On consistent stochastic processes in the Nelson-Siegel framework

T.K. Molenaars
Utrecht University, Department of Mathematics
RiskCo B.V., Financial Product Design and IT

Utrecht, June 12, 2012

Master's Thesis
Stochastics and Financial Mathematics

Supervisors:	Prof.dr. A.W. van der Vaart, VU University Amsterdam
	Dr. K. Dajani, Utrecht University
	Dr. M.A. Hemminga, RiskCo B.V.
Co-reader:	Dr. M.C.J. Bootsma, Utrecht University
Version:	Final

Preface

You are about to read the Master's thesis of Tomas Molenaars. This thesis was written as the final part of the Master's programme Stochastics and Financial Mathematics of Utrecht University (UU), VU University Amsterdam (VU) and University of Amsterdam (UvA).

I carried out my research at RiskCo B.V. in Utrecht. RiskCo's main activity is bridging the gap between financial product design and information technology. I was involved in the research and development group Forecasting Methodologies that develops approaches for the generation of financial and economic scenarios. These scenarios are used within Asset Liability Management and Solvency II calculations.

During my work on economic scenarios at RiskCo's prior to my Master's research, I encountered that interest rate is one of the most important, but also one of the most difficult aspects of generating economic scenarios. The interest rates do not only depend on the market (through the demand for capital), but are also highly correlated to governments' monetary policies. On the other hand, those same governments would like to be able to forecast the interest rates to construct their debt strategies. With interest rates being such an important, but difficult aspect of economic forecasting, this was the first main research subject of the research group.

After exploring several models and theories about interest rate forecasting with RiskCo's research group, I decided to use a topic we had encountered as a starting point for my work: jump processes in the Nelson-Siegel framework. From a business point of view my research may be of limited importance (for RiskCo), because I went for a purely theoretical mathematical approach. Nevertheless it was possible to do my research project at RiskCo's, of which this thesis is the result.

Acknowledgments I'm very thankful to Bert de Bock for giving me the opportunity to do my research at RiskCo's, which has been a very nice place to work. Besides Bert, I would like to thank my friend, classmate and colleague Nick Reinerink and colleague Marcus Hemminga for sharing ideas and thoughts.

I would also like to thank Aad van der Vaart for helping me to get the math correct. I appreciate the time you made free for me. Furthermore I will not forget the enthusiasm of Karma Dajani which helped me starting up my research project.

Last but not least I would like to thank Jöbke Janssen and Wouter Vink for simply being two very good friends. As if this were not enough reason to be acknowledged in this preface, I would like to thank you both for monitoring and being interested in the progress of my thesis. But I think that is of much less importance than being my friends.

Tomas Molenaars **Utrecht, May 2012**

Contents

Notation

In this section we give some definitions and notations we will use in the rest of this thesis. Those definitions are widely used, so this section is more about our notation. A more detailed explanation can be found in Chapter 8 of 'Stochastic Processes for Finance: Risk Management Tools' [3].

Definition 1. *A* zero-coupon bond *(or* discount bond*) is a contract which guarantees a pay-off of 1 euro at time T. The price of such a bond at time $t \leq T$ is denoted by $P(t,T)$. It is the amount we are willing to pay at time t to receive 1 euro at time T. Time T is called the* maturity.

By definition, $P(t,t)$ equals 1 for all t. The collection of all bond prices at time $t = 0$, $\{P(0,T)|T > 0\}$, is called the term structure of interest rates.

Definition 2. *The* yield to maturity*, $Y(t,T)$, is defined as the continuously compounded interest rate between times t and T such that $P(t,T) = e^{-Y(t,T)(T-t)}$, i.e. $Y(t,T) := -\frac{\log P(t,T)}{T-t}$.*

When we know the bond prices at given maturities, we can compute the yields to those maturities and vice versa.

Definition 3. *The* forward rate *for time T determined at time t is defined as $f(t,T) := -\frac{\partial}{\partial T}\log P(t,T)$.*

By integrating the forward rates and using that $P(t,t) = 1$, we get

$$\int_t^T f(t,s)\,\mathrm{d}s = -\log P(t,T) + \log P(t,t) = -\log P(t,T),$$

hence $P(t,T) = e^{-\int_t^T f(t,s)\,\mathrm{d}s}$ and $Y(t,T) = \frac{\int_t^T f(t,s)\,\mathrm{d}s}{T-t}$. The bond prices and yields can be computed from the forward rates and vice versa. It is obvious that those three definitions contain the same information. If we know one of these, we can compute the other two.

Using the forward rate, we define the short rate as follows:

Definition 4 (Short rate)**.** *The* short rate *at time t is defined as $r_t := \lim_{T\downarrow t} f(t,T)$.*

Definition 5 (Savings account)**.** *The* savings account*-process is defined as $B(t) := \exp\left(\int_0^t r_s\,\mathrm{d}s\right),\ 0 \leq t < \infty$.*

Let f be a function from $\mathbb{R} \times \mathbb{R}^d$ to $\mathbb{R}$. Denote by $D_x f(x,z)$ the partial derivative of function $f(x,z)$ to its first argument x and let $\nabla_z f(x,z)$ be the gradient of $f(x,z)$ where z is d-dimensional:

$$\nabla_z f(x,z) := \begin{pmatrix} \frac{\mathrm{d}f(x,z)}{\mathrm{d}z_1} \\ \vdots \\ \frac{\mathrm{d}f(x,z)}{\mathrm{d}z_d} \end{pmatrix}.$$

Chapter 1

Introduction

In 1987 Charles Nelson and Andrew Siegel published an article in the *Journal of Business* about the modeling of yield curves: 'Parsimonious Modeling of Yield Curves' [18]. Their aim was to construct a model of the yield curve that uses only a few parameters to describe the yield curve, while at the same time it had to be able to represent the various shapes a yield curve can have. The shapes generally associated with the yield curve are monotonic-, humped- and S-shaped. The Nelson-Siegel yield curve covers those shapes.

In 1994 Lars Svensson introduced an extended and more flexible version of the Nelson- Siegel yield curve by adding a second curvature term [22]. Nowadays, the Nelson-Siegel model either with or without the Svensson extension is widely used by central banks to estimate the yield curve (see ref. [15]).

Yield curve (or interest rate in general) forecasting is very important in bond portfolio management, for pricing derivatives and in risk management. D. Bolder stated that one central question in a government's debt strategy analysis is "how much of the federal government's annual borrowing needs should be financed with long-term coupon bonds versus short-term treasury bills?"[2]. One way to answer this question is to consider how various strategies perform under different interest rate outcomes. In order to do such an analysis, we need to have a model for the interest rates, to be able to generate interest rate scenarios.

In their article of 2006, Diebold and Li [9] proposed a very practical way to model the yield curve. They used neither the no-arbitrage approach (models like Hull-White [17]), nor the equilibrium approach (models like Vasicek [23], Cox et al. [7] and Duffie and Kan [10]). Instead, they fitted the yield curve using a Nelson-Siegel curve and interpreted the parameters as factors. By doing time-series analysis on the factors, they forecasted the yield curve. They called their method the *Dynamic Nelson-Siegel* (DNS). De Pooter examined this kind of forecasting by the Nelson-Siegel class with additional extensions [8].

Whereas the article of Diebold and Li was mainly written to propose a very practical way to forecast the yield curve, the theoretical background of the Nelson-Siegel dynamics had already been investigated by Filipović back in 1999 [13] and 2000 [14] and by Björk and Christensen in 1999 [1]. Filipović showed there exist no nontrivial interest rate model in the Nelson-Siegel framework when the parameters in the Nelson-Siegel framework are driven by a state space process, which provides an arbitrage-free interest rate model.

Where on the one hand people doubted whether the absence of arbitrage would give better forecasts (like Coroneo [6]), other people tried to make a theoretically rigid method, which still offers the flexibility of the Nelson-Siegel framework. Christensen, Diebold and Rudebusch [4, 5] derived a class of "affine arbitrage-free dynamic term structure models that approximate the widely used Nelson-Siegel yield curve specification". They introduced a yield-adjustment term to make the Nelson-Siegel term structure arbitrage-free. They called their models *arbitrage-free Nelson-Siegel* (AFNS) models.

In all the articles mentioned above continuous processes are used to fit and forecast the parameters. Continuous processes, by definition, do not allow the presence of jumps. However, the occurrence of a sudden change in interest rate, a jump, cannot be excluded.

At the end of April 2012, during the political conflict between the Dutch government and Geert Wilders, Fitch Ratings warned that the Netherlands faced losing the AAA-rating, because of the failing housing market and the lack of political action [12]. The consequences of such a devaluation cannot be predicted. Some people think interest rates on Dutch government bonds will increase with 50% if the Dutch government loses the AAA-rating, because traders will be less eager to lend money. This shows it makes sense to consider jumps in interest rates when we try to forecast them and use them to generate scenarios.

In 1996 Duffie and Kan already mentioned the possible perception of including jumps, see Chapter 11 of ref. [10]. Bolder confirmed that incorporation of jumps is an increase in reality. Moreover, he mentioned that this incorporation increases the complexity as well, by reducing the existence of closed form solutions for bond price functions [2]. Filipović mentioned the possibility to involve jumps to expand his results too [13].

The main line in this thesis is as follows. In Chapter 2 we introduce the Nelson-Siegel model and explore some of its properties. In Chapter 3 we define the consistent state space process and prove, following Filipović, there exists no nontrivial interest rate model driven by a consistent state space Itô process. In Chapter 4, we introduce a certain class of stochastic processes with independent jumps, which we called Independent jump processes, and stochastic calculus for jump processes. Furthermore we prove there exists no nontrivial interest rate model driven by a consistent state space Independent jump process (what has not been done before).

Chapter 2

Nelson-Siegel framework

In this chapter we introduce the Nelson-Siegel curves as constructed by Nelson and Siegel in 1987 [18]. We also introduce some notation to simplify expressions involving the Nelson-Siegel curves and we show some properties of the curves. Furthermore, we mention the way to estimate the parameters following DNS and possible issues with it.

2.1 The original Nelson-Siegel model

The model Nelson and Siegel constructed was parameterized by only four parameters, denoted by $\beta_0, \beta_1, \beta_2$ and λ. The (instantaneous) forward rate in the Nelson-Siegel model is given by:

$$f_t(\tau) := f(t, t+\tau) = \beta_0 + \beta_1 \cdot \exp(-\lambda\tau) + \beta_2 \cdot [\lambda\tau \exp(-\lambda\tau)], \tag{2.1}$$

where λ equals $1/\tau$ and τ equals m in the article of Nelson and Siegel. Integrating from 0 to τ and dividing by τ gives the yield as function to maturity τ (see page v):

$$y_t(\tau) := Y(t, t+\tau) = \beta_0 + \beta_1 \cdot \frac{1-e^{-\lambda\tau}}{\lambda\tau} + \beta_2 \cdot \left[\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right]. \tag{2.2}$$

This is the formula for the yield to maturity in the Nelson-Siegel framework, which is used in this thesis. Some people, including Filipović [13], replace $\beta_2 \cdot \lambda$ by one single parameter in (2.1).

In figure 2.1 we see a few examples of the different shapes the Nelson-Siegel curve can represent. The estimated curves for two values for λ as well as real yield data is plotted. The Nelson-Siegel fits the inverted yield curve (first and last figure) as good as the normal (second figure) and humped (third figure) yield curve. In section 2.2 we will talk about the estimation and the data used in more detail.

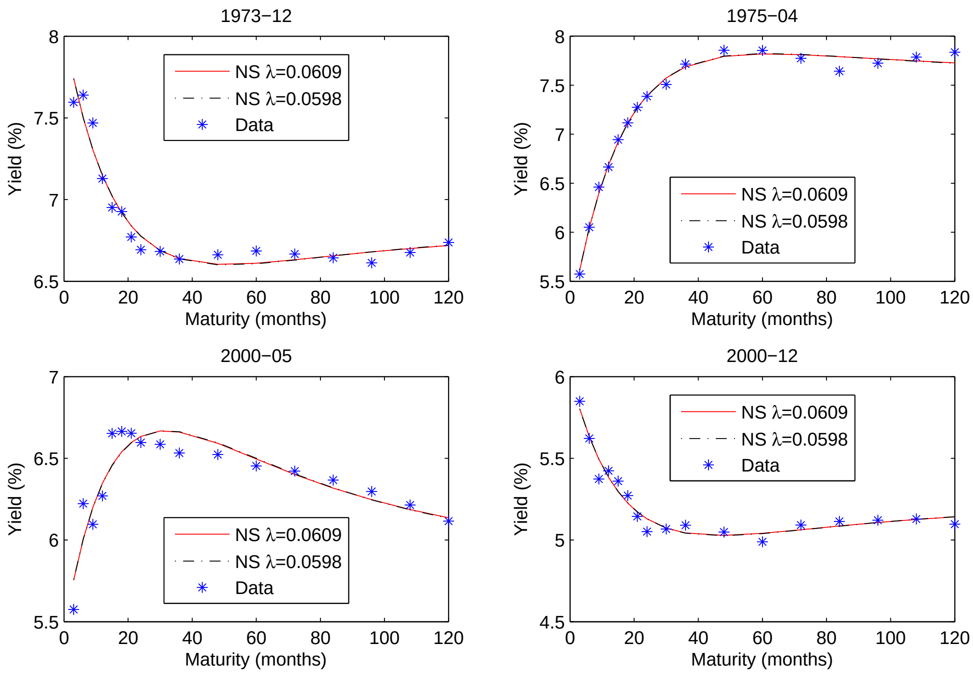

Figure 2.1: Yield curve data and Nelson-Siegel (NS) fits for two different λ's for four different dates.

2.1.1 Vector notation

To abbreviate the preceding formulas, we can rewrite them using vector notation. Write the inner product of two column vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$ as $\mathbf{a}^*\mathbf{b}$, where $\mathbf{a}^*$ is the transpose of $\mathbf{a}$. Define

$$\boldsymbol{\beta} := \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \ \mathbf{w}_t := \begin{pmatrix} 1 \\ e^{-\lambda t} \\ \lambda t e^{-\lambda t} \end{pmatrix}.$$

The dependence of $\mathbf{w}_t$ on the time is denoted by the subscript t. Note that $\boldsymbol{\beta}$ does not depend on t. Equation (2.1) can now be written as $f_t(\tau) = \mathbf{w}_\tau^*\boldsymbol{\beta}$ and (2.2) can be written as $y_t(\tau) = \mathbf{W}_\tau^*\boldsymbol{\beta}$, where $\mathbf{W}_t$ is defined as

$$\mathbf{W}_t := \begin{pmatrix} 1 \\ \frac{1-e^{-\lambda t}}{\lambda t} \\ \frac{1-e^{-\lambda t}}{\lambda t} - e^{-\lambda t} \end{pmatrix}.$$

Because $\boldsymbol{\beta}$ does not depend on t, the following holds: By definition (see page v) $y_t(\tau) = \frac{1}{\tau}\int_0^\tau f_t(s)\,\mathrm{d}s$, and using $f_t(\tau) = \mathbf{w}_\tau^*\boldsymbol{\beta}$ we can write

$$y_t(\tau) = \frac{1}{\tau}\int_0^\tau \mathbf{w}_s^*\boldsymbol{\beta}\,\mathrm{d}s = \left(\frac{1}{\tau}\int_0^\tau \mathbf{w}_s\,\mathrm{d}s\right)^* \boldsymbol{\beta}.$$

The second equality is valid because of the fact that $\boldsymbol{\beta}$ does not depend on t and the integral $\int_0^t \mathbf{w}_s \, \mathrm{d}s$ is defined componentwise:

$$\frac{1}{\tau}\int_0^\tau \mathbf{w}_s \, \mathrm{d}s = \frac{1}{\tau}\int_0^\tau \begin{pmatrix} 1 \\ e^{-\lambda s} \\ \lambda s e^{-\lambda s} \end{pmatrix} \mathrm{d}s := \begin{pmatrix} \frac{1}{\tau}\int_0^\tau 1 \, \mathrm{d}s \\ \frac{1}{\tau}\int_0^\tau e^{-\lambda s} \, \mathrm{d}s \\ \frac{1}{\tau}\int_0^\tau \lambda s e^{-\lambda s} \, \mathrm{d}s \end{pmatrix}.$$

This equals $\mathbf{W}_\tau$ (as expected).

2.1.2 Weight factor characterization

All parameters have different influences. If two parameters would have had the same influence, we would not need both. Because the parameters are constant in τ (time to maturity), their influence in the forward rates and yields depends on the weight factors. With weight factors, we mean the components of $\mathbf{w}_\tau$ and $\mathbf{W}_\tau$ respectively. The fact that the parameters β_0, β_1 and β_2 do not depend on t is already mentioned above; they were denoted by the time-independent vector $\boldsymbol{\beta}$. Remark that λ influences $\mathbf{w}_\tau$ and $\mathbf{W}_\tau$. (Later on we will write $\mathbf{w}_\tau(\lambda)$ and $\mathbf{W}_\tau(\lambda)$.)

The influence of $\boldsymbol{\beta}^i$ on $f_t(\tau)$ (respectively $y_t(\tau)$) is measured by the weight $\mathbf{w}_\tau^i$ (respectively $\mathbf{W}_\tau^i$). Let us investigate the weight functions. First, we can examine their limiting behavior, $t \downarrow 0$ and $t \to \infty$:

$$\lim_{t\downarrow 0} \mathbf{w}_t = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \lim_{t\downarrow 0} \mathbf{W}_t = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

$$\lim_{t\to\infty} \mathbf{w}_t = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \lim_{t\to\infty} \mathbf{W}_t = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

We see $\lim_{t\downarrow 0} \mathbf{w}_t = \lim_{t\downarrow 0} \mathbf{W}_t$ and $\lim_{t\to\infty} \mathbf{w}_t = \lim_{t\to\infty} \mathbf{W}_t$; the limiting behavior of the weight functions is the same.

For the forward rate and yield at time $t = 0$, only β_0 and β_1 contribute, whereas in the long run only β_0 contributes. In figure 2.2 the weights for β_0, β_1 and β_2 are plotted. The shapes of the weightfunctions are the same in both cases (forward rates and yields). The properties as described above are recognizable in the picture: the weight of β_1 is 1 for $t \downarrow 0$ and 0 for $t \to \infty$, whereas the weight of β_0 is 1 for both limits and the weight of β_2 is 0 in both limits.

From figure 2.2 we can understand the following characterization of the weights (as explained for example also in ref. [8]). The weight for β_0 is constant and it is the only weight that does not decay to zero in the limit. Therefore the component of β_0 is characterized as the long-term. The weight

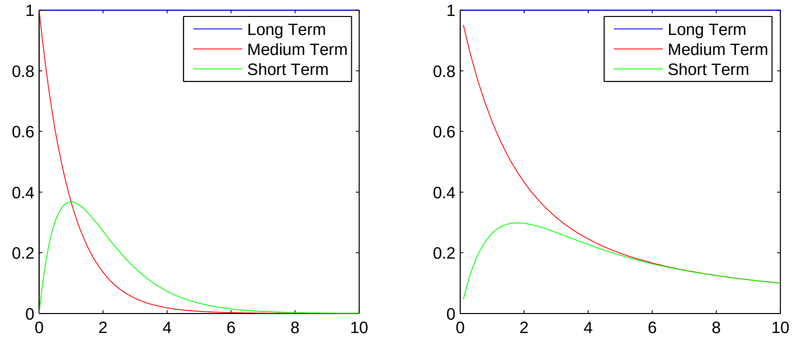

Figure 2.2: Left: components of $\mathbf{w}_t$, right: components of $\mathbf{W}_t$.

of β_1 is one in $t = 0$ but decreases to zero; therefore the component on β_1 is characterized as the short-term. λ determines the rate of decay. The weight of β_2 is the only term that starts at zero (and is therefore not short-term) and which decays to zero (and is therefore not long-term). The component on β_2 is therefore called medium-term. λ determines the maturity on which this weight reaches is maximum.

Diebold and Li give [9] a second interpretation of the factors. The long-term factor β_0 governs the yield curve *level*. To see this, note that changing β_0 moves the whole yield curve up or down: the loading of this factor is identical for all t. For the short-term factor β_1, note that a change of β_1 influences the short yields more than the long yields, therefore changing the *slope* of the curve. Finally, the medium-term factor β_2 does not influence the very short yields or very long yields, but will have effect on the medium-term yields, therefore influencing the yield curve *curvature*.

The slope of the yield curve is often defined as the difference between the long term yields and the short term yields. If one defines the yield curve slope (like Frankel and Lown [16]) as $\lim_{t\to\infty} y(t) - y(0)$, we get $\lim_{t\to\infty} y(t) - y(0) = \beta_0 - (\beta_0 + \beta_1) = -\beta_1$, so β_1 is indeed directly related to the slope.

2.2 Dynamic Nelson-Siegel

In their article 'Forecasting the term structure of government bond yields' [9], Diebold and Li introduce a dynamic model using the Nelson-Siegel curve. Their main idea is to interpret the parameters "as a three-dimensional parameter evolving dynamically". They consider the parameters as being some time series and this idea is reflected in the following notation:

$$y_t(\tau) = \beta_{0t} + \beta_{1t} \cdot \frac{1 - \exp(-\lambda\tau)}{\lambda\tau} + \beta_{2t} \cdot \left[\frac{1 - \exp(-\lambda\tau)}{\lambda\tau} - \exp(-\lambda\tau)\right]. \tag{2.3}$$

For a fixed t, this expression gives the yield curve; for fixed τ this expression gives the evolution of the yield of maturity τ over time.

In figure 2.4 we plotted the values for β_0, β_1 and β_2 for the data provided by Diebold and Li. The data set contains unsmoothed Fama-Bliss zero yields from U.S. Treasuries from January 1970 through December 2000.[1] By analyzing those time series, they try to predict the yield curve. Diebold and Li use the (fixed) value 0.0609 for λ. Recall that λ determines where the maximum of the medium-term weight factor is situated. Diebold and Li argue that this maximum is commonly situated between the two- and three-year maturity. They simply pick the average, 30 months, and state that $\lambda = 0.0609$ maximizes the medium-term weight factor at exactly 30 months. This is not true: the maximum of the function $f(x) = \frac{1-\exp(-x)}{x} - \exp(-x)$ is reached in $x = 1.79328$ which implies $\lambda = x/\tau = 1.79328/30 = 0.059776$. As is shown in figure 2.1, the difference between the fits with both λ's is very small.

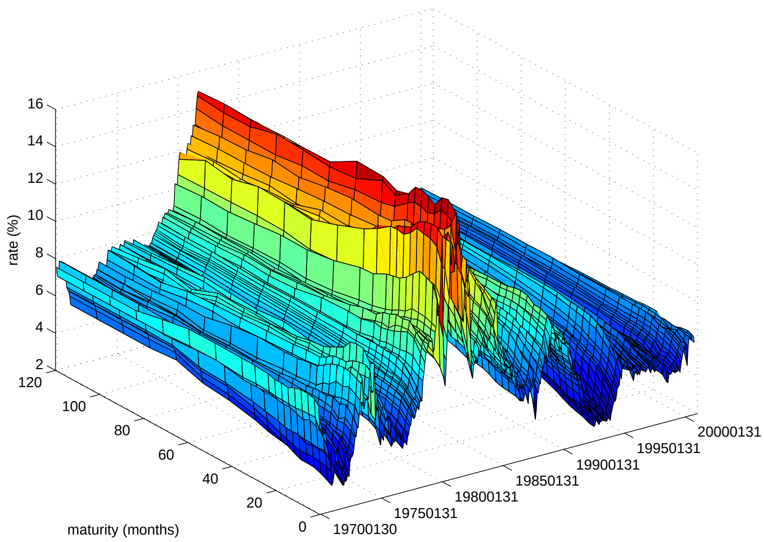

Figure 2.3: Data used by Diebold and Li, see page 7.

2.2.1 Estimation of the parameters

In this section, we explain the way Diebold and Li estimate the parameters in their article. Suppose we are given some yield curve data: a list of maturities $\{\tau_1, \dots, \tau_N\}$ and yields to those maturities $\{y_1, \dots, y_N\}$. Our aim is to fit a yield curve with the Nelson-Siegel form to this data. The system we want

[1]The data set can be downloaded from
http://www.ssc.upenn.edu/~fdiebold/papers/paper49/FBFITTED.txt

to solve is given by:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} \boldsymbol{W}^*_{\tau_1} \\ \vdots \\ \boldsymbol{W}^*_{\tau_N} \end{pmatrix} \boldsymbol{\beta} + \epsilon. \tag{2.4}$$

To simplify this expression, we denote the matrix with rows $\boldsymbol{W}^*_{\tau_1}, \ldots, \boldsymbol{W}^*_{\tau_N}$ by $\underline{\boldsymbol{W}}$ and $\boldsymbol{y}$ is the column vector of yields $y_1, \ldots, y_N$. The system (2.4) can now be written as:

$$\boldsymbol{y} = \underline{\boldsymbol{W}}\boldsymbol{\beta} + \epsilon.$$

The vector $\epsilon \in \mathbb{R}^N$ denotes the error made by estimating the data $\boldsymbol{y}$ by $\underline{\boldsymbol{W}}\boldsymbol{\beta}$. Remark that the matrix $\underline{\boldsymbol{W}}$ depends on λ, hence this is not a linear system in $\beta_0, \beta_1, \beta_2$ and λ! We should estimate the parameters using nonlinear least squares. By fixing λ in advance however, we make it a linear system. For $N > 3$ this is an overdetermined system and we can use Ordinary Least Squares to find an (approximate) solution. The Ordinary Least Squares solution is given by $\hat{\boldsymbol{\beta}} = (\underline{\boldsymbol{W}}^*\underline{\boldsymbol{W}})^{-1}\underline{\boldsymbol{W}}^*\boldsymbol{y}$.

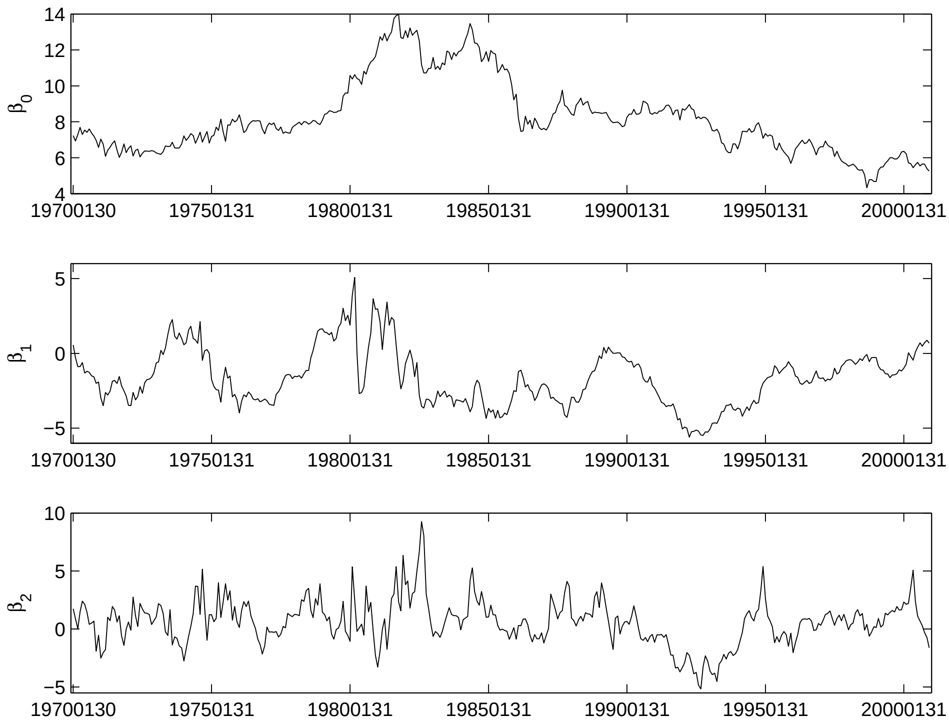

Figure 2.4: Estimated values for β_0, β_1 and β_2 from zero yields obtained from U.S. Treasury price quotes, data used by Diebold and Li.

2.2.2 Decay parameter

The decay parameter λ is, just like the β's, a parameter used to fit the curve as good as possible to the data. Fixing the λ during your whole analysis and during your forecast, like Diebold and Li do, is not be preferable from a 'best fit' point of view. Fixing λ simplifies the analysis because we can use Ordinary Least Squares as explained before. De Pooter however shows that it is not impossible to use estimation techniques to find the best (or better) λ during forecasting [8].

In our theoretical analysis in the following chapters we do not impose any restriction on λ: it will be treated in the same way as β_0, β_1 and β_2.

Chapter 3

Consistency: Itô process

After introducing the Nelson-Siegel framework, we are now going to introduce some math. In this chapter we introduce the class of *consistent state space processes* as used by Filipović. These processes provide an arbitrage-free interest rate model, in this case for the Nelson-Siegel forward rate family, when representing the parameters $\beta_0, \beta_1, \beta_2$ and λ of the Nelson-Siegel family. The consistent state space processes will be characterized by their parameters which will become clear later on.

We use the following version of the expression for the forward rates given by (2.1): We replace the parameters $\beta_0, \beta_1, \beta_2$ and λ by z_1, z_2, z_3 and z_4. Furthermore, we make in this notation the dependence of the forward rates on the parameters $z = (z_1, z_2, z_3, z_4)$ clear:

$$F(x, z) = z_1 + z_2 e^{-z_4 x} + z_3 z_4 x e^{-z_4 x}. \tag{3.1}$$

The notation we use coincides with the notation Nelson and Siegel originally used in ref. [9]. Remark that the notation of Filipović [13] slightly differs from the definition of Nelson and Siegel; Filipović replaced $z_3 z_4$ by z_3.

A good choice of the parameter $z \in \mathbb{R}^4$ gives today's term structure of interest rates, where $x \geq 0$ denotes the time to maturity. After choosing z_4, this is easily done by Ordinary Least Squares as explained in Section 2.2.1. Remark that the yield curve flattens for longer maturities, hence we restrict z to what we call the state space $\mathcal{Z} = \{z = (z_1, \ldots, z_4) \in \mathbb{R}^4 | z_4 > 0\}$. Using page v we have $y(x, z) = \frac{1}{x} \int_0^x F(\eta, z)\, \mathrm{d}\eta$ and hence the (structure of) bond prices are given by

$$G(x, z) = \exp(-y(x, z)x) = \exp\left(-\int_0^x F(\eta, z)\, \mathrm{d}\eta\right). \tag{3.2}$$

This is a function from $[0, \infty) \times \mathcal{Z}$ to $\mathbb{R}_+$. Remark that it is C^∞ in x because both the exponential function and $F(x, z)$ are C^∞.

When we view the parameters as time dependent and try to forecast them, we indirectly forecast a yield curve. The question rises, however, whether this prediction, assuming some dynamics of the parameters, is arbitrage-free. The next step is to assume the parameters z follow some state space process $Z := (Z_t)_{0\leq t<\infty}$ with values in $\mathcal{Z}$, and investigate whether $F(., Z)$ provides an arbitrage- free interest rate model.

The bond prices are functions of the yields (or forward rates) (see Definition 1). If we make the yield curve (or forward rate) dependent on some dynamic process, the bond prices will depend on this dynamic process too. Suppose we have an expression for the forward rate $F(x, z)$. The corresponding bond prices are then (see equation (3.2)) denoted by $G(x, z)$ and using this, the price at time t of a zero coupon bond with maturity T is given by

$$P(t, T) := G(T - t, Z_t),$$

where Z_t denotes the value of the process Z at time t.

We will discount the bond price with the process for the savings account, see page v: $B(t) := \exp\left(\int_0^t r_s\,\mathrm{d}s\right)$, where r_t is the short rate.

Furthermore we know that the discounted bond prices have to be martingales with respect to the risk free (martingale-)measure $\mathbb{P}$. This is the idea behind the following definition by Filipović:

Definition 3.1 (Consistency)**.** *The state space process Z is called* consistent *with the Nelson-Siegel family, if the discounted bond price is a $\mathbb{P}$-martingale, for all $T < \infty$, i.e. if*

$$\left(\frac{P(t, T)}{B(t)}\right)_{0\leq t\leq T} \tag{3.3}$$

is a $\mathbb{P}$-martingale, for all $T < \infty$.

Now we have defined the consistent state space process, we take for the state space process Z an Itô process (as defined in the next section) and formulate conditions on the parameters of the process to make it consistent. We prove that the interest rate models driven by a consistent Itô process are trivial, in line with Filipović in his paper [13].

3.1 Stochastic process without jumps

3.1.1 Multidimensional Itô process

The state space process we will use for this first part is defined below in Definition 3.4. It is a so called Itô process. Almost all stochastic processes without jumps are Itô processes ([20], p. 143).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $(\mathcal{F}_t)_{0\leq t<\infty}$ a filtration satisfying the usual conditions (see ref. [21]), and let $W = (W_t^1, \ldots, W_t^d)_{0\leq t<\infty}$ be standard d-dimensional $(\mathcal{F}_t)$-Brownian motion, $1 \leq d$.

First, let $d = 1$ to get the following definition of an Itô process which can be found in ref. [20], page 143.

Definition 3.2. *An* Itô process X_t *is a stochastic process of the form*

$$X_t = X_0 + \int_0^t b_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s, \tag{3.4}$$

where W *is 1-dimensional Brownian motion,* X_0 *is nonrandom (i.e.* $\mathcal{F}_0$*-measurable) and with* b *and* σ *progressively measurable*[1] *processes with values in* $\mathbb{R}$*, respectively* $\mathbb{R}$*, such that*

$$\int_0^t |b_s| \, \mathrm{d}s < \infty \text{ and } \int_0^t |\sigma_s|^2 \, \mathrm{d}s < \infty, \quad \mathbb{P}\text{-a.s.}, \quad \text{for all } t > 0.$$

We can also define the Itô process with a stochastic integral driven by a multidimensional Brownian motion:

Definition 3.3. *An* Itô process X_t *driven by a multidimensional Brownian motion is a stochastic process of the form*

$$X_t = X_0 + \int_0^t b_s \, \mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_s^j \, \mathrm{d}W_s^j, \tag{3.5}$$

where W *is d-dimensional Brownian motion,* X_0 *is nonrandom (i.e.* $\mathcal{F}_0$*-measurable) and with* b *and* σ *progressively measurable processes with values in* $\mathbb{R}$*, respectively* $\mathbb{R}^d$*, such that, for* $j = 1, \dots, d$*,*

$$\int_0^t |b_s| \, \mathrm{d}s < \infty \text{ and } \int_0^t |\sigma_s^j|^2 \, \mathrm{d}s < \infty, \quad \mathbb{P}\text{-a.s.}, \quad \text{for all } t > 0.$$

This definition can be extended to a definition for a multidimensional Itô process:

Definition 3.4 (Multidimensional Itô process)**.** *The* multidimensional Itô process $Z = (Z_t)_{0 \le t < \infty}$*,* $Z_t = (Z_t^1, \dots, Z_t^4)$ *is given by*

$$Z_t^i = Z_0^i + \int_0^t b_s^i \, \mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_s^{ij} \, \mathrm{d}W_s^j, \qquad i = 1, \dots, 4, \tag{3.6}$$

where W *is d-dimensional Brownian motion,* Z_0^i *is nonrandom (i.e.* $\mathcal{F}_0$*-measurable) and with* b *and* σ *progressively measurable processes with values in* $\mathbb{R}^4$*, respectively* $\mathbb{R}^{4 \times d}$*, such that, for* $i = 1, \dots, 4$*,* $j = 1, \dots, d$*,*

$$\int_0^t |b_s^i| \, \mathrm{d}s < \infty \text{ and } \int_0^t |\sigma_s^{ij}|^2 \, \mathrm{d}s < \infty, \quad \mathbb{P}\text{-a.s.}, \quad \text{for all } t > 0.$$

The conditions on b and σ are to be sure that the integrals in the right hand sides of (3.4), (3.5) and (3.6) are defined and the integral with respect to the Brownian motion is a martingale.

[1] A process X_t is called progressively measurable with respect to the filtration $\mathcal{F}$ if $\forall T \ge 0$, $(\omega, t) \mapsto X_t(\omega)$ considered as a map between $\Omega \times [0, T] \to \mathbb{R}$ is measurable with respect to $\mathcal{F}_T \otimes \mathcal{B}([0, T]) \to \mathcal{B}(\mathbb{R})$. Progressively measurableness implies adaptedness. See page 2 of Stochastic Integration [21].

3.1.2 Stochastic Calculus for Multidimensional Itô processes

Equation (3.6) for the multidimensional Itô process of Definition 3.4 can also be stated in its differential notation as

$$\mathrm{d}Z_t^i = b_t^i \,\mathrm{d}t + \sum_{j=1}^{d} \sigma_t^{ij} \,\mathrm{d}W_t^j. \tag{3.7}$$

From this definition is it easy to see what the quadratic (co)variation $[Z^i, Z^j]_t$ is (background information and more details about the quadratic (co)variation can be found in Appendix B):

$$\begin{aligned}\mathrm{d}[Z^i, Z^j]_t = \mathrm{d}Z_t^t \,\mathrm{d}Z_t^j &= b_t^i b_t^j \,\mathrm{d}t \,\mathrm{d}t + b_t^i \sum_{k=1}^{d} \sigma_t^{jk} \,\mathrm{d}t \,\mathrm{d}W_t^k \\ &+ b_t^j \sum_{k=1}^{d} \sigma_t^{ik} \,\mathrm{d}t \,\mathrm{d}W_t^k + \sum_{k=1}^{d} \sigma_t^{ik} \sum_{l=1}^{d} \sigma_t^{jl} \,\mathrm{d}W_t^k \,\mathrm{d}W_t^l.\end{aligned}$$

Because $\mathrm{d}t \,\mathrm{d}t = 0$, $\mathrm{d}t \,\mathrm{d}W_t^k = 0$, for all k and $\mathrm{d}W_t^i \,\mathrm{d}W_t^j = \mathbf{1}_{ij} \,\mathrm{d}t$, this equals

$$\mathrm{d}Z_t^i \,\mathrm{d}Z_t^j = \sum_{k=1}^{d} \sigma_t^{ik} \sum_{l=1}^{d} \sigma_t^{jl} \,\mathrm{d}W_t^k \,\mathrm{d}W_t^l = \sum_{k=1}^{d} \sigma_t^{ik} \sigma_t^{jk} \,\mathrm{d}t.$$

Hence the quadratic (co)variation of the multidimensional Itô process is given by

$$[Z^i, Z^j]_t = \int_0^t \sum_{k=1}^{d} \sigma_s^{ik} \sigma_s^{jk} \,\mathrm{d}s.$$

For a function $f(t, z)$ with $z \in \mathbb{R}^d$ for which the partial derivatives $\frac{\mathrm{d}f(t,z)}{\mathrm{d}t}$, $\frac{\mathrm{d}f(t,z)}{\mathrm{d}z^i}$ and $\frac{\partial^2 f(t,z)}{\partial z^i \partial z^j}$ for $1 \leq i, j \leq 4$ are defined and continuous, we can formulate an expression for the multidimensional Itô formula:

$$\mathrm{d}f(t, Z_t) = \frac{\mathrm{d}f(t, Z_t)}{\mathrm{d}t} \,\mathrm{d}t + \sum_{i=1}^{4} \frac{\mathrm{d}f(t, Z_t)}{\mathrm{d}z^i} \,\mathrm{d}Z_t^i + \sum_{i,j=1}^{4} \frac{1}{2} \frac{\partial^2 f(t, Z_t)}{\partial z^i \partial z^j} \,\mathrm{d}[Z^i, Z^j]_t.$$

When we substitute the equation for $\mathrm{d}Z_t$ (3.7) and use the fact that $\mathrm{d}[Z^i, Z^j]_t = \sum_{k=1}^{d} \sigma_t^{ik} \sigma_t^{jk} \,\mathrm{d}t$, we get:

$$\begin{aligned}\mathrm{d}f &= \frac{\mathrm{d}f}{\mathrm{d}t} \,\mathrm{d}t + \sum_{i=1}^{4} \frac{\mathrm{d}f}{\mathrm{d}z^i} (b_t^i \,\mathrm{d}t + \sum_{j=1}^{d} \sigma_t^{ij} \,\mathrm{d}W_t^j) + \sum_{i,j=1}^{4} \frac{1}{2} \frac{\partial^2 f}{\partial z^i \partial z^j} \sum_{k=1}^{d} \sigma_t^{ik} \sigma_t^{jk} \,\mathrm{d}t \\ &= \frac{\mathrm{d}f}{\mathrm{d}t} \,\mathrm{d}t + \left(\sum_{i=1}^{4} \frac{\mathrm{d}f}{\mathrm{d}z^i} b_t^i + \sum_{i,j=1}^{4} \frac{1}{2} \frac{\partial^2 f}{\partial z^i \partial z^j} \sum_{k=1}^{d} \sigma_t^{ik} \sigma_t^{jk} \right) \mathrm{d}t \\ &\qquad + \sum_{i=1}^{4} \frac{\mathrm{d}f}{\mathrm{d}z^i} \sum_{j=1}^{d} \sigma_t^{ij} \,\mathrm{d}W_t^j,\end{aligned} \tag{3.8}$$

where $f := f(t, Z_t)$ for simplicity.

Defining $a_t := \sigma_t \sigma_t^*$, such that $\sum_{k=1}^d \sigma_t^{ik} \sigma_t^{jk} = a_t^{ij}$, and defining

$$A_t(\omega) f(t,z) = b_t(\omega) \cdot \nabla_z f(t,z) + \frac{1}{2} \sum_{i,j=1}^{4} a_t^{ij}(\omega) \frac{\partial^2 f(t,z)}{\partial z^i \partial z^j}, \quad 0 \le t < \infty, \quad z \in \mathcal{Z},$$

we can write (3.8) as (again with $f := f(t, Z_t)$):

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}t}\,\mathrm{d}t + A_t f \;\,\mathrm{d}t + \sum_{i=1}^{4} \frac{\mathrm{d}f}{\mathrm{d}z^i} \sum_{j=1}^{d} \sigma_t^{ij}\,\mathrm{d}W_t^j, \quad 0 \le t < \infty. \tag{3.9}$$

This formula will be used in the proof of the theorem in the following section to simplify the expressions.

3.2 Consistency of the Itô process

In the following theorem, we state a condition using the forward rate curve for the Itô process $Z := (Z_t)_{0 \le t < \infty}$ to be consistent, following Definition 3.1. In this theorem we use the terminology *forward curve family* F: this is nothing else than the set of forward rate curves with expression $F(x, z)$ for $x \in \mathbb{R}_+$ and $z \in \mathcal{Z}$: $F = \{F(x,z) | x \in \mathbb{R}_+, z \in \mathcal{Z}\}$. This does not have to be the Nelson-Siegel forward curve.

Theorem 3.1. *Suppose $Z = (Z_t)_{0 \le t < \infty}$ follows the Itô process of Definition 3.4 with values in $\mathcal{Z}$. Then Z is consistent, following Definition 3.1, with the forward curve family F only if*

$$\begin{aligned} D_x F(x, Z_t) = b \cdot \nabla_z F(x, Z_t) + \frac{1}{2} \sum_{i,j=1}^{4} a^{ij} \frac{\partial^2}{\partial z^i \partial z^j} F(x, Z_t) \\ - \sum_{i,j=1}^{4} a^{ij} \left(\frac{\partial}{\partial z^i} F(x, Z_t) \int_0^x \frac{\partial}{\partial z^j} F(\eta, Z_t)\,\mathrm{d}\eta \right) \end{aligned}$$

for all $x \ge 0$, $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-a.s., where $a = \sigma\sigma^*$.

Proof. In our new framework, the bond price depends on the parameter-vector Z_t, hence we will use the expression for the multi-dimensional Itô formula for the bond price.

Suppose we have an expression for the forward rate $F(x, z)$ (remark that this notation relates to the definition of the forward rate (2.1) as $f_t(x) = F(x, Z_t)$). The corresponding bond prices are then (see equation (3.2)) denoted by $G(x, z)$ and using this, the price at time t of a zero coupon bond with maturity T is given by

$$P(t, T) := G(T - t, Z_t).$$

Combining this with (3.9) gives the following Itô formula in differential form:

$$\mathrm{d}P(t,T) = \big(-D_x G(T-t,Z_t) + A_t G(T-t,Z_t)\big)\,\mathrm{d}t + \sum_{i=1}^{4} \frac{\mathrm{d}G}{\mathrm{d}z^i} \sum_{j=1}^{d} \sigma_t^{ij}\,\mathrm{d}W_t^j,$$

where we write $G := G(T-t,Z_t)$ for simplicity. To simplify this equation further, write $\mathrm{d}W_t = (\mathrm{d}W_t^1,\dots,\mathrm{d}W_t^d)^*$ such that $\sum_{j=1}^{d} \sigma_t^{ij}\,\mathrm{d}W_t^j = (\sigma_t\,\mathrm{d}W_t)_i$ and $\sum_{i=1}^{4} \frac{\mathrm{d}G(T-t,Z_t)}{\mathrm{d}z^i}(\sigma_t\,\mathrm{d}W_t)_i = (\nabla_z G(T-t,Z_t))^* \sigma_t\,\mathrm{d}W_t$. In integral form, above equation can be written as:

$$\begin{aligned} P(t,T) = P(0,T) &+ \int_0^t (A_s G(T-s,Z_s) - D_x G(T-s,Z_s))\,\mathrm{d}s \\ &+ \int_0^t \nabla_z G(T-s,Z_s)^* \sigma_s\,\mathrm{d}W_s, \quad 0 \le t \le T, \quad \mathbb{P}\text{-}a.s.. \end{aligned} \tag{3.10}$$

Now the short rate $r_t = r(t,0)$ is, by Definition 4:

$$r(t,0) := \lim_{x\to 0} F(x,Z_t) = F(0,Z_t) = -D_x G(0,Z_t), \quad 0 \le t < \infty.$$

Furthermore, we have the definition of the process for the savings account:

$$B(t) := \exp\left(\int_0^t r(s,0)\,\mathrm{d}s\right), \qquad 0 \le t < \infty.$$

Remark that

$$\begin{aligned} \frac{\mathrm{d}(1/B(t))}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\exp\left(-\int_0^t r(s,0)\,\mathrm{d}s\right) \\ &= \exp\left(-\int_0^t r(s,0)\,\mathrm{d}s\right)\frac{\mathrm{d}}{\mathrm{d}t}\left(-\int_0^t r(s,0)\,\mathrm{d}s\right) = -\frac{r(t,0)}{B(t)}. \end{aligned}$$

Using this we derive the following expression for $\frac{1}{B(t)}$:

$$\begin{aligned} \frac{1}{B(t)} &= \frac{1}{B(t)} - 1 + 1 = \frac{1}{B(t)} - \frac{1}{B(0)} + 1 = \int_0^t \frac{\mathrm{d}(1/B(s))}{\mathrm{d}s}\,\mathrm{d}s + 1 \\ &= -\int_0^t \frac{r(s,0)}{B(s)}\,\mathrm{d}s + 1 = \int_0^t \frac{1}{B(s)} D_x G(0,Z_s)\,\mathrm{d}s + 1. \end{aligned}$$

Now we are able to investigate $\mathrm{d}\Big(\frac{P(t,T)}{B(t)}\Big)$:

$$\begin{aligned}
\mathrm{d}\left(\frac{P(t,T)}{B(t)}\right) &= \frac{1}{B(t)}\,\mathrm{d}P(t,T) + P(t,T)\,\mathrm{d}\left(\frac{1}{B(t)}\right) + \mathrm{d}\left(\frac{1}{B(t)}\right)\mathrm{d}P(t,T)\\
&= \frac{1}{B(t)}\Big(\big(A_t G(T-t,Z_t) - D_x G(T-t,Z_t)\big)\,\mathrm{d}t + \nabla_z G(T-t,Z_t)^* \sigma_t\,\mathrm{d}W_t\Big)\\
&\quad + G(T-t,Z_t)\frac{1}{B(t)} D_x G(0,Z_t)\,\mathrm{d}t + \frac{1}{B(t)} D_x G(0,Z_t)\,\mathrm{d}t\\
&\quad \Big(\big(A_t G(T-t,Z_t) - D_x G(T-t,Z_t)\big)\,\mathrm{d}t + \nabla_z G(T-t,Z_t)^* \sigma_t\,\mathrm{d}W_t\Big).
\end{aligned} \tag{3.11}$$

Using the fact that $\mathrm{d}t\,\mathrm{d}t = 0$ and $\mathrm{d}t\,\mathrm{d}W_t = 0$, we see that the last term equals zero. Define

$$H(t,T) := \frac{1}{B(t)}\big(A_t G(T-t,Z_t) - D_x G(T-t,Z_t) + D_x G(0,Z_t) G(T-t,Z_t)\big)$$

and

$$M(t,T) := \int_0^t \frac{1}{B(s)} \nabla_z G(T-s,Z_s)^* \sigma_s\,\mathrm{d}W_s,$$

which is is a local $\mathbb{P}$-martingale because it is an integral with respect to a (local) $\mathbb{P}$-martingale. Using the definition of $H(t,T)$ we can rewrite (3.11) as

$$\mathrm{d}\left(\frac{P(t,T)}{B(t)}\right) = H(t,T)\,\mathrm{d}t + \frac{1}{B(t)} \nabla_z G(T-t,Z_t)^* \sigma_t\,\mathrm{d}W_t.$$

Using this we arrive at the following identity:

$$\begin{aligned}
\frac{P(t,T)}{B(t)} - P(0,T) &= \int_0^t \mathrm{d}\left(\frac{P(s,T)}{B(s)}\right)\\
&= \int_0^t H(s,T)\,\mathrm{d}s + \int_0^t \frac{1}{B(s)} \nabla_z G(T-s,Z_s)^* \sigma_s\,\mathrm{d}W_s\\
&= \int_0^t H(s,T)\,\mathrm{d}s + M(t,T).
\end{aligned} \tag{3.12}$$

Let's suppose Z is consistent with the Nelson-Siegel family, i.e. $\Big(\frac{P(t,T)}{B(t)}\Big)_{0\le t\le T}$ is a $\mathbb{P}$-martingale, for all $T < \infty$. Then we know that $\int_0^t H(s,T)\,\mathrm{d}s$ is a local martingale (because it is the difference of two local martingales) which is continuous (because $H(s,T)$ is continuous) and of bounded variation. Therefore we know for $T < \infty$

$$\int_0^t H(s,T)\,\mathrm{d}s = 0, \quad \forall t \in [0,T], \quad \mathbb{P} - a.s.. \tag{3.13}$$

Because $H(t,T)$ depends on the stochastic process Z_t we write $H(t,T)(\omega)$ if we want to line out the dependence of $H(t,T)$ via $Z_t(\omega)$ on $\omega \in \Omega$.

Claim 3.2. *Equation* (3.13) *yields* $H(t,T)(\omega) = 0$ *for* $(t,\omega) \in [0,T) \times \Omega$, $\mathrm{d}t \otimes \mathrm{d}\mathbb{P} - a.s.$.

Proof. Denote by λ the lebesgue measure on $[0,T)$ and by by N the collection of points $(t,\omega) \in [0,T) \times \Omega$ where $H(t,T)(\omega) > 0$: $N := \{(t,\omega) \in [0,T) \times \Omega | H(t,T)(\omega) > 0\}$. We will show that the measure of N equals zero: $(\lambda \times \mathbb{P})(N) = 0$.

Let's have a look at

$$\int_N H(t,T)(\omega)\,\mathrm{d}t \otimes \mathrm{d}\mathbb{P}.$$

Because $H(s,T)(\omega)$ is positive on N, we can use Tonelli:

$$\int_N H(t,T)(\omega)\,\mathrm{d}t \otimes \mathrm{d}\mathbb{P} = \int_\Omega \left(\int_{N_\omega} H(t,T)(\omega)\,\mathrm{d}t \right) \mathrm{d}\mathbb{P},$$

where $N_\omega = \{t | (t,\omega) \in N\}$. Because by assumption $\int_0^t H(s,T)\,\mathrm{d}s = 0, \quad \forall t \in [0,T], \quad \mathbb{P}-a.s.$, we know $\int_{N_\omega} H(t,T)(\omega)\,\mathrm{d}t \overset{\mathbb{P}-a.s.}{=} 0$, hence $\int_N H(t,T)(\omega)\,\mathrm{d}t \otimes \mathrm{d}\mathbb{P} = 0$. Because $H(t,T)(\omega) > 0$ on N, it must hold that N has 0 $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-measure. To prove that also $N_- := \{(t,\omega) \in [0,T) \times \Omega | H(t,T)(\omega) < 0\}$ has measure 0, we follow the same argumentation applied to $-H(t,T)(\omega)$. We conclude $H(t,T)(\omega) = 0$ for $(t,\omega) \in [0,T) \times \Omega$, $\mathrm{d}t \otimes \mathrm{d}\mathbb{P} - a.s.$. □

Claim 3.2 holds for every $T < \infty$. Because $H(t,T)$ is continuous in T, we know that

$$H(t,t+x)(\omega) = 0, \quad \forall x \geq 0, \quad \text{for} \quad \mathrm{d}t \otimes \mathrm{d}\mathbb{P}\text{-}a.e.\ (t,\omega). \tag{3.14}$$

Because $B(t) > 0$ for all t, (3.14) yields

$$AG(x,Z) - D_x G(x,Z) + D_x G(0,Z) G(x,Z) = 0, \quad \forall x \geq 0, \quad \mathrm{d}t \otimes \mathrm{d}\mathbb{P}\text{-}a.s.. \tag{3.15}$$

Using the definition of the bond price $G(x,z), z \in \mathcal{Z}$, equation (3.2), we see:

$$\frac{\mathrm{d}G(x,z)}{\mathrm{d}z^i} = -\int_0^x \frac{\mathrm{d}}{\mathrm{d}z^i} F(\eta,z)\,\mathrm{d}\eta G(x,z),$$

$$\frac{\partial^2 G(x,z)}{\partial z^j \partial z^i} = \left(\int_0^x \frac{\partial}{\partial z^i} F(\eta,z)\,\mathrm{d}\eta \int_0^x \frac{\partial}{\partial z^j} F(\eta,z)\,\mathrm{d}\eta - \int_0^x \frac{\partial^2}{\partial z^j \partial z^i} \right) G(x,z)$$

and

$$D_x G(x,z) = -F(x,z) G(x,z).$$

Details can be found in Appendix A.1, equations (A.1), (A.2) and (A.3).

Equation (3.15) can now be written as

$$
\begin{aligned}
0 &= AG(x,Z) - D_x G(x,Z) + D_x G(0,Z)G(x,Z) \\
&= b\cdot\nabla_z G(x,Z) + \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}\frac{\partial^2 G(x,Z)}{\partial z^i \partial z^j} - D_x G(x,Z) + D_x G(0,Z)G(x,Z) \\
&= -\int_0^x b\cdot\nabla_z F(\eta,Z)\,\mathrm{d}\eta G(x,Z) \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}\left(\int_0^x \frac{\partial}{\partial z^i}F(\eta,Z)\,\mathrm{d}\eta \int_0^x \frac{\partial}{\partial z^j}F(\eta,Z)\,\mathrm{d}\eta - \int_0^x \frac{\partial^2}{\partial z^j \partial z^i}F(\eta,Z)\,\mathrm{d}\eta\right) \\
&\quad G(x,Z) + F(x,Z)G(x,Z) - F(0,Z)G(x,Z) \\
&= -\int_0^x AF(\eta,Z)\,\mathrm{d}\eta G(x,Z) \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}\left(\int_0^x \frac{\partial}{\partial z^i}F(\eta,Z)\,\mathrm{d}\eta \int_0^x \frac{\partial}{\partial z^j}F(\eta,Z)\,\mathrm{d}\eta\right) G(x,Z) \\
&\quad + F(x,Z)G(x,Z) - F(0,Z)G(x,Z), \quad \forall x \geq 0, \quad \mathrm{d}t\otimes\mathrm{d}\mathbb{P}\text{-}a.s..
\end{aligned}
$$

Dividing by $-G(x,Z)$ gives:

$$
\begin{aligned}
&\int_0^x AF(\eta,z)\,\mathrm{d}\eta - \frac{1}{2}\sum_{i,j=1}^{4} a^{ij}\left(\int_0^x \frac{\partial}{\partial z^i}F(\eta,Z)\,\mathrm{d}\eta \int_0^x \frac{\partial}{\partial z^j}F(\eta,Z)\,\mathrm{d}\eta\right) \\
&\qquad\qquad -F(x,Z) + F(0,Z) = 0, \quad \forall x\geq 0, \quad \mathrm{d}t\otimes\mathrm{d}\mathbb{P}\text{-}a.s..
\end{aligned}
$$

Differentiating this to x gives

$$
\begin{aligned}
&-\frac{1}{2}\sum_{i,j=1}^{4} a^{ij}\left(\frac{\partial}{\partial z^i}F(x,Z)\int_0^x \frac{\partial}{\partial z^j}F(\eta,Z)\,\mathrm{d}\eta + \int_0^x \frac{\partial}{\partial z^i}F(\eta,Z)\,\mathrm{d}\eta\frac{\partial}{\partial z^j}F(x,Z)\right) \\
&+ AF(\eta,Z)\,\mathrm{d}\eta - D_x F(x,Z) = 0, \quad \forall x\geq 0, \quad \mathrm{d}t\otimes\mathrm{d}\mathbb{P}\text{-}a.s., \qquad (3.16)
\end{aligned}
$$

which can be rewritten to the expression of Theorem 3.1. □

This theorem was derived using an arbitrary forward rate curve $F(x,z)$ depending on a process Z. We are going to apply this theorem to the Nelson-Siegel forward rate. The following corollary gives the explicit condition given by Theorem 3.1 in the Nelson-Siegel forward curve case.

Corollary 3.3 (Nelson-Siegel). *Suppose $Z = (Z_t)_{0\leq t<\infty}$ follows the Itô process of Definition 3.4 with values in $\mathcal{Z}$. Then Z is consistent with the Nelson-Siegel family only if for $\mathrm{d}t\otimes\mathrm{d}\mathbb{P}$-a.e. (t,ω) in $[0,\infty)\times\Omega$,*

$$0 = p_0(x) + p_1(x)e^{-z_4 x} + p_2(x)e^{-2z_4 x} \qquad (3.17)$$

for all $x \geq 0$, *where* $p_0(x), p_1(x)$ *and* $p_2(x)$ *are polynomials in* x *with coefficients containing* $b^i := b_t^i(\omega), a^{ij} := a_t^{ij}(\omega)$ *and* $z^i := Z_t^i(\omega)$ *for* $1 \leq i, j \leq 4$ *which are given by*

$$p_0(x) = -xa^{11} + b^1 - \frac{1}{z_4}(a^{12} + a^{13}) + \frac{z_2 + z_3}{z_4^2}a^{14}, \tag{3.18}$$

$$\begin{aligned} p_1(x) =& (z_2 - z_3)z_4 + z_3 z_4^2 x + b^2 + z_4 x b^3 + \left[(z_3 - z_2)\, x - z_3 z_4 x^2\right] b^4 \\ &+ \left[\frac{1}{z_4} - x\right] a^{12} + \left[\frac{1}{z_4} + x - z_4 x^2\right] a^{13} \\ &+ \left[-\frac{z_2 + z_3}{z_4^2} - \frac{z_2 + z_3}{z_4}x + (z_2 - 2z_3)\, x^2 + z_3 z_4 x^3\right] a^{14} - \frac{1}{z_4}a^{22} \\ &- \left[\frac{1}{z_4} + x\right] a^{23} + \left[\frac{z_2 + z_3}{z_4^2} + \left(\frac{z_2 - z_3}{z_4} - 1\right) x + z_3 x^2\right] a^{24} - xa^{33} \\ &+ \left[\left(1 + 2\frac{z_2}{z_4}\right) x + (z_3 - z_4)\, x^2\right] a^{34} + \Bigg[\frac{(z_3 - z_2)(z_3 + z_2)}{z_4^2}x \\ &+ \left(-\frac{(z_2 + z_3)\, z_3}{z_4} + \frac{z_2}{2} - z_3\right) x^2 + \frac{z_3 z_4}{2}x^3\Bigg] a^{44}, \end{aligned} \tag{3.19}$$

$$\begin{aligned} p_2(x) =& \frac{1}{z_4}\Bigg(a^{22} + [1 + 2z_4 x]\, a^{23} + \left[-\frac{z_2 + z_3}{z_4} - 2z_2 x - 2z_3 z_4 x^2\right] a^{24} \\ &+ \left[z_4 x + z_4^2 x^2\right] a^{33} + \left[-2z_2 x - 2z_2 z_4 x^2 - z_3 z_4 x^2 - 2z_3 z_4^2 x^3\right] a^{34} \\ &+ \left[\frac{z_2^2 - z_3^2}{z_4}x + \left(z_2^2 + z_2 z_3\right) x^2 + 2z_2 z_3 z_4 x^3 + z_3^2 z_4^2 x^4\right] a^{44}\Bigg). \end{aligned} \tag{3.20}$$

Proof. The proof is nothing more than expanding what the condition of Theorem 3.1 means when we take for F the Nelson-Siegel family. The Nelson-Siegel forward rates $F(x, z)$ are given by equation (3.1):

$$F(x, z) = z_1 + z_2 e^{-z_4 x} + z_3 z_4 x e^{-z_4 x}.$$

The (partial) derivatives of F are written out in Appendix A.2. Using that $a = a^*$ and combining all the terms, one shows the condition in Corollary 3.3. □

Equation (3.17) can only be satisfied for all $x \geq 0$ if each of the polynomials $p_0(x)$, $p_1(x)$ and $p_2(x)$ are equal to 0 for all $x \geq 0$. This is proved in Appendix C. This gives conditions on the b^i and a^{ij} and hence it determines the dynamics of Z. Before we state a new theorem, we prove the following lemmas which we will use in the proof:

Lemma 3.4. *Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = ax + b$, $a \in \mathbb{R}$, $b \in \mathbb{R}$. The solution for $f(x) = 0$ for all $x \geq 0$ is given by $a = 0$, $b = 0$.*

Proof. $b = f(0) = 0$ gives $b = 0$ and then $a + b = a = f(1) = 0$ gives $a = 0$. □

This lemma is a special case of the following lemma:

Lemma 3.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be given by a polynomial of degree $p > 0$: $f(x) = \sum_{k=0}^{p} c_k x^k$ with $c_k \in \mathbb{R}$. If $f(x) = 0$ for all $x \geq 0$, we know $c_k = 0$ for $0 \leq k \leq p$.*

Proof. First note that if $f(x) = 0$ for all $x \geq 0$, also $f^{(q)}(x) := \frac{\mathrm{d}^q}{\mathrm{d}x^q} f(x) = 0$ for $q \geq 0$ and for all $x \geq 0$. Secondly, $f^{(q)}(x) = \sum_{k=q}^{p} c_k \frac{\mathrm{d}^q}{\mathrm{d}x^q}(x^k) = \sum_{k=q}^{p} c_k \frac{k!}{(k-q)!} x^{k-q}$, hence $f^{(q)}(0) = q! c_q$, for $q \geq 0$. Because $f^{(q)}(x) = 0$ for $q \geq 0$ and for all $x \geq 0$, we know $q! c_q = 0$ for $q \geq 0$, hence $c_k = 0$ for $0 \leq k \leq p$. □

Lemma 3.6. *Let M be an $n \times m$ matrix with $M^{ij} \in \mathbb{R}$. Define $B = MM^*$. Suppose $B^{ll} = 0$ for some $1 \leq l \leq n$. Then $B^{lj} = B^{jl} = 0$ for all $1 \leq j \leq n$.*

Proof. Remark that B^{ij} can be written as $B^{ij} = \sum_{k=1}^{m} M^{ik} M^{jk}$. Now $B^{ll} = 0$ yields $0 = \sum_{k=1}^{m} M^{lk} M^{lk} = \sum_{k=1}^{m} (M^{lk})^2$. Because $M^{ij} \in \mathbb{R}$, we know that $(M^{ij})^2 \geq 0$ for all $1 \leq i \leq n, 1 \leq j \leq m$. From this it follows that $M^{lk} = 0$ for $1 \leq k \leq m$. Obviously then $B^{lj} = \sum_{k=1}^{m} M^{lk} M^{jk} = \sum_{k=1}^{m} 0 M^{jk} = 0$. Because $B^* = (MM^*)^* = M^{**} M^* = MM^* = B$, we know $B^{ij} = B^{ji}$, hence $B^{lj} = B^{jl} = 0$ for all $1 \leq j \leq n$. □

This is it for the preparation for the following theorem, which is a consequence of Corollary 3.3.

Theorem 3.7. *Suppose $Z = (Z_t)_{0 \leq t < \infty}$ follows the Itô process of Definition 3.4 with values in $\mathcal{Z}$. Let Z be consistent with the Nelson-Siegel family. Then Z_t is of the form*

$$
\begin{aligned}
Z_t^1 &= Z_0^1, \\
Z_t^2 &= Z_0^2 e^{-Z_0^4 t} + Z_0^3 Z_0^4 t e^{-Z_0^4 t}, \\
Z_t^3 &= Z_0^3 e^{-Z_0^4 t}, \\
Z_t^4 &= Z_0^4 + \int_0^{t \wedge \tau} c_s^1 \mathbf{1}_{\{Z_0^2 = Z_0^3 = 0\}} \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^{t \wedge \tau} \sigma_s^{4j} \mathbf{1}_{\{Z_0^2 = Z_0^3 = 0\}} \, \mathrm{d}W_s^j,
\end{aligned}
$$

with $c^1 \in \mathbb{R}$ and the stopping time $\tau := \inf\{s > 0 | Z_s^2 = Z_s^3 = 0\}$.

Proof. Z is consistent if equation (3.17) from Corollary 3.3 holds. As proven in Appendix C, equation (3.17) can only hold if $p_0(x)$, $p_1(x)$ and $p_2(x)$ equal zero for all x. Let's start with the first polynomial, $p_0(x)$.

The polynomial $p_0(x)$ is of the form of f in Lemma 3.4 (and 3.5), with $a = -a^{11}$ and $b = b^1 - \frac{1}{z_4}(a^{12} + a^{13}) + \frac{z_2+z_3}{z_4^2}a^{14}$. From the lemma(s) we know $a^{11} = 0$.

Because we have $a = \sigma\sigma^*$, $\sigma \in \mathbb{R}^{4\times d}$ and $a^{11} = 0$, we know by Lemma 3.6, $a^{1j} = a^{j1} = 0$ for $1 \leq j \leq 4$. What remains of the polynomial $p_0(x)$ is $p_0(x) = b^1$. Obviously $b^1 = 0$. These observations simplify $p_1(x)$ somewhat.

Remark that the degree of $p_2(x)$ depends on whether $z_3 = 0$ or not. Suppose $z_3 \neq 0$ and $z_2 \neq 0$. Then the degree in $p_2(x)$ is four and the coefficient of x^4 is given by $z_3^2 z_4 a^{44}$. Because $p_2(x) = 0$ for all $x \geq 0$, we know $z_3^2 z_4 a^{44} = 0$ and hence $a^{44} = 0$. By using Lemma 3.6 again, we know $a^{4j} = a^{j4} = 0$ for $1 \leq j \leq 4$.

The polynomial reduces to

$$p_2(x) = \frac{1}{z_4}\left(a^{22} + [1 + 2z_4x]\, a^{23} + \left[z_4x + z_4^2x^2\right] a^{33}\right).$$

The second order coefficient is z_4a^{33}, hence for the polynomial to be zero, $a^{33} = 0$. The order reduces to 1 with coefficient $2a^{23}$, hence also a^{23} and a^{32} equal zero. Now only $\frac{a^{22}}{z_4}$ is left, which means that $a^{22} = 0$. This implies $a = 0$ and hence $\sigma^{ij} = 0$. The polynomial $p_1(x)$ reduces to:

$$p_1(x) = (z_2 - z_3)z_4 + z_3z_4^2x + b^2 + z_4xb^3 + \left[(z_3 - z_2)\,x - z_3z_4x^2\right] b^4. \tag{3.21}$$

A similar argument as above gives that $b^4 = 0$ and we are left with

$$b^3 = -z_3z_4, \tag{3.22}$$

$$b^2 = (z_3 - z_2)z_4. \tag{3.23}$$

Until now we only assumed $z_3 \neq 0$ and $z_2 \neq 0$. It is worth considering what happens when $z_3 \neq 0$ and $z_2 = 0$. The following lemma will make things easier:

Lemma 3.8. *For $1 \leq i \leq 4$, it holds that $a^{ii}\mathbf{1}_{\{Z^i=0\}} = b^i\mathbf{1}_{\{Z^i=0\}} = 0$, $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-a.s..*

Proof. The Occupation times formula (Corollary 1.6, Chapter VI of Revuz and Yor [19]) gives the following: There is a $\mathbb{P}$-negligible set outside of which

$$\int_0^t \Phi(X_s)\,\mathrm{d}[X, X]_s = \int_{-\infty}^{\infty} \Phi(a)L_t^a\,\mathrm{d}a,$$

for every t and every positive Borel function Φ and L_t^a a local time. Take $X = Z^i$ for $i = 1, \ldots, 4$, then $[X, X]_s = [Z^i, Z^i]_s = \int_0^s a_t^{ii}\, \mathrm{d}t$ and take the positive Borel function $\Phi(z) = \mathbf{1}_0(z)$. Then by the formula

$$\int_0^t \mathbf{1}_0(Z_s^i) a_s^{ii}\, \mathrm{d}s = \int_{-\infty}^{\infty} \mathbf{1}_0(a) L_t^a\, \mathrm{d}a = \int_{\{0\}} L_t^a\, \mathrm{d}a = 0,$$

because the lebesgue measure of 0 is zero. Because this holds for every t, we know $a^{ii}\mathbf{1}_{\{Z^i=0\}} = 0, \mathrm{d}t \otimes \mathrm{d}\mathbb{P}\text{-}a.s.$, for $i = 1, \ldots, 4$.

Now let $\mathrm{d}Y_t = \mathbf{1}_{Z_t^i=0}\, \mathrm{d}Z_t^i$. Then $[Y, Y]_t = \int \mathbf{1}_{Z_s^i=0}[Z^i, Z^i]_s = \int \mathbf{1}_{Z_s^i=0} a_s^{ii}\, \mathrm{d}s = 0$. This implies that $Y_t \overset{a.s.}{=} Y_0$, or

$$\begin{aligned} 0 \overset{a.s.}{=} Y_t - Y_0 &= \int_0^t \mathbf{1}_{\{Z_s^i=0\}}\, \mathrm{d}Z_s \\ &= \int_0^t \mathbf{1}_{\{Z_s^i=0\}} b_s^i\, \mathrm{d}s + \sum_{j=1}^d \int_0^t \mathbf{1}_{\{Z_s^i=0\}} \sigma_s^{ij}\, \mathrm{d}W_s^j \\ &= \int_0^t \mathbf{1}_{\{Z_s^i=0\}} b_s^i\, \mathrm{d}s + 0. \end{aligned}$$

Again this holds for every t, so we know $b^i \mathbf{1}_{\{Z^i=0\}} = 0, \mathrm{d}t \otimes \mathrm{d}\mathbb{P}\text{-}a.s.$, for $i = 1, \ldots, 4$. □

The lemma tells us that when $z_2 = 0$, also $b_2 = 0$. The condition (3.23) gives $z_3 = 0$, which contradicts the assumption $z_3 \neq 0$ in the derivations of (3.23). Hence this cannot occur.

Let's consider the case that $z_3 = 0$. As above we have that $a^{11} = 0$. Now given that $z_3 = 0$, we immediately know b^3 and a^{33} are zero and hence (using Lemma 3.6) $a^{3j} = a^{j3} = 0$ for $1 \leq j \leq 4$. This reduces $p_2(x)$ to:

$$p_2(x) = \frac{1}{z_4} a^{22} - \left[\frac{z_2}{z_4^2} + 2\frac{z_2}{z_4} x \right] a^{24} + \left[\frac{z_2^2}{z_4^2} x + \frac{z_2^2}{z_4} x^2 \right] a^{44}.$$

The highest order coefficient to be zero holds that $a^{44} = 0$ (if $z_2 \neq 0$) and hence $a^{4j} = a^{j4} = 0$ for $1 \leq j \leq 4$. This gives $p_2(x) = \frac{1}{z_4} a^{22}$ hence $a^{22} = 0$. The matrix a^{ij} is zero again. We know immediately that $p_1(x) = z_2 z_4 + b^2 - z_2 x b^4$, hence $b^4 = 0$ and $b^2 = -z_2 z_4$.

If z_3 and z_2 equal zero, we know that $a^{ij} = 0$ and $b^k = 0$ for all $1 \leq i, j \leq 4$, except for $i = j = 4$, and $1 \leq k \leq 3$. In this case the polynomials are equal to zero, $p_1(x) = p_2(x) = 0$, regardless the choice for b^4 and a^{44}.

Summarizing:

$$\begin{aligned}
b^1 &= 0, \\
b^2 &= (z_3 - z_2) z_4, \\
b^3 &= -z_3 z_4, \\
b^4 &= c^1 \mathbf{1}_{\{z_2 = z_3 = 0\}}, \\
a^{ij} &= 0, \quad \text{for } (i,j) \neq (4,4), \\
a^{44} &= c^2 \mathbf{1}_{\{z_2 = z_3 = 0\}},
\end{aligned}$$

where $c^1 \in \mathbb{R}$ and $c^2 \in \mathbb{R}_{\geq 0}$ are arbitrary numbers.

Because Corollary 3.3 holds $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-*a.e.* (t, ω), the process Z (3.6) is now, up to indistinguishability, given by

$$\begin{pmatrix} Z_t^1 \\ Z_t^2 \\ Z_t^3 \\ Z_t^4 \end{pmatrix} = \begin{pmatrix} Z_0^1 \\ Z_0^2 \\ Z_0^3 \\ Z_0^4 \end{pmatrix} + \int_0^t \begin{pmatrix} 0 \\ \left(Z_s^3 - Z_s^2\right) Z_s^4 \\ -Z_s^3 Z_s^4 \\ c_s^1 \mathbf{1}_{\{Z_s^2 = Z_s^3 = 0\}} \end{pmatrix} \mathrm{d}s$$
$$+ \int_0^t \begin{pmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & 0 \\ \sigma_s^{41} \mathbf{1}_{\{Z_s^2 = Z_s^3 = 0\}} & \dots & \sigma_s^{4d} \mathbf{1}_{\{Z_s^2 = Z_s^3 = 0\}} \end{pmatrix} \begin{pmatrix} \mathrm{d}W_s^1 \\ \vdots \\ \mathrm{d}W_s^d \end{pmatrix} \tag{3.24}$$

or

$$\begin{aligned}
Z_t^1 &= Z_0^1, \\
Z_t^2 &= Z_0^2 + \int_0^t \left(Z_s^3 - Z_s^2\right) Z_s^4 \, \mathrm{d}s, \\
Z_t^3 &= Z_0^3 - \int_0^t Z_s^3 Z_s^4 \, \mathrm{d}s, \\
Z_t^4 &= Z_0^4 + \int_0^t c_s^1 \mathbf{1}_{\{Z_s^2 = Z_s^3 = 0\}} \, \mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_s^{4j} \mathbf{1}_{\{Z_s^2 = Z_s^3 = 0\}} \, \mathrm{d}W_s^j.
\end{aligned}$$

On $\Omega_0 := \{Z_0^2 = Z_0^3 = 0\}$, Z_t^2 and Z_t^3 remain zero (a.s.). It remains to show it on $\Omega_1 := \Omega \backslash \Omega_0$. Introduce the stopping time $\tau := \inf\{s > 0 | Z_s^2 = Z_s^3 = 0\}$. Obviously $\Omega_0 = \{\tau = 0\}$. As soon as Z_s^2 and Z_s^3 are zero, they remain zero. Hence $\mathbf{1}_{\{Z_s^2 = Z_s^3 = 0\}} = \mathbf{1}_{[\tau, \infty)}$. Define the stopped process $Y_t = Z_{t \wedge \tau}$, then:

$$Y_t^1 = Z_0^1,$$
$$Y_t^2 = Z_0^2 + \int_0^{t\wedge\tau} \left(Y_s^3 - Y_s^2\right) Y_s^4 \, \mathrm{d}s,$$
$$Y_t^3 = Z_0^3 - \int_0^{t\wedge\tau} Y_s^3 Y_s^4 \, \mathrm{d}s,$$
$$Y_t^4 = Z_0^4 + \int_0^{t\wedge\tau} c_s^1 \mathbf{1}_{\{Y_s^2=Y_s^3=0\}} \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^{t\wedge\tau} \sigma_s^{4j} \mathbf{1}_{\{Y_s^2=Y_s^3=0\}} \, \mathrm{d}W_s^j. \quad (3.25)$$

Remark that for $t < \tau$, $c_t^1 \mathbf{1}_{\{Y_s^2=Y_s^3=0\}}$ and $\sigma_s^{4j} \mathbf{1}_{\{Y_s^2=Y_s^3=0\}}$ are zero. Furthermore

$$\int_0^{t\wedge\tau} c_s^1 \mathbf{1}_{\{Y_s^2=Y_s^3=0\}} \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^{t\wedge\tau} \sigma_s^{4j} \mathbf{1}_{\{Y_s^2=Y_s^3=0\}} \, \mathrm{d}W_s^j$$
$$= \int_0^{t} c_s^1 \mathbf{1}_{[\tau,\infty]} \mathbf{1}_{[0,\tau]} \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^{t} \sigma_s^{4j} \mathbf{1}_{[\tau,\infty]} \mathbf{1}_{[0,\tau]} \, \mathrm{d}W_s^j$$
$$= \int_0^{t} c_s^1 \mathbf{1}_{[\tau]} \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^{t} \sigma_s^{4j} \mathbf{1}_{[\tau]} \, \mathrm{d}W_s^j = 0.$$

Hence Y_t^4 becomes $Y_t^4 = Z_0^4$, so Y_t of (3.25) becomes

$$Y_t^1 = Z_0^1,$$
$$Y_t^2 = Z_0^2 + \int_0^{t\wedge\tau} \left(Y_s^3 - Y_s^2\right) Z_0^4 \, \mathrm{d}s,$$
$$Y_t^3 = Z_0^3 - \int_0^{t\wedge\tau} Y_s^3 Z_0^4 \, \mathrm{d}s,$$
$$Y_t^4 = Z_0^4.$$

The solution of the equation for Y_t^3 is given by

$$Y_t^3 = Z_0^3 e^{-Z_0^4 (t\wedge\tau)},$$

which can easily been checked:

$$Z_0^3 - \int_0^{t\wedge\tau} Y_s^3 Z_0^4 \, \mathrm{d}s = Z_0^3 - \int_0^{t\wedge\tau} Z_0^3 e^{-Z_0^4 (s\wedge\tau)} Z_0^4 \, \mathrm{d}s$$
$$= Z_0^3 - Z_0^3 \int_0^{t\wedge\tau} e^{-Z_0^4 s} Z_0^4 \, \mathrm{d}s$$
$$= Z_0^3 - Z_0^3 \left[-e^{-Z_0^4 s}\right]_{s=0}^{s=t\wedge\tau}$$
$$= Z_0^3 e^{-Z_0^4 (t\wedge\tau)} = Y_t^3.$$

Furthermore it is easy to check that

$$Y_t^2 = Z_0^2 e^{-Z_0^4(t\wedge\tau)} + Z_0^3 Z_0^4 (t\wedge\tau) e^{-Z_0^4(t\wedge\tau)} :$$

$$\begin{aligned}
Z_0^2 + &\int_0^{t\wedge\tau} \left(Y_s^3 - Y_s^2\right) Z_0^4 \,\mathrm{d}s \\
&= Z_0^2 + \int_0^{t\wedge\tau} \left(Z_0^3 e^{-Z_0^4(s\wedge\tau)} - Z_0^2 e^{-Z_0^4(s\wedge\tau)} - Z_0^3 Z_0^4 (s\wedge\tau) e^{-Z_0^4(s\wedge\tau)}\right) Z_0^4 \,\mathrm{d}s \\
&= Z_0^2 + \left(Z_0^3 - Z_0^2\right) Z_0^4 \int_0^{t\wedge\tau} e^{-Z_0^4(s\wedge\tau)} \,\mathrm{d}s \\
&\qquad - Z_0^3 (Z_0^4)^2 \int_0^{t\wedge\tau} (s\wedge\tau) e^{-Z_0^4(s\wedge\tau)} \,\mathrm{d}s \\
&= Z_0^2 + \left(Z_0^3 - Z_0^2\right)\left(1 - e^{-Z_0^4(t\wedge\tau)}\right) \\
&\qquad - Z_0^3 \left(-Z_0^4 (t\wedge\tau) e^{-Z_0^4(t\wedge\tau)} + 1 - e^{-Z_0^4(t\wedge\tau)}\right) \\
&= Z_0^2 e^{-Z_0^4(t\wedge\tau)} + Z_0^3 Z_0^4 (t\wedge\tau) e^{-Z_0^4(t\wedge\tau)} = Y_t^2.
\end{aligned}$$

Putting everything together we have

$$\begin{aligned}
Y_t^1 &= Z_0^1, \\
Y_t^2 &= Z_0^2 e^{-Z_0^4(t\wedge\tau)} + Z_0^3 Z_0^4 (t\wedge\tau) e^{-Z_0^4(t\wedge\tau)}, \\
Y_t^3 &= Z_0^3 e^{-Z_0^4(t\wedge\tau)}, \\
Y_t^4 &= Z_0^4.
\end{aligned}$$

For $t \in [0,\tau]$, we have $Y_t = Z_t$. Remark that $Y_t > 0$ for all t, hence $\Omega_1 = \{\tau > 0\} = \{\tau = \infty\}$. Using this we get $Y_t = Z_{t\wedge\tau} = Z_t$ and

$$\begin{aligned}
Z_t^1 &= Z_0^1, \\
Z_t^2 &= Z_0^2 e^{-Z_0^4 t} + Z_0^3 Z_0^4 t e^{-Z_0^4 t}, \\
Z_t^3 &= Z_0^3 e^{-Z_0^4 t}, \\
Z_t^4 &= Z_0^4,
\end{aligned}$$

on Ω_1. Putting everything together we get

$$\begin{aligned}
Z_t^1 &= Z_0^1, \\
Z_t^2 &= Z_0^2 e^{-Z_0^4 t} + Z_0^3 Z_0^4 t e^{-Z_0^4 t}, \\
Z_t^3 &= Z_0^3 e^{-Z_0^4 t}, \\
Z_t^4 &= Z_0^4 + \int_0^{t\wedge\tau} c_s^1 \mathbf{1}_{\{Z_0^2 = Z_0^3 = 0\}} \,\mathrm{d}s + \sum_{j=1}^{d} \int_0^{t\wedge\tau} \sigma_s^{4j} \mathbf{1}_{\{Z_0^2 = Z_0^3 = 0\}} \,\mathrm{d}W_s^j,
\end{aligned}$$

which is the result of Theorem 3.7. □

We can of course substitute this expression for Z in the Nelson-Siegel forward curve to see what this really means:

Corollary 3.9. *The forward rates are non-random, they are $\mathcal{F}_0$-measurable.*

Proof. The Nelson-Siegel forward rate on Ω_0 is given by $F(x, Z_t) = Z_0^1$ (because $Z_t^2 = Z_t^3 = 0$), which is $\mathcal{F}_0$-measurable. The dynamics of Z_4^t have no influence. The Nelson-Siegel forward rate on Ω_1 is given by

$$\begin{aligned} F(x, Z_t) &= Z_t^1 + Z_t^2 e^{-Z_t^4 x} + Z_t^3 Z_t^4 x e^{-Z_t^4 x} \\ &= Z_0^1 + \left(Z_0^2 e^{-Z_0^4 t} + Z_0^3 Z_0^4 t e^{-Z_0^4 t} \right) e^{-Z_0^4 x} + Z_0^3 Z_t^4 e^{-Z_0^4 t} x e^{-Z_0^4 x} \\ &= Z_0^1 + Z_0^2 e^{-Z_0^4 (t+x)} + Z_0^3 Z_0^4 (t+x) e^{-Z_0^4 (t+x)} = F(t+x, Z_0), \end{aligned}$$

hence all the randomness remains $\mathcal{F}_0$-measurable. □

Corollary 3.9 tells us that the interest rate model is nonrandom, trivial. It only depends on the value of the process at time 0.

Chapter 4

Consistency: Jump process

In Chapter 3 we used stochastic processes without jumps, so called Itô processes. For this class of processes it is shown that the only Itô processes consistent with the Nelson-Siegel family provide a trivial interest rate model. In this chapter we will show there is an analogous result for stochastic processes involving jumps, from now on called *jump processes*. We will define these processes first.

4.1 Stochastic process with jumps

4.1.1 Poisson process

One of the most well known jump processes is the Poisson process. This section will give a short introduction. For more details, check for example ref. [20], Section 11.2.

There are different descriptions of the Poisson process possible. Hereby we use a definition that follows the construction in ref. [20] by Shreve.

Definition 4.1 (Poisson process). *Given a sequence of independent identically distributed exponential random variables $\tau_1, \tau_2, \dots$ with mean $\frac{1}{\lambda}$, define a* Poisson process *$N(t)$ with intensity $\lambda > 0$ as*

$$N(t) = \max\{n : \sum_{i=1}^{n} \tau_i \leq t\}, \tag{4.1}$$

i.e. when we interpret τ_k as the time between the kth jump and the $k-1$th jump, $N(t)$ counts the number of jumps that occur at or before time t.

As a consequence

$$P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \dots$$

and

$$P(N(t) - N(s) = k) = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)}, \quad k = 0, 1, \dots, \quad t > s.$$

It is easy to check that

$$E[N(t) - N(s)] = \sum_{k=0}^{\infty} k \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)}$$
$$= \lambda(t-s) \sum_{k=1}^{\infty} \frac{\lambda^{k-1}(t-s)^{k-1}}{(k-1)!} e^{-\lambda(t-s)} = \lambda(t-s).$$

Furthermore we know that for $t > s$, $N(t) - N(s)$ is independent of $\mathcal{F}_s$, where $\mathcal{F}_t$ is the σ-algebra containing all the information of $N(s)$ for $0 \leq s \leq t$.

4.1.2 Compensated Poisson process

Whereas the Poisson process is a pure jump process (the process does not change, unless it jumps), there is an extension of this process which can be shown to be a martingale. We need this process to be able to say something about integrals with respect to the Poisson process later on.

Definition 4.2 (Compensated Poisson process)**.** *Let $N(t)$ be a Poisson process with intensity $\lambda > 0$. The* compensated Poisson process $M(t)$ *is defined as*

$$M(t) = N(t) - \lambda t. \tag{4.2}$$

For this process, we prove the following:

Lemma 4.1. *The compensated Poisson process $M(t)$ is a martingale.*

Proof. Let $0 \leq s < t$ be given. We have

$$\begin{aligned} E[M(t)|\mathcal{F}_s] &= E[M(t) - M(s)|\mathcal{F}_s] + E[M(s)|\mathcal{F}_s] \\ &= E[N(t) - N(s) - \lambda t + \lambda s|\mathcal{F}_s] + M(s) \\ &= E[N(t) - N(s)] - \lambda(t-s) + M(s) = M(s), \end{aligned}$$

because $M(s)$ is $\mathcal{F}_s$-measurable, $N(t) - N(s)$ is independent of $\mathcal{F}_s$ and the expected value of $N(t) - N(s)$ is $\lambda(t-s)$. □

4.1.3 Jump process

Next we will use the Poisson process to extend the state space process from an Itô process to a jump process where the jumps are given by a Poisson process. Recall the definition of the multidimensional Itô process: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $(\mathcal{F}_t)_{0 \leq t < \infty}$ a filtration, satisfying the usual conditions [21], and let $W = (W_t^1, \ldots, W_t^d)_{0 \leq t < \infty}$ be standard

d-dimensional $(\mathcal{F}_t)$-Brownian motion, $1 \leq d$. The multidimensional Itô process $Z = (Z_t)_{0 \leq t < \infty}$, $Z_t = (Z_t^1, \ldots, Z_t^4)$, is given by

$$Z_t^i = Z_0^i + \int_0^t b_s^i \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^t \sigma_s^{ij} \, \mathrm{d}W_s^j, \qquad i = 1, \ldots, 4, \tag{4.3}$$

where Z_0^i is nonrandom ($\mathcal{F}_0$-measurable) and with b and σ progressively measurable processes with values in $\mathbb{R}^4$, respectively $\mathbb{R}^{4\times d}$, such that $\int_0^t |b_s^i| \, \mathrm{d}s < \infty$ and $\int_0^t |\sigma_s^{ij}|^2 \, \mathrm{d}s < \infty$, $\mathbb{P}$-*a.s.*, for all $t > 0$. This part is the continuous part of the process. We add to this the jump part $J = (J^1, \ldots, J^4)$ given by

$$J_t^i = \sum_{j=1}^{m} \mathbf{\Sigma}^{ij} N_t^j, \qquad i = 1, \ldots, 4, \tag{4.4}$$

where $N_t = (N_t^1, \ldots, N_t^m)$ is an m-dimensional vector with independent Poisson processes with parameters $\lambda_1, \ldots, \lambda_m > 0$ and $\mathbf{\Sigma} \in \mathbb{R}^{4\times m}$. The general jump process we define in this thesis is defined as follows.

Definition 4.3. *The* jump process $Z = (Z_t)_{0 \leq t < \infty}$, $Z_t = (Z_t^1, \ldots, Z_t^4)$, *is given by*

$$Z_t^i = (Z^c)_t^i + J_t^i, \qquad i = 1, \ldots, 4, \tag{4.5}$$

where $(Z^c)_t^i$ *is the continuous part of* Z *given by* (3.6)*:*

$$(Z^c)_t^i = Z_0^i + \int_0^t b_s^i \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^t \sigma_s^{ij} \, \mathrm{d}W_s^j, \qquad i = 1, \ldots, 4,$$

and J_t^i *is the jump part of* Z *given by* (4.4)*.*

In what follows we will use a somewhat easier jump process, where $N_t = (N_t^1, N_t^2, N_t^3)$, $\mathbf{\Sigma} \in \mathbb{R}^{4\times 3}$, $\mathbf{\Sigma}^{ij} = 0$ for $i \neq j$, and $\mathbf{\Sigma}^{11} = \alpha^1$, $\mathbf{\Sigma}^{22} = \alpha^2$ and $\mathbf{\Sigma}^{33} = \alpha^3$. This gives the following process, which we call the *Independent jump process*:

Definition 4.4 (Independent jump process)**.** *The* Independent jump process $Z = (Z_t)_{0 \leq t < \infty}$, $Z_t = (Z_t^1, \ldots, Z_t^4)$, *is given by*

$$Z_t^i = Z_0^i + \int_0^t b_s^i \, \mathrm{d}s + \sum_{j=1}^{d} \int_0^t \sigma_s^{ij} \, \mathrm{d}W_s^j + \alpha^i N_t^i \mathbf{1}_{\{i \neq 4\}}, \qquad i = 1, \ldots, 4, \tag{4.6}$$

where N_t^i *is a Poisson process with parameter* $\lambda_i \in \mathbb{R}_+$, $\alpha^i \in \mathbb{R}$, W *is* d*-dimensional Brownian motion,* Z_0^i *is nonrandom (*$\mathcal{F}_0$*-measurable) and with* b *and* σ *progressively measurable processes with values in* $\mathbb{R}^4$*, respectively* $\mathbb{R}^{4\times d}$*, such that for* $i = 1, \ldots, 4$ *and* $j = 1, \ldots, d$*,*

$$\int_0^t |b_s^i| \, \mathrm{d}s < \infty \text{ and } \int_0^t |\sigma_s^{ij}|^2 \, \mathrm{d}s < \infty, \quad \mathbb{P}\text{-a.s.}, \quad \text{for all } t > 0.$$

4.1.4 Stochastic Calculus for Jump processes

To be able to do some stochastic calculus, we have to analyze what the quadratic (co)variation of those jump processes is. See Appendix B for more information about the quadratic (co)variation. From standard Itô calculus, we know that

$$[(Z^c)^i, (Z^c)^j]_t = \int_0^t \sum_{k=1}^d \sigma_s^{ik} \sigma_s^{jk} \, \mathrm{d}s$$

as we have used before. As shown in ref. [20], Theorem 11.4.7, the quadratic variation of two processes X_1 and X_2 with jumps is given by $[X_1, X_2]_t = [X_1^c, X_2^c]_t + [J_1, J_2]_t$ where X_i^c is the continuous part and J_i the pure jump part of process i and $[J_i, J_j]_t$ is given by $[J_i, J_j]_t = \sum_{0<s\leq t} \Delta J_i(s) \Delta J_j(s)$. The remark on page 482 of [20] tells that in differential form,

$$\mathrm{d}X_1(t)\, \mathrm{d}X_2(t) = \mathrm{d}X_1^c(t)\, \mathrm{d}X_2^c(t) + \mathrm{d}J_1(t)\, \mathrm{d}J_2(t)$$

and

$$\mathrm{d}X_1^c(t)\, \mathrm{d}J_2(t) = \mathrm{d}X_2^c(t)\, \mathrm{d}J_1(t) = 0.$$

The processes $N^i(t)$ are right continuous. Define by $f(t-)$ the limit $\lim_{s\uparrow t} f(s)$ from the left. Define for a process X_t the jump size $\Delta X(t) = X(t) - X(t-)$. For a right continuous process X, it holds that $\Delta X(t) = 0$ if there is no jump and $\Delta X(t) = J$ if there is a jump with size J. For the poisson process it holds that $\Delta N(t) = N(t) - N(t-) = 1$ if there is a jump and $\Delta N(t) = 0$ when there is no jump. Furthermore, the poisson process can only have a finite number of jumps in every time interval. Therefore we mean by $\sum_{0<s\leq t} \Delta N(s)$ the sum over all jump times s of the process between 0 and t. This of course equals $N(t)$ in this case: $N(t) = \sum_{0<s\leq t} \Delta N(s)$.

Suppose we have a continuous function f with continuous first and second order partial derivatives, depending on a 1-dimensional Itô-process $X(t)$ with jumps given by

$$X(t) = X^c(t) + J(t),$$
$$X^c(t) = X(0) + \int_0^t b_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s,$$

where W_t is brownian motion and $J(t)$ is a Poisson process. The function f becomes discontinuous because of the discontinuity of $X(t)$. However, when there is no jump (i.e. in between jumps) we have

$$\mathrm{d}f(X(s)) = f'(X(s))\, \mathrm{d}X^c(s) + \frac{1}{2} f''(X(s)) \sigma_s^2 \, \mathrm{d}s, \tag{4.7}$$

like we have seen before. When there is a jump, from $X(s-)$ to $X(s)$, the process f will jump from $f(X(s-))$ to $f(X(s))$. This leads to the following identity which can be found at page 484 of ref. [20]:

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s))\,\mathrm{d}X^c(s) + \frac{1}{2}\int_0^t f''(X(s))\sigma_s^2\,\mathrm{d}s + \sum_{0<s\le t}\big[f(X(s)) - f(X(s-))\big]\,. \tag{4.8}$$

Remark 4.2. *In some textbooks, for example in ref. [11], Theorem 7.3.1, the second integral on the righthand side is with respect to the jump process, not only to the continuous part of the jump process. To compensate for the changes at the jump times, they have to add another term, involving the derivative of the function and the jumps in X:* $\int_0^t f'(X(s))\,\mathrm{d}X^c(s) = \int_0^t f'(X(s))\,\mathrm{d}X(s) - \sum_{0\le s<t} f'(X(s))(X(s)-X(s-))$.

This identity can be extended to the case that we have a function of a multidimensional Itô process. Theorem 11.5.4 in ref. [20] gives the two-dimensional Itô-Doeblin formula for processes with jumps. This can be extended to the following identity for higher dimensional processes $X(t) = \big(X_1(t),\dots,X_d(t)\big)$ with jumps:

$$f(t,X(t)) = f(0,X(0)) + \int_0^t \frac{\partial}{\partial s} f(s,X(s))\,\mathrm{d}s + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(s,X(s))\,\mathrm{d}X_i^c(s) + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(s,X(s))\,\mathrm{d}X_i^c\,\mathrm{d}X_j^c + \sum_{0<s\le t}\big[f(s,X(s)) - f(s,X(s-))\big]\,. \tag{4.9}$$

The next step is to analyze $\sum_{0<s\le t}\big[f(s,X(s)) - f(s,X(s-))\big]$. In order to do this, we start with analyzing $\Delta X(t)$. Remark that again we can write $X^c(t) + X^J(t)$ with $X^c(t)$ the continuous part of $X(t)$ and $X^J(t)$ the pure jump part:

$$X^c(t) = \big(X_1^c(t),\dots,X_d^c(t)\big)\,,$$
$$X^J(t) = \big(J_1(t),\dots,J_d(t)\big)\,.$$

Now $\Delta X(t) = X^c(t) + X^J(t) - (X^c(t-) + X^J(t-)) = X^J(t) - X^J(t-) = \Delta X^J(t)$ by continuity of $X^c(t)$. Note that $\Delta X^J(t) = 0$ if there is no jump at time t.

If we take $X^J(t) = \big(N_1(t),\dots,N_d(t)\big)$, i.e. $X^J(t)$ is the vector containing d independent Poisson processes, a jump in $X^J(t)$ occurs when one of the Poisson processes jumps. The following proposition tells that two of those jumps never occur at the same time, almost surely.

Lemma 4.3. *Let $N_1(t)$ and $N_2(t)$ be two independent Poisson processes with intensities λ_1 and λ_2. Then $P\big(\Delta N_1(t) = 1, \Delta N_2(t) = 1\big) = 0$ for all t.*

Proof. Jumps of the Poisson processes occur when an arrival occurs. The arrival times are defined as $S_n^i := \sum_{k=1}^n \tau_k^i$ with τ_k^i exponential distributed independent random variables with parameter λ_i, for $i = 1, 2$. For the probability we have

$$\begin{aligned}
&P\left(\Delta N_1(t) = 1, \Delta N_2(t) = 1\right) \\
&\quad = \sum_{k=1}^{\infty} P\left(\Delta N_1(t) = 1, \Delta N_2(t) = 1 | N_1(t) = k\right) P(N_1(t) = k) \\
&\quad = \sum_{k,l=1}^{\infty} P\left(\Delta N_1(t) = 1, \Delta N_2(t) = 1 | N_1(t) = k, N_2(t) = l\right) \\
&\qquad\qquad\qquad\qquad P(N_1(t) = k)P(N_2(t) = l) \\
&\quad = \sum_{k,l=1}^{\infty} P\left(S_k^1 = t, S_l^2 = t\right) P(N_1(t) = k)P(N_2(t) = l). \qquad (4.10)
\end{aligned}$$

The arrival times S_n^i are Erlang(n, λ_i) distributed. Their distribution function is given by

$$f_{S_n^i}(x; n, \lambda_i) = \frac{\lambda_i x^{n-1} e^{-\lambda_i x}}{(k-1)!}.$$

The probability $P\left(S_k^1 = t, S_l^2 = t\right)$ can now be computed as

$$\begin{aligned}
P\left(S_k^1 = t, S_l^2 = t\right) &= \int_0^{\infty} P\left(S_k^1 = t, S_l^2 = t | S_l^2 = s\right) f_{S_l^2}(s; l, \lambda_2)\,\mathrm{d}s \\
&= \int_0^{\infty} P\left(S_k^1 = s\right) f_{S_l^2}(s; l, \lambda_2)\,\mathrm{d}s \\
&= \int_0^{\infty} \left[P\left(S_k^1 \leq s\right) - P\left(S_k^1 < s\right)\right] f_{S_l^2}(s; l, \lambda_2)\,\mathrm{d}s. \qquad (4.11)
\end{aligned}$$

Now

$$P\left(S_k^1 \leq s\right) = \int_0^s f_{S_k^1}(\eta; k, \lambda_1)\,\mathrm{d}\eta = P\left(S_k^1 < s\right)$$

and hence equation (4.11) and (4.10) equal zero, hence

$$P\left(\Delta N_1(t) = 1, \Delta N_2(t) = 1\right) = 0.$$

This holds for every t, hence two jumps never occur at the same time, almost surely. □

Next to this, we can prove an even stronger Lemma, which could have been used to prove Lemma 4.3:

Lemma 4.4. *Let $N_1(t)$ be a Poisson processes with intensity $\lambda_1 > 0$. Then $P\left(\Delta N_1(t) = 1\right) = 0$ for all t.*

Proof. Again $\Delta N_1(t) = 1$ if there is an arrival at time t, so $P\left(\Delta N_1(t) = 1\right) = \sum_{n=1}^{\infty} P\left(S_n = t\right)$, where S_n is the time of the n-th arrival. Now $P\left(S_n = t\right) = P\left(S_n \leq t\right) - P\left(S_n < t\right)$ and because

$$P\left(S_k \leq s\right) = \int_0^s f_{S_k}(\eta; k, \lambda_1)\,\mathrm{d}\eta = P\left(S_k < s\right) \tag{4.12}$$

we have $P\left(\Delta N_1(t) = 1\right) = \sum_{n=1}^{\infty} P\left(S_n = t\right) = 0$. □

We are now able to rewrite $\sum_{0<s\leq t}\left[f(s, X(s)) - f(s, X(s-))\right]$: When we denote by Δ_k the occurrence of a jump of process $N_k(t)$, we can rewrite (4.14) as

$$\begin{aligned}\sum_{0<s\leq t} \Delta f(s, X(s)) &= \sum_{0<s\leq t}\left[f(s, X(s)) - f(s, X(s-))\right] \\ &\overset{a.s.}{=} \sum_{0<s\leq t}\sum_{k=1}^{d} \Delta_k f(s, X(s)). \end{aligned} \tag{4.13}$$

Remark 4.5. *In the case of independent jumps this can be written out as:*

$$\begin{aligned}&\sum_{0<s\leq t}\left[f(s, X(s)) - f(s, X(s-))\right] \\ &\quad= \sum_{0<s\leq t}\left[f(s, (X_1(s), \ldots, X_d(s))) - f(s, (X_1(s-), \ldots, X_d(s-)))\right] \\ &\quad\overset{a.s.}{=} \sum_{0<s\leq t}\left[f(s, (X_1(s), \ldots, X_d(s))) - f(s, (X_1(s-), \ldots, X_d(s)))\right] \\ &\qquad+ \sum_{0<s\leq t}\left[f(s, (X_1(s), \ldots, X_d(s))) - f(s, (X_1(s), X_2(s-), \ldots, X_d(s)))\right] \\ &\qquad\vdots \\ &\qquad+ \sum_{0<s\leq t}\left[f(s, (X_1(s), \ldots, X_d(s))) - f(s, (X_1(s), \ldots, X_d(s-)))\right],\end{aligned} \tag{4.14}$$

because a jump in N_k only causes a jump in X_k.

Using equation (4.13), we can write the Itô-Doeblin formula (4.9) as

$$\begin{aligned}f(t, X(t)) &\overset{a.s.}{=} f(0, X(0)) + \int_0^t \frac{\partial}{\partial s} f(s, X(s))\,\mathrm{d}s + \sum_{i=1}^{d}\int_0^t \frac{\partial}{\partial x_i} f(s, X(s))\,\mathrm{d}X_i^c(s) \\ &+ \frac{1}{2}\sum_{i,j=1}^{d}\int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(s, X(s))\,\mathrm{d}X_i^c\,\mathrm{d}X_j^c + \sum_{0<s\leq t}\sum_{k=1}^{d} \Delta_k f(s, X(s)).\end{aligned} \tag{4.15}$$

4.2 Consistency of the Independent jump process

In the following theorem, we state a condition using the forward curve for the Independent jump process to be consistent, following Definition 3.1.

Theorem 4.6 (Filipović with independent jumps)**.** *Suppose* $Z = (Z_t)_{0\leq t<\infty}$ *follows a jump process with independent jumps with values in* $\mathcal{Z}$*. Then* Z *is consistent with the forward curve family* F *only if*

$$D_xF(x, Z_t) = b \cdot \nabla_z F(x, Z_t) + \frac{1}{2}\sum_{i,j=1}^{4} a^{ij}\frac{\partial^2}{\partial z^i \partial z^j}F(x, Z_t)$$
$$- \sum_{i,j=1}^{4} a^{ij}\left(\frac{\partial}{\partial z^i}F(x, Z_t)\int_0^x \frac{\partial}{\partial z^j}F(\eta, Z_t)\,\mathrm{d}\eta\right) - D_x\sum_{k=1}^{3}\left(1 - e^{x\alpha_k \mathbf{W}_x^k(Z^4)}\right)\lambda_k,$$

for all $x \geq 0$*, where* $a = \sigma\sigma^*$*.*

Proof. Assume as given an Independent jump process Z_t following Definition 4.4: For $i = 1, \ldots, 4,\ \ 0 \leq t < \infty$,

$$Z_t^i = Z_0^i + \int_0^t b_s^i ds + \sum_{j=1}^{d}\int_0^t \sigma_s^{ij} dW_s^j + \alpha^i N_t^i \mathbf{1}_{\{i\neq 4\}}, \tag{4.16}$$

where $W_t = (W_t^1, \ldots, W_t^d)_{0\leq t<\infty}$ is standard d-dimensional Brownian motion, $1 \leq d$, and N_t^j is a Poisson process with parameter λ_j, for $1 \leq j \leq 4$. This process can be written as the sum of a continuous process and a pure jump process:

$$Z_t^i = (Z^c)_t^i + \alpha^i N_t^i \mathbf{1}_{\{i\neq 4\}}, \tag{4.17}$$

where $(Z^c)_t^i$ is the continuous part, consisting of $Z_0^i + \int_0^t b_s^i ds + \sum_{j=1}^d \int_0^t \sigma_s^{ij} dW_s^j$ and N_t^i is the pure jump part. Remark that the continuous part $(Z^c)_t^i$ in this case equals Z_t^i in (3.6).

In line with the proof of Theorem 3.1, we apply equation (4.9) to the bond prize $P(t, T) = G(T - t, Z_t)$ where the process Z_t is given by (4.16). The continuous part of the proof remains the same, hence we add the jump part to (3.10)

$$\begin{aligned} P(t,T) = P(0,T) &+ \int_0^t (A_s G(T-s, Z_s) - D_x G(T-s, Z_t))\,\mathrm{d}s \\ &+ \int_0^t \nabla_z G(T-s, Z_s)^* \sigma_s\,\mathrm{d}W_s \\ &+ \sum_{0<s\leq t}\left[P(s,T) - P(s-,T)\right],\ 0 \leq t \leq T,\ \mathbb{P}\text{-}a.s.. \end{aligned} \tag{4.18}$$

Let $P(t,T)$ be the bond prize implied by the Nelson-Siegel forward rate (3.1), so

$$\begin{aligned}P(t,T) &= \exp\left[-(T-t)\left[Z_t^1 + Z_t^2\left(\frac{1-e^{-Z_t^4(T-t)}}{Z_t^4(T-t)}\right)\right.\right.\\ &\qquad \left.\left.+Z_t^3\left(\frac{1-e^{-Z_t^4(T-t)}}{Z_t^4(T-t)} - e^{-Z_t^4(T-t)}\right)\right]\right]\\ &= \exp\left[-(T-t)\mathbf{Z}_t^*\mathbf{W}_{T-t}(Z_t^4)\right],\end{aligned} \tag{4.19}$$

where

$$\mathbf{Z}_t := \begin{pmatrix} Z_t^1 \\ Z_t^2 \\ Z_t^3 \end{pmatrix}, \mathbf{W}_t(\lambda) := \begin{pmatrix} \mathbf{W}_t^1(\lambda) \\ \mathbf{W}_t^2(\lambda) \\ \mathbf{W}_t^3(\lambda) \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1-e^{-\lambda t}}{\lambda t} \\ \frac{1-e^{-\lambda t}}{\lambda t} - e^{-\lambda t} \end{pmatrix}. \tag{4.20}$$

Now we are interested in investigating the jump part of (4.18):

$$\begin{aligned}\sum_{0<s\le t}\left[P(s,T)-P(s-,T)\right] &= \sum_{0<s\le t}\Delta P(s,T)\\ &\overset{a.s.}{=} \sum_{0<s\le t}\Delta_1 P(s,T)+\Delta_2 P(s,T)+\Delta_3 P(s,T)\\ &= \sum_{0<s\le t}\sum_{k=1}^{3}\Delta_k P(s,T).\end{aligned}$$

When there is a jump at time s in Z_t^i we know $Z_s^i = Z_{s-}^i + \Delta Z_s^i = Z_{s-}^i + \alpha_i \Delta N_s^i = Z_{s-}^i + \alpha_i$. Denote by $\mathbf{e}_k$ the k-th unit vector, then $\mathbf{Z}_{s-} = \mathbf{Z}_s - \alpha_k \mathbf{e}_k$ when there is a jump of N^k at time s. Now

$$\begin{aligned}\Delta_k P(s,T) &= P(s,T) - P(s-,T)\\ &= \exp\left[-(T-s)\mathbf{Z}_s^*\mathbf{W}_{T-s}(Z_s^4)\right] - \exp\left[-(T-s)\mathbf{Z}_{s-}^*\mathbf{W}_{T-s}(Z_s^4)\right]\\ &= \exp\left[-(T-s)\mathbf{Z}_s^*\mathbf{W}_{T-s}(Z_s^4)\right] - \exp\left[-(T-s)(\mathbf{Z}_s^* - \alpha_k\mathbf{e}_k^*)\mathbf{W}_{T-s}(Z_s^4)\right]\\ &= \exp\left[-(T-s)\mathbf{Z}_s^*\mathbf{W}_{T-s}(Z_s^4)\right]\left(1-\exp\left[(T-s)\alpha_k\mathbf{e}_k^*\mathbf{W}_{T-s}(Z_s^4)\right]\right)\\ &= P(s,T)\left(1-\exp\left[(T-s)\alpha_k\mathbf{W}_{T-s}^k(Z_s^4)\right]\right)\end{aligned}$$

when there is a jump of process N^k at time s, else $\Delta_k P(s,T) = 0$. Hence

$$\Delta_k P(s,T) = P(s,T)\left(1-\exp\left[(T-s)\alpha_k\mathbf{W}_{T-s}^k(Z_s^4)\right]\right)\Delta N_k(s). \tag{4.21}$$

Remark 4.7. *Above we used* $\mathbf{Z}_{s-} = \mathbf{Z}_s - \mathbf{e}_k$*. Of course we could also have substituted* $\mathbf{Z}_s = \mathbf{Z}_{s-} + \mathbf{e}_k$*. This would result in*

$$\Delta_k P(s,T) = P(s-,T)\left(\exp\left[-(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)\right] - 1\right)\Delta N_k(s). \tag{4.22}$$

Combining (4.21) *and* (4.22) *gives*

$$P(s,T)\Delta N_k(s) = \exp\left[(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)\right] P(s-,T)\Delta N_k(s),$$

which shows the relation between $P(s,T)\Delta N_k(s)$ *and* $P(s-,T)\Delta N_k(s)$*.*

Now we get

$$\begin{aligned}\sum_{0<s\le t}\left[P(s,T) - P(s-,T)\right] &= \sum_{0<s\le t}\sum_{k=1}^{3}\Delta_k P(s,T)\\ &= \sum_{0<s\le t}\sum_{k=1}^{3} P(s,T)\left(1-\exp\left[(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)\right]\right)\Delta N_k(s).\end{aligned} \tag{4.23}$$

Remark that $\sum_{0<s\le t} f(s)\Delta N(s) = \int_0^t f(s)\,\mathrm{d}N(s)$ hence we can write

$$\sum_{0<s\le t}\left[P(s,T) - P(s-,T)\right] = \sum_{k=1}^{3}\int_0^t P(s,T)\left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right)\mathrm{d}N_k(s). \tag{4.24}$$

Equation (4.18) now becomes:

$$\begin{aligned}P(t,T) = P(0,T) &+ \int_0^t \left(A_s G(T-s,Z_s) - D_x G(T-s,Z_t)\right)\mathrm{d}s\\ &+ \int_0^t \nabla_z G(T-s,Z_s)^* \sigma_s\,\mathrm{d}W_s\\ &+ \sum_{k=1}^{3}\int_0^t P(s,T)\left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right)\mathrm{d}N_k(s),\ 0\le t\le T,\ \mathbb{P}\text{-}a.s..\end{aligned}$$

In differential notation:

$$\begin{aligned}dP(t,T) = &\left(A_t G(T-t,Z_t) - D_x G(T-t,Z_t)\right)\mathrm{d}t + \nabla_z G(T-t,Z_t)^*\sigma_t\,\mathrm{d}W_t\\ &+ \sum_{k=1}^{3} P(t,T)\left(1 - e^{(T-t)\alpha_k \mathbf{W}^k_{T-t}(Z^4_t)}\right)\mathrm{d}N_k(t),\ 0\le t\le T,\ \mathbb{P}\text{-}a.s..\end{aligned}$$

And this is equal to

$$\mathrm{d}P(t,T) = (A_t G(T-t,Z_t) - D_x G(T-t,Z_t))\,\mathrm{d}t + \nabla_z G(T-t,Z_t)^* \sigma_t\,\mathrm{d}W_t$$
$$+ \sum_{k=1}^{3} G(T-t,Z_t)\left(1 - e^{(T-t)\alpha_k \mathbf{W}^k_{T-t}(Z^4_t)}\right)\mathrm{d}N_k(t),\ 0 \le t \le T,\ \mathbb{P}\text{-}a.s..$$

In the same way as in equation (3.11) we are able to investigate $\mathrm{d}\left(\frac{P(t,T)}{B(t)}\right)$:

$$\mathrm{d}\left(\frac{P(t,T)}{B(t)}\right) = \frac{1}{B(t)}\,\mathrm{d}P(t,T) + P(t,T)\,\mathrm{d}\left(\frac{1}{B(t)}\right) + \mathrm{d}\left(\frac{1}{B(t)}\right)\mathrm{d}P(t,T)$$
$$= \frac{1}{B(t)}\Bigg[\big(AG(T-t,Z_t) - D_x G(T-t,Z_t)\big)\,\mathrm{d}t + \nabla_z G(T-t,Z_t)^* \sigma_t\,\mathrm{d}W_t$$
$$+ G(T-t,Z_t)\sum_{k=1}^{3}\left(1 - e^{(T-t)\alpha_k \mathbf{W}^k_{T-t}(Z^4_t)}\right)\mathrm{d}N_k(t)\Bigg]$$
$$+ G(T-t,Z_t)\frac{1}{B(t)} D_x G(0,Z_t)\,\mathrm{d}t$$
$$+ \frac{1}{B(t)} D_x G(0,Z_t)\,\mathrm{d}t\,\Bigg[\big(AG(T-t,Z_t) - D_x G(T-t,Z_t)\big)\,\mathrm{d}t$$
$$+ \nabla_z G(T-t,Z_t)^* \sigma_t\,\mathrm{d}W_t + G(T-t,Z_t)\sum_{k=1}^{3}\left(1 - e^{(T-t)\alpha_k \mathbf{W}^k_{T-t}(Z^4_t)}\right)\mathrm{d}N_k(t)\Bigg]$$
$$= \frac{1}{B(t)}\Bigg[\big(AG(T-t,Z_t) - D_x G(T-t,Z_t)\big)\,\mathrm{d}t + \nabla_z G(T-t,Z_t)^* \sigma_t\,\mathrm{d}W_t$$
$$+ G(T-t,Z_t)\sum_{k=1}^{3}\left(1 - e^{(T-t)\alpha_k \mathbf{W}^k_{T-t}(Z^4_t)}\right)\mathrm{d}N_k(t)\Bigg]$$
$$+ G(T-t,Z_t)\frac{1}{B(t)} D_x G(0,Z_t)\,\mathrm{d}t.$$

Again because $\mathrm{d}t\,\mathrm{d}t = 0$, $\mathrm{d}t\,\mathrm{d}W_t = 0$ and $\mathrm{d}t\,\mathrm{d}N(t) = 0$. Now we have in integral form:

$$\frac{P(t,T)}{B(t)} - P(0,T) = \int_0^t \frac{1}{B(s)} \nabla_z G(T-s,Z_s)^* \sigma_s\,\mathrm{d}W_s$$
$$+ \int_0^t \frac{1}{B(s)}\left[AG(T-s,Z_s) - D_x G(T-s,Z_s) + G(T-s,Z_s) D_x G(0,Z_s)\right]\mathrm{d}s$$
$$+ \sum_{k=1}^{3}\int_0^t \frac{1}{B(s)} G(T-s,Z_s)\left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right)\mathrm{d}N_k(s). \tag{4.25}$$

Remark that if poisson process $N(t)$ has intensity λ, $N(t) - \lambda t$ is a martingale. Add and subtract to (4.25)

$$\sum_{k=1}^{3} \int_0^t \frac{1}{B(s)} G(T-s, Z_s) \left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right) \lambda_k \, \mathrm{d}s$$

to get

$$\begin{aligned}
\frac{P(t,T)}{B(t)} - P(0,T) &= \int_0^t \frac{1}{B(s)} \nabla_z G(T-s, Z_s)^* \sigma_s \, \mathrm{d}W_s \\
&+ \int_0^t \frac{1}{B(s)} \big[AG(T-s, Z_s) - D_x G(T-s, Z_s) + G(T-s, Z_s) D_x G(0, Z_s) \big] \, \mathrm{d}s \\
&+ \sum_{k=1}^{3} \int_0^t \frac{1}{B(s)} G(T-s, Z_s) \left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right) \mathrm{d}N_k(s) \\
&+ \sum_{k=1}^{3} \int_0^t \frac{1}{B(s)} G(T-s, Z_s) \left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right) \lambda_k \, \mathrm{d}s \\
&- \sum_{k=1}^{3} \int_0^t \frac{1}{B(s)} G(T-s, Z_s) \left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right) \lambda_k \, \mathrm{d}s \\
&= \int_0^t \frac{1}{B(s)} \nabla_z G(T-s, Z_s)^* \sigma_s \, \mathrm{d}W_s \\
&+ \int_0^t \frac{1}{B(s)} \big[AG(T-s, Z_s) - D_x G(T-s, Z_s) + G(T-s, Z_s) D_x G(0, Z_s) \big] \, \mathrm{d}s \\
&+ \sum_{k=1}^{3} \int_0^t \frac{1}{B(s)} G(T-s, Z_s) \left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right) \mathrm{d}\big[N_k(s) - \lambda_k s\big] \\
&+ \sum_{k=1}^{3} \int_0^t \frac{1}{B(s)} G(T-s, Z_s) \left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right) \lambda_k \, \mathrm{d}s. \qquad (4.26)
\end{aligned}$$

Now define

$$\begin{aligned}
M_2(t,T) := &\int_0^t \frac{1}{B(s)} \nabla_z G(T-s, Z_s)^* \sigma_s \, \mathrm{d}W_s \\
&+ \sum_{k=1}^{3} \int_0^t \frac{1}{B(s)} G(T-s, Z_s) \left(1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)}\right) \mathrm{d}\big[N_k(s) - \lambda_k s\big]
\end{aligned}$$

and

$$\begin{aligned}
H_2(t,T) := \frac{1}{B(t)} [&AG(T-t, Z_t) - D_x G(T-t, Z_t) + G(T-t, Z_t) D_x G(0, Z_t) \\
&+ G(T-t, Z_t) \sum_{k=1}^{3} \left(1 - e^{(T-t)\alpha_k \mathbf{W}^k_{T-t}(Z^4_t)}\right)] \lambda_k. \qquad (4.27)
\end{aligned}$$

$M_2(t,T)$ is a local $\mathbb{P}$-martingale because it is an integral with respect to a (local) $\mathbb{P}$-martingale. Now we have that equation (4.26) can be written as

$$\frac{P(t,T)}{B(t)} - P(0,T) = \int_0^t H_2(s,T)\,\mathrm{d}s + M_2(t,T). \tag{4.28}$$

In the same way as for the Itô process part, lets suppose Z is consistent with the Nelson-Siegel family, i.e. $\left(\frac{P(t,T)}{B(t)}\right)_{0\le t\le T}$ is a $\mathbb{P}$-martingale, for all $T<\infty$. Then we know $\int_0^t H_2(s,T)\,\mathrm{d}s$ is a local martingale. In this case, it is only right-continuous so we have to do a little bit more work to be able to conclude $\int_0^t H_2(s,T)\,\mathrm{d}s = 0$.

Fix $\omega\in\Omega$. Let $\tau_1,\dots,\tau_l$ be the jump times of process Z such that $\tau_0 := 0 < \tau_1 < \ldots < \tau_l < t =: \tau_{l+1}$. On each of the intervals $[\tau_k,\tau_{k+1})$, $H_2(s,T)$ is continuous. Now $\int_0^{t\wedge\tau_1} H_2(s,T)\,\mathrm{d}s$ is a local martingale which is continuous and of bounded variation. Therefore $\int_0^{t\wedge\tau_1} H_2(s,T)\,\mathrm{d}s = 0$. Now we can write $\int_0^t H_2(s,T)\,\mathrm{d}s = \int_{\tau_1}^{t\vee\tau_1} H_2(s,T)\,\mathrm{d}s$ which still is a local martingale. Applying the same argument gives $\int_{\tau_1}^{t\vee\tau_2} H_2(s,T)\,\mathrm{d}s = 0$. Applying this repeatedly we conclude $\int_0^t H_2(s,T)\,\mathrm{d}s = 0$. Applying Claim 3.14 on the continuous parts of the integral, we conclude, because $H_2(t,T)$ is continuous in T,

$$H_2(t,t+x)(\omega) = 0, \quad \forall x\ge 0, \quad \text{for} \quad \mathrm{d}t\otimes\mathrm{d}\mathbb{P}\text{-}a.e.\ (t,\omega). \tag{4.29}$$

Because $B(t)>0$ for all t, (4.29) yields

$$\begin{aligned} &AG(x,Z_t) - D_xG(x,Z_t) + G(x,Z_t)D_xG(0,Z_t) \\ &+ G(x,Z_t)\sum_{k=1}^{3}\left(1-e^{x\alpha_k \mathbf{W}_x^k(Z_t^4)}\right)\lambda_k, \quad \forall x\ge 0, \quad \mathrm{d}t\otimes\mathrm{d}\mathbb{P}\text{-}a.s.. \end{aligned} \tag{4.30}$$

Using again that the definition of the bond price $G(x,z), z\in\mathcal{Z}$, equation (3.2), we have (see Appendix A.1):

$$\frac{\mathrm{d}G(x,z)}{\mathrm{d}z^i} = -\int_0^x \frac{\mathrm{d}}{\mathrm{d}z^i}F(\eta,z)\,\mathrm{d}\eta G(x,z),$$

$$\frac{\partial^2 G(x,z)}{\partial z^j\partial z^i} = \left(\int_0^x \frac{\partial}{\partial z^i}F(\eta,z)\,\mathrm{d}\eta\int_0^x \frac{\partial}{\partial z^j}F(\eta,z)\,\mathrm{d}\eta - \int_0^x \frac{\partial^2}{\partial z^j\partial z^i}\right)G(x,z)$$

and

$$D_xG(x,z) = -F(x,z)G(x,z).$$

Equation (4.30) can now be written as

$$\begin{aligned}0 &= AG(x,Z) - D_xG(x,Z) + D_xG(0,Z)G(x,Z)\\&\quad + G(x,Z)\sum_{k=1}^{3}\left(1-e^{x\alpha_k \mathbf{W}_x^k(Z^4)}\right)\lambda_k\\&= -\int_0^x AF(\eta,Z)\,\mathrm{d}\eta G(x,Z)\\&\quad + \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}\left(\int_0^x \frac{\partial}{\partial z^i}F(\eta,Z)\,\mathrm{d}\eta \int_0^x \frac{\partial}{\partial z^j}F(\eta,Z)\,\mathrm{d}\eta\right)G(x,Z)\\&\quad + F(x,Z)G(x,Z) - F(0,Z)G(x,Z)\\&\quad + G(x,Z)\sum_{k=1}^{3}\left(1-e^{x\alpha_k \mathbf{W}_x^k(Z^4)}\right)\lambda_k, \quad \forall x\geq 0, \quad \mathrm{d}t\otimes \mathrm{d}\mathbb{P}\text{-}a.s..\end{aligned}$$

Dividing by $-G(x,Z)$ gives:

$$\begin{aligned}&\int_0^x AF(\eta,z)\,\mathrm{d}\eta - \frac{1}{2}\sum_{i,j=1}^{4} a^{ij}\left(\int_0^x \frac{\partial}{\partial z^i}F(\eta,Z)\,\mathrm{d}\eta \int_0^x \frac{\partial}{\partial z^j}F(\eta,Z)\,\mathrm{d}\eta\right)\\&\quad - F(x,Z) + F(0,Z) - \sum_{k=1}^{3}\left(1-e^{x\alpha_k \mathbf{W}_x^k(Z^4)}\right)\lambda_k = 0, \quad \forall x\geq 0, \quad \mathrm{d}t\otimes \mathrm{d}\mathbb{P}\text{-}a.s..\end{aligned}$$

Differentiating this to x gives

$$\begin{aligned}&-\frac{1}{2}\sum_{i,j=1}^{4} a^{ij}\left(\frac{\partial}{\partial z^i}F(x,Z)\int_0^x \frac{\partial}{\partial z^j}F(\eta,Z)\,\mathrm{d}\eta + \int_0^x \frac{\partial}{\partial z^i}F(\eta,Z)\,\mathrm{d}\eta \frac{\partial}{\partial z^j}F(x,Z)\right)\\&+ AF(\eta,Z)\,\mathrm{d}\eta - D_xF(x,Z) - D_x\sum_{k=1}^{3}\left(1-e^{x\alpha_k \mathbf{W}_x^k(Z^4)}\right)\lambda_k = 0, \qquad (4.31)\end{aligned}$$

$\forall x \geq 0$, $\mathrm{d}t\otimes \mathrm{d}\mathbb{P}$-*a.s.*, which can be rewritten to the expression of Theorem 4.6. □

This theorem is very similar to Theorem 3.1, except the addition of three terms involving the jump intensities λ_k and the elements of the matrix $\mathbf{\Sigma}$. In the following corollary we work out what Theorem 4.6 explicitly means in the Nelson-Siegel forward curve case.

Corollary 4.8 (Nelson-Siegel with independent jumps)**.** *Suppose $Z = (Z_t)_{0\leq t<\infty}$ follows the jump process of Definition 4.4 with values in $\mathcal{Z}$. Then Z is consistent with the Nelson-Siegel family only if for $\mathrm{d}t\otimes\mathrm{d}\mathbb{P}$-a.e. (t,ω) in $[0,\infty)\times\Omega$,*

$$\begin{aligned}0 =& p_0(x) + p_1(x)e^{-z_4x} + p_2(x)e^{-2z_4x}\\&+ \lambda_1\alpha_1 e^{\alpha_1 x} + \lambda_2\alpha_2 e^{\alpha_2\left(\frac{1-e^{-z_4x}}{z_4}\right)-z_4x} + \lambda_3\alpha_3 z_4 x e^{\alpha_3\left(\frac{1-e^{-z_4x}}{z_4}-xe^{-z_4x}\right)-z_4x},\end{aligned} \qquad (4.32)$$

for all $x \geq 0$, where $p_0(x), p_1(x)$ and $p_2(x)$ are polynomials in x with coefficients containing $b^i := b_t^i(\omega), a^{ij} := a_t^{ij}(\omega)$ and $z^i := Z_t^i(\omega)$ for $1 \leq i, j \leq 4$ which are given by (3.18), (3.19) and (3.20) respectively.

Proof. The condition follows from writing out the condition of Theorem 4.6. The proof is the same as the proof of Corollary 3.3, but then with addition of $-D_x \sum_{k=1}^3 \left(1 - e^{x\alpha_k W_x^k(z^4)}\right) \lambda_k$. Before we work this out, remember the relations as shown in section 2.1.1:

$$\frac{1}{\tau}\int_0^\tau \mathbf{w}_s(\lambda)\,\mathrm{d}s = \frac{1}{\tau}\int_0^\tau \begin{pmatrix} 1 \\ e^{-\lambda s} \\ \lambda s e^{-\lambda s} \end{pmatrix} \mathrm{d}s = \begin{pmatrix} \frac{1}{\tau}\int_0^\tau 1\,\mathrm{d}s \\ \frac{1}{\tau}\int_0^\tau e^{-\lambda s}\,\mathrm{d}s \\ \frac{1}{\tau}\int_0^\tau \lambda s e^{-\lambda s}\,\mathrm{d}s \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ \frac{1-e^{\lambda\tau}}{\lambda\tau} \\ \frac{1-e^{\lambda\tau}}{\lambda\tau} - e^{\lambda\tau} \end{pmatrix} =: \mathbf{W}_\tau(\lambda). \tag{4.33}$$

Hence, because α_k does not depend on x,

$$D_x\left(x\alpha_k \mathbf{W}_x(\lambda)\right) = \alpha_k \begin{pmatrix} 1 \\ e^{-\lambda x} \\ \lambda x e^{-\lambda x} \end{pmatrix} = \alpha_k \mathbf{w}_x(\lambda).$$

Now

$$\begin{aligned} -D_x \sum_{k=1}^3 \left(1 - e^{x\alpha_k \mathbf{W}_x^k(z^4)}\right)\lambda_k &= \sum_{k=1}^3 D_x e^{x\alpha_k \mathbf{W}_x^k(z^4)} \lambda_k \\ &= \sum_{k=1}^3 e^{x\alpha_k \mathbf{W}_x^k(z^4)} D_x\left(x\alpha_k \mathbf{W}_x^k(z^4)\right)\lambda_k \\ &= \sum_{k=1}^3 e^{x\alpha_k \mathbf{W}_x^k(z^4)} \alpha_k \mathbf{w}_x^k(z^4)\lambda_k, \end{aligned}$$

which can be written out as

$$\begin{aligned} -D_x \sum_{k=1}^3 \left(1 - e^{x\alpha_k \mathbf{W}_x^k(z^4)}\right)\lambda_k &= \lambda_1\alpha_1 e^{\alpha_1 x} + \lambda_2\alpha_2 e^{\alpha_2\left(\frac{1-e^{-z_4 x}}{z_4}\right) - z_4 x} \\ &\quad + \lambda_3\alpha_3 z_4 x e^{\alpha_3\left(\frac{1-e^{-z_4 x}}{z_4} - x e^{-z_4 x}\right) - z_4 x}. \end{aligned}$$

Adding this to the result of Corollary 3.3 gives Corollary 4.8. □

It would be nice to be able to give a theoretical proof that Corollary 4.8 can only be satisfied if $\alpha_1 = \alpha_2 = \alpha_3 = 0$, i.e. that Corollary 4.8 can only be satisfied if there is no jump part in the state space process. The rest of the analysis would be going like the analysis of the Itô process part. But this, in general, is not as easy as the proofs in Appendix C which are used in the analysis of the Itô process part. It looks evident that it has to be the case that $p_0(x) = p_1(x) = p_2(x) = \alpha_1 = \alpha_2 = \alpha_3 = 0$. Mathematica gives the desired result.

Because α_1, α_2 and α_3 have to be 0 to give a consistent state space process, we see that an Independent jump process, a state space process of the form of Definition 4.4, can't in fact have jumps if it has to be consistent. Therefore it will be a process as considered in Chapter 3 and we can draw the same conclusion:

Theorem 4.9. *Suppose $Z = (Z_t)_{0\leq t<\infty}$ follows the Independent jump process of Definition 4.4 with values in $\mathcal{Z}$. Let Z_t be consistent with the Nelson-Siegel family. Then Z_t is of the form*

$$\begin{aligned}
Z_t^1 &= Z_0^1 \\
Z_t^2 &= Z_0^2 e^{-Z_0^4 t} + Z_0^3 Z_0^4 t e^{-Z_0^4 t} \\
Z_t^3 &= Z_0^3 e^{-Z_0^4 t} \\
Z_t^4 &= Z_0^4 + \int_0^{t\wedge\tau} c_s^1 \mathbf{1}_{\{Z_0^2 = Z_0^3 = 0\}}\, \mathrm{d}s + \sum_{j=1}^{d} \int_0^{t\wedge\tau} \sigma_s^{4j} \mathbf{1}_{\{Z_0^2 = Z_0^3 = 0\}}\, \mathrm{d}W_s^j,
\end{aligned}$$

with $c^1 \in \mathbb{R}$ and the stopping time $\tau := \inf\{s > 0 | Z_s^2 = Z_s^3 = 0\}$.

Proof. Because for Z to be consistent α_1, α_2 and α_3 have to be 0, the Independent jump process is in fact just an Itô process following Definition 3.4. Therefore, the proof is the same as the proof of theorem 3.7. □

And we can formulate the same corollary as in Chapter 3:

Corollary 4.10. *The forward rates are non-random, they are $\mathcal{F}_0$-measurable.*

Proof. The Nelson-Siegel forward rate on Ω_0 is given by $F(x, Z_t) = Z_0^1$ (because $Z_t^2 = Z_t^3 = 0$), which is $\mathcal{F}_0$-measurable. The dynamics of Z_4^t have no influence. The Nelson-Siegel forward rate on Ω_1 is given by

$$\begin{aligned}
F(x, Z_t) &= Z_t^1 + Z_t^2 e^{-Z_t^4 x} + Z_t^3 Z_t^4 x e^{-Z_t^4 x} \\
&= Z_0^1 + \left(Z_0^2 e^{-Z_0^4 t} + Z_0^3 Z_0^4 t e^{-Z_0^4 t}\right) e^{-Z_0^4 x} + Z_0^3 Z_t^4 e^{-Z_0^4 t} x e^{-Z_0^4 x} \\
&= Z_0^1 + Z_0^2 e^{-Z_0^4 (t+x)} + Z_0^3 Z_0^4 (t+x) e^{-Z_0^4 (t+x)} = F(t+x, Z_0),
\end{aligned}$$

hence all the randomness remains $\mathcal{F}_0$-measurable. □

Corollary 4.10 tells us that the interest rate model coming from a consistent Independent jump process is nonrandom, trivial. It only depends on the value of the process at time 0, as we saw earlier in the Itô process case.

Chapter 5

Concluding remarks

The Nelson-Siegel model is used by many practitioners in the field. In the beginning it was just a method to fit the yield curve, nowadays people have developed methods using the Nelson-Siegel curve to predict the yield curve.

Filipović's work showed the lack of theoretical background of this model as a forecasting method based on continuous processes. In his line, we defined the consistent state space process: the process which, when representing the parameters of the Nelson-Siegel curve (or in general of a forward rate curve), turns the discounted bond price into a martingale (which can be seen as the no-arbitrage condition). First we considered an Itô process and using the definition of the consistent state space process, we derived conditions on the dynamics of the Itô process as the state space processes. We concluded there exists no nontrivial interest rate model driven by an Itô process consistent with the Nelson-Siegel family.

Secondly we extended his research by introducing jump processes and stochastic calculus for jump processes. We also derived conditions on the dynamics of the Independent jump process in order to represent a consistent state space process. It turned out that there exists no nontrivial interest rate model driven by an Independent jump process.

Based on my experience in the recent project, I believe it is relevant to investigate the mathematical background of (new) mathematical and financial techniques to obtain a more reliable framework for financial products. From my point of view, a theoretically rigid model which works well in practice is preferable. As mentioned before, some people think we don't need a mathematically or theoretically rigid method to forecast the yield curve. This is of course true, if forecasting is the only aim of the method. However, as soon as we want to say more about it, we need a solid mathematical foundation. Not only because this could support the assumptions we make in mathematical finance (like the concept of arbitrage-freeness), but also because we will be able to base new theories on it.

Further research

We only showed the dynamics of a consistent jump process in the Nelson-Siegel framework for a so-called Independent jump process. It would be nice to have a similar result for an arbitrary jump process, according to Definition 4.3, where the jump part is given by $J = (J^1, J^2, J^3, 0)$, $J_t^i = \sum_{j=1}^m \mathbf{\Sigma}^{ij} N_t^j$ where $N_t = (N_t^1, \ldots, N_t^m)$ is an m -dimensional vector with independent Poisson processes with parameters $\lambda_1, \ldots, \lambda_m > 0$ and $\mathbf{\Sigma} \in \mathbb{R}^{4\times m}$.

One can show this process leads to the following equivalence of Theorem 3.1 and Theorem 4.6:

Theorem (Filipović with jumps). *Suppose $Z = (Z_t)_{0\le t<\infty}$ follows a jump process as described above with values in $\mathcal{Z}$. Then Z is consistent with the forward curve family F family only if*

$$
\begin{aligned}
D_x F(x,Z) = b \cdot \nabla_z F(x,Z) + \frac{1}{2}\sum_{i,j=1}^{4} a^{ij} \frac{\partial^2}{\partial z^i \partial z^j} F(x,Z) \\
- \sum_{i,j=1}^{4} a^{ij} \left(\frac{\partial}{\partial z^i} F(x,Z) \int_0^x \frac{\partial}{\partial z^j} F(\eta, Z)\, \mathrm{d}\eta \right) \\
- D_x \sum_{k=1}^{m} \left(1 - e^{x \sum_{j=1}^3 \mathbf{\Sigma}^{jk} \mathbf{W}_x^j (Z^4)} \right) \lambda_k
\end{aligned}
$$

for all $x \ge 0$, where $a = \sigma\sigma^*$.

Proof. This proof is similar to the proof of Theorem 4.6. Outline of the proof is given in Appendix D.1. ☐

Corollary (Nelson-Siegel with jumps). *Suppose $Z = (Z_t)_{0\le t<\infty}$ follows the jump process the Theorem above, with values in $\mathcal{Z}$. Then Z is consistent with the Nelson-Siegel family only if for $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-a.e. (t,ω) in $[0,\infty)\times\Omega$,*

$$
\begin{aligned}
0 = p_0(x) + p_1(x) e^{-z_4 x} + p_2(x) e^{-2z_4 x} \\
+ \sum_{k=1}^{m} \lambda_k \left[e^{x \sum_{j=1}^3 \mathbf{\Sigma}^{jk} \mathbf{W}_x^j (z_4)} \sum_{i=1}^{3} \mathbf{\Sigma}^{ik} \mathbf{w}_x^i (z_4) \right],
\end{aligned}
$$

for all $x \ge 0$, where $p_0(x), p_1(x)$ and $p_2(x)$ are polynomials in x with coefficients containing $b^i := b_t^i(\omega), a^{ij} := a_t^{ij}(\omega)$ and $z^i := Z_t^i(\omega)$ for $1 \le i,j \le 4$ which are given by (3.18), (3.19) and (3.20) respectively.

Proof. This proof is similar to the proof of Corollary 4.8. See Appendix D.2 for more details. ☐

This is an equation with $3m - 3$ (the nonzero entries of $\mathbf{\Sigma}$ minus 3) more unknowns than the equation of Corollary 4.8. Mathematica didn't manage to solve this. A nice extension of my research would be to find a theoretical way to proof a that this equation also implies no existence of nontrivial interest rate models. I believe the possible corresponding processes providing a nontrivial interest rate model are at most very restricted, because of the complexity of the equation, which has to hold for all $x \geq 0$.

Apart from this 'arbitrary jump process', one could also think of other jump processes to simulate interest rate jumps in specific periods in time, like processes with clustered jumps, jump processes with time dependent intensities and jump processes with time dependent jump sizes. Instead of investigating all those cases separately, I think the analysis as proposed in this thesis can be extended to a more 'overall' result. It may work with other, arbitrary, (jump) processes, as long as they have a continuous compensator to be able to construct expressions like 3.12 and 4.28.

Another subject which is not explored in this thesis is the question whether (some of) those jump processes describe the data well, or in any case better than the continuous processes, to be able to extend the research of Diebold and Li [9].

Finally we come up with a subject for further research, which is not directly related to this thesis, but involves arbitrage-free-ness and jump processes. As mentioned in the introduction, Christensen, Diebold and Rudebusch adapted the Nelson-Siegel model to make it arbitrage-free [4, 5]. They only used continuous processes. Being convinced that including jumps in models is almost obligatory, another research subject might be to investigate what the impact of jumps is on their arbitrage-free Nelson-Siegel model.

Bibliography

[1] T. Björk and B. J. Christensen. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9:323–348, 1999.

[2] D. J. Bolder. Affine term-structure models: Theory and implementation. Working Papers 01-15, Bank of Canada, 2001.

[3] F. Boshuizen, A. W. van der Vaart, H. van Zanten, K. Banachewicz, and P. Zareba. Stochastic processes for finance: Risk management tools, July 2010.

[4] J. H. E. Christensen, F. X. Diebold, and G. D. Rudebusch. The affine arbitrage-free class of nelson-siegel term structure models. NBER Working Papers 13611, National Bureau of Economic Research, Inc, Nov. 2007.

[5] J. H. E. Christensen, F. X. Diebold, and G. D. Rudebusch. An arbitrage-free generalized nelson-siegel term structure model. *Econometrics Journal*, 12(3):C33–C64, November 2009.

[6] L. Coroneo, K. Nyholm, and R. Vidova-Koleva. How arbitrage-free is the nelson-siegel model? Working Paper Series 874, European Central Bank, Feb. 2008.

[7] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.

[8] M. de Pooter. Examining the nelson-siegel class of term structure models. Tinbergen Institute Discussion Papers 07-043/4, Tinbergen Institute, June 2007.

[9] F. X. Diebold and C. Li. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364, February 2006.

[10] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406, 1996.

[11] A. Etheridge. *A Course in Financial Calculus*. Cambridge University Press, 2002.

[12] A. Evans-Pritchard. Fitch doubts dutch AAA as property slump reaches 'coma'. *Telegraph*, 2012.

[13] D. Filipović. A note on the nelson-siegel family. *Mathematical Finance*, 9(4):349–359, 1999.

[14] D. Filipović. Exponential-polynomial families and the term structure of interest rates. *Bernoulli*, 6:1081–1107, 2000.

[15] B. for International Settlements. Monetary and E. Dept. *Zero-coupon Yield Curves: Technical Documentation.* BIS papers. Bank for International Settlements, Monetary and Economic Department, 2005.

[16] J. A. Frankel and C. S. Lown. An indicator of future inflation extracted from the steepness of the interest rate yield curve along its entire length. *The Quarterly Journal of Economics*, 109(2):517–30, May 1994.

[17] J. Hull and A. White. Pricing interest-rate-derivative securities. *Review of Financial Studies*, 3(4):573–92, 1990.

[18] C. R. Nelson and A. F. Siegel. Parsimonious modeling of yield curves. *Journal of Business*, 60(4):473–89, October 1987.

[19] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, 1994.

[20] S. E. Shreve. *Stochastic calculus for finance. II.* Springer Finance. Springer-Verlag, New York, 2004.

[21] P. J. C. Spreij. Stochastic integration, February 2011.

[22] L. E. O. Svensson. Estimating and interpreting forward interest rates. *NBER*, 1994.

[23] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188, November 1977.

Appendix A

Nelson-Siegel family: integrals and derivatives

A.1 Bond prices

Recall the definition of the bond price $G(x,z)$, $z \in \mathcal{Z}$, equation (3.2):

$$G(x,z) = \exp(-y(x,z)x) = \exp\left(-\int_0^x F(\eta,z)\,\mathrm{d}\eta\right).$$

Then we can calculate the following derivatives:

$$\begin{aligned}
\frac{\mathrm{d}G(x,z)}{\mathrm{d}z^i} &= \frac{\mathrm{d}}{\mathrm{d}z^i}\exp\left(-\int_0^x F(\eta,z)\,\mathrm{d}\eta\right) \\
&= \exp\left(-\int_0^x F(\eta,z)\,\mathrm{d}\eta\right)\frac{\mathrm{d}}{\mathrm{d}z^i}\left(-\int_0^x F(\eta,z)\,\mathrm{d}\eta\right) \\
&= -\int_0^x \frac{\mathrm{d}}{\mathrm{d}z^i}F(\eta,z)\,\mathrm{d}\eta G(x,z), \qquad \text{(A.1)}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 G(x,z)}{\partial z^j \partial z^i} &= \frac{\partial^2}{\partial z^j \partial z^i}\exp\left(-\int_0^x F(\eta,z)\,\mathrm{d}\eta\right) \\
&= \frac{\partial}{\partial z^j}\left[\exp\left(-\int_0^x F(\eta,z)\,\mathrm{d}\eta\right)\frac{\partial}{\partial z^i}\left(-\int_0^x F(\eta,z)\,\mathrm{d}\eta\right)\right] \\
&= \frac{\partial}{\partial z^j}\left[-\int_0^x \frac{\partial}{\partial z^i}F(\eta,z)\,\mathrm{d}\eta G(x,z)\right] \\
&= \left(-\int_0^x \frac{\partial^2}{\partial z^j \partial z^i}F(\eta,z)\,\mathrm{d}\eta\right)G(x,z) \\
&\quad + \left(-\int_0^x \frac{\partial}{\partial z^i}F(\eta,z)\,\mathrm{d}\eta\right)\left(-\int_0^x \frac{\mathrm{d}}{\mathrm{d}z^j}F(\eta,z)\,\mathrm{d}\eta\right)G(x,z) \\
&= \Bigg(\int_0^x \frac{\partial}{\partial z^i}F(\eta,z)\,\mathrm{d}\eta \int_0^x \frac{\partial}{\partial z^j}F(\eta,z)\,\mathrm{d}\eta
\end{aligned}$$

$$- \int_0^x \frac{\partial^2}{\partial z^j \partial z^i} F(\eta, z)\, \mathrm{d}\eta \Bigg) G(x,z) \quad \text{(A.2)}$$

and

$$\begin{aligned} D_x G(x,z) &= D_x \exp\left(- \int_0^x F(\eta,z)\, \mathrm{d}\eta \right) \\ &= \exp\left(- \int_0^x F(\eta,z)\, \mathrm{d}\eta \right) D_x \left(- \int_0^x F(\eta,z)\, \mathrm{d}\eta \right) \\ &= -F(x,z) G(x,z). \end{aligned} \quad \text{(A.3)}$$

A.2 Forward rates

The Nelson-Siegel forward rates $F(x,z)$ are given by equation (3.1):

$$F(x,z) = z_1 + z_2 e^{-z_4 x} + z_3 z_4 x e^{-z_4 x}.$$

The (partial) derivatives are then given by:

$$\begin{aligned} D_x F(x,z) &= -z_2 z_4 e^{-z_4 x} - z_3 z_4^2 x e^{-z_4 x} + z_3 z_4 e^{-z_4 x}, \\ \nabla_z F(x,z) &= \left(1, e^{-z_4 x}, z_4 x e^{-z_4 x}, -z_2 x e^{-z_4 x} - z_3 z_4 x^2 e^{-z_4 x} + z_3 x e^{-z_4 x}\right), \\ \frac{\partial^2 F(x,z)}{\partial z_i \partial z_j} &= 0 \text{ for } 1 \leq i,j \leq 3, \\ \frac{\partial}{\partial z_4} \nabla_z F(x,z) &= \Big(0, -x e^{-z_4 x}, x e^{-z_4 x} - x^2 z_4 e^{-z_4 x}, \\ &\qquad z_2 x^2 e^{-z_4 x} - z_3 x^2 e^{-z_4 x} + z_3 z_4 x^3 e^{-z_4 x} - z_3 x^2 e^{-z_4 x}\Big). \end{aligned}$$

Remark that $\frac{\partial^2 F(x,z)}{\partial z_i \partial z_j} = \frac{\partial^2 F(x,z)}{\partial z_j \partial z_i}$ for all $1 \leq i \leq 4$, hence

$$\frac{\partial^2 F(x,z)}{\partial z_i \partial z_4} = \left(\frac{\partial}{\partial z_4} \nabla_z F(x,z) \right)_i .$$

Because $\int_0^x e^{-z_4 \eta}\, \mathrm{d}\eta = \frac{1}{z_4}\left(1 - e^{-z_4 x}\right)$, $\int_0^x \eta e^{-z_4 \eta}\, \mathrm{d}\eta = \frac{1}{z_4^2} - (\frac{x}{z_4} + \frac{1}{z_4^2}) e^{-z_4 x}$ and $\int_0^x \eta^2 e^{-z_4 \eta}\, \mathrm{d}\eta = \frac{2}{z_4^3} - (\frac{x^2}{z_4} + \frac{2x}{z_4^2} + \frac{2}{z_4^3}) e^{-z_4 x}$, we get

$$\begin{aligned} \int_0^x \frac{\partial F(\eta,z)}{\partial z_1}\, \mathrm{d}\eta &= x, \\ \int_0^x \frac{\partial F(\eta,z)}{\partial z_2}\, \mathrm{d}\eta &= \frac{1}{z_4}\left(1 - e^{-z_4 x}\right), \\ \int_0^x \frac{\partial F(\eta,z)}{\partial z_3}\, \mathrm{d}\eta &= z_4 \left(\frac{1}{z_4^2} - (\frac{x}{z_4} + \frac{1}{z_4^2}) e^{-z_4 x}\right) = \frac{1}{z_4} - (x + \frac{1}{z_4}) e^{-z_4 x}, \\ \int_0^x \frac{\partial F(\eta,z)}{\partial z_4}\, \mathrm{d}\eta &= (z_3 - z_2) \left(\frac{1}{z_4^2} - \left(\frac{x}{z_4} + \frac{1}{z_4^2} \right) e^{-z_4 x} \right) \end{aligned}$$

$$
\begin{aligned}
&- z_3 z_4 \left(\frac{2}{z_4^3} - \left(\frac{x^2}{z_4} + \frac{2x}{z_4^2} + \frac{2}{z_4^3} \right) e^{-z_4 x} \right), \\
&= (z_3 - z_2) \left(\frac{1}{z_4^2} - \left(\frac{x}{z_4} + \frac{1}{z_4^2} \right) e^{-z_4 x} \right) \\
&- z_3 \left(\frac{2}{z_4^2} - \left(x^2 + \frac{2x}{z_4} + \frac{2}{z_4^2} \right) e^{-z_4 x} \right).
\end{aligned}
$$

Using these facts and combining terms, one can show Corollary 3.3.

Appendix B

Quadratic (co)variation

The quadratic covariation of two functions $f(s)$ and $g(s)$ which are defined for $0 \leq s \leq t$ up to time t is given by

$$[f, g]_t = \lim_{||\Pi|| \to 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)][g(t_{j+1}) - g(t_j)], \tag{B.1}$$

where Π is a partition of $[0, t]$, $\Pi = \{t_0, t_1, \ldots, t_n\}$ with $0 = t_0 < t_1 < \ldots < t_n = t$, and $||\Pi|| = \max_{j=0,\ldots,n-1}(t_{j+1} - t_j)$. The limit $||\Pi|| \to 0$ is taken as n, the points in the partition, going to infinity while the length of the longest subinterval $t_{j+1} - t_j$ goes to zero.

The quadratic variation is simply given by equation (B.1) with $g(t) = f(t)$:

$$[f, f]_t = \lim_{||\Pi|| \to 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2. \tag{B.2}$$

Sometimes the differential form of equations (B.1) and (B.2) are informally written as:

$$\mathrm{d}[f, g]_t = \mathrm{d}f(t)\,\mathrm{d}g(t), \quad \mathrm{d}[f, f](t) = \mathrm{d}f(t)\,\mathrm{d}f(t).$$

The following identities are widely known and can for example be found in 'Stochastic Calculus for Finance II' [20] (including the proofs). All identities including Brownian motion hold almost sure.

- Let $f(t) = t$, then $[f, f]_t = 0$. In the informal notation above, this tells us $\mathrm{d}t\,\mathrm{d}t = 0$.
- Let $f(t) = W_t$ be Brownian motion. Then $[W, W]_t = t$.
- Let W_t^1 and W_t^2 be independent Brownian motion, then $[W^i, W^j]_t = t\mathbf{1}_{ij}$. Informally: $\mathrm{d}[W^i, W^j]_t = \mathrm{d}W_t^i\,\mathrm{d}W_t^j = \mathbf{1}_{ij}\,\mathrm{d}t$.

- Let $f(t) = t$ and $g(t) = W_t$, then $[f, g]_t = [t, W]_t = 0$. Informally: $\mathrm{d}[t, W]_t = \mathrm{d}t\,\mathrm{d}W_t = 0$.

- The quadratic variation of an Itô process of the form

$$X_t = X_0 + \int_0^t b_s\,\mathrm{d}s + \int_0^t \sigma_s\,\mathrm{d}W_s,$$

  where W is 1-dimensional Brownian motion, X_0 is nonrandom (i.e. $\mathcal{F}_0$-measurable) and with b and σ adapted processes with values in $\mathbb{R}$, respectively $\mathbb{R}$, such that $\int_0^t |b_s|\,\mathrm{d}s < \infty$ and $\int_0^t |\sigma_s|^2\,\mathrm{d}s < \infty$, $\mathbb{P}$-*a.s.*, for all $t > 0$, is given by

$$[X, X]_t = \int_0^t (\sigma_s)^2\,\mathrm{d}s.$$

  In informal notation: $\mathrm{d}X_t\,\mathrm{d}X_t = (\sigma_t)^2\,\mathrm{d}t$.

  Remark that this result can be easily found by using the informal notation and the bullets above:

$$\begin{aligned}\mathrm{d}[X, X]_t &= \mathrm{d}X_t\,\mathrm{d}X_t = (b_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t)\,(b_t\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t)\\ &= b_t^2\,\mathrm{d}t\,\mathrm{d}t + 2b_t\sigma_t\,\mathrm{d}t\,\mathrm{d}W_t + (\sigma_t)^2\,\mathrm{d}W_t\,\mathrm{d}W_t\\ &= 0 + 0 + \sigma_t^2\,\mathrm{d}t = (\sigma_t)^2\,\mathrm{d}t.\end{aligned}$$

- Let X_t^1 and X_t^2 be two Itô processes:

$$\begin{aligned}X_t^1 &= X_0^1 + \int_0^t b_s^1\,\mathrm{d}s + \int_0^t \sigma_s^1\,\mathrm{d}W_s,\\ X_t^2 &= X_0^2 + \int_0^t b_s^2\,\mathrm{d}s + \int_0^t \sigma_s^2\,\mathrm{d}W_s,\end{aligned}$$

  where b_s^1, b_s^2 and σ_s^1, σ_s^2 satisfy the same conditions as above. Then the quadratic covariation is given by

$$[X^1, X^2]_t = \int_0^t \sigma_s^1\sigma_s^2\,\mathrm{d}s.$$

  Informally, this is written as: $\mathrm{d}X_t^1\,\mathrm{d}X_t^2 = \sigma_t^1\sigma_t^2\,\mathrm{d}t$.

- Let X_t^1 and X_t^2 be two Itô processes driven by two independent Brownian motions W_t^1 and W_t^2:

$$\begin{aligned}X_t^1 &= X_0^1 + \int_0^t b_s^1\,\mathrm{d}s + \int_0^t \sigma_s^{11}\,\mathrm{d}W_s^1 + \int_0^t \sigma_s^{12}\,\mathrm{d}W_s^2,\\ X_t^2 &= X_0^2 + \int_0^t b_s^2\,\mathrm{d}s + \int_0^t \sigma_s^{21}\,\mathrm{d}W_s^1 + \int_0^t \sigma_s^{22}\,\mathrm{d}W_s^2,\end{aligned}$$

where b_s^i and σ_s^{ij}, $i, j = 1, 2$, satisfy the same conditions as above. Then the quadratic covariation is given by

$$[X^k, X^j]_t = \int_0^t \sum_{i=1}^{2} \sigma_s^{ki} \sigma_s^{ji} \, \mathrm{d}s.$$

Informally, this is written as: $\mathrm{d}X_t^k \, \mathrm{d}X_t^j = \sum_{i=1}^{2} \sigma_t^{ki} \sigma_t^{ji} \, \mathrm{d}t$.

- Let $X_t^1, \ldots, X_t^l$ be l Itô processes driven by d-dimensional Brownian motion. Then

$$[X^k, X^j]_t = \int_0^t \sum_{i=1}^{d} \sigma_s^{ki} \sigma_s^{ji} \, \mathrm{d}s.$$

Informally, this is written as: $\mathrm{d}X_t^k \, \mathrm{d}X_t^j = \sum_{i=1}^{d} \sigma_t^{ki} \sigma_t^{ji} \, \mathrm{d}t$.

- Let $f(t) = J_t^1$ and $g(t) = J_t^2$ be two independent right continuous pure jump processes. Then

$$[J^1, J^2]_t = \sum_{0 < s \leq t} \Delta J_s^1 \Delta J_s^2.$$

- Add to the Itô processes $X_t^1, \ldots, X_t^l$ above the right continuous, pure jump processes $J_t^1, \ldots, J_t^l$ to get $Y_t^1 = X_t^1 + J_t^1, \ldots, Y_t^l = X_t^l + J_t^l$. Then

$$[Y^k, Y^j] = \int_0^t \sum_{i=1}^{d} \sigma_s^{ki} \sigma_s^{ji} \, \mathrm{d}s + \sum_{0 < s \leq t} \Delta J_t^k \Delta J_t^j,$$

and

$$[Y^k, Y^j] = [X^k, X^j] + [J^k, J^j],$$

and

$$[X^k, J^j] = [X^j, Y^k] = 0.$$

In informal notation: $\mathrm{d}Y_t^k \, \mathrm{d}Y_t^j = \mathrm{d}X_t^k \, \mathrm{d}X_t^j + \mathrm{d}J_t^k \, \mathrm{d}J_t^j$ and $\mathrm{d}X_t^k \, \mathrm{d}J_t^j = \mathrm{d}X_t^j \, \mathrm{d}J_t^k = 0$.

Appendix C

Exponential polynomials

Lemma C.1. *From*

$$0 = p_0(x) + p_1(x)e^{-x}, \quad \forall x \geq 0, \tag{C.1}$$

where $p_0(x)$ is a polynomial of degree q_0 and $p_1(x)$ is a polynomial of degree q_1, we conclude

$$p_0(x) = 0 \text{ and } p_1(x) = 0, \quad \forall x \geq 0.$$

Proof. Write $p_0(x) = \sum_{k=0}^{q_0} a_k x^k$ and $p_1(x) = \sum_{k=0}^{q_1} b_k x^k$ where q_0 and q_1 are the orders of p_0 and p_1 respectively. We know $e^{-x} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} x^k$. Define $c_k := \frac{(-1)^k}{k!}$. Now (C.1) can be written as

$$\begin{aligned} 0 &= \sum_{k=0}^{q_0} a_k x^k + \sum_{k=0}^{q_1} b_k x^k \sum_{k=0}^{\infty} c_k x^k \\ &= \sum_{k=0}^{\infty} \left(a_k + \sum_{j=0}^{k} b_j c_{k-j} \right) x^k, \end{aligned} \tag{C.2}$$

with $a_k = 0$ for $k > q_0$ and $b_j = 0$ for $j > q_1$. Equation (C.2) to hold, yields $a_k + \sum_{j=0}^{k} b_j c_{k-j} = 0$ for all k. For $k > q_0$, we have $a_k = 0$ hence this implies $\sum_{j=0}^{k} b_j c_{k-j} = 0$. Because $b_j = 0$ for $j > q_1$ we have

$$\sum_{j=0}^{q_1 \wedge k} b_j c_{k-j} = 0, \quad \forall k > q_0.$$

Because $q_0+q_1 > q_0 \vee q_1$ we can make the following system of q_1+1 equations:

$$\sum_{j=0}^{q_1} b_j c_{q_0+q_1-j} = 0,$$
$$\sum_{j=0}^{q_1} b_j c_{q_0+q_1-j+1} = 0,$$
$$\vdots$$
$$\sum_{j=0}^{q_1} b_j c_{q_0+q_1-j+q_1} = 0.$$

This can be written in matrix form as $\mathbf{Cb} = \mathbf{0}$ where $\mathbf{b} = (b_0, b_1, \ldots, b_{q_1})^*$ and

$$\mathbf{C} = \begin{pmatrix} c_{q_0+q_1} & c_{q_0+q_1-1} & \cdots & c_{q_0+q_1-q_1} \\ c_{q_0+q_1+1} & c_{q_0+q_1} & \cdots & c_{q_0+q_1-q_1} \\ \vdots & & & \vdots \\ c_{q_0+q_1+q_1} & c_{q_0+q_1+q_1-1} & \cdots & c_{q_0+q_1} \end{pmatrix}.$$

Furthermore, we know that the homogeneous equation $\mathbf{Cb} = \mathbf{0}$ has a non-trivial solution if det($\mathbf{C}$)=0. If det($\mathbf{C}$)=0, there exists no inverse of $\mathbf{C}$. But MATLAB shows that $\mathbf{C}$ has an inverse, so the only solution to $\mathbf{Cb} = \mathbf{0}$ is $\mathbf{b} = \mathbf{0}$, hence $p_1(x) = 0$. The inverse of $\mathbf{C}$ is not easy to determine. For given q_0 and q_1, one can of course get the inverse with MATLAB and check that it is the inverse. Now (C.1) immediately gives $p_0(x) = 0, \forall x \geq 0$, hence $a_k = 0, \forall k$. □

Lemma C.2. *From*

$$0 = p_0(x) + p_1(x)e^{-x} + p_2(x)e^{-2x}, \quad \forall x \geq 0, \tag{C.3}$$

where $p_0(x)$ is a polynomial of degree q_0, $p_1(x)$ is a polynomial of degree q_1 and $p_2(x)$ is a polynomial of degree q_2, we conclude

$$p_0(x) = 0, p_1(x) = 0 \text{ and } p_2(x) = 0, \quad \forall x \geq 0.$$

Proof. Multiply (C.3) by $e^x > 0$ to get

$$0 = p_0(x)e^x + p_1(x) + p_2(x)e^{-x}, \quad \forall x \geq 0.$$

Using the same notation as in Lemma C.1 , we can write

$$\begin{aligned} 0 &= \sum_{k=0}^{q_0} a_k x^k \sum_{k=0}^{\infty} d_k x^k + \sum_{k=0}^{q_1} b_k x^k + \sum_{k=0}^{q_2} c_k x^k \sum_{k=0}^{\infty} f_k x^k \\ &= \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} a_j d_{k-j} + b_k + \sum_{j=0}^{k} c_j f_{k-j} \right) x^k, \end{aligned} \tag{C.4}$$

with $d_k = \frac{1^k}{k!}$ and $f_k = \frac{(-1)^k}{k!}$. Equation (C.4) gives

$$\sum_{j=0}^{k} a_j d_{k-j} + b_k + \sum_{j=0}^{k} c_j f_{k-j} = 0, \quad \forall k \geq 0. \tag{C.5}$$

When we take $k > q_0 \vee q_1 \vee q_2$, we have $b_k = 0$ and

$$\sum_{j=0}^{q_0} a_j d_{k-j} + \sum_{j=0}^{q_2} c_j f_{k-j} = 0, \quad \forall k > q_0 \vee q_1 \vee q_2. \tag{C.6}$$

To determine the $q_0 + 1 + q_2 + 1$ coefficients of $p_0(x)$ and $p_2(x)$, we need $q_0 + 1 + q_2 + 1$ equations. Because $q_0 + q_1 + q_2 > q_0 \vee q_1 \vee q_2$ we can take the following equations:

$$\sum_{j=0}^{q_0} a_j d_{q_0+q_1+q_2-j} + \sum_{j=0}^{q_2} c_j f_{q_0+q_1+q_2-j} = 0,$$

$$\vdots$$

$$\sum_{j=0}^{q_0} a_j d_{2q_0+q_1+2q_2+2-j} + \sum_{j=0}^{q_2} c_j f_{2q_0+q_1+2q_2+2-j} = 0.$$

This system can be written as $\mathbf{Cb} = \mathbf{0}$ with

$$\mathbf{C} = \begin{pmatrix} d_{q_0+q_1+q_2} & \cdots & d_{q_1+q_2} & f_{q_0+q_1+q_2} & \cdots & f_{q_0+q_1} \\ d_{q_0+q_1+q_2+1} & \cdots & d_{q_1+q_2+1} & f_{q_0+q_1+q_2+1} & \cdots & f_{q_0+q_1+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ d_{2q_0+q_1+2q_2+1} & \cdots & d_{q_0+q_1+2q_2+1} & f_{2q_0+q_1+2q_2+1} & \cdots & f_{2q_0+q_1+q_2+1} \end{pmatrix}$$

and

$$\mathbf{b} = \left(a_0, \ldots, a_{q_0}, c_0, \ldots, c_{q_2}\right)^*.$$

Again, we know that the homogeneous equation $\mathbf{Cb} = \mathbf{0}$ has a non- trivial solution if det($\mathbf{C}$)=0. If det($\mathbf{C}$)=0, there exists no inverse of $\mathbf{C}$. But MATLAB shows that $\mathbf{C}$ has an inverse, so the only solution to $\mathbf{Cb} = \mathbf{0}$ is $\mathbf{b} = \mathbf{0}$, hence $p_0(x) = 0$ and $p_2(x) = 0$. The inverse of $\mathbf{C}$ is not easy to determine. For given q_0 and q_2, one can of course get the inverse with MATLAB and check that it is the inverse. Now (C.3) immediately gives $p_1(x) = 0$, for $x \geq 0$. □

Appendix D

Appendix - Concluding remarks

D.1 Proof of the theorem of the Concluding remarks

To prove the theorem in the Concluding remarks, Chapter 5, we mimic the proof of Theorem 4.6. The only difference is the jump part of the process under consideration.

Suppose we have a process according to Definition 4.3, where the jump part is given by $J = (J^1, J^2, J^3, 0)$, $J_t^i = \sum_{j=1}^m \mathbf{\Sigma}^{ij} N_t^j$ with $N_t = (N_t^1, \ldots, N_t^m)$ is an m-dimensional vector with independent Poisson processes with parameters $\lambda_1, \ldots, \lambda_m > 0$ and $\mathbf{\Sigma} \in \mathbb{R}^{4\times m}$.

Let $P(t,T)$ be the bond prize implied by the Nelson-Siegel forward rate (3.1), so

$$P(t,T) = \exp\left[-(T-t)\mathbf{Z}_t^* \mathbf{W}_{T-t}(Z_t^4)\right],$$

where

$$\mathbf{Z}_t := \begin{pmatrix} Z_t^1 \\ Z_t^2 \\ Z_t^3 \end{pmatrix}, \mathbf{W}_t(\lambda) := \begin{pmatrix} \mathbf{W}_t^1(\lambda) \\ \mathbf{W}_t^2(\lambda) \\ \mathbf{W}_t^3(\lambda) \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1-e^{-\lambda t}}{\lambda t} \\ \frac{1-e^{-\lambda t}}{\lambda t} - e^{-\lambda t} \end{pmatrix}.$$

Now we are interested in investigating the jump part of (4.18):

$$\sum_{0<s\leq t} \left[P(s,T) - P(s-,T)\right] = \sum_{0<s\leq t} \sum_{j=1}^m \Delta_j P(s,T).$$

When there is a jump at time s in N_t^k we know $Z_s^i = Z_{s-}^i + \Delta Z_s^i = Z_{s-}^i + \mathbf{\Sigma}^{ik} \Delta N_s^k = Z_{s-}^i + \mathbf{\Sigma}^{ik}$. Denote by $\mathbf{e}_k$ the k-th unit vector, then $\mathbf{Z}_{s-} =$

$\mathbf{Z}_s - \mathbf{\Sigma e}_k$ when there is a jump of N^k at time s. Now

$$\begin{aligned}
\Delta_k P(s,T) &= P(s,T) - P(s-,T) \\
&= \exp\left[-(T-s)\mathbf{Z}_s^*\mathbf{W}_{T-s}(Z_s^4)\right] - \exp\left[-(T-s)\mathbf{Z}_{s-}^*\mathbf{W}_{T-s}(Z_s^4)\right] \\
&= \exp\left[-(T-s)\mathbf{Z}_s^*\mathbf{W}_{T-s}(Z_s^4)\right] - \exp\left[-(T-s)(\mathbf{Z}_s^* - \mathbf{\Sigma e}_k)^*\mathbf{W}_{T-s}(Z_s^4)\right] \\
&= P(s,T)\left(1 - \exp\left[(T-s)\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}_{T-s}^j(Z_s^4)\right]\right)
\end{aligned}$$

when there is a jump of process N^k at time s, else $\Delta_k P(s,T) = 0$. Hence

$$\Delta_k P(s,T) = P(s,T)\left(1 - \exp\left[(T-s)\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}_{T-s}^j(Z_s^4)\right]\right)\Delta N_k(s).$$

Now we get

$$\begin{aligned}
\sum_{0<s\leq t}\left[P(s,T) - P(s-,T)\right] &= \sum_{0<s\leq t}\sum_{k=1}^{m}\Delta_k P(s,T) \\
&= \sum_{0<s\leq t}\sum_{k=1}^{m} P(s,T)\left(1 - \exp\left[(T-s)\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}_{T-s}^j(Z_s^4)\right]\right)\Delta N_k(s).
\end{aligned}$$

Remark that $\sum_{0<s\leq t} f(s)\Delta N(s) = \int_0^t f(s)\,\mathrm{d}N(s)$ hence we can write

$$\begin{aligned}
&\sum_{0<s\leq t}\left[P(s,T) - P(s-,T)\right] \\
&\qquad = \sum_{k=1}^{m}\int_0^t P(s,T)\left(1 - e^{(T-s)\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}_{T-s}^j(Z_s^4)}\right)\mathrm{d}N_k(s).
\end{aligned}$$

Equation (4.18) now becomes:

$$\begin{aligned}
P(t,T) &= P(0,T) + \int_0^t (A_s G(T-s,Z_s) - D_x G(T-s,Z_t))\,\mathrm{d}s \\
&+ \int_0^t \nabla_z G(T-s,Z_s)^*\sigma_s\,\mathrm{d}W_s \\
&+ \sum_{k=1}^{m}\int_0^t P(s,T)\left(1 - e^{(T-s)\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}_{T-s}^j(Z_s^4)}\right)\mathrm{d}N_k(s),\ 0\leq t\leq T,\ \mathbb{P}\text{-}a.s..
\end{aligned}$$

In differential notation:

$$\begin{aligned}
dP(t,T) &= (A_t G(T-t,Z_t) - D_x G(T-t,Z_t))\,\mathrm{d}t + \nabla_z G(T-t,Z_t)^*\sigma_t\,\mathrm{d}W_t \\
&+ \sum_{k=1}^{m} P(t,T)\left(1 - e^{(T-t)\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}_{T-s}^j(Z_s^4)}\right)\mathrm{d}N_k(t),\ 0\leq t\leq T,\ \mathbb{P}\text{-}a.s..
\end{aligned}$$

And this is equal to

$$\mathrm{d}P(t,T) = (A_t G(T-t,Z_t) - D_x G(T-t,Z_t))\,\mathrm{d}t + \nabla_z G(T-t,Z_t)^* \sigma_t \,\mathrm{d}W_t$$
$$+ \sum_{k=1}^{m} G(T-t,Z_t)\left(1 - e^{(T-t)\sum_{j=1}^{3} \mathbf{\Sigma}^{jk}\mathbf{W}^j_{T-s}(Z^4_s)}\right)\mathrm{d}N_k(t),\ \ 0 \le t \le T,\ \mathbb{P}\text{-}a.s..$$

Following the lines in the proof of Theorem 4.6 with $1 - e^{(T-s)\alpha_k \mathbf{W}^k_{T-s}(Z^4)}$ replaced by $1 - e^{(T-t)\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_{T-s}(Z^4_s)}$ gives the desired result. □

Remark D.1. *Theorem 4.6 is a special case of the theorem in Chapter 5. It is the case that $\mathbf{\Sigma}^{kk} = \alpha_k$ for $k = 1,2,3$, and $\mathbf{\Sigma}^{kj} = 0$ for $k \neq j$. The sum $\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_{T-s}(Z^4_s)$ will disappear and only $\alpha_k \mathbf{W}^k_{T-s}(Z^4_s)$ is left.*

D.2 Proof of the corollary of the Concluding remarks

To prove the corollary in the Concluding remarks, Chapter 5, we mimic the proof of Corollary 4.8. The only difference is the jump part of the process under consideration.

Remember

$$D_x\left(x\mathbf{W}_x(\lambda)\right) = \begin{pmatrix} 1 \\ e^{-\lambda x} \\ \lambda x e^{-\lambda x}\end{pmatrix} = \mathbf{w}_x(\lambda),$$

hence

$$D_x\left[x\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_x(\lambda)\right] = \sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{w}^j_x(\lambda).$$

Now

$$-D_x\sum_{k=1}^{m}\left(1 - e^{x\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_x(z^4)}\right)\lambda_k = \sum_{k=1}^{m} D_x\left[e^{x\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_x(z^4)}\lambda_k\right]$$
$$= \sum_{k=1}^{m}\left[e^{x\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_x(z^4)} D_x\left[x\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_x(z^4)\right]\lambda_k\right]$$
$$= \sum_{k=1}^{m}\left[e^{x\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{W}^j_x(z^4)}\sum_{j=1}^{3}\mathbf{\Sigma}^{jk}\mathbf{w}^j_x(z^4)\lambda_k\right].$$

Adding this to the result of Corollary 3.3 gives the corollary in the Concluding remarks. □

Remark D.2. *Remark that Corollary 4.8 is (of course) a special case of the corollary above. It is the case that* $\mathbf{\Sigma}^{kk} = \alpha_k$ *for* $k = 1, 2, 3,$ *and* $\mathbf{\Sigma}^{kj} = 0$ *for* $k \neq j$*.*

If we write out the result above we get

$$-D_x \sum_{k=1}^{m} \left(1 - e^{x \sum_{j=1}^{3} \mathbf{\Sigma}^{jk} \mathbf{W}_x^j(z^4)}\right) \lambda_k = \sum_{k=1}^{m} \left[e^{x \sum_{j=1}^{3} \mathbf{\Sigma}^{jk} \mathbf{W}_x^j(z^4)} \sum_{j=1}^{3} \mathbf{\Sigma}^{jk} \mathbf{w}_x^j(z^4) \lambda_k \right]$$

$$= \sum_{k=1}^{m} \left[e^{x\left(\mathbf{\Sigma}^{1k} + \mathbf{\Sigma}^{2k} \frac{1-e^{-z^4 t}}{z^4 t} + \mathbf{\Sigma}^{3k} \left(\frac{1-e^{-z^4 t}}{z^4 t} - e^{-z^4 t}\right)\right)} \right] \left(\mathbf{\Sigma}^{1k} + \mathbf{\Sigma}^{2k} e^{-z^4 t} + \mathbf{\Sigma}^{3k} z^4 t e^{-z^4 t}\right).$$

This equation has $3m - 3$ more unknowns than the equation in the Independent jump case.