv \int_{-\infty}^{\infty} e^{-2\pi i f t} g(t) dt = \int_{-\infty}^{\infty} e^{-2\pi i f t} A \sin(2\pi f_0 t) dt.$$

From Euler's formula (2.18):

$$\begin{aligned} \mathcal{G}(f) &= A \int_{-\infty}^{\infty} e^{-2\pi i f t} \frac{1}{2i} (e^{i2\pi f_0 t} - e^{-i2\pi f_0 t}) dt \\ &= \frac{1}{2i} A \left\{ \int_{-\infty}^{\infty} e^{-2\pi i f t} e^{i2\pi f_0 t} dt - \int_{-\infty}^{\infty} e^{-2\pi i f t} e^{-i2\pi f_0 t} dt \right\} \end{aligned}$$

$$= \frac{1}{2i} A \left\{ \int_{-\infty}^{\infty} e^{-2\pi i(f-f_0)t} dt - \int_{-\infty}^{\infty} e^{-2\pi i(f+f_0)t} dt \right\}.$$

Multiply $1 = i/i$ to the right hand side:

$$\begin{aligned} \mathcal{G}(f) &= \frac{i}{2i^2} A \left\{ \int_{-\infty}^{\infty} e^{-2\pi i(f-f_0)t} dt - \int_{-\infty}^{\infty} e^{-2\pi i(f+f_0)t} dt \right\} \\ &= \frac{i}{2} A \left\{ \int_{-\infty}^{\infty} e^{-2\pi i(f+f_0)t} dt - \int_{-\infty}^{\infty} e^{-2\pi i(f-f_0)t} dt \right\}. \end{aligned}$$

Using the identity (3.12) of Dirac's delta function, we obtain:

$$\begin{aligned} \mathcal{G}(f) &= \frac{i}{2} A \{ \delta(f + f_0) - \delta(f - f_0) \} \\ &= i \frac{A}{2} \delta(f + f_0) - i \frac{A}{2} \delta(f - f_0), \end{aligned}$$

which is two complex impulse functions at $f = \pm f_0$ which are not symmetric about $f = 0$.

Next, in terms of angular frequency ω from (3.4):

$$\mathcal{G}(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt = A \int_{-\infty}^{\infty} e^{i\omega t} \sin(\omega_0 t) dt.$$

Using Euler's formula (2.18):

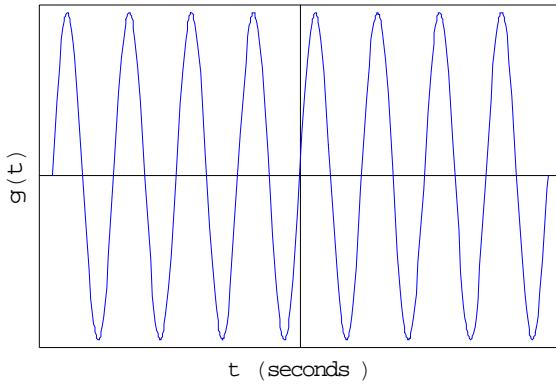
$$\begin{aligned} \mathcal{G}(\omega) &= A \int_{-\infty}^{\infty} e^{i\omega t} \left\{ \frac{1}{2i} (e^{i\omega_0 t} - e^{-i\omega_0 t}) \right\} dt = \frac{A}{2i} \int_{-\infty}^{\infty} e^{i\omega t} e^{i\omega_0 t} dt - \frac{A}{2i} \int_{-\infty}^{\infty} e^{i\omega t} e^{-i\omega_0 t} dt \\ &= \frac{A}{2i} \int_{-\infty}^{\infty} e^{i(\omega+\omega_0)t} dt - \frac{A}{2i} \int_{-\infty}^{\infty} e^{i(\omega-\omega_0)t} dt = \frac{Ai}{2i^2} \int_{-\infty}^{\infty} e^{i(\omega+\omega_0)t} dt - \frac{Ai}{2i^2} \int_{-\infty}^{\infty} e^{i(\omega-\omega_0)t} dt \\ &= \frac{Ai}{2} \int_{-\infty}^{\infty} e^{i(\omega-\omega_0)t} dt - \frac{Ai}{2} \int_{-\infty}^{\infty} e^{i(\omega+\omega_0)t} dt \end{aligned}$$

Use the identity (3.12) of Dirac's delta function:

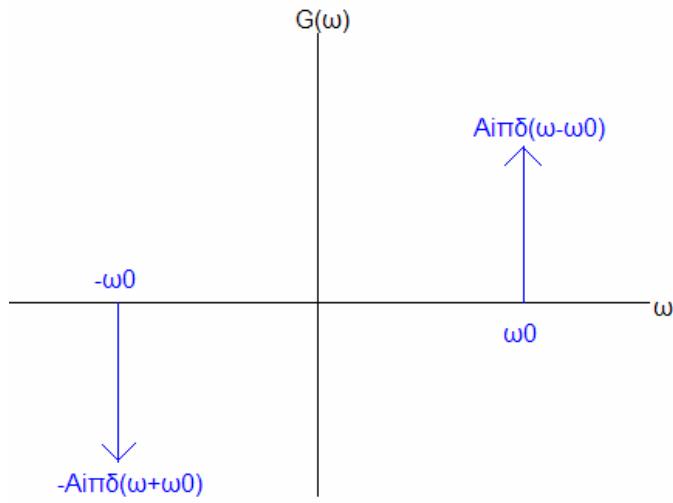
$$\begin{aligned} \mathcal{G}(\omega) &= \frac{Ai}{2} 2\pi \delta(\omega - \omega_0) - \frac{Ai}{2} 2\pi \delta(\omega + \omega_0) \\ &= Ai\pi\delta(\omega - \omega_0) - Ai\pi\delta(\omega + \omega_0). \end{aligned}$$

Figure 3.6 plots a sine wave $g(t) = A \sin(2\pi f_0 t) = A \sin(\omega_0 t)$ and its Fourier transforms

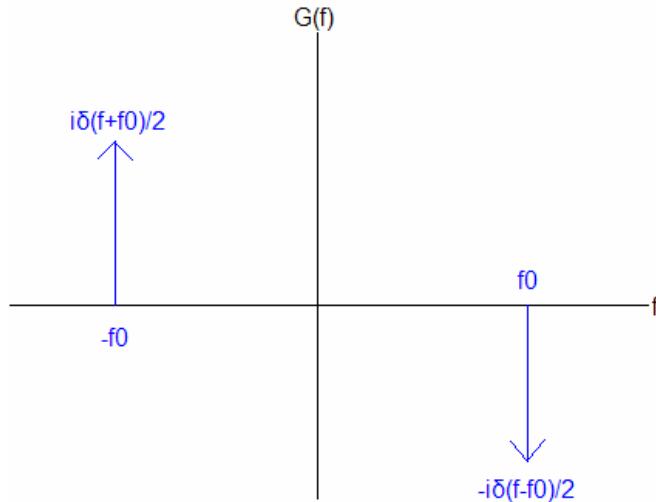
$$\mathcal{G}(\omega) = Ai\pi\delta(\omega - \omega_0) - Ai\pi\delta(\omega + \omega_0) \text{ and } \mathcal{G}(f) = i \frac{A}{2} \delta(f + f_0) - i \frac{A}{2} \delta(f - f_0).$$



A) Plot of a Sine Wave $g(t) = A \sin(2\pi f_0 t) = A \sin(\omega_0 t)$.



B) Plot of FT of $g(t)$ in Angular Frequency Domain, $\mathcal{G}(\omega)$.



C) Plot of FT of $g(t)$ in Frequency Domain, $\mathcal{G}(f)$.

Figure 3.6: Plot of a Sine Wave $g(t) = A \sin(2\pi f_0 t) = A \sin(\omega_0 t)$ and Fourier Transforms $\mathcal{G}(\omega)$ and $\mathcal{G}(f)$.

[3.3] Properties of Fourier Transform

We will discuss important properties of Fourier transform in this section starting from Dirac's delta function which is essential to the understanding of properties of Fourier transform.

[3.3.1] Dirac's Delta Function (Impulse Function)

Consider a function of the form with $n \in \mathbb{R}^+$:

$$h(x) = \frac{n}{\sqrt{\pi}} \exp(-n^2 x^2). \quad (3.11)$$

This function is plotted in Figure 3.7 for three different values for n . The function $h(x)$ becomes more and more concentrated around zero as the value of n increases. The function $h(x)$ has a unit integral:

$$\int_{-\infty}^{\infty} h(x) dx = \int_{-\infty}^{\infty} \frac{n}{\sqrt{\pi}} \exp(-n^2 x^2) dx = 1.$$

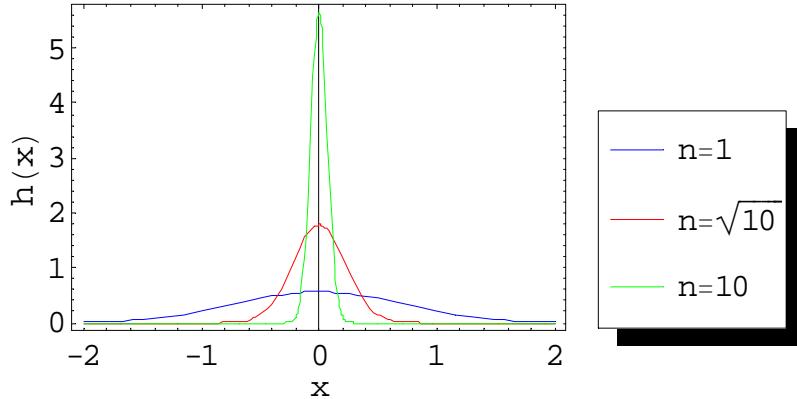


Figure 3.7: Plot of A Function $h(x)$ for $n = 1$, $n = 10^{1/2}$, and $n = 10$.

Dirac's delta function denoted by $\delta(x)$ can be considered as a limit of $h(x)$ when $n \rightarrow \infty$. In other words, $\delta(x)$ is a pulse of unbounded height and zero width with a unit integral:

$$\int_{-\infty}^{\infty} \delta(x) dx = 1.$$

Dirac's delta function $\delta(x)$ evaluates to 0 at all $x \in \mathbb{R}$ other than $x = 0$:

$$\delta(x) = \begin{cases} \delta(0) & \text{if } x=0 \\ 0 & \text{otherwise} \end{cases}$$

where $\delta(0)$ is undefined. $\delta(x)$ is called a generalized function not a function because of undefined $\delta(0)$. Therefore, $\delta(x)$ is a distribution with compact support $\{0\}$ meaning that $\delta(x)$ does not occur alone but occurs combined with any continuous functions $f(x)$ and is well defined only when it is integrated.

Dirac's delta function can be defined more generally by its sampling property. Suppose that a function $f(x)$ is defined at $x=0$. Applying $\delta(x)$ to $f(x)$ yields $f(0)$:

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0).$$

This is why Dirac's delta function $\delta(x)$ is called a functional because the use of $\delta(x)$ assigns a number $f(0)$ to a function $f(x)$. More generally for $a \in \mathbb{R}$:

$$\delta(x-a) = \begin{cases} \delta(0) & \text{if } x=a \\ 0 & \text{otherwise} \end{cases}$$

and:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a),$$

or for $\varepsilon > 0$:

$$\int_{a-\varepsilon}^{a+\varepsilon} f(x)\delta(x-a)dx = f(a).$$

$\delta(x)$ has identities such as:

$$\begin{aligned} \delta(ax) &= \frac{1}{|a|}\delta(x), \\ \delta(x^2 - a^2) &= \frac{1}{2|a|}[\delta(x+a) + \delta(x-a)]. \end{aligned}$$

Dirac's delta function $\delta(x)$ can be defined as the limit $n \rightarrow \infty$ of a class of delta sequences:

$$\delta(x) = \lim_{n \rightarrow \infty} \delta_n(x),$$

such that:

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \delta_n(x) f(x) dx = f(0),$$

where $\delta_n(x)$ is a class of delta sequences. Examples of $\delta_n(x)$ other than (3.8) are:

$$\begin{aligned}\delta_n(x) &= \begin{cases} n & \text{if } -1/2n < x < 1/2n \\ 0 & \text{otherwise} \end{cases}, \\ \delta_n(x) &= \frac{1}{2\pi} \int_{-n}^n \exp(iux) du, \\ \delta_n(x) &= \frac{1}{\pi x} \frac{e^{inx} - e^{-inx}}{2i}, \\ \delta_n(x) &= \frac{1}{2\pi} \frac{\sin[(n+1/2)x]}{\sin(x/2)}.\end{aligned}$$

[3.3.2] Useful Identity: Dirac's Delta Function

Dirac's delta function $\delta(x)$ has the following very useful identity which we have used many times before:

$$\delta(x-a) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-a)} d\omega. \quad (3.12)$$

PROOF¹

First step is to prove a proposition for all $d = 2, 3, 4, \dots$ and $j = 0, 1, 2, \dots, d-1$ (note that j depends on d):

$$\frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) = 1_{j=0} = \begin{cases} 1 & \text{if } j=0 \\ 0 & \text{otherwise} \end{cases} \quad (3.13)$$

First, we deal an informal proof of a proposition (3.13).

When $d = 2$ and $j = 0$:

$$\frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) = \frac{1}{2} \sum_{k=0}^{2-1} \exp\left(\frac{2\pi i}{2} 0k\right) = \frac{1}{2} \left\{ \exp\left(\frac{2\pi i}{2} 00\right) + \exp\left(\frac{2\pi i}{2} 01\right) \right\}$$

¹ This is based on "Option Pricing Using Integral Transforms" by Carr, P., Geman, H., Madan, D., and Yor, M.

$$= \frac{1}{2} \left\{ \exp(0) + \exp(0) \right\} = 1.$$

When $d = 2$ and $j = 1$:

$$\begin{aligned} \frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) &= \frac{1}{2} \sum_{k=0}^{2-1} \exp\left(\frac{2\pi i}{2} 1k\right) = \frac{1}{2} \left\{ \exp\left(\frac{2\pi i}{2} 10\right) + \exp\left(\frac{2\pi i}{2} 11\right) \right\} \\ &= \frac{1}{2} \left\{ \exp(0) + \exp(i\pi) \right\} = \frac{1}{2} \left\{ 1 - 1 \right\} = 0. \end{aligned}$$

When $d = 3$ and $j = 0$:

$$\begin{aligned} \frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) &= \frac{1}{3} \sum_{k=0}^{3-1} \exp\left(\frac{2\pi i}{3} 0k\right) \\ &= \frac{1}{3} \left\{ \exp\left(\frac{2\pi i}{3} 00\right) + \exp\left(\frac{2\pi i}{3} 01\right) + \exp\left(\frac{2\pi i}{3} 02\right) \right\} \\ &= \frac{1}{3} \left\{ \exp(0) + \exp(0) + \exp(0) \right\} = 1. \end{aligned}$$

When $d = 3$ and $j = 1$:

$$\begin{aligned} \frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) &= \frac{1}{3} \sum_{k=0}^{3-1} \exp\left(\frac{2\pi i}{3} 1k\right) \\ &= \frac{1}{3} \left\{ \exp\left(\frac{2\pi i}{3} 10\right) + \exp\left(\frac{2\pi i}{3} 11\right) + \exp\left(\frac{2\pi i}{3} 12\right) \right\} \\ &= \frac{1}{3} \left\{ \exp(0) + \exp\left(\frac{2\pi i}{3}\right) + \exp\left(\frac{4\pi i}{3}\right) \right\} \\ &= \frac{1}{3} \left\{ 1 + (0.866025i - 0.5) + (-0.866025i - 0.5) \right\} \\ &= 0. \end{aligned}$$

When $d = 3$ and $j = 2$:

$$\begin{aligned} \frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) &= \frac{1}{3} \sum_{k=0}^{3-1} \exp\left(\frac{2\pi i}{3} 2k\right) \\ &= \frac{1}{3} \left\{ \exp\left(\frac{2\pi i}{3} 20\right) + \exp\left(\frac{2\pi i}{3} 21\right) + \exp\left(\frac{2\pi i}{3} 22\right) \right\} \\ &= \frac{1}{3} \left\{ \exp(0) + \exp\left(\frac{4\pi i}{3}\right) + \exp\left(\frac{8\pi i}{3}\right) \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{3} \left\{ 1 + (-0.866025i - 0.5) + (0.866025i - 0.5) \right\} \\
&= 0.
\end{aligned}$$

□

(Formal) PROOF of a proposition (3.13)

Rewrite as the below:

$$\frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) = \frac{1}{d} \sum_{k=0}^{d-1} \beta^k, \quad (3.14)$$

where $\beta = \exp\left(\frac{2\pi i}{d} j\right)$. When $j = 0$, $\beta = \exp\left(\frac{2\pi i}{d} 0\right) = 1$. Thus, from (3.14):

$$\frac{1}{d} \sum_{k=0}^{d-1} \beta^k = \frac{1}{d} \sum_{k=0}^{d-1} 1^k = \frac{1}{d} \sum_{k=0}^{d-1} 1 = 1.$$

When $j \neq 0$, consider the term $\sum_{k=0}^{d-1} \beta^k$ which is a geometric series:

$$S_{d-1} = \sum_{k=0}^{d-1} \beta^k = 1 + \beta + \beta^2 + \beta^3 + \dots + \beta^{d-2} + \beta^{d-1}. \quad (3.15)$$

Multiply β to (3.12):

$$\beta S_{d-1} = \beta + \beta^2 + \beta^3 + \dots + \beta^{d-2} + \beta^{d-1} + \beta^d. \quad (3.16)$$

Subtract (3.16) from (3.15):

$$\begin{aligned}
(1 - \beta) S_{d-1} &= 1 - \beta^d \\
S_{d-1} &= \frac{1 - \beta^d}{1 - \beta}.
\end{aligned} \quad (3.17)$$

Note that for $j \neq 0$:

$$\beta^d = \exp\left(\frac{2\pi i}{d} j\right)^d = \exp(2\pi i j) = 1.$$

From (3.17):

$$S_{d-1} = \sum_{k=0}^{d-1} \beta^k = \frac{1-1}{(1-\beta)} = 0.$$

From (3.14):

$$\frac{1}{d} \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right) = \frac{1}{d} 0 = 0.$$

□

Now we have completed the proof of a proposition (3.13) and we use this now. Multiply d to both sides of a proposition (3.13):

$$d1_{j=0} = \sum_{k=0}^{d-1} \exp\left(\frac{2\pi i}{d} jk\right). \quad (3.18)$$

As the limit $d \rightarrow \infty$ and plug $j = x - a$, (3.18) becomes:

$$\delta(x-a) = \int_{-\infty}^{\infty} e^{i2\pi f(x-a)} df. \quad (3.19)$$

Convert frequency f into angular frequency ω by the equation (3.1) which is

$$\omega = \frac{2\pi}{T} = 2\pi f. \text{ From (3.19):}$$

$$\delta(x-a) = \int_{-\infty}^{\infty} e^{i\omega(x-a)} d \frac{\omega}{2\pi}.$$

This completes the proof of an identity (3.12).

□

[3.3.3] Linearity of Fourier Transform

Consider time domain functions $f(t)$ and $g(t)$ which have Fourier transforms $\mathcal{F}(\omega)$ and $\mathcal{G}(\omega)$ defined by the equation (3.4). Then:

$$\begin{aligned} \int_{-\infty}^{\infty} \{af(t) + bg(t)\} e^{i\omega t} dt &= a \int_{-\infty}^{\infty} f(t) e^{i\omega t} dt + b \int_{-\infty}^{\infty} g(t) e^{i\omega t} dt \\ &= a\mathcal{F}(\omega) + b\mathcal{G}(\omega). \end{aligned} \quad (3.20)$$

Or, we can write:

$$\mathcal{F}[af(t) + bg(t)](\omega) = a\mathcal{F}[f(t)](\omega) + b\mathcal{F}[g(t)](\omega).$$

Linearity means two things: homogeneity and additivity. Homogeneity of Fourier transform indicates that if the amplitude is changed in one domain by a , the amplitude in the other domain changes by an exactly the same amount a (i.e. scaling property):

$$\mathcal{F}[af(t)](\omega) = a\mathcal{F}[f(t)](\omega).$$

Additivity of Fourier transform indicates that an addition in one domain corresponds to an addition in other domain:

$$\mathcal{F}[f(t) + g(t)](\omega) = \mathcal{F}[f(t)](\omega) + \mathcal{F}[g(t)](\omega).$$

[3.3.4] FT of Even and Odd Functions

A function $g(x)$ is said to be even if for $x \in \mathbb{R}$:

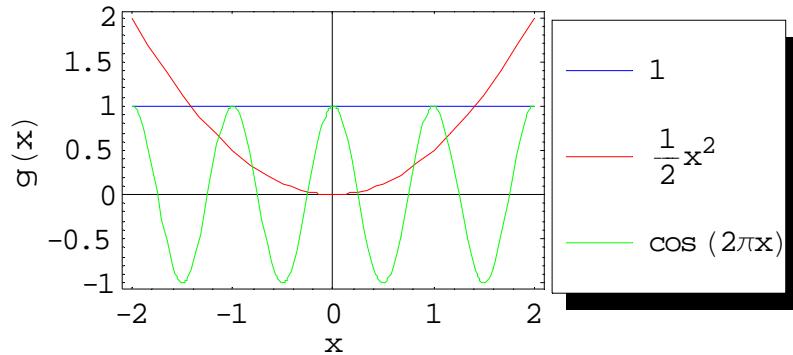
$$g(x) = g(-x), \quad (3.21)$$

which implies that even functions are symmetric with respect to vertical axis. Examples of even functions are illustrated in Panel A of Figure 3.8.

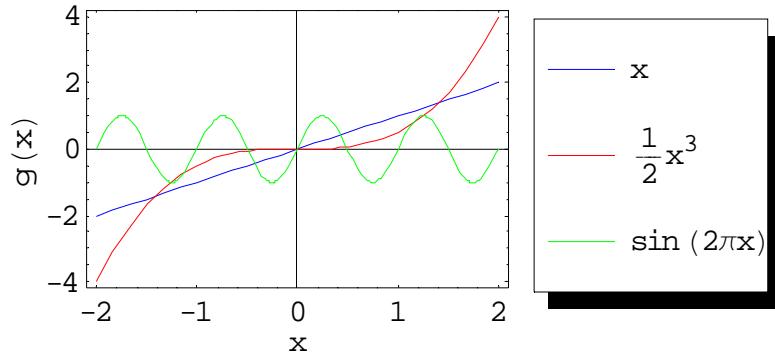
A function $g(x)$ is said to be odd if for $x \in \mathbb{R}$:

$$-g(x) = g(-x), \quad (3.22)$$

which implies that odd functions are symmetric with respect to the origin. Examples of odd functions are illustrated in Panel B of Figure 3.8.



A) Even Functions: $g(x) = 1$, $g(x) = \frac{1}{2}x^2$, and $g(x) = \cos(2\pi x)$.



B) Odd Functions: $g(x) = x$, $g(x) = \frac{1}{2}x^3$, and $g(x) = \sin(2\pi x)$.

Figure 3.8: Plot of Even and Odd Functions

There are several important properties of even and odd functions. The sum of even functions is even and the sum of odd functions is odd. The product of two even functions is even and the product of two odd functions is also even. The product of an even and an odd function is odd.

Let $\text{even}(x)$ be an even function and $\text{odd}(x)$ be an odd function. Integral properties of even and odd functions are:

$$\int_{-A}^A \text{odd}(x)dx = 0, \quad (3.23)$$

$$\int_{-A}^A \text{even}(x)dx = 2 \int_0^A \text{even}(x)dx. \quad (3.24)$$

Consider FT of an even function $g(t)$. From the definition (3.4) and use Euler's formula:

$$\mathcal{G}(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t)dt = \int_{-\infty}^{\infty} \cos(\omega t)g(t)dt + i \int_{-\infty}^{\infty} \sin(\omega t)g(t)dt. \quad (3.25)$$

Since the imaginary part $\int_{-\infty}^{\infty} \sin(\omega t)g(t)dt$ is zero (because $\sin(\omega t)g(t)$ is odd and use the integral property (3.23)), FT $\mathcal{G}(\omega)$ is real and symmetric with respect to $\omega = 0$. In other words, FT of an even function is also even.

Next, consider FT of an odd function $g(t)$. Since the term $\int_{-\infty}^{\infty} \cos(\omega t)g(t)dt$ in (3.25) becomes zero ($\cos(\omega t)g(t)$ is odd and the integral property (3.23)), FT is given as:

$$\mathcal{G}(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t)dt = i \int_{-\infty}^{\infty} \sin(\omega t)g(t)dt.$$

This means that FT of an odd function $\mathcal{G}(\omega)$ is complex and asymmetric with respect to $\omega = 0$. This property is also illustrated in the section 3.2.5 and 3.2.6.

[3.3.5] Symmetry of Fourier Transform

By the definition of an inverse Fourier transform (3.5):

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{F}(\omega) d\omega.$$

Change the variable in the integration to y :

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iyt} \mathcal{F}(y) dy.$$

Consider $2\pi f(-t)$:

$$2\pi f(-t) = \int_{-\infty}^{\infty} e^{iyt} \mathcal{F}(y) dy. \quad (3.26)$$

We can say that the right hand side of (3.26) is by definition the Fourier transform of a function $\mathcal{F}(y)$. Replace t by ω and y by t and (3.26) becomes:

$$2\pi f(-\omega) = \int_{-\infty}^{\infty} e^{it\omega} \mathcal{F}(t) dt. \quad (3.27)$$

The equation (3.27) is called a symmetry property of FT. It means that if a function $f(t)$ has a FT $\mathcal{F}(\omega)$, $\mathcal{F}(t)$ has a FT $2\pi f(-\omega)$. In other words, if $(f(t), \mathcal{F}(\omega))$ is a FT pair, $(\mathcal{F}(t), 2\pi f(-\omega))$ is another FT pair.

This symmetry property of FT can be shown with mathematically simpler form in the frequency domain f Hz (cycles/second). By the definition of an inverse FT (3.7):

$$g(t) \equiv \int_{-\infty}^{\infty} e^{2\pi i ft} \mathcal{G}(f) df.$$

Change the variable in the integration to y :

$$g(t) \equiv \int_{-\infty}^{\infty} e^{2\pi iy t} \mathcal{G}(y) dy.$$

Consider $g(-t)$:

$$g(-t) \equiv \int_{-\infty}^{\infty} e^{-2\pi iyt} \mathcal{G}(y) dy. \quad (3.28)$$

We can say that the right hand side of (3.28) is by definition the FT of a function $\mathcal{G}(y)$. Replace t by f and y by t and (3.28) becomes:

$$g(-f) \equiv \int_{-\infty}^{\infty} e^{-2\pi ift} \mathcal{G}(t) dt. \quad (3.29)$$

We can state in this case that if $(g(t), \mathcal{G}(f))$ is a FT pair, $(\mathcal{G}(t), g(-f))$ is another FT pair.

[3.3.6] Differentiation of Fourier Transform

By the definition of an inverse Fourier transform (3.5):

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{F}(\omega) d\omega.$$

Differentiate with respect to t :

$$\begin{aligned} \frac{\partial f(t)}{\partial t} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \mathcal{F}(\omega) \frac{\partial e^{-i\omega t}}{\partial t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega (-i\omega \mathcal{F}(\omega)) e^{-i\omega t} \\ &= -i\omega \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\omega) e^{-i\omega t} d\omega. \end{aligned} \quad (3.30)$$

By the definition of an inverse FT (3.5), the equation (3.30) becomes:

$$\frac{\partial f(t)}{\partial t} = -i\omega f(t). \quad (3.31)$$

Equation (3.31) tells us that FT of $\partial f(t) / \partial t$ is equal to a FT of $f(t)$ multiplied by $-i\omega$:

$$\mathcal{F}[\partial f(t) / \partial t](\omega) = -i\omega \mathcal{F}[f(t)](\omega). \quad (3.32)$$

Next, consider FT in frequency domain f . By the definition of an inverse FT (3.7):

$$g(t) \equiv \int_{-\infty}^{\infty} e^{2\pi ift} \mathcal{G}(f) df.$$

Differentiate with respect to t :

$$\frac{\partial g(t)}{\partial t} = 2\pi if \int_{-\infty}^{\infty} e^{2\pi ift} \mathcal{G}(f) df = 2\pi if g(t). \quad (3.33)$$

Equation (3.33) tells us that FT of $\partial g(t)/\partial t$ is equal to a Fourier transform of $g(t)$ multiplied by $2\pi if$:

$$\mathcal{F}[\partial g(t)/\partial t](f) = 2\pi if \mathcal{F}[g(t)](f).$$

[3.3.7] Time Scaling of Fourier Transform

Consider a time domain function $f(t)$ and its Fourier transform $\mathcal{F}(\omega)$ by the equation (3.4):

$$\mathcal{F}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} f(t) dt.$$

Then, FT of a function $f(at)$ (i.e. scaled by a real non-zero constant a) can be expressed in terms of $\mathcal{F}(\omega)$ as:

$$\frac{1}{|a|} \mathcal{F}\left(\frac{\omega}{a}\right) = \int_{-\infty}^{\infty} e^{i\omega t} f(at) dt. \quad (3.34)$$

PROOF

Set $at = s$. When $a > 0$:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{i\omega t} f(at) dt &= \int_{-\infty}^{\infty} e^{i\omega s/a} f(s) d\left(\frac{s}{a}\right) = \frac{1}{a} \int_{-\infty}^{\infty} e^{i(\omega/a)s} f(s) ds \\ &= \frac{1}{a} \mathcal{F}\left(\frac{\omega}{a}\right). \end{aligned}$$

When $a < 0$:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{i\omega t} f(at) dt &= \int_{-\infty}^{\infty} e^{i\omega s/a} f(s) d\left(\frac{s}{a}\right) = -\frac{1}{a} \int_{-\infty}^{\infty} e^{i(\omega/a)s} f(s) ds \\ &= -\frac{1}{a} \mathcal{F}\left(\frac{\omega}{a}\right). \end{aligned}$$

□

Similarly, FT of a function $g(at)$ can be expressed in terms of $\mathcal{G}(f)$ in frequency domain f as:

$$\frac{1}{|a|} \mathcal{G}\left(\frac{f}{a}\right) = \int_{-\infty}^{\infty} e^{-2\pi ift} g(at) dt.$$

[3.3.8] Time Shifting of Fourier Transform

Consider a function $f(t)$ and its Fourier transform $\mathcal{F}(\omega)$ by the definition (3.4):

$$\mathcal{F}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} f(t) dt .$$

Then, FT of a function $f(t-t_0)$ (i.e. time t is shifted by $t_0 \in \mathbb{R}$) can be expressed in terms of $\mathcal{F}(\omega)$ as:

$$\int_{-\infty}^{\infty} e^{i\omega t} f(t-t_0) dt = e^{i\omega t_0} \mathcal{F}(\omega) . \quad (3.35)$$

PROOF

Set $t - t_0 = t^*$:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{i\omega t} f(t-t_0) dt &= \int_{-\infty}^{\infty} e^{i\omega(t^*+t_0)} f(t^*) d(t^*+t_0) \\ &= e^{i\omega t_0} \int_{-\infty}^{\infty} e^{i\omega t^*} f(t^*) dt^* \\ &= e^{i\omega t_0} \mathcal{F}(\omega) . \end{aligned}$$

□

Next, consider FT in frequency domain f . FT of a time domain function $g(t)$ is defined by the definition (3.6) as:

$$\mathcal{G}(f) \equiv \int_{-\infty}^{\infty} e^{-2\pi ift} g(t) dt .$$

Then, FT of a function $g(t-t_0)$ (i.e. time t is shifted by $t_0 \in \mathbb{R}$) can be expressed in terms of $\mathcal{G}(f)$ as:

$$\int_{-\infty}^{\infty} e^{-2\pi ift} g(t-t_0) dt = e^{-2\pi ift_0} \mathcal{G}(f) . \quad (3.36)$$

PROOF

Set $t - t_0 = t^*$:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-2\pi ift} g(t-t_0) dt &= \int_{-\infty}^{\infty} e^{-2\pi if(t^*+t_0)} g(t^*) d(t^*+t_0) \\ &= e^{-2\pi ift_0} \int_{-\infty}^{\infty} e^{-2\pi ift^*} g(t^*) dt^* = e^{-2\pi ift_0} \mathcal{G}(f) . \end{aligned}$$

□

[3.3.9] Convolution: Time Convolution Theorem

Convolution of time domain functions $f(t)$ and $g(t)$ over a finite interval $[0, t]$ is defined as:

$$f * g \equiv \int_0^t f(\tau)g(t-\tau)d\tau . \quad (3.37)$$

Convolution of time domain functions $f(t)$ and $g(t)$ over an infinite interval $[-\infty, \infty]$ is defined as:

$$f * g \equiv \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{\infty} g(\tau)f(t-\tau)d\tau . \quad (3.38)$$

Convolution can be considered as an integral which measures the amount of overlapping of one function $g(t)$ when $g(t)$ is shifted over another function $f(t)$. Website by mathworld provides an excellent description of convolution with animation. For example, suppose $f(t)$ and $g(t)$ are Gaussian functions:

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{(t-\mu_1)^2}{2\sigma_1^2}\right\}, \\ g(t) &= \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left\{-\frac{(t-\mu_2)^2}{2\sigma_2^2}\right\}. \end{aligned}$$

Then, the convolution of two Gaussian functions is calculated as from (3.38):

$$f * g = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left\{-\frac{(t-(\mu_1 + \mu_2))^2}{2(\sigma_1^2 + \sigma_2^2)}\right\},$$

which is another Gaussian function. The convolution $f * g$ of two Gaussians for the case $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, and $\sigma_2 = 2$ is plotted in Figure 3.8.

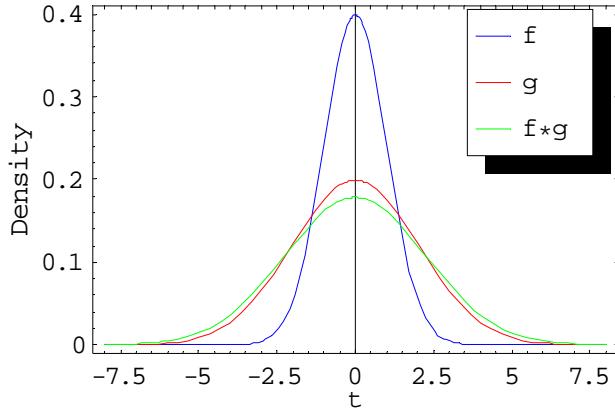


Figure 3.8: Plot of Two Gaussian functions f and g and their convolution $f * g$.

Consider time domain functions $f(t)$ and $g(t)$ with Fourier transforms $\mathcal{F}(\omega)$ and $\mathcal{G}(\omega)$. FT of the convolution of $f(t)$ and $g(t)$ in the time domain is equal to the multiplication in the angular frequency domain (called time-convolution theorem):

$$\mathcal{F}(f * g) \equiv \mathcal{F}\left(\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau\right) = \mathcal{F}(\omega)\mathcal{G}(\omega). \quad (3.39)$$

PROOF

$$\begin{aligned} \mathcal{F}(f * g) &\equiv \mathcal{F}\left(\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau\right) = \int_{-\infty}^{\infty} e^{i\omega t} \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau dt \\ &= \int_{-\infty}^{\infty} f(\tau)d\tau \left(\int_{-\infty}^{\infty} g(t-\tau)e^{i\omega t} dt \right). \end{aligned}$$

Use the time shifting property of Fourier transform of the equation (3.35):

$$\begin{aligned} \mathcal{F}(f * g) &= \int_{-\infty}^{\infty} f(\tau)d\tau \left(e^{i\omega\tau} \mathcal{G}(\omega) \right) = \mathcal{G}(\omega) \int_{-\infty}^{\infty} f(\tau)e^{i\omega\tau} d\tau \\ &= \mathcal{F}(\omega)\mathcal{G}(\omega). \end{aligned}$$

□

[3.3.10] Frequency-Convolution Theorem

Consider time domain functions $f(t)$ and $g(t)$ with Fourier transforms $\mathcal{F}(\omega)$ and $\mathcal{G}(\omega)$. Frequency-convolution theorem states that convolution in the angular frequency domain (scaled by $\frac{1}{2\pi}$) is equal to the multiplication in the time domain. In other words, FT of the product $f(t)g(t)$ in the time domain is equal to the convolution $\mathcal{F}(\omega)*\mathcal{G}(\omega)$ (scaled by $\frac{1}{2\pi}$) in the angular frequency domain:

$$\mathcal{F}[f(t)g(t)](\omega) = \frac{1}{2\pi} \mathcal{F}(\omega) * \mathcal{G}(\omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\varpi) \mathcal{G}(\omega - \varpi) d\varpi. \quad (3.40)$$

PROOF

There are several different ways to prove the frequency convolution theorem. But we prove this by showing that the inverse FT of the convolution $\mathcal{F} * \mathcal{G}$ in the angular frequency domain is equal to the multiplication $f(t)g(t)$ (scaled by 2π) in the time domain.

Following the definition of inverse FT (3.5):

$$\begin{aligned} \mathcal{F}_{\omega}^{-1}[\mathcal{F}(\omega) * \mathcal{G}(\omega)](t) &\equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{F}(\omega) * \mathcal{G}(\omega) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \int_{-\infty}^{\infty} \mathcal{F}(\varpi) \mathcal{G}(\omega - \varpi) d\varpi d\omega \\ &= \int_{-\infty}^{\infty} d\varpi \mathcal{F}(\varpi) \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{G}(\omega - \varpi) d\omega \right). \end{aligned}$$

Using the frequency shifting (modulation) property of FT (discussed in section 3.3.11):

$$\begin{aligned} \mathcal{F}_{\omega}^{-1}[\mathcal{F}(\omega) * \mathcal{G}(\omega)](t) &= \int_{-\infty}^{\infty} d\varpi \mathcal{F}(\varpi) e^{-i\varpi t} g(t) \\ &= g(t) \int_{-\infty}^{\infty} d\varpi \mathcal{F}(\varpi) e^{-i\varpi t} = g(t) 2\pi f(t). \end{aligned}$$

□

Next, consider FT in the frequency domain f . Following the definition of inverse FT (3.7):

$$\begin{aligned} \mathcal{F}_f^{-1}[\mathcal{F}(f) * \mathcal{G}(f)](t) &\equiv \int_{-\infty}^{\infty} e^{2\pi ift} \mathcal{F}(f) * \mathcal{G}(f) df \\ &= \int_{-\infty}^{\infty} e^{2\pi ift} \int_{-\infty}^{\infty} \mathcal{F}(f^*) \mathcal{G}(f - f^*) df * df \\ &= \int_{-\infty}^{\infty} df * \mathcal{F}(f^*) \left(\int_{-\infty}^{\infty} e^{2\pi ift} \mathcal{G}(f - f^*) df \right). \end{aligned}$$

Using the frequency shifting (modulation) property of FT (discussed in section 3.3.11):

$$\begin{aligned} \mathcal{F}_f^{-1}[\mathcal{F}(f) * \mathcal{G}(f)](t) &= \int_{-\infty}^{\infty} df * \mathcal{F}(f^*) e^{2\pi if^* t} g(t) \\ &= g(t) \int_{-\infty}^{\infty} df * \mathcal{F}(f^*) e^{2\pi if^* t} = f(t)g(t). \end{aligned}$$

□

[3.3.11] Frequency Shifting: Modulation

Consider time domain function $f(t)$ with Fourier transform $\mathcal{F}(\omega)$. If FT $\mathcal{F}(\omega)$ is shifted by $\omega_0 \in \mathbb{R}$ in the angular frequency domain, then the inverse FT $f(t)$ is multiplied by $e^{i\omega_0 t}$:

$$e^{-i\omega_0 t} f(t) = \mathcal{F}^{-1}(\omega - \omega_0), \quad (3.41)$$

$$\mathcal{F}(e^{-i\omega_0 t} f(t)) = \mathcal{F}(\omega - \omega_0). \quad (3.42)$$

PROOF

Let $s = \omega - \omega_0$. From the definition of an inverse Fourier transform (3.5):

$$\begin{aligned} \mathcal{F}^{-1}(\omega - \omega_0) &\equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-is\omega} \mathcal{F}(\omega - \omega_0) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(s+\omega_0)t} \mathcal{F}(s) d(s + \omega_0) \\ &= e^{-i\omega_0 t} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ist} \mathcal{F}(s) ds = e^{-i\omega_0 t} f(t). \end{aligned}$$

□

Similarly, if FT $\mathcal{F}(f)$ is shifted by $f_0 \in \mathbb{R}$ in the frequency domain, then the inverse FT $g(t)$ is multiplied by $e^{2\pi i f_0 t}$:

$$e^{2\pi i f_0 t} g(t) = \mathcal{F}^{-1}[\mathcal{G}(f - f_0)], \quad (3.43)$$

$$\mathcal{F}(e^{2\pi i f_0 t} g(t)) = \mathcal{G}(f - f_0). \quad (3.44)$$

[3.3.12] Parseval's Relation

Let $f(t)$ and $g(t)$ be L^2 -complex functions. A L^2 -function can be informally considered as a square integrable function (i.e. A function $f(t)$ is said to be square-integrable if $\int_{-\infty}^{\infty} |f(t)|^2 dt < \infty$.) Let $\mathcal{F}(\omega)$ and $\mathcal{G}(\omega)$ be the Fourier transforms of $f(t)$ and $g(t)$ defined by (3.4):

$$\begin{aligned} \mathcal{F}(\omega) &= \int_{-\infty}^{\infty} e^{i\omega t} f(t) dt, \\ \mathcal{G}(\omega) &= \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt. \end{aligned}$$

Let $\bar{g}(t)$ be a complex conjugate of $g(t)$ and $\bar{\mathcal{G}}(\omega)$ be a complex conjugate of $\mathcal{G}(\omega)$ ²:

² A complex conjugate of a complex number $z \equiv a + bi$ is $\bar{z} \equiv a - bi$.

$$\begin{aligned} |f(t)|^2 &= f(t)\bar{g}(t), \\ |\mathcal{F}(\omega)|^2 &= \mathcal{F}(\omega)\bar{\mathcal{G}}(\omega). \end{aligned}$$

Then, Parseval's relation is:

$$\begin{aligned} \int_{-\infty}^{\infty} |f(t)|^2 dt &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |\mathcal{F}(\omega)|^2 d\omega, \\ \int_{-\infty}^{\infty} f(t)\bar{g}(t)dt &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\omega)\bar{\mathcal{G}}(\omega)d\omega. \end{aligned} \quad (3.45)$$

Parseval's relation in the case of FT parameters $(a,b) = (1,1)$ indicates that there is a very simple relationship between the power of a signal function $f(t)$ computed in signal space t or transform space ω of the form (3.45).

Parseval's relation becomes simpler when considered in the frequency domain f instead of angular frequency domain ω . Let $\mathcal{F}(f)$ and $\mathcal{G}(f)$ be the Fourier transforms of $f(t)$ and $g(t)$ defined by (3.6):

$$\begin{aligned} \mathcal{F}(f) &\equiv \int_{-\infty}^{\infty} e^{-2\pi ift} f(t)dt, \\ \mathcal{G}(f) &\equiv \int_{-\infty}^{\infty} e^{-2\pi ift} g(t)dt. \end{aligned}$$

Let $\bar{g}(t)$ be a complex conjugate of $g(t)$ and $\bar{\mathcal{G}}(f)$ be a complex conjugate of $\mathcal{G}(f)$:

$$\begin{aligned} |f(t)|^2 &= f(t)\bar{g}(t), \\ |\mathcal{F}(f)|^2 &= \mathcal{F}(f)\bar{\mathcal{G}}(f). \end{aligned}$$

Then, Parseval's relation is:

$$\begin{aligned} \int_{-\infty}^{\infty} |f(t)|^2 dt &= \int_{-\infty}^{\infty} |\mathcal{F}(f)|^2 df, \\ \int_{-\infty}^{\infty} f(t)\bar{g}(t)dt &= \int_{-\infty}^{\infty} \mathcal{F}(f)\bar{\mathcal{G}}(f)df. \end{aligned} \quad (3.46)$$

This version of Parseval's relation means that the power of a signal function $f(t)$ is same whether it is computed in signal space t or in transform space f .

PROOF

Using inverse Fourier transforms of (3.5):

$$\begin{aligned}\int_{-\infty}^{\infty} |f(t)|^2 dt &= \int_{-\infty}^{\infty} f(t) \bar{g}(t) dt = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{F}(\omega) d\omega \right] \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega' t} \bar{\mathcal{G}}(\omega') d\omega' \right] dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\omega) \int_{-\infty}^{\infty} \bar{\mathcal{G}}(\omega') \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{it(\omega'-\omega)} dt \right] d\omega' d\omega.\end{aligned}$$

Use the identity of Dirac's delta function (3.12):

$$\delta(t-a) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(t-a)} d\omega.$$

Thus:

$$\begin{aligned}\int_{-\infty}^{\infty} |f(t)|^2 dt &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\omega) \int_{-\infty}^{\infty} \bar{\mathcal{G}}(\omega') \delta(\omega' - \omega) d\omega' d\omega \\ \int_{-\infty}^{\infty} |f(t)|^2 dt &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\omega) \bar{\mathcal{G}}(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\mathcal{F}(\omega)|^2 d\omega.\end{aligned}$$

□

[3.3.13] Summary of Important Properties of Fourier Transform

Table 3.1: Summary of Important Properties of Fourier Transform in Angular Frequency Domain ω Hz (radians/second)

Property	Time Domain Function $y(t)$	Fourier Transform $\mathcal{F}[y(t)](\omega)$
Linearity	$af(t) + bg(t)$	$a\mathcal{F}(\omega) + b\mathcal{G}(\omega)$
Even Function	$f(t)$ is even	$\mathcal{F}(\omega) \in \mathbb{R}$
Odd Function	$f(t)$ is odd	$\mathcal{F}(\omega) \in \mathbb{I}$
Symmetry	$\mathcal{F}(t)$	$2\pi f(-\omega)$
Differentiation	$\frac{df(t)}{dt}$	$-i\omega\mathcal{F}(\omega)$
	$\frac{d^k f(t)}{dt^k}$	$(-i\omega)^k \mathcal{F}(\omega)$
Time Scaling	$f(at)$	$\frac{1}{ a } \mathcal{F}\left(\frac{\omega}{a}\right)$

Time Shifting	$f(t - t_0)$	$e^{i\omega_0 t} \mathcal{F}(\omega)$
Convolution	$f * g \equiv \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$	$\mathcal{F}(\omega)\mathcal{G}(\omega)$
Multiplication	$f(t)g(t)$	$\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\varpi)\mathcal{G}(\omega-\varpi)d\varpi$
Modulation (Frequency Shifting)	$e^{-i\omega_0 t} f(t)$	$\mathcal{F}(\omega - \omega_0)$

Table 3.2: Summary of Important Properties of Fourier Transform in Frequency Domain f Hz (cycles/second)

Property	Time Domain Function $y(t)$	Fourier Transform $\mathcal{F}[y(t)](f)$
Linearity	$af(t) + bg(t)$	$a\mathcal{F}(f) + b\mathcal{G}(f)$
Even Function	$g(t)$ is even	$\mathcal{G}(f) \in \mathbb{R}$
Odd Function	$g(t)$ is odd	$\mathcal{G}(f) \in \mathbb{I}$
Symmetry	$\mathcal{G}(t)$	$\mathcal{G}(-f)$
Differentiation	$\frac{dg(t)}{dt}$	$2\pi if\mathcal{G}(f)$
	$\frac{d^k g(t)}{dt^k}$	$(2\pi if)^k \mathcal{G}(f)$
Time Scaling	$g(at)$	$\frac{1}{ a } \mathcal{G}\left(\frac{f}{a}\right)$
Time Shifting	$g(t - t_0)$	$e^{-2\pi ift_0} \mathcal{G}(f)$
Convolution	$f * g \equiv \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$	$\mathcal{F}(f)\mathcal{G}(f)$
Multiplication	$f(t)g(t)$	$\mathcal{F} * \mathcal{G}$

Modulation (Frequency Shifting)	$e^{2\pi i f_0 t} g(t)$	$\mathcal{G}(f - f_0)$
------------------------------------	-------------------------	------------------------

[3.4] Existence of the Fourier Integral

Thus far, we've assumed that Fourier integral and its inverse of the definitions (3.4) and (3.5) of a FT pair $(g(t), \mathcal{G}(\omega))$ do exist (i.e. they are well-defined for all functions).

Sufficient (but not necessary) condition for the existence of Fourier transform and its inverse is:

$$\int_{-\infty}^{\infty} |g(t)| dt < \infty, \quad (3.47)$$

which is an integrability condition.

[4] Characteristic Function

[4.1] Definition of a Characteristic Function

Let X be a random variable with its probability density function $\mathbb{P}(x)$. A characteristic function $\phi(\omega)$ with $\omega \in \mathbb{R}$ is defined as the Fourier transform of the probability density function $\mathbb{P}(x)$ using Fourier transform parameters $(a, b) = (1, 1)$. From the definition (3.4):

$$\phi(\omega) \equiv \mathcal{F}[\mathbb{P}(x)] \equiv \int_{-\infty}^{\infty} e^{i\omega x} \mathbb{P}(x) dx = E[e^{i\omega x}]. \quad (4.1)$$

Using the Euler's formula (2.15), $\phi(\omega)$ can be expressed as:

$$\phi(\omega) = E[e^{i\omega x}] = E[\cos(\omega x)] + iE[\sin(\omega x)].$$

Taylor series expansion of a real function $f(x)$ in one dimension about a point $x = b$ is given by:

$$f(x) = f(b) + f'(b)(x - b) + \frac{f''(b)}{2!}(x - b)^2 + \frac{f'''(b)}{3!}(x - b)^3 + \dots \quad (4.2)$$

Taylor series expansion of an exponential function $\exp(i\omega x)$ in one dimension about a point $x = 0$ is given by from (4.2) as:

$$\begin{aligned} \exp(i\omega x) &= \exp(i\omega 0) + \frac{\partial \exp(i\omega x)}{\partial x} \Big|_{x=0} (x - 0) + \frac{1}{2!} \frac{\partial^2 \exp(i\omega x)}{\partial x^2} \Big|_{x=0} (x - 0)^2 \\ &\quad + \frac{1}{3!} \frac{\partial^3 \exp(i\omega x)}{\partial x^3} \Big|_{x=0} (x - 0)^3 + \dots \\ \exp(i\omega x) &= \exp(i\omega 0) + i\omega \exp(i\omega x) \Big|_{x=0} (x - 0) + \frac{1}{2!} (i\omega)^2 \exp(i\omega x) \Big|_{x=0} (x - 0)^2 \\ &\quad + \frac{1}{3!} (i\omega)^3 \exp(i\omega x) \Big|_{x=0} (x - 0)^3 + \dots \\ \exp(i\omega x) &= 1 + i\omega x + \frac{1}{2!} (i\omega x)^2 + \frac{1}{3!} (i\omega x)^3 + \frac{1}{4!} (i\omega x)^4 + \dots \end{aligned} \quad (4.3)$$

Therefore, a characteristic function $\phi(\omega)$ can be rewritten as from equations (4.1) and (4.3) as:

$$\begin{aligned} \phi(\omega) &\equiv \int_{-\infty}^{\infty} e^{i\omega x} \mathbb{P}(x) dx \\ &= \int_{-\infty}^{\infty} \left(1 + i\omega x + \frac{1}{2!} (i\omega x)^2 + \frac{1}{3!} (i\omega x)^3 + \frac{1}{4!} (i\omega x)^4 + \dots \right) \mathbb{P}(x) dx \end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \mathbb{P}(x) dx + i\omega \int_{-\infty}^{\infty} x \mathbb{P}(x) dx + \frac{1}{2!} (i\omega)^2 \int_{-\infty}^{\infty} x^2 \mathbb{P}(x) dx + \frac{1}{3!} (i\omega)^3 \int_{-\infty}^{\infty} x^3 \mathbb{P}(x) dx + \dots \\
&= r_0 + i\omega r_1 - \frac{1}{2!} \omega^2 r_2 - \frac{1}{3!} i\omega^3 r_3 + \frac{1}{4!} \omega^4 r_4 + \dots \\
&= \sum_{n=0}^{\infty} \frac{(i\omega)^n}{n!} r_n,
\end{aligned} \tag{4.4}$$

where r_n is the n -th moment about 0 (called raw moment).

Probability density function $\mathbb{P}(x)$ can be obtained by inverse Fourier transform of the characteristic function using the equation (3.5):

$$\mathbb{P}(x) = \mathcal{F}^{-1}[\phi(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega x} \phi(\omega) d\omega. \tag{4.5}$$

If X is a discrete random variable with possible values $\{x_k\}_{k=0}^{\infty}$ and $\Pr\{X = x_k\} = a_k$, then $\phi(\omega)$ is obtained by a series expansion instead of an integration as:

$$\phi(\omega) \equiv \sum_{k=0}^{\infty} \exp(i\omega x_k) a_k. \tag{4.6}$$

[4.2] Properties of a Characteristic Function

Let X be a random variable and $\phi_X(\omega)$ be its characteristic function. A characteristic function $\phi_X(\omega)$ is: 1) bounded by 1 (i.e. $|\phi_X(\omega)| \leq 1$, $\omega \in \mathbb{R}$), 2) $\phi_X(0) = 1$, and 3) uniformly continuous in \mathbb{R} .

A statistical distribution is uniquely determined by its characteristic function, i.e. one-to-one relationship between distribution functions and characteristic functions. In other words, if two random variables X and Y have the same characteristic functions (i.e. $\phi_X(\omega) = \phi_Y(\omega)$), they have the same distribution.

If $\{X_k, k = 1, \dots, n\}$ are independent random variables, the characteristic function of their sum $X_1 + X_2 + \dots + X_n$ is the product of their characteristic functions:

$$\phi_{X_1+X_2+\dots+X_n}(\omega) = \prod_{k=1}^n \phi_{X_k}(\omega). \tag{4.7}$$

A random variable X has a symmetric probability density function $\mathbb{P}(x)$ if and only if its characteristic function $\phi_X(\omega)$ is a real-valued function, i.e.:

$$\phi_X(\omega) \in \mathbb{R} \text{ for } \omega \in \mathbb{R}.$$

For $a, b \in \mathbb{R}$:

$$\phi_{aX+b}(\omega) = e^{i\omega b} \phi_X(a\omega). \quad (4.8)$$

[4.3] Characteristic Exponent: Cumulant-Generating Function

A characteristic exponent of a random variable X , $\Psi_X(\omega)$, is defined as a log of a characteristic function $\phi_X(\omega)$:

$$\Psi_X(\omega) \equiv \ln \phi_X(\omega). \quad (4.9)$$

The n -th cumulant is defined as:

$$\text{cumulant}_n = \frac{1}{i^n} \frac{\partial^n \Psi_X(\omega)}{\partial \omega^n} \Big|_{\omega=0}. \quad (4.10)$$

Mean, variance, skewness, and excess kurtosis of the random variable X can be obtained from cumulants as follows:

$$\begin{aligned} \text{Mean of } X &= E[X] = \text{cumulant}_1, \\ \text{Variance of } X &= E[(X - E(X))^2] = \text{cumulant}_2, \\ \text{Skewness of } X &= \frac{E[(X - E(X))^3]}{\left(\sqrt{E[(X - E(X))^2]}\right)^3} = \frac{\text{cumulant}_3}{(\text{cumulant}_2)^{3/2}}, \\ \text{Excess kurtosis of } X &= \frac{E[(X - E(X))^4]}{\left(\sqrt{E[(X - E(X))^2]}\right)^4} - 3 = \frac{\text{cumulant}_4}{(\text{cumulant}_2)^2}. \end{aligned} \quad (4.11)$$

Let's consider one fundamental example. A normal random variable X with mean μ and variance σ^2 has a density:

$$\mathbb{P}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\}.$$

Its characteristic function can be calculated as from the definition (4.1):

$$\phi(\omega) \equiv \mathcal{F}[\mathbb{P}(x)] \equiv \int_{-\infty}^{\infty} e^{i\omega x} \mathbb{P}(x) dx = \exp(i\mu\omega - \frac{\sigma^2\omega^2}{2}).$$

Its characteristic exponent is:

$$\Psi(\omega) \equiv \ln \phi(\omega) = \ln \left\{ \exp(i\mu\omega - \frac{\sigma^2\omega^2}{2}) \right\} = i\mu\omega - \frac{\sigma^2\omega^2}{2}.$$

Cumulants are calculated using $\Psi(\omega)$ from (4.10):

$$\begin{aligned} \text{cumulant}_1 &= \mu, \\ \text{cumulant}_2 &= \sigma^2, \\ \text{cumulant}_3 &= 0, \\ \text{cumulant}_4 &= 0. \end{aligned}$$

This tells us that a normal random variable X has a mean μ and variance σ^2 , zero skewness, and zero excess kurtosis.

[4.4] Laplace Transform

For nonnegative random variables we replace the Fourier transform by Laplace transform in order to obtain characteristic functions. The (unilateral) Laplace transform \mathcal{L} of a function $f(x)$ is defined as:

$$\mathcal{L}[f(x)] \equiv \int_0^{\infty} f(x) e^{-\omega x} dx, \quad (4.12)$$

where $f(x)$ is defined for $x \geq 0$. Thus, the characteristic function of a nonnegative random variable X with its density function $\mathbb{P}(x)$ is given by:

$$\phi_X(\omega) = \mathcal{L}[\mathbb{P}(x)] \equiv \int_0^{\infty} \mathbb{P}(x) e^{-\omega x} dx. \quad (4.13)$$

[4.5] Relationship with Moment Generating Function

Let X be a random variable on \mathbb{R} and $\mathbb{P}(x)$ be its probability density function. A function $M(\omega)$ with $\omega \in \mathbb{R}$ is called a moment generating function if there exists an $h > 0$ for $|\omega| < h$ such that (i.e. if the expectation in (4.14) converges):

$$M(\omega) \equiv \int_{-\infty}^{\infty} e^{\omega x} \mathbb{P}(x) dx \equiv E[\exp(\omega x)]. \quad (4.14)$$

For a continuous random variable X , again using the equation (4.3):

$$\begin{aligned} M(\omega) &\equiv \int_{-\infty}^{\infty} e^{\omega x} \mathbb{P}(x) dx \\ &= \int_{-\infty}^{\infty} \left(1 + \omega x + \frac{1}{2!} (\omega x)^2 + \frac{1}{3!} (\omega x)^3 + \frac{1}{4!} (\omega x)^4 + \dots \right) \mathbb{P}(x) dx \\ &= 1 + \omega c_1 + \frac{1}{2!} \omega^2 c_2 + \frac{1}{3!} \omega^3 c_3 + \frac{1}{4!} \omega^4 c_4 + \dots, \end{aligned}$$

where c_n is the n -th central moment.

If $\{X_k, k = 1, \dots, n\}$ are independent random variables, the moment generating function of their sum $X_1 + X_2 + \dots + X_n$ is the product of their moment generating functions:

$$M_{X_1+X_2+\dots+X_n}(\omega) = \prod_{k=1}^n M_{X_k}(\omega). \quad (4.15)$$

Its proof is very simple:

$$\begin{aligned} M_{X_1+X_2+\dots+X_n}(\omega) &= E\left[\exp\left\{\omega(X_1 + X_2 + \dots + X_n)\right\}\right] \\ &= E\left[\exp\left\{\omega X_1 + \omega X_2 + \dots + \omega X_n\right\}\right] \\ &= E\left[\exp\{\omega X_1\} \exp\{\omega X_2\} \dots \exp\{\omega X_n\}\right] \\ &= E\left[\exp\{\omega X_1\}\right] E\left[\exp\{\omega X_2\}\right] \dots E\left[\exp\{\omega X_n\}\right] \\ &= M_{X_1}(\omega) M_{X_2}(\omega) \dots M_{X_n}(\omega) = \prod_{k=1}^n M_{X_k}(\omega). \end{aligned}$$

If the moment generating function $M(\omega)$ is differentiable at zero (defined on a neighborhood $[-\varepsilon, \varepsilon]$ of zero), then the n -th raw moments r_n can be obtained by:

$$r_n = \frac{\partial^n M(\omega)}{\partial \omega^n} \Big|_{\omega=0}. \quad (4.16)$$

Thus:

$$\begin{aligned} r_1 &= M'(0) = E[X], \\ r_2 &= M''(0) = E[X^2], \\ r_3 &= M'''(0) = E[X^3], \\ r_4 &= M''''(0) = E[X^4]. \end{aligned}$$

For example, the mean and variance of the random variable X are computed using raw moments as (discussed in detail in section 4.6):

$$\begin{aligned}\text{Mean of } X &= E[X] = r_1, \\ \text{Variance of } X &= E[X^2] - E[X]^2 = r_2 - r_1^2.\end{aligned}$$

A characteristic function is always well-defined since it is the Fourier transform of a probability measure. But because the integral in (4.14) ($\forall \omega \in \mathbb{R}$) may not converge for some (all) values of ω , a moment generating function is not always well-defined. When $M(\omega)$ is well-defined, the relationship between the moment generating function $M(\omega)$ and the characteristic function $\phi(\omega)$ is given by:

$$M(\omega) = \phi(-i\omega). \quad (4.17)$$

Let's consider one fundamental example. If $X \sim \text{Normal}(\mu, \sigma)$, its moment generating function can be calculated as following the definition (4.14):

$$M(\omega) \equiv \int_{-\infty}^{\infty} e^{\omega x} \mathbb{P}(x) dx \equiv E[\exp(\omega x)] = \exp\left(\mu\omega + \frac{\sigma^2\omega^2}{2}\right).$$

You can confirm that (4.17) is true for the normal case. Raw moments are calculated as the following:

$$\begin{aligned}r_1 &= M'(0) = \mu, \\ r_2 &= M''(0) = \mu^2 + \sigma^2, \\ r_3 &= M'''(0) = \mu^3 + 3\mu\sigma^2, \\ r_4 &= M''''(0) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4.\end{aligned}$$

Using these raw moments, central moments can be calculated as:

$$\begin{aligned}E[X] &= r_1 = \mu, \\ \text{Variance}[X] &= E[X^2] - E[X]^2 = r_2 - r_1^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2, \\ \text{Skewness}[X] &= \frac{E\{X - E[X]\}^3}{(\sqrt{E\{X - E[X]\}^2})^3} = \frac{2r_1^3 - 3r_1r_2 + r_3}{(r_2 - r_1^2)^{3/2}} = 0, \\ \text{Excess Kurtosis}[X] &= \frac{E\{X - E[X]\}^4}{(\sqrt{E\{X - E[X]\}^2})^4} - 3 \\ &= \frac{-6r_1^4 + 12r_1^2r_2 - 3r_2^2 - 4r_1r_3 + r_4}{(r_2 - r_1^2)^2} = 0.\end{aligned}$$

[4.6] Summary: How to Calculate Standardized Moments from Characteristic Function and Moment Generating function

Following is a summary of the relationship between standardized moments, $cumulant_n$, and raw moments r_n . Let X be a random variable. n -th cumulant and n -th raw moment are defined by (5.10) and (5.16):

$$cum_n = \frac{1}{i^n} \frac{\partial^n \Psi_X(\omega)}{\partial \omega^n} \Big|_{\omega=0},$$

$$r_n = \frac{\partial^n M(\omega)}{\partial \omega^n} \Big|_{\omega=0}.$$

Table 4.1: How to Calculate Standardized Moments from Characteristic Function and Moment Generating function

Moments	n -th cumulant	n -th raw moment
Mean	$E[X]$	r_1
Variance	$E\{X - E[X]\}^2$	$r_2 - r_1^2$
Skewness	$\frac{E\{X - E[X]\}^3}{\left(\sqrt{E\{X - E[X]\}^2}\right)^3}$	$\frac{cum_3}{cum_2^{3/2}} \frac{2r_1^3 - 3r_1r_2 + r_3}{(r_2 - r_1^2)^{3/2}}$
Excess Kurtosis	$\frac{E\{X - E[X]\}^4}{\left(\sqrt{E\{X - E[X]\}^2}\right)^4} - 3$	$\frac{cum_4}{cum_2^2} \frac{-6r_1^4 + 12r_1^2r_2 - 3r_2^2 - 4r_1r_3 + r_4}{(r_2 - r_1^2)^2}$

Note:

- 2-nd central moment of $X = E\{X - E[X]\}^2 = E[X^2] - E[X]^2 = Var[X] = cum_2 = r_2 - r_1^2$.
- 3-rd central moment of $X = E\{X - E[X]\}^3 = cum_3 = 2r_1^3 - 3r_1r_2 + r_3$.
- Skewness of $X = Skewness[X] = \frac{E\{X - E[X]\}^3}{\left(\sqrt{E\{X - E[X]\}^2}\right)^3} = \frac{cum_3}{cum_2^{3/2}} = \frac{2r_1^3 - 3r_1r_2 + r_3}{(r_2 - r_1^2)^{3/2}}$
- 4-th central moment of $X = E\{X - E[X]\}^4 = cum_4 + 3\{E\{X - E[X]\}^2\}^2 = cum_4 + 3cum_2^2 = (-6r_1^4 + 12r_1^2r_2 - 3r_2^2 - 4r_1r_3 + r_4) + 3(r_2 - r_1^2)^2$.

- Excess Kurtosis of $X = \frac{E\{X - E[X]\}^4}{\left(\sqrt{E\{X - E[X]\}^2}\right)^4} - 3 = \frac{cum_4 + 3cum_2^2}{cum_2^2} - 3$
 $= \frac{cum_4}{cum_2^2} = \frac{-6r_1^4 + 12r_1^2r_2 - 3r_2^2 - 4r_1r_3 + r_4}{(r_2 - r_1^2)^2}.$

[4.7] Examples of Characteristic Functions

Several examples of characteristic functions using the definition (4.1) for continuous distributions and (4.6) for discrete distributions are given in the Table 4.2.

Table 4.2: Examples of Characteristic Functions

Distribution	$\mathbb{P}(x)$	$\phi(\omega)$
Normal	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(X-\mu)^2}{2\sigma^2}\right\}$	$\exp(i\mu\omega - \frac{\sigma^2\omega^2}{2})$
Exponential	ae^{-ax}	$\frac{a}{a-i\omega}$
Gamma	$\frac{b^{-a}e^{-x/b}x^{a-1}}{\Gamma(a)}$	$(1-b\omega)^{-a}$
Poisson	$\frac{e^{-\lambda}\lambda^x}{x!}$	$\exp[\lambda(e^{ix} - 1)]$

[5] Discrete Fourier Transform (DFT)

DFT is a special case of (continuous) Fourier transform. The results obtained from FT and DFT are identical and the only difference is how we interpret these results.

[5.1] Intuitive Derivation of DFT: Approximation of FT

We first consider DFT of a time domain function $g(t)$ into an angular frequency domain ω Hz (radians/second). This follows the convention in physics. In the field of signal processing which is a major application of FT, frequency f Hz (cycles/second) is used instead of ω . But this difference is not important because ω and f are measuring the same thing (rotation speed/second) in different units and related by:

$$\omega = 2\pi f . \quad (5.1)$$

We saw in section 3 that continuous FT of $g(t)$ and inverse FT of $\mathcal{G}(\omega)$ using FT parameters $(a, b) = (1, 1)$ are defined as:

$$\mathcal{G}(\omega) \equiv \mathcal{F}_t[g(t)](\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega t} g(t) dt , \quad (5.2)$$

$$g(t) \equiv \mathcal{F}_{\omega}^{-1}[\mathcal{G}(\omega)](t) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{G}(\omega) d\omega . \quad (5.3)$$

The purpose of DFT is to approximate FT as close as possible by sampling a finite number of points N of a continuous time domain function $g(t)$ with time domain sampling interval Δt (seconds) and sampling a finite number of points N of a continuous FT $\mathcal{G}(\omega)$ with angular frequency sampling interval $\Delta\omega$ Hz. In other words, both the original continuous time domain function $g(t)$ and the original continuous FT $\mathcal{G}(\omega)$ are approximated by a sample of N points.

We begin with the notation. Let N be the number of discrete samples taken to approximate $g(t)$ and $\mathcal{G}(\omega)$. Let Δt (seconds/sample) be the time domain sampling interval which is the time increment between samples. Its inverse $f_s = 1/\Delta t$ (samples/second) is called a sampling rate. Let T be the total sampling time:

$$\Delta t \equiv T / N . \quad (5.4)$$

It is extremely important to mention that total sampling time T defined by (5.4) has absolutely nothing to do with period T of oscillation of a wave (i.e. the seconds (time) taken for the wave to complete one wavelength defined by the equation (2.3)). To avoid confusion, period T of oscillation of a wave is called a fundamental period.

If Δt is assumed to be 1, the total sampling time and the number of samples taken are same (i.e. for example, $T = 10$ seconds and $N = 10$ samples). If $\Delta t = 0.01$, 1 sample is

taken in every 0.01 second in a time domain which in turn means that 100 samples are taken every second (i.e. sampling rate $f_s \equiv 1/\Delta t = 100$ Hz).

The first step to DFT is to take N discrete samples of a continuous time domain function $g(t)$ at n -th sampling instant $t_n = n\Delta t$ (seconds) with $n = 0, 1, \dots, N-1$. When $T = 10$ seconds and $N = 10$ samples (i.e. time domain sampling interval $\Delta t = 1$), $g(t)$ is sampled at 0 second, 1 second, 2 seconds, ..., and 9 seconds. Let $g(t_n = n\Delta t)$ be the sampled values of $g(t)$. In a special case of $\Delta t = 1$:

$$g(t_n) = g(n). \quad (5.5)$$

We call this process as time domain sampling.

Next consider angular frequency domain (ω Hz) sampling. Let $\Delta\omega$ Hz (radians/second) be angular frequency sampling interval:

$$\Delta\omega \equiv \frac{2\pi}{N\Delta t} \equiv \frac{2\pi}{T}. \quad (5.6)$$

The second step to DFT is to take N samples of a continuous FT $\mathcal{G}(\omega)$ at k -th angular frequency sampling instant $\omega_k = k\Delta\omega$ (radians) with $k = 0, 1, \dots, N-1$. When $T = 10$ seconds and $N = 10$ samples (i.e. $\Delta t = 1$), $\mathcal{G}(\omega)$ is sampled at 0 radian, $\pi/5$ radians, $2\pi/5$ radians, $3\pi/5$ radians, ..., and $9\pi/5$ radians. Let $\mathcal{G}(\omega_k)$ be the sampled values of $\mathcal{G}(\omega)$:

$$\mathcal{G}(\omega_k) \equiv \mathcal{G}(k\Delta\omega) \equiv \mathcal{G}\left(k \frac{2\pi}{N\Delta t}\right) \equiv \mathcal{G}\left(k \frac{2\pi}{T}\right). \quad (5.7)$$

$\mathcal{G}(\omega_k)$ is called a spectrum of $g(t_n)$ at angular frequency ω_k and it is a complex number.

DFT defines the relationship between the sampled wave in time domain $g(t_n)$ and its spectrum at angular frequency ω_k , $\mathcal{G}(\omega_k)$, as:

$$\begin{aligned} \mathcal{G}(\omega_k \equiv k \frac{2\pi}{N\Delta t}) &\equiv \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\{i\omega_k t_n\} \\ \mathcal{G}\left(k \frac{2\pi}{N\Delta t}\right) &\equiv \sum_{n=0}^{N-1} g(n\Delta t) \exp\{2\pi i kn/N\}, \end{aligned} \quad (5.8)$$

and:

$$\begin{aligned} g(t_n \equiv n\Delta t) &\equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(\omega_k \equiv k \frac{2\pi}{N\Delta t}) \exp(-i\omega_k t_n) \\ g(n\Delta t) &\equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(k \frac{2\pi}{N\Delta t}) \exp(-2\pi i kn/N). \end{aligned} \quad (5.9)$$

As you can see, DFT replaces an infinite integral of FT with summation of N points.

[5.2] Definition of DFT

[5.2.1] Physicists' Definition of DFT

Likewise FT, there are several different definitions of DFT depending on the field of study. We first follow physicists' convention. We begin with the most general definition of DFT of a continuous time domain function $g(t)$ into angular frequency domain function $\mathcal{G}(\omega)$ and its inverse using DFT parameters (a,b) as:

$$\mathcal{G}(\omega_k \equiv k \frac{2\pi}{N\Delta t}) \equiv \frac{1}{N^{\frac{1-a}{2}}} \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\{ib\omega_k t_n\}, \quad (5.10)$$

$$g(t_n \equiv n\Delta t) \equiv \frac{1}{N^{\frac{1+a}{2}}} \sum_{k=0}^{N-1} \mathcal{G}(\omega_k \equiv k \frac{2\pi}{N\Delta t}) \exp(-ib\omega_k t_n). \quad (5.11)$$

When DFT parameters are $(a,b) = (1,1)$ for the purpose of calculating characteristic functions, the definitions (5.10) and (5.11) become the definitions (5.8) and (5.9).

[5.2.2] Signal Processing Definition of DFT

In the field of signal processing (and most of the textbooks about FT), frequency f Hz (cycles/second) is used instead of ω Hz (radians/second). Consider frequency domain f sampling. Let Δf Hz be frequency domain sampling interval (also called frequency resolution):

$$\Delta f \equiv \frac{1}{N\Delta t} \equiv \frac{1}{T}. \quad (5.12)$$

Take N samples of a continuous FT $\mathcal{G}(f)$ at k -th frequency sampling instant $f_k = k\Delta f$ Hz with $k = 0, 1, \dots, N-1$. When $T = 10$ seconds and $N = 10$ samples (i.e. $\Delta t = 1$), $\mathcal{G}(f)$ is sampled at 0 Hz, 1/10 Hz, 2/10 Hz, ..., and 9/10 Hz. Let $\mathcal{G}(f_k)$ be the sampled values of $\mathcal{G}(f)$:

$$\mathcal{G}(f_k) \equiv \mathcal{G}(k\Delta f) \equiv \mathcal{G}\left(k \frac{1}{N\Delta t}\right) \equiv \mathcal{G}\left(k \frac{1}{T}\right). \quad (5.13)$$

$\mathcal{G}(f_k)$ is called a spectrum of $g(t_n)$ at frequency f_k Hz and it is a complex number.

The general definition of DFT of a continuous time domain function $g(t)$ into frequency domain function $\mathcal{G}(f)$ and its inverse using DFT parameters (a, b) become (i.e. equivalent to (5.10) and (5.11)):

$$\mathcal{G}(f_k \equiv k \frac{1}{N\Delta t}) \equiv \frac{1}{N^{\frac{1-a}{2}}} \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\left\{2\pi i b kn / N\right\}, \quad (5.14)$$

$$g(t_n \equiv n\Delta t) \equiv \frac{1}{N^{\frac{1+a}{2}}} \sum_{k=0}^{N-1} \mathcal{G}(f_k \equiv k \frac{1}{N\Delta t}) \exp\left(-2\pi i b kn / N\right). \quad (5.15)$$

In the signal processing, DFT parameters $(1, -1)$ are used:

$$\mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \equiv \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\left\{-2\pi i kn / N\right\}, \quad (5.16)$$

$$g(t_n \equiv n\Delta t) \equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \exp\left(2\pi i kn / N\right). \quad (5.17)$$

[5.2.3] Requirement of DFT

FT and inverse FT defined by the equations (5.10) and (5.11) require both the time domain function $g(t_n \equiv n\Delta t)$ and the angular frequency domain function $\mathcal{G}(\omega_k \equiv k \frac{2\pi}{N\Delta t})$ to be periodic. This means for $\theta = 0, \pm 1, \pm 2, \dots$:

$$g(t_n \equiv n\Delta t) = g\left[(\theta N + n)\Delta t\right],$$

$$\mathcal{G}(\omega_k \equiv k \frac{2\pi}{N\Delta t}) = \mathcal{G}\left[\left(\theta N + k\right) \frac{2\pi}{N\Delta t}\right].$$

FT and inverse FT defined by the equations (5.16) and (5.17) require both the time domain function $g(t_n \equiv n\Delta t)$ and the frequency domain function $\mathcal{G}(f_k \equiv \frac{k}{N\Delta t})$ to be periodic. This means for $\theta = 0, \pm 1, \pm 2, \dots$:

$$g(t_n \equiv n\Delta t) = g\left[(\theta N + n)\Delta t\right],$$

$$\mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) = \mathcal{G}\left[\frac{(\theta N + k)}{N\Delta t}\right].$$

[5.2.4] Sampling Theorem and Nyquist Rule: How to Determine the Time Domain Sampling Interval Δt

At this point we are very familiar with the definition of DFT at least in terms of concept. DFT is simply a discrete approximation of FT by taking a N point sample both in the time domain t (seconds) and frequency (angular frequency) domain f (ω) Hz. When implementing DFT, the following very important question about the frequency of sampling arises. What is the appropriate time domain sampling interval Δt (seconds/sample), 1 second, 0.5 seconds, or 0.01 seconds? In terms of sampling rate $f_s = 1/\Delta t$ (samples/second), question becomes the following. What is the appropriate sampling rate f_s , 1 Hz, 2 Hz, or 100 Hz?

If time domain function $g(t)$ is not sampled at an appropriately high rate (i.e. you chose too large time domain sampling interval Δt or you chose too small sampling rate f_s Hz), it turns out that DFT yields distorted (overlapped) approximation of FT known as aliasing. According to the sampling theorem of Fourier transform, a continuous time domain function $g(t)$ can be uniquely determined by its sampled values $g(t_n \equiv n\Delta t)$ by

choosing the time domain sampling interval of $\Delta t = \frac{1}{2f_c}$ (seconds) if the FT $\mathcal{G}(f)$ of $g(t)$ is zero for all frequencies greater than f_c Hz (i.e. if the FT $\mathcal{G}(f)$ of $g(t)$ is band-limited at the frequency f_c Hz):

$$\mathcal{G}(f) = 0 \quad \text{if} \quad |f| > f_c.$$

The time domain sampling interval of $\Delta t = \frac{1}{2f_c}$ (seconds) is the maximum value of the interval without aliasing. In other words, aliasing occurs if $\Delta t > \frac{1}{2f_c}$ and aliasing does not occur if $\Delta t < \frac{1}{2f_c}$.

In terms of a sampling rate $f_s = 1/\Delta t$ Hz, the sampling theorem becomes the following. A continuous time domain function $g(t)$ can be uniquely determined by its sampled values $g(t_n \equiv n\Delta t)$ by choosing the sampling rate of:

$$f_s = 2f_c \text{ Hz}, \quad (5.18)$$

where f_c is the folding frequency of the waveform $g(t)$. The sampling rate of $f_s = 2f_c$ Hz is the minimum value of sampling rate without aliasing. In other words, aliasing occurs if $f_s < 2f_c$ and aliasing does not occur if $f_s \geq 2f_c$.

After simple rearrangement:

$$\begin{aligned}\Delta t &< \frac{1}{2f_c} \\ \frac{1}{\Delta t} &\equiv f_s < 2f_c \\ \frac{1}{2}f_s &< f_c.\end{aligned}$$

Nyquist rule states that when we sample a time domain function $g(t)$ with sampling rate f_s Hz (samples/second), its FT $\mathcal{G}(f)$ is reliable (i.e. without aliasing) only up to frequency $f_s/2$ Hz. The maximum frequency without aliasing $f_s/2 = f_c$ is called a folding frequency.

Nyquist rule can be stated in terms of k -th frequency sampling instant $f_k = k\Delta f$ Hz with $k = 0, 1, \dots, N-1$. We solve for the value of k such that:

$$k\Delta f = f_s/2.$$

Since $f_s = 1/\Delta t$ and $\Delta f = 1/N\Delta t$:

$$\begin{aligned}k \frac{1}{N\Delta t} &= \frac{1}{2} \frac{1}{\Delta t} \\ k &= \frac{N}{2}.\end{aligned}$$

This means that when we take a N point sample of a time domain function $g(t)$, its FT $\mathcal{G}(f)$ is reliable (i.e. without aliasing) only up to $k = N/2$ -th frequency sampling instant. In other words, only half of N point DFT outputs are reliable.

Consider DFT of a continuous time domain function $g(t)$ with the sampling rate $f_s = 200$ Hz (samples/second) for the total sampling time $T = 3$ (seconds). $g(t)$ is sampled in the time domain with the interval $\Delta t = 1/f_s = 1/200 = 0.005$ (seconds/sample). The number of samples taken is $N = T/\Delta t = 3/0.005 = 600$. Frequency resolution (frequency domain sampling interval) is $\Delta f \equiv 1/N\Delta t \equiv 1/T = 1/3$ Hz. Following Nyquist rule, FT $\mathcal{G}(f)$ is reliable (i.e. without aliasing) only up to frequency $f_s/2 = 200/2 = 100$ Hz. In other words, out of 600 DFT outputs $\mathcal{G}(f_k) \equiv \frac{k}{N\Delta t}$ with $k = 0, 1, 2, \dots, 599$ only the half $N/2 = 600/2 = 300$ are reliable.

Read the chapter 5 of Brigham (1988) which is an excellent book for more rigorous treatment of the sampling theorem.

[5.3] Examples of DFT

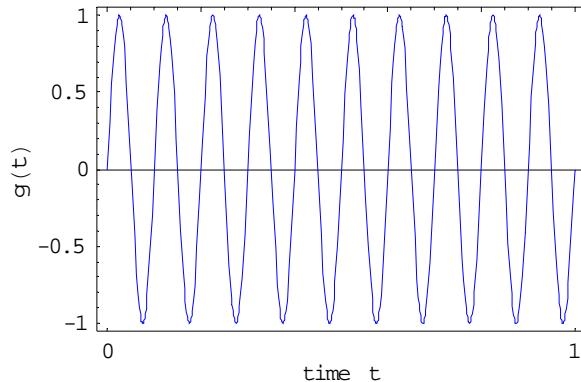
Above all, section 5.3.1 is the most important example. Using an example of a sine wave $g(t) = A \sin(2\pi f_0 t) = A \sin(\omega_0 t)$, following important features of DFT are explained in detail: 1) Approximation error inherent in DFT called leakage, 2) Leakage cannot always be reduced by increasing a sampling rate f_s Hz (cycles/second) with holding the number of samples taken N constant, 3) Guideline of choosing the optimal sampling rate f_s Hz when the folding frequency f_c of the waveform is known (Nyquist rule), and 4) The only remedy for leakage is to sample more points (i.e. higher N).

[5.3.1] Sine Wave $g(t) = A \sin(2\pi f_0 t) = A \sin(\omega_0 t)$

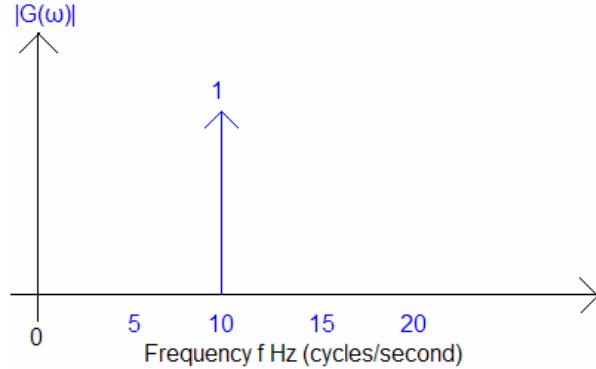
Consider 10 Hz (cycles/second) sine wave (i.e. fundamental frequency $f_0 = 10$) with amplitude $A = 1$ volt plotted in Panel A of Figure 5.1:

$$g(t) = \sin(2\pi 10t). \quad (5.19)$$

We saw in section 3.2.6 that the ideal FT of (5.19) is an impulse function of normalized magnitude (frequency spectrum) 1 volt at frequency $f = 10$ Hz plotted in Panel B of Figure 5.1 which implies that the frequency $f = 10$ Hz contains all the energy of the waveform. Normalized magnitude is normalized so that it equals the amplitude A of the waveform $g(t)$.



A) Plot of a 10 Hz sine wave $g(t) = \sin(2\pi 10t)$.



B) Plot of the ideal normalized FT of $g(t) = \sin(2\pi 10t)$.

Figure 5.1: Plot of a 10 Hz Sine Wave and Its Ideal Normalized FT

Next, we consider DFT. Since DFT is only an approximation of FT, the frequency spectrum produced by DFT is not as clean as the Panel B Figure 5.1 meaning that it contains a lot of noise. This degree of cleanliness (noise) depends on such factors as the sampling rate f_s Hz (or $\Delta t = 1/f_s$), the point of sample N , and the total sampling time T which are all related by:

$$f_s \equiv N/T. \quad (5.20)$$

First, we sample a time domain function (5.19) by a sampling rate of $f_s = 200$ Hz (samples/second) and the number of samples $N = 256$ (samples). In other words, we take a $N = 256$ point sample of (5.19) with a time domain sampling interval:

$$\Delta t = 1/f_s = 1/200 = 0.005 \text{ (seconds/sample)},$$

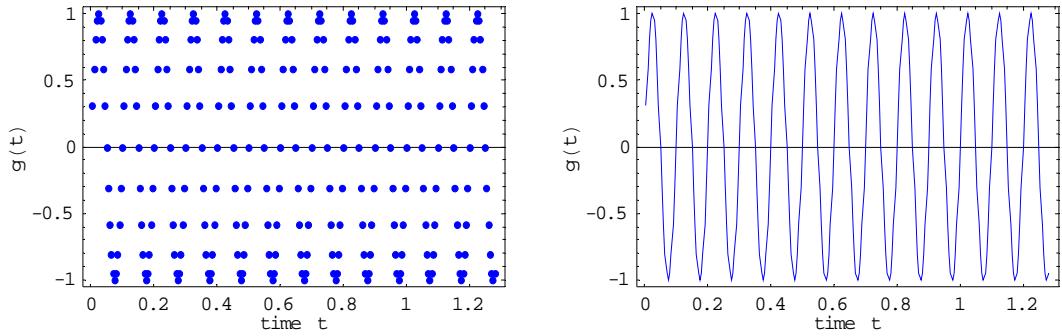
and total sampling time:

$$T = N\Delta t = 256 \times 0.005 = 1.28 \text{ (seconds)}.$$

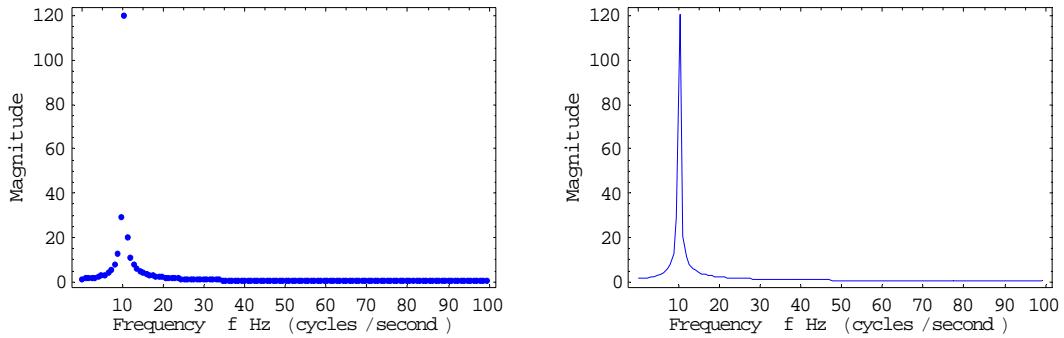
Sampled 10 Hz sine wave (5.19) is plotted in Panel A of Figure 5.2.

Secondly, consider frequency domain f sampling. Frequency resolution (frequency domain sampling interval) Δf Hz (cycles/second) is following (5.12):

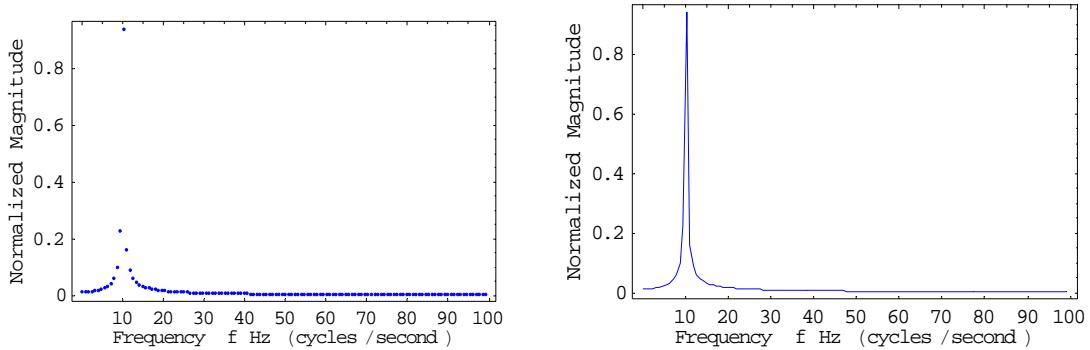
$$\Delta f \equiv \frac{1}{T \equiv N\Delta t} = \frac{1}{1.28} = 0.78125.$$



A) Plot of sampled 10 Hz sine wave $g(t) = \sin(2\pi 10t)$ on the left. Points are joined on the right.



B) Plot of DFT non-normalized frequency spectrum of sampled 10 Hz sine wave $g(t) = \sin(2\pi 10t)$ on the left. Points are joined on the right.



C) Plot of DFT normalized frequency spectrum of sampled 10 Hz sine wave $g(t) = \sin(2\pi 10t)$ on the left. Points are joined on the right.

Figure 5.2: Plot of Sampled 10 Hz Sine Wave and Its DFT Output

Following the signal processing definition of DFT (5.16), frequency spectrum of sampled 10 Hz sine wave (5.19) is plotted in Panel B of Figure 5.2. The frequency axis goes from 0 to 100 Hz because aliasing does not occur up to the folding frequency $f_s / 2 = 200 / 2 = 100$ Hz. Frequency spectrum in Panel C is normalized so that the normalized magnitude equals the amplitude A of the waveform $g(t)$ (i.e. in this example $A = 1$). This normalization is done by weighting the non-normalized magnitude by:

$$\frac{1}{N/2}.$$

Several important features should be noticed from the Panel C of Figure 5.2. Continuous (normalized) FT $\mathcal{G}(f)$ of a sine wave (5.19) is a unit impulse function at the frequency $f_0 = 10$ Hz (this fundamental frequency is also a folding frequency) as depicted in the Panel B of Figure 5.1. Its approximation by DFT is not an impulse function, but it is rather a spike with significant portion of the energy distributed around 10 Hz (i.e. positive energy in the interval $5 < f < 15$ Hz). In other words, the approximation by DFT leaks the energy around the original impulse. This approximation error inherent in DFT is called a leakage which occurs when the time domain sampling does not end at the phase of the sine wave as it started. As a consequence of leakage, the normalized maximum magnitude of DFT $\mathcal{G}(f)$ is less than 1 compared to the exactly 1 normalized maximum magnitude of an original impulse. Theoretically speaking, the only remedy for leakage is to take infinitely many samples (i.e. $N \rightarrow \infty$). But practically speaking, DFT always contains leakage because it is impossible to take infinitely many samples.

Some readers might think that the degree of leakage can be reduced by taking samples with higher sampling rate f_s Hz with no change in N . The answer is no. To show why this is the case, we redo DFT of a sine wave (5.19) with higher sampling rate $f_s = 1000$ Hz (5 times faster). First, we sample a time domain function (5.19) by a sampling rate of $f_s = 1000$ Hz (samples/second) and the number of samples $N = 256$ (samples). In other words, we take a $N = 256$ point sample of (5.19) with a time domain sampling interval:

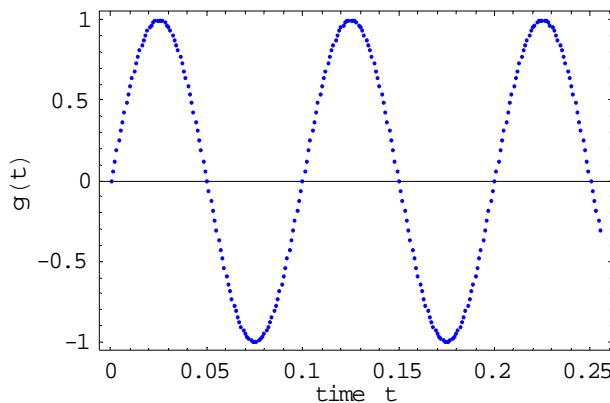
$$\Delta t = 1/f_s = 1/1000 = 0.001 \text{ (seconds/sample)},$$

and total sampling time:

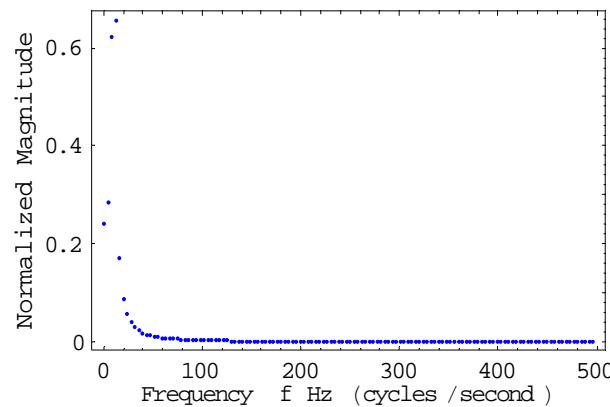
$$T = N\Delta t = 256 \times 0.001 = 0.256 \text{ (seconds)}.$$

Sampled 10 Hz sine wave (5.19) is plotted in Panel A of Figure 5.3. Secondly, consider frequency domain f sampling. Frequency resolution (frequency domain sampling interval) Δf Hz (cycles/second) is following (5.12):

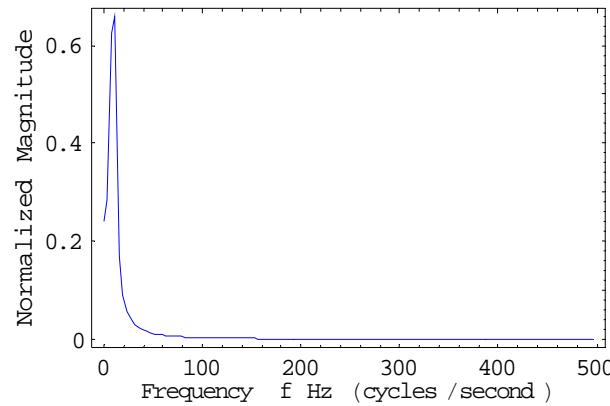
$$\Delta f \equiv \frac{1}{T \equiv N\Delta t} = \frac{1}{0.256} = 3.90625.$$



A) Plot of sampled 10 Hz sine wave with the sampling rate $f_s = 1000$ Hz.



B) Plot of DFT normalized frequency spectrum of sampled 10 Hz sine wave
 $g(t) = \sin(2\pi 10t)$ with the sampling rate $f_s = 1000$ Hz.



C) Points are joined.

Figure 5.3: Plot of Sampled 10 Hz Sine Wave and Its Normalized DFT Output with Sampling Rate $f_s = 1000$ Hz.

Following the signal processing definition of DFT (5.16), normalized frequency spectrum of sampled 10 Hz sine wave (5.19) with $f_s = 1000$ Hz is plotted in Panel B and C of

Figure 5.3. The frequency axis goes from 0 to 500 Hz because aliasing does not occur up to the folding frequency $f_s / 2 = 1000 / 2 = 500$ = 500 Hz according to Nyquist rule. As you realize now, leakage gets worse as a result of an increase in a sampling rate f_s from 200 to 1000 Hz. We can see that the normalized maximum magnitude of DFT $\mathcal{G}(f)$ is 0.6573267 (significantly less than 1 which is the original normalized magnitude) which in turn indicates that the degree of leakage of energy around 10 Hz is greater. In addition, the normalized maximum magnitude 0.6573267 occurs at 11.7188 Hz, although continuous normalized FT is a unit impulse function at the frequency $f_0 = 10$ Hz. The reason that the accuracy of the approximation by DFT gets poorer is the fact that the frequency resolution gets poorer from:

$$\Delta f \equiv \frac{1}{T \equiv N\Delta t} = \frac{1}{1.28} = 0.78125 \text{ Hz},$$

to:

$$\Delta f \equiv \frac{1}{T \equiv N\Delta t} = \frac{1}{0.256} = 3.90625 \text{ Hz},$$

as a result of an increase in a sampling rate f_s from 200 to 1000 Hz (i.e. because an increase in f_s is an decrease in Δt through the relationship $f_s \equiv \frac{1}{\Delta t}$). The bottom line is that the degree of leakage cannot necessarily be reduced by taking samples with higher sampling rate f_s Hz with no change in N .

The above example indicates that for a 10 Hz sine wave using 1000 Hz sampling rate f_s is too much (i.e. frequency resolution becomes too poor). So the next natural question arises that how can we determine the appropriate sampling rate f_s Hz? Following Nyquist rule of (5.18), the sampling rate of $f_s = 2f_c$ Hz is the minimum value of sampling rate without aliasing. As long as the folding frequency f_c of the waveform $g(t)$ is known (in our example $f_c = 10$ Hz), the minimum value of sampling rate without aliasing can be obtained by doubling f_c . Although this is a nice rule, its practical usefulness is very doubtful because in other than textbook examples the folding frequency f_c of the waveform $g(t)$ is not known a priori. Redo DFT of a sine wave (5.19) using the sampling rate $f_s = 50$ Hz. We take a $N = 256$ point sample of (5.19) with a time domain sampling interval:

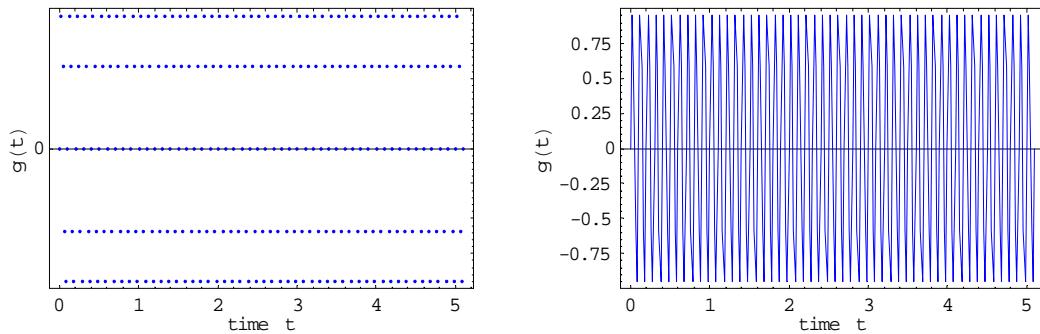
$$\Delta t = 1/f_s = 1/50 = 0.02 \text{ (seconds/sample)},$$

and total sampling time:

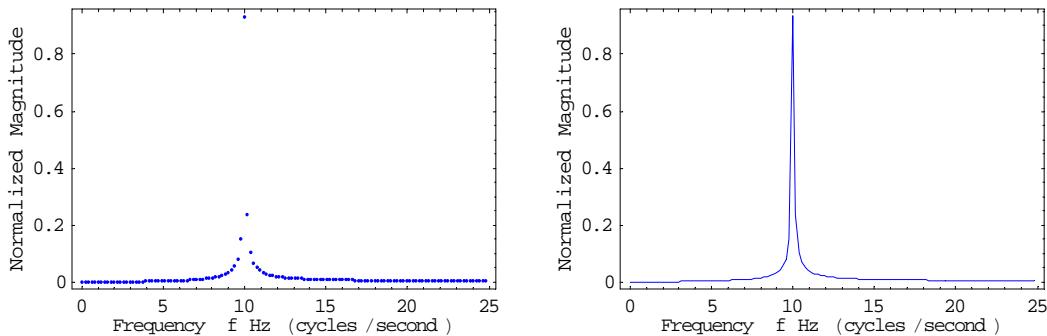
$$T = N\Delta t = 256 \times 0.02 = 5.12 \text{ (seconds).}$$

Sampled 10 Hz sine wave (5.19) is plotted in Panel A of Figure 5.4. Frequency resolution of DFT (frequency domain sampling interval) Δf Hz (cycles/second) is following (5.12):

$$\Delta f \equiv \frac{1}{T \equiv N\Delta t} = \frac{1}{5.12} = 0.1953125.$$



A) Plot of sampled 10 Hz sine wave $g(t) = \sin(2\pi 10t)$ using the sampling rate $f_s = 50$ Hz on the left. Points are joined on the right.



B) Plot of DFT normalized frequency spectrum of sampled 10 Hz sine wave $g(t) = \sin(2\pi 10t)$ using the sampling rate $f_s = 50$ Hz on the left. Points are joined on the right.

Figure 5.4: Plot of Sampled 10 Hz Sine Wave and Its Normalized DFT Output with Sampling Rate $f_s = 50$ Hz.

By reducing the sampling rate from $f_s = 1000$ Hz to $f_s = 50$ Hz, DFT approximates continuous FT (a unit impulse function at 10 Hz in this example) much better because the frequency resolution Δf increases from 3.90625 Hz to 0.1953125 Hz. Again, the important result is the fact that higher sampling rate f_s does not always reduce leakage (DFT inherent error) depending on the folding frequency f_c of the waveform.

The only way to reduce leakage is to sample more points. To illustrate this, we take a $N = 4 \times 256 = 1024$ point sample of (5.19) using the sampling rate $f_s = 50$ Hz with a time domain sampling interval:

$$\Delta t = 1/f_s = 1/50 = 0.02 \text{ (seconds/sample)},$$

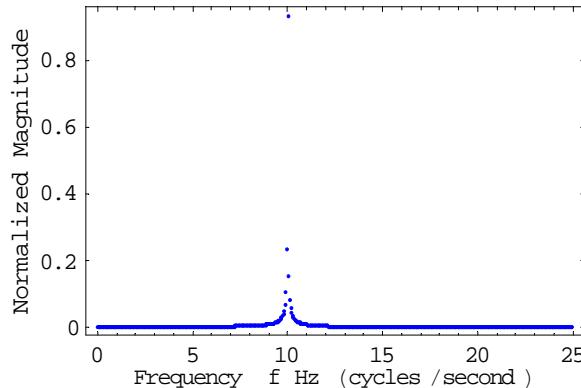
and total sampling time:

$$T = N\Delta t = 1024 \times 0.02 = 20.48 \text{ (seconds)}.$$

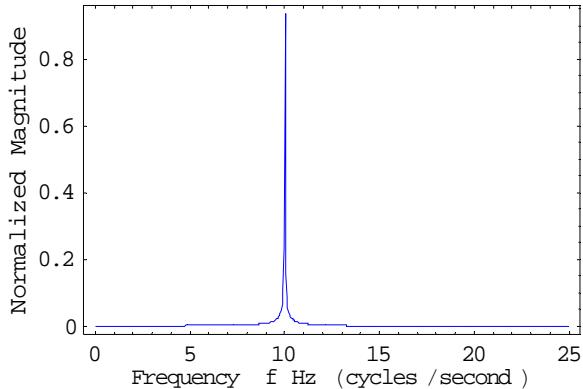
Sampled 10 Hz sine wave (5.19) is plotted in Figure 5.5. Frequency resolution of DFT (frequency domain sampling interval) Δf Hz (cycles/second) is following (5.12):

$$\Delta f \equiv \frac{1}{T \equiv N\Delta t} = \frac{1}{20.48} = 0.048828125.$$

Note that the frequency resolution Δf become 4 times finer because the number of samples taken N increased by four-fold. As a result, leakage has been dramatically reduced. In fact, in Figure 5.5 the peak magnitude is 0.935 at a frequency of 10.01 Hz which is really close (i.e. a good approximation) to the continuous FT of the peak magnitude 1 at 10 Hz.



A) Plot of $N = 1,024$ point-DFT normalized frequency spectrum of 10 Hz sine wave $g(t) = \sin(2\pi 10t)$. Sampling rate $f_s = 50$ Hz is used.



B) Points are joined.

Figure 5.5: Plot of N = 1,024 Point-DFT of a 10-Hz Sine Wave and Its Normalized DFT Output with Sampling Rate $f_s = 50$ Hz.

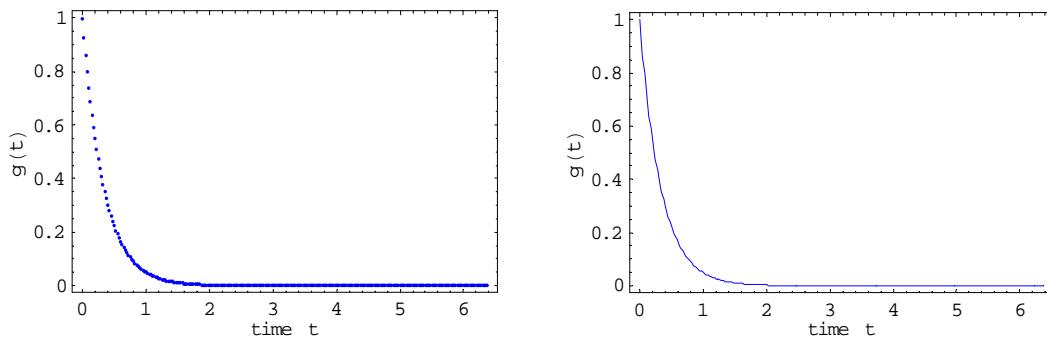
[5.3.2] Double-Sided Exponential

Consider a double-sided exponential function with $A, \alpha \in \mathbb{R}$:

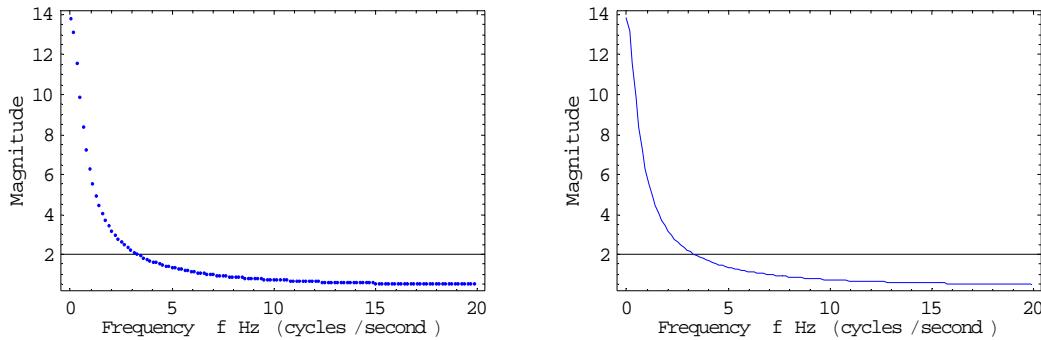
$$g(t) = Ae^{-\alpha|t|}.$$

Set $A = 1$ and $\alpha = 3$. Following the definition (5.16), $N = 256$ point DFT of a double-sided exponential function with $f_s = 40$ Hz (samples/second) sampling rate is performed on Figure 5.6. Time domain sampling interval Δt (seconds/sample), total sampling time T (seconds), and frequency resolution Δf Hz (cycles/second) are:

$$\begin{aligned}\Delta t &= 1/f_s = 1/40 = 0.025, \\ T &= N\Delta t = 256 \times 0.025 = 6.4, \\ \Delta f &\equiv 1/N\Delta t \equiv 1/T = 1/6.4 = 0.15625.\end{aligned}$$



A) Plot of sampled $g(t) = e^{-3|t|}$ using the sampling rate $f_s = 40$ Hz. Points are joined on the right.



B) Plot of $N = 256$ point DFT frequency spectrum of sampled $g(t) = e^{-3|t|}$ using the sampling rate $f_s = 40$ Hz. Points are joined on the right.

Figure 5.6: Plot of Sampled Double-Sided Exponential Function $g(t) = e^{-3|t|}$ and $N = 256$ Point DFT with Sampling Rate $f_s = 40$ Hz.

[5.3.3] Rectangular Pulse

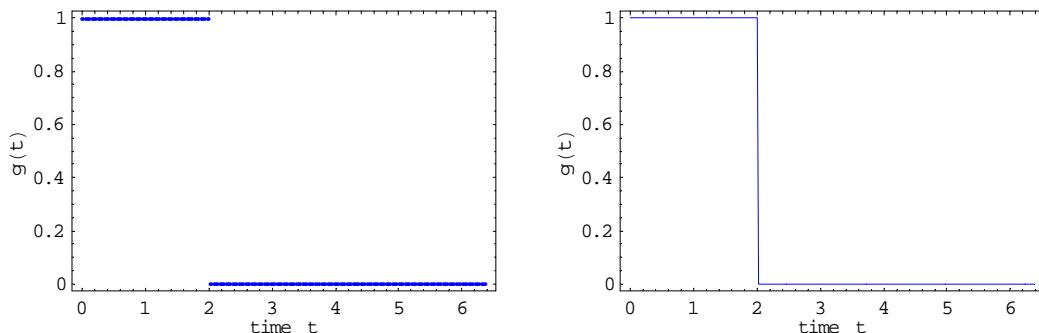
Consider a rectangular pulse with $A, T_0 \in \mathbb{R}$:

$$g(t) = \begin{cases} A & -T_0 \leq t \leq T_0 \\ 0 & |t| > T_0 \end{cases},$$

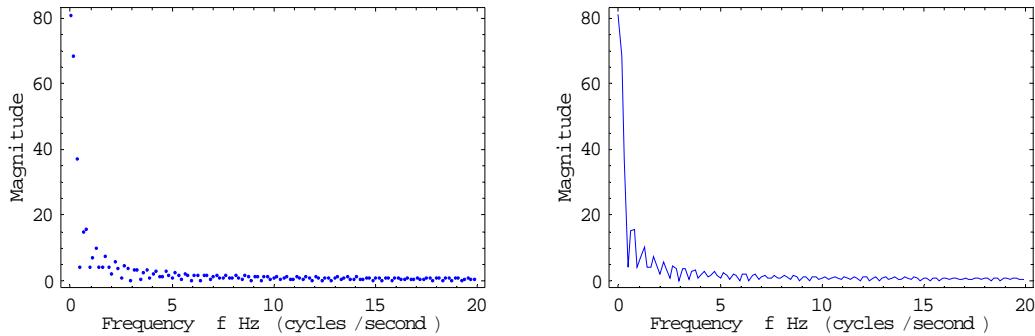
which is an even function of t (symmetric with respect to t).

Set $A = 1$ and $T_0 = 2$. Following the definition (5.16), $N = 256$ point DFT of a rectangular pulse with $f_s = 40$ Hz (samples/second) sampling rate is performed on Figure 5.7. Time domain sampling interval Δt (seconds/sample), total sampling time T (seconds), and frequency resolution Δf Hz (cycles/second) are:

$$\begin{aligned}\Delta t &= 1/f_s = 1/40 = 0.025, \\ T &= N\Delta t = 256 \times 0.025 = 6.4, \\ \Delta f &\equiv 1/N\Delta t \equiv 1/T = 1/6.4 = 0.15625.\end{aligned}$$



A) Plot of sampled rectangular pulse using the sampling rate $f_s = 40$ Hz. Points are joined on the right.



B) Plot of $N = 256$ point DFT frequency spectrum of sampled rectangular pulse using the sampling rate $f_s = 40$ Hz. Points are joined on the right.

Figure 5.7: Plot of Sampled Rectangular Pulse and $N = 256$ Point DFT with Sampling Rate $f_s = 40$ Hz.

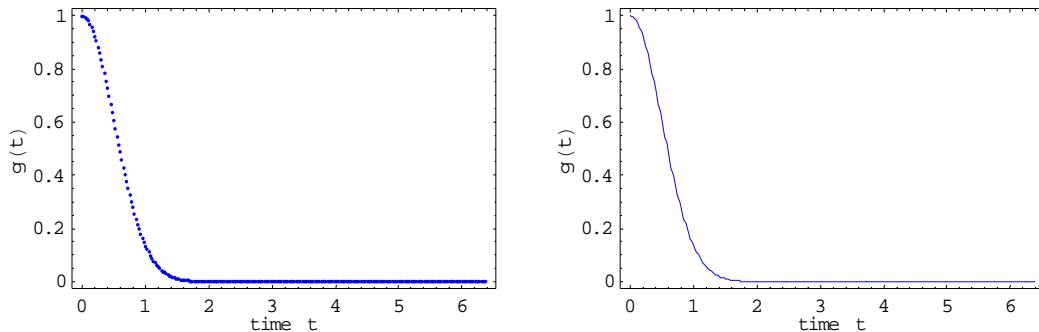
[5.3.4] Gaussian Function

Consider a Gaussian function with $A \in \mathbb{R}^+$:

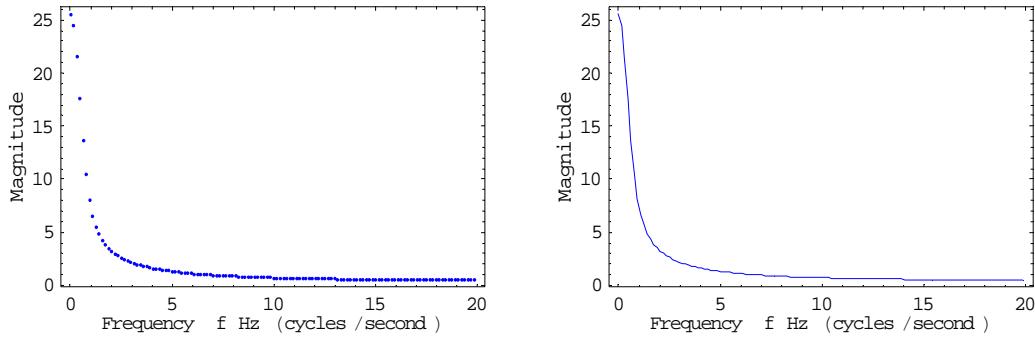
$$g(t) = e^{-At^2}.$$

Set $A = 2$. Following the definition (5.16), $N = 256$ point DFT of a Gaussian function with $f_s = 40$ Hz (samples/second) sampling rate is performed on Figure 5.8. Time domain sampling interval Δt (seconds/sample), total sampling time T (seconds), and frequency resolution Δf Hz (cycles/second) are:

$$\begin{aligned}\Delta t &= 1/f_s = 1/40 = 0.025, \\ T &= N\Delta t = 256 \times 0.025 = 6.4, \\ \Delta f &\equiv 1/N\Delta t \equiv 1/T = 1/6.4 = 0.15625.\end{aligned}$$



A) Plot of sampled $g(t) = e^{-2t^2}$ using the sampling rate $f_s = 40$ Hz. Points are joined on the right.



B) Plot of $N = 256$ point DFT frequency spectrum of sampled $g(t) = e^{-2t^2}$ using the sampling rate $f_s = 40$ Hz. Points are joined on the right.

Figure 5.8: Plot of Sampled Gaussian Function $g(t) = e^{-2t^2}$ and N = 256 Point DFT with Sampling Rate $f_s = 40$ Hz.

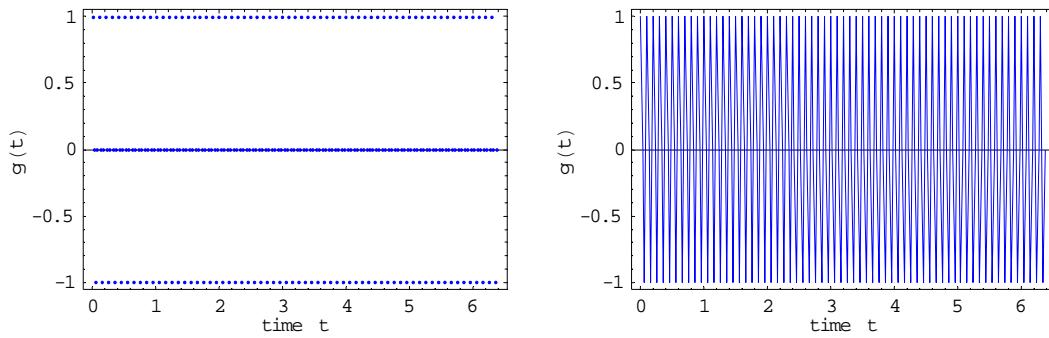
[5.3.5] **Cosine Wave** $g(t) = A \cos(2\pi f_0 t) = A \cos(\omega_0 t)$

Consider 10 Hz (cycles/second) cosine wave (i.e. fundamental frequency $f_0 = 10$) with amplitude $A = 1$ volt:

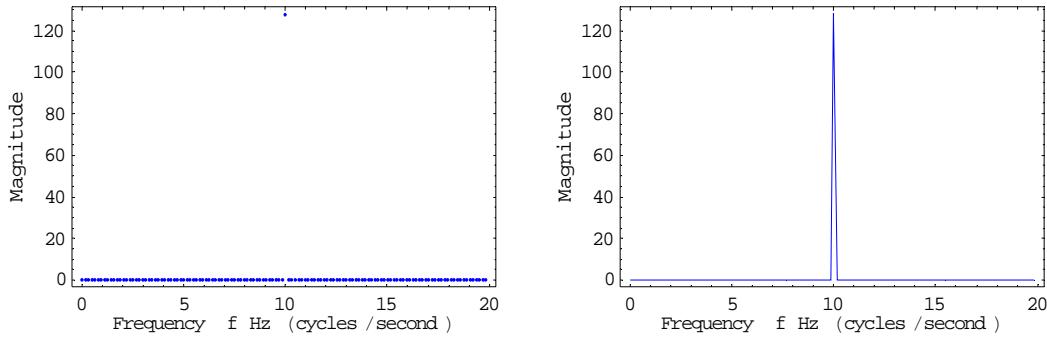
$$g(t) = \cos(2\pi 10t).$$

Following the definition (5.16), $N = 256$ point DFT of a cosine function with $f_s = 40$ Hz (samples/second) sampling rate is performed on Figure 5.9. Time domain sampling interval Δt (seconds/sample), total sampling time T (seconds), and frequency resolution Δf Hz (cycles/second) are:

$$\begin{aligned}\Delta t &= 1/f_s = 1/40 = 0.025, \\ T &= N\Delta t = 256 \times 0.025 = 6.4, \\ \Delta f &\equiv 1/N\Delta t \equiv 1/T = 1/6.4 = 0.15625.\end{aligned}$$



A) Plot of sampled $g(t) = \cos(2\pi 10t)$ using the sampling rate $f_s = 40$ Hz. Points are joined on the right.



B) Plot of $N = 256$ point DFT frequency spectrum of sampled $g(t) = \cos(2\pi 10t)$ using the sampling rate $f_s = 40$ Hz. Points are joined on the right.

Figure 5.9: Plot of Sampled Cosine Wave $g(t) = \cos(2\pi 10t)$ and $N = 256$ Point DFT with Sampling Rate $f_s = 40$ Hz.

[5.4] Properties of DFT

DFT retains all the properties of continuous FT since DFT is a special case of FT.

[5.4.1] Linearity of DFT

Suppose that the time domain sequence $\{f(t_n) : n = 0, 1, \dots, N-1\}$ and $\{g(t_n) : n = 0, 1, \dots, N-1\}$ have discrete Fourier transforms $\{\mathcal{F}(f_k) : k = 0, 1, \dots, N-1\}$ and $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16). Then:

$$a\mathcal{F}(f_k) + b\mathcal{G}(f_k) = \sum_{n=0}^{N-1} (af(t_n) + bg(t_n)) \exp\{-2\pi i kn/N\}. \quad (5.21)$$

PROOF

$$\begin{aligned} & \sum_{n=0}^{N-1} (af(t_n) + bg(t_n)) \exp\{-2\pi i kn/N\} \\ &= a \sum_{n=0}^{N-1} f(t_n) \exp\{-2\pi i kn/N\} + b \sum_{n=0}^{N-1} g(t_n) \exp\{-2\pi i kn/N\} \\ &= a\mathcal{F}(f_k) + b\mathcal{G}(f_k). \end{aligned}$$

[5.4.2] DFT of Even and Odd Functions

Let $even(x)$ be an even function and $odd(x)$ be an odd function. Integral properties of even and odd functions are:

$$\begin{aligned}\int_{-A}^A \text{odd}(x)dx &= 0, \\ \int_{-A}^A \text{even}(x)dx &= 2\int_0^A \text{even}(x)dx.\end{aligned}$$

If the time domain sequence $\{g(t_n) : n = 0, 1, \dots, N-1\}$ is even, i.e. $g(t_n) = g(-t_n)$, then its DFT $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16) is a real-valued even function:

$$\mathcal{G}(f_k) = \sum_{n=0}^{N-1} g(t_n) \cos\left(\frac{2\pi kn}{N}\right). \quad (5.22)$$

PROOF

Start from the DFT definition (5.16):

$$\mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \equiv \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\{-2\pi i kn / N\}.$$

Using the Euler's formula (2.16) of $e^{-ix} = \cos(x) - i \sin(x)$:

$$\begin{aligned}\mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) &\equiv \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \{\cos(2\pi kn / N) - i \sin(2\pi kn / N)\} \\ \mathcal{G}(f_k) &= \sum_{n=0}^{N-1} g(t_n) \cos(2\pi kn / N) - i \sum_{n=0}^{N-1} g(t_n) \sin(2\pi kn / N).\end{aligned}$$

Since the product of an even function $g(t_n)$ and an odd function $\sin(2\pi kn / N)$ is odd, from its integral property:

$$\sum_{n=0}^{N-1} g(t_n) \sin(2\pi kn / N) = 0.$$

Thus:

$$\mathcal{G}(f_k) = \sum_{n=0}^{N-1} g(t_n) \cos(2\pi kn / N),$$

which is a real-valued even function since it is a product of two real-valued even functions $g(t_n)$ and $\cos(2\pi kn / N)$.

Next consider an odd case. If the time domain sequence $\{g(t_n) : n = 0, 1, \dots, N-1\}$ is odd, i.e. $g(t_n) = -g(-t_n)$, then its DFT $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16) is a complex-valued odd function:

$$\mathcal{G}(f_k) = -i \sum_{n=0}^{N-1} g(t_n) \sin\left(\frac{2\pi kn}{N}\right). \quad (5.23)$$

PROOF

Start from the DFT definition (5.16):

$$\mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) = \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\{-2\pi i kn / N\}.$$

Using the Euler's formula (2.16) of $e^{-ix} = \cos(x) - i \sin(x)$:

$$\begin{aligned} \mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) &\equiv \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \{\cos(2\pi kn / N) - i \sin(2\pi kn / N)\} \\ \mathcal{G}(f_k) &= \sum_{n=0}^{N-1} g(t_n) \cos(2\pi kn / N) - i \sum_{n=0}^{N-1} g(t_n) \sin(2\pi kn / N). \end{aligned}$$

Since the product of an odd function $g(t_n)$ and an even function $\cos(2\pi kn / N)$ is odd, from its integral property:

$$\sum_{n=0}^{N-1} g(t_n) \cos(2\pi kn / N) = 0.$$

Thus:

$$\mathcal{G}(f_k) = -i \sum_{n=0}^{N-1} g(t_n) \sin(2\pi kn / N),$$

which is a complex-valued odd function.

[5.4.3] Symmetry of DFT

If the time domain sequence $\{g(t_n) : n = 0, 1, \dots, N-1\}$ has DFT $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16), then the sequence $\left\{ \frac{1}{N} \mathcal{G}(t_n) : n = 0, 1, \dots, N-1 \right\}$ has DFT

$\{g(-f_k) : k = 0, 1, \dots, N-1\}$. In other words, if $(g(t_n), \mathcal{G}(f_k))$ is a DFT pair,

$\left(\frac{1}{N}\mathcal{G}(t_n), g(-f_k)\right)$ is another DFT pair.

PROOF

By the definition of an inverse DFT (5.17):

$$g(t_n \equiv n\Delta t) \equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \exp(2\pi i kn/N).$$

Rewrite it by substituting $-n$ for n :

$$\begin{aligned} g(-t_n \equiv -n\Delta t) &\equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \exp(2\pi i k(-n)/N) \\ g(-t_n) &\equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f_k) \exp(2\pi i k(-n)/N). \end{aligned}$$

Exchange t_n with f_k and n with k , vice versa:

$$g(-f_k) \equiv \frac{1}{N} \sum_{n=0}^{N-1} g(t_n) \exp(-2\pi i nk/N).$$

From the definition of DFT (5.16), $\left(\frac{1}{N}\mathcal{G}(t_n), g(-f_k)\right)$ is a DFT pair.

[5.4.4] Time Shifting of DFT

Suppose that the time domain sequence $\{g(t_n) : n = 0, 1, \dots, N-1\}$ has DFT $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16). Then, DFT of the time domain sequence $\{g(t_n - t_0\Delta t) : n = 0, 1, \dots, N-1\}$ (i.e. time domain sampling moment t_n is shifted by $t_0\Delta t \in \mathbb{R}$) can be expressed in terms of $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ as:

$$\mathcal{G}(f_k) e^{-2\pi i k t_0 / N} \equiv \sum_{n=0}^{N-1} g(t_n - t_0\Delta t) \exp\{-2\pi i kn/N\}. \quad (5.24)$$

PROOF

By the definition of DFT (5.16):

$$\mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \equiv \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\{-i2\pi f_k t_n\} \equiv \sum_{n=0}^{N-1} g(t_n \equiv n\Delta t) \exp\{-2\pi i k n / N\}.$$

By setting $t_n - t_0 \Delta t = t^*$:

$$\begin{aligned} \sum_{n=0}^{N-1} g(t_n - t_0 \Delta t) \exp\{-i2\pi f_k t_n\} &= \sum_{n=0}^{N-1} g(t^*) \exp\{-i2\pi f_k (t^* + t_0 \Delta t)\} \\ &= \exp\{-i2\pi f_k t_0 \Delta t\} \sum_{n=0}^{N-1} g(t^*) \exp\{-i2\pi f_k t^*\} \\ &= \exp\{-i2\pi k t_0 / N\} \mathcal{G}(f_k). \end{aligned}$$

□

[5.4.5] Frequency Shifting: Modulation

Suppose that the time domain sequence $\{g(t_n) : n = 0, 1, \dots, N-1\}$ has DFT $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16). If DFT frequency f_k is shifted by $f_0 / N\Delta t \in \mathbb{R}$ Hz (i.e. $\mathcal{G}(f_k - \frac{f_0}{N\Delta t})$), then its inverse DFT $g(t_n)$ is multiplied by $\exp\{2\pi i n f_0 / N\}$:

$$e^{2\pi i n f_0 / N} g(t_n) = \mathcal{DFT}^{-1} \left[\mathcal{G}(f_k - \frac{f_0}{N\Delta t}) \right], \quad (5.25)$$

$$\mathcal{DFT} \left(e^{2\pi i n f_0 / N} g(t_n) \right) = \mathcal{G}(f_k - \frac{f_0}{N\Delta t}). \quad (5.26)$$

PROOF

From the definition of an inverse DFT of (5.17):

$$\begin{aligned} g(t_n \equiv n\Delta t) &\equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \exp(2\pi i k n / N), \\ g(t_n \equiv n\Delta t) &\equiv \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) \exp(i2\pi f_k t_n). \end{aligned}$$

By setting $f_k - \frac{f_0}{N\Delta t} = f^*$:

$$\frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f_k - \frac{f_0}{N\Delta t}) \exp(i2\pi f_k t_n) = \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f^*) \exp\left(i2\pi(f^* + \frac{f_0}{N\Delta t})t_n\right)$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f^*) \exp(i2\pi f^* t_n) \exp\left(i2\pi \frac{f_0}{N\Delta t} t_n\right) \\
&= \exp\left(i2\pi \frac{f_0}{N\Delta t} n\Delta t\right) \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}(f^*) \exp(i2\pi f^* t_n) \\
&= \exp(i2\pi n f_0 / N) g(t_n).
\end{aligned}$$

□

[5.4.6] Discrete Convolution: Time Convolution Theorem

Suppose that the time domain sequence $\{f(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ and $\{g(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ have discrete Fourier transforms $\{\mathcal{F}(f_k) : k = 0, 1, \dots, N-1\}$ and $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16). FT and inverse FT defined by the equations (5.16) and (5.17) require both the time domain function $g(t_n \equiv n\Delta t)$ and the frequency domain function $\mathcal{G}(f_k \equiv \frac{k}{N\Delta t})$ to be periodic. This means for $\theta = 0, \pm 1, \pm 2, \dots$:

$$\begin{aligned}
g(t_n \equiv n\Delta t) &= g[(\theta N + n)\Delta t], \\
\mathcal{G}(f_k \equiv \frac{k}{N\Delta t}) &= \mathcal{G}\left[\frac{(\theta N + k)}{N\Delta t}\right],
\end{aligned}$$

where N is the period.

Discrete convolution of time domain sequence $\{f(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ and $\{g(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ for a total of N discrete samples which is denoted as $f(n\Delta t) * g(n\Delta t)$ is defined as:

$$f(n\Delta t) * g(n\Delta t) \equiv \sum_{j=0}^{N-1} f(j\Delta t) h[(n-j)\Delta t]. \quad (5.27)$$

DFT of the discrete convolution of $\{f(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ and $\{g(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ in the time domain is equal to the multiplication in the frequency domain:

$$\mathcal{DFT}(f(n\Delta t) * g(n\Delta t)) \equiv \mathcal{DFT}\left(\sum_{j=0}^{N-1} f(j\Delta t) h[(n-j)\Delta t]\right) = \mathcal{F}(f_k) \mathcal{G}(f_k). \quad (5.28)$$

PROOF

Following the definition of DFT of (5.16):

$$\mathcal{DFT} \left(\sum_{j=0}^{N-1} f(j\Delta t) h[(n-j)\Delta t] \right) = \sum_{n=0}^{N-1} \sum_{j=0}^{N-1} f(j\Delta t) h[(n-j)\Delta t] \exp \left\{ -2\pi i k n / N \right\}.$$

Use the time shifting property of DFT of the equation (5.24):

$$\begin{aligned} \mathcal{DFT} \left(\sum_{j=0}^{N-1} f(j\Delta t) h[(n-j)\Delta t] \right) &= \sum_{j=0}^{N-1} f(j\Delta t) \sum_{n=0}^{N-1} h[(n-j)\Delta t] \exp \left\{ -2\pi i k n / N \right\} \\ &= \sum_{j=0}^{N-1} f(j\Delta t) \mathcal{G}(f_k) e^{-2\pi i k j / N} = \mathcal{G}(f_k) \sum_{j=0}^{N-1} f(j\Delta t) e^{-2\pi i k j / N} \\ &= \mathcal{F}(f_k) \mathcal{G}(f_k) \end{aligned}$$

□

[5.4.7] Discrete Frequency-Convolution Theorem

Suppose that the time domain sequence $\{f(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ and $\{g(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ have discrete Fourier transforms $\{\mathcal{F}(f_k) : k = 0, 1, \dots, N-1\}$ and $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (6.16). We assume $f(t_n \equiv n\Delta t)$, $g(t_n \equiv n\Delta t)$, $\mathcal{F}(f_k)$, and $\mathcal{G}(f_k)$ are all periodic functions. Discrete convolution of DFTs (frequency convolution) $\{\mathcal{F}(f_k) : k = 0, 1, \dots, N-1\}$ and $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ for a total of N discrete samples which is denoted as $\mathcal{F}(f_k) * \mathcal{G}(f_k)$ is defined as:

$$\mathcal{F}(f_k) * \mathcal{G}(f_k) = \sum_{j=0}^{N-1} \mathcal{F}\left(\frac{j}{N\Delta t}\right) \mathcal{G}\left(\frac{k}{N\Delta t} - \frac{j}{N\Delta t}\right). \quad (5.29)$$

Inverse DFT of the frequency convolution $\mathcal{F}(f_k) * \mathcal{G}(f_k)$ scaled by $1/N$ is equal to the multiplication in the time domain:

$$\mathcal{DFT}^{-1} \left[\frac{1}{N} \mathcal{F}(f_k) * \mathcal{G}(f_k) \right] = f(t_n) g(t_n). \quad (5.30)$$

In other words, DFT of a multiplication in the time domain $f(t_n) g(t_n)$ is equal to a convolution $\mathcal{F}(f_k) * \mathcal{G}(f_k)$ scaled by $1/N$ in the frequency domain:

$$\mathcal{DFT}(f(t_n) g(t_n)) = \frac{1}{N} \mathcal{F}(f_k) * \mathcal{G}(f_k). \quad (5.31)$$

PROOF

There are several different ways to prove the frequency convolution theorem. But we prove this by showing that the inverse DFT of the frequency convolution $\mathcal{F}(f_k) * \mathcal{G}(f_k)$ scaled by $1/N$ is equal to the multiplication in the time domain $f(t_n)g(t_n)$.

Following the definition of an inverse DFT (5.17):

$$\begin{aligned} & \frac{1}{N} \sum_{k=0}^{N-1} \left\{ \frac{1}{N} \mathcal{F}(f_k) * \mathcal{G}(f_k) \right\} \exp(2\pi i k n / N) \\ &= \frac{1}{N} \frac{1}{N} \sum_{k=0}^{N-1} \left\{ \sum_{j=0}^{N-1} \mathcal{F}\left(\frac{j}{N\Delta t}\right) \mathcal{G}\left(\frac{k-j}{N\Delta t}\right) \right\} \exp(2\pi i k n / N) \\ &= \frac{1}{N} \sum_{j=0}^{N-1} \mathcal{F}\left(\frac{j}{N\Delta t}\right) \left\{ \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{G}\left(\frac{k-j}{N\Delta t}\right) \exp(2\pi i k n / N) \right\}. \end{aligned}$$

Using the frequency shifting (modulation) property of DFT of the equation (5.25):

$$\begin{aligned} \frac{1}{N} \sum_{k=0}^{N-1} \left\{ \frac{1}{N} \mathcal{F}(f_k) * \mathcal{G}(f_k) \right\} \exp(2\pi i k n / N) &= \frac{1}{N} \sum_{j=0}^{N-1} \mathcal{F}\left(\frac{j}{N\Delta t}\right) e^{2\pi i j n / N} g(t_n) \\ &= g(t_n) \left[\frac{1}{N} \sum_{j=0}^{N-1} \mathcal{F}\left(\frac{j}{N\Delta t}\right) e^{2\pi i j n / N} \right] \\ &= f(t_n)g(t_n) \end{aligned}$$

□

[5.4.8] Parseval's Relation

Suppose that the time domain sequence $\{f(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ and $\{g(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ have discrete Fourier transforms $\{\mathcal{F}(f_k) : k = 0, 1, \dots, N-1\}$ and $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16). We assume $f(t_n \equiv n\Delta t)$, $g(t_n \equiv n\Delta t)$, $\mathcal{F}(f_k)$, and $\mathcal{G}(f_k)$ are all periodic functions.

Let $\bar{g}(t_n \equiv n\Delta t)$ be a complex conjugate of $g(t_n \equiv n\Delta t)$ and $\bar{\mathcal{G}}(f_k)$ be a complex conjugate of $\mathcal{G}(f_k)$:

$$\begin{aligned} |f(t_n \equiv n\Delta t)|^2 &= f(t_n \equiv n\Delta t) \bar{g}(t_n \equiv n\Delta t), \\ |\mathcal{F}(f_k)|^2 &= \mathcal{F}(f_k) \bar{\mathcal{G}}(f_k). \end{aligned}$$

Parseval's relation states that the power of a signal function $f(t_n \equiv n\Delta t)$ is same whether it is computed in signal (time) space t or transform (frequency) space f after taken care of the weight $1/N$:

$$\begin{aligned}\sum_{n=0}^{N-1} |f(t_n \equiv n\Delta t)|^2 &= \frac{1}{N} \sum_{k=0}^{N-1} |\mathcal{F}(f_k)|^2, \\ \sum_{n=0}^{N-1} f(t_n \equiv n\Delta t) \bar{g}(t_n \equiv n\Delta t) &= \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{F}(f_k) \bar{\mathcal{G}}(f_k).\end{aligned}\quad (5.32)$$

[5.4.9] Summary of DFT Properties

Suppose that the time domain sequence $\{f(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ and $\{g(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$ have discrete Fourier transforms $\{\mathcal{F}(f_k) : k = 0, 1, \dots, N-1\}$ and $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$ defined by the equation (5.16). We assume $f(t_n \equiv n\Delta t)$, $g(t_n \equiv n\Delta t)$, $\mathcal{F}(f_k)$, and $\mathcal{G}(f_k)$ are all periodic functions.

Table 5.1: Summary of DFT Properties

Property	$\{g(t_n \equiv n\Delta t) : n = 0, 1, \dots, N-1\}$	DFT $\{\mathcal{G}(f_k) : k = 0, 1, \dots, N-1\}$
Linearity	$af(t_n) + bg(t_n)$	$a\mathcal{F}(f_k) + b\mathcal{G}(f_k)$
Even Function	$g(t_n)$ is even	$\mathcal{G}(f_k) \in \mathbb{R}, \text{even}$
Odd Function	$g(t_n)$ is odd	$\mathcal{G}(f_k) \in \mathbb{I}, \text{odd}$
Symmetry	$\frac{1}{N} \mathcal{G}(t_n)$	$g(-f_k)$
Time Shifting	$g(t_n - t_0 \Delta t)$	$\mathcal{G}(f_k) e^{-2\pi i k t_0 / N}$
Convolution	$f(n\Delta t) * g(n\Delta t)$	$\mathcal{F}(f_k) \mathcal{G}(f_k)$
Multiplication	$f(t_n) g(t_n)$	$\frac{1}{N} \mathcal{F}(f_k) * \mathcal{G}(f_k)$
Modulation (Frequency Shifting)	$e^{2\pi i n f_0 / N} g(t_n)$	$\mathcal{G}(f_k - \frac{f_0}{N \Delta t})$

[6] Lévy Processes

In this sequel theorems and propositions are presented without proofs most of the time. This is obviously because we don't need them for our purpose (it is finance) and always inquisitive readers can read Cont and Tankov (2004) and Sato (1999) for more rigorous treatment of the subjects dealt in this section. We recommend for those readers without training in the field of set theory and measure theory to first read Appendix 1 and 2 for the basic knowledge.

[6.1] Definition of Lévy Process

A right-continuous with left limits (cadlag) or an adapted (non-anticipating) stochastic process $\{X_t; 0 \leq t < \infty\}$ on a space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in \mathbb{R} is said to be a Lévy process if it satisfies the following conditions:

- (1) Its increments are independent of the past: $X_u - X_t$ is independent of the filtration \mathcal{F}_t with $0 \leq t < u < \infty$. i.e. $\mathbb{P}(X_u - X_t | \mathcal{F}_t) = \mathbb{P}(X_u - X_t)$.
- (2) Its increments are stationary: $X_{t+h} - X_t$ has the same distribution as X_h . In other words, the distribution of increments does not depend on t (i.e. temporal homogeneity).
- (3) $X_0 = 0$ a.s.
- (4) X_t is continuous in probability: $\forall \varepsilon > 0, \lim_{h \rightarrow 0} \mathbb{P}(|X_{t+h} - X_t| \geq \varepsilon) = 0$.

Processes which satisfy conditions (1), (2), and (3) are said to be processes with stationary independent increments. Condition (4) is satisfied when conditions (1), (2), and (3) are satisfied. Condition (4) does not imply the continuous sample paths. Actually the very opposite is true meaning that most Lévy processes have discontinuous sample paths (i.e. except for one). Condition (4) means that if we are at time t , the probability of a jump at time t is zero because there is no uncertainty about the present. Jumps occur at random times.

From our knowledge, a right-continuous with left limits (cadlag in French) stochastic process, a non-anticipating stochastic process, and an adapted stochastic process define an identical process. A stochastic process $\{X_t; 0 \leq t \leq T\}$ is said to be non-anticipating with respect to the filtration $\{\mathcal{F}_t; 0 \leq t \leq T\}$ or \mathcal{F}_t -adapted if the value of X_t is revealed at time t for each $t \in [0, T]$. In other words, $\{X_t; 0 \leq t \leq T\}$ is said to be non-anticipating if it satisfies for $t \in [0, T]$:

- (1) Left limit of the process $X(t-) \equiv \lim_{s \rightarrow t, s < t} X(s)$ exists.
- (2) Right limit of the process $X(t+) \equiv \lim_{s \rightarrow t, s > t} X(s)$ exists.
- (3) $X(t) = X(t+)$.

Any continuous function is non-anticipating but non-anticipating functions allow discontinuities. Suppose t is a discontinuity point, the jump of X at t is:

$$\Delta X(t) = X(t) - X(t-).$$

A non-anticipating process $\{X_t; 0 \leq t \leq T\}$ can have a finite number of large jumps and countable number (possibly infinite) of small jumps.

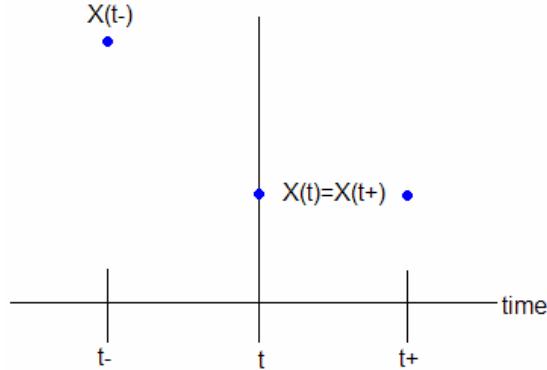


Figure 6.1: Illustration of Non-Anticipating Stochastic Process

Suppose that $X(t)$ is a stock price right now, $X(t-)$ a stock price 1 second ago, and $X(t+)$ a stock price 1 second from now. A stock price process X should be modeled as a non-anticipating process because at time $t-$ we cannot predict $X(t)$ (i.e. it is a future value), but at time $t+$ we already know $X(t)$ (i.e. it is a past value).

We saw the definition of a Lévy process. Next let's discuss infinite divisibility of a distribution. It turns out that we cannot separate Lévy processes from infinitely divisible distributions because Lévy processes are generated by infinitely divisible distributions.

A random variable Y is said to be divisible if it can be represented as the sum of two independent random variables with identical distributions:

$$Y = Y_1 + Y_2.$$

A random variable Y is said to be infinitely divisible if it can be represented as the sum of n independent random variables with identical distributions for any integer $n \geq 2$:

$$Y = Y_1 + Y_2 + \dots + Y_n.$$

Let $\phi(\omega)$ be the characteristic function of the distribution of the infinitely divisible random variable Y and $\phi_n(\omega)$ be the characteristic function of the common distribution of the n summands. Then, the relationship between $\phi(\omega)$ and $\phi_n(\omega)$ is:

$$\phi(\omega) = (\phi_n(\omega))^n \text{ and } \phi_n(\omega) = (\phi(\omega))^{1/n}. \quad (6.1)$$

Examples of infinitely divisible distributions are: the normal distribution, the gamma distribution, α -stable distributions, and the Poisson distribution.

If Y is a normal random variable, i.e. $Y \sim N(\mu, \sigma^2)$, its characteristic function is:

$$\phi(\omega) = \int_{-\infty}^{\infty} e^{i\omega Y} \mathbb{P}(Y) dY = \exp(i\mu\omega - \frac{\sigma^2\omega^2}{2}).$$

Then, the characteristic function for the identically distributed n summands of Y can be computed as using the above relation:

$$\begin{aligned} \phi_n(\omega) &= \{\phi(\omega)\}^{1/n} = \left\{ \exp(i\mu\omega - \frac{\sigma^2\omega^2}{2}) \right\}^{1/n} \\ &= \exp \left\{ i \left(\frac{\mu}{n} \right) \omega - \frac{(\sigma^2/n)\omega^2}{2} \right\}. \end{aligned}$$

Thus, the identically distributed n summands of $Y \sim N(\mu, \sigma^2)$ are also normally distributed with the mean μ/n and the variance σ^2/n (because a characteristic function uniquely determines a probability distribution):

$$Y = \sum_{k=0}^{n-1} Y_k, \quad Y_k \sim i.i.d. N(\mu/n, \sigma^2/n).$$

Another example is a Poisson case. If Z is a Poisson random variable (read Appendix 6 for the definition), i.e. $Z \sim \text{Poisson}(\lambda)$, its characteristic function is:

$$\phi(\omega) = \sum_{z=0}^{\infty} \left(\frac{e^{-\lambda} \lambda^z}{z!} \right) e^{i\omega z} = \exp[\lambda(e^{i\omega} - 1)].$$

Its identically distributed n summands follow Poisson law with the parameter λ/n since their characteristic functions take the form:

$$\phi_n(\omega) = \{\phi(\omega)\}^{1/n} = \left[\exp \left\{ \lambda(e^{i\omega} - 1) \right\} \right]^{1/n} = \exp \left\{ \frac{\lambda}{n} (e^{i\omega} - 1) \right\}.$$

A Lévy process $\{X_t; t \geq 0\}$ possesses this infinite divisibility, i.e. for every t increments of a Lévy process $X_{t+h} - X_t$ has an infinitely divisible law. Conversely, if \mathbb{P} is an infinitely divisible distribution, then there exists a Lévy process $\{X_t; t \geq 0\}$ where the

distribution of increments $X_{t+h} - X_t$ is governed by \mathbb{P} . This lemma is extremely important because it means that infinitely divisible distributions (normal, gamma, α -stable, and Poisson distribution) can generate Lévy processes.

$$\text{Lévy process } \{X_t; t \geq 0\} \iff \mathbb{P}(X_{t+h} - X_t) \in \text{Infinitely divisible distributions}$$

[6.2] Standard Brownian Motion Process: The Only Continuous Lévy Process Generated by A Normal Distribution

Some people misinterpret that Lévy processes are discontinuous (i.e. jump) processes. This is not true in a strict sense. It is true to state that most Lévy processes are discontinuous processes. But you can carefully go through the definition of Lévy processes again and you'll notice that no conditions require Lévy processes to possess discontinuous sample paths. A Lévy process can have a continuous sample path. The only example of a continuous Lévy process is a standard Brownian motion process.

A standard Brownian motion process $\{B_t; t \geq 0\}$ is a Lévy process on \mathbb{R} defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that:

- (1) $B_t \sim \text{Normal}(0, t)$.
- (2) There is $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$, i.e. $X_t(\omega)$ is continuous in t for every $\omega \in \Omega_0$.

[6.3] Poisson Process and Compound Poisson Process: Finite Activity Lévy Processes Generated by Poisson Distribution

Appendix 5 gives the in-depth background treatment of the definition and characteristics of an exponential random variable, a Poisson process.

Let $(\tau_i)_{i \geq 1}$ be a sequence of independent exponential random variables with parameter λ and $T_n = \sum_{i=1}^n \tau_i$. A Poisson process with intensity λ is:

$$N_t = \sum_{n \geq 1} 1_{t \geq T_n} .$$

The sample function N_t counts the number of random times (T_n) at which a specified event occurs during the time period between 0 and t where $(T_n - T_{n-1})_{n \geq 1}$ is an *i.i.d.* sequence of exponential variables. Therefore, each possible N_t is represented as a non-decreasing piecewise constant function.

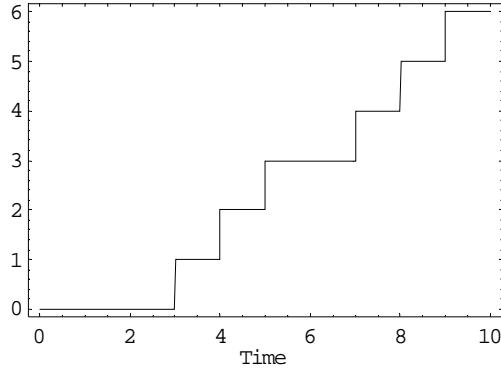


Figure 6.2: Sample Path of a Poisson Process with $\lambda = 1$ and $t = 10$. In this sample, the process has 6 jumps.

A Poisson process (N_t) is a Lévy process because it possesses the following properties:

- (1) Its increments are independent of the past: $N_u - N_t$ is independent of the filtration \mathcal{F}_t with $0 \leq t < u < \infty$. i.e. $\mathbb{P}(N_u - N_t | \mathcal{F}_t) = \mathbb{P}(N_u - N_t)$.
- (2) Its increments are stationary: $N_{t+h} - N_t$ has the same distribution as N_h . In other words, the distribution of increments does not depend on t (i.e. temporal homogeneity).
- (3) $N_0 = 0$ a.s.
- (4) N_t is continuous in probability: $\forall \varepsilon > 0, \lim_{h \rightarrow 0} \mathbb{P}(|N_{t+h} - N_t| \geq \varepsilon) = 0$.
- (5) A sample path of a Poisson process (N_t) is non-anticipating.
- (6) $N_t < \infty$ a.s. for any $t > 0$. A Poisson process has finite number of jumps.

Next, let's take a look at more general version of Poisson process called a compound Poisson process. A compound Poisson process $\{X_t; t \geq 0\}$ with intensity λ is defined as:

$$X_t = \sum_{i=1}^{N_t} Y_i, \quad (6.2)$$

where Y_i are *i.i.d.* jump sizes with the probability density function f . We assume (N_t) and $(Y_i)_{i \geq 1}$ are independent.

Since a compound Poisson process reduces to a Poisson process if $Y_i = 1$ (a Poisson process can be considered as a compound Poisson process with a constant jump size $Y_i = 1$), a compound Poisson process is a Lévy process with properties:

- (1) Its sample paths are piece-wise constant and non-anticipating functions.
- (2) The jump sizes $(Y_i)_{i \geq 1}$ are independent and identically distributed with the probability density function f .

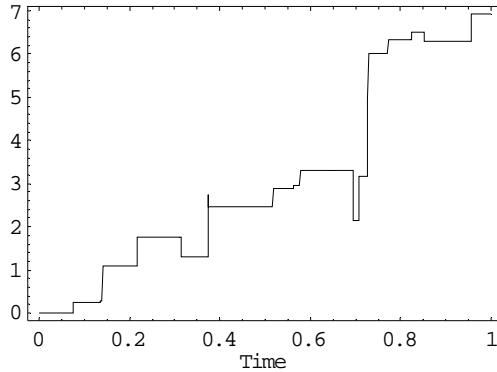


Figure 6.3: Sample Path of a Compound Poisson Process with $\lambda = 20$ and $t = 1$. Jump sizes are drawn from standard normal distribution. In this sample, the process has 19 jumps.

Note that a stochastic process $\{X_t; t \geq 0\}$ is a compound Poisson process, if and only if it is a Lévy process with piecewise constant functions. General Lévy process can be well approximated by a compound Poisson process because any cadlag functions can be approximated by a step function.

The characteristic function of a Poisson process can be obtained by a series expansion:

$$\phi(\omega) \equiv \sum_{z=0}^{\infty} \left\{ \frac{e^{-\lambda t} (\lambda t)^z}{z!} \right\} e^{i\omega z} = \exp \left\{ \lambda t (e^{i\omega} - 1) \right\}. \quad (6.3)$$

To obtain the characteristic function of a compound Poisson process $\{X_t; t \geq 0\}$ with intensity λ and the jump size distribution f , we condition the expectation on a Poisson process N_t and let f^* be the characteristic function of f (Cont and Tankov (2004)):

$$\begin{aligned} E[e^{i\omega X_t}] &= E[E[e^{i\omega X_t}] | N_t] = E[f^*(\omega)^{N_t}] = \sum_{n=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^n f^*(\omega)^n}{n!} \\ &= \exp[\lambda t (f^*(\omega) - 1)] = \exp\{t \lambda \int_{-\infty}^{\infty} (e^{i\omega x} - 1) f(dx)\}, \quad \forall \omega \in \mathbb{R}. \end{aligned} \quad (6.4)$$

We can interpret the characteristic function of a compound Poisson process as a superposition of independent Poisson processes with random jump sizes from distribution f .

Defining a new measure $\ell(dx) = \lambda f(dx)$ which is called a Lévy measure of the Lévy process $\{X_t; 0 \leq t\}$, the above formula can be rewritten as (this is a special case of Lévy-Khintchin representation which will be discussed later):

$$E[e^{i\omega X_t}] = \exp\left\{t \int_{-\infty}^{\infty} (e^{i\omega x} - 1) \ell(dx)\right\}, \quad \forall \omega \in \mathbb{R}. \quad (6.5)$$

The Lévy measure $\ell(dx)$ represents the arrival rate (i.e. total intensity) of jumps of sizes $[x, x+dx]$. In other words, we can interpret the Lévy measure $\ell(dx)$ of a compound Poisson process as the measure of the average number of jumps per unit of time. There are a couple of extremely important points which should be mentioned about this Lévy measure $\ell(dx)$. Lévy measure is a positive measure on \mathbb{R} , but it is not a probability measure since its total mass λ (in the compound Poisson case) does not have to equal 1:

$$\int \ell(dx) = \lambda \in \mathbb{R}^+.$$

Also, Lévy density $\ell(dx)$ of a Lévy process $\{X_t; t \geq 0\}$ is completely different from a probability density of Lévy process $\{X_t; t \geq 0\}$ which will be denoted as $\mathbb{P}(x_t)$ or $vg(x_t)$ in the VG model, for example. We emphasize not to confuse these two densities although they are related (this relation is briefly discussed in Cont and Tankov (2004)).

A Poisson process and a compound Poisson process (i.e. a piecewise constant Lévy process) are called finite activity Lévy processes since their Lévy measures $\ell(dx)$ are finite (i.e. the average number of jumps per unit time is finite):

$$\int_{-\infty}^{\infty} \ell(dx) < \infty.$$

[6.4] Lévy-Itô Decomposition and Infinite Activity Lévy Process

There is a very useful theorem called Lévy-Itô decomposition which basically states that any Lévy process can be represented as the sum of a Brownian motion with drift process (which is a continuous process) and a discontinuous jump process (i.e. compensated (centered) sum of independent jumps). We assume that these two components are independent.

Let $\{X_t; t \geq 0\}$ be a Lévy process on \mathbb{R} with Lévy measure ℓ which is the measure of expected number of jumps per unit of time whose sizes belong to any positive measurable set A . Lévy-Itô decomposition states:

(1) Lévy measure ℓ satisfies:

$$\int_{-\infty}^{\infty} 1(|x| \geq 1) \ell(dx) < \infty \text{ and } \int_{\mathbb{R}-\{0\}} x^2 1(|x| < 1) \ell(dx) < \infty.$$

The first condition means that Lévy processes X have finite number of large jumps (large jumps are defined as jumps with absolute values greater than 1).

Second condition states Lévy measure must be square-integrable around the origin.

- (2) There exists a drift b and a Brownian motion process with a diffusion coefficient σ , $(\sigma B_t)_{t \geq 0}$, such that:

$$X_t = bt + \sigma B_t + X_t^L + \lim_{S \downarrow 0} \widetilde{X}_t^S.$$

Following Lévy-Itô decomposition, the distribution of every Lévy process X_t is uniquely determined by its characteristic triplet (Lévy triplet) (b, σ, ℓ) of the process. Condition (2) means that any Lévy process X_t can be decomposed into a continuous (i.e. diffusion) part $bt + \sigma B_t$ and a discontinuous part $X_t^L + \lim_{S \downarrow 0} \widetilde{X}_t^S$. We arbitrarily define large jumps ΔX^L as those with absolute size greater than 1 (i.e. this does not have to be 1):

$$\Delta X^L \equiv |\Delta X| \geq 1,$$

and small jumps as those with absolute size between S and 1:

$$\Delta X^S \equiv S \leq |\Delta X| < 1.$$

X_t^L is the sum of finite number of large jumps during the time interval of $0 \leq r \leq t$:

$$X_t^L = \sum_{0 \leq r \leq t} \Delta X_r^L.$$

X_t^S is the sum of (possibly infinite in the limit $S \downarrow 0$) number of small jumps during the time interval of $0 \leq r \leq t$:

$$X_t^S = \sum_{0 \leq r \leq t} \Delta X_r^S.$$

In the limit case of $S \downarrow 0$, the process can have infinitely many small jumps, therefore X_t^S may not converge. In other words, Lévy measure ℓ can have a singularity at 0 (i.e. infinite arrival rate of small jumps at zero):

$$\int_{-\infty}^{\infty} \ell(dx) = \infty.$$

This type of Lévy process is called an infinite activity Lévy process. Convergence in the condition (2) can be obtained by replacing the jump integral by its compensated version \widetilde{X}_t^S . For more rigorous treatment, we recommend Cont and Tankov (2004).

[6.5] Lévy-Khintchin Representation

Lévy-Itô decomposition states that the characteristic function of a Lévy process can be expressed in terms of its characteristic triplet. Let $\{X_t; t \geq 0\}$ be a finite variation Lévy process on \mathbb{R} and (b, σ, ℓ) be its triplet. Then for any $\omega \in \mathbb{R}$, a characteristic function $\phi(\omega)$ and a characteristic exponent $\psi(\omega)$ of a finite variation Lévy process $\{X_t; t \geq 0\}$ can be expressed as:

$$\begin{aligned}\phi(\omega) &\equiv E[\exp(i\omega X_t)] = \exp(t\psi(\omega)), \\ \psi(\omega) &= ib\omega - \frac{\sigma^2\omega^2}{2} + \int_{-\infty}^{\infty} \{\exp(i\omega x) - 1\} \ell(dx).\end{aligned}\quad (6.6)$$

[6.6] Stable Processes

Consider a real-valued random variable X . Let $\phi_X(\omega)$ be its characteristic function. X is said to have stable distribution if for any $a > 0$, there exist $b(a) > 0$ and $c(a) \in \mathbb{R}$ such that:

$$\phi_X(\omega)^a = \phi_X(\omega b(a)) \exp(ic\omega), \quad \forall \omega \in \mathbb{R}. \quad (6.7)$$

X is said to have strictly stable distribution if for any $a > 0$, there exist $b(a) > 0$ and $c(a) \in \mathbb{R}$ such that:

$$\phi_X(\omega)^a = \phi_X(\omega b(a)), \quad \forall \omega \in \mathbb{R}. \quad (6.8)$$

At first look these definitions seem complicated, but what they mean is very simple. For example, if Y is a normal random variable, i.e. $Y \sim N(\mu, \sigma^2)$, its characteristic function is $\phi_Y(\omega) = \exp(i\mu\omega - \frac{\sigma^2\omega^2}{2})$. A normal random variable has stable distribution since it satisfies the equation (6.7) with $b(a) = a^{1/2}$ and $c = (-\sqrt{a} + a)\mu$:

$$\begin{aligned}\phi(\omega)^a &= \{\exp(i\mu\omega - \frac{\sigma^2\omega^2}{2})\}^a = \exp(i\mu\omega a - \frac{\sigma^2\omega^2 a}{2}) \\ \phi(\omega)^a &= \exp(i\mu\omega a - \frac{\sigma^2\omega^2 a}{2}) \exp(i\mu\omega\sqrt{a} - i\mu\omega\sqrt{a}) \\ \phi(\omega)^a &= \exp\{i\mu\omega\sqrt{a} - \sigma^2(\omega\sqrt{a})^2/2\} \exp(i\mu\omega a - i\mu\omega\sqrt{a}) \\ \phi(\omega)^a &= \phi_X(\omega\sqrt{a}) \exp\{i(a - \sqrt{a})\mu\omega\}.\end{aligned}$$

A normal random variable with $\mu = 0$ has strictly stable distribution since it satisfies the equation (6.8):

$$\begin{aligned}\phi(\omega)^a &= \{\exp(-\sigma^2 \omega^2 / 2)\}^a = \exp\left(-\frac{\sigma^2 \omega^2 a}{2}\right) = \exp\left\{-\frac{\sigma^2 (\omega \sqrt{a})^2}{2}\right\} \\ \phi(\omega)^a &= \phi_X(\omega \sqrt{a}).\end{aligned}$$

Let $\{X_t; t \geq 0\}$ be a stochastic process on \mathbb{R} . A stochastic process is said to be self-similar if for any $a > 0$, there exists $b > 0$ such that:

$$\{X_{at}; t \geq 0\} \stackrel{d}{=} \{bX_t; t \geq 0\}, \quad (6.9)$$

which we read as that two processes $\{X_{at}; t \geq 0\}$ and $\{bX_t; t \geq 0\}$ are identical in law.

It is said to be broad-sense self-similar if for any $a > 0$, there exists $b > 0$ and a function $c(t)$ such that:

$$\{X_{at}; t \geq 0\} \stackrel{d}{=} \{bX_t + c(t); t \geq 0\}. \quad (6.10)$$

A standard Brownian motion process $\{B_t; t \geq 0\}$ (which is a Lévy process), $B_t \sim \text{Normal}(0, t)$, possesses this self-similarity property since it satisfies the equation (6.9) for $\forall a > 0$:

$$\begin{aligned}\{B_{at}; t \geq 0\} &\sim \text{Normal}(0, at) \\ \{B_{at}; t \geq 0\} &\stackrel{d}{=} \{\sqrt{a}B_t; t \geq 0\}.\end{aligned}$$

A Brownian motion with drift $B_t + \gamma t$ is broad-sense self-similar since it satisfies the equation (6.10) for $\forall a > 0$:

$$\begin{aligned}\{B_{at} + \gamma at; t \geq 0\} &\sim \text{Normal}(\gamma at, at), \\ \{B_{at} + \gamma at; t \geq 0\} &\stackrel{d}{=} \{\sqrt{a}B_t + \gamma at; t \geq 0\}.\end{aligned}$$

Self-similarity means that any change of time scale for the self-similar process has the same effect as some change of spatial scale (also called a scale-invariance property).

Let $\{X_t; t \geq 0\}$ be a Lévy process. A Lévy process is said to be self-similar if for any $a > 0$, there exists $b(a) > 0$ such that:

$$\{X_{at}; t \geq 0\} \stackrel{d}{=} \{b(a)X_t; t \geq 0\}.$$

Following the Lévy-Khintchin representation, the characteristic function of a Lévy process can be expressed as $\phi(\omega) = E[\exp(i\omega X_t)] = \exp(t\psi(\omega))$, $\omega \in \mathbb{R}$. Thus, in terms of

characteristic function a Lévy process is said to be self-similar if for any $a > 0$, there exists $b(a) > 0$ such that:

$$\phi_X(\omega)^a = \phi_X(\omega b(a)),$$

which is the definition of a strictly stable distribution of the equation (6.8). This means that if a Lévy process is self-similar, then it is strictly stable. Formally, a Lévy process $\{X_t : t \geq 0\}$ on \mathbb{R} is self-similar (broad-sense self-similar), if and only if it is strictly stable (stable).

For every stable distribution, there exists a constant $0 < \alpha \leq 2$ such that $b(a) = a^{1/\alpha}$ in the equation (6.7). Let $\phi(\omega)$ be the characteristic function of a stable real-valued random variable X . Then if for any $a > 0$, there exists $0 < \alpha \leq 2$ and $c(a) \in \mathbb{R}$ such that:

$$\phi_X(\omega)^a = \phi_X(\omega a^{1/\alpha}) \exp(ic\omega), \quad \forall \omega \in \mathbb{R}.$$

α is called an index of stability and stable distributions with α are called α -stable distributions. Normal distributions are the only 2-stable distributions.

A strictly α -stable Lévy process (a self-similar Lévy process which is strictly stable with index of stability α) satisfies for any $a > 0$:

$$\{X_{at} : t \geq 0\} \stackrel{d}{=} \{a^{1/\alpha} X_t : t \geq 0\}. \quad (6.11)$$

Brownian motion process is an example of a strictly 2-stable Lévy process.

An α -stable Lévy process satisfies for any $a > 0$:

$$\exists c \in \mathbb{R} \text{ such that } \{X_{at} : t \geq 0\} \stackrel{d}{=} \{a^{1/\alpha} X_t + ct : t \geq 0\}.$$

A real-valued random variable is α -stable with $0 < \alpha < 2$ if and only if it is infinitely divisible with characteristic triplet $(b, \sigma = 0, \ell)$ and its Lévy measure is of the form¹:

$$\ell(x) = \frac{A}{x^{\alpha+1}} 1_{x>0} + \frac{B}{|x|^{\alpha+1}} 1_{x<0}, \quad (6.12)$$

where A and B are some positive constants.

¹ For more details, Samorodnitsky, G. and Taqqu, M., 1994, Stable Non-Gaussian Random Processes, Chapman & Hall: New York.

[7] Black-Scholes Model as an Exponential Lévy Model

Even if you were very experienced readers, we recommend you take a look at section 7.6. Other sections may be skipped.

[7.1] Standard Brownian Motion Process: A Lévy Process Generated by a Normal Distribution

As we saw in section 6.2, a standard Brownian motion process $\{B_t; t \geq 0\}$ is a stochastic process on \mathbb{R} defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that:

- (1) $B_{t+u} - B_t \forall u > 0 \sim \text{Normal}(0, u)$.
- (2) There is $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$, i.e. $X_t(\omega)$ is continuous in t for every $\omega \in \Omega_0$.
- (3) $B_{t+u} - B_t \forall u > 0$ are stationary and independent.
- (4) The process $\{B_t; t \geq 0\}$ begins with 0, i.e. $B_0 = 0$.

Stationary increments of the condition (3) mean that the distributions of increments $B_{t+u} - B_t$ do not depend on the time t , but they depend on the time-distance u of two observations (i.e. interval of time). For example, if you model a log stock price $\ln S_t$ as a Brownian motion (with drift) process, the distribution of increment in year 2004 for the next one year $\ln S_{2004+1} - \ln S_{2004}$ is the same as that in year 2050, $\ln S_{2050+1} - \ln S_{2050}$:

$$\ln S_{2004+1} - \ln S_{2004} \stackrel{d}{=} \ln S_{2050+1} - \ln S_{2050}.$$

The conditional probability of the event A given B is assuming $\mathbb{P}(B) > 0$:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

If A and B are independent events:

$$\mathbb{P}(A|B) = \mathbb{P}(A).$$

Independent increments of Brownian motion process of the condition (3) mean that when modeling a log stock price $\ln S_t$ as a Brownian motion (with drift) process, the probability distribution of a log stock price in year 2005 is not affected by whatever happens in year 2004 in the stock price (i.e. such as stock price crush):

$$\mathbb{P}(\ln S_{2005+1} - \ln S_{2005} | \ln S_{2004+1} - \ln S_{2004}) = \mathbb{P}(\ln S_{2005+1} - \ln S_{2005}).$$

These are the two main restrictions imposed by modeling a log stock price process using a Lévy process.

[7.2] Black-Scholes' Distributional Assumptions on a Stock Price

In traditional finance literature almost every financial asset price (stocks, currencies, interest rates) is assumed to follow some variations of Brownian motion with drift process. BS (Black-Scholes) models a stock price increment process in an infinitesimal time interval dt as a log-normal random walk process:

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \quad (7.1)$$

where the drift is μS_t , which is a constant expected return on a stock μ proportional to a stock price S_t and the volatility is σS_t , which is a constant stock price volatility σ proportional to a stock price S_t . The reason why the process (7.1) is called a log-normal random walk process will be explained very soon. Alternatively, we can state that BS models a percentage change in a stock price process in an infinitesimal time interval dt as a Brownian motion with drift process:

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu dt + \sigma dB_t, \\ \mathbb{P}\left(\frac{dS_t}{S_t}\right) &= \frac{1}{\sqrt{2\pi\sigma^2 dt}} \exp\left[-\frac{(dS_t/S_t - \mu dt)^2}{2\sigma^2 dt}\right]. \end{aligned} \quad (7.2)$$

Let S be a random variable whose dynamics is given by a Ito process:

$$dS = a(S, t)dt + b(S, t)dB,$$

and V be a function dependent on a random variable S and time t . The dynamics of $V(S, t)$ is given by an Ito formula:

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2} b^2 \frac{\partial^2 V}{\partial S^2} dt, \quad (7.3)$$

or in terms of a standard Brownian motion process B :

$$\begin{aligned} dV &= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} (adt + bdB) + \frac{1}{2} b^2 \frac{\partial^2 V}{\partial S^2} dt, \\ dV &= \left(\frac{\partial V}{\partial t} + a \frac{\partial V}{\partial S} + \frac{1}{2} b^2 \frac{\partial^2 V}{\partial S^2} \right) dt + b \frac{\partial V}{\partial S} dB. \end{aligned} \quad (7.4)$$

Dynamics of a log stock price process $\ln S_t$ can be obtained by applying (7.4) to (7.1) as:

$$d \ln S_t = \left(\frac{\partial \ln S_t}{\partial t} + \mu S_t \frac{\partial \ln S_t}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 \ln S_t}{\partial S_t^2} \right) dt + \sigma S_t \frac{\partial \ln S_t}{\partial S_t} dB_t.$$

Substituting $\frac{\partial \ln S_t}{\partial t} = 0$, $\frac{\partial \ln S_t}{\partial S_t} = \frac{1}{S_t}$, and $\frac{\partial^2 \ln S_t}{\partial S_t^2} = -\frac{1}{S_t^2}$ yields:

$$d \ln S_t = \left(\mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dB_t, \quad (7.5)$$

or:

$$\begin{aligned} \ln S_t - \ln S_0 &= \left(\mu - \frac{1}{2} \sigma^2 \right) (t - 0) + \sigma (B_t - B_0) \\ \ln S_t &= \ln S_0 + \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t. \end{aligned} \quad (7.6)$$

The equation (7.6) means that BS models a log stock price $\ln S_t$ as a Brownian motion with drift process whose probability density is given by a normal density:

$$\mathbb{P}(\ln S_t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left[-\frac{\left\{ \ln S_t - \left(\ln S_0 + (\mu - \frac{1}{2} \sigma^2)t \right) \right\}^2}{2\sigma^2 t}\right]. \quad (7.7)$$

Alternatively, the equation (7.6) means that BS models a log return $\ln(S_t / S_0)$ as a Brownian motion with drift process whose probability density is given by a normal density:

$$\begin{aligned} \ln(S_t / S_0) &= \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t, \\ \mathbb{P}(\ln(S_t / S_0)) &= \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left[-\frac{\left\{ \ln(S_t / S_0) - \left(\mu - \frac{1}{2} \sigma^2 \right) t \right\}^2}{2\sigma^2 t}\right]. \end{aligned} \quad (7.8)$$

An example of BS normal log return $\ln(S_t / S_0)$ density of (7.8) is illustrated in Figure 7.1. Of course, BS log return density is symmetric (i.e. zero skewness) and have zero excess kurtosis because it is a normal density.

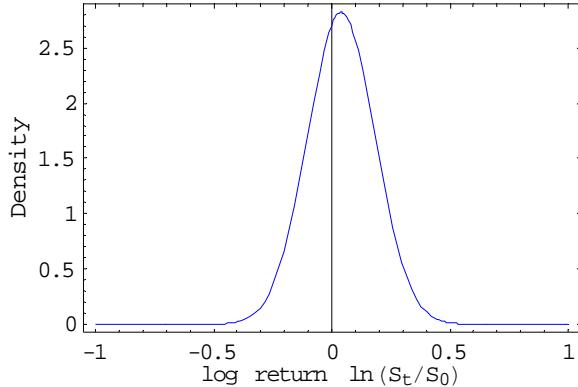


Figure 7.1: An Example of BS normal log return $\ln(S_t/S_0)$ Density. Parameters and variables fixed are $\mu = 0.1$, $\sigma = 0.2$, and $t = 0.5$.

Let y be a random variable. If the log of y is normally distributed with mean a and variance b^2 such that $\ln y \sim N(a, b^2)$, then y is a log-normal random variable whose density is a two parameter family (a, b) :

$$y \sim \text{Lognormal}(e^{\frac{a+b^2}{2}}, e^{2a+b^2}(e^{b^2} - 1)),$$

$$\mathbb{P}(y) = \frac{1}{y\sqrt{2\pi b^2}} \exp\left[-\frac{(\ln y - a)^2}{2b^2}\right].$$

From the equation (7.6), we can state that BS models a stock price S_t as a log-normally distributed random variable whose density is given by:

$$\mathbb{P}(S_t) = \frac{1}{S_t \sqrt{2\pi\sigma^2 t}} \exp\left[-\frac{\left\{\ln S_t - \left(\ln S_0 + (\mu - \frac{1}{2}\sigma^2)t\right)\right\}^2}{2\sigma^2 t}\right]. \quad (7.9)$$

Its annualized moments are calculated as:

$$\begin{aligned} \text{Mean}[S_t] &= S_0 e^\mu, \\ \text{Variance}[S_t] &= S_0^2 \left(e^{\sigma^2} - 1\right) e^{2\mu}, \\ \text{Skewness}[S_t] &= \left(e^{\sigma^2} + 2\right) \sqrt{e^{\sigma^2} - 1}, \\ \text{Excess Kurtosis}[S_t] &= -6 + 3e^{2\sigma^2} + 2e^{3\sigma^2} + e^{4\sigma^2}. \end{aligned}$$

An example of BS log-normal stock price density of (7.9) is illustrated in Figure 7.2. Notice that BS log-normal stock price density is positively skewed.

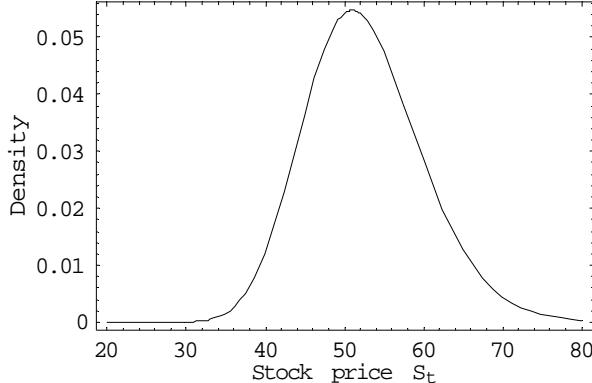


Figure 7.2: An Example of BS Log-Normal Density of a Stock Price. Parameters and variables fixed are $S_0 = 50$, $\mu = 0.1$, $\sigma = 0.2$, and $t = 0.5$.

Table 7.1
Annualized Moments of BS Log-Normal Density of A Stock Price in Figure 2.2

Mean	Standard Deviation	Skewness	Excess Kurtosis
55.2585	11.1632	0.614295	0.678366

From the equation (7.6), we can obtain an integral version equivalent of (7.1):

$$\begin{aligned} \exp[\ln S_t] &= \exp[\ln S_0 + \left(\mu - \frac{1}{2}\sigma^2 \right)t + \sigma B_t] \\ S_t &= \exp[\ln S_0] \exp \left[\left(\mu - \frac{1}{2}\sigma^2 \right)t + \sigma B_t \right] \\ S_t &= S_0 \exp \left[\left(\mu - \frac{1}{2}\sigma^2 \right)t + \sigma B_t \right]. \end{aligned} \quad (7.10)$$

Equation (7.10) means that BS models a stock price process S_t as a geometric process with the growth rate given by a Brownian motion with drift process:

$$\left(\mu - \frac{1}{2}\sigma^2 \right)t + \sigma B_t.$$

[7.3] Traditional Black-Scholes Option Pricing: PDE Approach by Hedging

Consider a portfolio P of the one long option position $V(S, t)$ on the underlying stock S written at time t and a short position of the underlying stock in quantity Δ to derive option pricing function.

$$P_t = V(S_t, t) - \Delta S_t. \quad (7.11)$$

Portfolio value changes in a very short period of time dt by:

$$dP_t = dV(S_t, t) - \Delta dS_t. \quad (7.12)$$

Stock price dynamics is given by a log-normal random walk process of the equation (7.1):

$$dS_t = \mu S_t dt + \sigma S_t dB_t. \quad (7.13)$$

Option price dynamics is given by applying Ito formula of the equation (7.3):

$$dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_t} dS_t + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} dt. \quad (7.14)$$

Now the change in the portfolio value can be expressed as by substituting (7.13) and (7.14) into (7.12):

$$dP_t = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_t} dS_t + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} dt - \Delta dS_t. \quad (7.15)$$

Setting $\Delta = \partial V / \partial S_t$ (i.e. delta hedging) makes the portfolio completely risk-free (i.e. the randomness dS_t has been eliminated) and the portfolio value dynamics of the equation (7.15) simplifies to:

$$dP_t = \left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} \right) dt. \quad (7.16)$$

Since this portfolio is perfectly risk-free, assuming the absence of arbitrage opportunities the portfolio is expected to grow at the risk-free interest rate r :

$$E[dP_t] = r P_t dt. \quad (7.17)$$

After substitution of (7.11) and (7.16) into (7.17) by setting $\Delta = \partial V / \partial S_t$, we obtain:

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} \right) dt = r \left(V - \frac{\partial V}{\partial S_t} S_t \right) dt.$$

After rearrangement, Black-Scholes PDE is obtained:

$$\frac{\partial V(S_t, t)}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V(S_t, t)}{\partial S_t^2} + rS_t \frac{\partial V(S_t, t)}{\partial S_t} - rV(S_t, t) = 0. \quad (7.18)$$

BS PDE is categorized as a linear second-order parabolic PDE. The equation (7.18) is a linear PDE because coefficients of the partial derivatives of $V(S_t, t)$ (i.e. $\sigma^2 S_t^2 / 2$ and rS_t) are not functions of $V(S_t, t)$ itself. The equation (7.18) is a second-order PDE because it involves the second-order partial derivative $\partial^2 V(S_t, t) / \partial S_t^2$. Generally speaking, a PDE of the form:

$$a + b \frac{\partial V}{\partial t} + c \frac{\partial V}{\partial S} + d \frac{\partial^2 V}{\partial S^2} + e \frac{\partial^2 V}{\partial t^2} + g \frac{\partial^2 V}{\partial t \partial S} = 0,$$

is said to be a parabolic type if:

$$g - 4de = 0. \quad (7.19)$$

The equation (7.18) is a parabolic PDE because it has $g = 0$ and $e = 0$ which satisfies the condition (7.19).

BS solves PDE of (7.18) with boundary conditions:

$$\begin{aligned} \max(S_T - K, 0) &\text{ for a plain vanilla call,} \\ \max(K - S_T, 0) &\text{ for a plain vanilla put,} \end{aligned}$$

and obtains closed-form solutions of call and put pricing functions. Exact derivation of closed-form solutions by solving BS PDE is omitted here (i.e. the original BS approach). Instead we will provide the exact derivation by a martingale asset pricing approach (this is much simpler) in the next section.

[7.4] Traditional Black-Scholes Option Pricing: Martingale Pricing Approach

Let $\{B_t; 0 \leq t \leq T\}$ be a standard Brownian motion process on a space $(\Omega, \mathcal{F}, \mathbb{P})$. Under actual probability measure \mathbb{P} , the dynamics of BS stock price process is given by equation (7.9) in the integral form (i.e. which is a geometric Brownian motion process):

$$S_t = S_0 \exp \left[\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right] \text{ under } \mathbb{P}. \quad (7.20)$$

BS model is an example of a complete model because there is only one equivalent martingale risk-neutral measure $\mathbb{Q} \sim \mathbb{P}$ under which the discounted asset price process $\{e^{-rt} S_t; 0 \leq t \leq T\}$ becomes a martingale. BS finds the equivalent martingale risk-neutral

measure $\mathbb{Q}_{BS} \sim \mathbb{P}$ by changing the drift of the Brownian motion process while keeping the volatility parameter σ unchanged:

$$S_t = S_0 \exp \left[\left(r - \frac{1}{2} \sigma^2 \right) t + \sigma B_t^{\mathbb{Q}_{BS}} \right] \text{ under } \mathbb{Q}_{BS}. \quad (7.21)$$

Note that $B_t^{\mathbb{Q}_{BS}}$ is a standard Brownian motion process on $(\Omega, \mathcal{F}, \mathbb{Q}_{BS})$ and the process $\{e^{-rt} S_t; 0 \leq t \leq T\}$ is a martingale under \mathbb{Q}_{BS} . Then, a plain vanilla call option price $C(t, S_t)$ which has a terminal payoff function $\max(S_T - K, 0)$ is calculated as:

$$C(t, S_t) = e^{-r(T-t)} E_t^{\mathbb{Q}_{BS}} [\max(S_T - K, 0)] = e^{-r\tau} E_t^{\mathbb{Q}_{BS}} [\max(S_T - K, 0)]. \quad (7.22)$$

Note that an expectation operator $E_t^{\mathbb{Q}_{BS}} [\cdot]$ is under a probability measure \mathbb{Q}_{BS} and conditional on time t . Let $\mathbb{Q}(S_T)$ (drop the subscript BS for simplicity) be a probability density function of S_T in a risk-neutral world. From the equation (7.9), a terminal stock price S_T is a log-normal random variable with its density of the form:

$$\mathbb{Q}(S_T) = \frac{1}{S_T \sqrt{2\pi\sigma^2\tau}} \exp \left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)\tau \right) \right\}^2}{2\sigma^2\tau} \right]. \quad (7.23)$$

Using (7.23), the expectation term in (7.22) can be rewritten as:

$$\begin{aligned} E_t^{\mathbb{Q}} [\max(S_T - K, 0)] &= \int_K^\infty (S_T - K) \mathbb{Q}(S_T) dS_T + \int_0^K (0) \mathbb{Q}(S_T) dS_T \\ E_t^{\mathbb{Q}} [\max(S_T - K, 0)] &= \int_K^\infty (S_T - K) \mathbb{Q}(S_T) dS_T. \end{aligned}$$

Using this, we can rewrite (7.22) as:

$$C(\tau, S_t) = e^{-r\tau} \int_K^\infty (S_T - K) \mathbb{Q}(S_T) dS_T. \quad (7.24)$$

Since S_T is a log-normal random variable with its density given by the equation (7.23):

$$\ln S_T \sim \text{Normal} \left(m \equiv \ln S_t + (r - \frac{1}{2}\sigma^2)\tau, \sigma^2\tau \right). \quad (7.25)$$

We use a change of variable technique from a log-normal random variable S_T to a standard normal random variable Z through:

$$Z \equiv \frac{\ln S_T - \left\{ \ln S_t + (r - \frac{1}{2}\sigma^2)\tau \right\}}{\sigma\sqrt{\tau}} \equiv \frac{\ln S_T - m}{\sigma\sqrt{\tau}} \sim Normal(0,1), \quad (7.26)$$

with:

$$\mathbb{Z}(Z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{Z^2}{2}\right].$$

From (7.26):

$$S_T = \exp(Z\sigma\sqrt{\tau} + m). \quad (7.27)$$

We can rewrite (7.24) as:

$$C(\tau, S_t) = e^{-r\tau} \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \left(\exp(Z\sigma\sqrt{\tau} + m) - K \right) \mathbb{Z}(Z) dZ,$$

and we express this with more compact form as:

$$C(\tau, S_t) = C_1 - C_2, \quad (7.28)$$

where $C_1 = e^{-r\tau} \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \exp(Z\sigma\sqrt{\tau} + m) \mathbb{Z}(Z) dZ$ and $C_2 = Ke^{-r\tau} \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \mathbb{Z}(Z) dZ$.

Consider C_1 :

$$\begin{aligned} C_1 &= e^{-r\tau} \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \exp(Z\sigma\sqrt{\tau}) \exp(m) \mathbb{Z}(Z) dZ \\ C_1 &= \exp(-r\tau) \exp\left(\ln S_t + (r - \frac{1}{2}\sigma^2)\tau\right) \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \exp(Z\sigma\sqrt{\tau}) \mathbb{Z}(Z) dZ \\ C_1 &= \exp\left(\ln S_t - \frac{1}{2}\sigma^2\tau\right) \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \exp(Z\sigma\sqrt{\tau}) \mathbb{Z}(Z) dZ \\ C_1 &= \exp\left(\ln S_t - \frac{1}{2}\sigma^2\tau\right) \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \exp(Z\sigma\sqrt{\tau}) \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{Z^2}{2}\right] dZ \\ C_1 &= \exp\left(\ln S_t - \frac{1}{2}\sigma^2\tau\right) \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{Z^2 - 2Z\sigma\sqrt{\tau}}{2}\right] dZ \\ C_1 &= \exp\left(\ln S_t - \frac{1}{2}\sigma^2\tau\right) \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(Z - \sigma\sqrt{\tau})^2 - \sigma^2\tau}{2}\right] dZ \end{aligned}$$

$$\begin{aligned}
C_1 &= \exp\left(\ln S_t - \frac{1}{2}\sigma^2\tau\right) \exp\left(\frac{1}{2}\sigma^2\tau\right) \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(Z-\sigma\sqrt{\tau})^2}{2}\right] dZ \\
C_1 &= \exp(\ln S_t) \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(Z-\sigma\sqrt{\tau})^2}{2}\right] dZ \\
C_1 &= S_t \int_{(\ln K-m)/\sigma\sqrt{\tau}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(Z-\sigma\sqrt{\tau})^2}{2}\right] dZ. \tag{7.29}
\end{aligned}$$

Use the following relationship:

$$\int_a^b \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(Z-c)^2}{2}\right] dZ = \int_{a-c}^{b-c} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{Z^2}{2}\right] dZ.$$

Equation (7.29) can be rewritten as:

$$C_1 = S_t \int_{(\ln K-m)/\sigma\sqrt{\tau}-\sigma\sqrt{\tau}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{Z^2}{2}\right] dZ. \tag{7.30}$$

Let $N(\cdot)$ be the standard normal cumulative density function. Using the symmetry of a normal density, (7.30) can be rewritten as:

$$\begin{aligned}
C_1 &= S_t \int_{-\infty}^{-(\ln K-m)/\sigma\sqrt{\tau}+\sigma\sqrt{\tau}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{Z^2}{2}\right] dZ \\
C_1 &= S_t N\left(\frac{-\ln K + m}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau}\right). \tag{7.31}
\end{aligned}$$

From (7.25), substitute for m . The equation (7.31) becomes:

$$\begin{aligned}
C_1 &= S_t N\left(\frac{-\ln K + \ln S_t + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau}\right) \\
C_1 &= S_t N\left(\frac{\ln\left(\frac{S_t}{K}\right) + (r - \frac{1}{2}\sigma^2)\tau + \sigma^2\tau}{\sigma\sqrt{\tau}}\right) = S_t N\left(\frac{\ln\left(\frac{S_t}{K}\right) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\right) \tag{7.32}
\end{aligned}$$

Next, consider C_2 in (7.28):

$$C_2 = Ke^{-r\tau} \int_{(\ln K - m)/\sigma\sqrt{\tau}}^{\infty} \mathbb{Z}(Z) dZ = Ke^{-r\tau} \int_{-\infty}^{-(\ln K - m)/\sigma\sqrt{\tau}} \mathbb{Z}(Z) dZ$$

$$C_2 = Ke^{-r\tau} N\left(\frac{-\ln K + m}{\sigma\sqrt{\tau}}\right). \quad (7.33)$$

From (7.25), substitute for m . The equation (7.33) becomes:

$$C_2 = Ke^{-r\tau} N\left(\frac{-\ln K + \ln S_t + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\right) = Ke^{-r\tau} N\left(\frac{\ln \frac{S_t}{K} + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}\right). \quad (7.34)$$

Substitute (7.32) and (7.34) into (7.28) and we obtain BS plain vanilla call option pricing formula:

$$C(\tau, S_t) = S_t N(d_1) - Ke^{-r\tau} N(d_2), \quad (7.35)$$

$$\text{where } d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} \text{ and } d_2 = \frac{\ln\frac{S_t}{K} + (r - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}} = d_1 - \sigma\sqrt{\tau}.$$

Following the similar method, BS plain vanilla put option pricing formula can be obtained as:

$$P(\tau, S_t) = Ke^{-r\tau} N(-d_2) - S_t N(-d_1). \quad (7.36)$$

We can conclude that both PDE approach and martingale approach give the same result. This is because in both approaches we move from a historical probability measure \mathbb{P} to a risk-neutral probability measure \mathbb{Q} . This is very obvious for martingale method. But in PDE approach because the source of randomness can be completely eliminated by forming a portfolio of options and underlying stocks, this portfolio grows at a rate equal to the risk-free interest rate. Thus, we switch to a measure \mathbb{Q} . For more details, we recommend Neftci (2000) pages 280-282 and 358-365.

[7.5] Alternative Interpretation of Black-Scholes Formula: A Single Integration Problem

Under an equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$ under which the discounted asset price process $\{e^{-rt} S_t; 0 \leq t \leq T\}$ becomes a martingale, a plain vanilla call and put option price which has a terminal payoff function $\max(S_T - K, 0)$ and $\max(K - S_T, 0)$ are calculated as:

$$C(t, S_t) = e^{-r(T-t)} E_t^{\mathbb{Q}} \left[\max(S_T - K, 0) \right] = e^{-r\tau} E_t^{\mathbb{Q}} \left[\max(S_T - K, 0) \right], \quad (7.37)$$

$$P(t, S_t) = e^{-r(T-t)} E_t^{\mathbb{Q}} \left[\max(K - S_T, 0) \right] = e^{-r\tau} E_t^{\mathbb{Q}} \left[\max(K - S_T, 0) \right]. \quad (7.38)$$

Note that an expectation operator $E_t^{\mathbb{Q}} [\cdot]$ is under a probability measure \mathbb{Q} and conditional on time t . Let $\mathbb{Q}(S_T)$ (drop the subscript BS for simplicity) be a probability density function of S_T under \mathbb{Q} . Using $\mathbb{Q}(S_T)$, (7.37) and (7.38) can be rewritten as:

$$\begin{aligned} E_t^{\mathbb{Q}} \left[\max(S_T - K, 0) \right] &= \int_K^{\infty} (S_T - K) \mathbb{Q}(S_T) dS_T + \int_0^K (0) \mathbb{Q}(S_T) dS_T \\ &= \int_K^{\infty} (S_T - K) \mathbb{Q}(S_T) dS_T, \end{aligned}$$

$$\begin{aligned} E_t^{\mathbb{Q}} \left[\max(K - S_T, 0) \right] &= \int_K^{\infty} (0) \mathbb{Q}(S_T) dS_T + \int_0^K (K - S_T) \mathbb{Q}(S_T) dS_T \\ &= \int_0^K (K - S_T) \mathbb{Q}(S_T) dS_T. \end{aligned}$$

Using these, we can rewrite (7.37) and (7.38) as:

$$C(\tau, S_t) = e^{-r\tau} \int_K^{\infty} (S_T - K) \mathbb{Q}(S_T) dS_T, \quad (7.39)$$

$$P(\tau, S_t) = e^{-r\tau} \int_0^K (K - S_T) \mathbb{Q}(S_T) dS_T. \quad (7.40)$$

BS assumes that a terminal stock price S_T is a log-normal random variable with its density of the form:

$$\mathbb{Q}(S_T) = \frac{1}{S_T \sqrt{2\pi\sigma^2\tau}} \exp\left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)\tau \right) \right\}^2}{2\sigma^2\tau}\right]. \quad (7.40)$$

Therefore, BS option pricing formula comes down to a very simple single integration problem:

$$C(\tau, S_t) = e^{-r\tau} \int_K^{\infty} (S_T - K) \frac{1}{S_T \sqrt{2\pi\sigma^2\tau}} \exp\left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)\tau \right) \right\}^2}{2\sigma^2\tau}\right] dS_T, \quad (7.41)$$

$$P(\tau, S_t) = e^{-r\tau} \int_0^K (K - S_T) \frac{1}{S_T \sqrt{2\pi\sigma^2\tau}} \exp\left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)\tau \right) \right\}^2}{2\sigma^2\tau}\right] dS_T. \quad (7.42)$$

This implies that as far as a risk-neutral density of the terminal stock price $\mathbb{Q}(S_T)$ is known, plain vanilla option pricing reduces to a simple integration problem.

[7.6] Black-Scholes Model as an Exponential Lévy Model

The equation (7.10) tells us that BS models a stock price process as a geometric Brownian motion process:

$$S_t = S_0 \exp \left[\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right].$$

BS model is an exponential Lévy model of the form:

$$S_t = S_0 e^{L_t},$$

where the stock price process $\{S_t : 0 \leq t \leq T\}$ is modeled as an exponential of a Lévy process $\{L_t : 0 \leq t \leq T\}$. Black and Scholes' choice of the Lévy process is a Brownian motion with drift (continuous diffusion process):

$$L_t = \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t. \quad (7.43)$$

The fact that an stock price S_t is modeled as an exponential of Lévy process L_t means that its log-return $\ln(\frac{S_t}{S_0})$ is modeled as a Lévy process such that:

$$\ln\left(\frac{S_t}{S_0}\right) = L_t = \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t.$$

BS model can be categorized as a continuous exponential Lévy model apparently because a Brownian motion process is continuous (i.e. no jumps). Later, we will deal with Merton jump-diffusion model (we call it Merton JD model) and variance gamma model by Madan, Carr, and Chang (1998) (we call it VG model). These are all exponential Lévy models of different types. Merton's choice of Lévy process X_t is a Brownian motion with drift process plus a compound Poisson jump process which has a path of continuous path with occasional jumps. Merton JD model can be categorized as a finite activity exponential Lévy model because the expected number of jumps per unit of time (i.e. intensity λ) is finite and small. In other words, the Lévy measure $\ell(dx)$ of Merton JD model is finite:

$$\int \ell(dx) < \infty.$$

VG model can be categorized as an infinite activity exponential Lévy model because the expected number of small jumps per unit of time is infinite:

$$\int \ell(dx) = \infty,$$

although the expected number of large jumps per unit of time is finite. VG model's Lévy process X_t is a subordinated Brownian motion process by a tempered α -stable subordinator (i.e. a normal tempered α -stable process) which is a pure jump process.

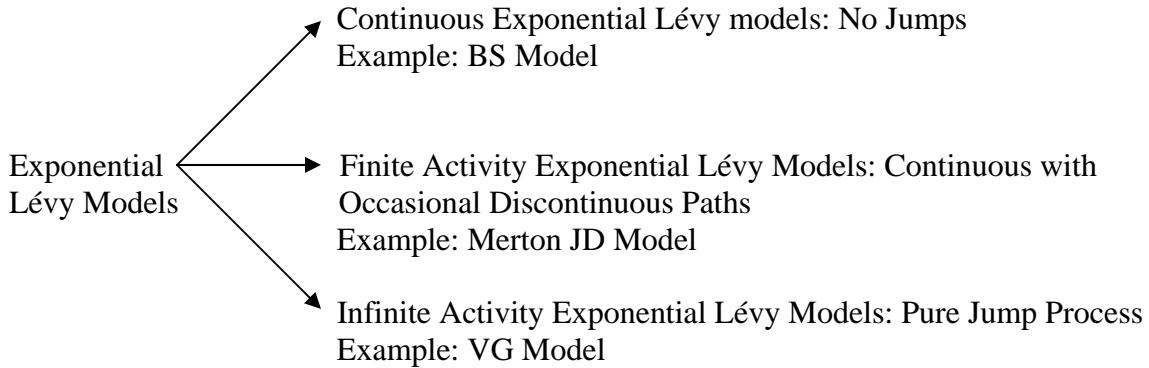


Figure 7.3: Category of Exponential Lévy Models

Recently, there are enormous publications regarding infinite activity exponential Lévy models which we intend to look into in near future.

[8] Option Pricing with Fourier Transform: Black-Scholes Example

In this section we present Fourier transform option pricing approach by Carr and Madan (1999). To illustrate the idea clearly and simply, we present BS model with Fourier transform pricing methodology.

[8.1] Motivation

Let $\mathbb{Q} \sim \mathbb{P}$ be an equivalent martingale measure under which the discounted asset price process $\{e^{-rt} S_t; 0 \leq t \leq T\}$ becomes a martingale and $\{\mathcal{F}_t; 0 \leq t \leq T\}$ be an information flow of the asset price S (i.e. filtration). In an arbitrage-free market, prices of any assets can be calculated as expected terminal payoffs under \mathbb{Q} discounted by a risk-free interest rate r :

$$e^{-rt} S_t = E^{\mathbb{Q}}[e^{-rT} S_T | \mathcal{F}_t], \\ S_t = e^{-r(T-t)} E^{\mathbb{Q}}[S_T | \mathcal{F}_t],$$

which are martingale conditions.

Let K be a strike price and T be an expiration of a contingent claim. Plain vanilla call and put option prices are computed as discounted risk-neutral conditional expectations of the terminal payoffs $(S_T - K)^+ \equiv \max(S_T - K, 0)$ and $(K - S_T)^+ \equiv \max(K - S_T, 0)$:

$$C(t, S_t) = e^{-r(T-t)} E^{\mathbb{Q}}[(S_T - K)^+ | \mathcal{F}_t], \quad (8.1)$$

$$P(t, S_t) = e^{-r(T-t)} E^{\mathbb{Q}}[(K - S_T)^+ | \mathcal{F}_t]. \quad (8.2)$$

Intrinsic value of a vanilla call (put) is defined as $(S_t - K)^+$ ($(K - S_t)^+$) which is the value of the call (put) exercised immediately. Obviously, the intrinsic value of out of the money option is zero. Current option price minus its intrinsic value $C(t, S_t) - (S_t - K)^+$ ($P(t, S_t) - (K - S_t)^+$) is called a time value of the option.

Let $\mathbb{Q}(S_T | \mathcal{F}_t)$ be a probability density function of a terminal asset price S_T under \mathbb{Q} conditional on \mathcal{F}_t . Using $\mathbb{Q}(S_T | \mathcal{F}_t)$, (8.1) and (8.2) can be rewritten as:

$$C(t, S_t) = e^{-r(T-t)} \left\{ \int_K^\infty (S_T - K) \mathbb{Q}(S_T | \mathcal{F}_t) dS_T + \int_0^K (0) \mathbb{Q}(S_T | \mathcal{F}_t) dS_T \right\} \\ = e^{-r(T-t)} \int_K^\infty (S_T - K) \mathbb{Q}(S_T | \mathcal{F}_t) dS_T, \quad (8.3)$$

$$P(t, S_t) = e^{-r(T-t)} \left\{ \int_K^\infty (0) \mathbb{Q}(S_T | \mathcal{F}_t) dS_T + \int_0^K (K - S_T) \mathbb{Q}(S_T | \mathcal{F}_t) dS_T \right\}$$

$$= e^{-r(T-t)} \int_0^K (K - S_T) \mathbb{Q}(S_T | \mathcal{F}_t) dS_T. \quad (8.4)$$

Black-Scholes (BS) assumes that a terminal stock price S_T conditional on \mathcal{F}_t is a log-normal random variable with its density given by:

$$\mathbb{Q}(S_T | \mathcal{F}_t) = \frac{1}{S_T \sqrt{2\pi\sigma^2(T-t)}} \exp \left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)(T-t) \right) \right\}^2}{2\sigma^2(T-t)} \right].$$

Therefore, BS option pricing formula comes down to single integration problem with respect to S_T since all parameters and variables are known:

$$C_{BS}(t, S_t) = e^{-rt} \int_K^\infty (S_T - K) \frac{1}{S_T \sqrt{2\pi\sigma^2\tau}} \exp \left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)\tau \right) \right\}^2}{2\sigma^2\tau} \right] dS_T, \quad (8.5)$$

$$P_{BS}(t, S_t) = e^{-rt} \int_0^K (K - S_T) \frac{1}{S_T \sqrt{2\pi\sigma^2\tau}} \exp \left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)\tau \right) \right\}^2}{2\sigma^2\tau} \right] dS_T. \quad (8.6)$$

This implies that as far as a conditional risk-neutral density of the terminal stock price $\mathbb{Q}(S_T | \mathcal{F}_t)$ is given, plain vanilla option pricing reduces to single integration problem by the equation (8.3) and (8.4).

But for general exponential Lévy models $\mathbb{Q}(S_T | \mathcal{F}_t)$ cannot be expressed using special functions of mathematics or is not known. Therefore, we cannot price plain vanilla options using (8.3) and (8.4). So how do we price options using general exponential Lévy models? The answer is to use a very interesting fact that characteristic functions of general exponential Lévy processes are always known in closed-forms or can be expressed in terms of special functions of mathematics although their probability densities are not. We saw in the section 4 that there is one-to-one relationship between a probability density and a characteristic functions (i.e. through Fourier transform) and both of which uniquely determine a probability distribution. If we can somehow rewrite (8.3) and (8.4) in terms of a characteristic function of $S_T | \mathcal{F}_t$ (i.e. log of $S_T | \mathcal{F}_t$ to be more precise) instead of its probability density $\mathbb{Q}(S_T | \mathcal{F}_t)$, we will be able to price options in general exponential Lévy models.

[8.2] Derivation of Call Price with Fourier Transform: Carr and Madan (1999)

For simplicity, assume $t = 0$ without loss of generality. From the equation (8.3), use a change of variable technique from S_T to $\ln S_T$:

$$C(T, K) = e^{-rT} \int_{\ln K}^{\infty} (e^{\ln S_T} - e^{\ln K}) \mathbb{Q}(\ln S_T | \mathcal{F}_0) d \ln S_T.$$

Let s_T be a log terminal stock price and k be a log strike price, i.e. $s_T \equiv \ln S_T$ and $k \equiv \ln K$. Thus, we have:

$$C(T, k) = e^{-rT} \int_k^{\infty} (e^{s_T} - e^k) \mathbb{Q}(s_T | \mathcal{F}_0) ds_T, \quad (8.7)$$

where $\mathbb{Q}(s_T) \equiv \mathbb{Q}(s_T | \mathcal{F}_0)$ (for simplicity) is a risk-neutral density of a log terminal stock price s_T conditional on filtration \mathcal{F}_0 . From the equation (4.1), a characteristic function of s_T is a Fourier transform of its density function $\mathbb{Q}(s_T)$:

$$\phi_T(\omega) \equiv \mathcal{F}[\mathbb{Q}(s_T)](\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega s_T} \mathbb{Q}(s_T) ds_T. \quad (8.8)$$

Consider a function $g(t)$. Sufficient (but not necessary) condition for the existence of Fourier transform and its inverse is the equation (3.47):

$$\int_{-\infty}^{\infty} |g(t)| dt < \infty. \quad (8.9)$$

We saw in section 7 that BS models a log terminal stock price s_T as a normal random variable with its normal density given by under \mathbb{Q} (from (7.7)):

$$\mathbb{Q}(s_T) = \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp \left[-\frac{\left\{ s_T - \left(s_0 + (r - \frac{1}{2}\sigma^2)T \right) \right\}^2}{2\sigma^2 T} \right]. \quad (8.10)$$

From the equations (8.8) and (8.10), a characteristic function of BS log terminal stock price s_T is easily obtained as:

$$\phi_T(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega s_T} \mathbb{Q}(s_T) ds_T = \exp \left[i \left\{ s_0 + (r - \frac{1}{2}\sigma^2)T \right\} \omega - \frac{(\sigma^2 T)\omega^2}{2} \right]. \quad (8.11)$$

When a call price is expressed in terms of a log strike price $k \equiv \ln K$ in the equation (8.7), k approaches $-\infty$ as a strike price K approaches 0 in the limit. Thus, from (7.8):

$$\begin{aligned} C(T, k) &= e^{-rT} \int_{-\infty}^{\infty} (e^{s_T} - e^{-\infty}) \mathbb{Q}(s_T | \mathcal{F}_0) ds_T = e^{-rT} \int_{-\infty}^{\infty} e^{s_T} \mathbb{Q}(s_T | \mathcal{F}_0) ds_T \\ C(T, k) &= e^{-rT} E^{\mathbb{Q}}[e^{s_T} | \mathcal{F}_0]. \end{aligned} \quad (8.12)$$

We know under equivalent martingale measure \mathbb{Q} :

$$E^{\mathbb{Q}}[S_T \equiv e^{s_T} | \mathcal{F}_0] = S_0 e^{rT}.$$

Equation (8.12) becomes:

$$C(T, k) = S_0.$$

Therefore, a call price $C(T, k)$ is not integrable (i.e. $C(T, k)$ does not satisfy (8.9)). Therefore, $C(T, k)$ cannot be Fourier transformed. To solve this problem, CM defines a modified call price as:

$$C_{\text{mod}}(T, k) \equiv e^{\alpha k} C(T, k), \quad (8.13)$$

where $C_{\text{mod}}(T, k)$ is expected to satisfy the integrability condition (8.9) by carefully choosing $\alpha > 0$:

$$\int_{-\infty}^{\infty} |C_{\text{mod}}(T, k)| dk < \infty.$$

Consider a FT of a modified call price by the FT definition (3.4):

$$\psi_T(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega k} C_{\text{mod}}(T, k) dk. \quad (8.14)$$

From (8.14), call price $C(T, k)$ can be obtained by an inverse FT (i.e. the definition (3.5)) of $\psi_T(\omega)$:

$$\begin{aligned} C_{\text{mod}}(T, k) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \psi_T(\omega) d\omega \\ e^{\alpha k} C(T, k) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \psi_T(\omega) d\omega \\ C(T, k) &= \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \psi_T(\omega) d\omega. \end{aligned} \quad (8.15)$$

Now CM derives an analytical expression of $\psi_T(\omega)$ in terms of a characteristic function $\phi_T(\omega)$. Substitute (8.13) into (8.14):

$$\psi_T(\omega) = \int_{-\infty}^{\infty} e^{i\omega k} e^{\alpha k} C(T, k) dk.$$

Substitute (8.7) and interchanging integrals yields:

$$\begin{aligned} \psi_T(\omega) &= \int_{-\infty}^{\infty} e^{i\omega k} e^{\alpha k} e^{-rT} \int_k^{\infty} (e^{s_T} - e^k) \mathbb{Q}(s_T) ds_T dk \\ &= \int_{-\infty}^{\infty} e^{-rT} \mathbb{Q}(s_T) \int_{-\infty}^s e^{i\omega k} (e^{s_T + \alpha k} - e^{(1+\alpha)k}) dk ds_T \\ &= \int_{-\infty}^{\infty} e^{-rT} \mathbb{Q}(s_T) \left(\frac{e^{(\alpha+1+i\omega)s_T}}{\alpha+i\omega} - \frac{e^{(\alpha+1+i\omega)s_T}}{\alpha+1+i\omega} \right) ds_T \\ &= \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega}. \end{aligned} \quad (8.16)$$

Thus, a call pricing function is obtained by substituting (8.16) into (8.15):

$$C(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} d\omega, \quad (8.17)$$

where $\phi_T(\cdot)$ is a characteristic function of a log terminal stock price s_T conditional on filtration \mathcal{F}_0 . We can interpret the equation (8.17) which is single numerical integration problem as a characteristic function $\phi_T(\cdot)$ equivalent of the equation (8.3). Table 8.1 illustrates this point.

Table 8.1: Comparison between Traditional and FT Option Pricing Formula

Option Pricing Method	Equation	Formula
Traditional	(8.3)	$C = e^{-r(T-t)} \int_K^{\infty} (S_T - K) \mathbb{Q}(S_T \mathcal{F}_t) dS_T$
Fourier Transform	(8.17)	$C(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} d\omega$ $\phi_T(\cdot) \equiv \int_{-\infty}^{\infty} e^{i\omega \ln(S_T)} \mathbb{Q}(\ln(S_T) \mathcal{F}_0) d\ln(S_T), k \equiv \ln K$

[8.3] How to Choose Decay Rate Parameter α : Carr and Madan (1999)

We mentioned earlier that when a call price is expressed in terms of a log strike price $k \equiv \ln K$ in the equation (8.7), k approaches $-\infty$ as a strike price K approaches 0 in the limit which is illustrated in Figure 8.1.

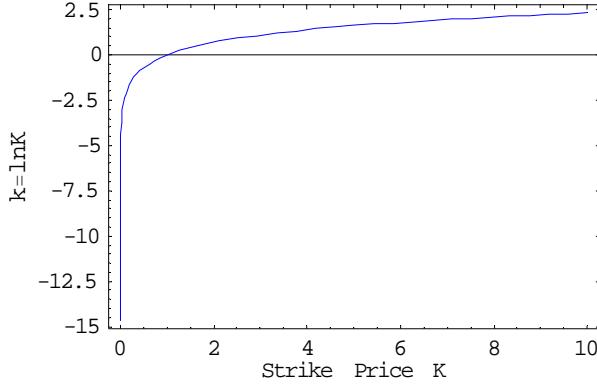


Figure 8.1: Relationship between a Strike Price K and a Log-Strike Price $k \equiv \ln K$.

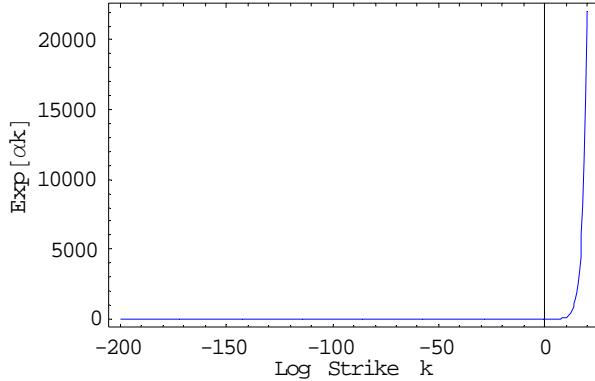


Figure 8.2: Plot of an Exponential Function $e^{0.5k}$ with Respect to Log-Strike Price $k \equiv \ln K$.

Thus, a call price function $C(T, k)$ becomes S_0 as $k \rightarrow -\infty$ discussed before. In order to make a call price integrable, CM multiplies an exponential function $e^{\alpha k}$ with $\alpha \in \mathbb{R}^+$ to $C(T, k)$ and obtains a modified call price $C_{\text{mod}}(T, k)$. As shown by Figure 8.2, the role of $e^{\alpha k}$ with $\alpha \in \mathbb{R}^+$ is to dampen the size of $C(T, k) = S_0$ in the limit $k \rightarrow -\infty$. But this in turn worsens the integrability condition for $k \in \mathbb{R}^+$ (i.e. positive log strike) axis. In order for a modified call price $C_{\text{mod}}(T, k) \equiv e^{\alpha k} C(T, k)$ to be integrable for both positive and negative k axis (i.e. square integrable), CM provides a sufficient condition:

$$\psi_T(0) < \infty.$$

From (8.16):

$$\psi_T(0) = \frac{e^{-rT} \phi_T(-(\alpha+1)i)}{\alpha^2 + \alpha}.$$

Therefore, the sufficient condition of the square-integrability of $C_{\text{mod}}(T, k)$ is:

$$\phi_T(-(\alpha+1)i) < \infty. \quad (8.18)$$

From the definition of a characteristic function (8.8):

$$\phi_T(\omega) = E[e^{i\omega s_T}] = E[e^{i\omega \ln S_T}] = E[(e^{\ln S_T})^{i\omega}] = E[S_T^{i\omega}]. \quad (8.19)$$

From (8.18) and (8.19):

$$\begin{aligned} \phi_T(-(\alpha+1)i) &< \infty \\ E[S_T^{i\{-(\alpha+1)i\}}] &< \infty \\ E[S_T^{\alpha+1}] &< \infty. \end{aligned} \quad (8.20)$$

CM suggests the use of (8.20) and the analytical expression of the characteristic function to determine an upper bound on α .

[8.4] Black-Scholes Model with Fourier Transform Pricing Method

By substituting a characteristic function of BS log terminal stock price s_T (i.e. the equation (8.11)) into the general FT pricing formula of the equation (8.17), we obtain BS-FT call pricing formula:

$$C_{BS-FT}(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} d\omega, \quad (8.21)$$

$$\text{with } \phi_T(\omega) = \exp \left[i \left\{ s_0 + \left(r - \frac{1}{2} \sigma^2 \right) T \right\} \omega - \frac{(\sigma^2 T) \omega^2}{2} \right].$$

Table 8.2: Comparison between Original BS and BS-FT Option Pricing Formula

Option Pricing Method
Original Black-Scholes

Original Black-Scholes

$$C = e^{-r(T-t)} \int_K^\infty (S_T - K) \mathbb{Q}(S_T | \mathcal{F}_t) dS_T$$

$$\mathbb{Q}(S_T | \mathcal{F}_t) = \frac{1}{S_t \sqrt{2\pi\sigma^2(T-t)}} \exp \left[-\frac{\left\{ \ln S_T - \left(\ln S_t + (r - \frac{1}{2}\sigma^2)(T-t) \right) \right\}^2}{2\sigma^2(T-t)} \right]$$

Black-Scholes with Fourier Transform Approach

$$C = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} d\omega$$

$$\phi_T(\omega) = \exp \left[i \left\{ s_0 + (r - \frac{1}{2}\sigma^2)T \right\} \omega - \frac{(\sigma^2 T)\omega^2}{2} \right], k = \ln K$$

We implement the BS-FT formula (8.21) with decay rate parameter $\alpha = 1$ and compare the result to the original BS call price. As illustrated by Figure 8.3, as a principle BS-FT call price and the original BS call price are identical. This is no surprise because the original BS formula and BS-FT formula in Table 8.2 are the same person with a different look. BS-FT formula is just frequency representation of the original BS formula.

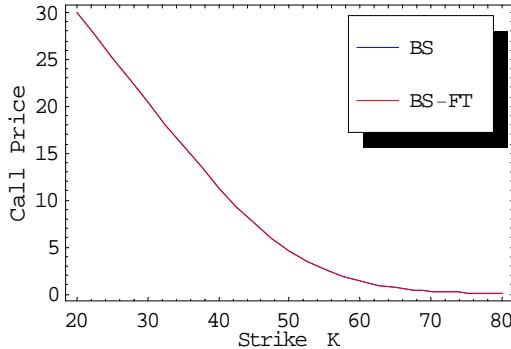


Figure 8.3: Original BS Call Price Vs. BS-FT with $\alpha = 1$ Call Price. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 1$.

Several points should be investigated further now. Firstly, CPU time should be discussed. Consider an ATM call option with $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 1$. Our BS-FT $\alpha = 1$ code needs 0.01 second CPU time while our original BS code needs zero seconds. Although you can state that BS-FT formula is slower than the original BS, speed is not an issue for most of the purposes.

Secondly, let's consider the choice of decay rate parameter α . We saw the selection procedure for α by Carr and Madan (1999) in section 8.3. But we found out that choice of α is not much important as long as $\alpha \in \mathbb{R}^+$. Consider an ATM vanilla call price with $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 0.25$. Its BS price is 2.16794.

Figure 8.4 and Table 8.3 indicates that for $\alpha \geq 0.05$ BS-FT price converges to the

original BS price. But it is possible that option prices using FT approach can be sensitive to the choice of α for models other than the BS, so at this point it is safer for us to state that for at least BS model the choice of α is of no importance (i.e. $\alpha=1$ is fine).

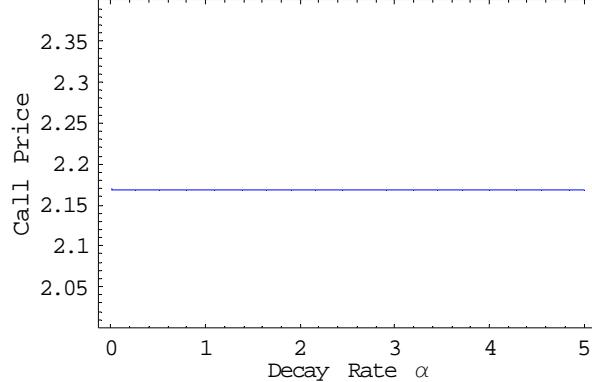


Figure 8.4: Vanilla BS-FT ATM Call Price as a Function of Decay Rate Parameter

a. Common parameters and variables fixed are $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 0.25$.

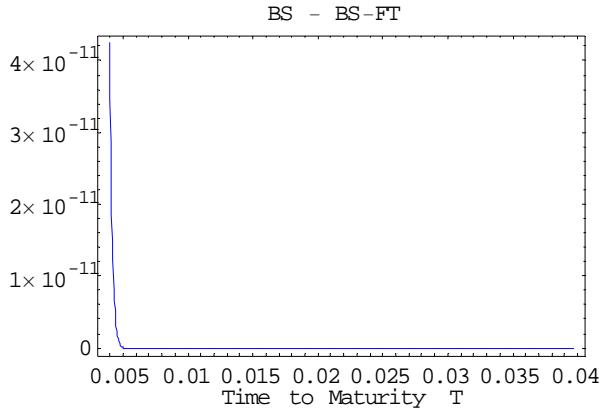
Table 8.3: BS-FT ATM Call Price for Different Values of Decay Rate Parameter a

α	0.01	0.05	0.1	0.5	1	2	5
Call Price	2.16972	2.16794	2.16794	2.16794	2.16794	2.16794	2.16794

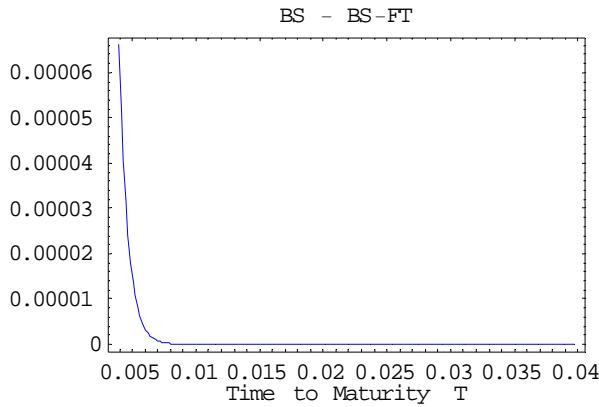
Thirdly, we discuss the FT option pricing with (8.21) for near maturity deep OTM and ITM calls. As CM (1999) points out, for near maturity deep OTM and ITM calls (puts), call (put) prices approach their intrinsic values $(S_t - K)^+$ as $t \rightarrow T$ ($(K - S_t)^+$). Since this makes the inverse Fourier transform integrand highly oscillatory, the numerical integration problem in (8.21) becomes slow and difficult. For example, using our code, in the case of an ATM call with $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$, BS-FT $\alpha=1$ call price does not face the difficulty in numerical integration at all even for one day to maturity $T = 1/252$. But in the case of a deep OTM call with $S_0 = 50$, $K = 80$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$, BS-FT $\alpha=1$ call price begins to experience difficulty in numerical integration around 37 trading days to maturity (i.e. $T < 38/252$). For a deep ITM call with $S_0 = 50$, $K = 20$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$, BS-FT $\alpha=1$ call price begins to experience difficulty around 5 days to maturity (i.e. $T < 5/252$). Of course, this depends on the value of α , your hardware, and your software, etc.

More important question is to investigate the amount of error caused by this numerical integration difficulty for the near maturity deep OTM and ITM options. Consider a call option with $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$. Figure 8.5 plots a series of the difference between the original BS price and BS-FT $\alpha=1$ price for the range of less than 10 trading days to maturity $1/252 \leq T \leq 10/252$. We find that despite the difficulty in the numerical integration of BS-FT price of (8.21) for the near maturity deep OTM (in Panel

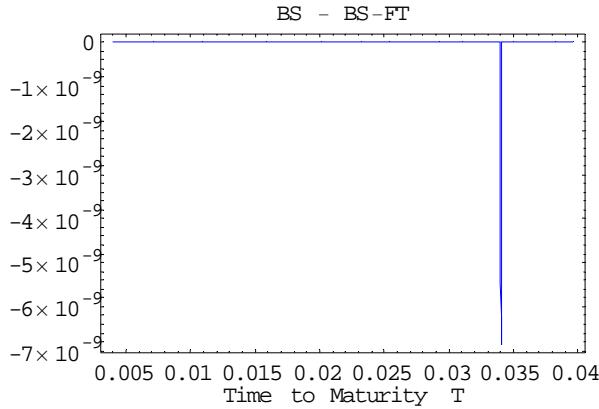
A) and deep ITM (in Panel B) call, our BS-FT code causes no significant errors in terms of pricing. For a near maturity ATM call (in Panel C) which faces no difficulty in numerical integration, BS-FT pricing error is negligible.



A) OTM call with $S_0 = 50$ and $K = 80$.



B) ITM call with $S_0 = 50$ and $K = 20$.



C) ATM Call with $S_0 = 50$ and $K = 50$.

Figure 8.5: Plot of BS Price minus BS-FT $\alpha=1$ Price for Near-Maturity Vanilla Call with Less Than 10 Trading Days to Maturity $1/252 \leq T \leq 10/252$. Common parameters and variables fixed are $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

We conclude this section by stating the following remarks. In the case of BS model, the FT call price of the equation (8.17) needs equivalent CPU time as the original BS formula and produces negligible pricing errors regardless of the maturity and the moneyness of the call.

[8.5] Derivation of Near-Maturity OTM Vanilla Option Pricing Function with Fourier Transform: Carr and Madan (1999)

Although the numerical integration difficulty of BS-FT call price of (8.21) for the near maturity deep OTM and deep ITM calls causes no significant pricing errors, CM (1999) provides the vanilla option pricing formula specifically designed for them which makes the inverse Fourier transform integrand less oscillatory and facilitates the numerical integration problem.

Let $z_T(k)$ be the time value of an OTM vanilla options with maturity T at the current time $t = 0$ without loss of generality. This implies that when the strike price is less than the spot stock price $K < S_0$ (i.e. $k < s_0$ in log-scale), $z_T(k)$ takes on the OTM put price. When the strike price is greater than the spot stock price $K > S_0$ (i.e. $k > s_0$ in log-scale), $z_T(k)$ takes on the OTM call price:

$$z_T(k) \equiv \begin{cases} P(0, S_0) - (K - S_0)^+ = P(0, S_0) & \text{if } K < S_0 \\ C(0, S_0) - (S_0 - K)^+ = C(0, S_0) & \text{if } K > S_0 \end{cases} .$$

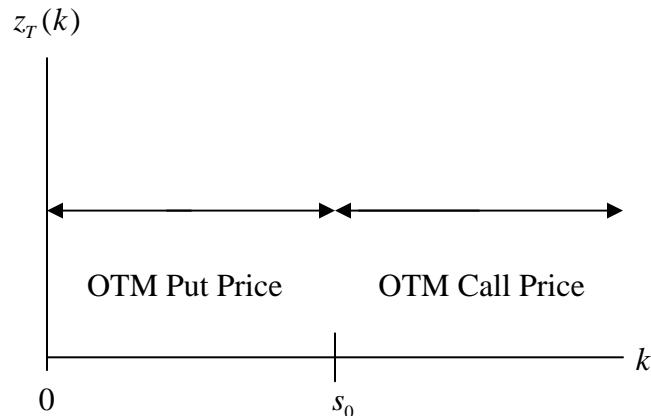


Figure 8.6: Illustration of an OTM Vanilla Option Price $z_T(k)$.

Assume the current stock price $S_0 = 1$ for the simplicity (i.e. taken care in the end) which in turn means $s_0 = 0$.

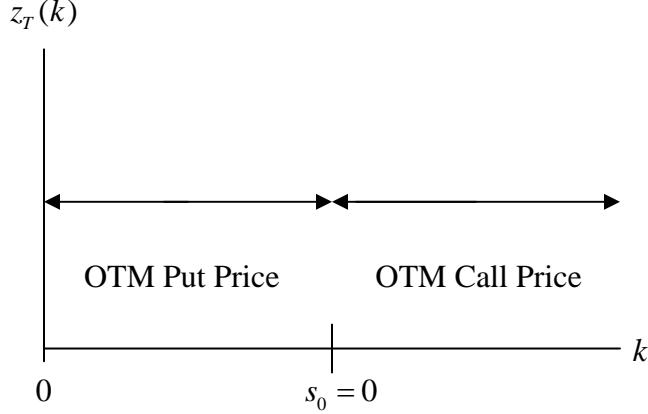


Figure 8.7: Illustration of an OTM Vanilla Option Price $z_T(k)$ Assuming $S_0 = 1$.

Under a risk-neutral measure \mathbb{Q} , $z_T(k)$ can be computed as:

$$z_T(k) = e^{-rT} \int_{-\infty}^{\infty} \left[(e^k - e^{s_T}) \mathbf{1}_{s_T < k, k < 0} + (e^{s_T} - e^k) \mathbf{1}_{s_T > k, k > 0} \right] \mathbb{Q}(s_T | \mathcal{F}_0) ds_T, \quad (8.22)$$

where $\mathbb{Q}(s_T) \equiv \mathbb{Q}(s_T | \mathcal{F}_0)$ is a risk-neutral density of a log terminal stock price s_T conditional on filtration \mathcal{F}_0 . $z_T(k)$ takes on a put payoff $(e^k - e^{s_T})$ if this option is currently OTM put $k < 0$ and if at maturity T this put finishes ITM $s_T < k$. $z_T(k)$ takes on a call payoff $(e^{s_T} - e^k)$ if this option is currently OTM call $k > 0$ and if at maturity T this call finishes ITM $s_T > k$.

Consider a FT of OTM vanilla option price $z_T(k)$ by the FT definition (3.4):

$$\zeta_T(\omega) = \int_{-\infty}^{\infty} e^{i\omega k} z_T(k) dk. \quad (8.23)$$

From (8.23), OTM vanilla option price $z_T(k)$ can be obtained by an inverse FT (i.e. the definition (3.5)) of $\zeta_T(\omega)$:

$$z_T(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \zeta_T(\omega) d\omega. \quad (8.24)$$

By substituting (8.22) into (8.23), CM shows that an analytical expression of $\zeta_T(\omega)$ in terms of a characteristic function $\phi_T(\omega)$ of a log terminal stock price s_T conditional on filtration \mathcal{F}_0 can be obtained as:

$$\zeta_T(\omega) = e^{-rT} \left(\frac{1}{1+i\omega} - \frac{e^{rT}}{i\omega} - \frac{\phi_T(\omega-i)}{\omega^2 - i\omega} \right). \quad (8.25)$$

To facilitate numerical integration, CM again considers a FT of OTM vanilla option price $z_T(k)$ modified with dampening function $\sinh(\alpha k)$ (compare with (8.23)):

$$\begin{aligned} \gamma_T(\omega) &= \int_{-\infty}^{\infty} e^{i\omega k} \sinh(\alpha k) z_T(k) dk \\ &= \int_{-\infty}^{\infty} e^{i\omega k} \frac{e^{\alpha k} - e^{-\alpha k}}{2} z_T(k) dk \\ &= \frac{\zeta_T(\omega - i\alpha) - \zeta_T(\omega + i\alpha)}{2}. \end{aligned} \quad (8.26)$$

From (8.26), OTM vanilla option price $z_T(k)$ can be obtained by an inverse FT of $\gamma_T(\omega)$:

$$z_T(k) = \frac{1}{\sinh(\alpha k)} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \gamma_T(\omega) d\omega. \quad (8.27)$$

From now on, we call the OTM near maturity option price of the equation (8.27) as FT/TV (i.e. TV indicates the time value approach). Table 8.4 compares the general FT pricing formula using the modified call price (i.e. the equation (8.17)) and FT/TV formula using the time value of OTM options which is specifically designed for near maturity options.

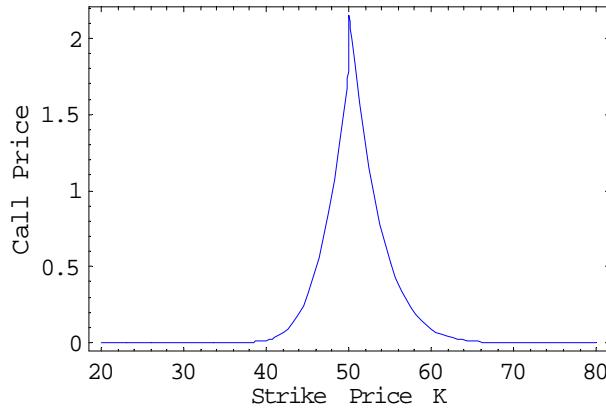
Table 8.4: FT Formula Vs. FT/TV Formula

Approach	Formula
FT	$C(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha + 1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha + 1)\omega} d\omega$ $\phi_T(.) \equiv \int_{-\infty}^{\infty} e^{i\omega \ln(S_T)} \mathbb{Q}(\ln(S_T) \mathcal{F}_0) d\ln(S_T), \quad k \equiv \ln K$
FT/TV	$z_T(k) \equiv \begin{cases} P(0, s_0) & \text{if } k < s_0 \\ C(0, s_0) & \text{if } k > s_0 \end{cases}$ $z_T(k) = \frac{1}{\sinh(\alpha k)} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \gamma_T(\omega) d\omega$ $\gamma_T(\omega) = \frac{\zeta_T(\omega - i\alpha) - \zeta_T(\omega + i\alpha)}{2}$

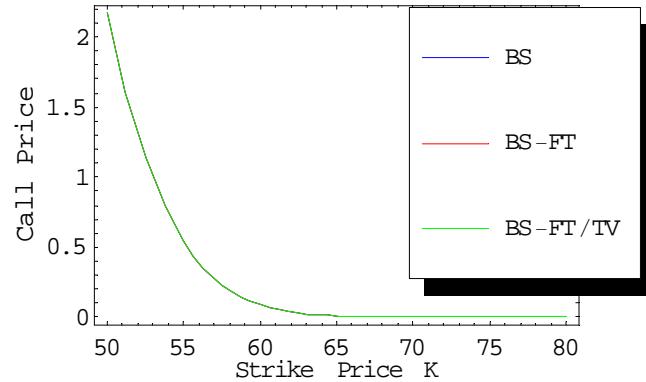
$$\zeta_T(\omega) = e^{-rT} \left(\frac{1}{1+i\omega} - \frac{e^{rT}}{i\omega} - \frac{\phi_T(\omega-i)}{\omega^2 - i\omega} \right)$$

$$\phi_T(.) \equiv \int_{-\infty}^{\infty} e^{i\omega \ln(S_T)} \mathbb{Q}(\ln(S_T) | \mathcal{F}_0) d \ln(S_T), \quad k \equiv \ln K$$

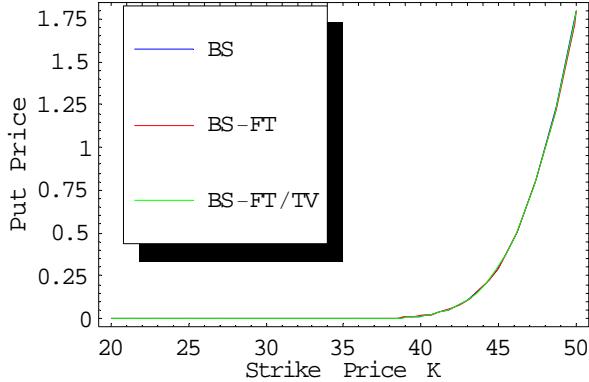
Next the performance of FT-TV formula of the equation (8.27) with decay rate parameter $\alpha=10$ (we will discuss the choice of α soon) is tested and compared to the original BS price and BS-FT $\alpha=1$ price for a range of different moneyness in Figure 8.8. Panel A plots $z_T(k)$ which reaches its maximum at ATM and takes the price of OTM put for $K < S_0$ and takes the price of OTM call for $K > S_0$. Panel B and C indicates that as a principle, all three formulae produce identical prices. Again, this is no surprise because all these are same thing with different looks.



A) Plot of $z_T(k)$.



B) For the range $50 \leq K \leq 80$.



C) For the range $20 \leq K \leq 50$.

Figure 8.8: BS Vs. BS-FT $\alpha = 1$ Vs. BS-FT/TV $\alpha = 10$. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 0.25$.

Next, CPU time and the accuracy of BS-FT/TV formula are discussed. Consider vanilla options with common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$. Table 8.5 to 8.7 show CPU time and the accuracy of BS-FT/TV $\alpha=10$ formula compared to other two formulae with three different maturities and with varying moneyness. In terms of CPU time, we notice that BS-FT/TV formula is by far the slowest although most of purposes this will not be an issue. BS-FT/TV $\alpha=10$ formula has marginally larger error than BS-FT $\alpha=1$ formula, but the size of the errors are trivial.

Table 8.5: CPU Time for an OTM Call with $K = 80$

Price is in the bracket.

Method	Time to Maturity		
	$T = 0.1$	$T = 0.5$	$T = 1$
BS	0 seconds (3.77524×10^{-14})	0 seconds (0.00152306)	0 seconds (0.0594469)
BS-FT $\alpha = 1$	0.08 seconds (3.97419×10^{-14})	0.02 seconds (0.00152306)	0.01 seconds (0.0594469)
BS/FT/OTM $\alpha = 10$	0.31 seconds (4.00513×10^{-14})	0.07 seconds (0.00152306)	0.03 seconds (0.0594469)

Table 8.6: CPU Time for an ATM Call with $K = 50$

Price is in the bracket.

Method	Time to Maturity		
	$T = 0.1$	$T = 0.5$	$T = 1$
BS	0 seconds	0 seconds	0 seconds

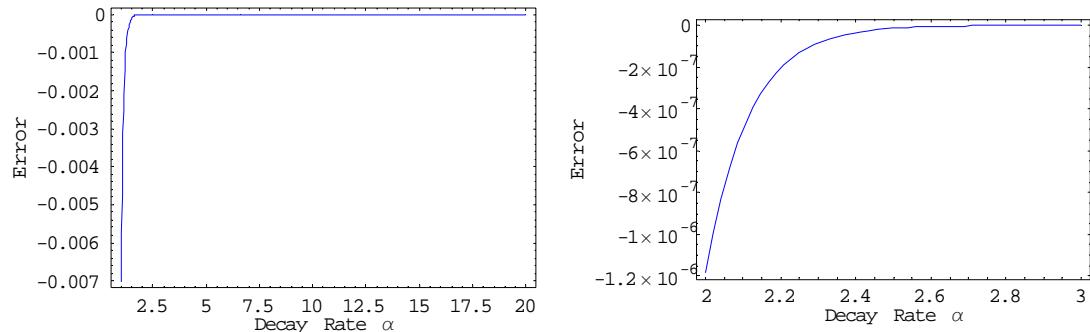
	(1.3331)	(3.15382)	(4.6135)
BS-FT $\alpha = 1$	0.01 seconds (1.3331)	0.01 seconds (3.15382)	0.01 seconds (4.6135)
BS/FT/OTM $\alpha = 10$	0.04 seconds (1.3331)	0.04 seconds (3.15382)	0.07 seconds (4.6135)

Table 8.7: CPU Time for an OTM Put with $K = 20$

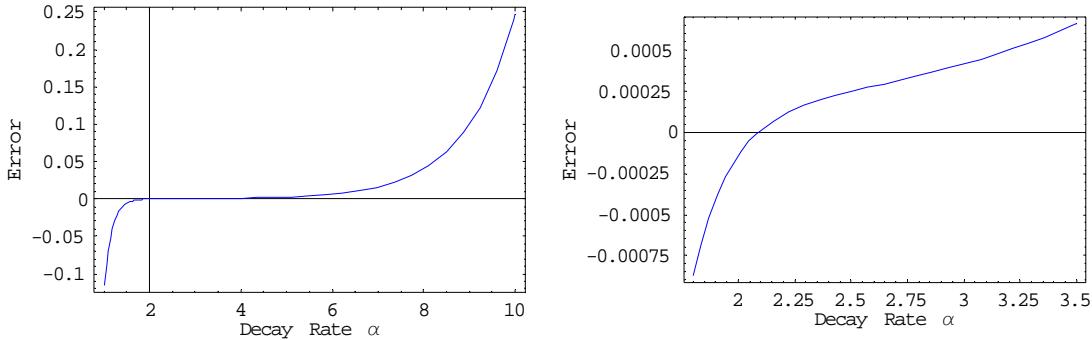
Price is in the bracket.

Method	Time to Maturity		
	$T = 0.1$	$T = 0.5$	$T = 1$
BS	0 seconds (0)	0 seconds (1.4492×10^{-11})	0 seconds (1.32586×10^{-6})
BS-FT $\alpha = 1$	0.04 seconds (-7.10543×10^{-15})	0.02 seconds (1.4481×10^{-11})	0.02 seconds (1.32586×10^{-6})
BS/FT/OTM $\alpha = 10$	0.3 seconds (-3.76588×10^{-13})	0.14 seconds (1.86517×10^{-11})	0.13 seconds (1.32587×10^{-6})

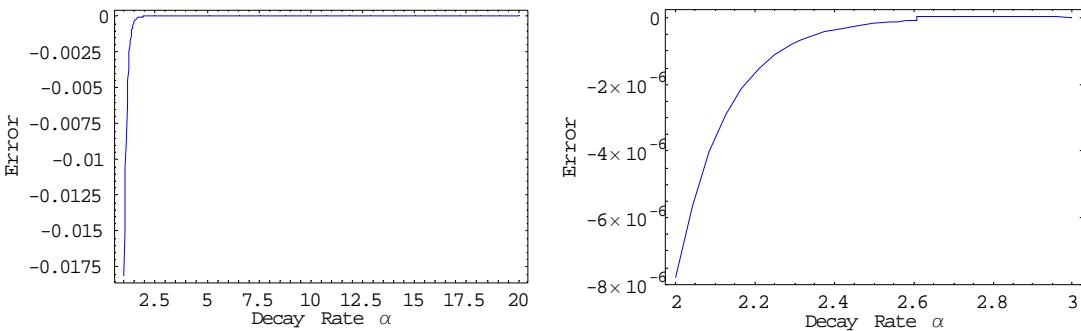
Next, the level of decay rate parameter α for BS-FT/TV formula is dealt. Figure 8.9 illustrates the pricing error of BS-FT/TV formula of a one day to maturity $T = 1/252$ option as a function of varying α . Panel A (for an OTM call) and C (for an ATM call) tells us that for $\alpha \geq 2$, BS-FT/TV formula has effectively zero error relative to BS price. But for an OTM put (Panel B), the error does not monotonically decrease as α rises. It seems that the value of $2 < \alpha < 2.25$ yields the negligible size of the error. Therefore, from now on, we always use $\alpha = 2.1$ when implementing BS-FT/TV formula.



A) For an OTM Call with $K = 80$.



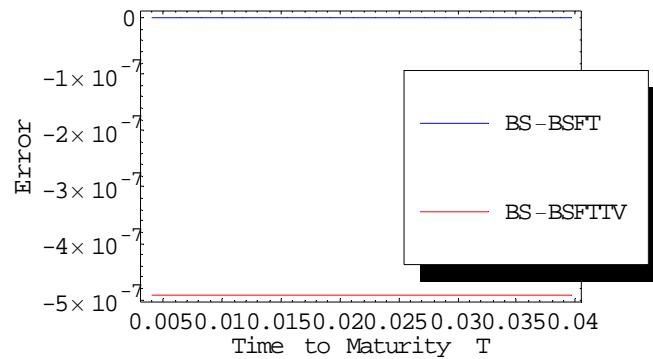
B) For an OTM Put with $K = 20$.



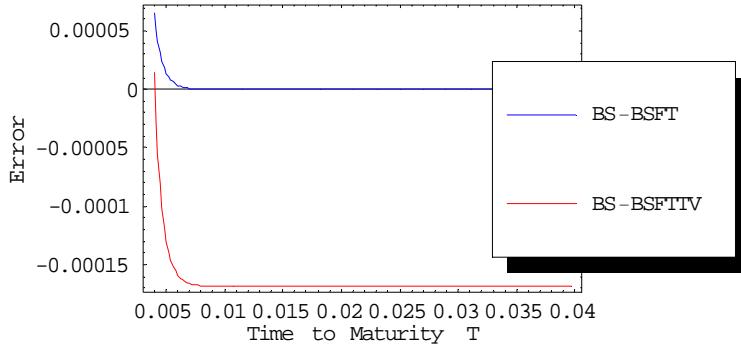
C) For an ATM Call with $K = 50$.

Figure 8.9: BS Price Minus BS-FT/TV Price for One Day to Maturity Option as a Function of Decay Rate Parameter α . Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 1/252$.

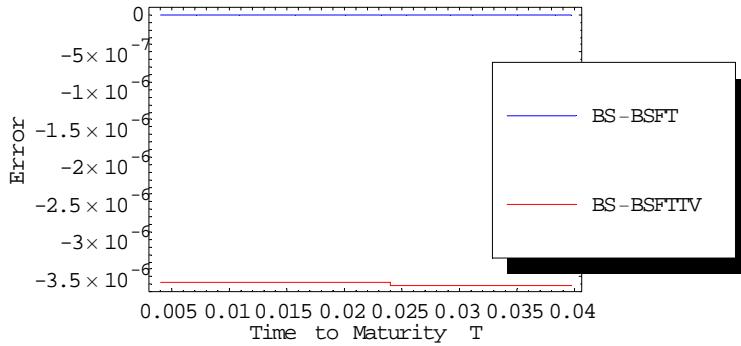
Next, we compare the performance between BS-FT $\alpha = 1$ formula and BS-FT/TV $\alpha = 2.1$ formula and let's see if BS-FT/TV $\alpha = 2.1$ formula improves the pricing accuracy for near maturity options. According to Figure 8.10, BS-FT/TV $\alpha = 2.1$ formula generally has larger error regardless of the moneyness except for one-day to maturity OTM put (in Panel B).



A) OTM call with $K = 80$.



B) OTM put with $K = 20$.



C) ATM Call with $K = 50$.

Figure 8.10: Plot of Error of BS-FT/TV $\alpha = 2.1$ Price Vs. Error of BS-FT $\alpha = 1$ Price with Less Than 10 Trading Days to Maturity $1/252 \leq T \leq 10/252$. Common parameters and variables fixed are $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

We conclude this section by stating the following remarks. Carr and Madan (1999) provides the FT/TV formula of the equation (8.27) in order to improve the pricing accuracy specifically for near maturity options compared to the FT formula of the equation (8.17). Firstly, FT/TV formula is slower than the original BS and FT formula, but this won't be an issue for pricing several options. Secondly, we recommend the use of $2 < \alpha < 2.25$ for the decay rate in FT/TV formula. Thirdly, contrary to the design by CM, FT/TV formula generally has larger error than FT $\alpha = 1$ formula. At this point, we are skeptical about the usefulness of FT/TV formula.

[8.6] Derivation of Call Pricing Function with Discrete Fourier Transform (DFT): Carr and Madan (1999)

There still remains one problem to be solved. Although the numerical integration difficulty of FT call price of (8.17) for the near maturity options causes no significant pricing errors, it makes the evaluation of FT price slow. This speed becomes an issue when calibrating hundreds or thousands of prices (i.e. also in Monte Carlo simulation).

To improve computational time, Carr and Madan (1999) apply discrete Fourier transform to approximate the equation (8.17). Note that our version of DFT call price formula is different from their original formula.

FT call price is (i.e. the equations (8.15) and (8.16)):

$$C_T(k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \psi_T(\omega) d\omega, \quad (8.28)$$

where $\psi_T(\omega) = \frac{e^{-rT} \phi_r(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega}$. The integral is an inverse FT of an angular frequency domain function $\psi_T(\omega)$ into signal space k function. As we saw in section 5, it can be approximated by DFT. Firstly, the infinite integral of (8.28) needs to be truncated as:

$$C_T(k) \approx \frac{e^{-\alpha k}}{2\pi} \int_{-\Omega/2}^{\Omega/2} e^{-i\omega k} \psi_T(\omega) d\omega. \quad (8.29)$$

Secondly, (8.29) can be discretized as:

$$C_T(k_n) \approx \frac{\exp(-\alpha k_n)}{2\pi} \sum_{j=0}^{N-1} w_j \left\{ \exp(-i\omega_j k_n) \psi_T(\omega_j) \Delta\omega \right\}. \quad (8.30)$$

Let us discuss the interpretation of (8.30). The purpose of a DFT (an inverse DFT) is to approximate the transform space function, i.e. in our case FT of a modified call price $\psi_T(\omega)$ (the signal space function, i.e. in our case a call price $C_T(k)$) as close as possible by sampling a finite number of points N of a signal space function $C_T(k)$ with signal space sampling interval Δk and by sampling a finite number of points N of a transform space function $\psi_T(\omega)$ with transform space sampling interval $\Delta\omega$. In other words, both the original continuous signal space function $C_T(k)$ and the original continuous FT $\psi_T(\omega)$ are approximated by a sample of N points.

Let N be the number of discrete samples taken to approximate $C_T(k)$ and $\psi_T(\omega)$. Let Δk be the signal space sampling interval which is the signal increment between signal space samples. Its inverse $f_s \equiv 1/\Delta k$ (samples/1 unit of k) is called a signal space sampling rate. Let K be the total sampling range in the signal space:

$$K \equiv N\Delta k. \quad (8.31)$$

If Δk is 1, the total sampling range in the signal space and the number of samples taken are same (i.e. for example, $K = 10$ dollars and $N = 10$ samples). If $\Delta k = 0.01$, 1 sample

is taken in every 0.01 unit of k in signal space which in turn means that 100 samples are taken per 1 unit of k (i.e. signal space sampling rate $f_s \equiv 1/\Delta k = 100$ Hz).

Consider a sampling in the signal space. Take N discrete samples of a signal space function $C_T(k)$ at n -th sampling instant $k_n \equiv -\frac{N\Delta k}{2} + n\Delta k$ with $n = 0, 1, \dots, N-1$. When $K = 10$ and $N = 10$ samples (i.e. signal space sampling interval $\Delta k = 1$), $k_n \equiv -5 + n$ which means that $C_T(k)$ is sampled at $k = -5, -4, -3, -2, -1, 0, 1, 2, 3$, and 4. Let $C_T(k_n)$ be the sampled values of $C_T(k)$.

Next, consider a sampling in the transform space (i.e. angular frequency ω domain). Let $\Delta\omega$ Hz (radians/1 unit of k) be the angular frequency domain sampling interval:

$$\Delta\omega \equiv \frac{2\pi}{N\Delta k} \equiv \frac{2\pi}{K}. \quad (8.32)$$

The total sampling range in the angular frequency domain Ω is:

$$\Omega \equiv N\Delta\omega. \quad (8.33)$$

Take N samples of a continuous FT $\mathcal{G}(\omega)$ at j -th angular frequency sampling instant:

$$\omega_j \equiv -\frac{N\Delta\omega}{2} + j\Delta\omega \equiv \Delta\omega \left(-\frac{N}{2} + j \right) \equiv \frac{2\pi}{N\Delta k} \left(-\frac{N}{2} + j \right) \equiv -\frac{\pi}{\Delta k} + j \frac{2\pi}{N\Delta k}, \quad (8.34)$$

with $j = 0, 1, \dots, N-1$. For example, when $K = 10$ and $N = 10$ samples (i.e. $\Delta k = 1$), $\mathcal{G}(\omega)$ is sampled at $\omega = -\pi, -4\pi/5, -3\pi/5, -2\pi/5, -\pi/5, 0, \pi/5, 2\pi/5, 3\pi/5, 4\pi/5$ (i.e. $\omega_j = -\pi + j\pi/5$). Let $\mathcal{G}(\omega_j)$ be the sampled values of $\mathcal{G}(\omega)$:

$$\mathcal{G}(\omega_j) \equiv \mathcal{G}\left(-\frac{N\Delta\omega}{2} + j\Delta\omega\right). \quad (8.35)$$

$\mathcal{G}(\omega_j)$ is called a spectrum of $g(k_n)$ at angular frequency ω_j and it is a complex number.

DFT and inverse DFT defines the relationship between the sampled signal space function $g(k_n)$ and its spectrum at angular frequency ω_j , $\mathcal{G}(\omega_j)$, as the following:

$$\mathcal{G}(\omega_j) \equiv \sum_{n=0}^{N-1} g(k_n) \exp\left\{i\omega_j k_n\right\}$$

$$\begin{aligned}\mathcal{G}(\omega_j) &\equiv \sum_{n=0}^{N-1} g(k_n) \exp \left\{ i \left(\frac{\pi N}{2} - \pi j - \pi n + \frac{2\pi j n}{N} \right) \right\} \\ \mathcal{G}(\omega_j) &\equiv \exp(-i\pi j) \exp(i\pi N/2) \sum_{n=0}^{N-1} \{g(k_n) \exp(-i\pi n)\} \exp(i2\pi j n/N),\end{aligned}\quad (8.36)$$

and:

$$\begin{aligned}g(k_n) &\equiv \frac{1}{N} \sum_{j=0}^{N-1} \mathcal{G}(\omega_j) \exp(-i\omega_j k_n) \\ g(k_n) &\equiv \frac{1}{N} \sum_{j=0}^{N-1} \mathcal{G}(\omega_j) \exp \left\{ -i \left(\frac{\pi N}{2} - \pi j - \pi n + \frac{2\pi j n}{N} \right) \right\} \\ g(k_n) &\equiv \exp(-i\pi N/2) \exp(i\pi n) \frac{1}{N} \sum_{j=0}^{N-1} \mathcal{G}(\omega_j) \exp(-i2\pi j n/N) \exp(i\pi j),\end{aligned}\quad (8.37)$$

where the following relationships are used:

$$\begin{aligned}k_n &\equiv -\frac{N\Delta k}{2} + n\Delta k \\ \omega_j &\equiv -\frac{N\Delta\omega}{2} + j\Delta\omega \equiv \Delta\omega \left(-\frac{N}{2} + j \right) \equiv \frac{2\pi}{N\Delta k} \left(-\frac{N}{2} + j \right) \equiv -\frac{\pi}{\Delta k} + j \frac{2\pi}{N\Delta k} \\ k_n \omega_j &\equiv \frac{\pi N}{2} - \pi j - \pi n + \frac{2\pi j n}{N}.\end{aligned}$$

Therefore, from the equation (8.30):

$$\begin{aligned}C_T(k_n) &\approx \frac{\exp(-\alpha k_n)}{2\pi} \sum_{j=0}^{N-1} w_j \{ \exp(-i\omega_j k_n) \psi_T(\omega_j) \Delta\omega \} \\ C_T(k_n) &\approx \frac{\exp(-\alpha k_n)}{2\pi} \sum_{j=0}^{N-1} w_j \left[\exp \left\{ -i \left(\frac{\pi N}{2} - \pi j - \pi n + \frac{2\pi j n}{N} \right) \right\} \psi_T(\omega_j) \frac{2\pi}{N\Delta k} \right] \\ C_T(k_n) &\approx \frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k} \\ &\quad \times \frac{1}{N} \sum_{j=0}^{N-1} w_j \{ \exp(i\pi j) \psi_T(\omega_j) \} \exp(-i2\pi j n/N),\end{aligned}\quad (8.38)$$

where $w_{j \in [0, N-1]}$ are weights for sampled points. For example, when trapezoidal rule is chosen:

$$w_j = \begin{cases} 1/2 & \text{for } j = 0 \text{ and } N-1 \\ 1 & \text{for others} \end{cases}.$$

For near-maturity OTM options, DFT/TV formula which is a DFT approximation of the formula (8.27) is given by simply replacing $\exp(-\alpha k_n)$ by $\frac{1}{\sinh(\alpha k_n)}$ and $\psi_T(\omega_j)$ by $\gamma_T(\omega_j)$ in the DFT call price (8.38) as the following:

$$\begin{aligned} z_T(k_n) &\approx \frac{\exp(i\pi n)\exp(-i\pi N/2)}{\sinh(\alpha k_n)\Delta k} \\ &\times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \gamma_T(\omega_j) \right\} \exp(-i2\pi jn/N), \end{aligned} \quad (8.39)$$

where $\gamma_T(\omega_j)$ is given by (8.26).

[8.7] Implementation and Performance of DFT Pricing Method with Black-Scholes Model

In this section, performance of DFT call price of the equation (8.38) is tested in the BS case. We will call this as BS-DFT. We implement the formula (8.38) with decay rate parameter $\alpha = 1$ and compare the result to the original BS call price and BS-FT $\alpha = 1$ call price under various settings. We will start from its implementation.

The first step to implement (8.38) is to choose the number of samples taken N and the signal space sampling interval Δk which we will call as log-strike space sampling interval, hereafter. Note that selecting Δk corresponds to selecting the frequency domain sampling interval $\Delta\omega$, because these are related by:

$$\Delta k \Delta \omega \equiv \frac{2\pi}{N}. \quad (8.40)$$

We set $N = 4096$ as suggested by CM. Note that because our code uses DFT not FFT, our choice of N does not need to be a power of 2. We use $\Delta k = 0.005$ which is a half the percentage point in the log-strike space. This corresponds to $\Delta\omega = 0.306796$ radians. The total sampling range in the log-strike space is $K = N\Delta k = 20.48$, its sampling rate is 200 samples per unit of k , and the total sampling range in the angular frequency domain is $\Omega = N\Delta\omega = 1256.64$.

Second step is to construct the $N = 4096$ point-sampling grid in the frequency domain using $\omega_j \equiv -\frac{N\Delta\omega}{2} + j\Delta\omega$. Table 8.8 illustrates this. Using this grid, obtain $N = 4096$ point-samples of DFT integrand in (8.32), i.e. $w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\}$. With common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 0.25$, sampled values of $w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\}$ are shown in Table 8.9.

Table 8.8: Angular Frequency Domain Sampling Grid

Index j	j -th angular frequency point ω_j (Hz)
0	-628.319
1	-628.012
2	-627.705
~	~
4095	627.705
4096	628.012

Table 8.9: Angular Frequency Domain Sampled Values

Index j	j -th frequency domain sample $w_j \{ \exp(i\pi j) \psi_T(\omega_j) \}$
0	$8.6137 \times 10^{-859} + 2.5868 \times 10^{-859} i$
1	$-8.5146 \times 10^{-859} - 1.2340 \times 10^{-857} i$
2	$-7.7265 \times 10^{-857} + 3.5391 \times 10^{-857} i$
~	~
4095	$-7.7265 \times 10^{-857} - 3.5391 \times 10^{-857} i$
4096	$-4.2573 \times 10^{-859} + 6.1702 \times 10^{-858} i$

Third step is to perform inverse DFT of Table 8.9 and multiply the result by $\frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k}$ in the equation (8.38). This amounts to the $N = 4096$

point-samples of call price $C_T(k_n)$ in log-strike space. Table 8.10 shows this result and it is plotted in Figure 8.7.

Table 8.10: $N = 4096$ Point-Samples of BS Call Price $C_T(k_n)$ in Log-Strike Space

Index n	n -th sample of call price $C_T(k_n)$
0	49.7525
1	49.752
~	~
2500	40.3338
~	~
2600	34.2249
~	~
2700	24.153
~	~
2800	7.6428
~	~
2900	0.0006067
~	~

4095	8.0931×10^{-17}
4096	8.0929×10^{-17}

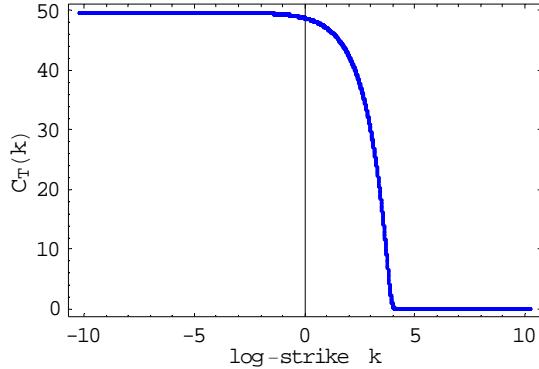


Figure 8.11: $N = 4096$ Point-Samples of Call Price $C_T(k_n)$ in Log-Strike Space

Fourth step is to match the log-strike price k grid with the strike price grid K as shown in Table 8.11.

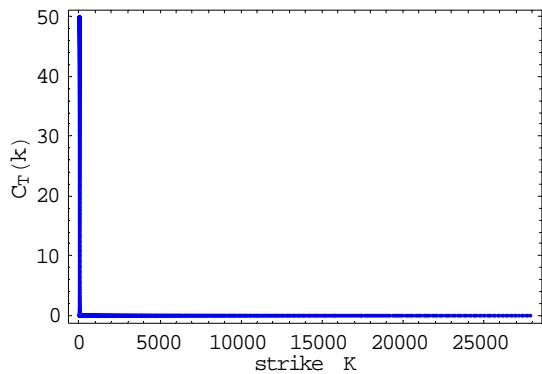
Table 8.11: $N = 4096$ Point Log-Strike Price (k) Grid and Strike Price (K) Grid

Index n	n -th log-strike price point $n\Delta k$	n -th strike price point $\exp[n\Delta k]$
0	-10.24	0.00003571
~		
1500	-2.745	0.0642
~		
2500	2.255	9.5353
~		
2600	2.755	15.721
~		
2700	3.255	25.9196
~		
2800	3.755	42.7342
~		
2900	4.255	70.4568
~		
4096	10.235	27861.5

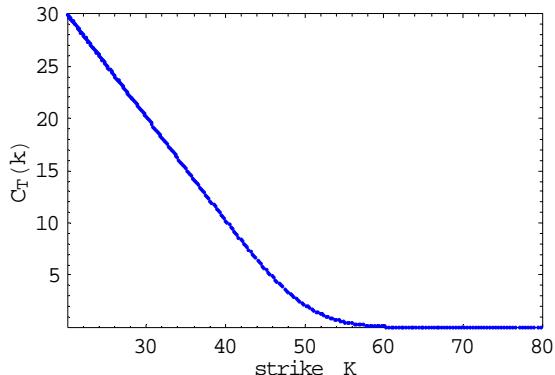
Fifth and final step is to match the strike price K grid with the computed call price as shown in the table 8.12 which is plotted in Figure 8.12. Panel A of Figure 8.12 is for the entire range of the strike price sampled, and Panel B is only for the range of our interest. Remember that we use the spot stock price $S_0 = 50$ and the maturity $T = 0.25$.

Table 8.12: $N = 4096$ Point Sampled BS-DFT Call Price with respect to Strike Price K_n

Index n	n -th strike price K_n	BS-DFT Call Price
0	0.00003571	49.7525
~		
1500	0.0642	49.6872
~		
2500	9.5353	40.3338
~		
2600	15.721	34.2249
~		
2700	25.9196	24.153
~		
2800	42.7342	7.6428
~		
2900	70.4568	0.0006067
~		
4096	27861.5	8.0929×10^{-17}



A) For the entire strike price range.



B) Our interest.

Figure 8.12: $N = 4096$ Point Sampled BS-DFT Call Price with respect to Strike Price K_n .

This completes the step-by-step procedure of DFT option price computation. Next, we investigate the several important properties of BS-DFT $\alpha = 1$ call price.

Firstly, let's consider the call price computed by the original BS, BS-FT $\alpha = 1$, BS-DFT $\alpha = 1$ with $N = 4096$ and $\Delta k = 0.005$ using common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 0.25$ with respect to a strike price K . The result is plotted in Figure 8.13 which indicates that as a principle these three approaches produce the identical price. Again, this is no surprise because all these are same thing with different looks and different precisions.

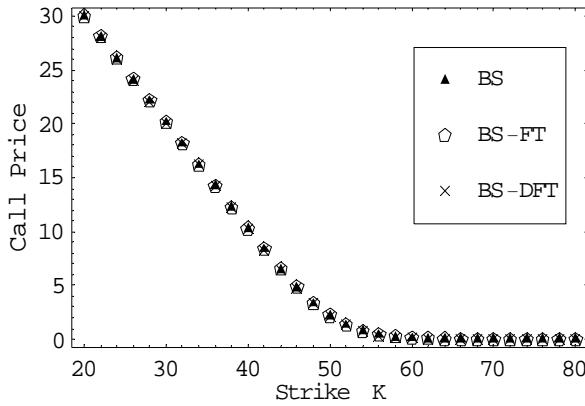


Figure 8.13: BS Vs. BS-FT $\alpha = 1$ Vs. BS-DFT $\alpha = 1$. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 0.25$.

Secondly, let's investigate the difference in CPU time. Consider calculating 100-point call prices for a range of strike price $1 \leq K \leq 100$ with interval 1 with common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 20/252$. Table 8.13 compares CPU time using the original BS, BS-FT $\alpha = 1$, and BS-DFT $\alpha = 1$ with $N = 4096$ and $\Delta k = 0.005$ and Figure 8.14 reports the prices. We notice that there is a significant improvement in the computational time by the use of DFT.

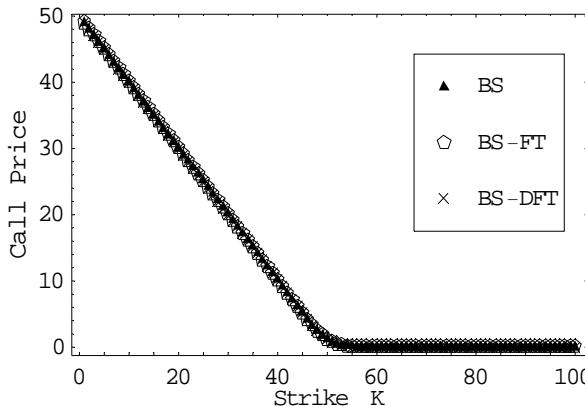


Figure 8.14: BS Vs. BS-FT $\alpha = 1$ Vs. BS-DFT $\alpha = 1$. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 20/252$.

Table 8.13 CPU Time for Calculating 100-point call prices for a range of strike price $1 \leq K \leq 100$ with interval 1 Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 20/252$.

Method	CPU Time
BS	0.03 seconds
BS-FT $\alpha = 1$	5.448 seconds
BS-DFT $\alpha = 1$ $N = 4096$, $\Delta k = 0.005$	0.491 seconds

Thirdly, we pay extra attention to the pricing errors of very near-maturity call prices because these are of Carr and Madan's interest. Consider a call option with common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$. Figures 8.15 to 8.17 plot three series of price differences for vanilla calls computed by the original BS, BS-FT $\alpha = 1$, and BS-DFT $\alpha = 1$ with $N = 4096$ and $\Delta k = 0.005$ as a function of time to maturity of less than a month $1/252 \leq T \leq 20/252$. Figure 8.15 tells us that for deep OTM calls, BS-DFT yields the pricing error around 6×10^{-8} which we interpret negligible. Figure 8.17 tells us that for deep ITM calls, BS-DFT price and BS price are virtually identical except for one day to maturity. Figure 8.16 is the most interesting case among the three. It tells us that as the maturity nears, the error of ATM BS-DFT price monotonically increases and the size of error is large but negligible. In contrast, ATM BS-FT price produces virtually no error. This finding is more clearly illustrated in Figure 8.17 where the $T = 1/252$ ATM pricing error (relative to BS) of BS-DFT $\alpha = 1$ with $N = 4096$ and $\Delta k = 0.005$ is plotted across different moneyness. We again realize that BS-DFT price is virtually identical to BS price for the deep ITM and OTM calls, but its approximation error becomes an issue for ATM call. We have done several experiments trying to reduce the size of this around ATM error by increasing the decay rate parameter α to 10 or by sampling more points $N = 8192$. But these attempts were futile.

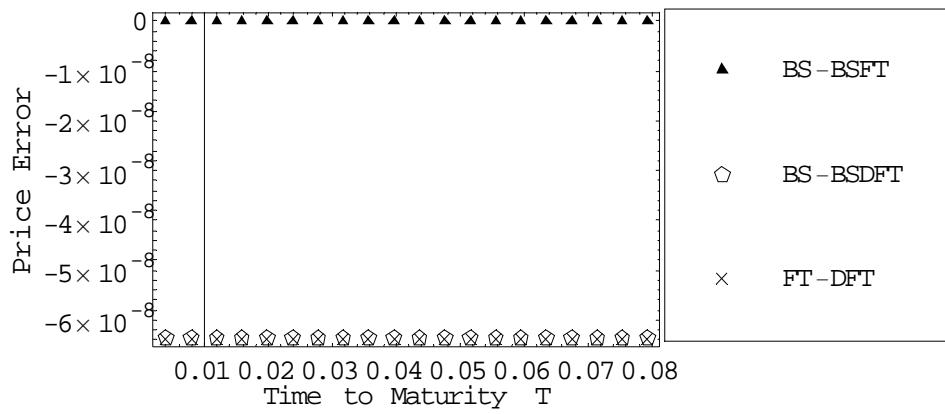


Figure 8.15: Plot of Price Error for Deep-OTM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 80$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

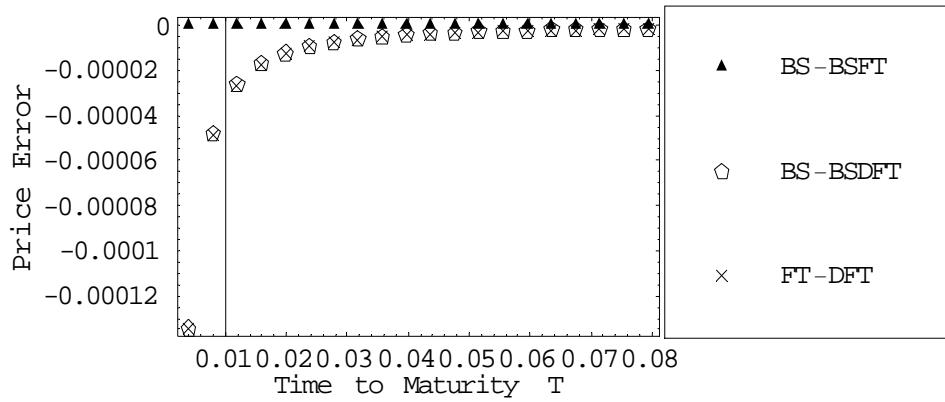


Figure 8.16: Plot of Price Error for ATM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

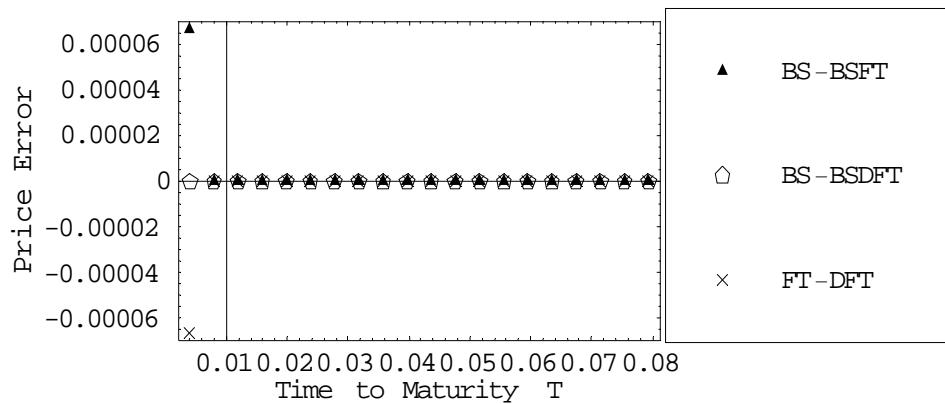


Figure 8.17: Plot of Price Error for Deep-ITM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 20$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

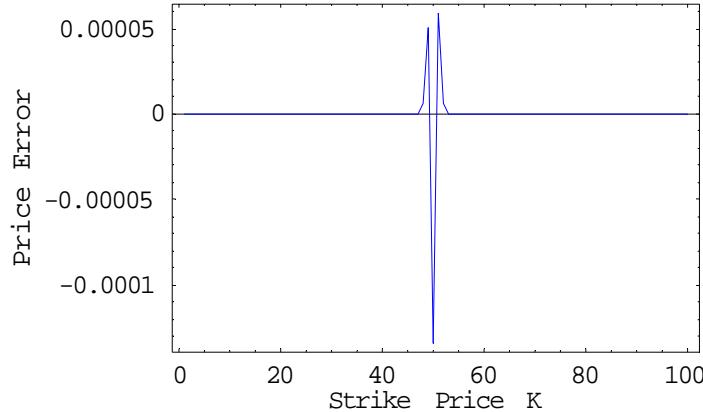


Figure 8.18: Plot of Price Error of BS-DFT Formula for 1-Day-to-Maturity ATM Call as a Function of Strike Price $0 \leq K \leq 100$. Common parameters and variables fixed are $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

We conclude this section by stating the following remarks. Our version of CM (1999) DFT call price formula, the equation (8.38), yields the price virtually identical to the original BS price for OTM and ITM calls even for extreme near-maturity case (i.e. $T = 1/252$) although the size of error is larger than BS-FT formula. But the error of BS-DFT price becomes large around (i.e. ± 3) ATM. In our example used, the maximum error is 0.0001345 which occurs at exactly ATM at $S_0 = K = 50$. Increasing the decay rate parameter α or sampling more points (i.e. larger N) cannot reduce the size of this error. But we can accept this size of error when considering the dramatic improvement in the CPU time to compute hundreds of prices.

[8.8] Summary of Formulae of Option Price with Fourier Transform

Table 8.14: Summary of Formulae of Option Price with Fourier Transform

Method	Formula
Traditional	$C = e^{-r(T-t)} \int_K^\infty (S_T - K) \mathbb{Q}(S_T \mathcal{F}_t) dS_T$
FT	$C(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha + 1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha + 1)\omega} d\omega$ $\phi_T(.) \equiv \int_{-\infty}^{\infty} e^{i\omega \ln(S_T)} \mathbb{Q}(\ln(S_T) \mathcal{F}_0) d\ln(S_T), \quad k \equiv \ln K$

$$\begin{aligned}
& \text{FT/TV} \\
z_T(k) &= \frac{1}{\sinh(\alpha k)} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \gamma_T(\omega) d\omega \\
\gamma_T(\omega) &= \frac{\zeta_T(\omega - i\alpha) - \zeta_T(\omega + i\alpha)}{2} \\
\zeta_T(\omega) &= e^{-rT} \left(\frac{1}{1+i\omega} - \frac{e^{rT}}{i\omega} - \frac{\phi_T(\omega - i)}{\omega^2 - i\omega} \right) \\
\phi_T(\cdot) &\equiv \int_{-\infty}^{\infty} e^{i\omega \ln(S_T)} \mathbb{Q}(\ln(S_T) | \mathcal{F}_0) d \ln(S_T), \quad k \equiv \ln K
\end{aligned}$$

$$\begin{aligned}
& \text{DFT} \\
C_T(k_n) &\approx \frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k} \\
&\times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\} \exp(-i2\pi jn/N)
\end{aligned}$$

$$\begin{aligned}
& \text{DFT/TW} \\
C_T(k_n) &\approx \frac{\exp(i\pi n) \exp(-i\pi N/2)}{\sinh(\alpha k_n) \Delta k} \\
&\times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \gamma_T(\omega_j) \right\} \exp(-i2\pi jn/N) \\
\gamma_T(\omega) &= \frac{\zeta_T(\omega - i\alpha) - \zeta_T(\omega + i\alpha)}{2} \\
\zeta_T(\omega) &= e^{-rT} \left(\frac{1}{1+i\omega} - \frac{e^{rT}}{i\omega} - \frac{\phi_T(\omega - i)}{\omega^2 - i\omega} \right) \\
\phi_T(\cdot) &\equiv \int_{-\infty}^{\infty} e^{i\omega \ln(S_T)} \mathbb{Q}(\ln(S_T) | \mathcal{F}_0) d \ln(S_T), \quad k \equiv \ln K
\end{aligned}$$

[9] Merton (1976) Jump-Diffusion Model

[9.1] Model Type

In this section the basic structure of Merton JD model is described without the derivation of the model which will be done in the next section.

Merton JD model is an exponential Lévy model of the form:

$$S_t = S_0 e^{X_t},$$

where the asset price process $\{S_t; 0 \leq t \leq T\}$ is modeled as an exponential of a Lévy process $\{X_t; 0 \leq t \leq T\}$. Merton's choice of the Lévy process is a Brownian motion with drift (continuous diffusion process) plus a compound Poisson process (discontinuous jump process) such that:

$$X_t = (\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t + \sum_{i=1}^{N_t} Y_i,$$

where $\{B_t; 0 \leq t \leq T\}$ is a standard Brownian motion process. The term

$(\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t$ is a Brownian motion with drift process and the term $\sum_{i=1}^{N_t} Y_i$ is a compound Poisson jump process. The only difference between the Black-Scholes and the Merton jump-diffusion is the addition of the term $\sum_{i=1}^{N_t} Y_i$. A compound Poisson jump process $\sum_{i=1}^{N_t} Y_i$ contains two sources of randomness. The first is the Poisson process dN_t with intensity (i.e. average number of jumps per unit of time) λ which causes the asset price to jump randomly (i.e. random timing). Once the asset price jumps, how much it jumps is also modeled random (i.e. random jump size). Merton assumes that log stock price jump size follows normal distribution, $(dx_i) \sim i.i.d. Normal(\mu, \delta^2)$:

$$f(dx_i) = \frac{1}{\sqrt{2\pi\delta^2}} \exp\left\{-\frac{(dx_i - \mu)^2}{2\delta^2}\right\}.$$

It is assumed that these two sources of randomness are independent of each other. By introducing three extra parameters λ , μ , and δ to the original BS model, Merton JD model tries to capture the (negative) skewness and excess kurtosis of the log return density $\mathbb{P}(\ln(S_t / S_0))$ which deviates from the BS normal log return density.

Lévy measure $\ell(dx)$ of a compound Poisson process is given by the multiplication of the intensity and the jump size density $f(dx)$:

$$\ell(dx) = \lambda f(dx).$$

A compound Poisson process (i.e. a piecewise constant Lévy process) is called finite activity Lévy process since its Lévy measure $\ell(dx)$ is finite (i.e. the average number of jumps per unit time is finite):

$$\int_{-\infty}^{\infty} \ell(dx) = \lambda < \infty.$$

The fact that an asset price S_t is modeled as an exponential of Lévy process X_t means that its log-return $\ln(\frac{S_t}{S_0})$ is modeled as a Lévy process such that:

$$\ln\left(\frac{S_t}{S_0}\right) = X_t = (\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t + \sum_{i=1}^{N_t} Y_i.$$

Let's derive the model.

[9.2] Model Derivation

In the jump-diffusion model, changes in the asset price consist of normal (continuous diffusion) component that is modeled by a Brownian motion with drift process and abnormal (discontinuous, i.e. jump) component that is modeled by a compound Poisson process. Asset price jumps are assumed to occur independently and identically. The probability that an asset price jumps during a small time interval dt can be written using a Poisson process dN_t as:

$$\Pr\{\text{an asset price jumps once in } dt\} = \Pr\{dN_t = 1\} \cong \lambda dt,$$

$$\Pr\{\text{an asset price jumps more than once in } dt\} = \Pr\{dN_t \geq 2\} \cong 0,$$

$$\Pr\{\text{an asset price does not jump in } dt\} = \Pr\{dN_t = 0\} \cong 1 - \lambda dt,$$

where the parameter $\lambda \in \mathbb{R}^+$ is the intensity of the jump process (the mean number of jumps per unit of time) which is independent of time t .

Suppose in the small time interval dt the asset price jumps from S_t to $y_t S_t$ (we call y_t as absolute price jump size). So the relative price jump size (i.e. percentage change in the asset price caused by the jump) is:

$$\frac{dS_t}{S_t} = \frac{y_t S_t - S_t}{S_t} = y_t - 1,$$

where Merton assumes that the absolute price jump size y_t is a nonnegative random variables drawn from lognormal distribution, i.e. $\ln(y_t) \sim i.i.d.N(\mu, \delta^2)$. This in turn

implies that $E[y_t] = e^{\mu + \frac{1}{2}\delta^2}$ and $E[(y_t - E[y_t])^2] = e^{2\mu + \delta^2}(e^{\delta^2} - 1)$. This is because if $\ln x \sim N(a, b)$, then $x \sim Lognormal(e^{\frac{a+1}{2}\delta^2}, e^{2a+\delta^2}(e^{\delta^2} - 1))$.

Merton's jump-diffusion dynamics of asset price which incorporates the above properties takes the SDE of the form:

$$\frac{dS_t}{S_t} = (\alpha - \lambda k)dt + \sigma dB_t + (y_t - 1)dN_t, \quad (9.1)$$

where α is the instantaneous expected return on the asset, σ is the instantaneous volatility of the asset return conditional on that jump does not occur, B_t is a standard Brownian motion process, and N_t is an Poisson process with intensity λ . Standard assumption is that (B_t) , (N_t) , and (y_t) are independent. The relative price jump size of S_t , $y_t - 1$, is lognormally distributed with the mean $E[y_t - 1] = e^{\mu + \frac{1}{2}\delta^2} - 1 \equiv k$ and the variance $E[(y_t - 1 - E[y_t - 1])^2] = e^{2\mu + \delta^2}(e^{\delta^2} - 1)$.¹ This may be confusing to some readers, so we will repeat it again. Merton assumes that the absolute price jump size y_t is a lognormal random variable such that:

$$(y_t) \sim i.i.d. Lognormal(e^{\mu + \frac{1}{2}\delta^2}, e^{2\mu + \delta^2}(e^{\delta^2} - 1)). \quad (9.2)$$

This is equivalent to saying that Merton assumes that the relative price jump size $y_t - 1$ is a lognormal random variable such that:

$$(y_t - 1) \sim i.i.d. Lognormal(k \equiv e^{\mu + \frac{1}{2}\delta^2} - 1, e^{2\mu + \delta^2}(e^{\delta^2} - 1)). \quad (9.3)$$

This is equivalent to saying that Merton assumes that the log price jump size $\ln y_t \equiv Y_t$ is a normal random variable such that:

$$\ln(y_t) \sim i.i.d. Normal(\mu, \delta^2). \quad (9.4)$$

This is equivalent to saying that Merton assumes that the log-return jump size $\ln(\frac{y_t S_t}{S_t})$ is a normal random variable such that:

¹For random variable x , $Variance[x-1] = Variance[x]$.

$$\ln\left(\frac{y_t S_t}{S_t}\right) = \ln(y_t) \equiv Y_t \sim i.i.d. Normal(\mu, \delta^2). \quad (9.5)$$

It is extremely important to note:

$$E[y_t - 1] = e^{\mu + \frac{1}{2}\delta^2} - 1 \equiv k \neq E[\ln(y_t)] = \mu,$$

because $E[y_t - 1] \neq E[\ln(y_t)] = E[\ln(y_t)]$.

The expected relative price change $E[dS_t / S_t]$ from the jump part dN_t in the time interval dt is $\lambda k dt$ since $E[(y_t - 1)dN_t] = E[y_t - 1]E[dN_t] = k\lambda dt$. This is the predictable part of the jump. This is why the instantaneous expected return on the asset αdt is adjusted by $-\lambda k dt$ in the drift term of the jump-diffusion process to make the jump part an unpredictable innovation:

$$\begin{aligned} E\left[\frac{dS_t}{S_t}\right] &= E[(\alpha - \lambda k)dt] + E[\sigma dB_t] + E[(y_t - 1)dN_t] \\ E\left[\frac{dS_t}{S_t}\right] &= (\alpha - \lambda k)dt + 0 + \lambda k dt = \alpha dt. \end{aligned}$$

Some researchers include this adjustment term for predictable part of the jump $-\lambda k dt$ in the drift term of the Brownian motion process leading to the following simpler (?) specification:

$$\begin{aligned} \frac{dS_t}{S_t} &= \alpha dt + \sigma dB_t + (y_t - 1)dN_t \\ B_t &\sim Normal\left(-\frac{\lambda k t}{\sigma}, t\right) \\ E\left[\frac{dS_t}{S_t}\right] &= \alpha dt + \sigma\left(-\frac{\lambda k dt}{\sigma}\right) + \lambda k dt = \alpha dt. \end{aligned}$$

But we choose to explicitly subtract $\lambda k dt$ from the instantaneous expected return αdt because we prefer to keep B_t as a standard (zero-drift) Brownian motion process. Realize that there are two sources of randomness in the jump-diffusion process. The first source is the Poisson Process dN_t , which causes the asset price to jump randomly. Once the asset price jumps, how much it jumps (the jump size) is also random. It is assumed that these two sources of randomness are independent of each other.

If the asset price does not jump in small time interval dt (i.e. $dN_t = 0$), then the jump-diffusion process is simply a Brownian motion motion with drift process:

$$\frac{dS_t}{S_t} = (\alpha - \lambda k)dt + \sigma dB_t.$$

If the asset price jumps in dt ($dN_t = 1$):

$$\frac{dS_t}{S_t} = (\alpha - \lambda k)dt + \sigma dB_t + (y_t - 1),$$

the relative price jump size is $y_t - 1$. Suppose that the lognormal random drawing y_t is 0.8, the asset price falls by 20%.

Let's solve SDE of (9.1). From (9.1), Merton jump-diffusion dynamics of an asset price is:

$$dS_t = (\alpha - \lambda k)S_t dt + \sigma S_t dB_t + (y_t - 1)S_t dN_t.$$

Appendix 10 gives the Itô formula for the jump-diffusion process as:

$$\begin{aligned} df(X_t, t) &= \frac{\partial f(X_t, t)}{\partial t} dt + b_t \frac{\partial f(X_t, t)}{\partial x} dt + \frac{\sigma_t^2}{2} \frac{\partial^2 f(X_t, t)}{\partial x^2} dt \\ &\quad + \sigma_t \frac{\partial f(X_t, t)}{\partial x} dB_t + [f(X_{t-} + \Delta X_t) - f(X_{t-})], \end{aligned}$$

where b_t corresponds to the drift term and σ_t corresponds to the volatility term of a jump-diffusion process $X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s + \sum_{i=1}^{N_t} \Delta X_i$. By applying this:

$$\begin{aligned} d \ln S_t &= \frac{\partial \ln S_t}{\partial t} dt + (\alpha - \lambda k)S_t \frac{\partial \ln S_t}{\partial S_t} dt + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 \ln S_t}{\partial S_t^2} dt \\ &\quad + \sigma S_t \frac{\partial \ln S_t}{\partial S_t} dB_t + [\ln y_t S_t - \ln S_t] \\ d \ln S_t &= (\alpha - \lambda k)S_t \frac{1}{S_t} dt + \frac{\sigma^2 S_t^2}{2} \left(-\frac{1}{S_t^2} \right) dt + \sigma S_t \frac{1}{S_t} dB_t + [\ln y_t + \ln S_t - \ln S_t] \\ d \ln S_t &= (\alpha - \lambda k)dt - \frac{\sigma^2}{2} dt + \sigma_t dB_t + \ln y_t \\ \ln S_t - \ln S_0 &= (\alpha - \frac{\sigma^2}{2} - \lambda k)(t - 0) + \sigma_t (B_t - B_0) + \sum_{i=1}^{N_t} \ln y_i \\ \ln S_t &= \ln S_0 + (\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma_t B_t + \sum_{i=1}^{N_t} \ln y_i \end{aligned}$$

$$\begin{aligned}\exp(\ln S_t) &= \exp\left\{\ln S_0 + (\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma_t B_t + \sum_{i=1}^{N_t} \ln y_i\right\} \\ S_t &= S_0 \exp\left\{\left(\alpha - \frac{\sigma^2}{2} - \lambda k\right)t + \sigma_t B_t\right\} \exp\left(\sum_{i=1}^{N_t} \ln y_i\right) \\ S_t &= S_0 \exp[(\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t] \prod_{i=1}^{N_t} y_i,\end{aligned}$$

or alternatively as:

$$S_t = S_0 \exp[(\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t + \sum_{i=1}^{N_t} \ln y_i].$$

Using the previous definition of the log price (return) jump size $\ln(y_i) \equiv Y_i$:

$$S_t = S_0 \exp[(\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t + \sum_{i=1}^{N_t} Y_i]. \quad (9.6)$$

This means that the asset price process $\{S_t; 0 \leq t \leq T\}$ is modeled as an exponential Lévy model of the form:

$$S_t = S_0 e^{X_t},$$

where X_t is a Lévy process which is categorized as a Brownian motion with drift (continuous part) plus a compound Poisson process (jump part) such that:

$$X_t = (\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t + \sum_{i=1}^{N_t} Y_i.$$

In other words, log-return $\ln(\frac{S_t}{S_0})$ is modeled as a Lévy process such that:

$$\ln(\frac{S_t}{S_0}) = X_t = (\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t + \sum_{i=1}^{N_t} Y_i.$$

Note that the compound Poisson jump process $\prod_{i=1}^{N_t} y_i = 1$ (in absolute price scale) or $\sum_{i=1}^{N_t} \ln y_i = \sum_{i=1}^{N_t} Y_i = 0$ (in log price scale) if $N_t = 0$ (i.e. no jumps between time 0 and t) or positive and negative jumps cancel each other out.

In the Black-Scholes case, log return $\ln(S_t / S_0)$ is normally distributed:

$$S_t = S_0 \exp\left[\left(\alpha - \frac{\sigma^2}{2}\right)t + \sigma B_t\right]$$

$$\ln\left(\frac{S_t}{S_0}\right) \sim \text{Normal}\left[\left(\alpha - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right].$$

But in jump-diffusion case, the existence of compound Poisson jump process $\sum_{i=1}^{N_t} Y_i$ makes log return non-normal. In Merton's case the simple distributional assumption about the log return jump size $(Y_i) \sim N(\mu, \delta^2)$ enables the probability density of log return $x_t = \ln(S_t / S_0)$ to be obtained as a quickly converging series of the following form:

$$\mathbb{P}(x_t \in A) = \sum_{i=0}^{\infty} \mathbb{P}(N_t = i) \mathbb{P}(x_t \in A | N_t = i)$$

$$\mathbb{P}(x_t) = \sum_{i=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^i}{i!} N(x_t; \left(\alpha - \frac{\sigma^2}{2} - \lambda k\right)t + i\mu, \sigma^2 t + i\delta^2) \quad (9.7)$$

where $N(x_t; \left(\alpha - \frac{\sigma^2}{2} - \lambda k\right)t + i\mu, \sigma^2 t + i\delta^2)$

$$= \frac{1}{\sqrt{2\pi(\sigma^2 t + i\delta^2)}} \exp\left[-\frac{\left[x_t - \left(\alpha - \frac{\sigma^2}{2} - \lambda k\right)t + i\mu\right]^2}{2(\sigma^2 t + i\delta^2)}\right].$$

The term $\mathbb{P}(N_t = i) = \frac{e^{-\lambda t} (\lambda t)^i}{i!}$ is the probability that the asset price jumps i times during

the time interval of length t . And $\mathbb{P}(x_t \in A | N_t = i) = N(x_t; \left(\alpha - \frac{\sigma^2}{2} - \lambda k\right)t + i\mu, \sigma^2 t + i\delta^2)$

is the Black-Scholes normal density of log-return assuming that the asset price jumps i times in the time interval of t . Therefore, the log-return density in the Merton jump-diffusion model can be interpreted as the weighted average of the Black-Scholes normal density by the probability that the asset price jumps i times.

By Fourier transforming the Merton log-return density function with FT parameters $(a, b,) = (1, 1)$, its characteristic function is calculated as:

$$\phi(\omega) = \int_{-\infty}^{\infty} \exp(i\omega x_t) \mathbb{P}(x_t) dx_t$$

$$= \exp \left[\lambda t \exp \left\{ \frac{1}{2} \omega (2i\mu - \delta^2 \omega) \right\} - \lambda t (1 + i\omega k) - \frac{1}{2} t \omega \{-2i\alpha + \sigma^2 (i + \omega)\} \right].$$

After simplification:

$$\phi(\omega) = \exp[t\psi(\omega)]$$

with the characteristic exponent (cumulant generating function):

$$\psi(\omega) = \lambda \left\{ \exp \left(i\omega \mu - \frac{\delta^2 \omega^2}{2} \right) - 1 \right\} + i\omega \left(\alpha - \frac{\sigma^2}{2} - \lambda k \right) - \frac{\sigma^2 \omega^2}{2}, \quad (9.8)$$

where $k \equiv e^{\frac{\mu+1}{2}\delta^2} - 1$. The characteristic exponent (9.8) can be alternatively obtained by substituting the Lévy measure of the Meron jump-diffusion model:

$$\begin{aligned} \ell(dx) &= \frac{\lambda}{\sqrt{2\pi\delta^2}} \exp \left\{ -\frac{(dx - \mu)^2}{2\delta^2} \right\} = \lambda f(dx) \\ f(dx) &\sim N(\mu, \delta^2) \end{aligned}$$

into the Lévy-Khintchin representation of the equation (6.6):

$$\begin{aligned} \psi(\omega) &= ib\omega - \frac{\sigma^2 \omega^2}{2} + \int_{-\infty}^{\infty} \{\exp(i\omega x) - 1\} \ell(dx) \\ \psi(\omega) &= ib\omega - \frac{\sigma^2 \omega^2}{2} + \int_{-\infty}^{\infty} \{\exp(i\omega x) - 1\} \lambda f(dx) \\ \psi(\omega) &= ib\omega - \frac{\sigma^2 \omega^2}{2} + \lambda \int_{-\infty}^{\infty} \{\exp(i\omega x) - 1\} f(dx) \\ \psi(\omega) &= ib\omega - \frac{\sigma^2 \omega^2}{2} + \lambda \left\{ \int_{-\infty}^{\infty} e^{i\omega x} f(dx) - \int_{-\infty}^{\infty} f(dx) \right\} \end{aligned}$$

Note that $\int_{-\infty}^{\infty} e^{i\omega x} f(dx)$ is the characteristic function of $f(dx)$:

$$\int_{-\infty}^{\infty} e^{i\omega x} f(dx) = \exp \left(i\mu\omega - \frac{\delta^2 \omega^2}{2} \right).$$

Therefore:

$$\psi(\omega) = ib\omega - \frac{\sigma^2\omega^2}{2} + \lambda \left\{ \exp\left(i\mu\omega - \frac{\delta^2\omega^2}{2}\right) - 1 \right\},$$

where $b = \alpha - \frac{\sigma^2}{2} - \lambda k$. This corresponds to (9.8). Characteristic exponent (9.8) generates cumulants as follows:

$$\begin{aligned} \text{cumulant}_1 &= \alpha - \frac{\sigma^2}{2} - \lambda k + \lambda\mu, \\ \text{cumulant}_2 &= \sigma^2 + \lambda\delta^2 + \lambda\mu^2, \\ \text{cumulant}_3 &= \lambda(3\delta^2\mu + \mu^3), \\ \text{cumulant}_4 &= \lambda(3\delta^4 + 6\mu^2\delta^2 + \mu^4). \end{aligned}$$

Annualized (per unit of time) mean, variance, skewness, and excess kurtosis of the log-return density $\mathbb{P}(x_t)$ are computed from above cumulants as follows:

$$\begin{aligned} E[x_t] &= \text{cumulant}_1 = \alpha - \frac{\sigma^2}{2} - \lambda \left(e^{\mu+\frac{1}{2}\delta^2} - 1 \right) + \lambda\mu \\ \text{Variance}[x_t] &= \text{cumulant}_2 = \sigma^2 + \lambda\delta^2 + \lambda\mu^2 \\ \text{Skewness}[x_t] &= \frac{\text{cumulant}_3}{(\text{cumulant}_2)^{3/2}} = \frac{\lambda(3\delta^2\mu + \mu^3)}{(\sigma^2 + \lambda\delta^2 + \lambda\mu^2)^{3/2}} \\ \text{Excess Kurtosis}[x_t] &= \frac{\text{cumulant}_4}{(\text{cumulant}_2)^{4/2}} = \frac{\lambda(3\delta^4 + 6\mu^2\delta^2 + \mu^4)}{(\sigma^2 + \lambda\delta^2 + \lambda\mu^2)^2}. \end{aligned} \quad (9.9)$$

We can observe several interesting properties of Merton's log-return density $\mathbb{P}(x_t)$. Firstly, the sign of μ which is the expected log-return jump size, $E[Y_t] = \mu$, determines the sign of skewness. The log-return density $\mathbb{P}(x_t)$ is negatively skewed if $\mu < 0$ and it is symmetric if $\mu = 0$ as illustrated in Figure 9.1.

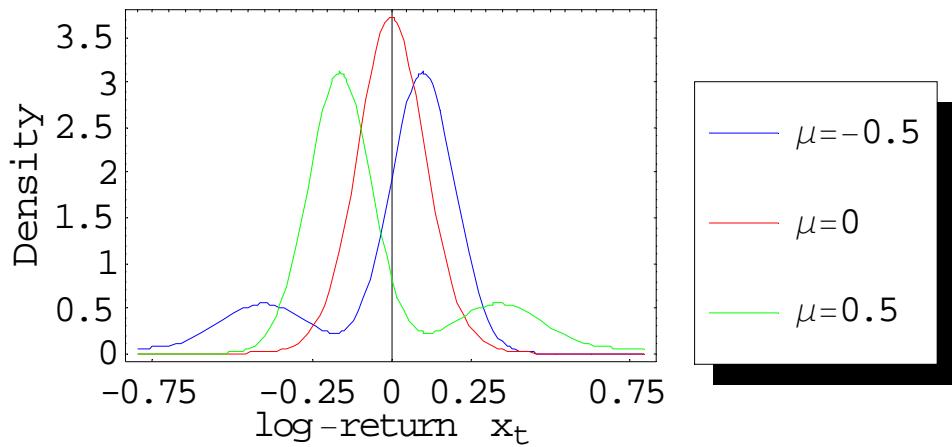


Figure 9.1: Merton's Log-Return Density for Different Values of μ . $\mu = -0.5$ in blue, $\mu = 0$ in red, and $\mu = 0.5$ in green. Parameters fixed are $\tau = 0.25$, $\alpha = 0.03$, $\sigma = 0.2$, $\lambda = 1$, and $\delta = 0.1$.

Table 9.1
Annualized Moments of Merton's Log-Return Density in Figure 9.1

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
$\mu = -0.5$	-0.0996	0.548	-0.852	0.864
$\mu = 0$	0.005	0.3742	0	0.12
$\mu = 0.5$	-0.147	0.5477	0.852	0.864

Secondly, larger value of intensity λ (which means that jumps are expected to occur more frequently) makes the density fatter-tailed as illustrated in Figure 9.2. Note that the excess kurtosis in the case $\lambda = 100$ is much smaller than in the case $\lambda = 1$ or $\lambda = 10$. This is because excess kurtosis is a standardized measure (by standard deviation).

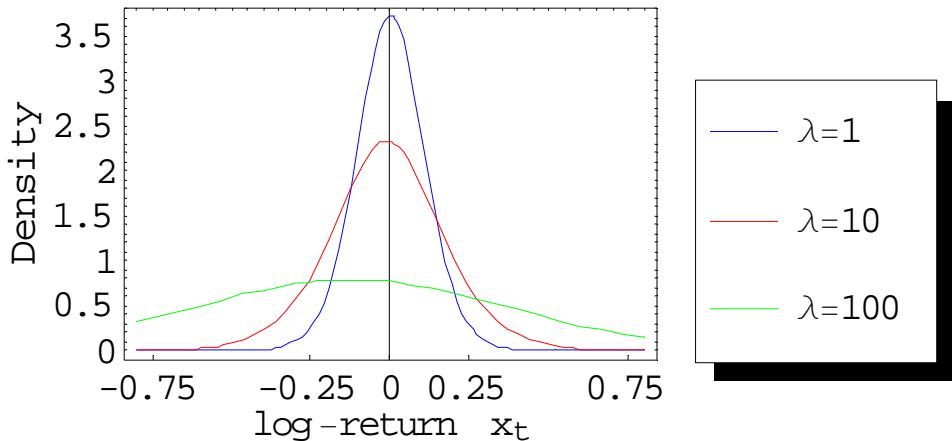


Figure 9.2: Merton's Log-Return Density for Different Values of Intensity λ . $\lambda = 1$ in blue, $\lambda = 10$ in red, and $\lambda = 100$ in green. Parameters fixed are $\tau = 0.25$, $\alpha = 0.03$, $\sigma = 0.2$, $\mu = 0$, and $\delta = 0.1$.

Table 9.2
Annualized Moments of Merton's Log-Return Density in Figure 9.2

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
$\lambda = 1$	0.00499	0.2236	0	0.12
$\lambda = 10$	-0.04012	0.3742	0	0.1531
$\lambda = 100$	-0.49125	1.0198	0	0.0277

Also note that Merton's log-return density has higher peak and fatter tails (more leptokurtic) when matched to the Black-Scholes normal counterpart as illustrated in Figure 9.3.

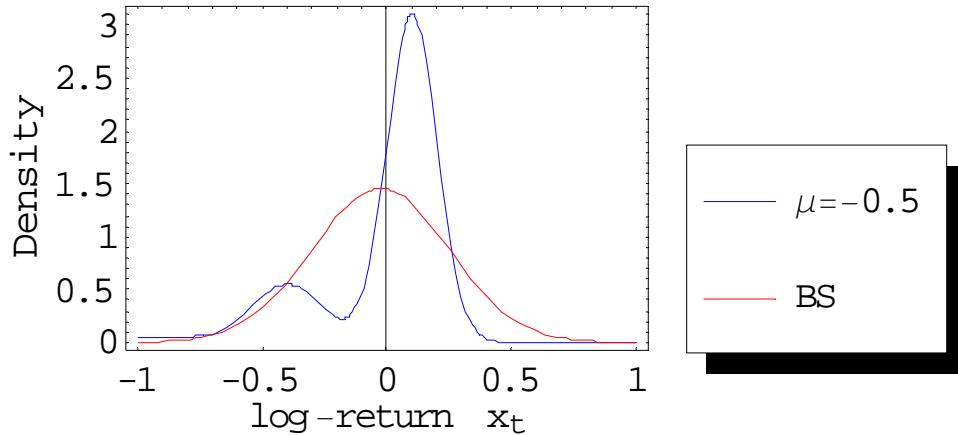


Figure 9.3: Merton Log-Return Density vs. Black-Scholes Log-Return Density (Normal). Parameters fixed for the Merton (in blue) are $\tau = 0.25$, $\alpha = 0.03$, $\sigma = 0.2$, $\lambda = 1$, $\mu = -0.5$, and $\delta = 0.1$. Black-Scholes normal log-return density is plotted (in red) by matching the mean and variance to the Merton.

Table 9.3
Annualized Moments of Merton vs. Black-Scholes Log-Return Density in Figure 9.3

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
Merton with $\mu = -0.5$	-0.0996	0.548	-0.852	0.864
Black-Scholes	-0.0996	0.548	0	0

[9.3] Log Stock Price Process for Merton Jump-Diffusion Model

Log stock price dynamics can be obtained from the equation (9.6) as:

$$\ln S_t = \ln S_0 + \left(\alpha - \frac{\sigma^2}{2} - \lambda k \right) t + \sigma B_t + \sum_{i=1}^{N_t} Y_i . \quad (9.10)$$

Probability density of log stock price $\ln S_t$ is obtained as a quickly converging series of the following form (i.e. conditionally normal):

$$\begin{aligned} \mathbb{P}(\ln S_t \in A) &= \sum_{i=0}^{\infty} \mathbb{P}(N_t = i) \mathbb{P}(\ln S_t \in A | N_t = i) \\ \mathbb{P}(\ln S_t) &= \sum_{i=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^i}{i!} N\left(\ln S_t; \ln S_0 + \left(\alpha - \frac{\sigma^2}{2} - \lambda k \right) t + i\mu, \sigma^2 t + i\delta^2\right), \end{aligned} \quad (9.11)$$

where:

$$\begin{aligned} &N\left(\ln S_t; \ln S_0 + \left(\alpha - \frac{\sigma^2}{2} - \lambda k \right) t + i\mu, \sigma^2 t + i\delta^2\right) \\ &= \frac{1}{\sqrt{2\pi(\sigma^2 t + i\delta^2)}} \exp\left[-\frac{\left\{ \ln S_t - \left(\ln S_0 + \left(\alpha - \frac{\sigma^2}{2} - \lambda k \right) t + i\mu \right) \right\}^2}{2(\sigma^2 t + i\delta^2)} \right]. \end{aligned}$$

By Fourier transforming (9.11) with FT parameters $(a, b) = (1, 1)$, its characteristic function is calculated as:

$$\begin{aligned} \phi(\omega) &= \int_{-\infty}^{\infty} \exp(i\omega \ln S_t) \mathbb{P}(\ln S_t) d \ln S_t \\ &= \exp\left[\lambda t \left(\exp(i\mu\omega - \frac{\delta^2\omega^2}{2}) - 1 \right) + i\omega \left(\ln S_0 + \left(\alpha - \frac{\sigma^2}{2} - \lambda k \right) t \right) - \frac{\sigma^2\omega^2}{2} t \right], \end{aligned} \quad (9.12)$$

where $k = e^{\frac{\mu+\delta^2}{2}} - 1$.

[9.4] Lévy Measure for Merton Jump-Diffusion Model

We saw in section 6.3 that Lévy measure $\ell(dx)$ of a compound Poisson process is given by the multiplication of the intensity and the jump size density $f(dx)$:

$$\ell(dx) = \lambda f(dx).$$

The Lévy measure $\ell(dx)$ represents the arrival rate (i.e. total intensity) of jumps of sizes $[x, x+dx]$. In other words, we can interpret the Lévy measure $\ell(dx)$ of a compound Poisson process as the measure of the average number of jumps per unit of time. Lévy measure is a positive measure on \mathbb{R} , but it is not a probability measure since its total mass λ (in the compound Poisson case) does not have to equal 1:

$$\int \ell(dx) = \lambda \in \mathbb{R}^+.$$

A Poisson process and a compound Poisson process (i.e. a piecewise constant Lévy process) are called finite activity Lévy processes since their Lévy measures $\ell(dx)$ are finite (i.e. the average number of jumps per unit time is finite):

$$\int_{-\infty}^{\infty} \ell(dx) < \infty.$$

In Merton jump-diffusion case, the log-return jump size is $(dx_i) \sim i.i.d. Normal(\mu, \delta^2)$:

$$f(dx_i) = \frac{1}{\sqrt{2\pi\delta^2}} \exp\left\{-\frac{(dx_i - \mu)^2}{2\delta^2}\right\}.$$

Therefore, the Lévy measure $\ell(dx)$ for Merton case can be expressed as:

$$\ell(dx) = \lambda f(dx) = \frac{\lambda}{\sqrt{2\pi\delta^2}} \exp\left\{-\frac{(dx - \mu)^2}{2\delta^2}\right\}. \quad (9.13)$$

An example of Lévy measure $\ell(dx)$ for the log-return $x_t = \ln(S_t / S_0)$ in the Merton jump-diffusion model is plotted in Figure 9.5. Each Lévy measure is symmetric (i.e. $\mu = 0$ is used) with total mass 1, 2, and 4 respectively.

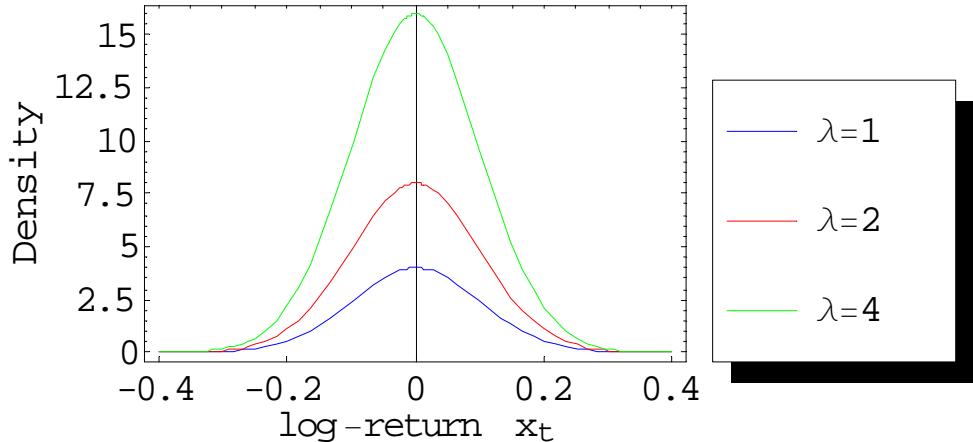


Figure 9.5: Lévy Measures $\ell(dx)$ for the Log-Return $x_t = \ln(S_t / S_0)$ in the Merton Jump-Diffusion Model for Different Values of Intensity λ . Parameters used are $\mu = 0$ and $\delta = 0.1$.

[9.5] Option (Contingent Claim) Pricing: PDE Approach by Hedging

Consider a portfolio P of the one long option position $V(S, t)$ on the underlying asset S written at time t and a short position of the underlying asset in quantity Δ to derive option pricing functions in the presence of jumps:

$$P_t = V(S_t, t) - \Delta S_t. \quad (9.14)$$

Portfolio value changes by in a very short period of time:

$$dP_t = dV(S_t, t) - \Delta dS_t. \quad (9.15)$$

Merton jump-diffusion dynamics of an asset price is given by equation (9.1) in the differential form as:

$$\begin{aligned} \frac{dS_t}{S_t} &= (\alpha - \lambda k)dt + \sigma dB_t + (y_t - 1)dN_t, \\ dS_t &= (\alpha - \lambda k)S_t dt + \sigma S_t dB_t + (y_t - 1)S_t dN_t. \end{aligned} \quad (9.16)$$

Appendix 8 gives the Itô formula for the jump-diffusion process as:

$$\begin{aligned} df(X_t, t) &= \frac{\partial f(X_t, t)}{\partial t} dt + b_t \frac{\partial f(X_t, t)}{\partial x} dt + \frac{\sigma_t^2}{2} \frac{\partial^2 f(X_t, t)}{\partial x^2} dt \\ &\quad + \sigma_t \frac{\partial f(X_t, t)}{\partial x} dB_t + [f(X_{t-} + \Delta X_t) - f(X_{t-})], \end{aligned}$$

where b_t corresponds to the drift term and σ_t corresponds to the volatility term of a jump-diffusion process $X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s + \sum_{i=1}^{N_t} \Delta X_i$. Apply this to our case of option price function $V(S, t)$:

$$\begin{aligned} dV(S_t, t) &= \frac{\partial V}{\partial t} dt + (\alpha - \lambda k)S_t \frac{\partial V}{\partial S_t} dt + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} dt + \sigma S_t \frac{\partial V}{\partial S_t} dB_t \\ &\quad + [V(y_t S_t, t) - V(S_t, t)]dN_t. \end{aligned} \quad (9.17)$$

The term $[V(y_t S_t, t) - V(S_t, t)]dN_t$ describes the difference in the option value when a jump occurs. Now the change in the portfolio value can be expressed as by substituting (9.16) and (9.17) into (9.15):

$$\begin{aligned}
dP_t &= dV(S_t, t) - \Delta dS_t \\
dP_t &= \frac{\partial V}{\partial t} dt + (\alpha - \lambda k) S_t \frac{\partial V}{\partial S_t} dt + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} dt + \sigma S_t \frac{\partial V}{\partial S_t} dB_t + [V(y_t S_t, t) - V(S_t, t)] dN_t \\
&\quad - \Delta \{(\alpha - \lambda k) S_t dt + \sigma S_t dB_t + (y_t - 1) S_t dN_t\} \\
dP_t &= \left\{ \frac{\partial V}{\partial t} + (\alpha - \lambda k) S_t \frac{\partial V}{\partial S_t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} - \Delta (\alpha - \lambda k) S_t \right\} dt + \left(\sigma S_t \frac{\partial V}{\partial S_t} - \Delta \sigma S_t \right) dB_t \\
&\quad + \{V(y_t S_t, t) - V(S_t, t) - \Delta (y_t - 1) S_t\} dN_t. \tag{9.18}
\end{aligned}$$

If there is no jump between time 0 and t (i.e. $dN_t = 0$), the problem reduces to Black-Scholes case in which setting $\Delta = \partial V / \partial S_t$ makes the portfolio risk-free leading to the following (i.e. the randomness dB_t has been eliminated):

$$\begin{aligned}
dP_t &= \left\{ \frac{\partial V}{\partial t} + (\alpha - \lambda k) S_t \frac{\partial V}{\partial S_t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} - \frac{\partial V}{\partial S_t} (\alpha - \lambda k) S_t \right\} dt + \left(\sigma S_t \frac{\partial V}{\partial S_t} - \frac{\partial V}{\partial S_t} \sigma S_t \right) dB_t \\
dP_t &= \left\{ \frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} \right\} dt.
\end{aligned}$$

This in turn means that if there is a jump between time 0 and t (i.e. $dN_t \neq 0$), setting $\Delta = \partial V / \partial S_t$ does not eliminate the risk. Suppose we decided to hedge the randomness caused by diffusion part dB_t in the underlying asset price (which are always present) and not to hedge the randomness caused by jumps dN_t (which occur infrequently) by setting $\Delta = \partial V / \partial S_t$. Then, the change in the value of the portfolio is given by from equation (9.18):

$$\begin{aligned}
dP_t &= \left\{ \frac{\partial V}{\partial t} + (\alpha - \lambda k) S_t \frac{\partial V}{\partial S_t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} - \frac{\partial V}{\partial S_t} (\alpha - \lambda k) S_t \right\} dt + \left(\sigma S_t \frac{\partial V}{\partial S_t} - \frac{\partial V}{\partial S_t} \sigma S_t \right) dB_t \\
&\quad + \left\{ V(y_t S_t, t) - V(S_t, t) - \frac{\partial V}{\partial S_t} (y_t - 1) S_t \right\} dN_t, \\
dP_t &= \left\{ \frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} \right\} dt + \left\{ V(y_t S_t, t) - V(S_t, t) - \frac{\partial V}{\partial S_t} (y_t - 1) S_t \right\} dN_t. \tag{9.19}
\end{aligned}$$

Merton argues that the jump component (dN_t) of the asset price process S_t is uncorrelated with the market as a whole. Then, the risk of jump is diversifiable (non-systematic) and it should earn no risk premium. Therefore, the portfolio is expected to grow at the risk-free interest rate r :

$$E[dP_t] = rP_t dt. \quad (9.20)$$

After substitution of (9.14) and (9.19) into (9.20) by setting $\Delta = \partial V / \partial S_t$:

$$\begin{aligned} E\left[\left\{\frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2}\right\} dt + \left\{V(y_t S_t, t) - V(S_t, t) - \frac{\partial V}{\partial S_t}(y_t - 1)S_t\right\} dN_t\right] &= r\{V(S_t, t) - \Delta S_t\} dt \\ \left(\frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2}\right) dt + E[V(y_t S_t, t) - V(S_t, t) - \frac{\partial V}{\partial S_t}(y_t - 1)S_t] E[dN_t] &= r\{V(S_t, t) - \frac{\partial V}{\partial S_t} S_t\} dt \\ \left(\frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2}\right) dt + E[V(y_t S_t, t) - V(S_t, t) - \frac{\partial V}{\partial S_t}(y_t - 1)S_t] \lambda dt &= r\{V(S_t, t) - \frac{\partial V}{\partial S_t} S_t\} dt \\ \frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} + \lambda E[V(y_t S_t, t) - V(S_t, t) - \frac{\partial V}{\partial S_t}(y_t - 1)S_t] &= r\{V(S_t, t) - \frac{\partial V}{\partial S_t} S_t\} \end{aligned}$$

Thus, the Merton jump-diffusion counterpart of Black-Scholes PDE is:

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} + rS_t \frac{\partial V}{\partial S_t} - rV + \lambda E[V(y_t S_t, t) - V(S_t, t)] - \lambda S_t \frac{\partial V}{\partial S_t} E[y_t - 1] = 0. \quad (9.21)$$

where the term $E[V(y_t S_t, t) - V(S_t, t)]$ involves the expectation operator and

$E[y_t - 1] = e^{\mu + \frac{1}{2}\delta^2} - 1 \equiv k$ (which is the mean of relative asset price jump size). Obviously, if jump is not expected to occur (i.e. $\lambda = 0$), this reduces to Black-Scholes PDE:²

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S_t^2}{2} \frac{\partial^2 V}{\partial S_t^2} + rS_t \frac{\partial V}{\partial S_t} - rV = 0.$$

Merton's simple assumption that the absolute price jump size is lognormally distributed (i.e. the log-return jump size is normally distributed, $Y_t \equiv \ln(y_t) \sim N(\mu, \delta^2)$) makes it possible to solve the jump-diffusion PDE to obtain the following price function of European vanilla options as a quickly converging series of the form:

$$\sum_{i=0}^{\infty} \frac{e^{-\bar{\lambda}\tau} (\bar{\lambda}\tau)^i}{i!} V_{BS}(S_t, \tau = T - t, \sigma_i, r_i), \quad (9.22)$$

where $\bar{\lambda} = \lambda(1+k) = \lambda e^{\mu + \frac{1}{2}\delta^2}$,

$$\sigma_i^2 = \sigma^2 + \frac{i\delta^2}{\tau},$$

² This equation not only contains local derivatives but also links together option values at discontinuous values in S. This is called non-local nature.

$$r_i = r - \lambda k + \frac{i \ln(1+k)}{\tau} = r - \lambda(e^{\frac{\mu+1}{2}\delta^2} - 1) + \frac{i(\mu + \frac{1}{2}\delta^2)}{\tau},$$

and V_{BS} is the Black-Scholes price without jumps.

Thus, Merton's jump-diffusion option price can be interpreted as the weighted average of the Black-Scholes price conditional on that the underlying asset price jumps i times to the expiry with weights being the probability that the underlying jumps i times to the expiry.

[9.6] Option (Contingent Claim) Pricing: Martingale Approach

Let $\{B_t; 0 \leq t \leq T\}$ be a standard Brownian motion process on a space $(\Omega, \mathcal{F}, \mathbb{P})$. Under actual probability measure \mathbb{P} , the dynamics of Merton jump-diffusion asset price process is given by equation (9.6) in the integral form:

$$S_t = S_0 \exp[(\alpha - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t + \sum_{k=1}^{N_t} Y_k] \text{ under } \mathbb{P}.$$

We changed the index from $\sum_{i=1}^{N_t} Y_i$ to $\sum_{k=1}^{N_t} Y_k$. This is trivial but readers will find the reason soon. Merton jump-diffusion model is an example of an incomplete model because there are many equivalent martingale risk-neutral measures $\mathbb{Q} \sim \mathbb{P}$ under which the discounted asset price process $\{e^{-rt} S_t; 0 \leq t \leq T\}$ becomes a martingale. Merton finds his equivalent martingale risk-neutral measure $\mathbb{Q}_M \sim \mathbb{P}$ by changing the drift of the Brownian motion process while keeping the other parts (most important is the jump measure, i.e. the distribution of jump times and jump sizes) unchanged:

$$S_t = S_0 \exp[(r - \frac{\sigma^2}{2} - \lambda k)t + \sigma B_t^{\mathbb{Q}_M} + \sum_{k=1}^{N_t} Y_k] \text{ under } \mathbb{Q}_M. \quad (9.23)$$

Note that $B_t^{\mathbb{Q}_M}$ is a standard Brownian motion process on $(\Omega, \mathcal{F}, \mathbb{Q}_M)$ and the process $\{e^{-rt} S_t; 0 \leq t \leq T\}$ is a martingale under \mathbb{Q}_M . Then, a European option price $V^{Merton}(t, S_t)$ with payoff function $H(S_T)$ is calculated as:

$$V^{Merton}(t, S_t) = e^{-r(T-t)} E^{\mathbb{Q}_M}[H(S_T) | \mathcal{F}_t]. \quad (9.24)$$

Standard assumption is $\mathcal{F}_t = S_t$, thus:

$$V^{Merton}(t, S_t) = e^{-r(T-t)} E^{\mathbb{Q}_M}[H(S_t \exp[(r - \frac{\sigma^2}{2} - \lambda k)(T-t) + \sigma B_{T-t}^{\mathbb{Q}_M} + \sum_{k=1}^{N_{T-t}} Y_k]) | S_t]$$

$$V^{Merton}(t, S_t) = e^{-r(T-t)} E^{\mathbb{Q}_M} [H(S_t \exp[(r - \frac{\sigma^2}{2} - \lambda k)(T-t) + \sigma B_{T-t}^{\mathbb{Q}_M} + \sum_{k=1}^{N_{T-t}} Y_k])]. \quad (9.25)$$

Poisson counter is (we would like to use index i for the number of jumps):

$$N_{T-t} = 0, 1, 2, \dots \equiv i.$$

And the compound Poisson process is distributed as:

$$\sum_{k=1}^{N_{T-t}} Y_k \sim Normal(i\mu, i\delta^2).$$

Thus, $V^{Merton}(t, S_t)$ can be expressed as from equation (9.25) (i.e. by conditioning on i):

$$\begin{aligned} V^{Merton}(t, S_t) \\ = e^{-r(T-t)} \sum_{i \geq 0} \mathbb{Q}_M(N_{T-t} = i) E^{\mathbb{Q}_M} [H(S_t \exp[(r - \frac{\sigma^2}{2} - \lambda k)(T-t) + \sigma B_{T-t}^{\mathbb{Q}_M} + \sum_{k=1}^i Y_k])]. \end{aligned}$$

Use $\tau = T - t$:

$$\begin{aligned} V^{Merton}(t, S_t) \\ = e^{-r\tau} \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} E^{\mathbb{Q}_M} \left[H \left(S_t \exp[\{r - \frac{\sigma^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\}\tau + \sigma B_\tau^{\mathbb{Q}_M} + \sum_{k=1}^i Y_k] \right) \right]. \quad (9.26) \end{aligned}$$

Inside the exponential function is normally distributed:

$$\begin{aligned} \{r - \frac{\sigma^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\}\tau + \sigma B_\tau^{\mathbb{Q}_M} + \sum_{k=1}^i Y_k \\ \sim Normal \left(\{r - \frac{\sigma^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\}\tau + i\mu, \sigma^2\tau + i\delta^2 \right). \end{aligned}$$

Rewrite it so that its distribution remains the same:

$$\begin{aligned} \{r - \frac{\sigma^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\}\tau + i\mu + \sqrt{\frac{\sigma^2\tau + i\delta^2}{\tau}} B_\tau^{\mathbb{Q}_M} \\ \sim Normal \left(\{r - \frac{\sigma^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\}\tau + i\mu, \sigma^2\tau + i\delta^2 \right). \end{aligned}$$

Now we can rewrite equation (9.24) as (we can do this operation because a normal density is uniquely determined by only two parameters: its mean and variance):

$$V^{Merton}(t, S_t) = e^{-rt} \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} E^{\mathbb{Q}_M} \left[H \left(S_t \exp \left[\left\{ r - \frac{\sigma^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\right\}\tau + i\mu + \sqrt{\frac{\sigma^2\tau + i\delta^2}{\tau}} B_{\tau}^{\mathbb{Q}_M} \right] \right) \right]$$

We can always add $(\frac{i\delta^2}{2\tau} - \frac{i\delta^2}{2\tau}) = 0$ inside the exponential function:

$$V^{Merton}(t, S_t) = e^{-rt} \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} \times \\ E^{\mathbb{Q}_M} \left[H \left(S_t \exp \left[\left\{ r - \frac{\sigma^2}{2} + \left(\frac{i\delta^2}{2\tau} - \frac{i\delta^2}{2\tau} \right) - \lambda(e^{\mu+\delta^2/2} - 1)\right\}\tau + i\mu + \sqrt{\sigma^2 + \frac{i\delta^2}{\tau}} B_{\tau}^{\mathbb{Q}_M} \right] \right) \right]$$

$$V^{Merton}(t, S_t) = e^{-rt} \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} \times \\ E^{\mathbb{Q}_M} \left[H \left(S_t \exp \left[\left\{ r - \frac{1}{2}(\sigma^2 + \frac{i\delta^2}{\tau}) + \frac{i\delta^2}{2\tau} - \lambda(e^{\mu+\delta^2/2} - 1)\right\}\tau + i\mu + \sqrt{\sigma^2 + \frac{i\delta^2}{\tau}} B_{\tau}^{\mathbb{Q}_M} \right] \right) \right]$$

Set $\sigma_i^2 = \sigma^2 + \frac{i\delta^2}{\tau}$ and rearrange:

$$V^{Merton}(t, S_t) = e^{-rt} \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} E^{\mathbb{Q}_M} \left[H \left(S_t \exp \left[\left\{ r - \frac{1}{2}\sigma_i^2 + \frac{i\delta^2}{2\tau} - \lambda(e^{\mu+\delta^2/2} - 1)\right\}\tau + i\mu + \sigma_i B_{\tau}^{\mathbb{Q}_M} \right] \right) \right]$$

$$= e^{-rt} \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} E^{\mathbb{Q}_M} \left[H \left(S_t \exp \left\{ i\mu + \frac{i\delta^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\tau \right\} \exp \left\{ \left(r - \frac{1}{2}\sigma_i^2 \right)\tau + \sigma_i B_{\tau}^{\mathbb{Q}_M} \right\} \right) \right].$$

Black-Scholes price can be expressed as:

$$V^{BS}(\tau = T - t, S_t; \sigma) = e^{-rt} E^{\mathbb{Q}_{BS}} [H \{S_t \exp(r - \frac{1}{2}\sigma^2)\tau + \sigma B_{\tau}^{\mathbb{Q}_{BS}}\}].$$

Finally, Merton jump-diffusion pricing formula can be obtained as a weighted average of Black-Scholes price conditioned on the number of jumps i :

$$V^{Merton}(t, S_t) = \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} V^{BS}(\tau, S_i \equiv S_t \exp \{i\mu + \frac{i\delta^2}{2} - \lambda(e^{\mu+\delta^2/2} - 1)\tau\}; \sigma_i \equiv \sqrt{\sigma^2 + \frac{i\delta^2}{\tau}}). \quad (9.27)$$

Alternatively:

$$\begin{aligned}
V^{Merton}(t, S_t) &= e^{-rt} \sum_{i \geq 0} \frac{e^{-\lambda\tau} (\lambda\tau)^i}{i!} E^{\mathbb{Q}_M} \left[H \left\{ S_t \exp \left\{ \left(r - \lambda(e^{\mu+\delta^2/2} - 1) + \frac{i\mu + i\delta^2/2}{\tau} - \frac{1}{2}\sigma_i^2 \right) \tau + \sigma_i B_\tau^{\mathbb{Q}_M} \right\} \right] \\
&= \sum_{i \geq 0} \frac{e^{-\bar{\lambda}\tau} (\bar{\lambda}\tau)^i}{i!} V^{BS}(\tau, S_t; \sigma_i \equiv \sqrt{\sigma^2 + \frac{i\delta^2}{\tau}}, r_i \equiv r - \lambda(e^{\mu+\delta^2/2} - 1) + \frac{i\mu + i\delta^2/2}{\tau}). \quad (9.28)
\end{aligned}$$

where $\bar{\lambda} = \lambda(1+k) = \lambda e^{\mu+\frac{1}{2}\delta^2}$. As you might notice, this is the same result as the option pricing formula derived from solving a PDE by forming a risk-free portfolio in equation (9.22). PDE approach and Martingale approach are different approaches but they are related and give the same result.

[9.7] Option Pricing Example of Merton Jump-Diffusion Model

In this section the equation (9.22) is used to price hypothetical plain vanilla options: current stock price $S_t = 50$, risk-free interest rate $r = 0.05$, continuously compounded dividend yield $q = 0.02$, time to maturity $\tau = 0.25$ years.

We need to be careful about volatility σ . In the Black-Scholes case, the t -period standard deviation of log-return x_t is:

$$\text{Standard Deviation}_{BS}(x_t) = \sigma_{BS} \sqrt{t}. \quad (9.29)$$

Equation (9.10) tells that the t -period standard deviation of log-return x_t in the Merton model is given as:

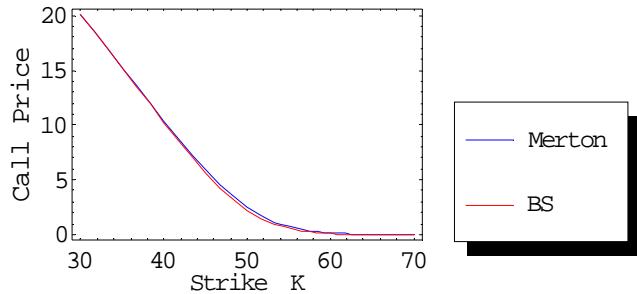
$$\text{Standard Deviation}_{Merton}(x_t) = \sqrt{(\sigma_{Merton}^2 + \lambda\delta^2 + \lambda\mu^2)t}. \quad (9.30)$$

This means that if we set $\sigma_{BS} = \sigma_{Merton}$, Merton jump-diffusion prices are always greater (or equal to) than Black-Scholes prices because of the extra source of volatility λ (intensity), μ (mean log-return jump size), δ (standard deviation of log-return jump) (i.e. larger volatility is translated to larger option price):

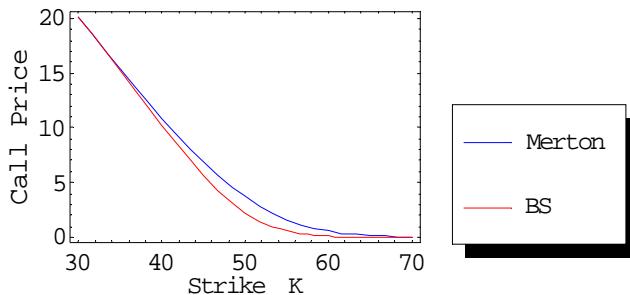
$$\begin{aligned}
\text{Standard Deviation}_{BS}(x_t) &\leq \text{Standard Deviation}_{Merton}(x_t) \\
\sigma_{BS} \sqrt{t} &\leq \sqrt{(\sigma_{Merton}^2 + \lambda\delta^2 + \lambda\mu^2)t}.
\end{aligned}$$

This very obvious point is illustrated in Figure 9.6 where diffusion volatility is set $\sigma_{BS} = \sigma_{Merton} = 0.2$. Note the followings: (1) In all four panels Merton jump-diffusion price is always greater (or equal to) than Black-Scholes price. (2) When Merton parameters (λ , μ , and δ)

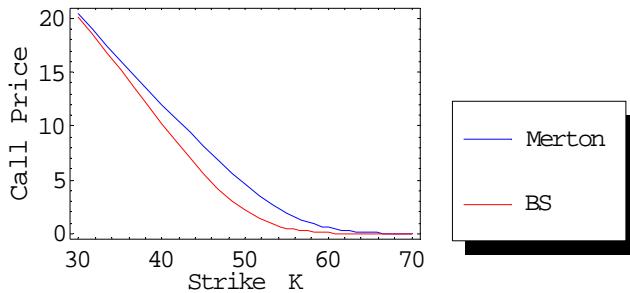
are small in Panel A, the difference between these two prices is small. (3) As intensity λ increases (i.e. increased expected number of jumps per unit of time), the t -period Merton standard deviation of log-return x_t increases (equation (9.30)) leading to the larger difference between Merton price and Black-Scholes price as illustrated in Panel B. (4) As Merton mean log-return jump size μ increases, the t -period Merton standard deviation of log-return x_t increases (equation (9.30)) leading to the larger difference between Merton price and Black-Scholes price as illustrated in Panel C. (5) As Merton standard deviation of log-return jump size δ increases, the t -period Merton standard deviation of log-return x_t increases (equation (9.30)) leading to the larger difference between Merton price and Black-Scholes price as illustrated in Panel D.



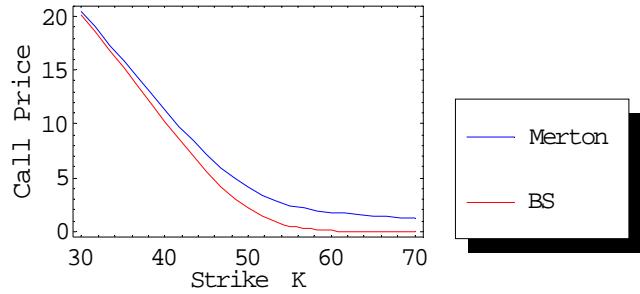
A) Merton parameters: $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$.



B) Merton parameters: $\lambda = 5$, $\mu = -0.1$, and $\delta = 0.1$.



C) Merton parameters: $\lambda = 1$, $\mu = -0.5$, and $\delta = 0.1$.



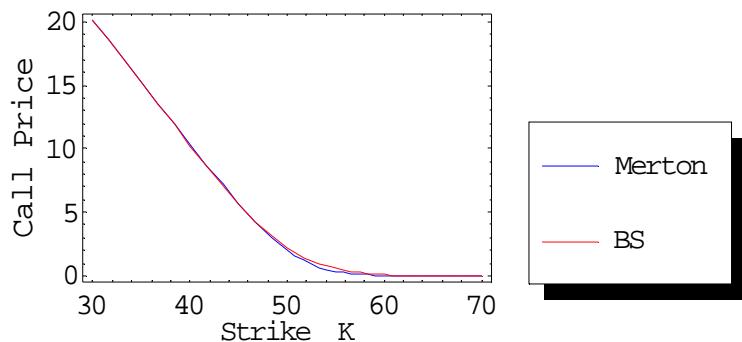
D) Merton parameters: $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.5$.

Figure 9.6: Merton Jump-Diffusion Call Price vs. Black-Scholes Call Price When Diffusion Volatility σ is same. Parameters and variables used are $S_t = 50$, $r = 0.05$, $q = 0.02$, $\tau = 0.25$, and $\sigma_{BS} = \sigma_{Merton} = 0.2$.

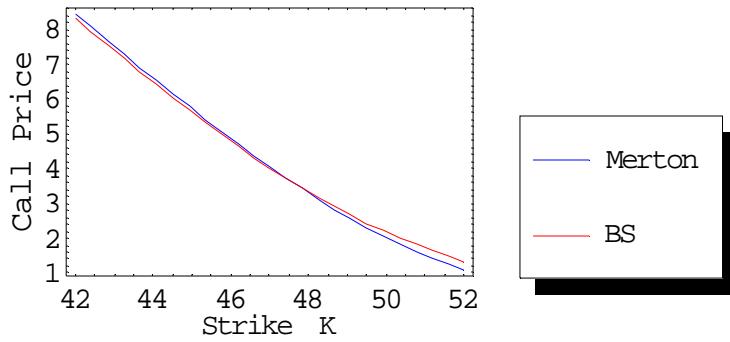
Next we consider a more delicate case where we restrict diffusion volatilities σ_{BS} and σ_{Merton} such that standard deviations of Merton jump-diffusion and Black-Scholes log-return densities are the same:

$$\text{Standard Deviation}_{BS}(x_t) = \text{Standard Deviation}_{Merton}(x_t), \\ \sigma_{BS} \sqrt{t} = \sqrt{(\sigma_{Merton}^2 + \lambda\delta^2 + \lambda\mu^2)t}.$$

Using the Merton parameters $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$ and Black-Scholes volatility $\sigma_{BS} = 0.2$, Merton diffusion volatility is calculated as $\sigma_{Merton} = 0.141421$. In this same standard deviation case, call price function is plotted in Figure 9.7 and put price function is plotted in Figure 9.8. It seems that Merton jump diffusion model overestimates in-the-money call and underestimates out-of-money call when compared to Black-Scholes model. And Merton jump diffusion model overestimates out-of-money put and underestimates in-the-money put when compared to Black-Scholes model.

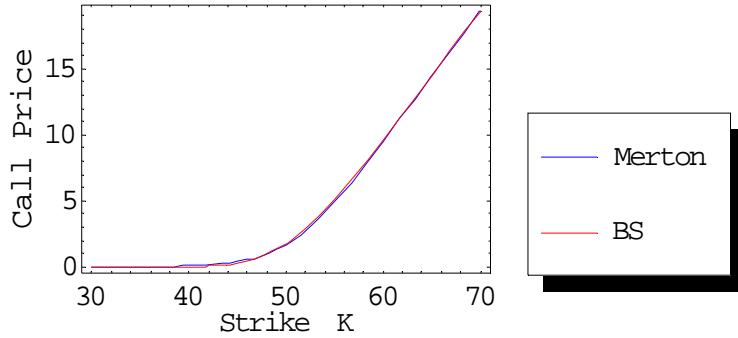


A) Range 30 to 70.

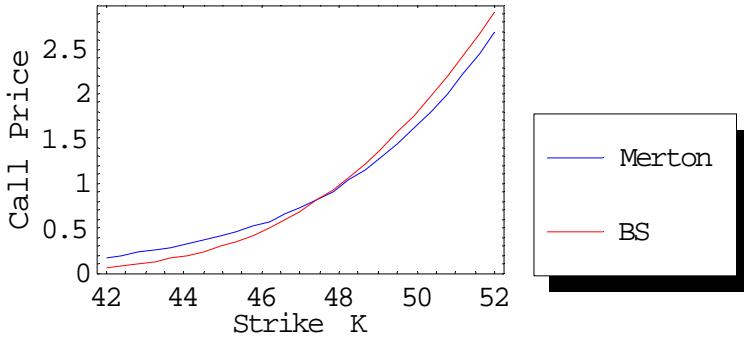


B) Range 42 to 52.

Figure 9.7: Merton Jump-Diffusion Call Price vs. Black-Scholes Call Price When Restricting Merton Diffusion Volatility σ_{Merton} . We set $\sigma_{BS} = 0.2$ and $\sigma_{Merton} = 0.141421$. Parameters and variables used are $S_t = 50$, $r = 0.05$, $q = 0.02$, $\tau = 0.25$.



A) Range 30 to 70.



B) Range 42 to 52.

Figure 9.8: Merton Jump-Diffusion Put Price vs. Black-Scholes Put Price When Restricting Merton Diffusion Volatility σ_{Merton} . We set $\sigma_{BS} = 0.2$ and $\sigma_{Merton} = 0.141421$. Parameters and variables used are $S_t = 50$, $r = 0.05$, $q = 0.02$, $\tau = 0.25$.

[10] Merton (1976) Jump-Diffusion Model with Fourier Transform Pricing

We saw in section 9 that the most important character of Merton JD model is its conditional Black-Scholes property. Because Merton JD model is a simple addition of compound Poisson jump process to the original BS model (i.e. geometric Brownian motion with drift), everything in the Merton JD model such as log return density $\ln(S_t / S_0)$ or even call price can be expressed as the weighted average of the BS (log return density or price) conditional on that the underlying asset price jumps i times to the expiry with weights being the probability that the underlying jumps i times to the expiry. Traditional pricing approach with Merton JD model is discussed in the section 9.5 (PDE approach by hedging) and 9.6 (martingale approach).

In this section we present Fourier transform option pricing approach with Merton JD model. You will see FT method is much more general and simpler.

[10.1] Merton JD Model with Fourier Transform Pricing Method

The first step to FT option pricing is to obtain the characteristic function of the log stock price $\ln S_t$. Similar arguments can be seen in section 9.3. Risk-neutral log stock price dynamics can be obtained from the equation (9.6) as:

$$\ln S_t = \ln S_0 + \left(r - \frac{\sigma^2}{2} - \lambda k \right) t + \sigma B_t + \sum_{i=1}^{N_t} Y_i, \quad (10.1)$$

where $k \equiv e^{\mu + \frac{1}{2}\delta^2} - 1$. Probability density of log stock price $\ln S_t$ is obtained as a quickly converging series of the following form (i.e. conditionally normal):

$$\begin{aligned} \mathbb{P}(\ln S_t \in A) &= \sum_{i=0}^{\infty} \mathbb{P}(N_t = i) \mathbb{P}(\ln S_t \in A | N_t = i) \\ \mathbb{P}(\ln S_t) &= \sum_{i=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^i}{i!} N\left(\ln S_t; \ln S_0 + \left(r - \frac{\sigma^2}{2} - \lambda k\right)t + i\mu, \sigma^2 t + i\delta^2\right), \end{aligned} \quad (10.2)$$

where:

$$\begin{aligned} &N\left(\ln S_t; \ln S_0 + \left(r - \frac{\sigma^2}{2} - \lambda k\right)t + i\mu, \sigma^2 t + i\delta^2\right) \\ &= \frac{1}{\sqrt{2\pi(\sigma^2 t + i\delta^2)}} \exp\left[-\frac{\left\{\ln S_t - \left(\ln S_0 + \left(r - \frac{\sigma^2}{2} - \lambda k\right)t + i\mu\right)\right\}^2}{2(\sigma^2 t + i\delta^2)}\right]. \end{aligned}$$

Let $s_t \equiv \ln S_t$ as before. By Fourier transforming (10.2) with FT parameters $(a, b,) = (1, 1)$, its characteristic function is calculated as:

$$\begin{aligned}\phi(\omega) &= \int_{-\infty}^{\infty} e^{i\omega s_t} \mathbb{P}(s_t) ds_t \\ &= \exp \left[\lambda t \left(\exp(i\mu\omega - \frac{\delta^2\omega^2}{2}) - 1 \right) + i\omega \left(s_0 + (r - \frac{\sigma^2}{2} - \lambda k)t \right) - \frac{\sigma^2\omega^2}{2}t \right]. \quad (10.3)\end{aligned}$$

By substituting a characteristic function of Merton log terminal stock price s_T , the equation (10.3), into the general FT pricing formula of the equation (8.17), we obtain Merton-FT call pricing formula:

$$C_{Merton-FT}(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha + 1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha + 1)\omega} d\omega, \quad (10.4)$$

where:

$$\phi_T(\omega) = \exp \left[\lambda T \left(\exp(i\mu\omega - \frac{\delta^2\omega^2}{2}) - 1 \right) + i\omega \left(s_0 + (r - \frac{\sigma^2}{2} - \lambda k)T \right) - \frac{\sigma^2\omega^2}{2}T \right].$$

We implement the Merton-FT formula (10.4) with decay rate parameter $\alpha = 1$ and compare the result to the original Merton call price of the equation (9.22) using common parameters and variables fixed $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$. As illustrated by Figure 10.1, as a principle Merton-FT call price and the original Merton call price are identical. This is no surprise because the original Merton formula (9.22) and Merton-FT formula (10.4) are the same person with a different look. Merton-FT formula is just frequency representation of the original Merton formula.

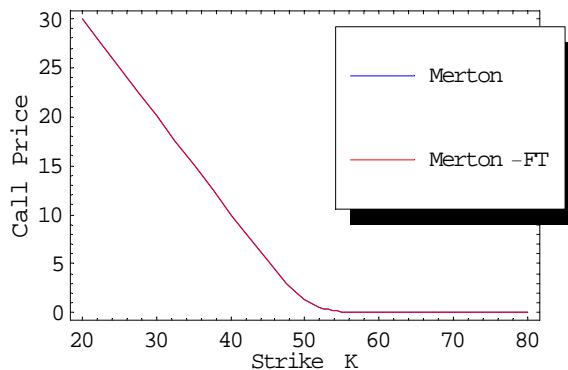


Figure 10.1: Original Merton Call Price Vs. Merton-FT with $\alpha = 1$ Call Price.

Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$.

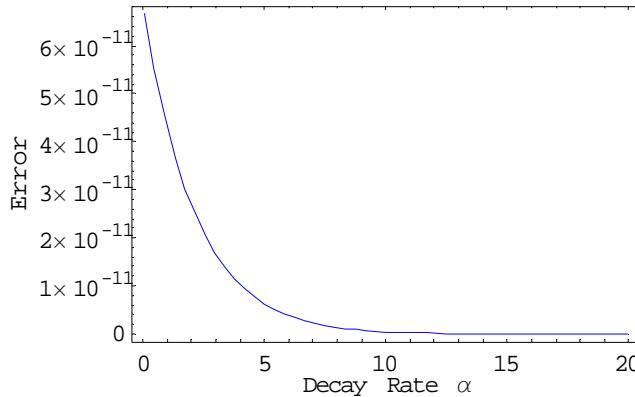
Next, CPU time should be discussed. Consider call options with $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 20/252$ as a function of varying strike price K . We use Merton JD parameters $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$. Table 10.1 reveals that although Merton-FT formula is slower than the original Merton, speed is not an issue for most of the purposes.

Table 10.1: CPU Time for Calls with Different Moneyness

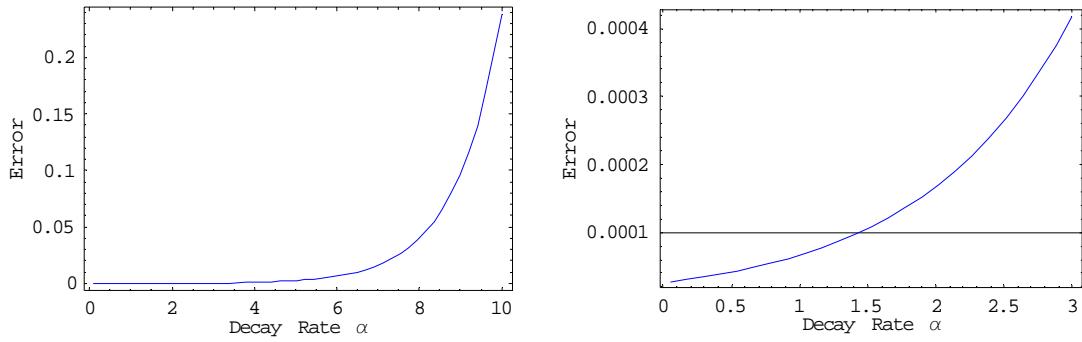
Price is in the bracket.

Method	Strike		
	$K = 20$	$K = 50$	$K = 80$
Merton	0.01 seconds (29.9999)	0.01 seconds (1.32941)	0.01 seconds (1.19634×10^{-7})
Merton-FT $\alpha = 1$	0.06 seconds (29.9999)	0.02 seconds (1.32941)	0.12 seconds (1.19634×10^{-7})

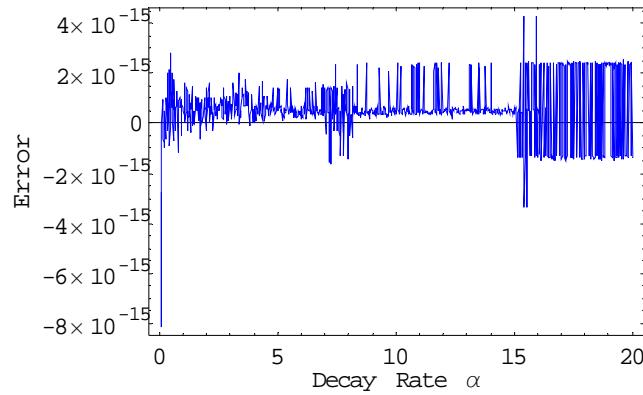
Secondly, we investigate the level of decay rate parameter α for Merton-FT formula. Figure 10.2 illustrates the pricing error (i.e. relative to the original Merton) of Merton-FT formula of a one day to maturity $T = 1/252$ call as a function of varying α . Panel A (for an OTM call) and C (for an ATM call) tells us that for $0.05 \leq \alpha \leq 20$, Merton-FT formula has effectively zero error relative to the original Merton price. But it seems for an ITM call (Panel B), the error monotonically increases as α rises. Therefore, from now on, we always use $\alpha = 1$ when implementing Merton-FT formula.



A) For an OTM Call with $K = 80$.



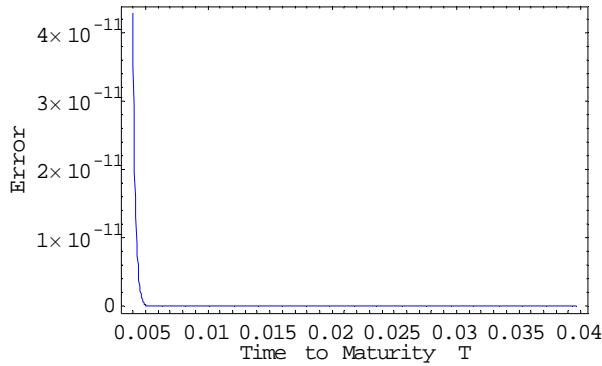
B) For an ITM Call with $K = 20$.



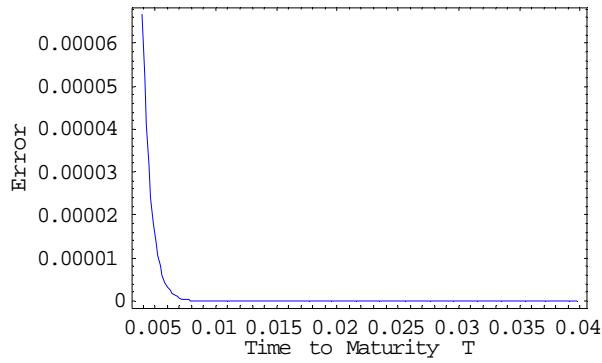
C) For an ATM Call with $K = 50$.

Figure 10.2: Original Merton Price Minus Merton-FT Price for One Day to Maturity Call as a Function of Decay Rate Parameter α . Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 1/252$, $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$.

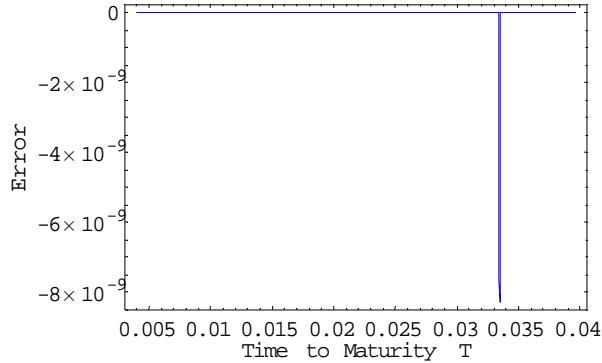
Thirdly, we investigate the amount of error caused by the difficulty in numerical integration in (10.4) for the near maturity calls. Consider a call option with $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$. Figure 10.3 plots a series of the difference between the original Merton price and Merton-FT $\alpha = 1$ price for the range of less than 10 trading days to maturity $1/252 \leq T \leq 10/252$. We find that despite the difficulty in the numerical integration of Merton-FT price of (10.4) for the near maturity deep OTM (in Panel A) and deep ITM (in Panel B) call, our Merton-FT code yields effectively zero error in terms of pricing.



A) OTM call with $K = 80$.



B) ITM call with $S_0 = 50$ and $K = 20$.



C) ATM Call with $S_0 = 50$ and $K = 50$.

Figure 10.3: Plot of Merton Price minus Merton-FT $\alpha = 1$ Price for Near-Maturity Vanilla Call with Less Than 10 Trading Days to Maturity $1/252 \leq T \leq 10/252$.

Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$.

We summarize this section. Although Merton-FT call price of the equation (10.4) is slower than the original Merton price formula, it produces negligible pricing errors

regardless of the maturity and the moneyness of the call. We recommend the use of $\alpha = 1$ for its decay rate.

[10.2] Discrete Fourier Transform (DFT) Call Pricing Formula with Merton Jump-Diffusion Model

Although the numerical integration difficulty of Merton-FT call price of (10.4) for the near maturity options yields virtually zero pricing error, it makes the evaluation slow. This speed becomes an issue when calibrating hundreds or thousands of prices (i.e. also in Monte Carlo simulation). To improve computational time, we apply our version of DFT call price formula in Merton case.

From the equation (8.38), Merton-DFT call price is given as:

$$C_T(k_n) \approx \frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k} \times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\} \exp(-i2\pi jn/N), \quad (10.5)$$

where $w_{j \in [0, N-1]}$ are trapezoidal rule weights:

$$w_j = \begin{cases} 1/2 & \text{for } j = 0 \text{ and } N-1 \\ 1 & \text{for others} \end{cases},$$

and:

$$\psi_T(\omega) = \frac{e^{-rT} \phi_T(\omega - (\alpha + 1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha + 1)\omega},$$

with $\phi_T(\cdot)$ given by (10.3).

[10.3] Implementation and Performance of DFT Pricing Method with Merton Jump-Diffusion Model

In this section, performance of Merton-DFT $\alpha = 1$ call price of the equation (10.5) is tested by comparing results to the original Merton call price and Merton-FT $\alpha = 1$ call price of the equation (10.4) under various settings. Merton-DFT $\alpha = 1$ call price is implemented using $N = 4096$ samples and log strike space sampling interval $\Delta k = 0.005$. This corresponds to angular frequency domain sampling interval of $\Delta\omega = 0.306796$ radians, the total sampling range in the log strike space is $K = N\Delta k = 20.48$, its sampling rate is 200 samples per unit of k , and the total sampling range in the angular frequency domain is $\Omega = N\Delta\omega = 1256.64$.

Firstly, let's investigate the difference in price and CPU time. Consider calculating 100-point call prices for a range of strike price $1 \leq K \leq 100$ with interval 1 with common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$. Figure 10.4 reports the price and Table 10.2 compares CPU time. We notice that call prices are virtually identical (i.e. they are supposed to be identical) and the use of DFT significantly improves the computational time although it is slower than the original Merton price.

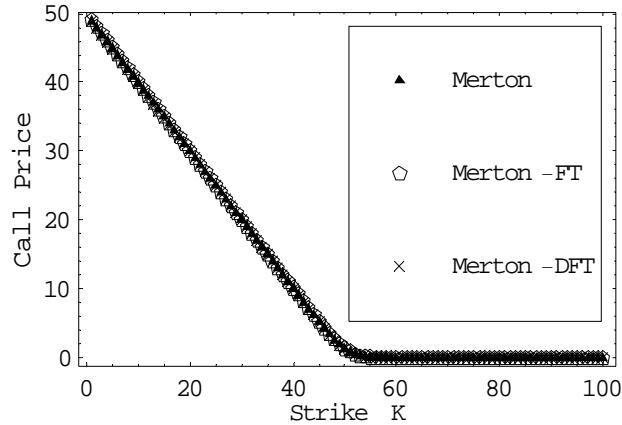


Figure 10.4: Merton Vs. Merton-FT $\alpha = 1$ Vs. Merton-DFT $\alpha = 1$. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$.

Table 10.2 CPU Time for Calculating 100-point call prices for a range of strike price $1 \leq K \leq 100$ with interval 1 Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$.

Method	CPU Time
Merton	0.14 seconds
Merton-FT $\alpha = 1$	7.22 seconds
Merton-DFT $\alpha = 1$ $N = 4096$, $\Delta k = 0.005$	1.31 seconds

Secondly, we pay extra attention to the pricing errors of very near-maturity calls. Consider a call option with common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$. Merton jump-diffusion parameters are set as $\lambda = 1$, $\mu = -0.1$, and $\delta = 0.1$. Figures 10.5 to 10.7 plot three series of price differences for vanilla calls computed by the original Merton, Merton-FT $\alpha = 1$, and Merton-DFT $\alpha = 1$ with $N = 4096$ and $\Delta k = 0.005$ as a function of time to maturity of less than a month $1/252 \leq T \leq 20/252$. Figure 10.5 tells us that for deep OTM calls, Merton-DFT yields

the pricing error around 6×10^{-8} which we interpret negligible. Figure 10.7 tells us that for deep ITM calls, Merton-DFT price is virtually identical to the original Merton except for one day to maturity. Very similar to the BS case, Figure 10.6 tells us that as the maturity nears, the error of ATM Merton-DFT price monotonically increases and the size of error is large (compared to ITM and OTM errors) but still negligible. In contrast, ATM Merton-FT price produces virtually no error. This finding is more clearly illustrated in Figure 10.7 where the $T = 1/252$ ATM pricing error of Merton-DFT $\alpha = 1$ with $N = 4096$ and $\Delta k = 0.005$ is plotted across different moneyness. We again realize that Merton-DFT price is virtually identical to BS price for the deep ITM and OTM calls, but its approximation error becomes an issue for ATM call. We have done several experiments trying to reduce the size of this around ATM error by increasing the decay rate parameter α to 10 or by sampling more points $N = 8192$. But these attempts were futile.

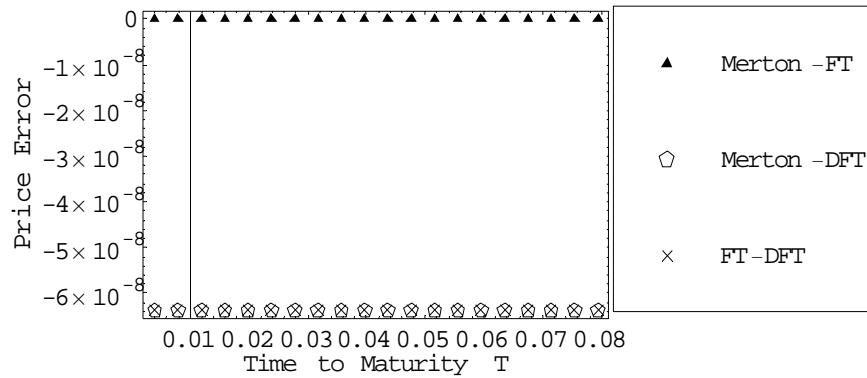


Figure 10.5: Plot of Price Error for Deep-OTM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 80$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

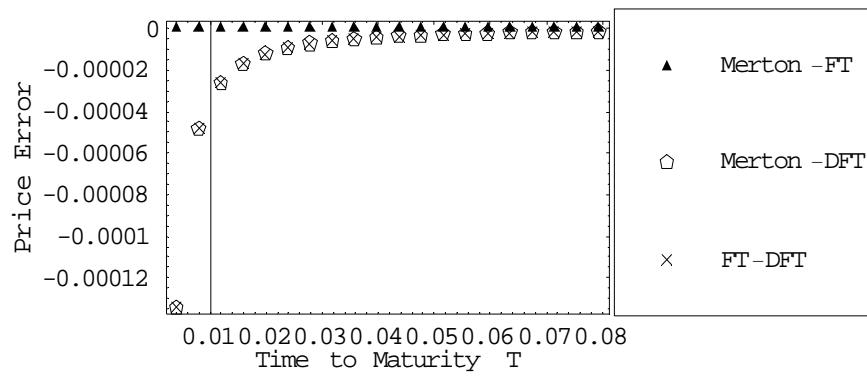


Figure 10.6: Plot of Price Error for ATM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

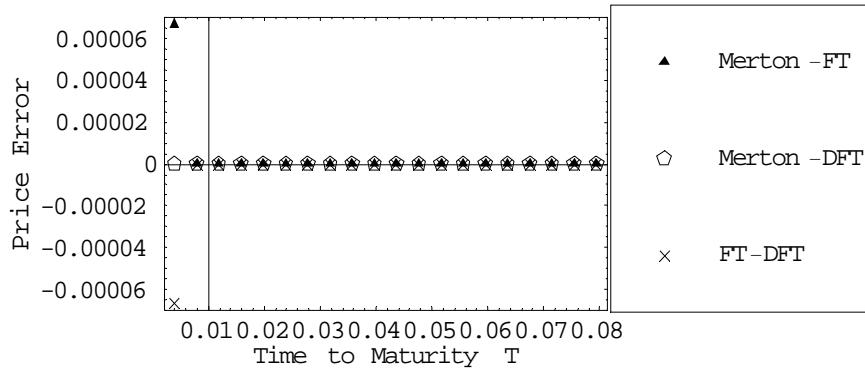


Figure 10.7: Plot of Price Error for Deep-ITM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 20$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

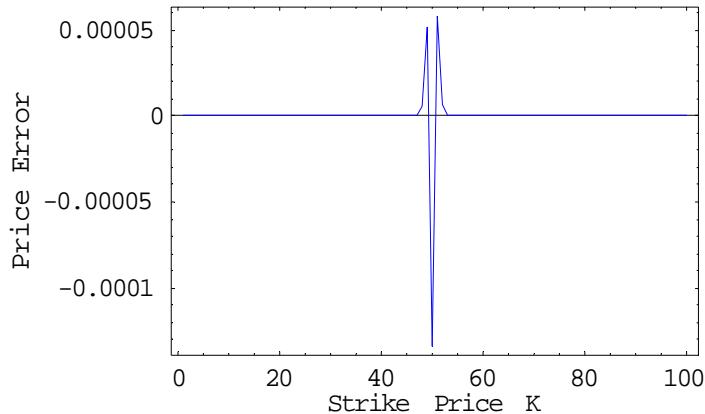


Figure 10.8: Plot of Price Error of Merton-DFT Formula for 1-Day-to-Maturity $T = 1/252$ ATM Call as a Function of Strike Price $0 \leq K \leq 100$. Common parameters and variables fixed are $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$.

We conclude this section by stating the following remarks. Our Merton-DFT $\alpha = 1$ call price formula, the equation (10.5), yields the price virtually identical to the original Merton price for OTM and ITM calls even for extreme near-maturity case (i.e. $T = 1/252$) although the size of error is larger than Merton-FT formula. But the error of Merton-DFT price becomes large around (i.e. ± 3) ATM. In our example used, the maximum error is 0.0001343 which occurs at exactly ATM at $S_0 = K = 50$. Increasing the decay rate parameter α or sampling more points (i.e. larger N) cannot reduce the size of this error. But we can accept this size of error when considering the dramatic improvement in the CPU time to compute hundreds of prices.

[10.4] Summary of Formulae of Option Price with Fourier Transform in Merton Jump-Diffusion Model

Table 10.10: Summary of Formulae for Merton Jump-Diffusion Model

Method	Formula
Original Merton	$C_{Merton} = \sum_{j=0}^{\infty} \frac{e^{-\bar{\lambda}T} (\bar{\lambda}T)^j}{j!} C_{BS}(S_0, T, \sigma_j, r_j)$ $\bar{\lambda} = \lambda(1+k) = \lambda \exp\left(\mu + \frac{1}{2}\delta^2\right), \quad \sigma_j^2 = \sigma^2 + \frac{j\delta^2}{T}$ $r_j = r - \lambda k + \frac{i \ln(1+k)}{T} = r - \lambda \left(e^{\mu + \frac{1}{2}\delta^2} - 1 \right) + \frac{i \left(\mu + \frac{1}{2}\delta^2 \right)}{T}$
FT	$C_{Merton-FT}(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} d\omega$ $\phi_T(\omega) = \exp\left[\lambda T \left(\exp(i\mu\omega - \frac{\delta^2\omega^2}{2}) - 1 \right) + i\omega \left(s_0 + (r - \frac{\sigma^2}{2} - \lambda k)T \right) - \frac{\sigma^2\omega^2}{2} T \right]$ $k \equiv e^{\mu + \frac{1}{2}\delta^2} - 1$
DFT	$C_T(k_n) \approx \frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k}$ $\times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\} \exp(-i2\pi jn/N)$ $\psi_T(\omega_j) = \frac{e^{-rT} \phi_T(\omega_j - (\alpha+1)i)}{\alpha^2 + \alpha - \omega_j^2 + i(2\alpha+1)\omega_j}$

[11] Variance Gamma (VG) Model by Madan, Carr, and Chang (1998)

[11.1] Model Type

In this section the basic structure of VG model is described without the derivation of the model which will be done in the next section.

VG model is an exponential Lévy model of the form:

$$S_t = S_0 e^{L_t},$$

where the asset price process $\{S_t; 0 \leq t \leq T\}$ is modeled as an exponential of a Lévy process $\{L_t; 0 \leq t \leq T\}$. Choice of the Lévy process by Madan, Carr, and Chang (1998) is a VG process, $VG(x_t; \theta_{\mathbb{P}}, \sigma_{\mathbb{P}}, \kappa_{\mathbb{P}})$, plus a drift:

$$L_t \equiv \left\{ m + \frac{1}{\kappa_{\mathbb{P}}} \ln \left(1 - \theta_{\mathbb{P}} \kappa_{\mathbb{P}} - \frac{\sigma_{\mathbb{P}}^2 \kappa_{\mathbb{P}}}{2} \right) \right\} t + VG(x_t; \theta_{\mathbb{P}}, \sigma_{\mathbb{P}}, \kappa_{\mathbb{P}}).$$

A VG process $VG(x_t; \theta_{\mathbb{P}}, \sigma_{\mathbb{P}}, \kappa_{\mathbb{P}})$ is defined as a stochastic process $\{X_t; 0 \leq t \leq T\}$ created by random time changing (i.e. subordinating) a Brownian motion with drift process $\theta t + \sigma B_t$ by a tempered 0-stable subordinator (i.e. a gamma subordinator) $\{S_t; 0 \leq t \leq T\}$ with unit mean rate:

$$X_t \equiv \theta(S_t) + \sigma B_{S_t}.$$

The CF of a VG process can be easily obtained by the use of the subordination theorem described in the section 11.2.1 and its probability density function is obtained in closed form.

A VG process $VG(x_t; \theta_{\mathbb{P}}, \sigma_{\mathbb{P}}, \kappa_{\mathbb{P}})$ is characterized as a pure jump Lévy process with infinite arrival rate of jumps. In other words, the Lévy measure of a VG process has an infinite integral:

$$\int_{-\infty}^{\infty} \ell(x) dx = \infty.$$

This means that a VG process has infinitely many small jumps but a finite number of large jumps.

Introduction of two extra parameters by the VG model captures the (negative) skewness and excess kurtosis of the log return density $\mathbb{P}(\ln(S_t / S_0))$ which deviates from the BS normal log return density. One is variance rate parameter κ which controls the degree of

the randomness of the subordination. Larger κ implies the fatter tails of the log return density $\mathbb{P}(\ln(S_t / S_0))$. The other is the drift parameter of the subordinated Brownian motion process θ which captures the skewness of the log return density.

Let's derive the model.

[11.2] Model Derivation

[11.2.1] Subordination Theorem of Lévy Processes

A transformation of a stochastic process to a new stochastic process through random time change by an increasing Lévy process (subordinator) independent of the original process is called subordination. Time is an increasing sequence of numbers: 0, 1, 2, 3, ... or 1, 10, 23, 30, 50, So is subordinator because subordinator is an increasing (i.e. non-decreasing) Lévy process. This non-decreasing property of a subordinator makes it a possible candidate for a time indicator.

Let $\{S_t; 0 \leq t \leq T\}$ be a subordinator with its characteristic triplet $(b_s, \sigma_s = 0, \ell_s)$. Any increasing function is of finite variation in one dimension and following the Lévy-Khintchin representation in section 6.5, the MGF (moment generating function) of S_t is:

$$M_s(\omega) \equiv E[\exp(\omega S_t)] = \exp\{t\mathcal{L}_s(\omega)\},$$

where the Laplace exponent is given by:

$$\mathcal{L}_s(\omega) = b_s \omega + \int_0^\infty (e^{\omega x} - 1) \ell_s(dx). \quad (11.1)$$

We used a relationship between the MGF M_x and the CF (characteristic function) given by $M_x(t) = \phi_x(-it)$ which holds when M_x is well-defined. Laplace transform is used instead of Fourier transform because Lévy measure of an increasing Lévy process is concentrated on the positive real-axis (i.e. by definition). $\mathcal{L}_s(\omega)$ satisfies the following:

$$\begin{aligned} \int_0^\infty 1_{dx \geq 1} \rho(dx) &< \infty \text{ (i.e. a finite variation condition),} \\ b_s &\geq 0 \text{ (i.e. a positive drift),} \end{aligned}$$

which are the conditions of being an increasing Lévy process.

Let $\{X_t; 0 \leq t \leq T\}$ be a Lévy process on \mathbb{R} with its characteristic triplet (b, σ, ℓ) and its CF given by the Lévy-Khintchin representation:

$$\phi_x(\omega) \equiv E[\exp(i\omega X_t)] = \exp\{t\psi_x(\omega)\},$$

where the CE (characteristic exponent) is given by:

$$\psi_X(\omega) = ib\omega - \frac{\sigma^2\omega^2}{2} + \int_{-\infty}^{\infty} \{\exp(i\omega x) - 1\} \ell(dx). \quad (11.2)$$

We assume $\{X_t; 0 \leq t \leq T\}$ and $\{S_t; 0 \leq t \leq T\}$ are independent.

Define a new stochastic process $\{Z_t; 0 \leq t \leq T\}$ by random time changing (i.e. subordinating) the original Lévy process $\{X_t; 0 \leq t \leq T\}$:

$$Z_t \equiv X_{S_t}. \quad (11.3)$$

Then, the process $\{Z_t; 0 \leq t \leq T\}$ is a Lévy process on \mathbb{R} with its CF given by:

$$\phi_Z(\omega) \equiv E\{\exp(i\omega Z_t)\} = \exp\{t\mathcal{L}_S(\psi_X(\omega))\}. \quad (11.4)$$

This is an extremely powerful and useful theorem since we can obtain the CF of a random time-changed Lévy process Z_t by simply substituting the CE of an original Lévy process $\psi_X(\omega)$ into the Laplace exponent of the subordinator $\mathcal{L}_S(\omega)$. The characteristic triplet of a random time-changed Lévy process Z_t , (b_Z, σ_Z, ℓ_Z) , is given by:

$$\begin{aligned} b_Z &= b_S b + \int_0^\infty \ell_s(ds) \int 1_{|x| \leq 1} x \mathbb{P}_X(dx), \\ \sigma_Z^2 &= b_S^2 \sigma^2, \\ \ell_Z(dx) &= b_S \ell(dx) + \int_0^\infty \mathbb{P}_X(dx) \ell_s(ds), \end{aligned} \quad (11.5)$$

where $\mathbb{P}_X()$ is the probability density function of X_t . Transformation of a Lévy process $\{X_t; 0 \leq t \leq T\}$ into another Lévy process $\{Z_t; 0 \leq t \leq T\}$ is called subordination by the subordinator $\{S_t; 0 \leq t \leq T\}$.

[11.2.2] Tempered α -Stable Subordinator: General Case

For the reason of pure mathematical tractability of the process, we restrict our attention to the tempered α -stable subordinator. Let $\{S_t; 0 \leq t \leq T\}$ be a tempered α -stable subordinator which is Esscher transformed stable subordinator (i.e. whose tails of both probability density and Lévy density are damped symmetrically) and $(b_S, \sigma_S = 0, \ell_S)$ be its characteristic triplet, then:

- (1) $\{S_t; 0 \leq t \leq T\}$ is a α -stable process with the index of stability $0 < \alpha < 1$.

- (2) $\int_{-\infty}^0 \ell_s(dx) = 0$: $\{S_t; 0 \leq t \leq T\}$ has no negative jump. In other words, Lévy measure concentrated on the positive real-axis.
- (3) $b_s \geq 0$: Positive drift.
- (4) Lévy measure is three-parameter family and is simply a tempered version of the Lévy measure of a stable subordinator of the equation (6.12):

$$\ell_s(x) = \frac{c \exp(-\lambda x)}{x^{\alpha+1}} 1_{x>0}, \quad (11.6)$$

where $c > 0$ and $\lambda > 0$. The intensity of jumps of all sizes is controlled by the parameter c and the decay rate of large jumps is controlled by λ . And the relative importance of small jumps in the trajectory is determined by α .

Let $\{S_t; 0 \leq t \leq T\}$ be a tempered α -stable subordinator with zero drift for simplicity (i.e. $b_s = 0$) in the general case $0 < \alpha < 1$. Following the Lévy-Khintchin representation in section 6.5, the MGF of S_t is given by (11.1) as:

$$M_s(\omega) \equiv E[\exp(\omega S_t)] = \exp\{t\mathcal{L}_s(\omega)\},$$

where the Laplace exponent is given by:

$$\begin{aligned} \mathcal{L}_s(\omega) &= b_s \omega + \int_0^\infty (e^{\omega x} - 1) \ell_s(dx) \\ &= \int_0^\infty (e^{\omega x} - 1) \frac{c e^{-\lambda x}}{x^{\alpha+1}} 1_{x>0} dx \\ &= c \{(\lambda - \omega)^\alpha - \lambda^\alpha\} \Gamma(-\alpha). \end{aligned} \quad (11.7)$$

Using the MGF, the mean, variance, skewness, kurtosis of the tempered α -stable subordinator with zero drift $b_s = 0$ in the general case are obtained as:

$$\begin{aligned} E[S_t] &= m_1 = -ct\alpha\lambda^{\alpha-1}\Gamma[-\alpha], \\ Var[S_t] &= m_2 - m_1^2 = ct(\alpha-1)\alpha\lambda^{\alpha-2}\Gamma[-\alpha], \\ Skewness[S_t] &= \frac{2m_1^3 - 3m_1m_2 + m_3}{(m_2 - m_1^2)^{3/2}} = \frac{2-\alpha}{\lambda\sqrt{ct(\alpha-1)\alpha\lambda^{\alpha-2}\Gamma[-\alpha]}}, \\ Excess\ Kurtosis[S_t] &= \frac{-6m_1^4 + 12m_1^2m_2 - 3m_2^2 - 4m_1m_3 + m_4}{(m_2 - m_1^2)^2} = \frac{\lambda^{-\alpha}(6-5\alpha+\alpha^2)}{ct(\alpha-1)\alpha\Gamma[-\alpha]}. \end{aligned} \quad (11.8)$$

Note that the probability density of the tempered α -stable subordinator with zero drift $b_s = 0$ in the general case $0 < \alpha < 1$ is not known.

[11.2.3] Gamma Subordinator (Process): Special Case of Tempered α -Stable Subordinator When $\alpha = 0$

Consider a special case when the stability index takes zero, i.e. $\alpha = 0$. A tempered 0-stable subordinator with zero drift $b_s = 0$ is called a gamma subordinator. Lévy density of the gamma subordinator S_t is obtained by setting $\alpha = 0$ in the equation (11.6) as:

$$\ell_S(x) = \frac{ce^{-\lambda x}}{x} 1_{x>0}. \quad (11.9)$$

Following the Lévy-Khintchin representation of the equation (11.1), the MGF of the gamma subordinator S_t is:

$$M_S(\omega) \equiv E[\exp(\omega S_t)] = \exp\{t\mathcal{L}_S(\omega)\},$$

where the Laplace exponent is given by substituting (11.9):

$$\begin{aligned} \mathcal{L}_S(\omega) &= b_s \omega + \int_0^\infty (e^{\omega x} - 1) \ell_S(dx) \\ &= \int_0^\infty (e^{\omega x} - 1) \frac{ce^{-\lambda x}}{x} dx \\ &= c \ln\left(\frac{\lambda}{\lambda - \omega}\right). \end{aligned}$$

Therefore, the MGF of the gamma subordinator S_t is explicitly calculated as the following:

$$\begin{aligned} M_S(\omega) &\equiv \exp\{t\mathcal{L}(\omega)\} = \exp\left\{tc \ln\left(\frac{\lambda}{\lambda - \omega}\right)\right\} \\ &= \exp\left\{\ln\left(\frac{\lambda}{\lambda - \omega}\right)^{tc}\right\} = \left(\frac{\lambda}{\lambda - \omega}\right)^{tc} \\ &= \left(\frac{1}{1 - \omega\lambda^{-1}}\right)^{tc} = \frac{1}{(1 - \omega\lambda^{-1})^{ct}}. \end{aligned} \quad (11.10)$$

Note that the equation (11.10) cannot be obtained by setting $\alpha = 0$ in (11.7).

The probability density of the gamma subordinator S_t follows a gamma distribution:

$$Gamma_t(g) = \frac{\lambda^{ct}}{\Gamma(ct)} g^{ct-1} e^{-\lambda g} 1_{g>0}. \quad (11.11)$$

Using moment generating function (11.10) the mean, variance, skewness, and excess kurtosis of the gamma subordinator are calculated as:

$$\begin{aligned} E[S_t] &= \frac{ct}{\lambda}, \\ Var[S_t] &= \frac{ct}{\lambda^2}, \\ Skewness[S_t] &= \frac{2}{\sqrt{ct}}, \\ Excess\ Kurtosis[S_t] &= \frac{6}{ct}. \end{aligned} \tag{11.12}$$

Note that MCC (1998) uses the different parameterization such that:

$$\begin{aligned} \mu &= \text{mean rate} = \frac{c}{\lambda} \quad \text{and} \quad \kappa = \text{variance rate} = \frac{c}{\lambda^2}, \\ (c &= \frac{\mu^2}{\kappa} \quad \text{and} \quad \lambda = \frac{\mu}{\kappa}). \end{aligned} \tag{11.13}$$

Don't be confused by this different parameterization. We and MCC (1998) are the same. Alternatively, (11.12) can be obtained by setting $\alpha = 0$ in the equation (11.8).

A gamma process is an infinite activity Lévy process because a gamma distribution possesses the infinite divisibility and its Lévy measure (11.9) has an infinite integral. This means that a gamma process has infinitely many small jumps (i.e. infinite arrival rate of small jumps) and a finite number of large jumps (i.e. $\ell_s(dx)$ is concentrated at the origin). A gamma process is a pure jump process since it has no continuous (Brownian motion) component.

[11.2.4] Re-Parameterizing Tempered α -Stable Subordinator Using Its Scale-Invariance (Self-Similarity) Property

A tempered α -stable subordinator $S_t(\alpha, \lambda, c)$ possesses a very important property called the scale-invariance since the process S at scale $w^\alpha t$ has the same law as the process S at scale t after appropriate rescaling. In other words, if $\{S_t(\alpha, \lambda, c); 0 \leq t \leq T\}$ is a tempered α -stable subordinator with zero drift $b_s = 0$ and with parameters α , λ , and c , then for every positive constant w :

$$wS_t(\alpha, \lambda, c) \underset{w^\alpha t}{=} S_{w^\alpha t}(\alpha, \lambda / w, c),$$

which is equivalent to stating that the process S is self-similar (read section 6.6 for the definition). Remember that if a Lévy process (i.e. includes subordinator) is self-similar,

then it is strictly stable. This indicates by definition that a tempered α -stable subordinator possesses this scale-invariance property. We saw in section 6.6 that a standard Brownian motion process B_t is a strictly 2-stable Lévy process:

$$wB_t \stackrel{d}{=} B_{w^2 t}.$$

Because of this scale invariance property of a tempered α -stable subordinator $\{S_t(\alpha, \lambda, c); t \geq 0\}$ and a standard Brownian motion process B_t , we consider only tempered α -stable subordinators with $E[S_t] = t$ in subordinated models. In other words, we re-parameterize a tempered α -stable subordinator so that it has the mean t (i.e. called unit mean rate). Remember that in the general case $0 < \alpha < 1$ the mean of a tempered α -stable subordinator is given by the equation (11.8). The re-parameterization is done by setting:

$$c = \frac{1}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{\kappa} \right)^{1-\alpha} \quad \text{and} \quad \lambda = \frac{1-\alpha}{\kappa}, \quad (11.14)$$

such that:

$$E[S_t] = -\frac{1}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{\kappa} \right)^{1-\alpha} t \alpha \left(\frac{1-\alpha}{\kappa} \right)^{\alpha-1} \Gamma[-\alpha] = t,$$

where use $\frac{\Gamma(-c)}{\Gamma(1-c)} = -\frac{1}{c}$. And the Lévy measure is re-parameterized from three-parameter measure (11.6) to two-parameter measure:

$$\ell_s(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{\kappa} \right)^{1-\alpha} \frac{\exp\{- (1-\alpha)x/\kappa\}}{x^{1+\alpha}}, \quad (11.15)$$

where α is the stability index and $\kappa > 0$ is a variance rate which is equal to the variance of these subordinators at time 1, i.e. $Var[S_1] = \kappa$. Thus, κ determines the degree of the randomness of the time change (i.e. subordination). When $\kappa = 0$, the process is deterministic.

Following the Lévy-Khinchin representation of the equation (11.1), the MGF of the re-parameterized tempered α -stable subordinator S_t is:

$$M_s(\omega) \equiv E[\exp(\omega S_t)] = \exp\{t\mathcal{L}_s(\omega)\},$$

where the Laplace exponent is given by substituting (11.15):

$$\begin{aligned}
\mathcal{L}_S(\omega) &= b_S \omega + \int_0^\infty (e^{\omega x} - 1) \ell_S(dx) \\
&= \int_0^\infty (e^{\omega x} - 1) \frac{1}{\Gamma(1-\alpha)} \left(\frac{1-\alpha}{\kappa}\right)^{1-\alpha} \frac{\exp\{-(1-\alpha)x/\kappa\}}{x^{1+\alpha}} dx \\
&= \frac{1-\alpha}{\kappa\alpha} \left\{ 1 - \left(1 - \frac{\kappa\omega}{1-\alpha} \right)^\alpha \right\}.
\end{aligned} \tag{11.16}$$

From the MGF, the variance, skewness, and excess kurtosis of the tempered α -stable subordinator with unit mean rate are calculated as:

$$\begin{aligned}
Var[S_t] &= \kappa t, \\
Skewness[S_t] &= \frac{2-\alpha}{1-\alpha} \sqrt{\frac{\kappa}{t}}, \\
Excess\ Kurtosis[S_t] &= \frac{\kappa \alpha^2 - 5\alpha + 6}{t (1-\alpha)^2}.
\end{aligned} \tag{11.17}$$

Note that the probability density of the tempered α -stable subordinator with unit mean rate and zero drift $b_S = 0$ in the general case $0 < \alpha < 1$ is not known.

[11.2.5] Gamma Subordinator (Process): Special Case of Tempered α -Stable Subordinator with Unit Mean Rate When $\alpha = 0$

Again, consider a special case when $\alpha = 0$, the tempered α -stable subordinator with unit mean rate becomes a gamma subordinator with unit mean rate $E[S_t] = t$. Its Lévy density can be expressed by setting $\alpha = 0$ in the equation (11.15) as:

$$\ell_S(x) = \frac{1}{\kappa} \frac{\exp(-x/\kappa)}{x} 1_{x>0}. \tag{11.18}$$

Because the Laplace exponent of the gamma subordinator with unit mean rate cannot be obtained by simply substituting $\alpha = 0$ in (11.16), we need to derive it step-by-step. The MGF of the gamma subordinator S_t with unit mean rate can be expressed as using Lévy-Khinchin representation:

$$M_S(\omega) \equiv E[\exp(\omega S_t)] = \exp\{t\mathcal{L}_S(\omega)\},$$

where the Laplace exponent is given as using (11.18):

$$\mathcal{L}_S(\omega) = \int_0^\infty (e^{\omega x} - 1) \ell_S(x) dx = \int_0^\infty (e^{\omega x} - 1) \frac{1}{\kappa} \frac{\exp(-x/\kappa)}{x} dx$$

$$= \frac{-\log \kappa - \log \left(\frac{1}{\kappa} - \omega \right)}{\kappa}. \quad (11.19)$$

The mean, variance, skewness, and excess kurtosis of the gamma subordinator with unit mean rate are calculated as using MGF or by setting $\alpha = 0$ in (11.17):

$$\begin{aligned} E[S_t] &= t, \\ Var[S_t] &= \kappa t, \\ Skewness[S_t] &= 2\sqrt{\frac{\kappa}{t}}, \\ Excess\ Kurtosis[S_t] &= \frac{6\kappa}{t}. \end{aligned} \quad (11.20)$$

Consider obtaining the probability density of the gamma subordinator S_t with unit mean rate. The equation (11.11) is the probability density of the gamma subordinator S_t with the mean rate c/λ (check the equation (11.12)):

$$Gamma_t(g) = \frac{\lambda^{ct}}{\Gamma(ct)} g^{ct-1} e^{-\lambda g} 1_{g>0}.$$

Re-parameterization terms for the purpose of having unit mean rate for the gamma subordinator case are obtained by setting $\alpha = 0$ in the equation (11.14):

$$c = \frac{1}{\kappa} \text{ and } \lambda = \frac{1}{\kappa}.$$

Obviously, the mean rate after this re-parameterization is:

$$\frac{c}{\lambda} = \frac{1/\kappa}{1/\kappa} = 1.$$

Thus, the probability density of a gamma subordinator S_t with unit mean rate $E[S_t] = t$ has a gamma density of the following form:

$$\begin{aligned} Gamma_t(g) &= \frac{(1/\kappa)^{(1/\kappa)t}}{\Gamma((1/\kappa)t)} g^{(1/\kappa)t-1} e^{-(1/\kappa)g} 1_{g>0} \\ &= \frac{\left(\frac{1}{\kappa}\right)^{t/\kappa}}{\Gamma(t/\kappa)} g^{\frac{t}{\kappa}-1} e^{-g/\kappa} 1_{g>0}. \end{aligned} \quad (11.21)$$

[11.2.6] Subordinated Brownian Motion Process with Tempered α -Stable Subordinator with Unit Mean Rate: Normal Tempered α -Stable Process

Firstly, we deal with a general case. Let $\{B_t; 0 \leq t \leq T\}$ be a standard Brownian motion process and $\theta t + \sigma B_t \sim N(\theta t, \sigma^2 t)$ be a Brownian motion with drift process. Define a new stochastic process $\{X_t; 0 \leq t \leq T\}$ by random time changing (i.e. subordinating) a Brownian motion with drift process $\theta t + \sigma B_t$ by a tempered α -stable subordinator with unit mean rate $\{S_t; 0 \leq t \leq T\}$:

$$X_t \equiv \theta(S_t) + \sigma B_{S_t}.$$

Then, the stochastic process $\{X_t; t \geq 0\}$ is said to be a normal tempered α -stable process. The CF of a normal tempered α -stable process can be obtained by the use of the subordination theorem of the equation (11.4). As the first step, obtain the CF of the original Lévy process which is a Brownian motion with drift process $\theta t + \sigma B_t$:

$$\phi_B(\omega) = E[e^{i\omega(\theta t + \sigma B_t)}] = \exp(i\theta\omega t - \frac{\sigma^2\omega^2}{2}t) = \exp\{t\psi_B(\omega)\},$$

where the characteristic exponent is:

$$\psi_B(\omega) = i\theta\omega - \frac{\sigma^2\omega^2}{2}. \quad (11.22)$$

We saw that the Laplace exponent of the tempered α -stable subordinator S_t with $E[S_t] = t$ can be expressed as in the general case $0 < \alpha < 1$ by (11.16). Following the subordination theorem of the equation (11.2), the CF of the normal tempered α -stable process $X_t \equiv \theta(S_t) + \sigma B_{S_t}$ can be obtained by a substitution of (11.22) into (11.16):

$$\phi_X(\omega) \equiv E\{\exp(i\omega X_t)\} = \exp\{t\mathcal{L}_S(\psi_B(\omega))\},$$

with its characteristic exponent:

$$\begin{aligned} \psi_X(\omega) &= \mathcal{L}_S(\psi_B(\omega)) = \mathcal{L}_S\left(i\theta\omega - \frac{\sigma^2\omega^2}{2}\right) \\ &= \frac{1-\alpha}{\kappa\alpha} \left[1 - \left\{ 1 + \frac{\kappa \left(\frac{\sigma^2\omega^2}{2} - i\theta\omega \right)}{1-\alpha} \right\}^\alpha \right]. \end{aligned} \quad (11.23)$$

Using the characteristic exponent (11.23), the mean, variance, skewness, and excess kurtosis of the normal tempered α -stable process with unit mean rate are calculated as in the general case $0 < \alpha < 1$:

$$\begin{aligned} E[X_t] &= \theta t, \\ Var[X_t] &= (\kappa\theta^2 + \sigma^2)t, \\ Skewness[X_t] &= \frac{\kappa\theta(\kappa\theta^2(\alpha-2) + 3\sigma^2(\alpha-1))}{\alpha-1}t, \\ Excess\ Kurtosis[X_t] &= \frac{\kappa(\kappa^2\theta^4(\alpha^2-5\alpha+6) + 6\kappa\theta^2\sigma^2(\alpha^2-3\alpha+2) + 3\sigma^4(\alpha-1)^2)}{(\alpha-1)^2(\kappa\theta^2 + \sigma^2)^2}t. \end{aligned} \quad (11.24)$$

Cont and Tankov (2004) shows that the Lévy density of the normal tempered α -stable process is given by (again we assume the drift of the subordinator is zero for simplicity, i.e. $b_s = 0$):

$$\ell(x) = \exp\left(\frac{\theta}{\sigma^2}x\right) \frac{F(\alpha, \kappa, \sigma, \theta)}{|x|^{\alpha+\frac{1}{2}}} K_{\alpha+\frac{1}{2}}\left(\frac{|x|\sqrt{\theta^2 + \frac{2}{\kappa}\sigma^2(1-\alpha)}}{\sigma^2}\right), \quad (11.25)$$

where $F(\alpha, \kappa, \sigma, \theta) = \left\{\theta^2 + \frac{2}{\kappa}\sigma^2(1-\alpha)\right\}^{\frac{\alpha-1}{2}} \frac{2}{\Gamma(1-\alpha)\sqrt{2\pi\sigma^2}} \left(\frac{1-\alpha}{\kappa}\right)^{1-\alpha}$ and K is a modified Bessel function of the second kind (Appendix 7 gives its details).

Note that because the probability density of tempered α -stable subordinator is not known in the general case (i.e. known only for $\alpha = 0$ and $1/2$), the probability density of the normal tempered α -stable process is not known in the general case $0 < \alpha < 1$, either.

[11.2.7] Variance Gamma (VG) Process: Subordinated Brownian Motion Process with Tempered 0-Stable Subordinator with Unit Mean Rate: Normal Tempered 0-Stable Process

Consider a special case when $\alpha = 0$. Define a new stochastic process $\{X_t; 0 \leq t \leq T\}$ by random time changing (i.e. subordinating) a Brownian motion with drift process $\theta t + \sigma B_t$ by a tempered $\alpha = 0$ -stable subordinator $\{S_t; 0 \leq t \leq T\}$ with unit mean rate:

$$X_t \equiv \theta(S_t) + \sigma B_{S_t}.$$

Then, the stochastic process $\{X_t; 0 \leq t \leq T\}$ is a normal tempered 0-stable process which is called a variance gamma (VG) process. Because the CF of VG process cannot be obtained by simply substituting $\alpha = 0$ in (11.23), we need to do this step-by-step.

Following the subordination theorem of the equation (11.2), the characteristic function of the VG (normal tempered 0-stable) process $X_t \equiv \theta(S_t) + \sigma B_{S_t}$ can be obtained by substituting the CE of the Brownian motion with drift process (11.22) into the Laplace exponent of the gamma subordinator $\{S_t; 0 \leq t \leq T\}$ with unit mean rate (11.19):

$$\phi_X(\omega) \equiv E\{\exp(i\omega X_t)\} = \exp\{t\mathcal{L}_S(\psi_B(\omega))\},$$

with its characteristic exponent:

$$\psi_X(\omega) = \mathcal{L}_S(\psi_B(\omega)) = \mathcal{L}\left(i\theta\omega - \frac{\sigma^2\omega^2}{2}\right) = -\frac{1}{\kappa} \ln\left(1 + \frac{\sigma^2\omega^2\kappa}{2} - i\theta\kappa\omega\right). \quad (11.26)$$

The characteristic function of the VG process $\phi_X(\omega)$ can be explicitly calculated as follows (corresponding to MCC (1998) equation 7):

$$\begin{aligned} \phi_X(\omega) &= \exp\{t\mathcal{L}_S(\psi_B(\omega))\} = \exp\left[t\left\{-\frac{1}{\kappa} \ln\left(1 + \frac{\sigma^2\omega^2\kappa}{2} - i\theta\kappa\omega\right)\right\}\right] \\ \ln \phi_X(\omega) &= -\frac{t}{\kappa} \ln\left(1 + \frac{\sigma^2\omega^2\kappa}{2} - i\theta\kappa\omega\right) = \ln\left(1 + \frac{\sigma^2\omega^2\kappa}{2} - i\theta\kappa\omega\right)^{-\frac{t}{\kappa}} \\ \phi_X(\omega) &= \left(1 + \frac{\sigma^2\omega^2\kappa}{2} - i\theta\kappa\omega\right)^{-\frac{t}{\kappa}}. \end{aligned} \quad (11.27)$$

Consider obtaining a probability density function of a VG process $\{X_t; 0 \leq t \leq T\}$. Start from the fact that a Brownian motion with drift $W_t \equiv \theta t + \sigma B_t$ has a normal density:

$$\mathbb{P}(W_t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left\{-\frac{(W_t - \theta t)^2}{2\sigma^2 t}\right\}.$$

Because of the subordination structure $X_t \equiv \theta(S_t) + \sigma B_{S_t}$, by conditioning on the fact that the realized value of the random time change by a gamma subordinator S_t with unit mean rate is $S_t = g$, the conditional probability density of VG process can be expressed as:

$$VG(x_t | t = S_t = g) = \frac{1}{\sqrt{2\pi\sigma^2}g} \exp\left\{-\frac{(x_t - \theta g)^2}{2\sigma^2 g}\right\}.$$

Therefore, the unconditional density of the VG process is calculated by multiplying out the gamma probability density of the equation (11.21):

$$VG(x_t) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}g} \exp\left\{-\frac{(x_t - \theta g)^2}{2\sigma^2 g}\right\} \frac{(1/\kappa)^{t/\kappa}}{\Gamma(t/\kappa)} g^{\frac{t}{\kappa}-1} e^{-g/\kappa} dg, \quad (11.28)$$

which corresponds to MCC (1998) equation 6. After tedious algebra:

$$VG(x_t) = \frac{\sqrt{2} \exp\left(\frac{\theta}{\sigma^2} x_t\right)}{\kappa^{t/\kappa} \sigma \sqrt{\pi} \Gamma(t/\kappa)} \left(\frac{x_t^2}{\frac{2\sigma^2}{\kappa} + \theta^2} \right)^{\frac{t-1}{2\kappa}} K_{\frac{t-1}{\kappa-2}} \left(\sqrt{\frac{x_t^2 \left(\frac{2\sigma^2}{\kappa} + \theta^2 \right)}{\sigma^2}} \right), \quad (11.29)$$

which is MCC (1998) equation 23 and K is a modified Bessel function of the second kind.

Using the CF (11.27), cumulants of the VG process are calculated as:

$$\begin{aligned} cumulant_1 &= \theta t, \\ cumulant_2 &= (\kappa\theta^2 + \sigma^2)t, \\ cumulant_3 &= (2\theta^3\kappa^2 + 3\sigma^2\theta\kappa)t, \\ cumulant_4 &= (3\sigma^4\kappa + 6\theta^4\kappa^3 + 12\sigma^2\theta^2\kappa^2)t. \end{aligned}$$

From these cumulants (or by setting $\alpha = 0$ in the equation (11.24)), annualized (per unit of time) mean, variance, skewness, and excess kurtosis of the variance gamma process are calculated as (corresponding to MCC (1998) equations 18, 19, and 20):

$$\begin{aligned} E[X_t] &= \theta, \\ Var[X_t] &= \kappa\theta^2 + \sigma^2, \\ Skewness[X_t] &= 2\theta^3\kappa^2 + 3\sigma^2\theta\kappa, \\ Excess\ Kurtosis[X_t] &= \frac{3\sigma^4\kappa + 6\theta^4\kappa^3 + 12\sigma^2\theta^2\kappa^2}{(\kappa\theta^2 + \sigma^2)^2}. \end{aligned} \quad (11.30)$$

Note that these standardized moments are not equivalent to those of log-return density of VG model. These are simply the moments of VG density. This point will be explained in section 11.4. Several very interesting properties of VG density should be worth

mentioning. If $\theta = 0$ (i.e. subordination of zero-drift Brownian motion process with gamma subordinator), then:

$$\begin{aligned} E[X_t] &= 0, \\ \text{Var}[X_t] &= \sigma^2, \\ \text{Skewness}[X_t] &= 0, \\ \text{Excess Kurtosis}[X_t] &= 3\kappa, \end{aligned}$$

which indicates that VG density is symmetric (zero-skewness) and its excess kurtosis (i.e. tail behavior) is determined by the variance rate κ of the gamma subordinator S_t .

Generally, the sign of the drift parameter of the Brownian motion with drift process θ determines the skewness of VG density as illustrated in Figure 11.1. The case $\theta = -0.2$ is a mirror image of the case $\theta = 0.2$.

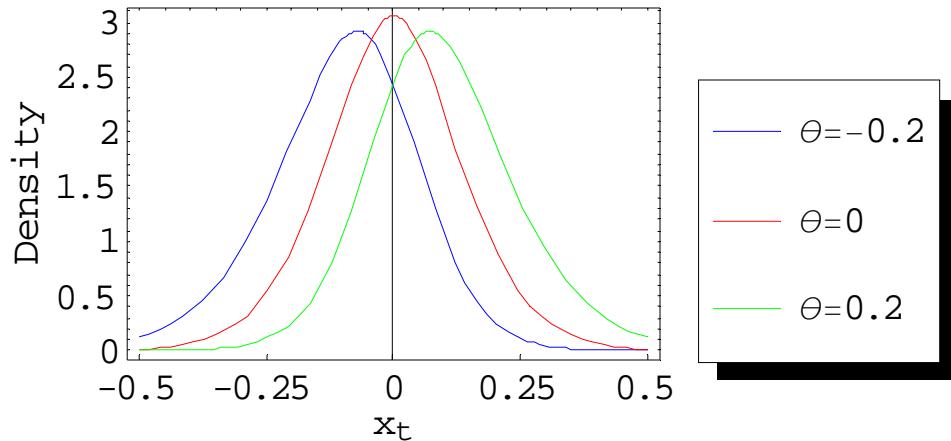


Figure 11.1: VG Density for Different Values of Drift Parameter of the Brownian Motion Process θ . Parameters fixed are $t = 0.5$, $\sigma = 0.2$, and $\kappa = 0.1$.

Table 11.1
Annualized Moments of VG Density in Figure 11.1

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
$\theta = -0.2$	-0.2	0.209762	-0.00256	0.352066
$\theta = 0$	0	0.2	0	0.3
$\theta = 0.2$	0.2	0.209762	0.00256	0.352066

Secondly, larger value of variance rate κ (i.e. the variance rate of the gamma subordinator S_t with unit mean rate and determines the degree of randomness of the subordination) makes the density fatter-tailed and higher-peaked as illustrated in Figure 11.2.

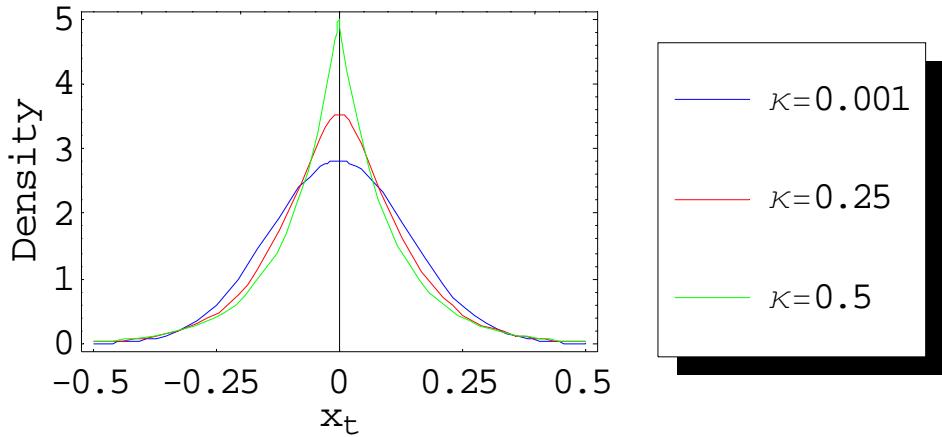


Figure 11.2: VG Density for Different Values of Variance Rate Parameter κ .
Parameters fixed are $t = 0.5$, $\sigma = 0.2$, and $\theta = 0$.

Table 11.2
Annualized Moments of VG Density in Figure 11.2

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
$\kappa = 0.001$	0	0.2	0	0.003
$\kappa = 0.25$	0	0.2	0	0.75
$\kappa = 0.5$	0	0.2	0	1.5

Also note that VG density has higher peak and fatter tails (more leptokurtic) when matched to the Normal density as illustrated in Figure 11.3.

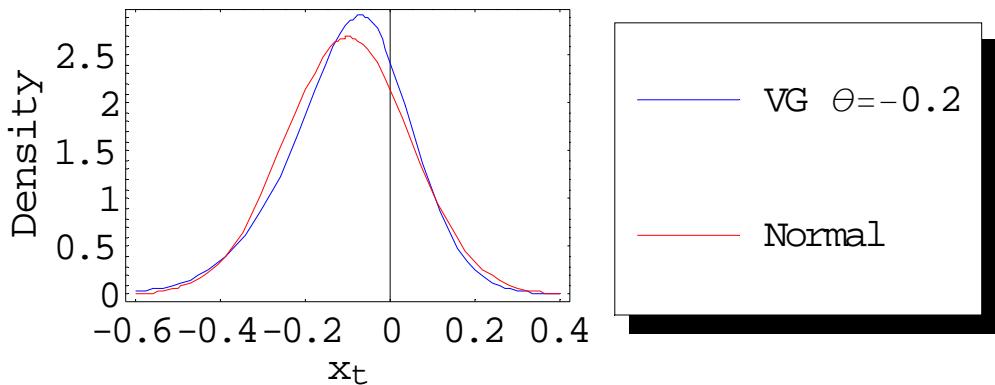


Figure 11.3: VG Density vs. Normal Density. Parameters fixed for the VG density are $t = 0.5$, $\theta = -0.2$, $\sigma = 0.2$, and $\kappa = 0.1$. Normal density is plotted by matching the mean and variance to the VG.

Table 11.3
Annualized Moments of VG Density in Figure 11.3

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
VG with $\theta = -0.2$	-0.2	0.209762	-0.00256	0.352066
Normal	-0.2	0.209762	0	0

[11.3] Lévy Measure for Variance Gamma (VG) Process

We saw that the Lévy measure for the normal tempered α -stable process in the general case $0 < \alpha < 1$ is given by the equation (11.25). The Lévy density of the VG process can be obtained by setting $\alpha = 0$ in (11.25) as:

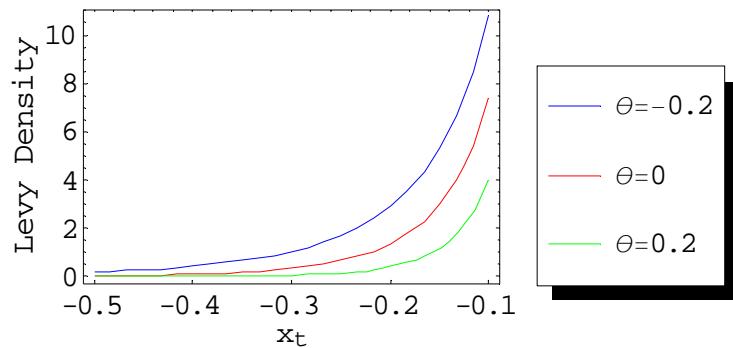
$$\ell(x) = \frac{1}{\kappa|x|} \exp\left\{Ax - B|x|\right\}, \quad (11.31)$$

where $A = \frac{\theta}{\sigma^2}$ and $B = \sqrt{\frac{2\sigma^2}{\kappa} + \theta^2}$ and (11.31) corresponds to MCC (1998) equation 14.

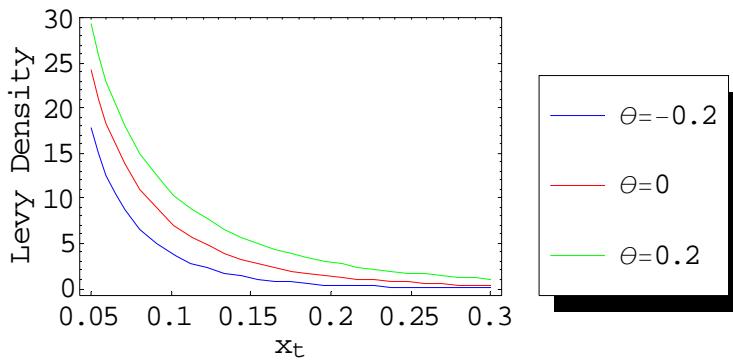
Tails of both Lévy density of the equation (11.31) and probability density of the equation (11.29) of the VG process are exponentially decayed with rates $\lambda_+ = B - A$ (for the upper tail) and $\lambda_- = B + A$ (for the lower tail).

A VG process is an infinite activity Lévy process since the integral of Lévy measure (11.31), $\int_{-\infty}^{\infty} \ell(x)dx$, does not converge in $[0, \infty]$ (i.e. an infinite integral). This means that a VG process has infinitely many small jumps but a finite number of large jumps inheriting the infinite arrival rate of jumps from the gamma subordinator. In other words, the Lévy measure $\ell(x)$ of the VG process is concentrated around zero which will be illustrated soon.

An example of Lévy measure $\ell(x)$ for the VG process is plotted in Figure 11.4 where panel A plots the lower tail and panel B plots the upper tail because $\ell(x)$ becomes complex infinity when $x = 0$ (i.e. infinite arrival rate). Generally, the sign of the drift parameter of the Brownian motion with drift process θ determines the skewness of VG Lévy density as illustrated and the case $\theta = -0.2$ is a mirror image of the case $\theta = 0.2$. This means that if $\theta = 0$, we have a symmetric VG Lévy density. Larger values of the variance rate parameter κ for the gamma subordinator with unit variance rate (i.e. higher degree of randomness of the subordination) leads to lower exponential decay rate of the Lévy measure $\ell(x)$ (i.e. lower B in the above), thus the tails of Lévy measure becomes fatter indicating higher probability of large jumps. This point is illustrated in Figure 11.5.

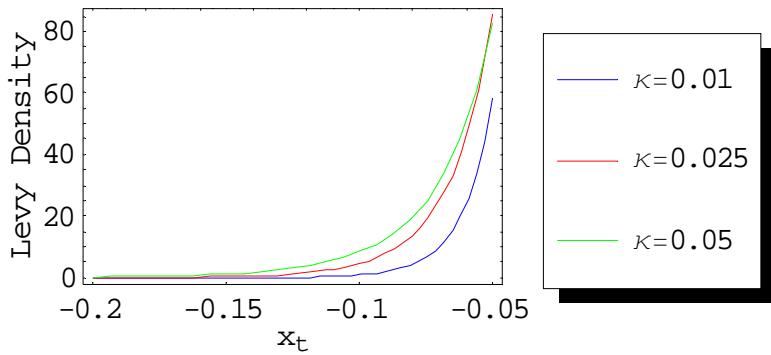


A) Lower tail.

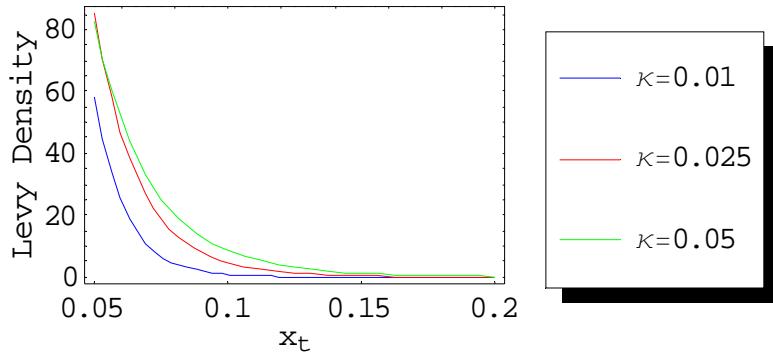


B) Upper tail.

Figure 11.4: Lévy Measure $\ell(x)$ for VG Process with Different Values for the Drift Parameter of the Brownian Motion Process θ . Parameters fixed are $t = 0.5$, $\sigma = 0.2$, and $\kappa = 0.5$.



A) Lower tail.



B) Upper tail.

Figure 11.5: Lévy Measure $\ell(x)$ for VG Process with Different Values for the Variance Rate Parameter κ . Parameters fixed are $t = 0.5$, $\theta = 0$, and $\sigma = 0.2$.

[11.4] Modeling Stock Price Dynamics with VG Process and VG Log-Return Density

We saw in section 7.6 that in the BS case, a stock price process $\{S_t; 0 \leq t \leq T\}$ is modeled as an exponential Lévy process of the form:

$$S_t = S_0 e^{X_t},$$

where $\{X_t; 0 \leq t \leq T\}$ is a Lévy process. BS choose a Brownian motion with drift (continuous diffusion process) as the Lévy process:

$$X_t \equiv \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t.$$

The fact that a stock price S_t is modeled as an exponential of Lévy process X_t means that its log-return $\ln(\frac{S_t}{S_0})$ is modeled as a Lévy process such that:

$$\ln\left(\frac{S_t}{S_0}\right) = X_t = \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t.$$

BS model can be categorized as the only continuous exponential Lévy model apparently because a Brownian motion process is the only continuous (i.e. no jumps) Lévy process.

Now by replacing a standard Brownian motion process B_t with a VG process $VG(x_t; \theta, \sigma, \kappa)$ and κ being the variance rate of the gamma subordinator, stock price

dynamics under historical probability measure \mathbb{P} can be expressed as (MCC (1998) equation 21):

$$S_t = S_0 \exp \left[\left\{ m + \frac{1}{\kappa_{\mathbb{P}}} \ln \left(1 - \theta_{\mathbb{P}} \kappa_{\mathbb{P}} - \frac{\sigma_{\mathbb{P}}^2 \kappa_{\mathbb{P}}}{2} \right) \right\} t + VG(x_t; \theta_{\mathbb{P}}, \sigma_{\mathbb{P}}, \kappa_{\mathbb{P}}) \right], \quad (11.32)$$

where m is the instantaneous rate of return on the stock under \mathbb{P} .

A stock price process $\{S_t; 0 \leq t \leq T\}$ is not a martingale under \mathbb{P} because risk-averse investors expect S_t to grow at a rate greater than the constant risk-free interest rate r :

$$E_t^{\mathbb{P}} [e^{-r\Delta} S_{t+\Delta}] > S_t.$$

Thus, for the purpose of option pricing we convert non-martingales into martingales by changing the probability measure. In other words, we will try to find an equivalent probability measure \mathbb{Q} (called risk-neutral measure) under which the stock price discounted by the risk-free interest rate becomes martingale:

$$E_t^{\mathbb{Q}} [e^{-r\Delta} S_{t+\Delta}] = S_t.$$

This risk-neutral dynamics is given by:

$$S_t = S_0 \exp \left[\left\{ r + \frac{1}{\kappa_{\mathbb{Q}}} \ln \left(1 - \theta_{\mathbb{Q}} \kappa_{\mathbb{Q}} - \frac{\sigma_{\mathbb{Q}}^2 \kappa_{\mathbb{Q}}}{2} \right) \right\} t + VG(x_t; \theta_{\mathbb{Q}}, \sigma_{\mathbb{Q}}, \kappa_{\mathbb{Q}}) \right]. \quad (11.33)$$

In VG model, log-return $z_t \equiv \ln(S_t / S_0)$ is modeled as under historical probability measure \mathbb{P} :

$$z_t = \left\{ m + \frac{1}{\kappa_{\mathbb{P}}} \ln \left(1 - \theta_{\mathbb{P}} \kappa_{\mathbb{P}} - \frac{\sigma_{\mathbb{P}}^2 \kappa_{\mathbb{P}}}{2} \right) \right\} t + VG(x_t; \theta_{\mathbb{P}}, \sigma_{\mathbb{P}}, \kappa_{\mathbb{P}}). \quad (11.34)$$

VG process x_t has VG density of the form of equation (11.29). Use the relationship (11.34):

$$x_t = z_t - \left\{ m + \frac{1}{\kappa_{\mathbb{P}}} \ln \left(1 - \theta_{\mathbb{P}} \kappa_{\mathbb{P}} - \frac{\sigma_{\mathbb{P}}^2 \kappa_{\mathbb{P}}}{2} \right) \right\} t, \quad (11.35)$$

where the term inside the curly bracket is deterministic. Therefore, log-return $z_t \equiv \ln(S_t / S_0)$ density under historical probability measure \mathbb{P} in VG model can be written as (MCC (1998) equation 23):

$$VG(z_t) = \frac{\sqrt{2} \exp\left(\frac{\theta}{\sigma^2} x_t\right)}{\kappa^{t/\kappa} \sigma \sqrt{\pi} \Gamma(t/\kappa)} \left(\frac{x_t^2}{\frac{2\sigma^2}{\kappa} + \theta^2} \right)^{\frac{t-1}{2\kappa-4}} K_{\frac{t-1}{\kappa-2}} \left(\sqrt{x_t^2 \left(\frac{2\sigma^2}{\kappa} + \theta^2 \right)} \right), \quad (11.36)$$

where $x_t = z_t - \left\{ m + \frac{1}{\kappa} \ln \left(1 - \theta \kappa - \frac{\sigma^2 \kappa}{2} \right) \right\} t$. This answers to the statement we made in section 11.2 that those standardized moments of a VG process x_t in the equation (11.30) are not equivalent to those of log-return $z_t \equiv \ln(S_t / S_0)$ density of VG model because of a correction term $-\left\{ m + \frac{1}{\kappa} \ln \left(1 - \theta \kappa - \frac{\sigma^2 \kappa}{2} \right) \right\} t$. VG log-return density $VG(z_t)$ does not have a nice closed form expression for the standardized moments like those for a VG process. Thus, we will simply numerically calculate those when necessary. By taking care of this correction term we plot several figures here for the purpose of illustration. Several very interesting properties of VG log-return density $VG(z_t)$ should be worth mentioning. If $\theta = 0$ (i.e. subordination of zero-drift Brownian motion process with gamma subordinator), VG log-return density is symmetric (zero-skewness). Generally, the sign of the drift parameter of the Brownian motion with drift process θ determines the skewness of VG log-return density as illustrated in Figure 11.6. The case $\theta = -0.2$ is a mirror image of the case $\theta = 0.2$.

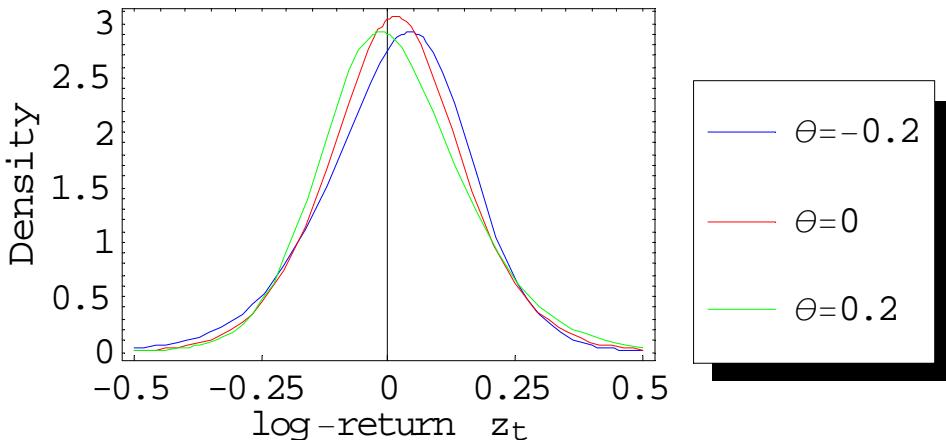


Figure 11.6: VG Log-Return Density for Different Values of Drift Parameter of the Brownian Motion Process θ . Parameters fixed are $t = 0.5$, $\sigma = 0.2$, $\kappa = 0.1$, and $m = 0.05$.

Table 11.4
Annualized Moments of VG Log-Return Density in Figure 11.6

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
$\theta = -0.2$	0.028399	0.209762	-0.392262	0.704132
$\theta = 0$	0.02998	0.2	0	0.6
$\theta = 0.2$	0.275439	0.209762	0.392262	0.704132

Secondly, larger value of variance rate κ (which is the variance rate of the gamma subordinator with $E[S_t] = t$ and determines how random the subordination is) makes the density fatter-tailed and higher-peaked as illustrated in Figure 11.7.

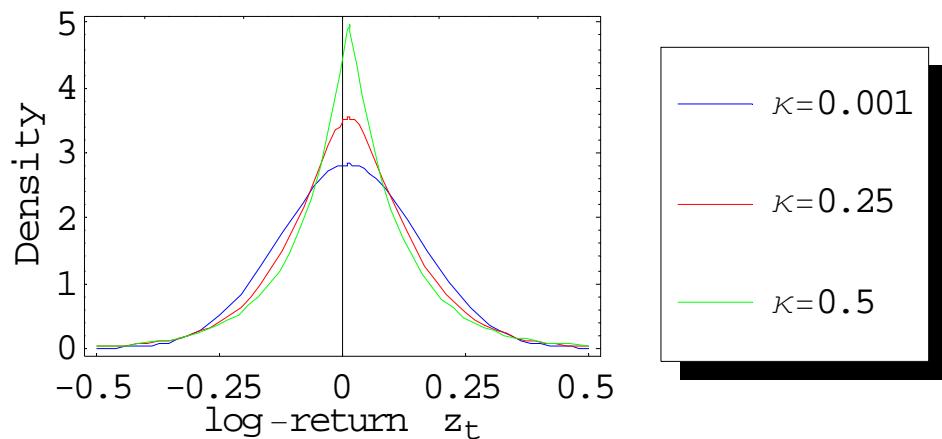


Figure 11.7: VG Log-Return Density for Different Values of Variance Rate Parameter κ . Parameters fixed are $t = 0.5$, $\theta = 0$, $\sigma = 0.2$, and $m = 0.05$.

Table 11.5
Annualized Moments of VG Log-Return Density in Figure 11.7

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
$\kappa = 0.001$	0.0299998	0.2	0	0.006
$\kappa = 0.25$	0.0299498	0.2	0	1.5
$\kappa = 0.5$	0.0298993	0.2	0	3

Also note that VG log return density has higher peak and fatter tails (more leptokurtic) when matched to the BS normal log return density as illustrated in Figure 11.8.

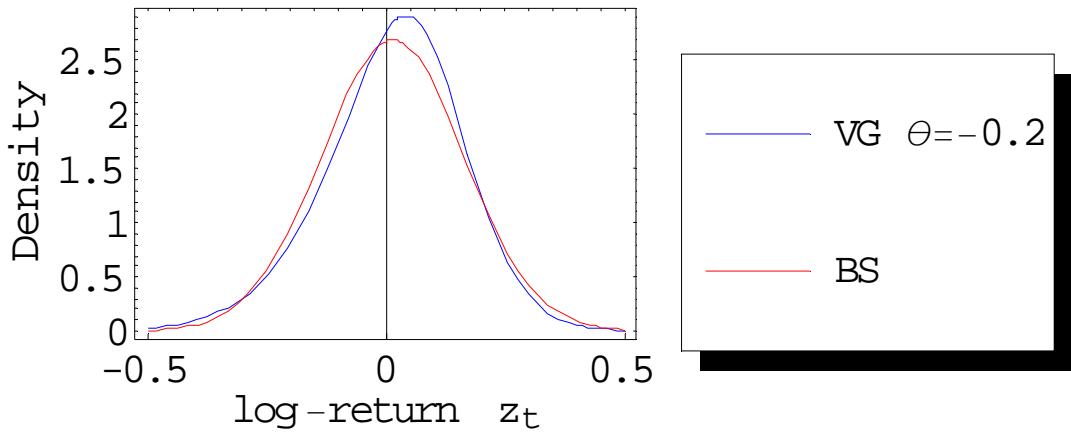


Figure 11.8: VG Log-Return Density vs. Black-Scholes Normal Log-Return Density.
 Parameters fixed for the VG log-return density are $t = 0.5$, $\theta = -0.2$, $\sigma = 0.2$, $\kappa = 0.1$, and $m = 0.05$. Black-Scholes normal density is plotted by matching the mean and variance to the VG.

Table 11.6
Annualized Moments of VG Log-Return Density in Figure 11.8

Model	Mean	Standard Deviation	Skewness	Excess Kurtosis
VG with $\theta = -0.2$	0.0283992	0.209762	-0.392262	0.704132
BS	0.0283992	0.209762	0	0

[11.5] Option Pricing with VG Model

We saw that the risk-neutral dynamics of a stock price is given by the equation (11.33):

$$S_t = S_0 \exp \left[\left\{ r + \frac{1}{\kappa_{\mathbb{Q}}} \ln \left(1 - \theta_{\mathbb{Q}} \kappa_{\mathbb{Q}} - \frac{\sigma_{\mathbb{Q}}^2 \kappa_{\mathbb{Q}}}{2} \right) \right\} t + VG(x_t; \theta_{\mathbb{Q}}, \sigma_{\mathbb{Q}}, \kappa_{\mathbb{Q}}) \right],$$

where $\{x_t; 0 \leq t \leq T\}$ is a VG process on a space $(\Omega, \mathcal{F}, \mathbb{Q})$ and r is the constant risk-free interest rate. Then, a European call option price $C_{VG}(S_t; \tau = T - t)$ is calculated as:

$$C_{VG}(S_t; \tau) = e^{-r\tau} E^{\mathbb{Q}} \left[\max(S_T - K, 0) \middle| \mathcal{F}_t \right].$$

We saw in section 11.2.7 that because of the subordination structure $X_t \equiv \theta(S_t) + \sigma B_{S_t}$ of the VG process $\{X_t; 0 \leq t \leq T\}$, the probability density of VG process can be expressed as

the conditionally normal by conditioning on the fact that the realized value of the random time change by a gamma subordinator S_t with unit mean rate is $S_t = g$:

$$VG(x_t | t = S_t = g) = \frac{1}{\sqrt{2\pi\sigma^2}g} \exp\left\{-\frac{(x_t - \theta g)^2}{2\sigma^2 g}\right\}.$$

Using this conditional normality of VG process, the call price conditioned on the fact that the realized value of the random time change by a gamma subordinator S_t with unit mean rate is $S_t = g$ can be obtained as a BS type formula (equation 6.5 in Madan and Milne (1991)):

$$\begin{aligned} C_{VG}|g = S_t &= S_t \left(1 - \frac{\kappa(\alpha + s)^2}{2}\right)^{\frac{\tau}{\kappa}} \exp\left(\frac{(\alpha + s)^2 g}{2}\right) N\left(\frac{d}{\sqrt{g}} + (\alpha + s)\sqrt{g}\right) \\ &- K \exp(-r\tau) \left(1 - \frac{\kappa\alpha^2}{2}\right)^{\frac{\tau}{\kappa}} \exp\left(\frac{\alpha^2 g}{2}\right) N\left(\frac{d}{\sqrt{g}} + \alpha\sqrt{g}\right), \end{aligned} \quad (11.37)$$

where N is a cumulative distribution function for a standard normal variable. And the parameterizations are follows:

$$\begin{aligned} \zeta &= -\frac{\theta}{\sigma^2}, \quad s = \frac{\sigma}{\sqrt{1 + \frac{\theta^2\kappa}{2\sigma^2}}}, \quad \alpha = \zeta s = -\frac{\theta}{\sqrt{\sigma^2 + \frac{\theta^2\kappa}{2}}}, \quad c_1 = \frac{\kappa}{2}(\alpha + s)^2, \quad c_2 = \frac{\kappa}{2}\alpha^2, \\ \frac{1-c_1}{1-c_2} &= 1 + \kappa(\theta - \frac{\sigma^2}{2}), \text{ and } d = \frac{1}{s} \left[\ln\left(\frac{S_t}{K}\right) + r(T-t) + \frac{T-t}{\kappa} \ln\left(\frac{1-c_1}{1-c_2}\right) \right]. \end{aligned}$$

Unconditional call price is then obtained by integrating the conditional price $C_{VG}|g$ with respect to gamma density of the equation (11.21):

$$C_{VG}(S_t; \tau) = \int_0^\infty (C_{VG}|g) \frac{(1/\kappa)^{t/\kappa}}{\Gamma(t/\kappa)} g^{\frac{t}{\kappa}-1} e^{-g/\kappa} dg. \quad (11.38)$$

MCC (1998) defines a user specified function $\Psi(a, b, \gamma)$:

$$\Psi(a, b, \gamma) = \int_0^\infty N\left(\frac{a}{\sqrt{u}} + b\sqrt{u}\right) \frac{u^{\gamma-1} \exp(-u)}{\Gamma(\gamma)} du. \quad (11.39)$$

After numerous changes of variables, a European call price for VG model is obtained as (MCC (1998) theorem 2):

$$C_{VG}(S_t; \tau) = S_t \Psi \left(d \sqrt{\frac{1-c_1}{\kappa}}, (\alpha + s) \sqrt{\frac{\kappa}{1-c_1}} \frac{\tau}{\kappa} \right) - K \exp(-rt) \Psi \left(d \sqrt{\frac{1-c_2}{\kappa}}, \alpha s \sqrt{\frac{\kappa}{1-c_2}} \frac{\tau}{\kappa} \right). \quad (11.40)$$

Because the equation (11.39) is not a closed-form function, the call price (11.40) is not a closed-form option pricing function. In order to obtain a closed-form option pricing function, MCC (1998) rewrites (11.39) in terms of the modified Bessel function of the second kind K and the confluent hypergeometric function of two variables $\Phi(\alpha, \beta, \gamma; x, y)$ as follows:

$$\begin{aligned} & \Psi(a, b, \gamma) \\ &= \frac{c^{\gamma+1/2} \exp(\text{sign}(a)c)(1+u)^\gamma}{\sqrt{2\pi}\Gamma(\gamma)\gamma} K_{\gamma+1/2}(c) \Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -\text{sign}(a)c(1+u)) \\ & - \text{sign}(a) \frac{c^{\gamma+1/2} \exp(\text{sign}(a)c)(1+u)^{1+\gamma}}{\sqrt{2\pi}\Gamma(\gamma)(1+\gamma)} K_{\gamma-1/2}(c) \Phi(1+\gamma, 1-\gamma, 2+\gamma; \frac{1+u}{2}, -\text{sign}(a)c(1+u)) \\ & + \text{sign}(a) \frac{c^{\gamma+1/2} \exp(\text{sign}(a)c)(1+u)^\gamma}{\sqrt{2\pi}\Gamma(\gamma)\gamma} K_{\gamma-1/2}(c) \Phi(\gamma, 1-\gamma, 1+\gamma; \frac{1+u}{2}, -\text{sign}(a)c(1+u)), \end{aligned}$$

where:

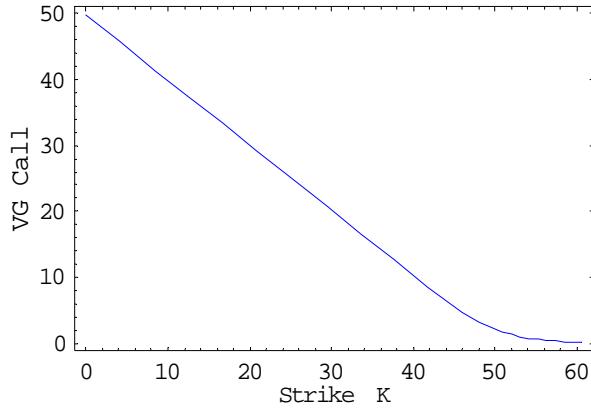
$$\begin{aligned} \Phi(\alpha, \beta, \gamma; x, y) &= \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 u^{\alpha-1} (1-u)^{\gamma-\alpha-1} (1-ux)^{-\beta} \exp(uy) du, \\ c &= |a| \sqrt{2+b^2}, \\ u &= \frac{b}{\sqrt{2+b^2}}. \end{aligned} \quad (11.41)$$

We were unsuccessful implementing the, so called, closed form VG call price of the equation (11.40) with (11.41). We believe this is due to the confluent hypergeometric function of two variables $\Phi(\alpha, \beta, \gamma; x, y)$ having the singularity at $u=1$. Therefore, we use the numerical VG call price of the equation (11.40) with (11.39).

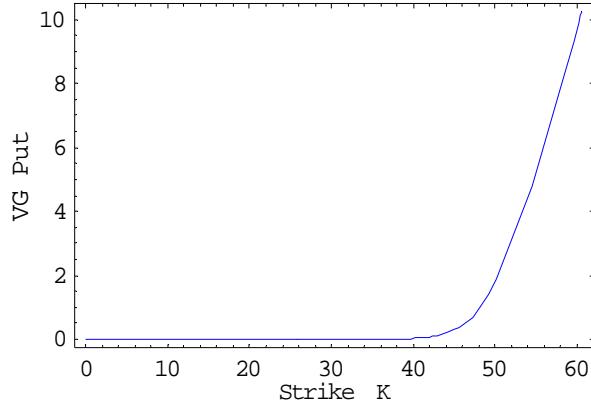
[11.6] Option Pricing Example of VG Model

In this section we calculate the numerical VG call prices of the equation (11.40) with (11.39) and compare them with BS counterparts. Consider pricing a plain vanilla call with common parameters and variables fixed $S_t = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $\tau = 0.25$. VG parameters are set at $\theta = -0.1$ and $\kappa = 0.1$, unless mentioned otherwise.

The first thing we noticed when implementing the numerical VG call price is that its failure to achieve convergence since the integrand (11.39) has a non-integrable singularity at $u = 0$. Practical problem with this singularity is that the numerical VG call price cannot be calculated for deep OTM calls with $60 < K$ as illustrated in Figure 11.9 where the numerical VG call prices are computed for the range of the strike price $0 \leq K \leq 100$. The level of out of moneyness of calls whose prices can be computed with the numerical VG formula depends on the VG parameters θ and κ . This is a big problem for the purpose of calibration because we are mainly interested in OTM options due to their high liquidity. For puts, the numerical VG price cannot be calculated for deep ITM puts. But this won't be a huge problem.



A) Call.

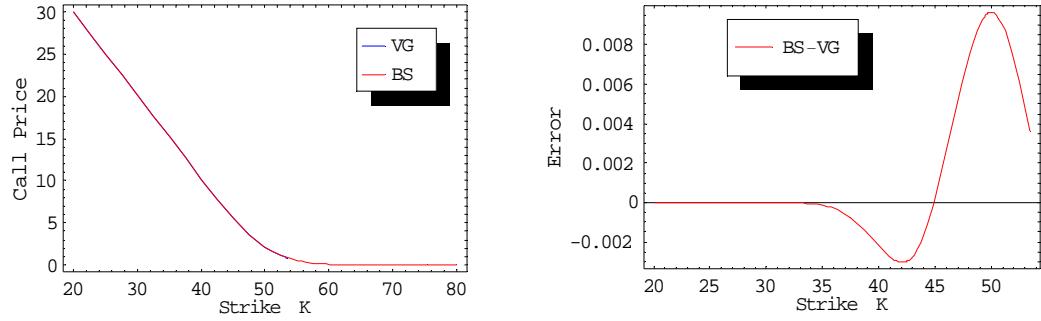


B) Put.

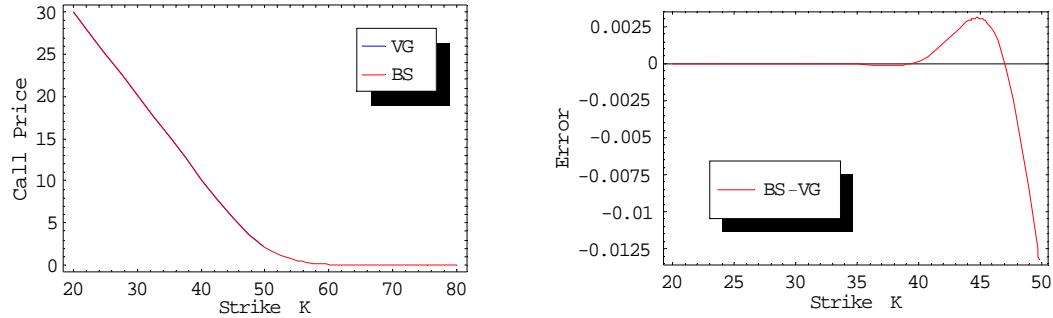
Figure 11.9: Numerical VG Prices. $S_t = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $\tau = 0.25$. VG parameters are set at $\theta = -0.1$ and $\kappa = 0.1$.

Figure 11.10 compares numerical VG call prices with BS counterparts. Note the followings: (1) When VG parameters (θ and κ) are small, these two prices cannot be distinguished by naked eyes in Panel A left. When the difference is plotted on the right, the numerical VG price underprices ATM calls by approximately 0.008. (2) The size of the drift parameter of the subordinated Brownian motion process θ has very little impact on the numerical VG call price. As the size of θ increases from -0.01 to -0.3 (i.e.

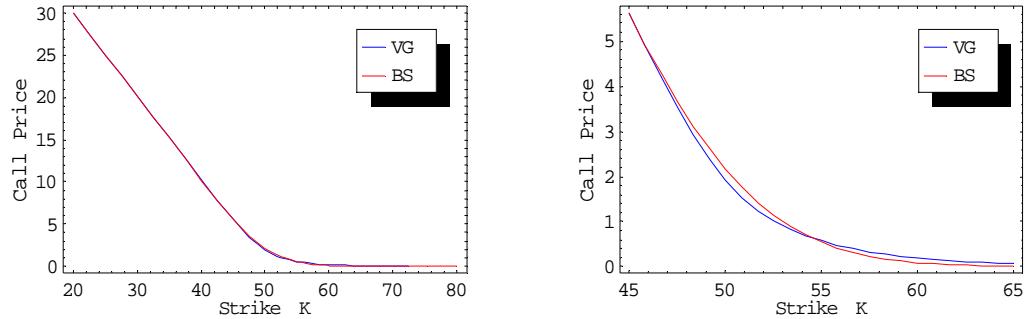
increased skewness of log-return density), the numerical VG price and the BS price are virtually identical as illustrated in Panel B where the numerical VG price overprices ATM call by approximately 0.0125. (3) As the variance rate parameter of the gamma subordinator κ increases (i.e. increased randomness associated with time change), VG model underprices ATM call and overprices OTM call as illustrated in Panel C. Panel D shows the distinct character of the numerical VG call price is that the overvaluation of OTM call.

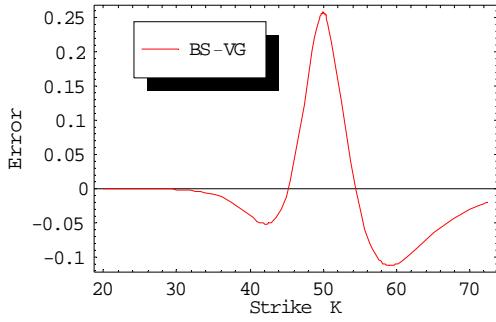


A) VG parameters: $\theta = -0.01$ and $\kappa = 0.01$.

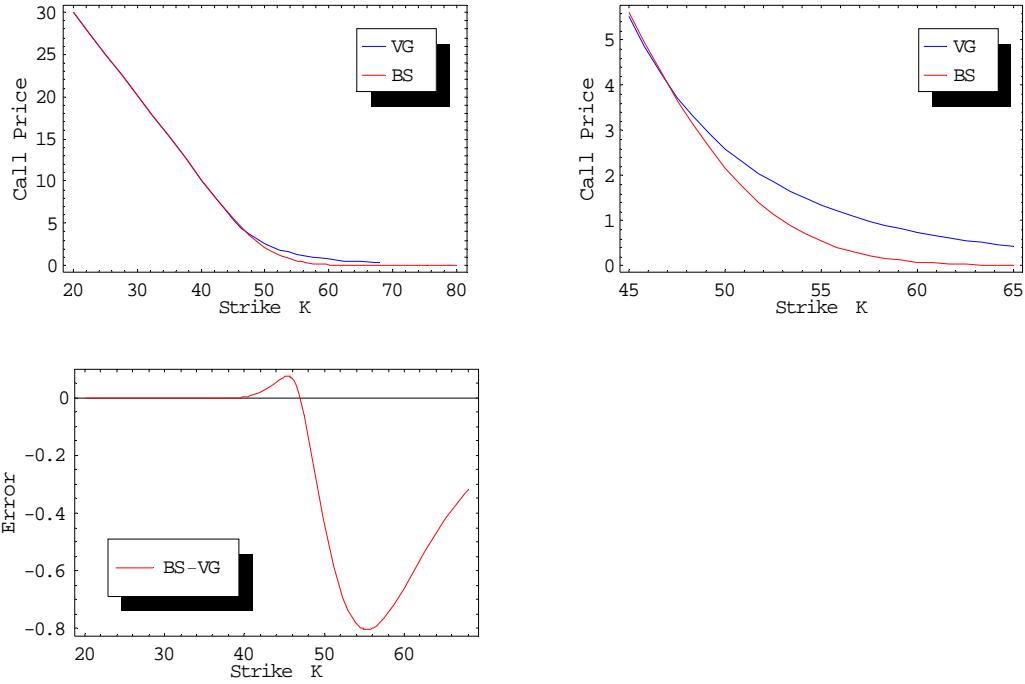


B) VG parameters: $\theta = -0.3$ and $\kappa = 0.01$.





C) VG parameters: $\theta = -0.01$ and $\kappa = 0.3$.

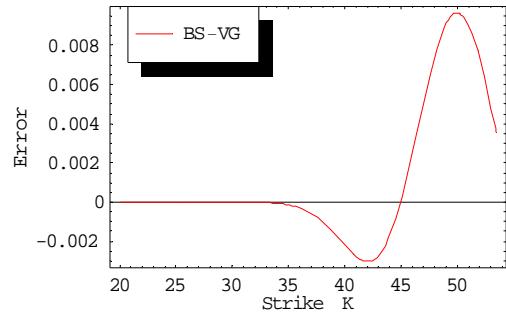
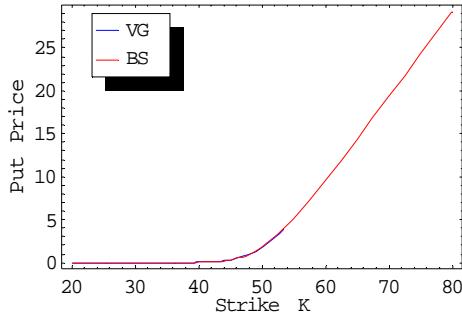


D) VG parameters: $\theta = -0.3$ and $\kappa = 0.3$.

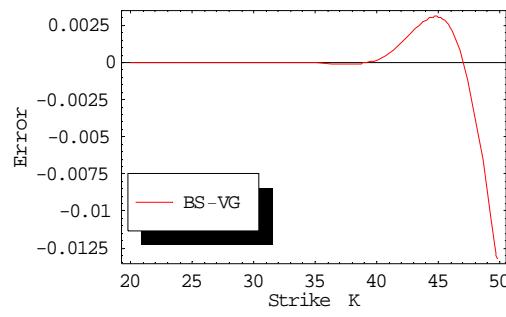
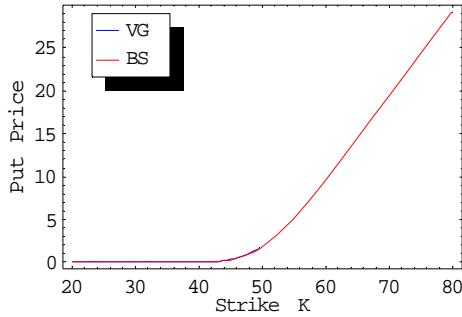
Figure 11.10: VG Call Price vs. Black-Scholes Call Price. Parameters and variables used are $S_t = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $\tau = 0.25$.

Figure 11.11 compares numerical VG put prices with BS counterparts. Note the followings: (1) When VG parameters (θ and κ) are small, these two prices cannot be distinguished by naked eyes in Panel A left. When the difference is plotted on the right, the numerical VG price slightly underprices ATM puts by approximately 0.008. (2) The size of the drift parameter of the subordinated Brownian motion process θ has very little impact on the numerical VG put price. As the size of θ increases from -0.01 to -0.3 (i.e. increased skewness of log-return density), the numerical VG price and the BS price are virtually identical as illustrated in Panel B where the numerical VG price overprices ATM put slightly by approximately 0.0125. (3) As the variance rate parameter of the gamma subordinator κ increases (i.e. increased randomness associated with time

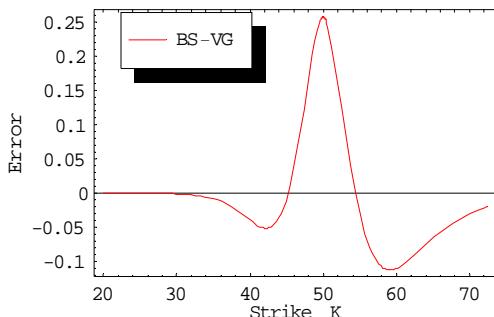
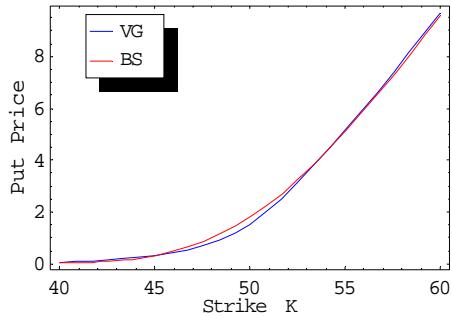
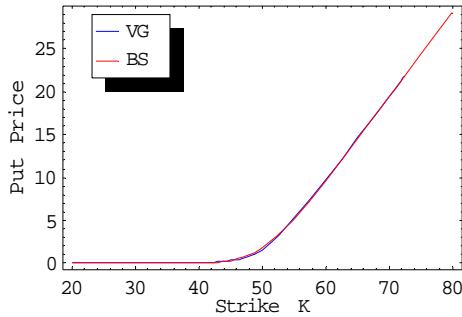
change), VG model underprices ATM puts and overprices ITM calls as illustrated in Panel C. Panel D shows the distinct character of the numerical VG put price is that the overvaluation of ITM put.



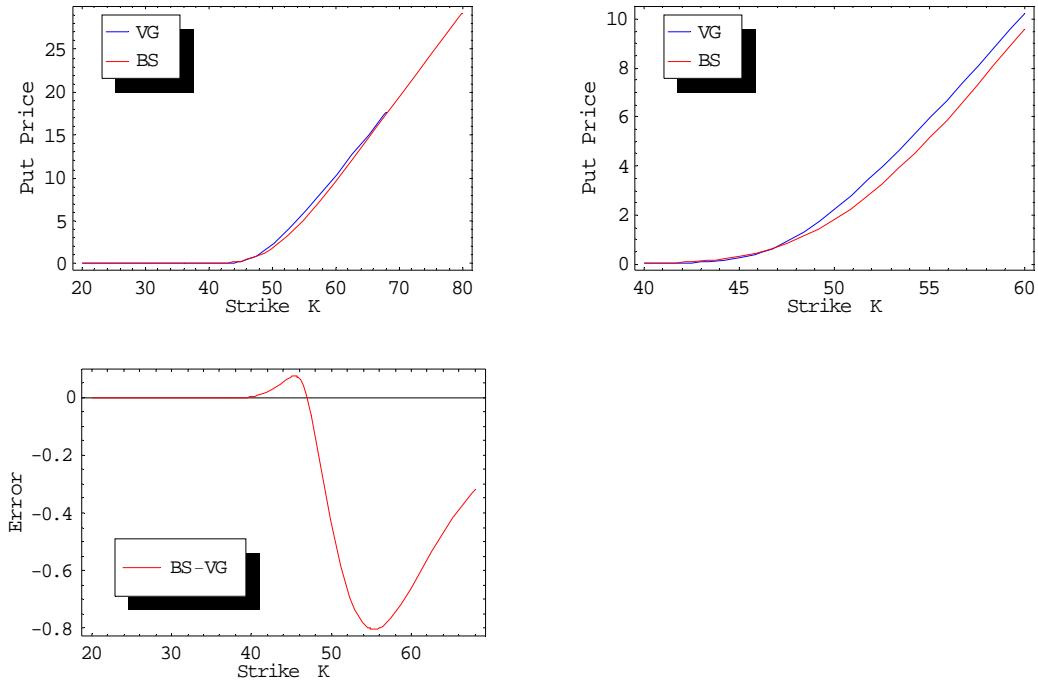
A) VG parameters: $\theta = -0.01$ and $\kappa = 0.01$.



B) VG parameters: $\theta = -0.3$ and $\kappa = 0.01$.



C) VG parameters: $\theta = -0.01$ and $\kappa = 0.3$.



D) VG parameters: $\theta = -0.3$ and $\kappa = 0.3$.

Figure 11.11: VG Put Price vs. Black-Scholes Put Price. Parameters and variables used are $S_t = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $\tau = 0.25$.

[11.7] Lévy Measure of VG Log-Return z_t

We saw in section 11.3 that the Lévy measure of the VG process is given by the equation (11.31). We saw in section 11.4 that the log-return $z_t \equiv \ln(S_t / S_0)$ and a VG process x_t are related by the equation (11.35). Therefore, the Lévy measure of the VG log-return z_t can be expressed as:

$$\ell(z) = \frac{1}{\kappa|x|} \exp\left\{Ax - B|x|\right\}, \quad (11.43)$$

where:

$$x = z - \left\{ m + \frac{1}{\kappa_{\mathbb{P}}} \ln \left(1 - \theta_{\mathbb{P}} \kappa_{\mathbb{P}} - \frac{\sigma_{\mathbb{P}}^2 \kappa_{\mathbb{P}}}{2} \right) \right\},$$

$$A = \frac{\theta_{\mathbb{P}}}{\sigma_{\mathbb{P}}^2},$$

$$B = \sqrt{\frac{2\sigma_{\mathbb{P}}^2}{\kappa_{\mathbb{P}}} + \theta_{\mathbb{P}}^2}.$$

An example of Lévy measure of the VG log-return z_t is plotted in Figure 11.12.

Generally, the sign of the drift parameter of the subordinated Brownian motion with drift process θ determines the skewness of Lévy measure $\ell(z)$ of the VG log-return z_t as illustrated and the case $\theta = -0.2$ is a mirror image of the case $\theta = 0.2$. This means that if $\theta = 0$, we have a symmetric Lévy measure $\ell(z)$. Larger values of the variance rate parameter κ for the gamma subordinator with unit variance rate (i.e. higher degree of randomness of the subordination) leads to lower exponential decay rate of the Lévy measure $v(dx)$, thus the tails of Lévy measure becomes fatter indicating higher probability of large jumps. This point is illustrated in Figure 11.13.

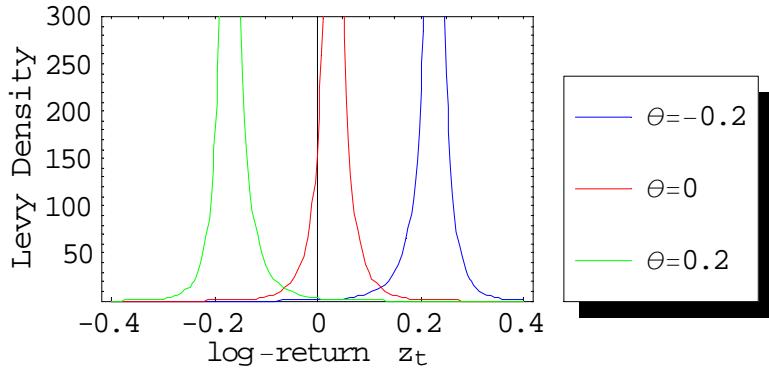


Figure 11.12: Lévy Measures of VG Log-Return z_t with Different Values for the Drift Parameter of the Subordinated Brownian Motion Process θ . Parameters fixed are $\sigma = 0.2$, $\kappa = 0.1$, and $m = 0.05$.

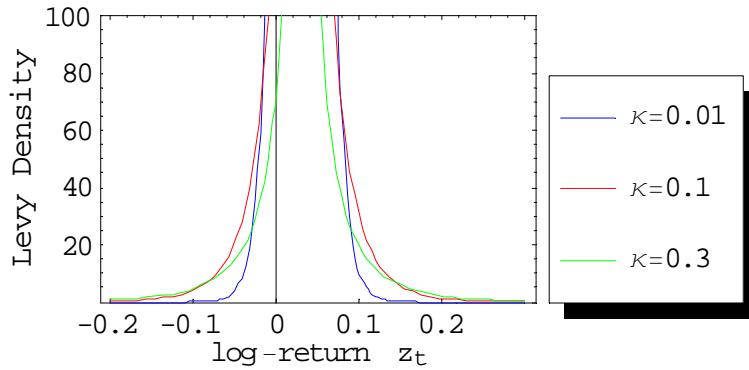


Figure 11.13: Lévy Measures of VG Log-Return z_t with Different Values for the Variance Rate Parameter of the Gamma Subordinator κ . Parameters fixed are $\sigma = 0.2$, $\theta = 0$, and $m = 0.05$.

[12] VG (Variance Gamma) Model with Fourier Transform Pricing

We saw in section 11 that the most important character of VG model is its conditional normal property. Because of the subordination structure $X_t \equiv \theta(S_t) + \sigma B_{S_t}$ of the VG process $\{X_t; 0 \leq t \leq T\}$, the probability density of VG process can be expressed as the conditionally normal by conditioning on the fact that the realized value of the random time change by a gamma subordinator S_t with unit mean rate is $S_t = g$. Similarly, the conditional call price can be obtained as a BS type formula of (11.37) and the unconditional call price is then obtained by integrating the conditional price with respect to gamma density. As a result, a call price can be expressed as single numerical integration problem of the equation (11.40) with (11.39).

But as we realize MCC (1998) use very complicated process of numerous changes of variables in order to obtain a closed form call price of the equation (11.40) with (11.41) and to express the call price in terms of special functions of mathematics. This procedure is not only cumbersome, but also deficient of generality.

As far as the numerical VG call price is concerned, it is simpler to implement, but it has a critical problem of failing to achieve convergence for deep OTM calls since the integrand (11.39) has a non-integrable singularity at $u = 0$.

Thus, we are in desperate need of more general and simpler method of pricing for VG model and other exponential Lévy models. In this section we present our version of Fourier transform option pricing with VG model. We would like readers to appreciate its generality and simplicity.

[12.1] VG Model with Fourier Transform Pricing Method

The first step to FT option pricing is to obtain the CF of the log stock price $\ln S_t$. Risk-neutral log stock price dynamics can be obtained from the equation (11.33) as:

$$\ln S_t = \ln S_0 + \left\{ r + \frac{1}{\kappa_Q} \ln \left(1 - \theta_Q \kappa_Q - \frac{\sigma_Q^2 \kappa_Q}{2} \right) \right\} t + VG(x_t; \theta_Q, \sigma_Q, \kappa_Q). \quad (12.1)$$

VG process x_t has VG density of the form of equation (11.29). After rearrangement:

$$x_t = \ln S_t - \left[\ln S_0 + \left\{ r + \frac{1}{\kappa_Q} \ln \left(1 - \theta_Q \kappa_Q - \frac{\sigma_Q^2 \kappa_Q}{2} \right) \right\} t \right], \quad (12.2)$$

where the term inside the square bracket is deterministic. Therefore, risk-neutral log stock price density in VG model can be written as:

$$VG(\ln S_t) = \frac{\sqrt{2} \exp\left(\frac{\theta}{\sigma^2} x_t\right)}{\kappa^{t/\kappa} \sigma \sqrt{\pi} \Gamma(t/\kappa)} \left(\frac{x_t^2}{\frac{2\sigma^2}{\kappa} + \theta^2} \right)^{\frac{t}{2\kappa} - \frac{1}{4}} K_{\frac{t}{\kappa} - \frac{1}{2}} \left(\frac{\sqrt{x_t^2 \left(\frac{2\sigma^2}{\kappa} + \theta^2 \right)}}{\sigma^2} \right), \quad (12.3)$$

where $x_t = \ln S_t - \left[\ln S_0 + \left\{ r + \frac{1}{\kappa_Q} \ln \left(1 - \theta_Q \kappa_Q - \frac{\sigma_Q^2 \kappa_Q}{2} \right) \right\} t \right]$. By Fourier transforming

(12.3) with FT parameters $(a, b,) = (1, 1)$, its CF can be calculated. But this direct computation of CF is difficult because it involves the Bessel function.

Consider a following example. The CF function of a multiplicative standard Brownian motion process $\{\sigma B_t; 0 \leq t \leq T\}$ is:

$$\begin{aligned} \phi_B(\omega) &\equiv \int_{-\infty}^{\infty} e^{i\omega B_t} \mathbb{P}(B_t) dB_t = \int_{-\infty}^{\infty} e^{i\omega B_t} \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{B_t^2}{2\sigma^2 t}\right) dB_t \\ &= \exp\left(-\frac{\sigma^2 t \omega^2}{2}\right). \end{aligned}$$

When the drift term At is added, the CF function of a Brownian motion with drift process $\{W_t \equiv At + \sigma B_t; 0 \leq t \leq T\}$ is:

$$\begin{aligned} \phi_W(\omega) &\equiv \int_{-\infty}^{\infty} e^{i\omega W_t} \mathbb{P}(W_t) dW_t = \int_{-\infty}^{\infty} e^{i\omega W_t} \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left\{-\frac{(W_t - At)^2}{2\sigma^2 t}\right\} dW_t \\ &= \exp\left(iAt\omega - \frac{\sigma^2 t \omega^2}{2}\right) = \exp(iAt\omega) \exp\left(-\frac{\sigma^2 t \omega^2}{2}\right). \end{aligned}$$

This implies that the CF of a Brownian motion with drift process can be obtained by simple multiplication of the CF function of a multiplicative standard Brownian motion process and $\exp(iAt\omega)$. Let's use this logic in our context. We derived the CF of a VG process x_t which is given by the equation (11.27):

$$\phi_x(\omega) = \left(1 + \frac{\sigma^2 \omega^2 \kappa}{2} - i\theta \kappa \omega \right)^{-\frac{t}{\kappa}}. \quad (11.27)$$

From the VG risk-neutral log stock price dynamics of the equation (12.1), by taking care of the drift term $At \equiv \ln S_0 + \left\{ r + \frac{1}{\kappa_Q} \ln \left(1 - \theta_Q \kappa_Q - \frac{\sigma_Q^2 \kappa_Q}{2} \right) \right\} t$, the CF of log stock price can be expressed as:

$$\phi_{\ln S_t}(\omega) = \exp \left[i\omega \left\{ \ln S_0 + \left(r + \frac{1}{\kappa_Q} \ln(1 - \theta_Q \kappa_Q - \frac{\sigma_Q^2 \kappa_Q}{2}) \right) t \right\} \right] \left(1 + \frac{\sigma^2 \omega^2 \kappa}{2} - i\theta \kappa \omega \right)^{-\frac{t}{\kappa}}.$$

Set $t = T$ and $\ln S_T \equiv s_T$. And drop \mathbb{Q} :

$$\phi_T(\omega) = \exp \left[i\omega \left\{ s_0 + \left(r + \frac{1}{\kappa} \ln(1 - \theta \kappa - \frac{\sigma^2 \kappa}{2}) \right) T \right\} \right] \left(1 + \frac{\sigma^2 \omega^2 \kappa}{2} - i\theta \kappa \omega \right)^{-\frac{T}{\kappa}}. \quad (12.4)$$

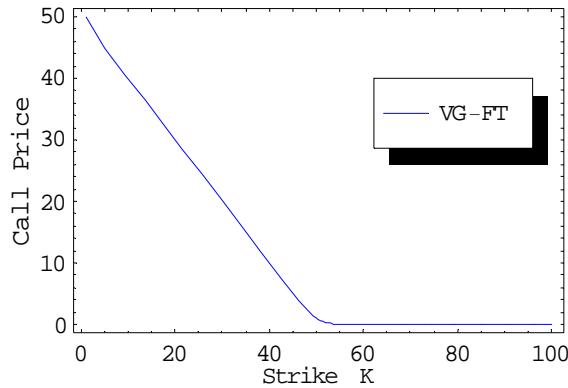
By substituting the above CF into the general FT call pricing formula of the equation (8.17), we obtain VG-FT call pricing formula:

$$C_{VG-FT}(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha + 1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha + 1)\omega} d\omega, \quad (12.5)$$

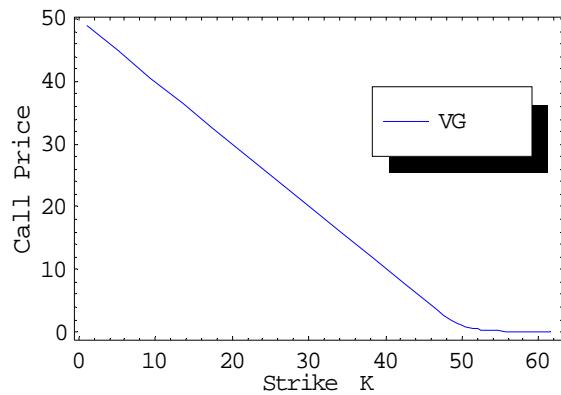
where:

$$\phi_T(\omega) = \exp \left[i\omega \left\{ s_0 + \left(r + \frac{1}{\kappa} \ln(1 - \theta \kappa - \frac{\sigma^2 \kappa}{2}) \right) T \right\} \right] \left(1 + \frac{\sigma^2 \omega^2 \kappa}{2} - i\theta \kappa \omega \right)^{-\frac{T}{\kappa}}.$$

We implement the VG-FT formula (12.5) with decay rate parameter $\alpha = 1.5$ and compare the result to the numerical VG call price using common parameters and variables fixed $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$. Figure 12.1 demonstrates two important points. One is that as previously mentioned, the numerical VG price cannot be calculated for deep OTM calls because they fail to converge which is shown in Panel B. The other is that although the VG-FT price can be calculated for deep OTM calls but they also fail to achieve convergence at the acceptable level of accuracy. On top of this, the additional issue with VG-FT price is that it oscillates severely for the deep ITM and OTM calls as illustrated by Figure 12.2. Especially for the deep OTM calls, this oscillation yields negative call prices. This interesting oscillation pattern for deep OTM calls occur for the BS-FT call price with much smaller degree. Merton-FT call price, in contrast, has no oscillation.

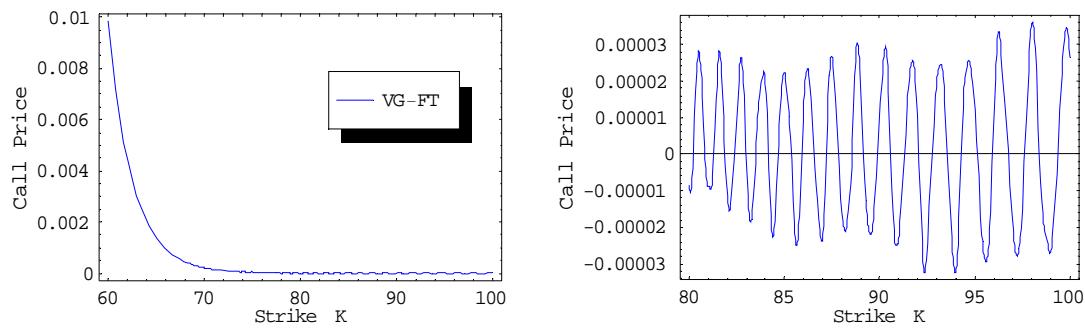


A) VG-FT $\alpha = 1.5$ call price.

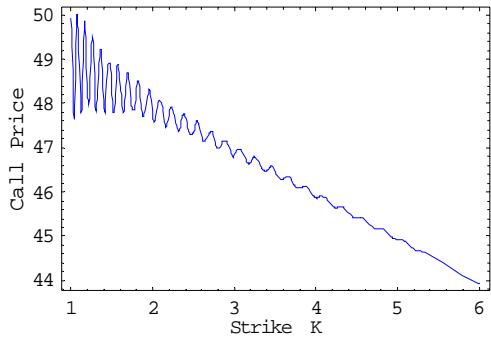


B) Numerical VG call price.

Figure 12.1: VG-FT $\alpha = 1.5$ Call Price Vs. Numerical VG Call Price. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$.



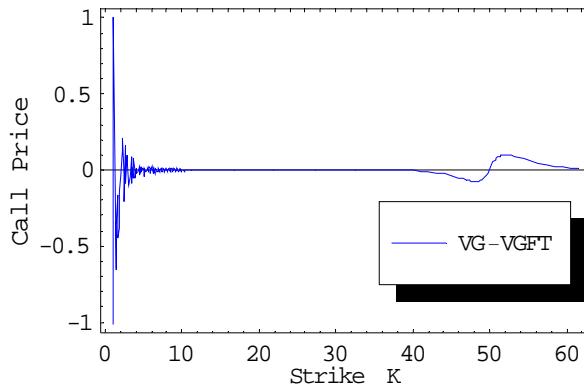
A) Deep OTM call price for $60 \leq K \leq 100$ on the left and for $80 \leq K \leq 100$ on the right.



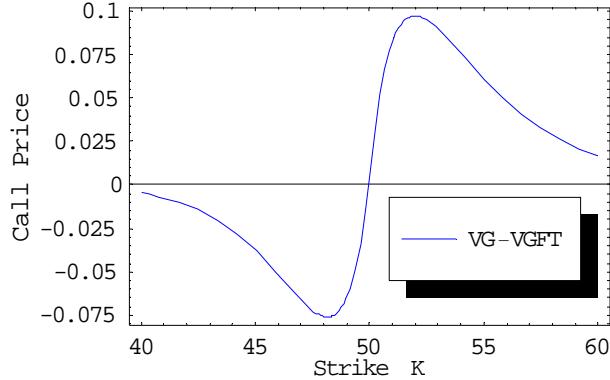
B) Deep ITM call price for $1 \leq K \leq 6$.

Figure 12.2: Oscillation of VG-FT $\alpha = 1.5$ Price for Deep ITM and Deep OTM Calls.
 Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$.

As a principle, VG-FT $\alpha = 1.5$ call price and the numerical VG call price are same (i.e. they are supposed to be). Figure 12.3 plots the difference between the numerical VG and VG-FT call price for the range of strikes with which the numerical VG price can be computed. Panel A shows that the above mentioned oscillation of VG-FT price causes the price difference to oscillate for deep ITM calls of $1 \leq K \leq 5$. For example, for $K = 1$ call the numerical VG price is 48.9247 while VG-FT price is 49.9419 leading to an error of 1.0173. But deep ITM calls are not most researchers' interests. More important error occurs around ATM calls of the strike $45 \leq K \leq 55$ which is illustrated in Panel B. The maximum size of the error is about 0.1.



A) For the range of strike price $1 \leq K \leq 100$.



B) For the range of strike price $40 \leq K \leq 60$

Figure 12.3: Numerical VG Call Price Minus VG-FT $\alpha = 1.5$ Call Price. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$.

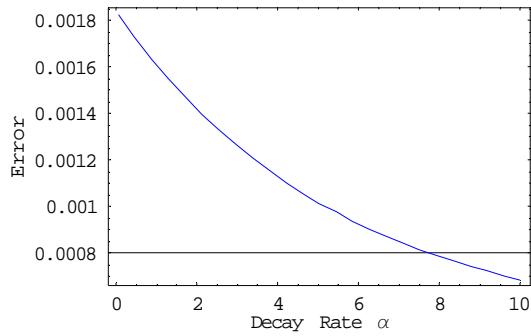
Next, we discuss CPU time considering calls with $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, and $T = 20/252$ as a function of varying strike price K . VG parameters are set as $\theta = -0.1$ and $\kappa = 0.1$. Table 12.1 reveals that VG-FT formula is slower than the numerical VG formula, but speed is not an issue for most of the pricing purposes.

Table 12.1: CPU Time for Calls with Different Moneyness

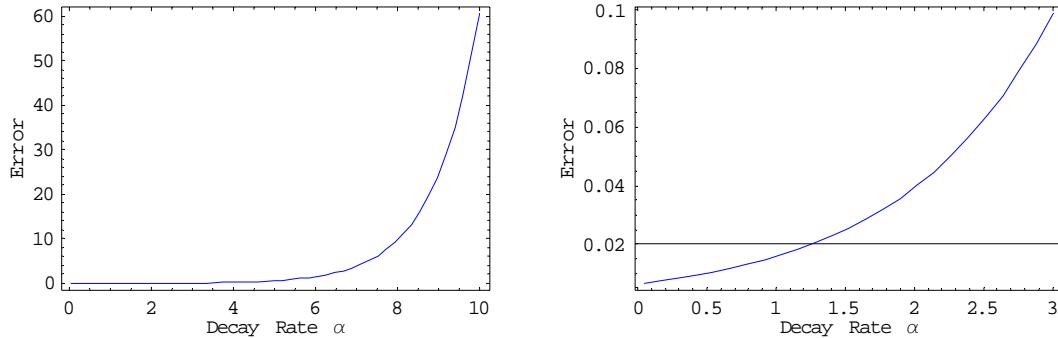
Price is in the bracket.

Method	Strike		
	$K = 20$	$K = 50$	$K = 80$
VG (numerical)	0.08 seconds (29.9999)	0.05 seconds (1.04663)	0.11 seconds (NA)
VG-FT $\alpha = 1$	0.17 seconds (30)	0.09 seconds (1.04107)	0.16 seconds $(-8.41091 \times 10^{-6})$

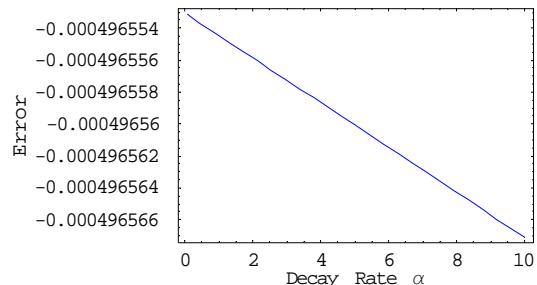
We investigate the level of decay rate parameter α for VG-FT formula. Figure 12.4 illustrates the pricing error of VG-FT price relative to the numerical VG price for a one day to maturity $T = 1/252$ call as a function of varying α and Figure 12.5 is for one month to maturity $T = 20/252$ call. First of all, Panel A and C of Figure 12.4 and 12.5 tell us that the size of the error does not vary significantly for $0.05 \leq \alpha \leq 10$ in the case of OTM and ATM call. Small exception is for OTM call with $T = 1/252$ in Panel A of Figure 12.4 where larger α makes the error smaller. While Panel B in Figure 12.4 and 5 shows that smaller α is better for deep ITM calls. Judging from these findings, we recommend the choice of $1 \leq \alpha \leq 2$ for the purpose of general pricing (i.e. ATM, OTM, and ITM calls) which includes the suggested value by MCC (1998) of $\alpha = 1.5$.



A) For an OTM Call with $K = 60$. Numerical VG price cannot be calculated for deep OTM calls with $K = 70$ or 80 .

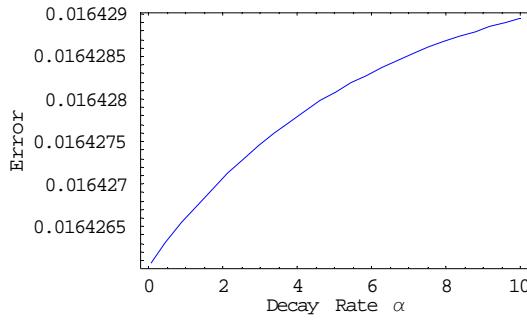


B) For an ITM Call with $K = 20$.

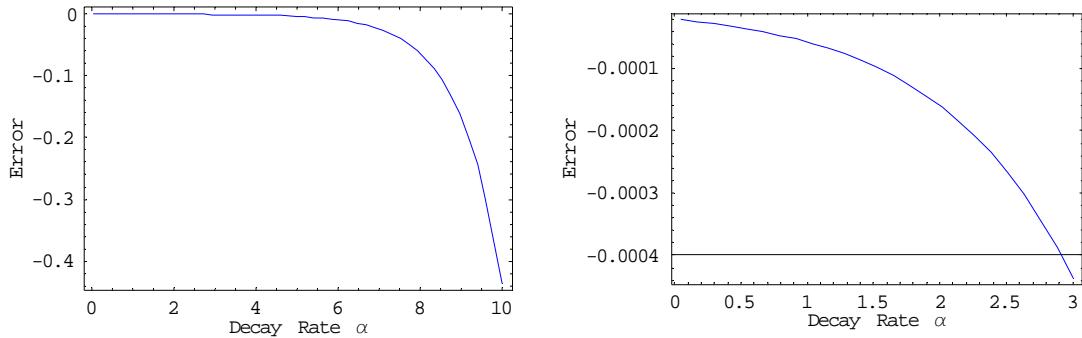


C) For an ATM Call with $K = 50$.

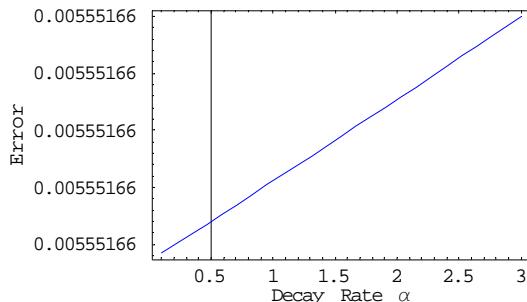
Figure 12.4: Numerical VG Price Minus VG-FT Price for One Day to Maturity Call as a Function of Decay Rate Parameter $0.05 \leq \alpha \leq 10$. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 1/252$, $\theta = -0.1$, and $\kappa = 0.1$.



A) For an OTM Call with $K = 60$. Numerical VG price cannot be calculated for deep OTM calls with $K = 70$ or 80 .



B) For an ITM Call with $K = 20$.



C) For an ATM Call with $K = 50$.

Figure 12.5: Numerical VG Price Minus VG-FT Price for Call with $T = 20/252$ as a Function of Decay Rate Parameter $0.05 \leq \alpha \leq 10$. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $\theta = -0.1$, and $\kappa = 0.1$.

Which is better, VG-FT price of the equation (12.5) or the numerical VG price (11.40) with (11.39)? It is a tough call. Numerical VG call price has a critical problem of failing to achieve convergence for deep OTM calls since the integrand (11.39) has a non-integrable singularity at $u = 0$. VG-FT formula calculates the prices for calls regardless of the moneyness, but it also has convergence problem and oscillation problem for deep ITM and OTM calls which yields negative prices. Can DFT solve these problems

associated with FT price? Let's see in the next section. The value of the decay rate parameter $\alpha = 1.5$ suggested by MCC (1998) seems to be appropriate.

[12.2] Discrete Fourier Transform (DFT) Call Pricing Formula with VG Model

To improve computational time of the VG-FT call price, we apply our version of DFT call price formula.

VG-DFT call price can be simply obtained by substituting the CF of VG log stock price of the equation (12.5) into the general DFT call price of the equation (8.38):

$$C_T(k_n) \approx \frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k} \times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\} \exp(-i2\pi jn/N), \quad (12.6)$$

where $w_{j \in [0, N-1]}$ are trapezoidal rule weights:

$$w_j = \begin{cases} 1/2 & \text{for } j = 0 \text{ and } N-1 \\ 1 & \text{for others} \end{cases},$$

and:

$$\psi_T(\omega) = \frac{e^{-rT} \phi_T(\omega - (\alpha + 1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha + 1)\omega},$$

with $\phi_T(\cdot)$ given by (12.4).

[12.3] Implementation and Performance of DFT Pricing Method with VG Model

In this section, performance of VG-DFT $\alpha = 1.5$ call price of the equation (12.6) is tested by comparing results to the VG-FT $\alpha = 1.5$ call price of the equation (12.5) under various settings. VG-DFT $\alpha = 1.5$ call price is implemented using $N = 4096$ samples and log strike space sampling interval $\Delta k = 0.005$. This corresponds to angular frequency domain sampling interval of $\Delta\omega = 0.306796$ radians, the total sampling range in the log strike space is $K = N\Delta k = 20.48$, its sampling rate is 200 samples per unit of k , and the total sampling range in the angular frequency domain is $\Omega = N\Delta\omega = 1256.64$.

Firstly, let's investigate the difference in price and CPU time. Consider calculating 100-point call prices for a range of strike price $1 \leq K \leq 100$ with interval 1 with common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$. Figure 12.6 reports the price and Table 12.2 compares CPU time. We notice

that prices are basically same (i.e. they are supposed to be same because DFT is an approximation of FT) and the use of DFT significantly improves the computational time.

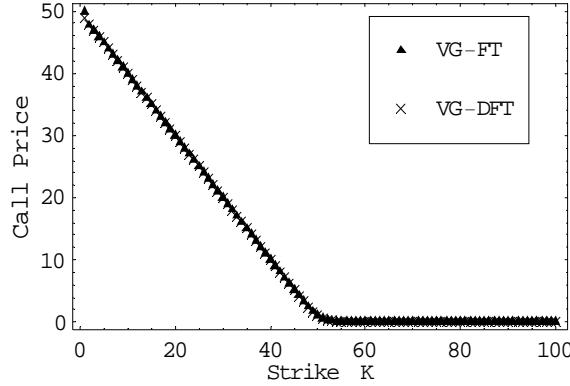
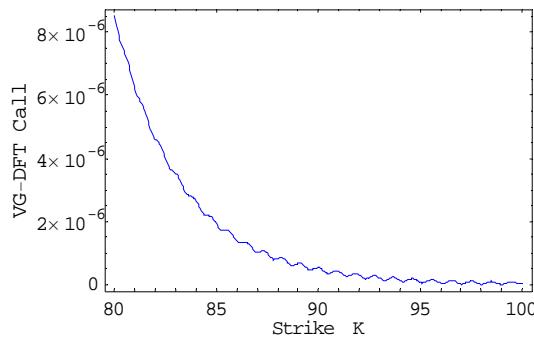


Figure 12.6: VG-FT $\alpha = 1.5$ Vs. VG-DFT $\alpha = 1.5$. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$.

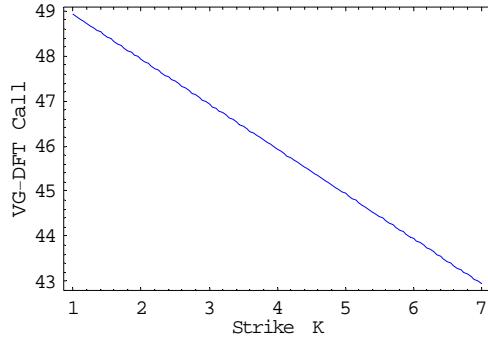
Table 12.2 CPU Time for Calculating 100-point call prices for a range of strike price $1 \leq K \leq 100$ with interval 1 Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$.

Method	CPU Time
VG-FT $\alpha = 1.5$	15.512 seconds
VG-DFT $\alpha = 1.5$ $N = 4096$, $\Delta k = 0.005$	0.571 seconds

Secondly, we saw in section 12.1 that VG-FT price has an oscillation pattern for deep ITM and OTM calls. We find that VG-DFT price has almost zero oscillation for deep ITM calls and very small (i.e. negligible) oscillation for deep OTM calls demonstrated by Figure 12.7.



A) Deep OTM call price for $80 \leq K \leq 100$.



B) Deep ITM call price for $1 \leq K \leq 7$.

Figure 12.7: Oscillation of VG-DFT $\alpha = 1.5$ Price for Deep ITM and Deep OTM Calls. Common parameters and variables fixed are $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, $q = 0.02$, $T = 20/252$, $\theta = -0.1$, and $\kappa = 0.1$.

Thirdly, we pay attention to the price difference between VG-DFT and VG-FT especially for very near-maturity calls. Consider a call option with common parameters and variables $S_0 = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$. VG parameters are set as $\theta = -0.1$, and $\kappa = 0.1$. Figures 12.8 to 12.10 plot three series of price differences computed by the VG-FT $\alpha = 1.5$ and VG-DFT $\alpha = 1.5$ with $N = 4096$ and $\Delta k = 0.005$ as a function of time to maturity of less than a month $1/252 \leq T \leq 20/252$. These figures demonstrate that, as we expected, the price difference tends to increases as the maturity nears regardless of the moneyness of the call. We consider the maximum price difference of the size 0.025 negligible.

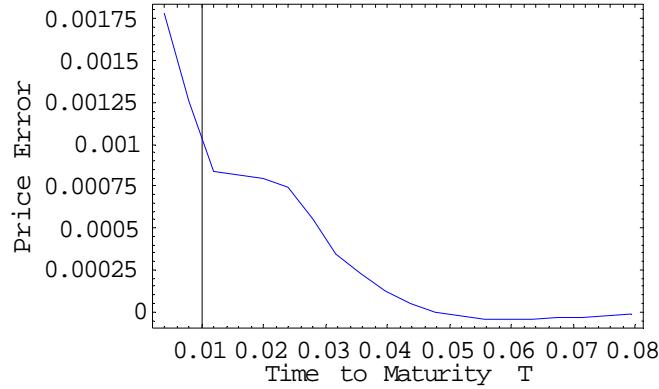


Figure 12.8: Plot of VG-FT $\alpha = 1.5$ Price Minus VG-DFT $\alpha = 1.5$ Price for Deep-OTM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 80$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$, $\theta = -0.1$, and $\kappa = 0.1$.

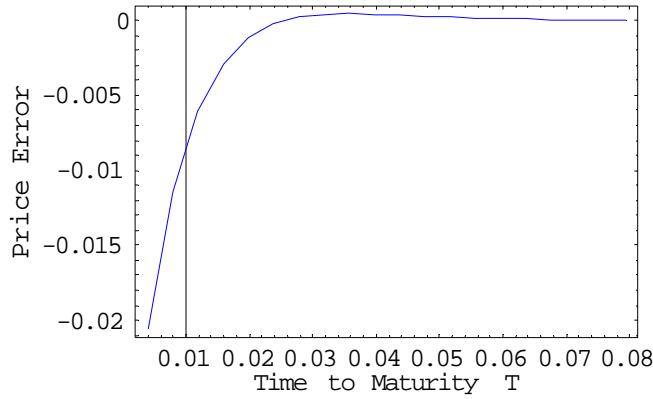


Figure 12.9: Plot of VG-FT $\alpha = 1.5$ Price Minus VG-DFT $\alpha = 1.5$ Price for ATM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$, $\theta = -0.1$, and $\kappa = 0.1$.

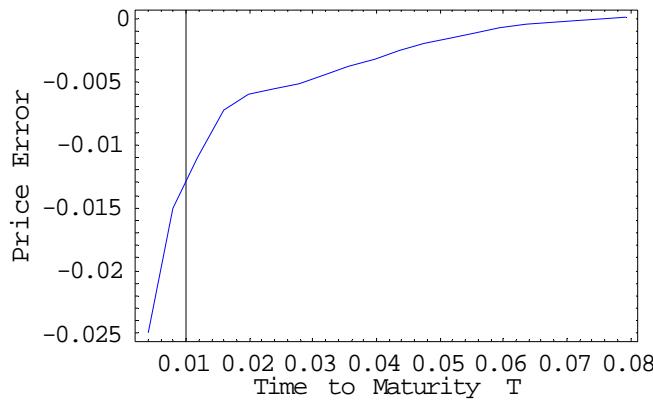
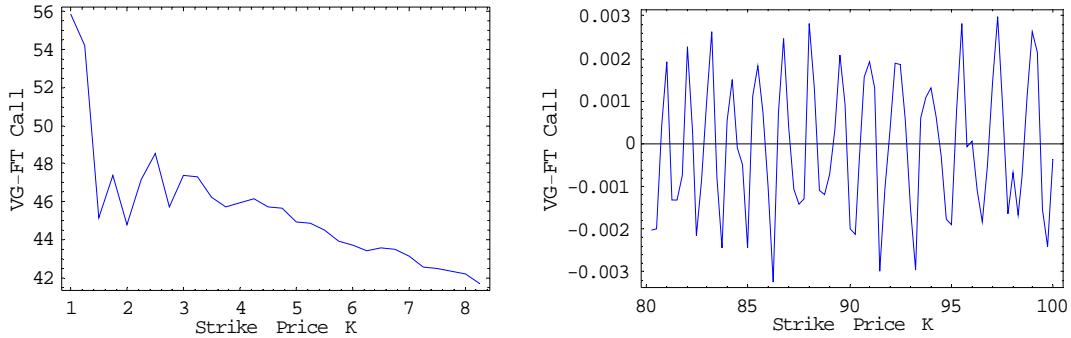
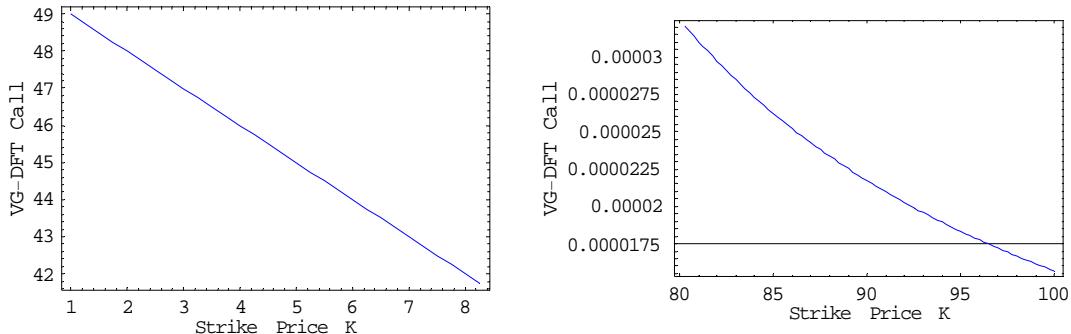


Figure 12.10: Plot of VG-FT $\alpha = 1.5$ Price Minus VG-DFT $\alpha = 1.5$ Price for Deep-ITM Call as a Function of Time to Maturity $1/252 \leq T \leq 20/252$. Common parameters and variables fixed are $S_0 = 50$, $K = 20$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$, $\theta = -0.1$, and $\kappa = 0.1$.

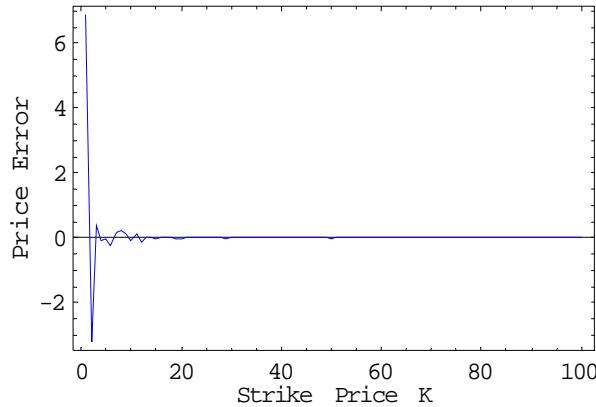
Figure 12.11 clarifies the superiority of VG-DFT price over VG-FT price for one day to maturity call. Panel A exhibits the oscillation of VG-FT price for deep ITM and OTM calls. In contrast, VG-DFT price has almost zero oscillation shown in Panel B. Panel C reminds us that as a principle these two prices are same despite the large deviation for deep ITM calls which is not so important.



A) Oscillation of VG-FT prices for Deep ITM and OTM calls.



B) Almost zero oscillation of VG-DFT prices for Deep ITM and OTM calls.



C) VG-FT Call Price – VG-DFT Call Price.

Figure 12.11: VG-DFT $\alpha = 1.5$ Price Vs. VG-FT $\alpha = 1.5$ Price for 1-Day-to-Maturity ($T = 1/252$). Common parameters and variables fixed are $S_0 = 50$, $K = 50$, $\sigma = 0.2$, $r = 0.05$, and $q = 0.02$, $\theta = -0.1$, and $\kappa = 0.1$.

We conclude this section by stating the following remarks. As a principle, VG-DFT $\alpha = 1.5$ price and VG-FT $\alpha = 1.5$ price are same. Which is better? The clear-cut answer is VG-DFT price. One reason is that VG-DFT needs significantly smaller CPU time to compute hundreds of option prices which is very important for the calibration. Another reason is that VG-DFT price does not have the oscillation pattern observed with VG-FT price for deep ITM and OTM options.

[12.4] Summary of Formulae of Option Price with Fourier Transform in VG Model

Table 12.3: Summary of Formulae for VG Model

Method	Formula
Numerical VG	$C_{VG}(S_t; \tau) = S_t \Psi \left(d \sqrt{\frac{1-c_1}{\kappa}}, (\alpha + s) \sqrt{\frac{\kappa}{1-c_1}} \frac{\tau}{\kappa} \right)$ $-K \exp(-rt) \Psi \left(d \sqrt{\frac{1-c_2}{\kappa}}, \alpha s \sqrt{\frac{\kappa}{1-c_2}} \frac{\tau}{\kappa} \right)$ $\Psi(a, b, \gamma) = \int_0^\infty N \left(\frac{a}{\sqrt{u}} + b\sqrt{u} \right) \frac{u^{\gamma-1} \exp(-u)}{\Gamma(\gamma)} du$
FT	$C_{VG-FT}(T, k) = \frac{e^{-\alpha k}}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega k} \frac{e^{-rT} \phi_T(\omega - (\alpha+1)i)}{\alpha^2 + \alpha - \omega^2 + i(2\alpha+1)\omega} d\omega$ $\phi_T(\omega) = \exp \left[i\omega \left\{ s_0 + \left(r + \frac{1}{\kappa} \ln(1 - \theta\kappa - \frac{\sigma^2 \kappa}{2}) \right) T \right\} \right] \left(1 + \frac{\sigma^2 \omega^2 \kappa}{2} - i\theta\kappa\omega \right)^{-\frac{T}{\kappa}}$
DFT	$C_T(k_n) \approx \frac{\exp(-\alpha k_n) \exp(i\pi n) \exp(-i\pi N/2)}{\Delta k}$ $\times \frac{1}{N} \sum_{j=0}^{N-1} w_j \left\{ \exp(i\pi j) \psi_T(\omega_j) \right\} \exp(-i2\pi jn/N)$ $\psi_T(\omega_j) = \frac{e^{-rT} \phi_T(\omega_j - (\alpha+1)i)}{\alpha^2 + \alpha - \omega_j^2 + i(2\alpha+1)\omega_j}$

[13] Conclusion

We implemented the general FT call price (8.17) and our version of the general DFT call price (8.38) for three different types of exponential Lévy models. DFT call price is implemented using $N = 4096$ samples and log strike space sampling interval $\Delta k = 0.005$. This corresponds to angular frequency domain sampling interval of $\Delta \omega = 0.306796$ radians, the total sampling range in the log strike space is $K = N\Delta k = 20.48$, its sampling rate is 200 samples per unit of k , and the total sampling range in the angular frequency domain is $\Omega = N\Delta \omega = 1256.64$.

In the classical Black-Scholes model, our version of DFT $\alpha = 1$ call price yields the price virtually identical to the original BS price for OTM and ITM calls even for extreme near-maturity case (i.e. $T = 1/252$) although the size of error is larger than BS-FT formula. But the error of BS-DFT price becomes relatively large around (i.e. ± 3) ATM. In our example used, the maximum error is 0.0001345 which occurs at exactly ATM at $S_0 = K = 50$. Increasing the decay rate parameter α or sampling more points (i.e. larger N) cannot reduce the size of this error. But we can accept this size of error when considering the dramatic improvement in the CPU time to compute hundreds of prices.

The result of Merton JD model is the same as that of the BS. Our Merton-DFT $\alpha = 1$ call price yields the price virtually identical to the original Merton price for OTM and ITM calls even for extreme near-maturity case (i.e. $T = 1/252$) although the size of error is larger than Merton-FT formula. But the error of Merton-DFT price becomes relatively large around (i.e. ± 3) ATM. In our example used, the maximum error is 0.0001343 which occurs at exactly ATM at $S_0 = K = 50$.

In the VG model as a principle, VG-DFT $\alpha = 1.5$ price and VG-FT $\alpha = 1.5$ price are same. As we expected, the price difference tends to increases as the maturity nears regardless of the moneyness of the call. We consider the maximum price difference of the size approximately 0.025 negligible. Which is better? The clear-cut answer is VG-DFT price. One reason is that VG-DFT needs significantly smaller CPU time to compute hundreds of option prices which is very important for the calibration. Another reason is that VG-DFT price does not have the oscillation pattern observed with VG-FT price for deep ITM and OTM options.

We hope readers appreciate the excellence of FT and DFT option pricing in the sense of its simplicity and generality and this sequel would be a good starting point.

Appendix

[A.1] Set Theory: Notation and Basics¹

- \mathbb{R} : Collection of all real numbers. Intervals in \mathbb{R} are denoted via each endpoint. A square bracket [indicates its inclusion and an open bracket (indicates its exclusion. For example, $[c, d] = \{x \in \mathbb{R} : c \leq x < d\}$. Unbounded intervals are described using ∞ and $-\infty$. Examples are $(-\infty, a) = \{x \in \mathbb{R} : x < a\}$ and $[0, \infty) = \{x \in \mathbb{R} : x \geq 0\} = \mathbb{R}^+$.
- \mathbb{R}^+ : Collection of nonnegative elements of \mathbb{R} .
- \mathbb{R}^d : d -dimensional Euclidian space. Its elements $x = (x_k)_{k=1,\dots,d}$ and $y = (y_k)_{k=1,\dots,d}$ are column vectors with d real components.
- \mathbb{N} : Collection of all positive integers.
- \mathbb{Z} : Collection of all integers.
- \mathbb{Q} : Collection of all rational numbers.
- Ω : Universal set (a set of all scenarios).
- \emptyset : Empty set. It has no members.
- $x \in A$: Membership. The element x is a member of the set A .
- $A \subset B$: Set inclusion. Every member of A is a member of B .
- $\{x \in A : P(x)\}$: The set of elements of A with property P .
- $\mathcal{P}(A)$: The set of all subsets (power set) of A .
- $A \cap B = \{x : x \in A \text{ and } x \in B\}$. Intersection.
- $A \cup B = \{x : x \in A \text{ or } x \in B\}$. Union.
- A^C : The complement of A . The elements of Ω which are not members of A .
$$A^C = \Omega \setminus A$$
- $B \setminus A$: The difference between the set A and B . $B \setminus A = \{x \in B : x \notin A\} = B \cap A^C$
- \exists : There exists.
- \forall : For all.
- $\bigcap_{\alpha \in \Lambda} A_\alpha = \{x : x \in A_\alpha \text{ for all } \alpha \in \Lambda\} = \{x : \forall \alpha \in \Lambda, x \in A_\alpha\}$
- $\bigcup_{\alpha \in \Lambda} A_\alpha = \{x : x \in A_\alpha \text{ for some } \alpha \in \Lambda\} = \{x : \exists \alpha \in \Lambda, x \in A_\alpha\}$
- de Morgan's law: $(\bigcup_{\alpha} A_\alpha)^C = \bigcap_{\alpha} A_\alpha^C$, $(\bigcap_{\alpha} A_\alpha)^C = \bigcup_{\alpha} A_\alpha^C$
- Disjoint sets: A and B are disjoint if $A \cap B = \emptyset$.

¹ Based on Capinski, Kopp, and Kopp (2004).

- Pairwise disjoint sets: A family of sets $(A_\alpha)_{\alpha \in \Lambda}$ is pairwise disjoint if $A_\alpha \cap A_\beta = \emptyset$ whenever $\alpha \neq \beta$ ($\alpha, \beta \in \Lambda$) .
- Cartesian product: Cartesian product $C \times D$ of sets C and D is the set of ordered pairs $C \times D = \{(c, d) : c \in C, d \in D\}$.
- A function $f : C \rightarrow D$: A subset of $C \times D$ where each first coordinate determines the second. Its scope is described by its domain $\mathcal{D}_f = \{c \in C : \exists d \in D, (c, d) \in f\}$ and range $\mathcal{R}_f = \{d \in D : \exists c \in C, (c, d) \in f\}$. A function f associates elements of D with those of C , such that each $c \in C$ has at most one image $d \in D$. This is written as $d = f(c)$.
- Indicator function : Indicator function 1_C of the set C is the function

$$1_C(x) = \begin{cases} 1 & \text{for } x \in C \\ 0 & \text{for } x \notin C. \end{cases}$$

- Upper bound: u is an upper bound for a set $A \subset \mathbb{R}$ if $a \leq u$ for all $a \in A$.
- Lower bound: l is a lower bound for a set $A \subset \mathbb{R}$ if $l \leq a$ for all $a \in A$.
- $\sup A$: Supremum (least upper bound) of A . Minimum of all upper bounds.
- $\inf A$: Infimum (greatest lower bound) of A . Maximum of all lower bounds.

[A.2] Measure²

[A.2.1] Null Sets

Consider an interval to define the length of a set. Let H be a bounded interval $H = [a, b]$, $H = [a, b)$, $H = (a, b]$, or $H = (a, b)$. In each case the length of H is $l(H) = b - a$. A one element set is null since $l(\{b\}) = l([b, b]) = 0$ (The length of a single point is 0). Next, consider the length of a finite set $B = \{1, 5, 20\}$. The length of a finite set is 0 since $l(B) = l(1) + l(5) + l(20) = 0$. Thus the length of a set can be calculated by adding the lengths of its pieces.

Definition: Null set

A null set $A \subseteq \mathbb{R}$ is a set that may be covered by a sequence of intervals of arbitrarily small total length, that is, a sequence $\{H_n : n \geq 1\}$ of intervals given any $\varepsilon > 0$ such that

$$\begin{aligned} A &\subseteq \bigcup_{k=1}^{\infty} H_k \\ \sum_{n=1}^{\infty} l(H_n) &< \varepsilon. \end{aligned}$$

Any one-element set $\{x\}$ is null. For example, when $H_1 = (x - \frac{\varepsilon}{8}, x + \frac{\varepsilon}{8})$ and $H_n = [0, 0]$ for $n \geq 2$,

² Based on Capinski, Kopp, and Kopp (2004).

$$\sum_{n=1}^{\infty} l(H_n) = l(H_1) = \frac{\varepsilon}{4} < \varepsilon.$$

More generally, any countable set $A = \{x_1, x_2, \dots\}$ is null.

Theorem: Union of a Sequence of Null Sets

Let $(N_k)_{k \geq 1}$ be a sequence of null sets, then their union

$$N = \bigcup_{k=1}^{\infty} N_k$$

is also null.

Any countable set is null since it is quite sparse when compared with an interval and it does not contribute to its length.

[A.2.2] Outer Measure

Definition: (Lebesgue) Outer Measure

The (Lebesgue) outer measure of any set $A \subseteq \mathbb{R}$ is the non-negative real number

$$m^*(A) = \inf Z_A$$

where

$$Z_A = \left\{ \sum_{n=1}^{\infty} l(H_n) : H_n \text{ are intervals, } A \subseteq \sum_{n=1}^{\infty} H_n \right\}.$$

The outer measure is the infimum (greatest lower bound) of length of all possible covers of A .

Theorem

$A \subseteq \mathbb{R}$ is a null set if and only if $m^*(A) = 0$.

Note that $m^*(\emptyset) = 0$, $m^*(\{x\}) = 0$ for any $x \in \mathbb{R}$, $m^*(\mathbb{Q}) = 0$, and $m^*(Y) = 0$ for any countable Y .

Proposition

If $A \subset B$, $m^*(A) \leq m^*(B)$.

This means that m^* is monotone: the larger the set, the greater its outer measure.

Theorem

The outer measure of an interval is equal to its length.

Theorem

Outer measure is countably subadditive, i.e. for any sequence of sets $\{B_n\}$

$$m^*\left(\bigcup_{n=1}^{\infty} B_n\right) \leq \sum_{n=1}^{\infty} m^*(B_n).$$

When an interval is shifted along the real line, there is no change in its length:

$$l([c, d]) = l([c+u, d+u]) = d - c.$$

Proposition

Outer measure is translation invariant,

$$m^*(A) = m^*(A+u)$$

where $A \subset \mathbb{R}$, $u \subset \mathbb{R}$, and $A+u = \{a+u : a \in A\}$.

[A.2.3] Lebesgue Measurable Sets and Lebesgue Measure

Consider the class of good sets for which outer measure is countably additive:

$$m^*\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} m^*(B_n)$$

for pairwise disjoint (B_n) .

Definition: Lebesgue Measurability

A set $B \subseteq \mathbb{R}$ is (Lebesgue) measurable if for every set $A \subseteq \mathbb{R}$ we have

$$m^*(A) = m^*(A \cap B) + m^*(A \cap B^C)$$

and we write $B \in \mathcal{M}$.

Theorem

Any null set is measurable. Any interval is measurable.

Theorem: Fundamental Properties of the Class \mathcal{M} of All Lebesgue-Measurable Subsets

- (i) $\mathbb{R} \in \mathcal{M}$.
 - (ii) If $B \in \mathcal{M}$, then $B^C \in \mathcal{M}$.
 - (iii) If $B_n \in \mathcal{M}$ for all $n = 1, 2, \dots$, then $\bigcup_{n=1}^{\infty} B_n \in \mathcal{M}$.
- If $B_n \in \mathcal{M}$ for all $n = 1, 2, \dots$ and $B_j \cap B_k = \emptyset$ for $j \neq k$, then

$$m^*\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} m^*(B_n).$$

A family of sets is a σ -field if it contains the base set and is closed under complements and countable unions. Thus, theorem above means that \mathcal{M} is a σ -field. A $[0, \infty]$ -valued

function defined on a σ -field is called a measure if it satisfies the theorem above for pairwise disjoint sets.

More general approach to measure theory is to begin with the above theorem as axioms. Start with a measure space $(\Omega, \mathcal{F}, \mu)$ where Ω is an abstractly given set (a set of given scenarios), \mathcal{F} is a σ -field of subsets of Ω , and $\mu: \mathcal{F} \mapsto [0, \infty]$ is a function satisfying the above theorem.

Proposition

If $B_k \in \mathcal{M}$, $k = 1, 2, \dots$, then

$$B = \bigcap_{k=1}^{\infty} B_k \in \mathcal{M}.$$

\mathcal{M} is closed under countable unions, countable intersections, and complements. \mathcal{M} contains intervals and all null sets.

Summary: Lebesgue Measure

Lebesgue measure $m: \mathcal{M} \mapsto [0, \infty]$ is a countably additive set function defined on the σ -field \mathcal{M} of measurable sets. Lebesgue measure of an interval equals its length. Lebesgue measure of a null set is zero.

[A.2.4] σ -field³

Let Ω be an arbitrary set and \mathcal{F} be a family of subsets of Ω , and A be a set of \mathbb{R}^d . \mathcal{F} is said to be a σ -field on Ω , if it satisfies:

- (1) $\Omega \in \mathcal{F}$, $\emptyset \in \mathcal{F}$ (It contains empty set.)
- (2) If $A_n \in \mathcal{F}$ for $n = 1, 2, \dots$ (disjoint A_n), then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ and $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$.
Stability under unions and intersections.
- (3) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ (It contains the complements of every element.).

(Ω, \mathcal{F}) is called a measurable space.

Let B be a subset of A , i.e. $B \subset A$, and the measure of a subset B be a positive finite (possibly infinite) number, $\mu(B) \in [0, \infty]$. B is said to be a measurable set. Properties of measurable sets include:

1. An empty set \emptyset has measure 0, i.e. $\mu(\emptyset) = 0$.
2. Additive property: If B and C are disjoint measurable sets, the measure of the union $B \bigcup C$ is $\mu(B \bigcup C) = \mu(B) + \mu(C)$.

³ Based on Cont and Tankov (2004).

3. σ -additivity: Let $(B_n)_{n \in \mathbb{N}}$ be a infinite sequence of disjoint measurable subsets.

Then:

$$\mu\left(\bigcup_{n \geq 1} B_n\right) = \sum_{n \geq 1} \mu(B_n).$$

- 4. It is possible that the measure of a set or subset is infinite, i.e. $\mu(A) = \infty$ and $\mu(B) = \infty$.
- 5. If a set A is finite (i.e. $\mu(A) < \infty$) for any measurable set B , the measure of the complement B^c ($B \bigcup B^c = A$) is $\mu(B^c) = \mu(A) - \mu(B)$.

[A.2.5] Borel σ -Field

Definition: Borel σ -field

$\mathcal{B} = \bigcap \{ \mathcal{F} : \mathcal{F}$ is a σ -field containing all intervals}.

Borel σ -field is the collection of all Borel sets on \mathbb{R}^d which is denoted by $\mathcal{B}(\mathbb{R}^d)$. Borel σ -field is the σ -field generated by the open sets in \mathbb{R}^d . In other words, it is the smallest σ -field containing all open sets in \mathbb{R}^d . A real-valued function $f(x)$ on \mathbb{R}^d is said to be measurable if it is $\mathcal{B}(\mathbb{R}^d)$ -measurable.

[A.2.6] Probability

Lebesgue measure restricted to $[0,1]$ is a probability measure. We select a number from $[0,1]$ at random, restrict Lebesgue measure m to the interval $B = [0,1]$, and consider the σ -field of $\mathcal{M}_{[0,1]}$ of measurable subsets of $[0,1]$. $m_{[0,1]}$ is a probability measure because it is a measure on $\mathcal{M}_{[0,1]}$ with total mass 1.

Definition: Probability Space and Probability Measure $(\Omega, \mathcal{F}, \mathbb{P})$

Let Ω be an arbitrary set (a set of scenarios), \mathcal{F} be a σ -field of subsets of Ω , and \mathbb{P} be a mapping from \mathcal{F} into \mathbb{R} . Let A be a set of \mathbb{R}^d . A triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a probability space and \mathbb{P} is said to be a probability measure if it satisfies the conditions:

- (1) $\mathbb{P}(\Omega) = 1$. (A probability space is a measure space with total mass 1.)
- (2) $0 \leq \mathbb{P}(A) \leq 1$ and $\mathbb{P}(\emptyset) = 0$.
- (3) If $A_n \in \mathcal{F}$ for $n = 1, 2, \dots$ and they are disjoint (i.e. $A_n \cap A_m = \emptyset$ for $n \neq m$), then $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

Let B be an event which is a measurable set $B \in \mathcal{F}$. A probability measure \mathbb{P} assigns a probability between 0 and 1 to each event:

$$\begin{aligned}\mathbb{P}: \mathcal{F} &\rightarrow [0,1] \\ B &\mapsto \mathbb{P}(B).\end{aligned}$$

If two probability measures \mathbb{P} and \mathbb{Q} on (Ω, \mathcal{F}) define the same impossible events, they are said to be equivalent:

$$\mathbb{P} \sim \mathbb{Q} \Leftrightarrow [\forall B \in \mathcal{F}, \mathbb{P}(B) = 0 \Leftrightarrow \mathbb{Q}(B) = 0].$$

Definition: Random Variable

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A mapping X from Ω into \mathbb{R}^d , i.e.

$X : \Omega \rightarrow \mathbb{R}^d$, is said to be an \mathbb{R}^d -valued random variable if it is \mathcal{F} -measurable:

$$\{\omega : X(\omega) \in B\} \in \mathcal{F} \text{ for each } B \in \mathcal{B}(\mathbb{R}^d).$$

We can interpret that $X(\omega)$ is the outcome of the random variable when the scenario ω happens. Probability measures on $\mathcal{B}(\mathbb{R}^d)$ are said to be distributions (laws) on \mathbb{R}^d .

Definition: Property A Almost Surely (a.s.)

A random variable X is said to have a property A almost surely (or with probability 1), if there is $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that $X(\omega)$ has the property A for every $\omega \in \Omega_0$.

Definition: Conditional Probability

Let $\mathbb{P}(B) > 0$. The conditional probability of the event A given B is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Given some disjoint hypotheses, the probability of an event can be calculated by means of conditional probabilities:

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|O_i)\mathbb{P}(O_i)$$

where O_i are pairwise disjoint events such that $\bigcup_{i=1}^{\infty} O_i \in \Omega$ and $\mathbb{P}(O_i) \neq 0$.

Definition: Independence

Let X_j be an \mathbb{R}^{d_j} -valued random variable for $j = 1, \dots, n$. The family $\{X_1, \dots, X_n\}$ is said to be independent if, for every $B_j \in \mathcal{B}(\mathbb{R}^{d_j})$:

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \dots \mathbb{P}(X_n \in B_n).$$

We often say that X_1, \dots, X_n are independent rather than saying that the family $\{X_1, \dots, X_n\}$ is independent.

[A.3] Stochastic Process

A stochastic process $\{X_t\}$ is a family $\{X_t\}_{t \in [0, \infty)}$ of random variables on \mathbb{R}^d defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For any increasing sequence of time $0 \leq t_1 < t_2 < \dots < t_n$, $\mathbb{P}(X(t_1) \in B_1, \dots, X(t_n) \in B_n)$ determines a probability measure on $\mathcal{B}((\mathbb{R}^d)^n)$. Sample function (path) of $\{X_t\}$ is $X(t, \omega)$.

Let $\{X_t\}$ and $\{Y_t\}$ be two stochastic processes. $\{X_t\}$ and $\{Y_t\}$ are said to be identical in law:

$$\{X_t\} \stackrel{d}{=} \{Y_t\},$$

if the systems of their finite-dimensional distributions are identical.

A stochastic process $\{X_t\}$ on \mathbb{R}^d is said to be continuous in probability (stochastically continuous) if, for every $t \geq 0$ and $\varepsilon > 0$:

$$\lim_{s \rightarrow t} \mathbb{P}(|X_s - X_t| > \varepsilon) = 0.$$

[A.3.1] Filtration (Information Flow)

An increasing family of σ -fields $(\mathcal{F}_t)_{t \in [0, T]} : \forall t \geq s \geq 0, \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ is called a filtration or information flow on $(\Omega, \mathcal{F}, \mathbb{P})$. We can interpret \mathcal{F}_t as the information known at time t . \mathcal{F}_t increases as time progresses.

[A.3.2] Non-anticipating (Adapted) Process

A stochastic process $(X_t)_{t \in [0, T]}$ is said to be non-anticipating with respect to the filtration $(\mathcal{F}_t)_{t \in [0, T]}$ or \mathcal{F}_t -adapted if the value of X_t is revealed at time t for each $t \in [0, T]$.

[A.4] Martingales

[A.4.1] General Concept

Consider a trend of a time series of a stochastic process. A stochastic process is said to be a martingale if its time series have no trend. A process with increasing trend is called a submartingale and a process with decreasing trend is called a supermartingale.

Definition: Martingale

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with an filtration (information flow) \mathcal{F}_t . A cadlag process $(X_t)_{t \in [0, T]}$ is said to be a martingale with respect to its filtration \mathcal{F}_t and the probability measure \mathbb{P} if X is nonanticipating (adapted to \mathcal{F}_t), $E[|X_t|]$ is finite for any $t \in [0, T]$ and:

$$\forall s > t, \quad E[X_s | \mathcal{F}_t] = X_t.$$

The best prediction of a martingale's future value is its present value. The definition of martingale makes sense only when the underlying probability measure P and the information flow $(\mathcal{F}_t)_{t \in [0, T]}$ have been specified.

The fundamental property of a martingale process is that its future variations are completely unpredictable with the information flow \mathcal{F}_t :

$$\forall u > 0, \quad E[X_{t+u} - X_t | \mathcal{F}_t] = E[X_{t+u} | \mathcal{F}_t] - E[X_t | \mathcal{F}_t] = X_t - X_t = 0.$$

Example of Continuous-Time Martingale 1)⁴

Let $(X_t)_{t \in [0, T]}$ be a continuous stochastic process whose increments are normally distributed. Let Δ be a small time interval. Then:

$$\begin{aligned} \Delta X_t &\sim N(\mu\Delta, \sigma^2\Delta) \\ \forall s \neq t, \quad E[(\Delta X_s - \mu\Delta)(\Delta X_t - \mu\Delta)] &= 0. \end{aligned}$$

Consider an expectation with the filtration \mathcal{F}_t and with respect to a probability distribution described above:

$$\begin{aligned} \forall u > 0, \quad X_{t+u} &= X_t + \int_t^{t+u} dX_v \\ E[X_{t+u} | \mathcal{F}_t] &= E[X_t + \int_t^{t+u} dX_v | \mathcal{F}_t] = E[X_t | \mathcal{F}_t] + E[\int_t^{t+u} dX_v | \mathcal{F}_t] = X_t + \mu u \neq X_t. \end{aligned}$$

Obviously, $(X_t)_{t \in [0, T]}$ is not a martingale. But the process $(Y_t)_{t \in [0, T]}$:

$$Y_t = X_t - \mu u,$$

is a martingale since:

$$E[Y_{t+u} | \mathcal{F}_t] = E[X_{t+u} - \mu u | \mathcal{F}_t] = E[X_t + (X_{t+u} - X_t) - \mu u | \mathcal{F}_t]$$

⁴ Based on Neftci (2000).

$$= E[X_t | \mathcal{F}_t] + E[X_{t+u} - X_t | \mathcal{F}_t] + E[-\mu u | \mathcal{F}_t] = X_t + \mu u - \mu u = X_t.$$

Example of Continuous-Time Martingale 2) Standard Brownian Motion Process

Consider a standard Brownian motion process $(B_t)_{t \in [0, T]}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Standard Brownian motion is a continuous adapted process with properties that $B_0 = 0$, increments $B_t - B_s = 0$ for $0 \leq s < t$ are independent of \mathcal{F}_s and are normally distributed with mean 0 and variance $t - s$. Obviously, $(B_t)_{t \in [0, T]}$ is a martingale:

$$\forall u > 0, \quad B_{t+u} = B_t + \int_t^{t+u} dB_v$$

$$E[B_{t+u} | \mathcal{F}_t] = E[B_t + \int_t^{t+u} dB_v | \mathcal{F}_t] = E[B_t | \mathcal{F}_t] + E[\int_t^{t+u} dB_v | \mathcal{F}_t] = B_t + 0 = B_t.$$

[A.4.2] Martingale Asset Pricing

Most of financial asset prices are not martingales (i.e. are not completely unpredictable). Consider a risky stock price S_t at time t and let r be the risk-free interest rate. In a small time interval Δ risk-averse investors expect S_t to grow at some positive rate. This can be written as under actual probability measure \mathbb{P} :

$$E_t^{\mathbb{P}}[S_{t+\Delta}] > S_t.$$

Obviously, $S_{t+\Delta}$ is not a martingale. To be more precise risk-averse investors expect S_t to grow at a rate greater than r :

$$E_t^{\mathbb{P}}[e^{-r\Delta} S_{t+\Delta}] > S_t.$$

The stock price discounted by the risk-free interest rate $e^{-r\Delta} S_{t+\Delta}$ is not martingale under \mathbb{P} .

But interestingly, non-martingales can be converted to martingales by changing the probability measure. We will try to find an equivalent probability measure \mathbb{Q} (called risk-neutral measure) under which the stock price discounted by the risk-free interest rate becomes martingale:

$$E_t^{\mathbb{Q}}[e^{-r\Delta} S_{t+\Delta}] = S_t.$$

[A.4.3] Continuous Martingales, Right Continuous Martingales, Square-Integrable Martingales

Let $\{X_t; 0 \leq t\}$ be a continuous martingale. Continuous martingales have continuous trajectories, when $\Delta \rightarrow 0$:

$$P(\Delta S_t > \varepsilon) \rightarrow 0 \quad \text{for all } \varepsilon > 0.$$

It tells us when the time step becomes extremely small, the probability that the stock price changes by some amount approaches zero. Example is a standard Brownian motion process.

In contrast to continuous martingales, right continuous martingales basically change by jumps. Compensated Poisson process is an example of right continuous martingales.

A continuous martingale $\{X_t; 0 \leq t\}$ is said to be a continuous square integrable martingale if X has a finite second moment:

$$E[X_t^2] < \infty.$$

[A.5] Poisson Process⁵

[A.5.1] Exponential Distribution

Suppose a positive random variable X follows an exponential distribution with parameter $\lambda > 0$. Its probability density function is:

$$f_X(x) = \lambda e^{-\lambda x} \quad \text{for } x > 0.$$

The distribution function is:

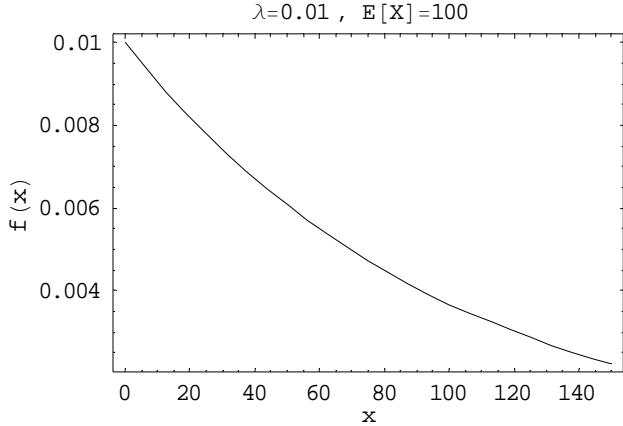
$$F_X(x) = \Pr(X \leq x) = 1 - e^{-\lambda x} \quad \text{for } x > 0.$$

Its mean and variance are:

$$E[X] = \frac{1}{\lambda} \quad \text{and} \quad \text{Var}[X] = \frac{1}{\lambda^2}.$$

For example, the probability density function of an exponential random variable X with $\lambda = 0.01$ is plotted below.

⁵ Based on Cont and Tankov (2004).



The exponential distribution has an important feature called memory-less property. Suppose that the arrival time of a large earthquake, X , follows an exponential distribution with mean 100 years. Consider a situation where $s = 10$ years have passed since the last large earthquake and let $\{X - s | X > s\}$ be the remaining arrival time. Since x and s are both positive,

$$\begin{aligned} \Pr\{X - s > x | X > s\} &= \frac{\Pr\{X - s > x \text{ and } X > s\}}{\Pr\{X > s\}} = \frac{\Pr\{X - s > x\}}{\Pr\{X > s\}} = \frac{\Pr\{X > x + s\}}{\Pr\{X > s\}} \\ &= \frac{1 - F_X(x+s)}{1 - F_X(s)} = \frac{1 - (1 - e^{-\lambda(x+s)})}{1 - F_X(s)} = \frac{e^{-\lambda x} e^{-\lambda s}}{1 - F_X(s)} = \frac{(1 - F_X(x))(1 - F_X(s))}{1 - F_X(s)} = 1 - F_X(x) \\ &= \Pr\{X > x\}. \end{aligned}$$

So the expected remaining arrival time of a next large earthquake is 100 years, i.e. the exponential random variable X does not remember that 10 years have passed since the last large earthquake.

Memory-less property of exponential distribution

If a random variable X follows an exponential distribution,

$$\forall x, s > 0, \quad \Pr\{X > x + s | X > s\} = \Pr\{X > x\}.$$

If X is a random time, the distribution of $X - s$ given $X > s$ is the same as the distribution of X itself.

[A.5.2] Poisson Distribution

If $(\tau_i)_{i \geq 1}$ are independent exponential random variables with parameter λ , the random variable for any $t > 0$:

$$N_t = \inf\{n \geq 1, \sum_{i=1}^n \tau_i > t\},$$

follows a Poisson distribution with parameter λt :

$$\forall n \in \mathbb{N}, \mathbb{P}(N_t = n) = \frac{e^{-\lambda t} (\lambda t)^n}{n!}.$$

If X_1 and X_2 are independent Poisson variables with parameters λ_1 and λ_2 , then $X_1 + X_2$ follows a Poisson distribution with parameter $\lambda_1 + \lambda_2$. This indicates that a Poisson random variable X is infinitely divisible, i.e. X can be divided into an arbitrary number of *i.i.d.* random variables:

$$\text{if } X \sim \text{Poisson}(\lambda), \\ \forall n \in \mathbb{N}, \quad X = \sum_{i=1}^n \frac{X_i}{n}, \quad \frac{X_i}{n} \sim \text{i.i.d. Poisson}\left(\frac{\lambda}{n}\right).$$

[A.5.3] Compensated Poisson Process

A compensated Poisson process $(\widetilde{N}_t)_{t \geq 0}$ is a centered version of a Poisson process $(N_t)_{t \geq 0}$:

$$\widetilde{N}_t = N_t - \lambda t.$$

The mean and variance of a compensated Poisson process are:

$$E[\widetilde{N}_t] = E[N_t - \lambda t] = \lambda t - \lambda t = 0, \\ \text{Var}[\widetilde{N}_t] = E[\widetilde{N}_t^2] = E[(N_t - \lambda t)^2] = \text{Var}[N_t] = \lambda t.$$

Increments of \widetilde{N} are independent and \widetilde{N} is a martingale:

$$E[\widetilde{N}_t | \widetilde{N}_s, s \leq t] = E[\widetilde{N}_t - \widetilde{N}_s + \widetilde{N}_s | \widetilde{N}_s] = E[\widetilde{N}_t - \widetilde{N}_s] + \widetilde{N}_s \\ = E[\widetilde{N}_t] - E[\widetilde{N}_s] + \widetilde{N}_s = 0 - 0 + \widetilde{N}_s = \widetilde{N}_s.$$

$(\lambda t)_{t \geq 0}$ called the compensator of a Poisson process $(N_t)_{t \geq 0}$ is the quantity which needs to be subtracted in order to make the process a martingale.

Unlike a Poisson process, a compensated Poisson process is not integer-valued and not a counting process.

A compensated Poisson process $(\widetilde{N}_t)_{t \geq 0}$ behaves like a Brownian motion after rescaling it by $1/\lambda$ because:

$$E\left[\frac{\widetilde{N}_t}{\lambda}\right] = 0 \quad \text{and} \quad \text{Var}\left[\frac{\widetilde{N}_t}{\lambda}\right] = t.$$

Formally, when the intensity λ of its jumps becomes larger, the compensated Poisson process converges in distribution to a Brownian motion:

$$\left(\frac{\widetilde{N}_t}{\lambda}\right)_{t \in [0, T]} \Rightarrow (B_t)_{t \in [0, T]} \quad \text{when } \lambda \rightarrow \infty.$$

[A.6] Other Distributions Used

[A.6.1] Gamma Function

Let c be a complex number. The integral called gamma function:

$$\Gamma(c) = \int_0^\infty t^{c-1} e^{-t} dt$$

converges absolutely, if the real part of c is positive. By integration by parts:

$$\Gamma(c+1) = c\Gamma(c).$$

Since $\Gamma(1) = 1$, for all natural numbers n :

$$\Gamma(n+1) = n!.$$

Properties of a gamma function which are often used in the process of calculation are:

$$\begin{aligned} \Gamma(1-c) &= -c\Gamma(-c), \\ \frac{\Gamma(-c)}{\Gamma(1-c)} &= -\frac{1}{c}, \\ \frac{\Gamma(1+c)}{\Gamma(c)} &= c, \\ \frac{\Gamma(2+c)}{\Gamma(c)} &= \frac{\Gamma(2+c)}{\Gamma(1+c)} \frac{\Gamma(1+c)}{\Gamma(c)} = (1+c)c. \end{aligned}$$

[A.6.2] Incomplete Gamma Function

The incomplete gamma function is defined by an indefinite integral of the same integrand $t^{c-1} e^{-t}$. For $x \in \mathbb{R}^+$ and complex variable c , if the real part of c is positive:

$$\Gamma(c, x) = \int_x^\infty t^{c-1} e^{-t} dt,$$

$$\gamma(c, x) = \int_0^x t^{c-1} e^{-t} dt,$$

where $\Gamma(c, x)$ is when the lower limit of integration is variable and $\gamma(c, x)$ is when upper limit of the integration is variable. Obviously, we have following relationship:

$$\begin{aligned}\Gamma(c) &= \gamma(c, x) + \Gamma(c, x) \\ \Gamma(c) &= \Gamma(c, 0) \\ \gamma(c, x) &\rightarrow \Gamma(c) \quad \text{as } x \rightarrow \infty.\end{aligned}$$

[A.6.3] Gamma Distribution

Let X be a gamma distributed random variable with the shape parameter $c > 0$ and the scale parameter (controls the tail behavior) $\lambda > 0$. Its probability density function can be expressed as:

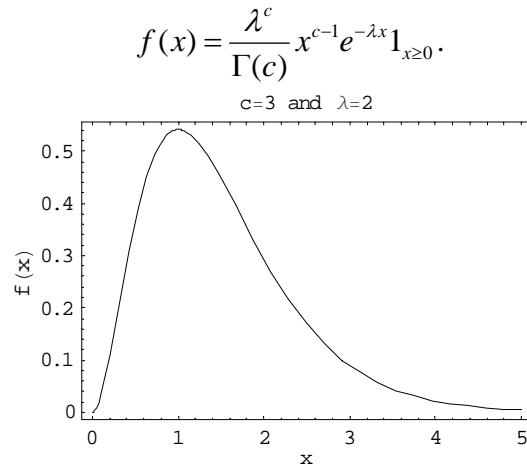


Figure A.6.1: The Gamma Density with $c = 3$ and $\lambda = 2$

Its distribution function in terms of the incomplete gamma function is

$$F(x) = \int_0^\infty f(z) dz = \frac{\gamma(c, \lambda x)}{\Gamma(c)} 1_{x \geq 0}.$$

Its characteristic function $\phi(\omega)$ for any $\omega \in \mathbb{R}$ is calculated as

$$\phi(\omega) = \mathcal{L}_x[f(x)](\omega) = \int_0^\infty e^{i\omega x} f(x) dx = E[e^{i\omega X}] = \frac{1}{(1 - i\omega\lambda^{-1})^c}.$$

The mean, variance, skewness, and excess kurtosis of a gamma random variable are follows:

$$E[X] = \frac{c}{\lambda},$$

$$Var[X] = \frac{c}{\lambda^2},$$

$$Skewness[X] = \frac{2}{\sqrt{c}},$$

$$Excess\ Kurtosis[X] = \frac{6}{c}.$$

A gamma distributed random variable X with the shape parameter $c > 0$ and the scale parameter $\lambda > 0$ possesses the following properties:

- o Suppose X_1, X_2, \dots, X_n are independent gamma distributed random variables with parameters $(c_1, \lambda), (c_2, \lambda), \dots, (c_n, \lambda)$, then $\sum_{i=1}^n X_i \sim Gamma(\sum_{i=1}^n c_i, \lambda)$.
- o If $c = 1$, the gamma distribution reduces to an exponential distribution with parameter λ .

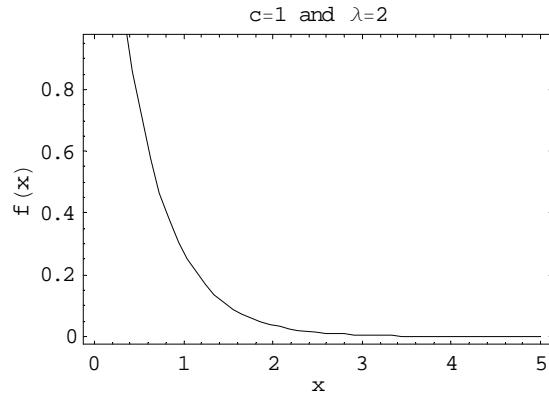


Figure A.6.2: The Gamma Density with $c = 1$ and $\lambda = 2$

[A.7] Modified Bessel Functions: Modified Bessel Function of the First Kind $I_v(z)$ and Second Kind $K_v(z)$

Modified Bessel differential equation is the second-order ordinary differential equation of the form

$$z^2 \frac{d^2 y}{dz^2} + z \frac{dy}{dz} - (z^2 + v^2)y = 0.$$

Its solutions can be expressed for $n = 0, 1, 2, \dots$,

$$\begin{aligned} y &= a_1 J_n(-iz) + a_2 Y_n(-iz) \\ &= b_1 I_n(z) + b_2 K_n(z), \end{aligned}$$

where $J_v(z)$ is called a Bessel function of the first kind of order v , $Y_v(z)$ a Bessel function of the second kind of order v , $I_v(z)$ a modified Bessel function of the first kind of order v , $K_v(z)$ a modified Bessel function of the second kind of order v , and $v \in \mathbb{R}^+$. The modified Bessel function of the first kind $I_v(z)$ admits the series expansion

$$I_v(z) = \left(\frac{1}{2}z\right)^v \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(k+v+1)} = \sum_{k=0}^{\infty} \frac{(z/2)^{v+k}}{k!\Gamma(k+v+1)} \quad \text{for } z \geq 0 \text{ and } v \geq 0,$$

which is the solution of the modified Bessel differential equation which is bounded when $z \rightarrow 0$. For $n = 0, 1, 2, \dots$

$$I_{-n}(z) = I_n(z).$$

$I_v(z)$ has asymptotic behavior

$$\begin{aligned} I_v(z) &= \frac{e^z}{\sqrt{2\pi z}} \left[1 + O\left(\frac{1}{z}\right) \right] \quad \text{when } z \rightarrow +\infty, \\ I_v(z) &\sim \frac{(z/2)^v}{\Gamma(v+1)} \quad \text{when } z \rightarrow 0. \end{aligned}$$

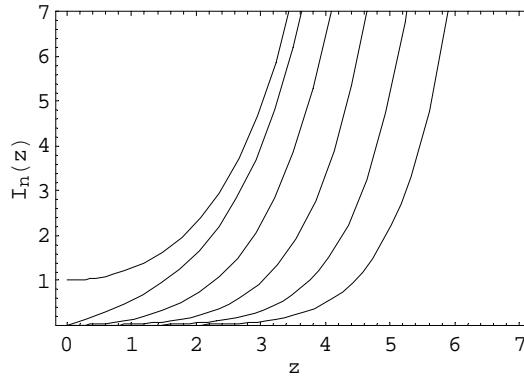


Figure A.7.1: Plot of the modified Bessel functions of the first kind. $I_v(z)$ for $v = 1, 2, 3, 4, 5$ from left to right.

The modified Bessel function of the second kind $K_v(z)$ admits the series expansion

$$K_v(z) = \frac{\pi}{2} \frac{I_{-v}(z) - I_v(z)}{\sin(\pi v)} \quad \text{for } z \geq 0 \text{ and } v \geq 0,$$

which is the solution of the modified Bessel differential equation which is bounded when $z \rightarrow +\infty$. For all orders v :

$$K_{-\nu}(z) = K_\nu(z).$$

$I_\nu(z)$ has asymptotic behavior:

$$\begin{aligned} K_\nu(z) &= e^{-z} \sqrt{\frac{\pi}{2z}} [1 + o(\frac{1}{z})] \quad \text{when } z \rightarrow +\infty, \\ K_0(z) &\sim -\log z \quad \text{for } \nu = 0, \\ K_\nu(z) &\sim \frac{1}{2} \Gamma(\nu) (z/2)^{-\nu} \quad \text{for } \nu > 0. \end{aligned}$$

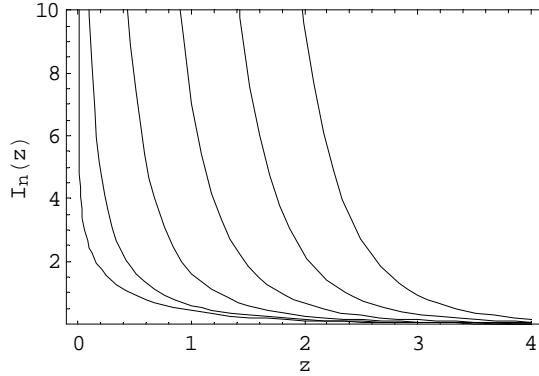


Figure A.7.2: Plot of the modified Bessel functions of the second kind. $K_\nu(z)$ for $\nu = 1, 2, 3, 4, 5$ from left to right.

[A.8] Itô Formula

[A.8.1] Itô Formula for Brownian Motion

Let $B_t \sim \text{Normal}(0, t)$ be a standard Brownian motion and $f(B)$ be an arbitrary function of a Brownian motion. Consider a very small time interval h . A Taylor series expansion of a function $f(B_{t+h})$ about a point B_t is given by:

$$\begin{aligned} f(B_{t+h}) - f(B_t) &= \frac{df(B_t)}{dB} (B_{t+h} - B_t) + \frac{d^2 f(B_t)}{dB^2} \frac{1}{2!} (B_{t+h} - B_t)^2 \\ &\quad + \frac{d^3 f(B_t)}{dB^3} \frac{1}{3!} (B_{t+h} - B_t)^3 + \dots \end{aligned}$$

Thus:

$$\begin{aligned} &\sum_{j=1}^n (f(B_{t+jh}) - f(B_{t+(j-1)h})) \\ &= (f(B_{t+h}) - f(B_t)) + (f(B_{t+2h}) - f(B_{t+h})) + \dots + (f(B_{t+nh}) - f(B_{t+(n-1)h})) \end{aligned}$$

$$\begin{aligned}
&= f(B_{t+nh}) - f(B_t) \\
&= \frac{df(B_t)}{dB}(B_{t+h} - B_t) + \frac{d^2f(B_t)}{dB^2} \frac{1}{2!}(B_{t+h} - B_t)^2 + \frac{d^3f(B_t)}{dB^3} \frac{1}{3!}(B_{t+h} - B_t)^3 + \dots \\
&\quad + \frac{df(B_{t+h})}{dB}(B_{t+2h} - B_{t+h}) + \frac{d^2f(B_{t+h})}{dB^2} \frac{1}{2!}(B_{t+2h} - B_{t+h})^2 + \frac{d^3f(B_{t+h})}{dB^3} \frac{1}{3!}(B_{t+2h} - B_{t+h})^3 + \dots \\
&\quad + \frac{df(B_{t+(n-1)h})}{dB}(B_{t+nh} - B_{t+(n-1)h}) + \frac{d^2f(B_{t+(n-1)h})}{dB^2} \frac{1}{2!}(B_{t+nh} - B_{t+(n-1)h})^2 \\
&\quad \quad + \frac{d^3f(B_{t+(n-1)h})}{dB^3} \frac{1}{3!}(B_{t+nh} - B_{t+(n-1)h})^3 + \dots \\
&= \sum_{j=1}^n \frac{df(B_{t+(j-1)h})}{dB}(B_{t+jh} - B_{t+(j-1)h}) + \frac{1}{2!} \sum_{j=1}^n \frac{d^2f(B_{t+(j-1)h})}{dB^2}(B_{t+jh} - B_{t+(j-1)h})^2 \\
&\quad \quad + \frac{1}{3!} \sum_{j=1}^n \frac{d^3f(B_{t+(j-1)h})}{dB^3}(B_{t+jh} - B_{t+(j-1)h})^3 + \dots
\end{aligned}$$

Use the approximation:

$$\frac{d^2f(B_{t+(j-1)h})}{dB^2} = \frac{d^2f(B_t)}{dB^2} \quad \text{and} \quad \frac{d^3f(B_{t+(j-1)h})}{dB^3} = \frac{d^3f(B_t)}{dB^3}.$$

Thus,

$$\begin{aligned}
f(B_{t+nh}) - f(B_t) &= \sum_{j=1}^n \frac{df(B_{t+(j-1)h})}{dB}(B_{t+jh} - B_{t+(j-1)h}) \\
&\quad + \frac{1}{2!} \frac{d^2f(B_t)}{dB^2} \sum_{j=1}^n (B_{t+jh} - B_{t+(j-1)h})^2 + \frac{1}{3!} \frac{d^3f(B_t)}{dB^3} \sum_{j=1}^n (B_{t+jh} - B_{t+(j-1)h})^3 + \dots
\end{aligned}$$

Note the following:

$$\begin{aligned}
\sum_{j=1}^n \frac{df(B_{t+(j-1)h})}{dB}(B_{t+jh} - B_{t+(j-1)h}) &\equiv \int_t^{t+\Delta t} \frac{df}{dB} dB, \\
\frac{1}{2!} \frac{df^2(B_t)}{dB^2} \sum_{j=1}^n (B_{t+jh} - B_{t+(j-1)h})^2 &= \frac{1}{2} \frac{df^2}{dB^2} \int_t^{t+\Delta t} (dB)^2 = \frac{1}{2} \frac{df^2}{dB^2} \Delta t,^6 \\
\frac{1}{3!} \frac{df^3(B_t)}{dB^3} \sum_{j=1}^n (B_{t+jh} - B_{t+(j-1)h})^3 + \dots &\approx 0.
\end{aligned}$$

Therefore, we have the integral version of Itô formula for the Brownian motion:

⁶ In the mean square limit $\int_0^t (dB)^2 = t$.

$$f(B_{t+\Delta t}) - f(B_t) = \int_t^{t+\Delta t} \frac{df(B_\tau)}{dB} dB_\tau + \frac{1}{2} \int_t^{t+\Delta t} \frac{df^2(B_\tau)}{dB^2} d\tau,$$

or in terms of time interval between 0 and t :

$$f(B_t) = f(B_0) + \int_0^t \frac{df(B_s)}{dB} dB_s + \frac{1}{2} \int_0^t \frac{df^2(B_s)}{dB^2} ds.$$

Its differential version is:

$$df = \frac{df}{dB} dB + \frac{1}{2} \frac{d^2 f}{dB^2} dt.$$

Consider a simple example of $f(B) = B^2$ where:

$$\frac{df}{dB} = 2B \quad \text{and} \quad \frac{d^2 f}{dB^2} = 2.$$

Thus the function $f(B)$ satisfies the stochastic differential equation (SDE):

$$df = 2B dB + dt.^7$$

[A.8.2] Wimott's (1998) Rule of Thumb of Itô Formula for Brownian Motion

A Taylor series expansion of a function $f(B + dB)$ about a point B is given by⁸:

$$\begin{aligned} f(B + dB) - f(B) &= \frac{df(B)}{dB} dB + \frac{1}{2} \frac{d^2 f(B)}{dB^2} dB^2, \\ df &= \frac{df(B)}{dB} dB + \frac{1}{2} \frac{d^2 f(B)}{dB^2} dB^2. \end{aligned}$$

Since in the mean square limit $\int_0^t (dX)^2 = t$, setting $dB^2 = dt$ yields the Itô formula:

$$df = \frac{df(B)}{dB} dB + \frac{1}{2} \frac{d^2 f(B)}{dB^2} dt.$$

This is of course technically wrong, but it is very useful.

[A.8.3] Itô Formula for Brownian Motion with Drift

⁷ In ordinary calculus (if B were deterministic variable), $df = 2B dB$.

⁸ Ignore higher order terms.

Consider a function of a random variable $V(S)$ and S is a Brownian motion with drift:

$$dS = a(S)dt + b(S)dB,$$

where the drift parameter a and the volatility parameter b depends on S . Using Wilmott's rule of thumb, the function $V(S)$ satisfies the SDE:

$$dV = \frac{dV}{dS}dS + \frac{1}{2}b^2 \frac{d^2V}{dS^2}dt.$$

In terms of a Brownian motion B , the function $V(S)$ satisfies the SDE:

$$\begin{aligned} dV &= \frac{dV}{dS}(a(S)dt + b(S)dB) + \frac{1}{2}b(S)^2 \frac{d^2V}{dS^2}dt \\ dV &= \left(a(S) \frac{dV}{dS} + \frac{1}{2}b(S)^2 \frac{d^2V}{dS^2} \right) dt + b(S) \frac{dV}{dS} dB. \end{aligned}$$

[A.8.4] Itô Formula for Brownian Motion with Drift in Higher Dimensions

Consider a function of a random variable $V(S, t)$ and S is a Brownian motion with drift

$$dS = a(S, t)dt + b(S, t)dB,$$

where the drift parameter a and the volatility parameter b depends on S and time t . The function $V(S, t)$ satisfies the SDE:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}b^2 \frac{\partial^2 V}{\partial S^2}dt.$$

An example is that if S is a Brownian motion with drift of the form⁹:

$$dS = \mu Sdt + \sigma SdB,$$

the function $V(S, t)$ satisfies the SDE:

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt.$$

[A.8.5] Itô Formula for Jump-Diffusion (Finite Activity Lévy) Processes

⁹ Geometric Brownian motion.

Let X be a jump-diffusion process of the form:

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s + \sum_{i=1}^{N_t} \Delta X_i,$$

where $\int_0^t b_s ds$ is a sum of a drift term, $\int_0^t \sigma_s dB_s$ is a stochastic integral of a (multiplicative) Brownian motion process with $E[\int_0^T \sigma_t^2 dt] < \infty$, and $\sum_{i=1}^{N_t} \Delta X_i$ is a compound Poisson Process. Drift and volatility processes b and σ are continuous and nonanticipating.

Let $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ be any $C^{1,2}$ function. Then, the integral version of Itô formula for jump-diffusion processes is:

$$\begin{aligned} f(X_t, t) - f(X_0, 0) &= \int_0^t \left[\frac{\partial f(X_s, s)}{\partial s} + b_s \frac{\partial f(X_s, s)}{\partial x} \right] ds \\ &\quad + \frac{1}{2} \int_0^t \sigma_s^2 \frac{\partial^2 f(X_s, s)}{\partial x^2} ds + \int_0^t \frac{\partial f(X_s, s)}{\partial x} \sigma_s dB_s \\ &\quad + \sum_{\{i \geq 1, T_i \leq t\}} [f(X_{T_i-} + \Delta X_i) - f(X_{T_i-})]. \end{aligned}$$

Its differential version becomes:

$$\begin{aligned} df(X_t, t) &= \frac{\partial f(X_t, t)}{\partial t} dt + b_t \frac{\partial f(X_t, t)}{\partial x} dt + \frac{\sigma_t^2}{2} \frac{\partial^2 f(X_t, t)}{\partial x^2} dt \\ &\quad + \sigma_t \frac{\partial f(X_t, t)}{\partial x} dB_t + [f(X_{t-} + \Delta X_t) - f(X_{t-})]. \end{aligned}$$

For more details, see Cont and Tankov (2004).

[A.8.6] Itô Formula for General (Finite and Infinite Activity) Scalar Lévy Processes

Let $\{X_t; t \geq 0\}$ be a general scalar Lévy process with its characteristic triplet (σ^2, ν, γ) and $f : \mathbb{R} \rightarrow \mathbb{R}$ be any C^2 function. Then, the integral version of Itô formula is:

$$\begin{aligned} f(X_t) &= f(0) + \int_0^t \frac{\sigma^2}{2} \frac{\partial^2 f(X_s)}{\partial x^2} ds + \int_0^t \frac{\partial f(X_{s-})}{\partial x} dX_s \\ &\quad + \sum_{\substack{0 \leq s \leq t \\ \Delta X_s \neq 0}} [f(X_{s-} + \Delta X_s) - f(X_{s-}) - \Delta X_s \frac{\partial f(X_{s-})}{\partial x}]. \end{aligned}$$

Its differential version is:

$$df(X_t) = \frac{\sigma^2}{2} \frac{\partial^2 f(X_t)}{\partial x^2} dt + \frac{\partial f(X_{t-})}{\partial x} dX_t + f(X_t) - f(X_{t-}) - \Delta X_t \frac{\partial f(X_{t-})}{\partial x}.$$

The integral version of Itô formula in higher dimension is:

$$\begin{aligned} f(X_t, t) - f(X_0, 0) &= \int_0^t \frac{\partial f(X_{s-}, s)}{\partial x} dX_s + \int_0^t \left[\frac{\partial f(X_s, s)}{\partial s} + \frac{\sigma^2}{2} \frac{\partial^2 f(X_s, s)}{\partial x^2} \right] ds \\ &\quad + \sum_{\substack{0 \leq s \leq t \\ \Delta X_s \neq 0}} [f(X_{s-} + \Delta X_s, s) - f(X_{s-}, s) - \Delta X_s \frac{\partial f(X_{s-}, s)}{\partial x}]. \end{aligned}$$

For more details, see Cont and Tankov (2004).

Bibliography

- Ball, C. A. and Torous, W. N., 1985, "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing," *The Journal of Finance* 40, 155-173.
- Black, F. and Scholes, M., 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 3.
- Breeden, D. T., and Litzenberger, R. H., 1978, "Prices of State-Contingent Claims Implicit in Option Prices," *Journal of Business*, Vol 51 Issue 4.
- Brigham, E. O., 1988, *The Fast Fourier Transform and Its Applications*, Prentice Hall.
- Capinski, M., Kopp, E., and Kopp, P. E., 2004, *Measure, Integral and Probability*, Springer-Verlag.
- Carr, P., Madan, D. B., 1998, "Option Valuation Using the Fast Fourier Transform," *Journal of Computational Finance* 2, 61-73.
- Carr, P., Géman, H., Madan, D. B., Wu, L., and Yor, M., "Option Pricing Using Integral Transforms," Unpublished work for seminar.
- Cont, R., Bouchaud, J.-P., and Potters, M., 1997, "Scaling in Stock Market Data: Stable Laws and Beyond," Appeared in Dubrulle, B., Graner, F., and Sornette, D., *Scale Invariance and Beyond*, Berlin, Springer.
- Cont, R. and Tankov, P., 2004, "Calibration of Jump-Diffusion Option Pricing Models: A Robust Non-Parametric Approach," *Rapport Interne* 490, CMAP, Ecole Polytechnique.
- Cont, R. and Tankov, P., 2004, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC Financial Mathematics Series.
- Harrison, J. M. and Pliska, S.R., 1981, "Martingales and Stochastic Integrals In the Theory of Continuous Trading," *Stochastic Processes and Applications* 11, 215-260.
- Harrison, J. M. and Pliska, S.R., 1983, "A Stochastic Calculus Model of Continuous Trading: Complete Markets," *Stochastic Processes and Applications* 15, 313-316.
- Madan, D. B., Carr, P., and Chang, E., 1998, "The Variance Gamma Process and Option Pricing," *European Finance Review* 2, 79-105.

Madan, D. B. and Milne, F., 1991, "Option Pricing with VG Martingale Components," Mathematical Finance 1 (4), 39-55.

Madan, D. B. and Seneta, E., 1990, "The Variance Gamma (V.G.) Model for Stock Market Returns," Journal of Business, Vol 63, No 4.

mathworld.wolfram.com

Matsuda, K., 2004, "Dynamics of Risk-Neutral Densities Implied By Option Prices," Working Paper, Graduate School and University Center of the City University of New York.

Merton, R. C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," Journal of Economic Theory, 3, 373-413.

Merton, R. C., 1976, "Option Pricing When Underlying Stock Returns Are Continuous," Journal of Financial Economics, 3, 125-144.

Navas, J. F., 2000, "A Note on Jump-Diffusion Process for Asset Retuens,"

Neftci, S. N., 2000, An Introduction to the Mathematics of Financial Derivatives, Academic Press.

Raible, S., 1998, "Lévy Processes in Finance: Theory, Numerics and Empirical Facts," PhD Thesis, Universität Freiburg.

Sato, K., 1999, Lévy process and Infinitely Divisible Distributions, Cambridge University Press.

Wilmott, P., 1998, Derivatives, John Wiley & Sons