2

A General Multi-Factor Model of the Term Structure of Interest Rates and the Principles of Characteristic Functions

2.1 An Extended Jump-Diffusion Term-Structure Model

The evolution of the yield curve can be described in various ways. For instance, it is possible to use such quantities as zero-bond prices, instantaneous forward rates and short interest rates, respectively, to build the term structure of interest rates. If the transformation law from one quantity to the other is known, the choice of the independent variable is just a matter of convenience.

In this thesis, we attempt to model the dynamics of the instantaneous interest rate, denoted hereafter by $r(\mathbf{x}_t)$, in order to construct our derivatives pricing framework. This instantaneous interest rate $r(\mathbf{x}_t)$ is also often referred to as the short-term interest rate or short rate, respectively, and characterizes the risk-free rate for borrowing or lending money over the infinitesimal time period $[t, t + \mathrm{d}t]$. Since we model the dynamics in a continuous trading environment, the relevant processes are described via stochastic differential equations.

The economy we consider has the trading interval $[0, T]$. The uncertainty under the physical probability measure is completely specified by the filtered probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. In this formulation Ω denotes the complete set of all possible outcome elements $\omega \in \Omega$. The information available in the economy is contained within the filtration $(\mathfrak{F})_{t\geq 0}$, such that the level of uncertainty is resolved over the trading interval with respect to the information filtration. The last term, completing the probability space, is called the *real-world* probability measure $\mathbb{P}$ on $(\Omega, \mathfrak{F})$, since it reflects the *real-world* probability law of the data.

We model the dynamic behavior of the term structure in the spirit of Duffie and Kan (1996) and Duffie, Pan and Singleton (2000), to preserve an exponential-affine structure of the characteristic function. However, we extend the framework in Duffie and Kan (1996) to allow for N different trigger processes[6], which offers more flexibility. The term structure is then modeled by a multi-factor structural Markov model of M factors, represented by a random vector $\mathbf{x}_t$, which solves the multivariate stochastic differential equation,

$$\mathrm{d}\mathbf{x}_t = \begin{pmatrix} \mathrm{d}x_t^{(1)} \\ \mathrm{d}x_t^{(2)} \\ \vdots \\ \mathrm{d}x_t^{(M-1)} \\ \mathrm{d}x_t^{(M)} \end{pmatrix} = \boldsymbol{\mu}^{\mathbb{P}}(\mathbf{x}_t)\,\mathrm{d}t + \boldsymbol{\Sigma}(\mathbf{x}_t)\,\mathrm{d}\mathbf{W}_t^{\mathbb{P}} + \mathbf{J}\,\mathrm{d}\mathbf{N}(\boldsymbol{\lambda}^{\mathbb{P}}t). \tag{2.1}$$

The coefficient vector $\boldsymbol{\mu}^{\mathbb{P}}(\mathbf{x}_t)$ has the affine structure

$$\boldsymbol{\mu}^{\mathbb{P}}(\mathbf{x}_t) = \boldsymbol{\mu}_0^{\mathbb{P}} + \boldsymbol{\mu}_1^{\mathbb{P}}\mathbf{x}_t \tag{2.2}$$

with $(\boldsymbol{\mu}_0^{\mathbb{P}}, \boldsymbol{\mu}_1^{\mathbb{P}}) \in \mathbb{R}^M \times \mathbb{R}^{M\times M}$ and the variance-covariance matrix $\boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Sigma}(\mathbf{x}_t)'$ suffices the relation

$$\boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Sigma}(\mathbf{x}_t)' = \boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_1\mathbf{x}_t, \tag{2.3}$$

where $\boldsymbol{\Sigma}_0 \in \mathbb{R}^{M\times M}$ is a matrix and $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{M\times M\times M}$ is a third order tensor. The vector $\mathbf{W}_t^{\mathbb{P}}$ in equation (2.1) represents M orthogonal Wiener processes. Thus, we have[7]

$$\mathbb{E}^{\mathbb{P}}(\,\mathrm{d}\mathbf{W}_t^{\mathbb{P}}\,\mathrm{d}\mathbf{W}_t^{\mathbb{P}'}) = \mathbf{I}_M\,\mathrm{d}t$$

with $\mathbf{I}_M$ as the $M \times M$ identity matrix.

As mentioned above, we extend the ordinary diffusion model[8] with N independent Poisson processes, condensed in the vector $\mathbf{N}(\boldsymbol{\lambda}^{\mathbb{P}}t)$. This vector process acts with constant and positive intensities[9] $\boldsymbol{\lambda}^{\mathbb{P}}$. We allow for every

[6] Chacko and Das (2002) model also the term structure with help of different Poisson processes. However, their approach consider a subordinated short rate.

[7] If not indicated otherwise, we subsequently use the shorthand notation $\mathbb{E}[\,\cdot\,]$ for the expression $\mathbb{E}[\,\cdot\,|\mathfrak{F}_t]$.

[8] This would be the original model approach presented in Duffie and Kan (1996).

[9] This exponential-affine model can be easily extended to stochastic jump intensities of the form $\boldsymbol{\lambda}^{\mathbb{P}}(\mathbf{x}_t) = \boldsymbol{\lambda}_0^{\mathbb{P}} + \boldsymbol{\lambda}_1^{\mathbb{P}}\mathbf{x}_t$. See Chapter 10.

particular factor in $\mathbf{x}_t$ an amount of N different jumps drawn from a jump amplitude matrix $\mathbf{J} \in \mathbb{R}^{M \times N}$. Hence, the distribution functions of the particular jump amplitudes are given within the matrix $\boldsymbol{\nu}(\mathbf{J})$. Finally, all jump amplitudes in $\mathbf{J}$ are independent of the state of the vector $\mathbf{x}_t$[10].

To preserve the exponential-affine structure of any derivatives contract based on $r(\mathbf{x}_t)$ and $\mathbf{x}_t$, respectively, all random sources, the Brownian motions $\mathbf{W}_t^{\mathbb{P}}$, intensities $\boldsymbol{\lambda}^{\mathbb{P}}$ and jump amplitudes $\mathbf{J}$ are mutually independent. As a direct consequence of the independence of $\mathbf{J}$ and $\mathbf{x}_t$, there is no chance to generate an arbitrage opportunity according to available information before the particular jump occurs. Hence, given a jump time t^*, we have formally $\mathbf{J} \in \mathfrak{F}_{t^*-}$. Therefore, if a jump occurs at time t^*, nobody is able to predict the exact jump amplitude and cannot gain an arbitrarily large profit with certainty.

In this thesis, the choice of jump amplitudes in $\mathbf{J}$ can draw on three different types of distribution. These are:

- Exponentially distributed jumps.
- Normally distributed jumps.
- Gamma distributed jumps.

These jump distributions and the resulting jump transforms, which are used in our pricing mechanism, are covered in Chapter 7.

Basically, we prefer to model the term structure in terms of the instantaneous short interest rate $r(\mathbf{x}_t)$[11], because in this framework all fundamental quantities are properly defined as the expectation of some functionals on the underlying process $r(\mathbf{x}_t)$. Accordingly, we are able to construct an arbitrage-free economy and simultaneously guarantee a consistent pricing methodol-

[10] From a technical point of view, it is either possible to introduce a dependence on $\mathbf{x}_t$ for the jump intensity together with independent random jump amplitudes or a dependence on $\mathbf{x}_t$ for the jump amplitude together with constant jump intensities. See Zhou (2001), p. 4.

[11] Other approaches are possible, e.g. the direct approach as used in Schöbel (1987) and Briys, Crouhy and Schöbel (1991) or modeling the forward-rate process as done in Heath, Jarrow and Morton (1992).

ogy[12]. The drawback of this approach is that we might not be able to explain perfectly the entire term structure extracted from observed bond market prices and therefore must content ourselves with a best fit scenario.

The literature distinguishes between two approaches in modeling the short interest rate in a multidimensional framework. Firstly, we can identify a strategy, which we call henceforth the *subordinated* modeling approach. Here, the short rate is modeled as

$$r\left(\mathbf{x}_t\right) = w_0 + w_1 x_t^{(1)}(x_t^{(2)}, \ldots, x_t^{(M)}).$$

Consequently, the other $M-1$ stochastic factors are subordinated loadings, containing e.g. a stochastic volatility and/or a stochastic mean[13]. Apart from the stochastic variable $x_t^{(1)}$, we also consider the deterministic parameters w_0 and w_1 in modeling the short rate. Indeed, there are other factors, which can possibly have some other economic meaning worth to be included in the interest-rate model.

The second method in modeling short rates, which we call the *additive* modeling approach, is to represent r_t as a weighted sum over $\mathbf{x}_t$, formally given by

$$r\left(\mathbf{x}_t\right) = w_0 + \mathbf{w}'\mathbf{x}_t,$$

[12] This means that all derivative prices are based on the same price of risk. See Culot (2003), Section 2.1.

[13] In Brennan and Schwartz (1979), Brennan and Schwartz (1980), and Brennan and Schwartz (1982) the short-rate process is subordinated by a stochastic long-term rate. Beaglehole and Tenney (1991) discuss a two-factor interest-rate model with a stochastic long-term mean component and Fong and Vasicek (1991a) introduce a short-rate model with stochastic volatility. A model where the short rate depends on a stochastic inflation factor is modeled in Pennacchi (1991). Kellerhals (2001) analyzes an interest-rate model with a stochastic market price of risk component. In Balduzzi, Das, Foresi and Sundaram (1996), the authors present a short-rate model with a stochastic mean and volatility component.

where $\mathbf{w}$ is a $M \times 1$ vector containing separate weights for the corresponding factor loadings in $\mathbf{x}_t$[14]. However, this model approach possibly entails difficulties in explaining the economic meaning of the variables $\mathbf{x}_t$[15].

2.2 Technical Preliminaries

Before we proceed any further, we have to discuss some general results and principles of stochastic analysis, which are commonly used in financial engineering, namely the prominent Itô's Lemma and the equally famous Feynman-Kac Theorem. These two principles play a major role in diffusion theory and are well connected. Since we consider discontinuous jumps in our model setup, we have to use extended versions of these two results. At first we have to state some regularity conditions on the jump-diffusion process, in order to guarantee their application.

Definition 2.2.1 (Regularity Conditions for Jump-Diffusion Processes). *If the vector process* $\mathbf{x}_t$ *represents a multivariate jump-diffusion, the parameter coefficients* $\boldsymbol{\mu}(\mathbf{x}_t), \boldsymbol{\Sigma}(\mathbf{x}_t)$ *have to satisfy the following technical conditions*[16] *for all* $t \geq 0$

- $\|\boldsymbol{\mu}(\mathbf{x}_t^a) - \boldsymbol{\mu}(\mathbf{x}_t^b)\| \leq A_1 \|\mathbf{x}_t^a - \mathbf{x}_t^b\|$
- $\|\boldsymbol{\Sigma}(\mathbf{x}_t^a)) - \boldsymbol{\Sigma}(\mathbf{x}_t^b)\| \leq A_2 \|\mathbf{x}_t^a - \mathbf{x}_t^b\|$
- $\|\boldsymbol{\mu}(\mathbf{x}_t^a)\| \leq A_1 \left(1 + \|\mathbf{x}_t^a\|\right)$
- $\|\boldsymbol{\Sigma}(\mathbf{x}_t^a))\| \leq A_2 \left(1 + \|\mathbf{x}_t^a\|\right)$

where $\mathbf{x}_t^a, \mathbf{x}_t^b \in \mathbb{R}^M$ *are two vectors containing different realizations of* $\mathbf{x}_t$ *and the constants* $A_1, A_2 < \infty$ *denote some scalar barriers. Additionally, we need*

[14] Langetieg (1980) models the short rate as an additive process consisting of two correlated Ornstein-Uhlenbeck processes. In Beaglehole and Tenney (1991) an additive, multivariate quadratic Gaussian interest-rate model is given. Longstaff and Schwartz (1992) and Chen and Scott (1992) model the interest-rate process as the sum of two uncorrelated Square-Root processes.

[15] A comprehensive discussion on this topic is given in Piazzesi (2003).

[16] The first two conditions are known as the Lipschitz conditions, the latter two represent the growth or polynomial growth conditions. See, for example, Karlin and Taylor (1981).

for the jump components the integral $\int_{\mathbb{R}} e^{cJ_{mn}} \, \mathrm{d}\nu(J_{mn})$ *to be well defined for every* $J_{mn} \in \mathbf{J}$ *and some constant* $c \in \mathbb{C}$.

If the conditions posed above are met, we are able to apply both Itô's Lemma and the Feynman-Kac Theorem.

We start with Itô's Lemma. This lemma enables us to determine the stochastic process driving some function $f(\mathbf{x}_t, t, T)$, depending on time t and a stochastic (vector) variable, e.g. the process $\mathbf{x}_t$ given in equation (2.1). The variables t and $\mathbf{x}_t$, respectively, are hereafter denoted as the independent variables. The coefficients $\boldsymbol{\mu}(\mathbf{x}_t)$ and $\boldsymbol{\lambda}$ used in this section have no superscripts, because the principles introduced here hold in general.

Theorem 2.2.2 (Itô Formula for Jump-Diffusion Processes[17]). *Assume the function* $f(\mathbf{x}_t, t, T)$ *is at least twice differentiable in* $\mathbf{x}_t$ *and once differentiable in* t. *Then the canonical decomposition of the stochastic differential equation for* $f(\mathbf{x}_t, t, T)$ *is given by*

$$\begin{aligned} \mathrm{d}f(\mathbf{x}_t, t, T) = & \bigg(\frac{\partial f(\mathbf{x}_t, t, T)}{\partial t} + \boldsymbol{\mu}(\mathbf{x}_t)' \frac{\partial f(\mathbf{x}_t, t, T)}{\partial \mathbf{x}_t} \\ & + \frac{1}{2} \operatorname{tr} \left[\boldsymbol{\Sigma}(\mathbf{x}_t) \boldsymbol{\Sigma}(\mathbf{x}_t)' \frac{\partial^2 f(\mathbf{x}_t, t, T)}{\partial \mathbf{x}_t \partial \mathbf{x}_t'} \right] \bigg) \mathrm{d}t \\ & + \frac{\partial f(\mathbf{x}_t, t, T)}{\partial \mathbf{x}_t'} \boldsymbol{\Sigma}(\mathbf{x}_t) \, \mathrm{d}\mathbf{W}_t \\ & + \left(\mathbf{f}(\mathbf{x}_t, \mathbf{J}, t, T)' - f(\mathbf{x}_t, t, T) \right) \mathrm{d}\mathbf{N}(\boldsymbol{\lambda} t), \end{aligned} \tag{2.4}$$

where the function $\mathbf{f}(\mathbf{x}_t, \mathbf{J}, t, T)$ *contains all jump components with elements* $(\mathbf{f}(\mathbf{x}_t, \mathbf{J}, t, T))_n = f(\mathbf{x}_t + \mathbf{j}_n, t, T)$ *and* $\mathbf{j}_n \in \mathbb{R}^M$ *contains as* m*th element* J_{mn} *of the amplitude matrix* $\mathbf{J}$.

Another key result which we use extensively is the Feynman-Kac theorem. This theorem provides us with a tool to determine the system of partial differential equations (PDEs), given an expectation.

[17] See, Kushner (1967), p. 15, for the jump-extended version of Itô's lemma.

Theorem 2.2.3 (Feynman-Kac). *If the restrictions in definition 2.2.1 hold, we have the expectation*

$$f(\mathbf{x}_t, t, T) = \mathbb{E}\left[e^{-\int_t^T h(\mathbf{x}_s, s)\,\mathrm{d}s} f(\mathbf{x}_T, T, T) \right], \tag{2.5}$$

solving the partial differential equation

$$\begin{aligned} \frac{\partial f(\mathbf{x}_t, t, T)}{\partial t} + \boldsymbol{\mu}(\mathbf{x}_t)' \frac{\partial f(\mathbf{x}_t, t)}{\partial \mathbf{x}_t} + \frac{1}{2} \operatorname{tr}\left[\boldsymbol{\Sigma}(\mathbf{x}_t) \boldsymbol{\Sigma}(\mathbf{x}_t)' \frac{\partial^2 f(\mathbf{x}_t, t, T)}{\partial \mathbf{x}_t \partial \mathbf{x}_t'} \right] \\ + \mathbb{E}_{\mathbf{J}} \left[\mathbf{f}(\mathbf{x}_t, \mathbf{J}, t, T)' - f(\mathbf{x}_t, t, T) \right] \boldsymbol{\lambda} = h(\mathbf{x}_t, t) f(\mathbf{x}_t, t, T), \end{aligned} \tag{2.6}$$

with boundary condition[18]

$$f(\mathbf{x}_T, T, T) = G(\mathbf{x}_T) \tag{2.7}$$

and $\mathbf{f}(\mathbf{x}_t, \mathbf{J}, t, T)$ *as defined in theorem 2.2.2.*

In diffusion theory, the function $h(\mathbf{x}_t, t)$ is commonly addressed to as the killing rate of the expectation[19] and can be interpreted as some short rate. Since we use equivalently as killing rate a short rate characterized by the time constant coefficients w_0 and $\mathbf{w}$ we set the relation

$$h(\mathbf{x}_t, t) = r(\mathbf{x}_t).$$

As we will see, these two principles are the fundamental tools in obtaining the solutions for our upcoming valuation problems, especially in calculating the general characteristic function of a stochastic process, which is discussed in the next sections.

2.3 The Risk-Neutral Pricing Approach

So far, the stochastic behavior of the state vector $\mathbf{x}_t$ was assumed to be modeled under the *real-world* probability measure $\mathbb{P}$. This probability measure depends on the investor's assessment of the market and therefore cannot be

[18] The operator $\mathbb{E}_{\mathbf{J}}[\,\cdot\,]$ denotes the expectation with respect to the jump sizes $\mathbf{J}$.

[19] See, for example, Øksendal (2003), p. 145.

used in calculating unique derivatives prices[20]. However, for valuation purposes we need to derive contract prices under the condition of an arbitrage-free market[21], which will be shown in this section.

According to the seminal papers of Harrison and Kreps (1979) and Harrison and Pliska (1981), it is a well known and rigorously proved fact, if one can find at least one equivalent martingale measure with respect to $\mathbb{P}$, then the observed market is arbitrage-free and therefore a derivatives pricing framework can be established. Thus, we establish the link between this equivalent martingale measure $\mathbb{Q}$, also known as the risk-neutral probability measure[22], and the probability measure $\mathbb{P}$ in this section.

Since we are dealing with M stochastic factors, primarily integrated in the short rate $r(\mathbf{x}_t)$, which are all non-tradable goods, we are confronted with an incomplete market. In contrast to other model frameworks in which factors represent prices of tradable goods, we encounter a somewhat more difficult situation to end up in a consistent arbitrage-free pricing approach[23]. Foremost, we need to introduce for every source of uncertainty a market price of risk reflecting the risk aversion of the market. The common procedure in this case is to choose a particular equivalent martingale measure, sometimes also called the pricing measure which determines the appropriate numeraire to be applied[24]. Having chosen the numeraire, which has the function of a denominator of the expected contingent claim and determines the martingale condition for the expectation, we afterwards have to extract yields for different maturities of zero-bond prices. In the next step the model prices of zero bonds

[20] See, for example, Musiela and Rutkowski (2005), p. 10.

[21] The arbitrage-free approach is also known as the partial equilibrium approach. Including preferences of investors, i.e. working with utility functions would be a general equilibrium approach. Schöbel (1995) gives a detailed overview of both approaches.

[22] The terminology can be justified, since in a risk-neutral world, where all market participants act under a risk-neutral utility behavior, the probability measures $\mathbb{P}$ and $\mathbb{Q}$ coincide. See, for example, Duffie (2001), p. 108.

[23] This statement holds only for tradable goods modeled by pure diffusion processes. Otherwise, due to the jump uncertainty one has again to implement some variable compensating jump risk. See Merton (1976).

[24] This can be for example the money market account or zero-coupon bond prices. See Dai and Singleton (2003), pp. 635-637.

are calibrated with respect to this empirical yield curve. In the calibration process for these parameters, two separate approaches can be utilized[25]. In the first approach one computes the particular model parameters under the $\mathbb{P}$ measure together with the different market prices of risk. The other method would be to calibrate the model onto the parameters under the objective measure $\mathbb{Q}$. A problem which is common to all model frameworks, where the instantaneous interest rate $r(\mathbf{x}_t)$ is used to describe the term structure of interest rates is that in general the given yield curve is not matched perfectly. Hence, we rather want an arbitrage-free model, which might not be able to explain perfectly all observed yields, but to state a model with an internally consistent stochastic environment.

In the upcoming subsections, we will first give an outline how the risk-neutral measure is defined and how the particular coefficients under this probability measure $\mathbb{Q}$ can be derived for our affine term-structure model. Due to the jump-diffusion framework, we also focus on the topic that our martingale measure should consider for discontinuous price shocks.

2.3.1 Arbitrage and the Equivalent Martingale Measure

Before we start with the formulation of our option-pricing methodology, we need to ensure the existence of an arbitrage-free pricing system. A very useful insight for this delicate matter is given in the above mentioned work of Harrison and Kreps (1979) and Harrison and Pliska (1981). Using measure theory, they judge the market to be arbitrage free enabling the consistent calculation of derivative prices if at least one equivalent martingale measure can be found, corresponding to the physical measure $\mathbb{P}$. Hence, using the money market account as numeraire in order to derive $\mathbb{Q}$, the price of a derivative contract would be just the discounted expectation of its terminal payoff $G(\mathbf{x}_T)$[26]. So our first step is to define the relevant conditions for an equivalent martingale measure.

---

[25] See Duffie, Pan and Singleton (2000), p. 1354.

[26] See, for example, Geman, Karoui and Rochet (1995) and Dai and Singleton (2003), p. 635.

Definition 2.3.1 (Equivalent Probability Measure). *Two probability measures* $\mathbb{P}$ *and* $\mathbb{Q}$ *are equivalent, if for any event* $\mathcal{A}$, $\mathbb{P}(\mathcal{A}) > 0$ *if and only if* $\mathbb{Q}(\mathcal{A}) > 0$.

According to definition 2.3.1, the equivalent probability measure $\mathbb{Q}$ must only agree on the same null sets given by $\mathbb{P}$. The next property we need, in order to obtain the probability measure $\mathbb{Q}$, is the martingale property.

Definition 2.3.2 (Martingale Property). *A stochastic process* $f(\mathbf{x}_t, t)$ *is a martingale under the probability measure* $\mathbb{Q}$ *if and only if the equality*

$$f(\mathbf{x}_t, t, T) = \mathbb{E}^{\mathbb{Q}}\left[f(\mathbf{x}_T, T, T)\right] \tag{2.8}$$

holds for any $t \leq T$.

This last definition ensures the fair game ability of our interest-rate market. Combining definitions 2.3.1 and 2.3.2 lead us to the equivalent martingale measure $\mathbb{Q}$ with respect to $\mathbb{P}$. Thus, to be a fair game, respectively a martingale, the probability measure $\mathbb{Q}$ transforms the probability law for $\mathbf{x}_t$, leaving the null sets of $\mathbb{P}$ untouched. In the next subsection we show the transition of the probability law from the real-world measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$.

2.3.2 Derivation of the Risk-Neutral Coefficients

Having found the formal conditions of an equivalent martingale measure, we now want to derive the transformation rule from measure $\mathbb{P}$ to $\mathbb{Q}$. This rule, also called the Radon-Nikodym derivative $\xi(\mathbf{x}_t, t, T)$, is represented by

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathfrak{F}_t} = \frac{\xi(\mathbf{x}_T, T, T)}{\xi(\mathbf{x}_t, t, T)}. \tag{2.9}$$

In order to derive the risk-neutral coefficients, we adopt the corresponding pricing-kernel methodology. Doing this, the pricing kernel or Radon-Nikodym derivative $\xi(\mathbf{x}_t, t, T)$, belongs itself to the class of exponential-affine functions of $\mathbf{x}_t$[27]. The principle of risk-neutrality implies for the state-price kernel an

[27] See, for example, Dai and Singleton (2003), p. 642.

expected discount rate equal to the instantaneous risk-free rate $r(\mathbf{x}_t)$. Thus, we need the equation

$$\mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\xi(\mathbf{x}_t,t,T)}{\xi(\mathbf{x}_t,t,T)}\right] = -r(\mathbf{x}_t)\,\mathrm{d}t, \tag{2.10}$$

to hold. Using this type of state-price kernel, we have the discounted expectation of an interest-rate derivatives price to fulfill the definition of a martingale as described in theorem 2.3.2. Consequently, ensuring the expectation made above holds and considering the systematic risk factors, we choose the specific form of $\xi(\mathbf{x}_t,t,T)$ to satisfy

$$\frac{\mathrm{d}\xi(\mathbf{x}_t,t,T)}{\xi(\mathbf{x}_t,t,T)} = -r(\mathbf{x}_t)\,\mathrm{d}t - \boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)'\,\mathrm{d}\mathbf{W}^{\mathbb{P}} - \boldsymbol{\Lambda}_{\boldsymbol{\lambda}}'\left(\mathrm{d}\mathbf{N}(\boldsymbol{\lambda}^{\mathbb{P}}t) - \boldsymbol{\lambda}^{\mathbb{P}}\,\mathrm{d}t\right). \tag{2.11}$$

The vectors $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)$ and $\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}$ compensate the sources of risk under the risk-neutral measure $\mathbb{Q}$ for the vector of Brownian motions and the vector of Poisson processes, respectively. The vector $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)$ is characterized by the two relations[28]

$$\begin{aligned}\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)'\,\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t) &= l_0 + \mathbf{l}_1'\mathbf{x}_t\\ \boldsymbol{\Sigma}(\mathbf{x}_t)\,\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t) &= \mathbf{s}_0 + \mathbf{s}_1\mathbf{x}_t\end{aligned}$$

with $l_0 \in \mathbb{R}$, $\mathbf{l}_1, \mathbf{s}_0 \in \mathbb{R}^M$, and $\mathbf{s}_1 \in \mathbb{R}^{M\times M}$. Defining $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)$ like this, we ensure the exponential-affine structure in the pricing kernel $\xi(\mathbf{x}_t,t,T)$. In contrast to the constant, N-dimensional vector $\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}$, we need to establish in $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)$ a dependence on the state vector $\mathbf{x}_t$ because of a possibly non-zero matrix $\boldsymbol{\Sigma}_1$[29]. Thus, if a particular factor $x_t^{(m)}$ has a constant volatility coefficient, meaning its volatility does not depend on any element in $\mathbf{x}_t$, there is either no dependence on $\mathbf{x}_t$ for the respective element in the the vector $\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)$ and vice versa. Since $\boldsymbol{\lambda}^{\mathbb{P}}$ is the vector of expected arrival rates, we have with

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{d}\mathbf{N}(\boldsymbol{\lambda}^{\mathbb{P}}t) - \boldsymbol{\lambda}^{\mathbb{P}}\,\mathrm{d}t\right] = \mathbf{0}_N,$$

a $\mathbb{P}$-martingale, representing a vector of compensated Poisson processes[30].

[28] Compare, for example, with Duffie, Pan and Singleton (2000), Culot (2003), and Dai and Singleton (2003).

[29] Dealing with a Square-Root process, we cannot set the particular market price of risk to a constant value, see Cox, Ingersoll and Ross (1985b), Section 5.

[30] A compensated Poisson process can be roughly seen as a discontinuous equivalent of a Brownian motion. See, for example, Karatzas and Shreve (1991), p. 12.

As a consequence of this incomplete market, the vectors $\mathbf{\Lambda}_{\mathbf{\Sigma}}(\mathbf{x}_t)$ and $\mathbf{\Lambda}_{\boldsymbol{\lambda}}$ are not uniquely defined. Therefore, the pricing kernel itself is not uniquely defined either and we have to determine these risk price vectors with a calibration of yields generated by the model to the empirical yield curve as mentioned earlier. We assume this calibration to depend on the yields of traded zero-coupon bonds $P(\mathbf{x}_t, t, T)$ with different times to maturities[31]. Suppressing unnecessary notations for convenience and applying Itô's Lemma, we get the following SDE for the $\mathbb{P}$-dynamics of a zero-coupon bond

$$\mathrm{d}P(\mathbf{x}_t, t, T) = \mu_P\,\mathrm{d}t + \boldsymbol{\sigma}'_P\,\mathrm{d}\mathbf{W}^{\mathbb{P}} + \mathbf{J}_P\,\mathrm{d}\mathbf{N}(\lambda^{\mathbb{P}}t) \tag{2.12}$$

with drift, diffusion and jump components[32]

$$\begin{aligned} \mu_P = &\frac{\partial P(\mathbf{x}_t, t, T)}{\partial t} + \boldsymbol{\mu}^{\mathbb{P}}(\mathbf{x}_t)'\frac{\partial P(\mathbf{x}_t, t, T)}{\partial \mathbf{x}_t} \\ &+ \frac{1}{2}\mathrm{tr}\left[\mathbf{\Sigma}(\mathbf{x}_t)\mathbf{\Sigma}(\mathbf{x}_t)'\frac{\partial^2 P(\mathbf{x}_t, t, T)}{\partial \mathbf{x}_t \partial \mathbf{x}'_t}\right], \end{aligned} \tag{2.13}$$

$$\boldsymbol{\sigma}_P = \mathbf{\Sigma}(\mathbf{x}_t)\frac{\partial P(\mathbf{x}_t, t, T)}{\partial \mathbf{x}_t}, \tag{2.14}$$

$$\mathbf{J}_P = \mathbf{P}(\mathbf{x}_t, \mathbf{J}, t, T)' - P(\mathbf{x}_t, t, T)\,. \tag{2.15}$$

On the other hand, we impose the martingale condition for traded contracts, which is due to the chosen numeraire,

$$\begin{aligned} P(\mathbf{x}_t, t, T) = &\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(\mathbf{x}_s)\,\mathrm{d}s} P(\mathbf{x}_T, T, T)\right] \\ = &\mathbb{E}^{\mathbb{P}}\left[\frac{\xi(\mathbf{x}_T, T, T)}{\xi(\mathbf{x}_t, t, T)} P(\mathbf{x}_T, T, T)\right]. \end{aligned} \tag{2.16}$$

Multiplying this last equation with $\xi(\mathbf{x}_t, t, T)$, which is known at time t and therefore a certain quantity, we consequently have $\xi(\mathbf{x}_t, t, T)P(\mathbf{x}_t, t, T)$ to be a martingale and the infinitesimal increment $\mathrm{d}\left(\xi(\mathbf{x}_t, t, T)P(\mathbf{x}_t, t, T)\right)$ to be a local martingale[33]. According to Theorem 2.2.2 we have

[31] Since coupon bonds are commonly traded, zero-bond values can be synthetically generated by coupon stripping.

[32] $\mathbf{P}(\mathbf{x}_t, \mathbf{J}, t, T)$ has the equivalent definition as $\mathbf{f}(\mathbf{x}_t, \mathbf{J}, t, T)$ with all calculations made with respect to $P(\mathbf{x}_t, t, T)$. See Theorem 2.2.2.

[33] The existence of a *local* martingale under the new measure $\mathbb{Q}$ is sufficient for the no-arbitrage condition. See Delbaen and Schachermayer (1995) and Øksendal (2003) Section 12.1., respectively.

$$
\begin{aligned}
\mathrm{d}(\xi(\mathbf{x}_t,t,T)P(\mathbf{x}_t,t,T))& \\
&= \xi(\mathbf{x}_t,t,T)\,\mathrm{d}P(\mathbf{x}_t,t,T) + P(\mathbf{x}_t,t,T)\,\mathrm{d}\xi(\mathbf{x}_t,t,T)\\
&\quad + \mathrm{d}P(\mathbf{x}_t,t,T)\,\mathrm{d}\xi(\mathbf{x}_t,t,T)\\
&= \xi(\mathbf{x}_t,t,T)\mu_P\,\mathrm{d}t + \xi(\mathbf{x}_t,t,T)\boldsymbol{\sigma}'_P\,\mathrm{d}\mathbf{W}^{\mathbb{P}}\\
&\quad + \xi(\mathbf{x}_t,t,T)\mathbf{J}_P\mathbf{N}(\lambda^{\mathbb{P}})\\
&\quad - P(\mathbf{x}_t,t,T)\,\xi(\mathbf{x}_t,t,T)r(\mathbf{x}_t)\,\mathrm{d}t\\
&\quad - P(\mathbf{x}_t,t,T)\,\xi(\mathbf{x}_t,t,T)\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)'\,\mathrm{d}\mathbf{W}^{\mathbb{P}}\\
&\quad - P(\mathbf{x}_t,t,T)\,\xi(\mathbf{x}_t,t,T)\boldsymbol{\Lambda}'_{\boldsymbol{\lambda}}\left(\mathrm{d}\mathbf{N}(\boldsymbol{\lambda}^{\mathbb{P}}t) - \boldsymbol{\lambda}^{\mathbb{P}}\,\mathrm{d}t\right)\\
&\quad - \xi(\mathbf{x}_t,t,T)\boldsymbol{\sigma}'_P\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}(\mathbf{x}_t)\,\mathrm{d}t - \xi(\mathbf{x}_t,t,T)\mathbf{J}_P\mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}\,\mathrm{d}t.
\end{aligned}
\tag{2.17}
$$

In the last equation, we used for the infinitesimal time increments the relation

$$\mathrm{d}t\,\mathrm{d}t = 0,$$

and for the vector of uncorrelated Brownian motions

$$\mathrm{d}\mathbf{W}^{\mathbb{P}}\,\mathrm{d}\mathbf{W}^{\mathbb{P}'} = \mathbf{I}_M\,\mathrm{d}t.$$

Similarly, the corresponding expression for the vector of independent Poisson processes is

$$\mathrm{d}\mathbf{N}(\boldsymbol{\lambda}^{\mathbb{P}}t)\,\mathrm{d}\mathbf{N}(\boldsymbol{\lambda}^{\mathbb{P}}t)' = \mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}\,\mathrm{d}t,$$

where $\mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}$ represents a matrix consisting of the diagonal elements

$$\mathrm{diag}\left[\mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}\right] = \boldsymbol{\lambda}^{\mathbb{P}},$$

and zeros otherwise. In the next step, we divide for notational ease all coefficients of the zero-bond SDE (2.12) by $P(\mathbf{x}_t,t,T)$. Hence, we use hereafter the normalized coefficients,

$$
\begin{aligned}
\tilde{\mu}_P &= \frac{\mu_P}{P(\mathbf{x}_t,t,T)},\\
\tilde{\boldsymbol{\sigma}}_P &= \frac{\boldsymbol{\sigma}_P}{P(\mathbf{x}_t,t,T)},\\
\tilde{\mathbf{J}}_P &= \frac{\mathbf{J}_P}{P(\mathbf{x}_t,t,T)}.
\end{aligned}
$$

Combining condition (2.16) and equation (2.17), and keeping in mind that under $\mathbb{P}$-dynamics, the Brownian motions and the compensated Poisson processes in equation (2.11) are martingales, we get for the expectation

$$\mathbb{E}^{\mathbb{P}}\left[\frac{\mathrm{d}\left(\xi(\mathbf{x}_t,t,T)P\left(\mathbf{x}_t,t,T\right)\right)}{\xi(\mathbf{x}_t,t,T)P\left(\mathbf{x}_t,t,T\right)}\right] = \tilde{\mu}_P\,\mathrm{d}t + \mathbb{E}_{\mathbf{J}}\left[\tilde{\mathbf{J}}_P\right]\boldsymbol{\lambda}^{\mathbb{P}}\,\mathrm{d}t - r\left(\mathbf{x}_t\right)\,\mathrm{d}t - \tilde{\boldsymbol{\sigma}}'_P\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}\left(\mathbf{x}_t\right)\,\mathrm{d}t - \mathbb{E}_{\mathbf{J}}\left[\tilde{\mathbf{J}}_P\right]\mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}\,\mathrm{d}t \equiv 0. \tag{2.18}$$

If we now solve equation (2.18) for the modified drift coefficient $\tilde{\mu}_P$, subsequently eliminating all $\mathrm{d}t$ terms, we eventually end up with the relation

$$\tilde{\mu}_P = r\left(\mathbf{x}_t\right) + \tilde{\boldsymbol{\sigma}}'_P\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}\left(\mathbf{x}_t\right) + \mathbb{E}_{\mathbf{J}}\left[\tilde{\mathbf{J}}_P\right]\left(\mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}\boldsymbol{\Lambda}_{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^{\mathbb{P}}\right), \tag{2.19}$$

which means that the rate of return of a zero bond must be equal to the risk free short rate plus some terms reflecting the particular risk premiums of the different sources of uncertainty.

We are now ready to identify the corresponding formal expressions under $\mathbb{Q}$-dynamics of the coefficient parameters $\boldsymbol{\mu}^{\mathbb{P}}$ and $\boldsymbol{\lambda}^{\mathbb{P}}$. Comparing equation (2.13) with (2.19) lead us to the fundamental partial differential equation for zero-bond prices[34]

$$\frac{\partial P\left(\mathbf{x}_t,t,T\right)}{\partial t} + \frac{\partial P\left(\mathbf{x}_t,t,T\right)}{\partial \mathbf{x}'_t}\left(\boldsymbol{\mu}^{\mathbb{P}} - \boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}\left(\mathbf{x}_t\right)\right) + \frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Sigma}(\mathbf{x}_t)'\frac{\partial^2 P\left(\mathbf{x}_t,t,T\right)}{\partial \mathbf{x}_t\partial \mathbf{x}'_t}\right] + \mathbb{E}_{\mathbf{J}}\left[\mathbf{J}_P\right]\left(\boldsymbol{\lambda}^{\mathbb{P}} - \mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}\right) = r\left(\mathbf{x}_t\right)P\left(\mathbf{x}_t,t,T\right). \tag{2.20}$$

According to equation (2.20), together with Itô's Lemma, and the Feynman-Kac representation, we are able to express the risk-neutral parameters as

$$\boldsymbol{\mu}^{\mathbb{Q}} = \boldsymbol{\mu}^{\mathbb{P}} - \boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Lambda}_{\boldsymbol{\Sigma}}\left(\mathbf{x}_t\right) = \boldsymbol{\mu}_0^{\mathbb{Q}} + \boldsymbol{\mu}_1^{\mathbb{Q}}\mathbf{x}_t, \tag{2.21}$$

$$\boldsymbol{\lambda}^{\mathbb{Q}} = \boldsymbol{\lambda}^{\mathbb{P}} - \mathbf{I}_N^{\boldsymbol{\lambda}^{\mathbb{P}}}\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}. \tag{2.22}$$

Since the jump intensities $\boldsymbol{\lambda}^{\mathbb{Q}}$ have to be positive, we need $\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}$ small enough to ensure the positiveness of the jump intensities under the risk-neutral measure $\mathbb{Q}$ given the intensity vector $\boldsymbol{\lambda}^{\mathbb{P}}$. The constant coefficients in the variance-covariance matrix (2.3) remain unchanged under the new measure $\mathbb{Q}$. This

[34] Once the risk-neutral coefficients for the interest-rate process are determined, equation (2.20) can be used to price any European contingent claim by exchanging the terminal condition and replacing $P\left(\mathbf{x}_t,t,T\right)$ with the particular function representing the price of the derivative security to be calculated.

phenomenon is often referred to as the diffusion invariance principle, although this terminology is not completely correct. We want to emphasize that the variations of the Brownian motions only coincide under both measures $\mathbb{P}$ and $\mathbb{Q}$, if the variance-covariance matrix exclusively exhibits constant coefficients[35]. Otherwise, we are implicitly dealing with a different time-dependent variance-covariance matrix, since the vector $\mathbf{x}_t$ experiences a drift correction and therefore affects the relation given in equation (2.3). Consequently, the probability transformation law of the process $\mathbf{x}_t$ from $\mathbb{P}$ to $\mathbb{Q}$ does not only contain a drift compensation. Moreover, besides the jump intensity correction, the very shape of the probability density itself can be changed, due to the implicitly altered variations of the diffusion terms.

Hence, calibrating the theoretical term-structure model to zero-bond yields, whether estimating the parameters of the left or the right sides of equations 2.21 and 2.22, results in the following SDE governing the particular factors under risk-neutral dynamics

$$\mathrm{d}\mathbf{x}_t = \boldsymbol{\mu}^{\mathbb{Q}}(\mathbf{x}_t)\,\mathrm{d}t + \boldsymbol{\Sigma}(\mathbf{x}_t)\,\mathrm{d}\mathbf{W}_t^{\mathbb{Q}} + \mathbf{J}\,\mathrm{d}\mathbf{N}\left(\boldsymbol{\lambda}^{\mathbb{Q}}t\right), \tag{2.23}$$

which we use in the subsequent sections as starting point for our calculations.

2.4 The Characteristic Function

In this section, we first give a brief overview of the abilities of characteristic functions and show afterwards how the characteristic function of an exponential-affine process, as given in equation (2.1), can be derived. We generalize the principle of building characteristic functions for some scalar process $g(\mathbf{x}_t)$, which is essential for our derivatives pricing technique. Since characteristic functions play a major part in our derivation of semi closed-form solutions for interest-rate derivatives, we discuss also some of their fundamental properties.

Before we introduce the characteristic function itself, we first need to state a definition of Fourier Transformations of some deterministic variable x[36].

[35] In this case, we would deal with the matrix $\boldsymbol{\Sigma}(\mathbf{x})\boldsymbol{\Sigma}(\mathbf{x})' = \boldsymbol{\Sigma}_0$.

[36] In the literature, there seems to exist various definitions for this type of transformation. Thus, we want to clarify the issue by giving a straightforward definition

This concept belongs to the field of integral transformations[37] and is a widely used tool in engineering disciplines, especially in signal processing.

Definition 2.4.1 (General one-dimensional Fourier Transformation and its Inversion). *We define the Fourier Transformation $\mathcal{F}^x[\,\cdot\,]$ of some function $f(x)$ with respect to the independent variable x as*

$$\mathcal{F}^x[f(x)] = \int_{-\infty}^{\infty} e^{\imath z x} f(x)\,\mathrm{d}x = \hat{f}(z), \tag{2.24}$$

where $z \in \mathbb{C}$ denotes the transform variable in Fourier space, satisfying the restriction $\mathrm{Im}(z) \in (\underline{\chi}, \overline{\chi})$ with $\underline{\chi}$ and $\overline{\chi}$ denoting some lower and upper boundaries guaranteeing the existence of the Fourier Transformation, $\imath = \sqrt{-1}$ as the standard imaginary unit, and $\hat{f}(z)$ as the shorthand notation for the Fourier Transformation of $f(x)$ with respect to its argument x.

Accordingly, the inverse transformation operator $\mathcal{F}^{-1}[\cdot]$ is then defined by

$$\mathcal{F}^{-1}[\hat{f}(z)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\imath z x} \hat{f}(z)\,\mathrm{d}z = f(x). \tag{2.25}$$

Due to the exponential character of the Fourier Transformation, we need to establish in equation (2.25) a normalization factor of 2π. The terminology *general* one-dimensional Fourier Transformation, in contrast to an *ordinary* one-dimensional Fourier Transformation, is used because we do not limit the transformation variable z to be on the real line[38]. Thus, we allow z to be complex-valued, which makes equation (2.24) and (2.25) a line integral, performed parallel to the real line. Note that both the transform and its inverse

in this section. In financial studies our definition according to equation (2.24) of a Fourier Transformation seems to be commonly accepted. See, for example, Carr and Madan (1999), Bakshi and Madan (2000) and Raible (2000). On the other hand in engineering sciences, the opposite definition of a Fourier Transformation and its inverse operation does exist. See, for example, Duffy (2004).

[37] Other popular integral transformations are e.g. the Laplace transformation or the z-transformation. A comprehensive discussion of the Laplace Transformation is given in Doetsch (1967).

[38] Hence, the equivalent expression *complex* Fourier Transformation is sometimes used in the literature.

operation have to take place on the same strip going through $\mathrm{Im}(z)$, in order to reconstruct the original function $f(x)$.

The advantage in performing this *general* Fourier Transformation is the possibility to derive image functions in cases where the *ordinary* transform approach would fail, e.g. for functions which are unbounded[39]. However, in these cases, the *general* approach enables us to derive solutions for their Fourier Transformations. For example, if we want to compute the Fourier Transformation of a function[40]

$$G(x) = \max(e^x - K, 0),$$

the *ordinary* transformation approach appears to be useless, since

$$\mathcal{F}^x[G(x)] \to \infty.$$

Performing a *general* transformation, in this case within the strip $\mathrm{Im}(z) \in (1, \infty)$, we get[41]

$$\mathcal{F}^x[G(x)] = \frac{K^{1+\imath z}}{\imath z(1+\imath z)}, \tag{2.26}$$

where $\mathrm{Im}(z)$ can be fixed at every value within the above mentioned strip to derive the original function by applying the inverse Fourier Transformation. The different contours in Fourier space of the transformed payoff function given in equation (2.26) are depicted in Figure 2.1. Having derived the fundamental technique to compute Fourier Transformations, which is an essential part in this thesis, we go further and have a look at Fourier Transformations of density functions of stochastic variables, which are commonly known as characteristic functions.

Definition 2.4.2 (Scalar Characteristic Functions). *We define the scalar characteristic function* $\psi^{x^{(m)}}(\mathbf{x}_t, z, w_0, \mathbf{w}, t, T)$ *as the expected value of the terminal condition* $G(\mathbf{x}_T) = e^{\imath z x_T^{(m)}}$*, given the state* $\mathbf{x}_t$ *at time* $t \leq T$*. This can be expressed more formally as*

[39] This is the case for most payoff structures of option contracts, e.g. plain vanilla call or put options.

[40] This function represents, for instance, the payoff function of a plain vanilla call option in an asset pricing environment, where x is the natural logarithm of the underlying asset price.

[41] In Section 5.3, Fourier Transformations are derived in detail for different types of payoff functions.

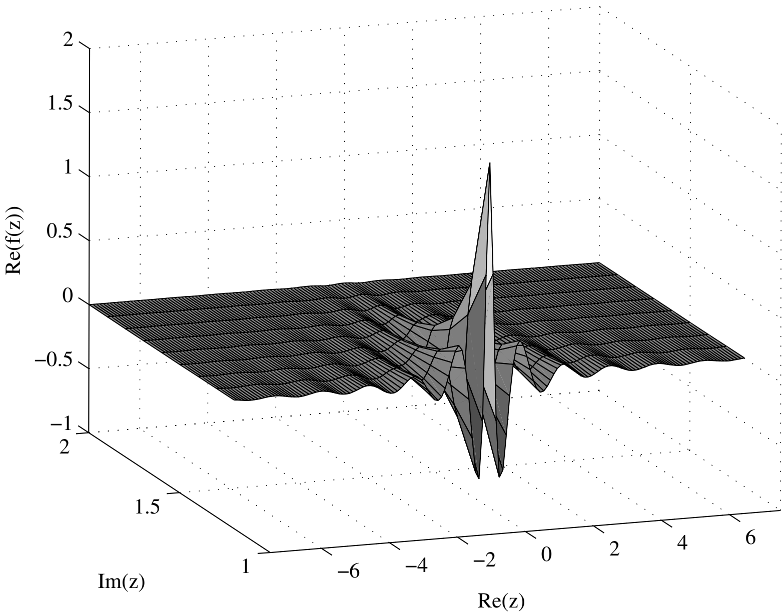

Fig. 2.1. Different contours of the Fourier transform in equation (2.26) for a strike of 90 units.

$$\begin{aligned}\psi^{x^{(m)}}(\mathbf{x}_t, z, w_0, \mathbf{w}, t, T) =& \mathbb{E}\left[e^{-\int_t^T r(\mathbf{x}_s)\,\mathrm{d}s + \imath z x_T^{(m)}}\right] \\ =& \int_{\mathbb{R}^M} e^{\imath z x_T^{(m)}} p(\mathbf{x}_t, \mathbf{x}_T, w_0, \mathbf{w}, t, T)\,\mathrm{d}\mathbf{x}_T,\end{aligned} \tag{2.27}$$

for all $m = 1, \ldots, M$. In the last equality of equation (2.27), the function $p(\mathbf{x}_t, \mathbf{x}_T, w_0, \mathbf{w}, t, T)$ represents the (discounted) transition probability density, starting with an initial state $\mathbf{x}_t$ and ending up in time T at $\mathbf{x}_T$. The continuous discounting is conducted with respect to $r(\mathbf{x}_{t^*})$ for $t > t^* \geq T$.

Obviously, if the stochastic process consists only of one variable x_t, the characteristic function $\psi^x(x_t, z, 0, 0, t, T)$ is then just the Fourier Transformation of the particular transition density function $p(x_t, x_T, 0, 0, t, T)$. Although the transform operation in equation (2.27) is performed with respect to the terminal state of one single random variable $x_T^{(m)}$, we have to consider the state of the vector $\mathbf{x}_t$ as an argument of the characteristic function. In fact, since we are looking at the overall expectation, equation (2.27) is generally built as the M-dimensional integral over the entire state vector $\mathbf{x}_T$[42]. Therefore, we are also able to apply the definition presented above of building a characteristic function for the more general case

$$g(\mathbf{x}_T) = g_0 + \mathbf{g}'\mathbf{x}_T \tag{2.28}$$

with $g_0 \in \mathbb{R}$ and $\mathbf{g} \in \mathbb{R}^M$. This implies, as long as $g(\mathbf{x}_T)$ is a linear combination of the elements in $\mathbf{x}_T$ that only one single transformation variable z necessary. Hence, if we are able to build the characteristic function for the scalar $g(\mathbf{x}_T)$[43], there is only a one-dimensional integral for the inverse operation to be performed, independent of the number of state variables included in $g(\mathbf{x}_T)$. Note, this powerful result will be used in our multi-factor framework. Equipped with these definitions we state next some general and important properties of Fourier Transformations on which we rely in our thesis.

Proposition 2.4.3 (Important Properties of Characteristic Functions and Fourier Transformations). *Let $\alpha, \beta, x, y \in \mathbb{R}$, and $f(x), g(y)$ some real-valued functions with Fourier transforms $\hat{f}(z), \hat{g}(z)$ and Fourier Transformation variable $z \in \mathbb{C}$. Then the following relations hold:*

1. *Linearity:*
$$\mathcal{F}^x[\alpha f(x) + \beta g(x)] = \alpha \hat{f}(z) + \beta \hat{g}(z).$$
2. *Differentiation:*
$$\mathcal{F}^x\left[\frac{\mathrm{d}^\alpha f(x)}{\mathrm{d}x^\alpha}\right] = (\imath z)^\alpha \hat{f}(z).$$
3. *Convolution:*
$$\mathcal{F}^x[f(x) * g(x)] = \hat{f}(z)\hat{g}(z).$$
4. *Symmetry:*
$$\pi f(x) = \int_0^\infty e^{-\imath z x} \hat{f}(z)\,\mathrm{d}z = \int_{-\infty}^0 e^{-\imath z x} \hat{f}(z)\,\mathrm{d}z.$$

[42] If $x_t^{(m)}$ would be no subordinated process and independent from all other state variables, equation (2.27) could still utilize the joint density function $p(\mathbf{x}_t, \mathbf{x}_T, w_0, \mathbf{w}, t, T)$ due to the possible discount factor including $r(\mathbf{x}_t)$.

[43] For example, calculating the general characteristic function for the short rate $r(\mathbf{x}_t)$ itself, we set $g(\mathbf{x}_T) = r(\mathbf{x}_T)$.

5. Relation of the Moment-Generating and the Characteristic Function:

$$\mathbb{E}\left[x^{\alpha}\right] = (-\imath)^{\alpha} \left. \frac{\mathrm{d}^{\alpha}\psi^{x}(x_t, z, 0, \mathbf{0}_M, t, T)}{\mathrm{d}z^{\alpha}} \right|_{z=0} .$$

Taking a second glance at Figure 2.1, we are able to justify the symmetry of the Fourier Transformation (2.26) of a real-valued function, mentioned in Proposition 2.4.3. Furthermore, one can clearly identify the dampening property of the characteristic function which is essential in developing a numerical algorithm to compute derivative prices. In the following, we show how the characteristic function for a scalar function $g(\mathbf{x}_T)$ is derived within the exponential-affine framework. Following Bakshi and Madan (2000), we interpret the characteristic function as a hypothetical contingent claim. Taking more elaborated payoff structures into account, we have to extend the list of permissible arguments for the characteristic function. The more general representation of the characteristic function, which we use hereafter is $\psi^{g(\mathbf{x})}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, t, T)$ with the complex-valued payoff representation at maturity T,

$$\psi^{g(\mathbf{x})}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, T, T) = e^{\imath z g(\mathbf{x}_T)}. \tag{2.29}$$

As discussed in the last section, we have to consider that all contingent claims need to be priced under the risk-neutral probability measure $\mathbb{Q}$. Hence, all prices are derived as discounted expectations. Consequently, the discounted expectation of the general form of the terminal condition can be represented as

$$\psi^{g(\mathbf{x})}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, t, T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(\mathbf{x}_s)\,\mathrm{d}s + \imath z g(\mathbf{x}_T)}\right]. \tag{2.30}$$

However, we need to compute discounted expectations, e.g. for vanilla zero-bond calls, or undiscounted expectations, e.g. in the case of futures instruments. Hence, for futures-style contracts, w_0 equals zero and $\mathbf{w}$ is a zero valued vector[44].

In calculating European derivative prices, we rather need the general characteristic function $\psi^{g(\mathbf{x})}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, t, T)$ than the special case of the

[44] The characteristic marking to market for standardized futures-style contracts results in the non-existence of a discount factor in the pricing formula and the relevant PDE, respectively, of such a contract under the risk-neutral measure $\mathbb{Q}$.

characteristic function without considering any discount factor, which is just $\psi^{g(\mathbf{x})}(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, t, T)$, where $\mathbf{0}_M$ represents a $M \times 1$ vector containing exclusively zeros. Applying Theorem 2.2.3 to our hypothetical claim with a solution according to equation (2.30), we take advantage of the Feynman-Kac representation to derive the partial differential equation. Simplifying and suppressing unnecessary notation, we write henceforth $\psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau) \equiv \psi^{g(\mathbf{x})}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, t, T)$ and then get the partial differential equation

$$\begin{aligned}
&\frac{\partial \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)}{\partial t} + \boldsymbol{\mu}^{\mathbb{Q}}(\mathbf{x}_t)' \frac{\partial \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)}{\partial \mathbf{x}_t} \\
&+ \frac{1}{2} \mathrm{tr} \left[\boldsymbol{\Sigma}(\mathbf{x}_t) \boldsymbol{\Sigma}(\mathbf{x}_t)' \frac{\partial^2 \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)}{\partial \mathbf{x}_t \partial \mathbf{x}_t'} \right] \\
&+ \mathbb{E}_{\mathbf{J}} \left[\boldsymbol{\psi}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau)' - \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau) \right] \boldsymbol{\lambda}^{\mathbb{Q}} \\
&\qquad = \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau) r(\mathbf{x}_t),
\end{aligned} \tag{2.31}$$

where the complex-valued vector $\boldsymbol{\psi}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau)$ contains all jump components with particular elements $(\boldsymbol{\psi}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau))_n = \psi(\mathbf{x}_t + \mathbf{j}_n, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)$. The vector $\mathbf{j}_n \in \mathbb{R}^M$ contains as mth element the random variable J_{mn} of the amplitude matrix $\mathbf{J}$. Every contingent claim or function dependent on $\mathbf{x}_t$, an arbitrage-free environment presupposed, has to satisfy the same Partial differential equation structure as given in equation (2.31). For example, the corresponding risk-neutral transition density for the characteristic function $\psi(\mathbf{x}_t, z, w_0, \mathbf{w}, w_0, \mathbf{w}, \tau)$, with $g(\mathbf{x}_T) = r(\mathbf{x}_T)$, which is actually $p(r(\mathbf{x}_t), r(\mathbf{x}_T), w_0, \mathbf{w}, t, T)$ need to satisfy the same partial differential equation as the characteristic function itself[45]. The only difference between them would be the particular terminal payoff condition. Hence, solving the above partial differential equation for $p(r(\mathbf{x}_t), r(\mathbf{x}_T), w_0, \mathbf{w}, t, T)$, we would impose the Dirac delta function as the relevant terminal condition, having its density mass exclusively concentrated in an infinite spike for $r(\mathbf{x}_T)$ at time T. Solving equation (2.31) together with this type of boundary condition can be quite challenging and is in many cases just impossible[46]. Thus, it is feasible to first solve equation (2.31) for the general characteristic function, with its smooth and continuous boundary function at T, and afterwards do some sort

[45] See Heston (1993), p. 331.

[46] A prominent example is given with the stochastic volatility model of Heston (1993), for which no closed-form representation of the transition density of the underlying equity log-price variable exists.

of normalized integration, the inverse Fourier Transformation, probably in a numerical manner, to get the desired result. Proceeding like this is a very elegant way to find some semi-analytic solution. In contrast, if we want to interpret the terminal payoff function in equation (2.29) as a hypothetical futures-style contract, with solution

$$\psi(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \tau) = \mathbb{E}^{\mathbb{Q}}\left[e^{\imath z g(\mathbf{x}_T)}\right], \tag{2.32}$$

we have a slightly different partial differential equation. In this case the dynamic behavior of $\psi(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \tau)$ is described by the slightly altered PDE

$$\begin{aligned} &\frac{\partial \psi(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \tau)}{\partial t} + \boldsymbol{\mu}^{\mathbb{Q}}(\mathbf{x}_t)' \frac{\partial \psi(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \tau)}{\partial \mathbf{x}_t} \\ &+ \frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Sigma}(\mathbf{x}_t)' \frac{\partial^2 \psi(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \tau)}{\partial \mathbf{x}_t \partial \mathbf{x}_t'}\right] \\ &+ \mathbb{E}_{\mathbf{J}}\left[\boldsymbol{\psi}(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \mathbf{J}, \tau)' - \psi(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \tau)\right]\boldsymbol{\lambda}^{\mathbb{Q}} \\ &= 0, \end{aligned} \tag{2.33}$$

Hence, the only difference to PDE (2.31) is that the right hand side is now equal to zero to contribute the missing discount rate. Moreover, we can use this futures-style characteristic function $\psi(\mathbf{x}_t, z, 0, \mathbf{0}_M, g_0, \mathbf{g}, \tau)$ to obtain the particular values of the undiscounted transition density function. Thus, to compute the probability density function of the short rate $r(\mathbf{x}_t)$, we use this futures-style solution of the characteristic function together with the identity $g(\mathbf{x}_t) = r(\mathbf{x}_t)$.

Consequently, using a *separation of variables* approach, the partial differential equations in (2.31) and (2.33) can be decoupled into a system of ordinary differential equations. Therefore, we assume for $\psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)$ the exponential-affine structure

$$\psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau) = e^{a(z,\tau) + \mathbf{b}(z,\tau)'\mathbf{x}_t + \imath z g_0}, \tag{2.34}$$

with the scalar and complex-valued coefficient function $a(z, \tau)$ and

$$\mathbf{b}(z,\tau) = \begin{pmatrix} \tilde{b}^{(1)}(z,\tau) \\ \tilde{b}^{(2)}(z,\tau) \\ \vdots \\ \tilde{b}^{(M)}(z,\tau) \end{pmatrix} + \imath z \begin{pmatrix} g^{(1)} \\ g^{(2)} \\ \vdots \\ g^{(M)} \end{pmatrix} = \tilde{\mathbf{b}}(z,\tau) + \imath z \mathbf{g},$$

denotes some complex-valued coefficient vector. In the next step we plug the required expressions of the candidate function (2.34) into equation (2.31). Starting with the time derivative, we get

$$\begin{aligned} &\frac{\partial \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)}{\partial t} \\ &\qquad = -\left(a(z,\tau)_\tau + \mathbf{b}(z,\tau)'_\tau \mathbf{x}_t\right) \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau), \end{aligned} \tag{2.35}$$

where $a(z,\tau)_\tau$ and $\mathbf{b}(z,\tau)_\tau$ are the first derivatives with respect to the time to maturity variable τ. The gradient vector with respect to the state variables $\mathbf{x}_t$ is given by

$$\frac{\partial \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)}{\partial \mathbf{x}_t} = \mathbf{b}(z,\tau)\psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau), \tag{2.36}$$

the Hesse matrix is

$$\frac{\partial^2 \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)}{\partial \mathbf{x}_t \partial \mathbf{x}_t'} = \mathbf{b}(z,\tau)\mathbf{b}(z,\tau)'\psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau), \tag{2.37}$$

and the jump component in equation (2.31) can be derived as

$$\begin{aligned} &\mathbb{E}_{\mathbf{J}}\left[\boldsymbol{\psi}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau)' - \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)\right] = \\ &\quad \mathbb{E}_{\mathbf{J}}\left[\boldsymbol{\psi}^*(z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau)' - 1\right] \psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau), \end{aligned} \tag{2.38}$$

with the normalized vector

$$\begin{aligned} \boldsymbol{\psi}^*(z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau) &= \frac{\boldsymbol{\psi}(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau)}{\psi(\mathbf{x}_t, z, w_0, \mathbf{w}, g_0, \mathbf{g}, \tau)} \\ &= \begin{pmatrix} e^{\mathbf{b}(z,\tau)'\mathbf{J}_1} \\ e^{\mathbf{b}(z,\tau)'\mathbf{J}_2} \\ \vdots \\ e^{\mathbf{b}(z,\tau)'\mathbf{J}_N} \end{pmatrix}. \end{aligned} \tag{2.39}$$

In this affine framework, it can be easily checked that the normalized amplitude vector $\boldsymbol{\psi}^*(z, w_0, \mathbf{w}, g_0, \mathbf{g}, \mathbf{J}, \tau)$ is independent of the actual state of $\mathbf{x}_t$, which results in the special form given by equation (2.39). Therefore, we are able to express the system of ODEs resulting from equations (2.31) and (2.33), respectively, and the affine form proposed in (2.34) in terms of the risk-neutral coefficients derived in Section 2.3.2. According to Theorem 2.2.3, the ODE which has to be solved for the scalar coefficient $a(z,\tau)$ is then

$$\begin{aligned} a(z,\tau)_\tau = -w_0 + \boldsymbol{\mu}_0^{\mathbb{Q}\prime}\mathbf{b}(z,\tau) + \frac{1}{2}\mathbf{b}(z,\tau)'\boldsymbol{\Sigma}_0\mathbf{b}(z,\tau) \\ + \mathbb{E}_{\mathbf{J}}\left[\boldsymbol{\psi}^*(z,w_0,\mathbf{w},g_0,\mathbf{g},\mathbf{J},\tau)' - 1\right]\boldsymbol{\lambda}^{\mathbb{Q}}, \end{aligned} \tag{2.40}$$

whereas for the vector coefficient $\mathbf{b}(z,\tau)$ we have to solve

$$\mathbf{b}(z,\tau)_\tau = -\mathbf{w} + \boldsymbol{\mu}_1^{\mathbb{Q}\prime}\mathbf{b}(z,\tau) + \frac{1}{2}\mathbf{b}(z,\tau)'\boldsymbol{\Sigma}_1\mathbf{b}(z,\tau), \tag{2.41}$$

with boundary conditions $a(z,0) = 0$ and $\mathbf{b}(z,0) = \imath z\mathbf{g}$, respectively. The parameters w_0 and $\mathbf{w}$, determine whether we consider a discount rate or not for the characteristic function. The mth element of $\mathbf{b}(z,\tau)'\boldsymbol{\Sigma}_1\mathbf{b}(z,\tau)$ can be computed as $\sum_{i,j} b(z,\tau)_i\,(\boldsymbol{\Sigma}_1)_{ijm}\,b(z,\tau)_j$[47]. Moreover, we want to emphasize that the trace operator is circular, meaning the equality

$$\operatorname{tr}\left[\boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Sigma}(\mathbf{x}_t)'\mathbf{b}(z,\tau)\mathbf{b}(z,\tau)'\right] = \operatorname{tr}\left[\mathbf{b}(z,\tau)'\boldsymbol{\Sigma}(\mathbf{x}_t)\boldsymbol{\Sigma}(\mathbf{x}_t)'\mathbf{b}(z,\tau)\right] \tag{2.42}$$

holds. Obviously, the right hand side of this last equation represents a scalar and therefore we are able to neglect the trace operator in equation (2.40) and equation (2.41), respectively.

In order to calculate derivatives prices, the coefficients $a(z,\tau)$ and $\mathbf{b}(z,\tau)$ need not exhibit closed-form solutions in any case. There are several scenarios conceivable, e.g. the time integrated expectations of the jump amplitudes have no closed-form representations, or the processes themselves have such complicated structures that there simply does not exist a closed-form solution of the coefficients $a(z,\tau)$ or $\mathbf{b}(z,\tau)$ of the characteristic function. However, if we are able to represent $a(z,\tau)$ and $\mathbf{b}(z,\tau)$ in terms of their ordinary differential equations (2.40) and (2.41), solutions can be efficiently obtained via a Runge-Kutta solver and appropriately integrated within our numerical pricing procedure, such that time consuming Monte-Carlo studies for the pricing of European interest-rate derivatives can be avoided.

[47] See Duffie, Pan and Singleton (2000), p. 1351.