FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Thesis for the Master degree in Insurance Mathematics.
Department of Mathematical Sciences, University of Copenhagen

Speciale for cand.act graden i Forsikringsmatematik.
Institut for Matematiske Fag, Københavns Universitet

Affine Processes in Life Insurance Mathematics
– With a View Towards Solvency II

Kristian Buchardt

Mogens Steffensen, University of Copenhagen
Thomas Møller, PFA Pension

September 1, 2011

Post hand-in version with minor typos corrected.
Version date: September 11, 2011

Abstract

Affine processes have been of great interest in financial mathematics, and with the introduction of Solvency II in life insurance, the need for stochastic modelling of the valuation basis has led to the entry of affine processes into the literature on life insurance mathematics. The thesis begins with a qualitative discussion of Solvency II and the implications for life insurance liabilities. The thesis then considers multidimensional time-inhomogeneous continuous affine processes which are subsequently applied in the valuation of life insurance liabilities, providing a foundation for stochastic modelling of the valuation basis. This leads to the introduction of generalised forward rates, and a theorem is proven that allows for the use of dependent interest and transition rates in e.g. a survival model or a model including the surrender option. Using the theory of affine processes it is also shown how one can model the mortality rate using a CIR process, such that the distribution of the forward rate can be found, allowing for simple simulation of the Solvency II capital requirement. It is shown that the distributional result also holds for certain time-inhomogeneous CIR processes.

Resumé

Affine processer har været af stor betydning i finansieringsmatematik, og med Solvens II har behovet for stokastisk modellering af opgørelsesgrundlaget ledt til introduktionen af affine processer i livsforsikringslitteraturen. Specialet begynder med en kvalitativ diskussion af Solvens II og følgerne ved opgørelse af livsforsikringsforpligtelser. Derefter betragtes multidimensionelle tidsinhomogene kontinuerte affine processer og deres anvendelser ved modellering af opgørelsesgrundlag for livsforsikringsforpligtelser. Dette fører til introduktionen af såkaldte generaliserede forward rater, og en sætning bevises der tillader opstilling modeller med afhængighed mellem rente og overgangsrater, f.eks. en overlevelsesmodel eller en model med genkøbsmodellering. Teorien for affine processer giver mulighed for modellering af dødelighedsraten ved anvendelse af en CIR process, og det vises hvordan forward dødeligheden er en transformation af en kendt fordeling. Simpel simulation af Solvens II kapitalkravet er derved mulig. Det vises at fordelingsresultatet for CIR processen kan udvides til visse tids-inhomogene CIR processer.

Contents

Introduction

Valuation of liabilities in life insurance is a classic problem, with a well-established theory once a valuation basis has been chosen. It has always been a central problem how to choose the valuation basis such that it is on the safe side, starting with the theorem of Lidstone (1905) [13], and how to assess the sensitivity of the liabilities with respect to changes in the valuation basis. With the introduction of Solvency II these questions have gained new interest. Recent research on these topics include, for sensitivity analysis, Kalashnikov and Norberg (2003) [12] and Christiansen (2008) [3], and for worst-case scenarios, Christiansen (2010) [4] and Christiansen and Steffensen (2011) [5].

In Solvency II the focus is essentially on the loss over a 1-year time horizon. Hence, the focus is not on the distribution of the present value of the future payments, but the distribution of the change in expectation over a 1-year time horizon. A proper model for life insurance liabilities requires stochastic modelling of the valuation basis, that is, a stochastic model for the interest and transition rates. Börger (2010) [2] carries out such a study quantitatively for mortality risk, and compares the findings with those of the standard model of Solvency II. In Section 1 of this thesis, a study of the mortality risk is carried out on a qualitative scale, where the mortality rate is modelled as a Gaussian process. Simple models with closed form results for the loss are considered, and a simple comparison with the standard model is carried out.

For more sophisticated modelling, stochastic modelling of both interest rates and mortality rates have been proposed in the literature, see e.g. Dahl and Møller (2006) [6]. This is done with specific affine processes for the interest and mortality rate, which are assumed to be independent, and especially the time-inhomogeneous Cox-Ingersol-Ross model has been applied. Affine processes are interesting because of the simple form of the characteristic function, and they have been studied and applied in financial mathematics before their introduction to life insurance mathematics. A general mathematical treatment of continuous time-homogeneous affine processes is given in Filipovic and Mayerhofer (2009) [8] and Filipovic (2010) [7]. In this thesis, the starting point for stochastic modelling is general multidimensional time-inhomogeneous continuous affine processes.

Stochastic modelling of the interest rate is an important discipline in financial mathematics. For the study of a stochastic interest rate, the *forward interest rate* has proven to play an essential role, and this is also the case for applications in life insurance mathematics. For the study of stochastic mortality rates in life insurance, it has been proposed (Milevsky and Promislow (2001) [15]) to copy the concept of the forward interest rate and in a similar way define the *forward mortality rate*. This has had some success, and an interesting discussion on forward mortality rates is found in Norberg (2010) [18]. In this thesis, continuous affine processes are studied and applied in life insurance. This leads to the introduction of the concept of *generalised forward rates*, where the interest and transition rates are allowed to be dependent affine processes. Earlier attempts at defining forward mortality rates for dependent interest and mortality rates have met legitimate criticism in [18]. We show that some of the proposed definitions from the literature of forward mortality rates are not well-defined, and demonstrate how the generalised forward rates introduced here seem to meet parts of the criticism. An important observation about the generalised forward rates is that the forward interest rate changes when studied together with a dependent mortality rate.

Structure of the Thesis Section 1 contains a simple qualitative discussion of Solvency II and mortality risk. Section 1.1 presents the basic setup and defines the *loss at time 1*, which is the quantity of interest when discussing the Solvency II capital requirement. In Section 1.2 we present some simplistic models setting the scene for a qualitative discussion of mortality risk. Sections 1.4 and 1.5 contain a discussion of the stress scenarios of the standard model, and compares those with results from the proposed simple models.

Section 2 is devoted to a general treatment of multidimensional time-inhomogeneous continuous affine processes. This treatment is an extension of the time-homogeneous treatment in [7]. In Sections 2.2–2.4 we find necessary and sufficient conditions on the parameters to ensure that a process is affine. In addition, the Riccati equations describing the characteristic and moment generating functions are found.

Section 3 takes as starting point a multidimensional affine process modelling the interest and/or transition rates, and quantities needed for valuation of life insurance contracts are found. Survival probabilities and discount factors are studied in Section 3.1. Section 3.2 presents the main results of the thesis, and a (to the author's knowledge) new theorem, Theorem 3.4, is proven. The result in part allows for pricing of term insurance contracts in a setting of dependent affine interest and mortality rates, and in part motivates the definition of the *generalised forward rates*. Section 3.3 contains a discussion on the relation between the generalised forward rates and earlier definitions of forward rates for dependent interest and mortality rates.

Examples of the application of continuous affine processes as valuation basis for life insurance contracs are studied in Section 4. In Section 4.1 it is demostrated how one can model the mortality such that the distribution of the forward mortality rate can be found, using that the CIR process has a known conditional distribution. Theorem 4.2 allows for an extension of the results of Maghsoodi (1996) [14] to certain time-inhomogeneous CIR processes. Generalised forward rates are applied to life insurance contracts in Section 4.3, and shows the usefulness of the concept. Modelling the interest and mortality rate as dependent affine processes, we find the prospective reserve and a Thiele differential equation where the generalised forward rates appear in the formulas in the ways one would expect. Finally, Section 4.4 shows how one can find the *loss at time 1* within the setup developed in Sections 2–4.

Acknowledgements I would like to thank my supervisors, Mogens Steffensen and Thomas Møller, for their great interest and guidance. Also, a lot of people have provided me help and support and I would like to thank the following; Ernst Hansen, Jostein Paulsen, Martin Jacobsen, Damir Filipovic, Oluf Hansen, Nina Munkholt, Lars Frederik Brandt Henriksen, Ketil Tvermosegaard and last but not least, Ann-Sophie Søgaard.

Notation

Throughout the thesis, the following notation is used.

For a complex number $z = x + \mathrm{i}y \in \mathbb{C}$ where $x, y \in \mathbb{R}$, let

$$\Re z = x, \quad \Im z = y.$$

Let $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$. Then

$$\begin{aligned} K_{+} &= \{z \in K \mid \Re z \geq 0\}, \\ K_{-} &= \{z \in K \mid \Re z \leq 0\}, \\ K_{++} &= \{z \in K \mid \Re z > 0\}, \\ K_{--} &= \{z \in K \mid \Re z < 0\}. \end{aligned}$$

For $x, y \in \mathbb{R}$, let

$$\begin{aligned} x \wedge y &= \min\{x, y\}, \\ x \vee y &= \max\{x, y\}. \end{aligned}$$

A process $(X(t))_{t \in J}$ for an interval J is denoted $\mathbf{X}$.

1 Solvency II for Life Insurance Liabilities

To measure solvency in the Solvency II regime, one needs to consider the "available capital" after one year. This is random, depending on the development of the world during the year. To be somewhat sure that one can survive another year, the solvency ruleset requires you to have enough capital, such that the probability of still having available capital after one year is greater than 99.5%. In the following it is made precise how this can be interpreted mathematically, when considering the life insurance liabilities only.

1.1 Basic Setup

Consider a classic life insurance contract, where the insurer promises a payment stream $\mathrm{d}B(t)$, i.e. $B(t)$ is the total payments by time t. (We assume that $B(t)$ is constant from a certain point of time almost surely.) Let $r(t)$ be the (possibly stochastic) instantaneous interest rate at time t. Then the present value at time t of the future guaranteed payments is

$$PV(t) = \int_t^\infty e^{-\int_t^s r(u)\,\mathrm{d}u}\,\mathrm{d}B(s).$$

Assume that the life of the insured is modelled by a Markov chain, $\mathbf{Z} = (Z(t))_{t\in[0,\infty)}$, in the finite state space $\mathcal{J}$, and denote by $\mathbf{F}^Z = (\mathcal{F}^Z(t))_{t\in[0,\infty)}$ the filtration generated by this process. The transition intensities of $\mathbf{Z}$ are defined as follows. Let $N_{jk}(t)$, with $j,k \in \mathcal{J}$, be the total number of transitions from state j to state k at time t. By [19], there exists a predictable process $\mathbf{\Lambda}_{jk}$ with respect to $\mathbf{F}^Z$ such that $N_{jk}(t) - \Lambda_{jk}(t)$ is a martingale. The process $\mathbf{\Lambda}_{jk}$ is referred to as the predictable compensator. If there exists a process $\tilde{\lambda}_{jk}(t)$ such that $\Lambda_{jk}(t) = \int_0^t \tilde{\lambda}_{jk}(s)\,\mathrm{d}s$, i.e. $\mathbf{\Lambda}_{jk}$ has density $\tilde{\lambda}_{jk}$ with respect to the Lebesgue measure, then $\tilde{\lambda}_{jk}$ is the transition intensity.

Note that the distribution of $\mathbf{Z}$ may be defined by the transition intensities. It is intuitively clear that one can write $\tilde{\lambda}_{jk}(t) = 1_{\{Z(t-)=j\}}\lambda_{jk}(t)$, and it is seen that $\tilde{\lambda}_{jk}$ is stochastic. In classic life insurance, the quantities λ_{jk} are not stochastic. In "modern" life insurance mathematics, one is also interested in uncertainty arising from incorrect prediction of mortality. To this end, we model λ_{jk} as a stochastic process itself, and we call this the *transition rate* (in contrast to the *transition intensity*, $\tilde{\lambda}_{jk}$). The distribution of $(N_{jk})_{j,k\in\mathcal{J}}$, and hence of $\mathbf{Z}$, is thus specified, and conditioning on the transition rates, the usual life insurance results hold true.

Let the interest rate be stochastic, and let $\mathbf{F}^Y = (\mathcal{F}^Y(t))_{t\in[0,\infty)}$ be the filtration generated by the transition rates and the interest rate. (We think of a general stochastic process $\mathbf{Y}$ that describes the economic-demographic environment including the interest rate and the transition rates.) We are thus considering the filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, P)$ where $\mathbf{F} = (\mathcal{F}(t))_{t\in[0,\infty)} = (\mathcal{F}^Z(t) \vee \mathcal{F}^Y(t))_{t\in[0,\infty)}$.

Now, the insurer puts aside capital of size $V(t) = \mathrm{E}\left[PV(t) \mid \mathcal{F}(t)\right]$ (possibly using another measure Q, say, to represent the market prices of the interest rate and transition rates). Additionally, at time 0 the insurer must have some extra capital, x_{SCR}, in order to meet the solvency capital requirement. Denote by $S(t)$ the total capital at time t, and note that at time 0,

$$S(0) = V(0) + x_{\text{SCR}}.$$

What is the economic situation after one year? After the first year, i.e. at time $t = 1$, we have paid out an amount of value $\int_0^1 e^{\int_s^1 r(u)\,\mathrm{d}u}\,\mathrm{d}B(s)$ at time 1. Gaining interest on the capital at time 0, the total capital at time 1 is then

$$S(1) = e^{\int_0^1 r(u)\,\mathrm{d}u} S(0) - \int_0^1 e^{\int_s^1 r(u)\,\mathrm{d}u}\,\mathrm{d}B(s).$$

Here, the investment of the capital in something else than the “bank account” is ignored i.e. it is interpreted as a risk on the assets and thus not treated here. The value of the liabilities at time 1 are given by $V(1)$. The Solvency II regime then requires that x_{SCR} is of a size, such that

$$P(S(1) - V(1) < 0) \leq 0.5\%,$$

and we are interested in the smallest x_{SCR} such that this holds true. Rewriting the inequality,

$$\begin{aligned}
&P(S(1) - V(1) < 0)\\
&= P\left(e^{\int_0^1 r(u)\,\mathrm{d}u} S(0) - \int_0^1 e^{\int_s^1 r(u)\,\mathrm{d}u}\,\mathrm{d}B(s) - V(1) < 0\right)\\
&= P\left(e^{-\int_0^1 r(u)\,\mathrm{d}u}\,\mathrm{E}\left[PV(1) \mid \mathcal{F}(1)\right] + \int_0^1 e^{-\int_0^s r(u)\,\mathrm{d}u}\,\mathrm{d}B(s) - V(0) > x_{\text{SCR}}\right)\\
&= P\left(\mathrm{E}\left[PV(0) \mid \mathcal{F}(1)\right] - \mathrm{E}\left[PV(0) \mid \mathcal{F}(0)\right] > x_{\text{SCR}}\right).
\end{aligned}$$

We define the *loss at time 1* evaluated at time 0 by the stochastic variable,

$$\widetilde{L}(1) = \mathrm{E}\left[PV(0) \mid \mathcal{F}(1)\right] - \mathrm{E}\left[PV(0) \mid \mathcal{F}(0)\right].$$

Then the Solvency capital requirement is reduced to finding a quantile in the distribution of $\widetilde{L}(1)$.

Now $\widetilde{L}(1)$ is a difference between the reserve now, and the reserve we would have if we knew everything about what would happen during the following year, i.e. all individual deaths and the development of the economic-demographic environment. It can also be interpreted in another way, namely as the sum of two risks that we will call the "unsystematic" risk and the "systematic" risk. Consider $\widetilde{L}(1)$, and note that

$$\begin{aligned}\widetilde{L}(1) = &\mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(1) \vee \mathcal{F}^Y(1)\right] - \mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(1)\right] \\ &+ \mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(1)\right] - \mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(0)\right].\end{aligned} \tag{1.1}$$

Consider the first line of (1.1). This can be written as

$$\mathrm{E}\left[PV(0) \mid \mathcal{F}(1)\right] - \mathrm{E}\left[\mathrm{E}\left[PV(0) \mid \mathcal{F}(1)\right] \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(1)\right].$$

An interpretation is that given information about the intensities until time 1, the difference of the actual behavior of $\mathbf{Z}$ from the expected behavior of $\mathbf{Z}$ is measured. In a large portfolio of identical contracts on different lifes (i.e. contracts that conditional on $\mathcal{F}^Y(\infty)$ are *iid*), this difference becomes relatively smaller, as the law of large numbers can be applied: Let n be the number of contracts and let $PV_i(0)$ be the present value of contract i. Taking average and applying the law of large numbers, the first part

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[PV_i(0) \mid \mathcal{F}(1)\right]$$

converges to the expectation, $\mathrm{E}\left[\mathrm{E}\left[PV(0) \mid \mathcal{F}(1)\right] \mid \mathcal{F}^Y(1)\right]$ which is the last part. Thus, by the usual law of large number argument the insurance companies build their business upon, this risk is manageable. This risk is often referred to as "unsystematic" risk because it affects contracts differently.

Consider now the second line of (1.1). This can be written as

$$\mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(1)\right] - \mathrm{E}\left[\mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(1)\right] \mid \mathcal{F}(0)\right].$$

The same argument cannot be applied here: If we should do that, we would need n *worlds* with different realisations of the economic-demographic environment, $\mathbf{Y}$. In other words, this difference cannot be "averaged out". Instead, this difference can be interpreted as the loss arising when updating the interest and transition rates with the new knowledge gained from year 0 to year 1. This affects all contracts in the same way, and this risk is often referred to as "systematic" risk.

In the Solvency II regime one considers only the systematic risk.

Definition 1.1. *The loss at time t arising from the systematic risk is given by,*

$$L(t) = \mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(t)\right] - \mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(0) \vee \mathcal{F}^Y(0)\right]. \tag{1.2}$$

1.2 Naive Models

This section contains some simple and idealistic models where some interesting points can be observed. The calculations are carried out with elementary methods, some of which would be considerably simpler applying Itô calculus.

Example 1.2. Consider a so-called pure endowment for an x-year old with remaining life-time T_x. The payment function is $B(t) = 1_{\{T_x \geq n\}} 1_{\{t \geq n\}}$ and the present value is

$$PV(t) = e^{-\int_t^n r(u)\,\mathrm{d}u} 1_{\{T_x \geq n\}}.$$

Assume that the interest rate r is deterministic. First note that

$$\mathrm{E}\left[PV(0) \mid \mathcal{F}^Z(s) \vee \mathcal{F}^Y(t)\right] = 1_{\{T_x \geq s\}} \cdot {}_{n-s}p_{x+s}^t \cdot e^{-\int_0^n r(u)\,\mathrm{d}u}.$$

Here, ${}_{n-s}p_{x+s}^t$ is the usual survival probability, but conditioned on knowing the economic-demographic environment until time t,

$${}_{n-s}p_{x+s}^t = P\left(T_x \geq n \mid T_x \geq s, \mathcal{F}^Y(t)\right).$$

We find,

$$\begin{aligned}
\widetilde{L}(1) &= 1_{\{T_x \geq 1\}} \cdot {}_{n-1}p_{x+1}^1 \cdot e^{-\int_0^n r(u)\,\mathrm{d}u} - {}_n p_x^1 \cdot e^{-\int_0^n r(u)\,\mathrm{d}u} \\
&\quad + {}_n p_x^1 \cdot e^{-\int_0^n r(u)\,\mathrm{d}u} - {}_n p_x^0 \cdot e^{-\int_0^n r(u)\,\mathrm{d}u} \\
&= e^{-\int_0^n r(u)\,\mathrm{d}u} {}_{n-1}p_{x+1}^1 \left(1_{\{T_x \geq 1\}} - {}_1 p_x^1\right) \\
&\quad + e^{-\int_0^n r(u)\,\mathrm{d}u} \left({}_n p_x^1 - {}_n p_x^0\right).
\end{aligned} \tag{1.3}$$

Here, the first line of (1.3) is the unsystematic risk corresponding to the first line of (1.1) and the second line of (1.3) is the systematic risk corresponding to the second line of (1.1). We see that the unsystematic risk arises from the individual deaths, and can be averaged away. The systematic risk arises from different valuation of the future liabilities because of the increased knowledge about the future mortality after one year. ○

As previously noted Solvency II focuses on the systematic risk, thus disregarding the unsystematic risk. In the above Example 1.2, it is then a matter of finding distributional results for $e^{-\int_0^n r(u)\,\mathrm{d}u}\left({}_np_x^1 - {}_np_x^0\right)$. The stochastic part here is ${}_np_x^1$, which is the probability of survival to time n, with economic-demographic information up to time 1.

An interesting point is this: Solvency II does not focus on finding a quantile for the quantity $PV(0)$ (which would be a very interesting thing indeed). Solvency II quantifies the change in liabilities arising from the development in, and change in expectations of, the interest and transition rates, because of one year's developments in the economic-demographic environment. This is reflected in $L(1)$, and it is essentially in the distribution of this quantity, that we need to find a quantile.

Before we continue Example 1.2, we show a simple result about integrals of the Brownian motion.

Lemma 1.3. *Let* $\mathbf{W} = (W(t))_{t\in\mathbb{R}_+}$ *be a standard Brownian motion. Then, for* $k \in \mathbb{R}$ *and* $t \geq 0$,

$$\int_0^t W(s)\,\mathrm{d}s + kW(t) \sim \mathrm{N}\left(0, \tfrac{1}{3}t^3 + k^2 t + kt^2\right).$$

Remark 1.4. Using the moment generating function for the Gaussian distribution, one obtains

$$\mathrm{E}\left[e^{\alpha\left(\int_0^t W(s)\,\mathrm{d}s + kW(t)\right)}\right] = e^{\frac{1}{2}\alpha^2\left(\frac{1}{3}t^3 + k^2 t + kt^2\right)}.$$

◇

Proof. Notice that the integral $\int_0^t W(s)\,\mathrm{d}s$ is well defined, since on $[0,t]$, W is a.s. continuous and in particular it is a classic Riemann integral. The characteristic function is

$$\begin{aligned}
\varphi(\alpha) &= \mathrm{E}\left[e^{i\alpha\left(\int_0^t W(s)\,\mathrm{d}s + kW(t)\right)}\right] \\
&= \mathrm{E}\left[e^{i\alpha\left(\lim_{n\to\infty}\frac{t}{n}\sum_{i=1}^n W(t\frac{i}{n}) + kW(t)\right)}\right] \\
&= \lim_{n\to\infty}\mathrm{E}\left[e^{i\alpha\left(\frac{t}{n}\sum_{i=1}^n W(t\frac{i}{n}) + kW(t)\right)}\right] \\
&= \lim_{n\to\infty} e^{-\frac{1}{2}\alpha^2\sigma_n^2},
\end{aligned}$$

where we first applied dominated convergence, and then used that a finite sum of dependent Gaussian stochastic variables is again Gaussian, and in this case, it obviously has mean 0. Now we just need to determine the variance $\sigma_\infty^2 = \lim_{n\to\infty}\sigma_n^2$. We consider the stochastic variable, and write $W_i = W(t\frac{i}{n})$,

$$\begin{aligned}
\frac{t}{n}\sum_{i=1}^{n} W_i + kW_n &= \frac{t}{n}\sum_{i=1}^{n}(1+\tfrac{nk}{t}1_{\{i=n\}})W_i \\
&= \frac{t}{n}\sum_{i=1}^{n}\left(\sum_{j=0}^{i-1}(1+\tfrac{nk}{t}1_{\{i=n\}})\left(W_{j+1}-W_j\right)\right) \\
&= \frac{t}{n}\sum_{j=0}^{n-1}\sum_{i=j+1}^{n}(1+\tfrac{nk}{t}1_{\{i=n\}})\left(W_{j+1}-W_j\right) \\
&= \frac{t}{n}\sum_{j=0}^{n-1}(\tfrac{nk}{t}+n-j)\left(W_{j+1}-W_j\right) \\
&\sim \mathrm{N}\left(0, \frac{t^2}{n^2}\sum_{j=0}^{n-1}(\tfrac{nk}{t}+n-j)^2\frac{t}{n}\right),
\end{aligned}$$

where we have remembered that $W_0 = 0$ and used that the last sum consists of independent Gaussian stochastic variables. We thus obtain

$$\begin{aligned}
\sigma_\infty^2 &= \lim_{n\to\infty}\frac{t^3}{n^3}\sum_{j=0}^{n-1}\left(\frac{n^2k^2}{t^2} + (n^2+j^2-2nj) + 2\frac{nk}{t}(n-j)\right) \\
&= \lim_{n\to\infty}\frac{t^3}{n^3}\left(\frac{n^3k^2}{t^2} + n^3 + \left(\frac{(n-1)^3}{3} + \frac{(n-1)^2}{2} + \frac{n-1}{6}\right)\right. \\
&\qquad \left. - 2n\frac{n(n-1)}{2} + 2\frac{n^3k}{t} - 2\frac{nk}{t}\frac{n(n-1)}{2}\right) \\
&= tk^2 + \frac{1}{3}t^3 + t^2k.
\end{aligned}$$

□

Example 1.5. (Example 1.2 continued) In this setup, $\mathcal{J}$ contains two states, *alive* and *dead*, and the only non-negative transition rate is from state *alive* to *dead*. We now model the transition rate, i.e. the mortality rate, as a Brownian motion. Let, for an x-year old,

$$\mu_x(t) = \mu_x^\circ(t) + \sigma W(t),$$

be the mortality rate at time t, for some standard non-stochastic mortality μ_x°, a volatility $\sigma > 0$ and a standard Brownian motion $\mathbf{W}$ which is adapted to $\mathbf{F}^Y$. Note that the mortality intensity with this model can become negative. Depending on σ this happens with very low probability and the model is thus interesting as a naive model. We find, for $u \geq t$,

$$\begin{aligned}
{}_{n-t}p_{x+t}^{u} &= P(T_x \geq n \mid T_x \geq t, \mathcal{F}^Y(u)) \\
&= \mathrm{E}\left[1_{\{T_x \geq n\}} \mid T_x \geq t, \mathcal{F}^Y(u)\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[1_{\{T_x \geq n\}} \mid T_x \geq t, \mathcal{F}^Y(\infty)\right] \middle| T_x \geq t, \mathcal{F}^Y(u)\right] \\
&= \mathrm{E}\left[e^{-\int_t^n \mu_x(s)\,\mathrm{d}s} \middle| \mathcal{F}^Y(u)\right] \\
&= e^{-\int_t^n \mu_x^\circ(s)\,\mathrm{d}s - \int_t^u \sigma W(s)\,\mathrm{d}s}\, \mathrm{E}\left[e^{-\int_u^n \sigma W(s)\,\mathrm{d}s} \middle| \mathcal{F}^Y(u)\right] \\
&= e^{-\int_t^n \mu_x^\circ(s)\,\mathrm{d}s - \int_t^u \sigma W(s)\,\mathrm{d}s - (n-u)\sigma W(u)}\, \mathrm{E}\left[e^{-\int_u^n \sigma(W(s)-W(u))\,\mathrm{d}s}\right]
\end{aligned}$$

Using Lemma 1.3, with $k = 0$, one obtains

$$\mathrm{E}\left[e^{-\int_u^n \sigma(W(s)-W(u))\,\mathrm{d}s}\right] = \mathrm{E}\left[e^{-\int_0^{n-u} \sigma W(s)\,\mathrm{d}s}\right] = e^{\frac{1}{2}\sigma^2 \frac{1}{3}(n-u)^3},$$

using the moment generating function for the Gaussian distribution. In total,

$${}_{n-t}p_{x+t}^{u} = \exp\left\{-\int_t^n \mu_x^\circ(s)\,\mathrm{d}s - \int_t^u \sigma W(s)\,\mathrm{d}s - (n-u)\sigma W(u) + \frac{1}{6}\sigma^2(n-u)^3\right\}. \tag{1.4}$$

Inserting this into the loss (1.3),

$$\begin{aligned}
&e^{\int_0^n r(s)\,\mathrm{d}s}\widetilde{L}(1) \\
&= e^{-\int_1^n \mu_x^\circ(s)\,\mathrm{d}s - (n-1)\sigma W(1) + \frac{1}{6}\sigma^2(n-1)^3}\left(1_{(T_x \geq 1)} - e^{-\int_0^1 (\mu_x^\circ(s) + \sigma W(s))\,\mathrm{d}s}\right) \\
&\quad + e^{-\int_0^n \mu_x^\circ(s)\,\mathrm{d}s - \int_0^1 \sigma W(s)\,\mathrm{d}s - (n-1)\sigma W(1) + \frac{1}{6}\sigma^2(n-1)^3} - e^{-\int_0^n \mu_x^\circ(s)\,\mathrm{d}s + \frac{1}{6}\sigma^2 n^3}.
\end{aligned}$$

If we continue considering only the systematic loss i.e. the second line of (1.3) (that is, we apply the law of the large numbers argument carried out above, and can thus replace $1_{(T_x \geq 1)}$ by ${}_1p_x^1$), then

$$\begin{aligned}
L(1) &= e^{-\int_0^n (r(s) + \mu_x^\circ(s))\,\mathrm{d}s}\left(e^{-\int_0^1 \sigma W(s)\,\mathrm{d}s - (n-1)\sigma W(1) + \frac{1}{6}\sigma^2(n-1)^3} - e^{\frac{1}{6}\sigma^2 n^3}\right) \\
&= e^{-\int_0^n (r(s) + \mu_x^\circ(s))\,\mathrm{d}s}\left(e^{\frac{1}{6}\sigma^2(n-1)^3} X - e^{\frac{1}{6}\sigma^2 n^3}\right),
\end{aligned}$$

i.e. it has representation $f(X)$ where f is an increasing function, and, by Lemma 1.3,

$$X \sim \operatorname{logN}\left(0, \frac{1}{3}\sigma^2 + \sigma^2 n(n-1)\right).$$

We conclude,

$$x_{\text{SCR}} \geq f\left(e^{\sqrt{\frac{1}{3}\sigma^2+\sigma^2 n(n-1)}\Phi^{-1}(1-\alpha)}\right),$$

where α is the confidence level, which in the Solvency II case is 0.5%, and Φ^{-1} is the inverse standard Gaussian distribution function. ○

We repeat the finding from (1.4) and apply Lemma 1.3 to find the distribution of ${}_{n-t}p^u_{x+t}$.

Corollary 1.6. *Consider an x-year old at time 0 and assume the mortality rate at time t is* $\mu_x(t) = \mu_x^\circ(t) + \sigma W(t)$ *for a Brownian motion* $\mathbf{W}$ *and a deterministic* $\mu_x^\circ(t)$*. Then, for* $n \geq u \geq t$,

$${}_{n-t}p^u_{x+t} = \exp\left\{-\int_t^n \mu_x^\circ(s)\,\mathrm{d}s - \int_t^u \sigma W(s)\,\mathrm{d}s - (n-u)\sigma W(u) + \frac{1}{6}\sigma^2(n-u)^3\right\},$$

and conditional on $\mathcal{F}^Y(t)$ *the distribution is,*

$$\begin{aligned}{}_{n-t}p^u_{x+t} \mid \mathcal{F}^Y(t) \sim \operatorname{logN}\Bigg(-\int_t^n \mu_x^\circ(s)\,\mathrm{d}s + \frac{1}{6}\sigma^2(n-u)^3 - (n-t)\sigma W(t),\\ \frac{1}{3}\sigma^2(u-t)^3 + \sigma^2(n-u)(u-t)(n-t)\Bigg).\end{aligned}$$

Proof. The distributional result follows by rewriting (1.4) and calculating the variance explicitly. □

This is of interest in simple cases like the pure endowment from Example 1.2 and 1.5, where the systematic loss $L(1)$, the second line of (1.3), is essentially a monotone transformation of this distribution. Then it is particularly easy to find the Solvency Capital Requirement, SCR.

Even though it is not obvious, the above examples include longevity models. The mortality could be of the form $\mu^\circ(x) = (\alpha + \beta e^{\gamma x})\, e^{-\kappa x}$ which is a Gompertz-Makeham mortality with an overall yearly decrease of a factor $e^{-\kappa}$.

Considering Example 1.5 above, one sees that a particular complication arises from the realisation of the first year's real mortality which contributes with the quantity $\int_t^u \sigma W(s)\,\mathrm{d}s$. If we assume it to be known already, we would get a simpler result, which could prove advantageous in more complicated models.

Example 1.7. Consider a life annuity that begins payment at some time n with a payment rate $b(t)$, i.e. $dB(t) = 1_{(T_x \geq t)} 1_{(t \geq n)} b(t)\,\mathrm{d}t$. The loss arising from the systematic risk, defined in Definition 1.1, is found using Corollary 1.6

$$L(1) = \int_n^\infty e^{-\int_0^s (r(u)+\mu_x^\circ(u))\,\mathrm{d}u} \left(e^{-\int_0^1 \sigma W(u)\,\mathrm{d}u - (s-1)\sigma W(1) + \frac{1}{6}\sigma^2 (s-1)^3} - e^{\frac{1}{6}\sigma^2 s^3} \right) \mathrm{d}s.$$

The distribution or even the quantile is not easily found here. We see that the stochastic part, $\int_0^1 \sigma W(u)\,\mathrm{d}u + (s-1)\sigma W(1)$, depends on s, the integration variable. Thus, the loss is not just a monotone transformation of a stochastic variable, but a complicated transformation of the two-dimensional stochastic variable $\left(\int_0^1 \sigma W(u)\,\mathrm{d}u, \sigma W(1)\right)$.

We now make the simplifying assumption that the development in the mortality does not affect the mortality in the first year, i.e. assume now,

$$\mu_x(t) = \mu_x^\circ(t) + 1_{(t \geq t_0)} \sigma W(t),$$

where $t_0 = 1$ is the obvious choice. By calculations similar to the ones in Example 1.5 and Lemma 1.3, we find, for $t \leq u \leq t_0$,

$$_{n-t}p_{x+t}^u = e^{-\int_t^n \mu_x^\circ(s)\,\mathrm{d}s}\, \mathrm{E}\left[e^{-(n-t_0)\sigma W(t_0)} \mid \mathcal{F}^Y(u) \right] e^{\frac{1}{6}\sigma^2 (n-t_0)^3}.$$

Using this, with $t_0 = 1$, the loss reduces to

$$L(1) = \int_n^\infty e^{-\int_0^s (r(u)+\mu_x^\circ(u))\,\mathrm{d}u + \frac{1}{6}\sigma^2 (s-1)^3} \left(e^{-(s-1)\sigma W(1)} - e^{\frac{1}{2}\sigma^2 (s-1)^2} \right) \mathrm{d}s.$$

We still have an integral (convolution) of dependent log-normal distributions, but it is easy to find quantiles: This is a decreasing function of $W(1)$, and we find,

$$x_{\mathrm{SCR}} \geq \int_n^\infty e^{-\int_0^s (r(u)+\mu_x^\circ(u))\,\mathrm{d}u + \frac{1}{6}\sigma^2 (s-1)^3} \left(e^{-(s-1)\sigma \Phi^{-1}(\alpha)} - e^{\frac{1}{2}\sigma^2 (s-1)^2} \right) \mathrm{d}s,$$

where α is the confidence level, e.g. 0.5%. ○

We now show a variation on Lemma 1.3.

Lemma 1.8. *Assume for $t \geq 0$ that $f : [0,t] \mapsto \mathbb{R}$ is discontinuous in only finitely many points and has limits everywhere (i.e. for $s \in (0,t)$, $f(s-), f(s+), f(0+)$ and $f(t-)$ exist in $\mathbb{R}$). Let $\mathbf{W}$ be a Brownian motion. Then, if*

$$\sigma^2 = 2 \int_0^t f(s) \int_0^s u f(u)\,\mathrm{d}u\,\mathrm{d}s$$

exists,

$$\int_0^t f(s)W(s)\,\mathrm{d}s \sim N\left(0, \sigma^2\right).$$

Remark 1.9. Using the moment generating function,

$$\mathrm{E}\left[e^{\alpha \int_0^t f(s)W(s)\,\mathrm{d}s}\right] = e^{\frac{1}{2}\alpha^2\sigma^2}.$$

◇

Proof. Since **W** is continuous, $f\mathbf{W}$ is piecewise Riemann integrable because f, and thus $f\mathbf{W}$, is piecewise continuous with finite limits. Then $f\mathbf{W}$ is Riemann integrable. We follow the proof of Lemma 1.3, and find the characteristic function to be

$$\mathrm{E}\left[e^{-i\alpha \int_0^t f(s)W(s)\,\mathrm{d}s}\right] = e^{-\frac{1}{2}\alpha^2 \lim_{n\to\infty}\sigma_n^2},$$

where σ_n^2 is the variance of the Riemann sum of the integral. With $f_i = f\left(t\frac{i}{n}\right)$ and $W_i = W\left(t\frac{i}{n}\right)$, we find

$$\begin{aligned}
\sigma_n^2 &= \mathrm{Var}\left(\frac{t}{n}\sum_{i=1}^n f_i W_i\right)\\
&= \frac{t^2}{n^2}\sum_{i=1}^n\sum_{j=1}^n f_i f_j \,\mathrm{Cov}\left(W_i, W_j\right)\\
&= \frac{t^2}{n^2}\sum_{i=1}^n\sum_{j=1}^n f_i f_j \left(\frac{ti}{n} \wedge \frac{tj}{n}\right)\\
&\xrightarrow[n\to\infty]{} \int_0^t\int_0^t f(s)f(u)\left(s\wedge u\right)\,\mathrm{d}u\,\mathrm{d}s\\
&= \int_0^t \left(\int_0^s uf(s)f(u)\,\mathrm{d}u + \int_s^t sf(s)f(u)\,\mathrm{d}u\right)\,\mathrm{d}s\\
&= \int_0^t f(s)\int_0^s uf(u)\,\mathrm{d}u\,\mathrm{d}s + \int_0^t f(u)\int_0^u sf(s)\,\mathrm{d}s\,\mathrm{d}u\\
&= 2\int_0^t f(s)\int_0^s uf(u)\,\mathrm{d}u\,\mathrm{d}s\\
&= \sigma_\infty^2.
\end{aligned}$$

□

This result can be used to study the examples above, now with the mortality rate of the form

$$\mu_x(t) = \mu_x^\circ(t)\left(1 + 1_{(t\geq t_0)}\sigma W(t)\right). \tag{1.5}$$

This model bears a stronger similarity to the Solvency II approach, where the stress scenarios are obtained by multiplication of factors onto the mortality rate.

Corollary 1.10. *Let μ_x be defined as in (1.5), with $t_0 = u$, where μ_x° satisfies the conditions set up for f in Lemma 1.8. For $n > u \geq t \geq 0$, the conditional survival probability ${}_{n-t}p^u_{x+t}$ has representation*

$${}_{n-t}p^u_{x+t} = \exp\left\{-\int_t^u \mu_x^\circ(\tau)\,\mathrm{d}\tau - (1+\sigma W(u))\int_u^n \mu_x^\circ(\tau)\,\mathrm{d}\tau + \frac{1}{2}\sigma^2\nu^2\right\},$$

where ν^2 is from Lemma 1.8,

$$\nu^2 = 2\int_0^{n-u} \mu_x^\circ(u+\tau)\int_0^\tau s\mu_x^\circ(u+s)\,\mathrm{d}s\,\mathrm{d}\tau.$$

The distribution is given by,

$${}_{n-t}p^u_{x+t} \sim \mathrm{logN}\left(-\int_t^n \mu_x^\circ(\tau)\,\mathrm{d}\tau + \frac{1}{2}\sigma^2\nu^2\ ,\ u\sigma^2\left(\int_u^n \mu_x^\circ(\tau)\,\mathrm{d}\tau\right)^2\right).$$

Proof. Straightforward calculations, using that

$$\left(W(u+\tau) - W(u) \mid \mathcal{F}^Y(u)\right)_{\tau\in[0,\infty)}$$

is a standard Brownian motion,

$$\begin{aligned}
{}_{n-t}p^u_{x+t} &= \mathrm{E}\left[e^{-\int_t^n \mu_x(\tau)\,\mathrm{d}\tau}\,\middle|\,\mathcal{F}^Y(u)\right] \\
&= e^{-\int_t^n \mu_x^\circ(\tau)\,\mathrm{d}\tau}\,\mathrm{E}\left[e^{-\int_u^n \mu_x^\circ(\tau)\sigma W(\tau)\,\mathrm{d}\tau}\,\middle|\,\mathcal{F}^Y(u)\right] \\
&= e^{-\int_t^n \mu_x^\circ(\tau)\,\mathrm{d}\tau - \sigma W(u)\int_u^n \mu_x^\circ(\tau)\,\mathrm{d}\tau}\,\mathrm{E}\left[e^{-\sigma\int_0^{n-u}\mu_x^\circ(u+\tau)(W(u+\tau)-W(u))\,\mathrm{d}\tau}\,\middle|\,\mathcal{F}^Y(u)\right] \\
&= e^{-\int_t^u \mu_x^\circ(\tau)\,\mathrm{d}\tau - (1+\sigma W(u))\int_u^n \mu_x^\circ(\tau)\,\mathrm{d}\tau}\,e^{\frac{1}{2}\sigma^2\left(2\int_0^{n-u}\mu_x^\circ(u+\tau)\int_0^\tau s\mu_x^\circ(u+s)\,\mathrm{d}s\,\mathrm{d}\tau\right)}.
\end{aligned}$$

The distributional result follows directly. □

The conditional survival probability, ${}_{n-t}p^u_{x+t}$, consists of three factors. The first, $e^{-\int_t^u \mu_x^\circ(\tau)\,\mathrm{d}\tau}$, is the survival probability during the first year. The second part, $e^{-(1+\sigma W(u))\int_u^n \mu_x^\circ(\tau)\,\mathrm{d}\tau}$, is the survival probability beyond time u, which is stochastic. The third, $e^{\frac{1}{2}\sigma^2\nu^2}$, is a *loading* that emerges because we consider the mean of a log normal stochastic variable.

1.3 Forward Mortality Rate

A popular quantity is the forward mortality rate, which is a term borrowed from finance mathematics.

Definition 1.11. *The* forward mortality rate *at time u for the mortality rate at time t is an $\mathcal{F}^Y(u)$-measurable stochastic variable denoted $f_{x,u}(t)$, where x is the age at time 0. It is defined by*

$$f_{x,u}(t) = \begin{cases} -\frac{\partial}{\partial t} \log \mathrm{E}\left[e^{-\int_u^t \mu_x(\tau)\,\mathrm{d}\tau} \mid \mathcal{F}^Y(u)\right] & , t > u \\ \mu_x(t) & , t \leq u. \end{cases}$$

The definition is pleasant, since, for $u \leq t$,

$$\begin{aligned} e^{-\int_0^t f_{x,u}(s)\,\mathrm{d}s} &= e^{-\int_0^u f_{x,u}(s)\,\mathrm{d}s} e^{-\int_u^t f_{x,u}(s)\,\mathrm{d}s} \\ &= e^{-\int_0^u \mu_x(s)\,\mathrm{d}s} e^{\int_u^t \frac{\partial}{\partial s} \log \mathrm{E}\left[e^{-\int_u^s \mu_x(\tau)\,\mathrm{d}\tau} | \mathcal{F}^Y(u)\right] \mathrm{d}s} \\ &= e^{-\int_0^u \mu_x(s)\,\mathrm{d}s} e^{\log \mathrm{E}\left[e^{-\int_u^t \mu_x(\tau)\,\mathrm{d}\tau} | \mathcal{F}^Y(u)\right] - \log(1)} \\ &= \mathrm{E}\left[e^{-\int_0^t \mu_x(s)\,\mathrm{d}s} \mid \mathcal{F}^Y(u)\right]. \end{aligned}$$

Also, a very useful result for the forward rate applications in life insurance mathematics is the following.

Lemma 1.12. *Let μ_x be a non-negative mortality rate and consider an open and bounded time interval $J \subset \mathbb{R}_+$. If there exists a stochastic and time independent bound for μ_x on J, i.e.*

$$\sup_{t \in J} \mu_x(t,\omega) \leq X(\omega) \quad \text{for } \omega \in \Omega,$$

such that X conditional on $\mathcal{F}^Y(u)$ has finite expectation, then for $t \in J$, $u \leq t$ and $s < t$, the forward rate satisfies

$$e^{-\int_s^t f_{x,u}(\tau)\,\mathrm{d}\tau} f_{x,u}(t) = \mathrm{E}\left[e^{-\int_s^t \mu_x(\tau)\,\mathrm{d}\tau} \mu_x(t) \mid \mathcal{F}^Y(u)\right].$$

Proof. The forward rate satisfies $e^{-\int_s^t f_{x,u}(\tau)\,\mathrm{d}\tau} = \mathrm{E}\left[e^{-\int_s^t \mu_x(\tau)\,\mathrm{d}\tau} \mid \mathcal{F}^Y(u)\right]$. Differentiating with respect to t yields,

$$-e^{-\int_s^t f_{x,u}(\tau)\,\mathrm{d}\tau} f_{x,u}(t) = \frac{\partial}{\partial t} \mathrm{E}\left[e^{-\int_s^t \mu_x(\tau)\,\mathrm{d}\tau} \mid \mathcal{F}^Y(u)\right].$$

In order to swich differentiation and integration, it is, by [10], Theorem 8.14, sufficient that for all $t \in J$ and $\omega \in \Omega$ it holds that $\left|\frac{\partial}{\partial t} e^{-\int_s^t \mu_x(\tau,\omega)\,\mathrm{d}\tau}\right| \leq \widetilde{X}(\omega)$ for some integrable stochastic variable $\widetilde{X}$. This is satisfied, since

$$\left|\frac{\partial}{\partial t} e^{-\int_s^t \mu_x(\tau)\,\mathrm{d}\tau}\right| = \left|e^{-\int_s^t \mu_x(\tau)\,\mathrm{d}\tau}\mu_x(t)\right| \leq \mu_x(t) \leq X.$$

Thus differentiation and expectation can be interchanged, and multiplying with -1 yields,

$$\begin{aligned} e^{-\int_s^t f_{x,u}(\tau)\,\mathrm{d}\tau} f_{x,u}(t) &= -\,\mathrm{E}\left[\frac{\partial}{\partial t} e^{-\int_s^t \mu_x(\tau)\,\mathrm{d}\tau} \mid \mathcal{F}^Y(u)\right] \\ &= \mathrm{E}\left[e^{-\int_s^t \mu_x(\tau)\,\mathrm{d}\tau}\mu_x(t) \mid \mathcal{F}^Y(u)\right]. \end{aligned}$$

□

The results from Corollary 1.6 and 1.10 can be expressed in terms of forward mortality rates. They are obtained by differentiation with respect to n. In Corollary 1.6, we found the forward mortality rate

$$f_{x,u}^{(1)}(t) = \mu_x^\circ(t) + \sigma W(t \wedge u) - 1_{\{t \geq u\}}\frac{1}{2}\sigma^2(t-u)^2,$$

and in Corollary 1.10 we found the forward mortality rate,

$$f_{x,u}^{(2)}(t) = \mu_x^\circ(t)\left(1 + 1_{\{t \geq u\}}\sigma W(u) - 1_{\{t \geq u\}}\sigma^2 \int_u^t (s-u)\mu_x^\circ(s)\,\mathrm{d}s\right). \tag{1.6}$$

1.4 Stress Scenarios

To find x_{SCR} is a matter of finding a confidence interval for the loss after one year. By considering the mortality only, in a simple model, this can be reduced to finding confidence bands for the forward mortality rate, $f_{x,1}$. This is the case if the loss after one year is a monotone function of the stochastic part forward mortality rate.

An example is a pure endowment with forward mortality rate (1.6), where

$$\mu_x(t) = \mu_x^\circ(t)\left(1 + 1_{\{t \geq t_0\}}\sigma W(t)\right),$$

(see (1.5)). The systematic loss is in that case a monotone transformation of ${}_n p_x^1 = e^{-\int_x^n f_{x,1}^{(2)}(t)\,\mathrm{d}t}$ which is a monotone transformation of the stochastic part, $W(1)$.

In the Solvency II regime, one tries to approximate confidence bands for the forward mortality rate $f_{x,1}(t)$ by *stressing* the mortality rate. In the QIS5 specification, it is done by a 20% increase and decrease of the mortality. Thus, a confidence band for $f_{x,1}(t)$ is defined by,

$$\mu_x^\circ(t)\,(1-0.2) \le f_{x,1}(t) \le \mu_x^\circ(t)\,(1+0.2)\,.$$

We can compare with the stochastic mortality rate model specified by (1.5), repeated above. In this model a $(1-\alpha)$-confidence band is given by,

$$\begin{aligned} &\mu_x^\circ(t)\left(1 - 1_{\{t\ge t_0\}}\sigma t_0 \Phi\left(1-\tfrac{\alpha}{2}\right) - 1_{\{t\ge t_0\}}\sigma^2 \textstyle\int_{t_0}^t (s-t_0)\mu_x^\circ(s)\,\mathrm{d}s\right) \\ &\qquad\qquad \le f_{x,1}(t) \le \\ &\mu_x^\circ(t)\left(1 + 1_{\{t\ge t_0\}}\sigma t_0 \Phi\left(1-\tfrac{\alpha}{2}\right) - 1_{\{t\ge t_0\}}\sigma^2 \textstyle\int_{t_0}^t (s-t_0)\mu_x^\circ(s)\,\mathrm{d}s\right). \end{aligned}$$

To qualitatively compare the two methods, first assume that $t \ge t_0 = 1$. Then choose σ such that $\sigma\Phi\left(1-\frac{\alpha}{2}\right) \simeq 0.2$. For $\alpha = 99.5\%$, this means $\sigma \simeq 0.0712$. Then the confidence band becomes,

$$\mu_x^\circ(t)\left(1 \pm 0.2 - 0.0051\int_1^t (s-1)\mu_x^\circ(s)\,\mathrm{d}s\right).$$

This is almost the same as the Solvency II confidence band. However, the last term is only present in our model. What is this term, and of what size is it? The term arises because of the skewed distribution that arises when a symmetric distribution is transformed via the exponential function. If the mortality drops a bit, it has a larger impact than if it increases a bit, because the mortality is transformed. The size of the term is small, because σ^2 is small. However, it can become large, for large t. Consider the special case depicted in Figure 1. This is the G82M mortality for a 30 year old, extended with a yearly 2% decrease in mortality. It is specified by,

$$\mu_{30}^\circ(t) = \left(0.0005 + 0.000075858 \cdot 1.09144^{(t+30)}\right)(1-0.02)^t.$$

For most t, except the large ones, the stress in the QIS5 specifications and the confidence band from our model are identical. For t larger than 45 the confidence band in the model decreases, because the last term becomes large.

Our simple model does not reproduce the stress scenarios specified in QIS5. This is however not (only) a critique of the QIS5 specification. Our model is not realistic, because it includes negative mortality rates. The Brownian motion has a

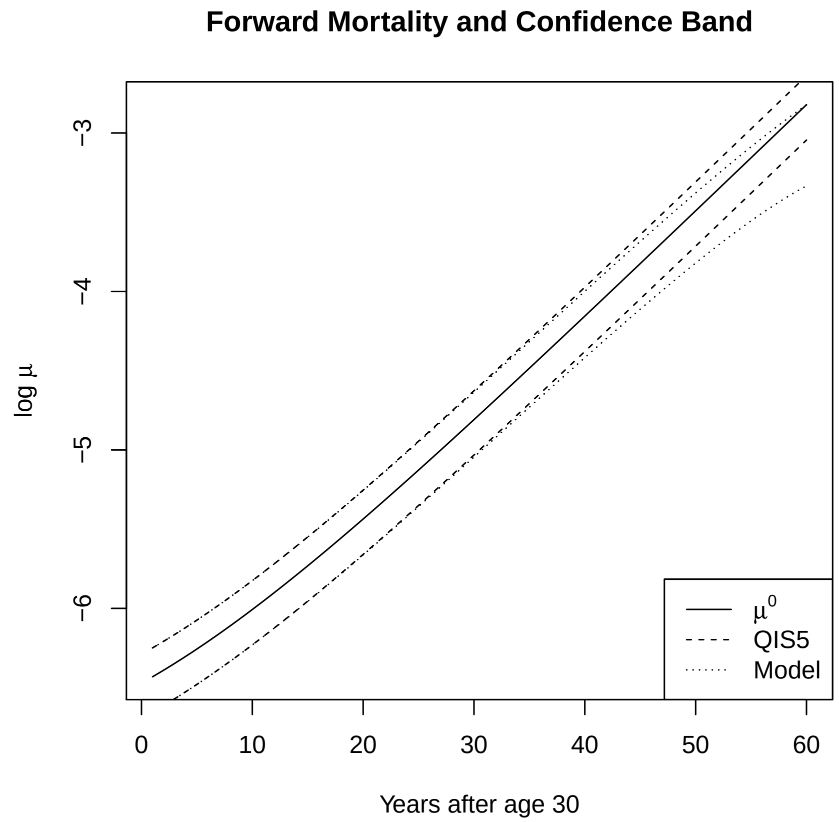

Figure 1: μ°_{30} *and confidence bands for the forward mortality rate* $f_{30,1}$ *for a male aged 30, based on the extended Gompertz-Makeham with G82 parameters and yearly mortality improvements of 2%. The confidence bands are almost identical until age 75 (time 45). The difference between the confidence bands does not arise because of the large age, but because it is far into the future. The y-axis has a logarithmic scale.*

positive probability of becoming negative of arbitrary size. If a more realistic model was used, that disallowed negative mortalities, the last term would be smaller, and then the two methods would be even more alike.

If one should conclude anything, it must be that the QIS5 stress is a very good approximation to the model set up here.

1.5 Stochastic Yearly Improvements

Definition 1.13. *For an x-year old at time 0, the* extended Gompertz-Makeham*, or the* Gompertz-Makeham model with longevity*, is defined by the parametrisation,*

$$\mu_x(t) = \left(\alpha + \beta e^{\gamma\cdot(t+x)}\right) e^{-\kappa(t)},$$

for parameters α, β, γ and the function $\kappa(t)$.

The extended Gompertz-Makeham mortality model is often split up into the *current mortality*, i.e. the Gompertz-Makeham part, $\alpha+\beta e^{\gamma\cdot(t+x)}$, and the *longevity* part, the yearly improvement $e^{-\kappa(t)}$. Often $\kappa(t)$ is a linear function, $\kappa(t) = \kappa t$.

In the linear case, there are 4 parameters to determine, and it is the uncertainty in these parameters that lead us to present a stochastic model such as (1.5). However, carrying out the actual estimation, one finds that it is rather easy to estimate the *current mortality*, i.e. α, β and γ. The difficult part is to predict the future mortality improvement, i.e. to estimate κ. For a more realistic and still simple stochastic model, one could propose to make κ stochastic, instead of multiplying the whole mortality by a stochastic process, which is similar to making α and β stochastic.

Example 1.14. Consider a binomial model for a stochastic κ. After one year, we realise if we are in state *up* or *down*, i.e. for $\kappa_u \geq \kappa_d$,

$$P(\kappa = \kappa_u) = 1 - P(\kappa = \kappa_d) = 1 - p_d = p_u,$$

with κ being $\mathcal{F}^Y(1)$-measurable, but not $\mathcal{F}^Y(1-\varepsilon)$-measurable, for $\varepsilon > 0$. Then, for $t \geq 1$,

$$f_{x,1}(t) = -\frac{\partial}{\partial t} \log \mathrm{E}\left[e^{-\int_1^t \mu_x(\tau)\,\mathrm{d}\tau} \mid \mathcal{F}^Y(1)\right] = \mu_x(t),$$

because the stochastic part, κ, is $\mathcal{F}^Y(1)$-measurable. Now, if $p_u \in \left(\frac{\alpha}{2}, 1-\frac{\alpha}{2}\right)$, which is very reasonable, then a $(1-\alpha)$-confidence area for κ is $[\kappa_d, \kappa_u]$ (actually $\{\kappa_d, \kappa_u\}$ is sufficient), which defines the confidence band for the forward mortality, $f_{x,1}$, for $t \geq 1$,

$$\left(\alpha + \beta e^{\gamma\cdot(t+x)}\right) e^{-\kappa_u t} \leq f_{x,1}(t) \leq \left(\alpha + \beta e^{\gamma\cdot(t+x)}\right) e^{-\kappa_d t}.$$

This should be seen in contrast to the QIS5 stress,

$$\left(\alpha + \beta e^{\gamma\cdot(t+x)}\right) e^{-\kappa t} q_d \leq f_{x,1}(t) \leq \left(\alpha + \beta e^{\gamma\cdot(t+x)}\right) e^{-\kappa t} q_u,$$

for $q_u = 1.2$ and $q_d = 0.8$. There is a qualitative difference which is important here. In our model, where the longevity effect is stochastic, the confidence band becomes wider with time, in contrast to the confidence band defined in QIS5 which do not widen with time.

To illustrate the difference, we consider the α, β and γ parameters from the G82M mortality table, and let

$$\begin{aligned}\kappa_u &= -\log(1 - 0.03),\\ \kappa_d &= -\log(1 - 0.01),\end{aligned}$$

i.e. a yearly mortality improvement of 3% or 1%. The confidence bands based on each method are depicted in Figure 2.

Another way to illustrate the difference is to consider the conditional survival probabilities. They are calculated in three cases and shown in Table 1. For large time horizons, the binomial model with a stochastic longevity parameter gets wider confidence bands. The difference between the models will, in most cases, have a large effect for contracts like the life annuity, which can often span time periods longer than 50 years. ○

Figure 2: *μ°_{30} and confidence bands for the forward mortality rate $f_{30,1}$ for a male aged 30, based on the extended Gompertz-Makeham with G82 parameters and yearly mortality improvements of 2%. The confidence bands have a different structure, and the stochastic longevity model gets a wider confidence band in the far future.*

$_{70-t}p^{1}_{x+t}$	20		30		60	
QIS5	79%	85%	76%	83%	78%	84%
Binomial	72%	87%	74%	84%	80%	82%

Table 1: *Confidence intervals for the probability of a 20, 30 or 60 year old reaching age 70. The confidence intervals for a 20 or 30 year old is wider with the binomial model while for a 60 year old, the confidence interval is wider with the QIS5 stress method.*

2 Continuous Affine Processes

We wish to extend our setup to a broader class of mortality models. In the previous section, we studied models where

$$ {}_{n-t}p^t_{x+t} = \mathrm{E}\left[e^{-\int_t^n \mu_x(\tau)\,\mathrm{d}\tau} \middle| \mathcal{F}^Y(t) \right] = e^{A_x(t,n)+B_x(t,n)W(t)}, $$

i.e. models where the conditional survival probability is an exponential affine function in the stochastic variable $W(t)$, at time t. These models are special cases of the broader class of models called *affine processes*.

We wish to examine these processes, and give them a theoretical treatment. We restrict ourselves to continuous stochastic processes, that is, we do not include processes with jumps. In this section the dependence of the age, x, is suppressed.

For $d \geq 1$ we consider the class of d-dimensional adapted processes $\mathbf{X} = (X(t))_{t\in\mathbb{R}_+}$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}(t))_t, P)$. We require $\mathbf{X}$ to be the solution of a stochastic differential equation,

$$ \begin{aligned} \mathrm{d}X(t) &= b(t, X(t))\,\mathrm{d}t + \rho(t, X(t))\,\mathrm{d}W(t), \\ X(0) &= x \in \mathcal{X}. \end{aligned} \tag{2.1} $$

Here, $\mathcal{X} \subset \mathbb{R}^d$ is a non-empty closed state space. The drift vector $b : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}^d$ is assumed to be measurable and locally Lipschitz continuous, and the matrix $\rho : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}^{d\times d}$ is assumed be measurable and such that the diffusion matrix

$$ a(t, x) = \rho(t, x)\rho(t, x)^\top $$

is locally Lipschitz continuous. Also $\mathbf{W}$ is a d-dimensional Brownian motion and $(\mathcal{F}(t))_t$ is the filtration generated by $\mathbf{W}$. The filtration $(\mathcal{F}(t))_t$ is assumed to satisfy the usual conditions.

In the rest of the thesis, the following assumption is adopted.

Assumption 2.1. *The diffusion process* $\mathbf{X}$ *satisfying by* (2.1) *is assumed to have a solution for all starting times* $t_0 \geq 0$ *and all* $x \in \mathcal{X}$*, i.e. there exists a solution to*

$$ \mathrm{d}X(t) = b(t_0 + t, X(t))\,\mathrm{d}t + \rho(t_0 + t, X(t))\,\mathrm{d}W(t), \quad X(0) = x. $$

We write $\mathbf{X}^x$ for the solution with $X(0) = x$, and we have $t_0 = 0$ unless otherwise specified. Often we suppress the notation of x in $\mathbf{X}^x$ and we just write $\mathbf{X}$.

We now give a definition of affine processes.

Definition 2.2. *A stochastic process* $\mathbf{X}$ *satisfying* (2.1) *and Assumption 2.1 is called* affine *if the* $\mathcal{F}(t)$*-conditional characteristic function of* $X^x(T)$ *is exponential affine for all* $x \in \mathcal{X}$ *and* $t, T \geq 0$ *with* $t \leq T$*. That is, there exist functions* ϕ *and* ψ *such that for all* $x \in \mathcal{X}$*,* $u \in \mathrm{i}\mathbb{R}^d$ *and* $t \leq T$*,*

$$\mathrm{E}\left[e^{u^\top X^x(T)} \middle| \mathcal{F}(t) \right] = e^{\phi(t,T,u) + \psi(t,T,u)^\top X^x(t)}, \tag{2.2}$$

where $\phi(t,T,u)$ *is* $\mathbb{C}$*-valued, and* $\psi(t,T,u)$ *is* $\mathbb{C}^d$*-valued, and with jointly continuous* t*-derivatives.*

Often the formula (2.2) is referred to as the *affine transformation formula.*

Remark 2.3. Note that $\phi(T,T,u) = 0$ and $\psi(T,T,u) = u$. Also, using that for any $z \in \mathbb{C}$, $|e^z| = e^{\Re z}$, and Remark A.12, we get

$$e^{\Re\left(\phi(t,T,u) + \psi(t,T,u)^\top X^x(t)\right)} = \left| \mathrm{E}\left[e^{u^\top X^x(T)} \middle| \mathcal{F}(t) \right] \right| \leq 1,$$

which means that $\Re\left(\phi(t,T,u) + \psi(t,T,u)^\top X(t)\right) \leq 0$. ◇

Remark 2.4. It is allowed to have $t \geq T$ in the sense that one can consider the stopped process $\mathbf{X} = (X(T \wedge t))_{t \in \mathbb{R}_+}$ and for $t \geq T$ let $(\phi(t,T,u), \psi(t,T,u)) = (0, u)$. Then (2.2) still holds. ◇

The treatment of affine processes in this thesis is inspired by Filipovic [7], Chapter 10, most of which was first published in the article by Filipovic and Mayerhofer [8]. There, multidimensional time-homogeneous continuous affine processes are examined, and the treatment presented here is an attempt to extend some of the results to the time-inhomogeneous case.

Sections 2.2 and 2.4 are, with their theorems and proofs, analogous to the treatment in Chapter 10, Section 1 and 2, in [7]. The main difference is the generalisation to the time-inhomogeneous case which brings certain complications to some of the proofs. Also, the proofs here are presented in greater detail. Sections 2.3 and 3.1 are also largely inspired by Chapter 10 in [7]. The content of Section 3.2 is, to the author's knowledge, new results not present in existing litterature.

The work in this thesis differs from that in [7] essentially because the definition of affine processes is more general. In the definition presented in [7], the equation (2.2) reads

$$\mathrm{E}\left[e^{u^\top X^x(T)} \middle| \mathcal{F}(t) \right] = e^{\phi(T-t,u) + \psi(T-t,u)^\top X^x(t)}. \tag{2.3}$$

The difference is, that ϕ and ψ can only depend on the times t and T through $T - t$.

2.1 Comparison With The Time-Homogeneous Case

Time-inhomogeneous processes can be viewed as special cases of multidimensional time-homogeneous processes. To see this, consider the d-dimensional time-inhomogeneous process $\mathbf{X}$ satisfying (2.1). Let $\boldsymbol{\Upsilon}$ be the time-process, i.e $\mathrm{d}\Upsilon(t) = \mathrm{d}t$. Then $\widetilde{\mathbf{X}} = (\boldsymbol{\Upsilon}, \mathbf{X})$ is a time-homogeneous process with,

$$\begin{aligned} \mathrm{d}\widetilde{X}(t) &= \begin{bmatrix} 1 \\ b(\widetilde{X}(t)) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 & 0 \\ 0 & \rho(\widetilde{X}(t)) \end{bmatrix} \mathrm{d}W(t), \\ \widetilde{X}(0) &= (t_0, x) \in \mathbb{R}_+ \times \mathcal{X}. \end{aligned} \tag{2.4}$$

Notice that Assumption 2.1 about $\mathbf{X}$ ensures that $\widetilde{\mathbf{X}}$ has a solution for all start times $t_0 \in \mathbb{R}_+$, as well as start points $x \in \mathcal{X}$. We refer to this extension of the process as "time-homogenisation".

This realisation would at first sight make our attempt to extend the results of [7] to the time-inhomogeneous case trivial. However, in [7] a result analogous to formula (2.9) of Theorem 2.6 below is stated, with the difference being that the parameters are time-independent. A consequence of this is that the parameters b and a have to be affine in the time t. As the following example shows, this is not necessary for the affine transformation formula to hold.

Example 2.5. Recall the Hull-White time-inhomogeneous Vasiček process,

$$\mathrm{d}X(t) = (\gamma(t) - \delta(t)X(t))\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W(t).$$

This is a 1-dimensional process, and we consider the case where only δ is time-dependent. We let $\delta : [0, T] \to \mathbb{R}_+$ be continuous and let $\gamma \in \mathbb{R}$ and $\sigma \in \mathbb{R}_{++}$. As a time-inhomogeneous process it satisfies the stochastic differential equation

$$\mathrm{d}X(t) = (\gamma - \delta(t)X(t))\,\mathrm{d}t + \sigma\,\mathrm{d}W(t), \quad X(0) = x \in \mathcal{X}, \tag{2.5}$$

with state space $\mathcal{X} = \mathbb{R}$.

We now show that $\mathbf{X}$ is affine by Definition 2.2, by finding functions ϕ and ψ such that (2.2) holds. To find such functions, we first assume that they exist and define $M(t) = e^{\phi(t,T,u)+\psi(t,T,u)X(t)}$. Now $\mathbf{M}$ is a martingale, and Itô's formula is applied,

$$
\begin{aligned}
\frac{\mathrm{d}M(t)}{M(t)} &= \left(\frac{\partial}{\partial t}\phi(t,T,u) + X(t)\frac{\partial}{\partial t}\psi(t,T,u)\right)\mathrm{d}t + \psi(t,T,u)\,\mathrm{d}X(t) \\
&\quad + \frac{1}{2}\psi^2(t,T,u)\,\mathrm{d}\left\langle X(t), X(t)\right\rangle \\
&= \left(\frac{\partial}{\partial t}\phi(t,T,u) + X(t)\frac{\partial}{\partial t}\psi(t,T,u) + \psi(t,T,u)(\gamma - \delta(t)X(t))\right. \\
&\quad \left. + \frac{1}{2}\psi^2(t,T,u)\sigma^2\right)\mathrm{d}t + \psi(t,T,u)\sigma\,\mathrm{d}W(t).
\end{aligned}
$$

Here, the quadratic variation is found according to the standard results stated in Section A.4.

Since $\mathbf{M}$ is assumed to be a martingale, the drift term must equal zero,

$$
\begin{aligned}
0 &= \frac{\partial}{\partial t}\phi(t,T,u) + X(t)\frac{\partial}{\partial t}\psi(t,T,u) + \psi(t,T,u)(\gamma - \delta(t)X(t)) \\
&\quad + \frac{1}{2}\psi^2(t,T,u)\sigma^2 \\
&= X(t)\left(\frac{\partial}{\partial t}\psi(t,T,u) - \delta(t)\psi(t,T,u)\right) \\
&\quad + \frac{\partial}{\partial t}\phi(t,T,u) + \gamma\psi(t,T,u) + \frac{1}{2}\sigma^2\psi^2(t,T,u).
\end{aligned}
$$

Since the above has to hold for any $X(t) \in \mathcal{X}$, and $\mathbf{X}$ in $\mathcal{X}$ is non-degenerate[1], we conclude that the two last lines each must be zero, i.e.

$$
\begin{aligned}
\frac{\partial}{\partial t}\psi(t,T,u) &= \delta(t)\psi(t,T,u), \\
\frac{\partial}{\partial t}\phi(t,T,u) &= -\gamma\psi(t,T,u) - \frac{1}{2}\sigma^2\psi^2(t,T,u).
\end{aligned}
$$

Hence, combined with the boundary conditions $\phi(T,T,u) = 0$ and $\psi(T,T,u) = u$, we have a set of differential equations. Since δ is continuous on the compact interval $[0,T]$, it is integrable with finite integral, and a solution for ψ is,

$$
\psi(t,T,u) = ue^{-\int_t^T \delta(s)\,\mathrm{d}s}, \tag{2.6}
$$

[1] That $\mathbf{X}$ in $\mathcal{X}$ is non-degenerate is a bit unprecise in this regard; What we mean is that for all $t \in (0,T]$, there does not exist a point $x_t \in \mathcal{X}$ such that $X(t) = x_t$ a.s.

which can be easily checked by differentiation. Then ϕ is an integral of a continuous function on a compact interval, and it exists with representation,

$$\phi(t,T,u) = \gamma u \int_t^T e^{-\int_s^T \delta(\tau)\,\mathrm{d}\tau}\,\mathrm{d}s + \frac{1}{2}\sigma^2 u^2 \int_t^T e^{-2\int_s^T \delta(\tau)\,\mathrm{d}\tau}\,\mathrm{d}s. \tag{2.7}$$

Now, drop the martingale assumption about $\mathbf{M}$. With ψ and ϕ as defined in (2.6) and (2.7), we see that $\mathbf{M}$ is a local martingale.

Since $-\int_t^T \delta(s)\,\mathrm{d}s \in \mathbb{R}$ and $u \in \mathrm{i}\mathbb{R}$, it is seen from (2.6) that $\psi(t) \in \mathrm{i}\mathbb{R}$. Also, the first term in (2.7) is imaginary, and since $u^2 \in \mathbb{R}_-$ the second term is real and negative. Thus, for all $0 \le t \le T$, $X(t) \in \mathbb{R}$ and $u \in \mathrm{i}\mathbb{R}$,

$$|M(t)| = e^{\Re(\phi(T-t,u)+\psi(T-t,u)X_1(t))} \in (0,1],$$

and we see that $\mathbf{M}$ is uniformly bounded. This allows us to conclude that $\mathbf{M}$ is a martingale, such that (2.2) is satisfied and $\mathbf{X}$ is affine.

We now examine whether this process is affine using the definition in Filipovic, i.e. (2.3). At first sight, we see that only if δ is a constant, i.e. time-independent, ϕ and ψ depend on t and T only through $T-t$. Thus for time-dependent δ, $\mathbf{X}$ is not itself affine using the definition in (2.3).

Thus, the process needs to be time-homogenised, and we do that to examine if it is then affine. Let $\widetilde{\mathbf{X}} = (\boldsymbol{\Upsilon}, \mathbf{X})$ with $\mathrm{d}\Upsilon(t) = \mathrm{d}t$, i.e.

$$\begin{aligned}
\mathrm{d}\widetilde{X}(t) &= \begin{bmatrix} 1 \\ \gamma - \delta(\Upsilon(t))X(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 & 0 \\ 0 & \sigma \end{bmatrix} \mathrm{d}W(t), \\
\widetilde{X}(0) &= (t_0, x) \in \mathcal{X},
\end{aligned}$$

where $\mathcal{X} = \mathbb{R}_+ \times \mathbb{R}$. We adopt the approach above, and start by defining $M(t) = e^{\phi(T-t,u)+\psi(T-t,u)^\top \widetilde{X}(t)}$. Now $\psi = (\psi_1, \psi_2)^\top$ is 2-dimensional. Applying Itô's formula, we get

$$\begin{aligned}
\frac{\mathrm{d}M(t)}{M(t)} &= -\left(\frac{\partial \phi}{\partial t}(T-t,u) + \frac{\partial \psi}{\partial t}(T-t,u)^\top \widetilde{X}(t)\right)\mathrm{d}t + \psi(T-t,u)^\top\,\mathrm{d}\widetilde{X}(t) \\
&\quad + \frac{1}{2}\psi_2^2(T-t,u)\,\mathrm{d}\langle X(t)\rangle \\
&= \Bigg(-\frac{\partial \phi}{\partial t}(T-t,u) - \frac{\partial \psi_1}{\partial t}(T-t,u)\Upsilon(t) - \frac{\partial \psi_2}{\partial t}(T-t,u)X(t) \\
&\quad + \psi_1(T-t,u) + \psi_2(T-t,u)(\gamma - \delta(\Upsilon(t))X(t)) \\
&\quad + \frac{1}{2}\psi_2^2(T-t,u)\sigma^2 \Bigg)\mathrm{d}t + \psi_2(T-t,u)\sigma\,\mathrm{d}W(t).
\end{aligned}$$

Here, the notation $\frac{\partial f}{\partial t}(T-t,u)$ is used for the function $\frac{\partial}{\partial t}(f(t,u))$ evaluated in $(T-t,u)$.

Letting the drift term equal 0, one obtains

$$\begin{aligned} 0 = &-X(t)\left(\frac{\partial \psi_2}{\partial t}(T-t,u) + \psi_2(T-t,u)\delta(\Upsilon(t))\right) \\ &- \Upsilon(t)\frac{\partial \psi_1}{\partial t}(T-t,u) \\ &+ \psi_2(T-t,u)\gamma - \frac{\partial \phi}{\partial t}(T-t,u) + \psi_1(T-t,u) + \frac{1}{2}\psi_2^2(T-t,u)\sigma^2. \end{aligned} \tag{2.8}$$

We have $\Upsilon(t) = t_0 + t$. The above has to hold for all $t \in [0,T]$ and $(t_0,x) \in \mathcal{X}$, thus for all $\Upsilon(t) \in [t,\infty)$ and $X(t) \in \mathbb{R}$. Especially, the first line has to equal 0, and unless δ is constant, this is only possible for $\psi_2(t,u) = 0$. However, this is not a solution, since we need $\psi(T-t,u) = u$ for $t = T$. Thus there does not exist ϕ and ψ such that $\mathbf{M}$ is a martingale, and $\widetilde{\mathbf{X}}$ is not affine when using (2.3) as the definition. ○

The example shows that the theory in [7] about affine time-homogeneous processes does not include time-inhomogeneous processes in general. As mentioned before the example, the parameters a and b of a time-inhomogeneous affine process need to be affine in t (or a function of the time t that satisfies an autonomous differential equation) for it to be a special case of the time-homogeneous affine processes treated in [7]. The drift in (2.5) is not affine in t since the time-dependent part $\delta(t)$ is multiplied on to $X(t)$. Thus any time-dependence in δ is not allowed. Alternatively, if the process is specified such that $\gamma(t)$ is the time-dependent part and $\gamma(t)$ is the solution to an autonomous differential equation, then the process (2.5) could be time-homogenised with parameters a and b that is affine in $\gamma(t)$.

There are two ways to extend the results in [7] to include time-inhomogeneous processes with parameters that are not restricted to being affine in functions of the time t. The first approach is to treat time-inhomogeneous processes for themselves, and not consider them as a special case of time-homogeneous processes. That works, as we saw in the first part of the example. The other approach is to treat time-inhomogeneous processes as a special case of the time-homogeneous processes, and then allow ϕ and ψ to depend on the starting value $x \in \mathcal{X}$. If that is the case there would exist ϕ and ψ such that (2.8) holds, and the time-homogenised version would be affine. In this thesis, the first approach is taken. The main reason is that (2.9) of Theorem 2.6 below then holds. As seen, it need not hold (for the non-stochastic dimensions) if ϕ and ψ are allowed to depend on the starting values, and then a characterisation of affine processes seems harder to obtain.

2.2 Necessary and Sufficient Conditions

Definition 2.2 is not very constructive. Natural questions are, what kind of processes are affine? And if a process is affine, then how does one find the functions ϕ and ψ? Luckily, there are answers to these questions. A stochastic process is affine if and only if the drift and diffusion are affine. If that is the case, the functions ϕ and ψ can be found as solutions to a system of differential equations of the Riccati type. We refer to this system of differential equations as the "Riccati equations".

The theorem below formalises the mentioned results on affine processes. The theorem is presented in [7], Theorem 10.1, for the time-homogeneous case. The generalisation to the time-inhomogeneous case stated here is almost straightforward, and the main contribution is the detailed proof, which is presented in less detail in [7].

Theorem 2.6. *Let* $\mathbf{X}$ *be a stochastic process satisfying* (2.1) *and Assumption 2.1.*

Part I: *Assume that* $\mathbf{X}$ *is affine. Then the drift* $b(t,x)$ *and the diffusion matrix* $a(t,x)$ *are affine in* x*, that is*

$$\begin{aligned} a(t,x) &= a(t) + \sum_{i=1}^{d} x_i \alpha_i(t), \\ b(t,x) &= b(t) + \sum_{i=1}^{d} x_i \beta_i(t) = b(t) + \mathcal{B}(t)x, \end{aligned} \tag{2.9}$$

with $a(t), \alpha_1(t), \dots, \alpha_d(t)$ *being* $d \times d$*-dimensional matrices and* $b(t), \beta_1(t), \dots, \beta_d(t)$ *being* d*-dimensional vectors. Then* $\mathcal{B}(t) = (\beta_1(t), \dots, \beta_d(t))$*, i.e.* $\mathcal{B}(t)$ *is a* $d \times d$ *matrix with the* i*th column vector being* $\beta_i(t)$*. Also,* ϕ *and* $\psi = (\psi_1, \dots, \psi_d)^\top$ *solve the Riccati equations,*

$$\begin{aligned} \frac{\partial}{\partial t}\phi(t,T,u) &= -\frac{1}{2}\psi(t,T,u)^\top a(t)\psi(t,T,u) - b(t)^\top \psi(t,T,u), \\ \phi(T,T,u) &= 0 \\ \frac{\partial}{\partial t}\psi_i(t,T,u) &= -\frac{1}{2}\psi(t,T,u)^\top \alpha_i(t)\psi(t,T,u) - \beta_i(t)^\top \psi(t,T,u), \quad 1 \le i \le d \\ \psi(T,T,u) &= u, \end{aligned} \tag{2.10}$$

for $0 \le t \le T$*.*

Part II: *Assume* $a(t,x)$ *and* $b(t,x)$ *are affine in* x *as specified in* (2.9) *and that there exists a solution* (ϕ,ψ) *to the Riccati equations* (2.10) *such that*

$$\Re\left\{\phi(t,T,u)+\psi(t,T,u)^\top x\right\}\le 0, \tag{2.11}$$

for all $0\le t\le T$, $u\in \mathrm{i}\mathbb{R}^d$ *and* $x\in\mathcal{X}$*. Then* $\mathbf{X}$ *is affine.*

Remark 2.7. In the Riccati equations (2.10), ϕ is determined directly from ψ by

$$\phi(t,T,u)=\int_t^T\left(\frac{1}{2}\psi(s,T,u)^\top a(s)\psi(s,T,u)+b(s)^\top\psi(s,T,u)\right)\,\mathrm{d}s.$$

$\diamond$

Proof. *Part I:* Assume $\mathbf{X}$ is affine. Then there exists ϕ and ψ such that

$$M(t)=e^{\phi(t,T,u)+\psi(t,T,u)^\top X(t)}$$

is a complex valued martingale for $u\in\mathrm{i}\mathbb{R}$. Applying Itô's formula, one obtains,

$$\begin{aligned}
\frac{\mathrm{d}M(t)}{M(t)} &= \left(\frac{\partial}{\partial t}\phi(t,T,u)+\sum_{i=1}^d\frac{\partial}{\partial t}\psi_i(t,T,u)X_i(t)\right)\,\mathrm{d}t+\sum_{i=1}^d\psi_i(t,T,u)\,\mathrm{d}X_i(t)\\
&\quad+\frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d\psi_i(t,T,u)\psi_j(t,T,u)\,\mathrm{d}\langle X_i(t),X_j(t)\rangle\\
&=\left(\frac{\partial}{\partial t}\phi(t,T,u)+\frac{\partial}{\partial t}\psi(t,T,u)^\top X(t)+\psi(t,T,u)^\top b(t,X(t))\right)\,\mathrm{d}t\\
&\quad+\psi(t,T,u)^\top\rho(t,X(t))\,\mathrm{d}W(t)\\
&\quad+\frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d\sum_{k=1}^d\psi_i(t,T,u)\rho_{ik}(t,X(t))\rho_{jk}(t,X(t))\psi_j(t,T,u)\,\mathrm{d}t\\
&=\left(\frac{\partial}{\partial t}\phi(t,T,u)+\frac{\partial}{\partial t}\psi(t,T,u)^\top X(t)+\psi(t,T,u)^\top b(t,X(t))\right.\\
&\quad\left.+\frac{1}{2}\psi(t,T,u)^\top a(t,X(t))\psi(t,T,u)\right)\,\mathrm{d}t\\
&\quad+\psi(t,T,u)^\top\rho(t,X(t))\,\mathrm{d}W(t).
\end{aligned}$$

By Assumption 2.1 the process can be started in all points $(t_0,x)\in\mathbb{R}_+\times\mathcal{X}$, thus the drift must equal zero for all $T\ge t\ge 0$, $x\in\mathcal{X}$ and $u\in\mathrm{i}\mathbb{R}^d$, i.e.

$$\begin{aligned}
0&=\frac{\partial}{\partial t}\phi(t,T,u)+\frac{\partial}{\partial t}\psi(t,T,u)^\top x+\psi(t,T,u)^\top b(t,x)\\
&\quad+\frac{1}{2}\psi(t,T,u)^\top a(t,x)\psi(t,T,u).
\end{aligned}\tag{2.12}$$

Choose now x, y and z such that $x, y, x+z, y+z \in \mathcal{X}$. Inserting $x+z$ and x into (2.12) respectively, and subtracting, one obtains

$$\begin{aligned} 0 = {} & \frac{\partial}{\partial t}\psi(t,T,u)^\top z + \psi(t,T,u)^\top \left(b(t,x+z) - b(t,x)\right) \\ & + \frac{1}{2}\psi(t,T,u)^\top \left(a(t,x+z) - a(t,x)\right)\psi(t,T,u). \end{aligned}$$

This can be done with the points $y+z$ and y as well. Doing that, and subtracting from the above, yields

$$\begin{aligned} 0 = {} & \psi(t,T,u)^\top \left[b(t,x+z) - b(t,x) - (b(t,y+z) - b(t,y))\right] \\ & + \frac{1}{2}\psi(t,T,u)^\top \left[a(t,x+z) - a(t,x) - (a(t,y+z) - a(t,y))\right]\psi(t,T,u). \end{aligned}$$

Letting $t = T$ such that $\psi(t,t,u) = u$ and remembering that this has to hold for all $u \in \mathrm{i}\mathbb{R}^d$, the following equations are obtained,

$$\begin{aligned} 0 &= b(t,x+z) - b(t,x) - (b(t,y+z) - b(t,y)) \\ 0 &= \left[a(t,x+z) - a(t,x) - (a(t,y+z) - a(t,y))\right]u. \end{aligned}$$

From this, we conclude that $a(t,\cdot)$ and $b(t,\cdot)$ are affine on $\mathcal{X}$ of the form (2.9).

Inserting (2.9) into (2.12),

$$\begin{aligned} 0 = {} & \frac{\partial}{\partial t}\phi(t,T,u) + \frac{\partial}{\partial t}\psi(t,T,u)^\top x + \psi(t,T,u)^\top \left(b(t) + \mathcal{B}(t)x\right) \\ & + \frac{1}{2}\psi(t,T,u)^\top \left(a(t) + \sum_{i=1}^{d} x_i \alpha_i(t)\right)\psi(t,T,u), \end{aligned}$$

and since this has to hold for all $x \in \mathcal{X}$, one obtains

$$\begin{aligned} \frac{\partial}{\partial t}\phi(t,T,u) &= -\psi(t,T,u)^\top b(t) - \frac{1}{2}\psi(t,T,u)^\top a(t)\psi(t,T,u) \\ \frac{\partial}{\partial t}\psi(t,T,u)^\top x &= -\sum_{i=1}^{d}\left(\psi(t,T,u)^\top \beta_i(t)x_i + \frac{1}{2}\psi(t,T,u)^\top x_i\alpha_i(t)\psi(t,T,u)\right). \end{aligned}$$

This is the differential equation system (2.10), if we remember the boundary conditions from Remark 2.3.

Part II: Assume $a(t,\cdot)$ and $b(t,\cdot)$ are affine, i.e. they are of the form (2.9), and that there exists a solution (ϕ,ψ) to the Riccati equations (2.10) satisfying (2.11). Then define

$$M(t) = e^{\phi(t,T,u) + \psi(t,T,u)^\top X(t)}.$$

Since ϕ and ψ satisfy the Riccati equations, **M** is a local martingale. Also, by (2.11), **M** is uniformly bounded, thus a true martingale. We conclude that (2.2) holds, and that **X** is affine. □

The theorem provides an easy way to check if a process is affine, and the affine transformation formula holds, namely if the parameters a and b are affine of the form (2.9). However, it is not all functions a and b of the form (2.9) that will lead to affine processes. It is also required that there exists a solution to the Riccati equations (2.10) for the process to be affine, which is not trivial. Also, it is required that

$$a(t,x) = a(t) + \sum_{i=1}^{d} x_i \alpha_i(t)$$

is symmetric and positive semi-definite, since $a(t,x) = \rho(t,x)\rho(t,x)^\top$.

To properly characterise affine processes such that the question of a solution to the Riccati equations are answered, specific state spaces are considered. That is, we find the set of admissible parameters

$$a(t), (\alpha_i(t))_{i=1,\dots,d}, b(t), \mathcal{B}(t),$$

such that **X** is affine, for specific kind of state spaces. The following section, Section 2.3, addresses general stochastic processes, and not only affine proceses. It contains a characterisation of the parameters that ensure that a process belongs to a specific state space. The section after, Section 2.4, utilises this to find the set of admissible parameters for affine processes, resulting in Theorem 2.12.

2.3 The State Space of a Process

When does a stochastic process belong to a specific state space? Considering general state spaces H of the form

$$H = \bigcap_{i=1}^{m} H_i, \quad H_i = \left\{ x \in \mathbb{R}^d \mid u_i^\top x \geq 0 \right\},$$

for vectors $u_1, \dots, u_m$, the question is addressed for general stochastic processes that satisfies a stochastic differential equation, thus not only affine processes. Often $u_i = e_i$, the unit vector with ith entry equal to 1, and the rest equal to 0, and in that case, H_i is the set of vectors where the ith entry is non-negative.

The following lemma is the essential result of the section, and identifies necessary and sufficient conditions on the parameters of a stochastic process, ensuring that the process belongs to a specific state space. The lemma considers time-homogeneous processes, and is presented in [7], Lemma 10.11, for $m = 1$. The contribution here is the, somewhat simple, extension of the proof to $m > 1$, together with the detailed presentation of the proof. Below, in Corollary 2.9, the result is extended to time-inhomogeneous processes, which is not proved in [7].

Lemma 2.8. *Let* $\mathbf{X}$ *be a time-homogeneous d-dimensional diffusion process, defined by*

$$\begin{aligned} \mathrm{d}X(t) &= b(X(t))\,\mathrm{d}t + \rho(X(t))\,\mathrm{d}W(t), \\ X(0) &= x, \end{aligned}$$

where b *and* ρ *are continuous. We assume that* $\mathbf{X}$ *is defined on a state space* $\mathcal{X} \subset \mathbb{R}^d$*. For some* $m \geq 1$*, let* $H \subset \mathcal{X}$ *be a set of the form* $H = \bigcap_{i=1}^m H_i$*, where*

$$H_i = \left\{ x \in \mathbb{R}^d \mid u_i^\top x \geq 0 \right\},$$

for vectors $u_1, \dots, u_m$*,* $i = 1, \dots, m$*. Let* ∂H_i *denote the boundary of* H_i*, i.e.* $\partial H_i = \left\{ x \in \mathbb{R}^d \mid u_i^\top x = 0 \right\}$*.*

Then, $X^x(t) \in H$ *for all start values* $x \in H$ *if and only if the system of equations*

$$\begin{aligned} u_i^\top b(x) &\geq 0, \\ u_i^\top a(x) u_i &= 0, \end{aligned} \tag{2.13}$$

holds for all $x \in \partial H_i \cap H$ *and* $i = 1, \dots, m$*. Here,* $a(x) = \rho(x)\rho(x)^\top$*.*

Proof. First assume that $X^x(t) \in H$ for all $x \in H$. Let $i \in \{1, \dots, m\}$, and let $x \in \partial H_i \cap H$. Then $\mathbf{X}^x$ exists, and we can consider the process $u_i^\top X^x(t)$ (denoted $u_i^\top X(t)$ from now on),

$$u_i^\top X(t) = \int_0^t u_i^\top b(X(s))\,\mathrm{d}s + \int_0^t u_i^\top \rho(X(s))\,\mathrm{d}W(s). \tag{2.14}$$

For $K > 0$, define the stopping time τ_1 by

$$\tau_1 = \inf \left\{ t \in \mathbb{R}_+ \mid \left| u_i^\top b(X(t)) \right| > K \text{ or } u_i^\top a(X(t)) u_i > K \right\},$$

with $\inf \emptyset = +\infty$. Then, for $t \geq 0$, we have,

$$\left|u_i^\top b(X^{\tau_1}(t))\right| \leq K, \quad u_i^\top a(X^{\tau_1}(t))u_i = \left\|u_i^\top \rho(X^{\tau_1}(t))\right\|^2 \leq K,$$

and the stopped process $\int_0^{t\wedge\tau_1} u_i^\top \rho(X(s))\,\mathrm{d}W(s)$ is a true martingale. Thus

$$\mathrm{E}\left[u_i^\top X^{\tau_1}(t)\right] = \mathrm{E}\left[\int_0^{\tau_1\wedge t} u_i^\top b(X(s))\,\mathrm{d}s\right].$$

Now, assume for contradiction that $u_i^\top b(x) < 0$. Let $0 < \varepsilon < -u_i^\top b(x)$, and define the stopping time

$$\tau_2 = \inf\left\{t \in \mathbb{R}_+ \mid u_i^\top b(X(t)) > -\varepsilon\right\}.$$

By continuity of b and X, we have $\tau_2 > 0$. Now

$$\mathrm{E}\left[u_i^\top X^{\tau_1\wedge\tau_2}(t)\right] = \mathrm{E}\left[\int_0^{\tau_1\wedge\tau_2\wedge t} u_i^\top b(X(s))\,\mathrm{d}s\right] \leq -\varepsilon\,\mathrm{E}\left[\tau_1\wedge\tau_2\wedge t\right].$$

We can choose K such that $\tau_1 > 0$, thus $\tau_1\wedge\tau_2\wedge t > 0$ and we have a contradiction. We conclude that $u_i^\top b(x) \geq 0$.

Let $C > 0$ and define $Y(t) = -C\int_0^t u_i^\top \rho(X(s))\,\mathrm{d}W(s)$. Let $Z(t)$ be the Doléans-Dade exponential of $\mathbf{Y}$,

$$Z(t) = e^{Y(t)-\frac{1}{2}\langle Y(t)\rangle},$$

implying that $\mathrm{d}Z(t) = Z(t)\,\mathrm{d}Y(t)$ and $\mathbf{Z}$ is a local martingale. We have

$$\mathrm{d}Z(t) = -CZ(t)u_i^\top \rho(X(t))\,\mathrm{d}W(t).$$

We are interested in the process $u_i^\top X(t)Z(t)$. An application of Itô's formula yields

$$u_i^\top X(t)Z(t) = \int_0^t u_i^\top X(s)\,\mathrm{d}Z(s) + \int_0^t Z(s)\,\mathrm{d}\left(u_i^\top X(s)\right) + \int_0^t \mathrm{d}\langle u_i^\top X(s), Z(s)\rangle.$$

We find

$$\begin{aligned}\langle u_i^\top X(t), Z(t)\rangle &= \int_0^t (-C)\,Z(s)u_i^\top \rho(X(s))\left(u_i^\top \rho(X(s))\right)^\top\,\mathrm{d}s \\ &= -C\int_0^t Z(s)u_i^\top a(X(s))u_i\,\mathrm{d}s.\end{aligned}$$

Using (2.14), we have, for some local martingale $\mathbf{M}$,

$$u_i^\top X(t)Z(t) = \int_0^t Z(s)\left(u_i^\top b(X(s)) - Cu_i^\top a(X(s))u_i\right)\,\mathrm{d}s + M(t).$$

Since $X^{\tau_1}(t)$ is bounded, then $Y^{\tau_1}(t)$ is bounded, and therefore $Z^{\tau_1}(t)$ and also $u_i^\top X^{\tau_1}(t)Z^{\tau_1}(t)$ are bounded. By Remark A.14 $\mathbf{M}^{\tau_1}$ is a martingale, and

$$\mathrm{E}\left[u_i^\top X^{\tau_1}(t)Z^{\tau_1}(t)\right] = \mathrm{E}\left[\int_0^{\tau_1\wedge t} Z(s)\left(u_i^\top b(X(s)) - Cu_i^\top a(X(s))u_i\right)\,\mathrm{d}s\right].$$

Now, assume for contradiction that $u_i^\top a(x)u_i > 0$. Let $0 < \varepsilon < u_i^\top a(x)u_i$ and define the stopping time

$$\tau_3 = \inf\left\{t \in \mathbb{R}_+ \mid u_i^\top a(X(t))u_i < \varepsilon\right\}.$$

By continuity of a and $\mathbf{X}$, $\tau_3 > 0$. Let $C > \frac{K}{\varepsilon}$ and note that

$$\begin{aligned}\mathrm{E}\left[u_i^\top X^{\tau_1\wedge\tau_3}(t)Z^{\tau_1\wedge\tau_3}(t)\right] &= \mathrm{E}\left[\int_0^{\tau_1\wedge\tau_3\wedge t} Z(s)\left(u_i^\top b(X(s)) - Cu_i^\top a(X(s))u_i\right)\,\mathrm{d}s\right]\\ &\le (K - C\varepsilon)\,\mathrm{E}\left[\int_0^{\tau_1\wedge\tau_3\wedge t} Z(s)\,\mathrm{d}s\right] \qquad (2.15)\\ &< 0.\end{aligned}$$

Here, we used that $Z(t) > 0$ and that $\tau_1 \wedge \tau_3 \wedge t > 0$, making the expectation in (2.15) strictly larger than 0. We have a contradiction, since $u_i^\top X(t)Z(t) \ge 0$, and we conclude that $u_i^\top a(x)u_i = 0$.

Finally, since i and x were arbitrary, we conclude that (2.13) must hold for all $x \in \partial H_i \cap H$, for $i = 1, \ldots, m$.

Now assume that the equation system (2.13) holds for all $x \in \partial H_i \cap H$, $i = 1, \ldots, m$. Define $\tilde{b}$ and $\tilde{\rho}$ by

$$\begin{aligned}\tilde{b}(x) &= \begin{cases} b(x), & x \in H,\\ 0, & x \notin H,\end{cases}\\ \tilde{\rho}(x) &= \begin{cases} \rho(x), & x \in H,\\ 0, & x \notin H,\end{cases}\end{aligned} \qquad (2.16)$$

and consider the process $\widetilde{\mathbf{X}}$ defined by

$$\mathrm{d}\widetilde{X}(t) = \tilde{b}(\widetilde{X}(t))\,\mathrm{d}t + \tilde{\rho}(\widetilde{X}(t))\,\mathrm{d}W(t),$$

with $\widetilde{X}(0) = X(0) = x$. Now, let $i \in \{1, \ldots, m\}$ and define the stopping time

$$\sigma_{\delta,\varepsilon} = \inf \left\{ t \geq \delta \mid u_i^\top X(t) < -\varepsilon < u_i^\top X(t-\delta),\ \forall s \in [t-\delta, t) : u_i^\top X(s) \leq 0 \right\},$$

for $\delta, \varepsilon > 0$, with $\inf \emptyset = +\infty$. In words, it is the first time the process is below ε and has been below 0 in a timespan of δ. Let $\delta, \varepsilon \in \mathbb{Q}_{++}$ and consider the process on the set $\{\sigma_{\delta,\varepsilon} < \infty\}$. On the interval $[t-\delta, t]$, we have $\widetilde{X}(t) \in \partial H_i \cap H$ or $\widetilde{X}(t) \in \complement H$, and by (2.16) and (2.13) one sees that

$$\begin{aligned} 0 &> u_i^\top \widetilde{X}(\sigma_{\delta,\varepsilon}) - u_i^\top \widetilde{X}(\sigma_{\delta,\varepsilon} - \delta) \\ &= \int_{\sigma_{\delta,\varepsilon}-\delta}^{\sigma_{\delta,\varepsilon}} u_i^\top b(\widetilde{X}(s))\,\mathrm{d}s + \int_{\sigma_{\delta,\varepsilon}-\delta}^{\sigma_{\delta,\varepsilon}} u_i^\top \rho(\widetilde{X}(s))\,\mathrm{d}W(s) \\ &\geq 0. \end{aligned}$$

Thus, the stochastic differential equation is not satisfied on the set, and $P(\sigma_{\delta,\varepsilon} < +\infty) = 0$. By sigma additivity, since $\mathbb{Q}_{++}^2$ is countable,

$$P\left(\bigcup_{(\delta,\varepsilon)\in\mathbb{Q}_{++}^2} \{\sigma_{\delta,\varepsilon} < +\infty\} \right) = 0.$$

Now, for all $\delta, \varepsilon \in \mathbb{R}_{++}$ where $\sigma_{\delta,\varepsilon} < \infty$, there exists $\tilde{\delta}, \tilde{\varepsilon} \in \mathbb{Q}_{++}$ such that $\sigma_{\tilde{\delta},\tilde{\varepsilon}} < \infty$. Thus

$$P\left(\bigcup_{(\delta,\varepsilon)\in\mathbb{R}_{++}^2} \{\sigma_{\delta,\varepsilon} < +\infty\} \right) = 0.$$

In other terms, for all $i = 1, \ldots, m$ and $t \in \mathbb{R}_+$, $\widetilde{X}(t) \in H_i$, i.e. $\widetilde{X}(t) \in H$ a.s.

For all start values $x \in H$, one sees from the stochastic differential equations defining $\mathbf{X}$ and $\widetilde{\mathbf{X}}$ that $\mathbf{X}^x = \widetilde{\mathbf{X}}^x$ a.s. We conclude that $X^x(t) \in H$ for all $t \geq \mathbb{R}_+$ and $x \in H$ a.s. □

From the lemma it is seen, that if (2.13) is satisfied for a stochastic process $\mathbf{X}$ with state space $\mathcal{X}$, we can set the state space $\mathcal{X} = H$ and be sure that $X(t) \in H$ for $t \in \mathbb{R}_+$.

The lemma can easily be extended to time-inhomogeneous processes, as is stated in the following corollary. Assumption 2.1, which holds throughout the thesis, is essential to this extension.

Corollary 2.9. *Let* $\mathbf{X}$ *be defined by* (2.1)*, satisfying Assumption 2.1. For some* $m \geq 1$*, let* $H = \bigcap_{i=1}^{m} H_i$ *where*

$$H_i = \left\{ x \in \mathbb{R}^d \mid u_i^\top x \geq 0 \right\},$$

for vectors $u_1, \dots u_m$*,* $i = 1, \dots, m$*.*

Then, $X^x(t) \in H$ *for all start values* $x \in H$ *and time points* $t \in \mathbb{R}_+$ *if and only if the equation system*

$$\begin{aligned} u_i^\top b(t,x) &\geq 0, \\ u_i^\top a(t,x) u_i &= 0, \end{aligned} \tag{2.17}$$

holds for all $t \geq 0$*,* $x \in \partial H_i \cap H$ *and* $i = 1, \dots, m$*.*

Proof. Let $\widetilde{\mathbf{X}}$ be the time-homogenised process from (2.4). Let, for $i = 1, \dots, m$,

$$v_i = \begin{bmatrix} 0 \\ u_i \end{bmatrix},$$

and let $\widetilde{H} = \bigcap_{i=1}^{m+1} \widetilde{H}_i$ where, for $i = 1, \dots, m$,

$$\widetilde{H}_i = \left\{ x \in \mathbb{R}^{d+1} \mid v_i^\top x \geq 0 \right\},$$

and $\widetilde{H}_{m+1} = \mathbb{R}_+ \times \mathbb{R}^d$ (corresponding to $v_{m+1} = (1, 0, \dots, 0)^\top$). From Lemma 2.8 we conclude that $\widetilde{\mathbf{X}} \in \widetilde{H}$, thus $\mathbf{X} \in H$, if and only if (2.17) is satisfied. □

A comment should be made on the start values. The process is in the state space H for all start values, if and only if the conditions on the parameters hold. If only some start values are considered, then the result does not necessarily holds. I.e. even if the conditions (2.17) are not satisfied, the process might still belong to the state space for some start values. However, if the conditions (2.17) are satisfied, then for any start value in the state space, the process belongs to the state space. Thus, the conditions are always sufficient.

The corollary is not only interesting when considering stochastic processes. It is also applicable to deterministic functions solving a differential equation system, which is seen by letting $a(t,x) = 0$. As it is presented in Corollary 2.9, the boundary condition must be in the left endpoint of the interval considered, but with a modification, a similar result can be obtained if the endpoint is in the right endpoint of the interval considered. The corollary is only stated one way, providing a sufficient condition for the solution being in H.

Corollary 2.10. *Let H be as in Corollary 2.9 and let J be an interval containing T. Assume the function $f : J \to \mathbb{R}^d$ exists, where f is satisfies the differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = b(t, f(t)), \quad f(T) = u,$$

for $u \in H$. If

$$u_i^\top b(t, x) \leq 0,$$

for all $t \in (-\infty, T] \cap J$, $x \in \partial H_i \cap H$ and $i = 1, \dots, m$, then $f(t) \in H$ for all $t \in (-\infty, T] \cap J$.

Proof. Let $g(t) = f(T - t)$. Then $g(0) = u$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t) = -b(T - t, f(T - t)) = -b(T - t, g(t)).$$

For $i \in \{1, \dots, m\}$, $x \in \partial H_i \cap H$ and $t \geq 0$ such that $T - t \in J$, we have $T - t \in (-\infty, T] \cap J$, thus

$$-u_i^\top b(T - t, x) \geq 0.$$

By Corollary 2.9, $g(t) \in H$ for $t \geq 0$ such that $T - t \in J$, thus $f(t) \in H$ for $t \in (-\infty, T] \cap J$. □

Corollary 2.9 is the main result of this section, and is used to find the set of admissible parameters in the following section. Also, the corollaries are applied to deterministic functions that satisfy differential equations, primarily the Riccati equations in the proof of Theorem 2.12.

2.4 Canonical State Space

Here and in the following we consider the canonical state space

$$\mathcal{X} = \mathbb{R}_+^m \times \mathbb{R}^n, \tag{2.18}$$

with $d = m + n \geq 1$ and $m, n \geq 0$, and find admissible parameters. The state space covers most practical applications, although one can find examples of other state spaces.

It is convenient to introduce some further notation. Let

$$I = \{1, \dots, m\}, \quad J = \{m+1, \dots, d\}.$$

For a vector γ, a matrix Γ and index sets M and N, let

$$\gamma_N = (\gamma_i)_{i \in N}, \quad \Gamma_{NM} = (\Gamma_{ij})_{i \in N, j \in M}.$$

Before stating the main theorem, we realise that Corollary 2.9 is applicable to the canonical state space (2.18), resulting in a slightly different version of the equations (2.17). We put this in the following remark.

Remark 2.11. Let $u_i = e_i$ for $i = 1, \dots, m$, where e_i is the unit vector where the ith entry is equal to 1 and all other entries are equal to 0. Then

$$H = \bigcap_{i=1}^{m} H_i = \mathbb{R}_+^m \times \mathbb{R}^n,$$

and from (2.17) we have $b_i(t,x) \geq 0$ and $a_{ii}(t,x) = 0$ when $x_i = 0$. Since a is required to be symmetric and positive semi-definite, then, using Lemma A.1, the equations (2.17) read

$$\begin{aligned} b_i(t,x) &\geq 0, \\ a_{ik}(t,x) = a_{ki}(t,x) &= 0, \quad k \in \{1, \dots, d\}, \end{aligned} \tag{2.19}$$

for all $x \in \mathbb{R}_+^m \times \mathbb{R}^n$ with $e_i^\top x = x_i = 0$, $i \in I$ and $t \in \mathbb{R}_+$. ◇

The conditions (2.19) holds exactly when the process $\mathbf{X}^x$ is in the canonical state space for all start values x. By combining the results here with the affine parameters from (2.9), this is used, in part, to show the following theorem. The theorem states necessary and sufficient conditions on the parameters, for when $\mathbf{X}$ is an affine process in the canonical state space, and it is a powerful tool for determining what processes are affine when some of the dimensions are positive-valued.

The theorem is analogous to Theorem 10.2 in [7], with the difference being the extension to time-inhomogeneous parameters. This results in a slighty more general version of the Riccati equations (2.21), which in turn demands a non-trivial extension[2] of the proof. The proof is also carried out in detail below, which is not the case in [7].

[2]For the time-homogeneous case, the matrix $\mathcal{B}$ is constant, and in that case the solution ψ_J is guaranteed with a closed form expression.

Theorem 2.12. *Let* **X** *be the solution to* (2.1)*, satisfying Assumption 2.1. Then* **X** *is affine on the canonical state space* $\mathbb{R}_+^m \times \mathbb{R}^n$ *if and only if* $a(t,x)$ *and* $b(t,x)$ *are affine of the form* (2.9)*, for continuous parameter functions* $a(t)$, $(\alpha_i(t))_{i=1,\dots,d}$, $b(t)$, $\mathcal{B}(t)$ *satisfying*

$$\begin{aligned} a(t), \alpha_i(t) &\text{ are symmetric and positive semi-definite, } i \in I, \\ a_{II}(t) &= 0, \\ a_{IJ}(t) = a_{JI}(t)^\top &= 0, \\ \alpha_j(t) &= 0, \quad j \in J, \\ \alpha_{i,kl}(t) = \alpha_{i,lk}(t) &= 0, \quad k \in I \setminus \{i\}, i \in I, l \in \{1, \dots, d\}, \\ b(t) &\in \mathbb{R}_+^m \times \mathbb{R}^n, \\ \mathcal{B}_{IJ}(t) &= 0, \\ \beta_{i,k}(t) &\geq 0, \quad k \in I \setminus \{i\}, i \in I, \end{aligned} \tag{2.20}$$

for all $t \in \mathbb{R}_+$.

In this case, the Riccati equations (2.10) *simplify to,*

$$\begin{aligned} \frac{\partial}{\partial t}\phi(t,T,u) &= -\frac{1}{2}\psi_J(t,T,u)^\top a_{JJ}(t)\psi_J(t,T,u) - b(t)^\top \psi(t,T,u), \\ \phi(T,T,u) &= 0, \\ \frac{\partial}{\partial t}\psi_i(t,T,u) &= -\frac{1}{2}\psi(t,T,u)^\top \alpha_i(t)\psi(t,T,u) - \beta_i(t)^\top \psi(t,T,u), \quad i \in I \\ \frac{\partial}{\partial t}\psi_J(t,T,u) &= -\mathcal{B}_{JJ}(t)^\top \psi_J(t,T,u), \\ \psi(T,T,u) &= u, \end{aligned} \tag{2.21}$$

and there exists a unique solution $(\phi(\cdot,T,u), \psi(\cdot,T,u)) : \mathbb{R}_+ \to \mathbb{C}_- \times \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ *for all* $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ *and* $T > 0$.

Before proving the theorem, the conditions on the parameters are exemplified for the case $d = 3$, $m = 2$, $n = 1$. For the drift,

$$b(t) = \begin{bmatrix} + \\ + \\ * \end{bmatrix}, \quad \mathcal{B}(t) = \begin{bmatrix} \mathcal{B}_{II}(t) & \mathcal{B}_{IJ}(t) \\ \mathcal{B}_{JI}(t) & \mathcal{B}_{JJ}(t) \end{bmatrix} = \begin{bmatrix} * & + & 0 \\ + & * & 0 \\ * & * & * \end{bmatrix},$$

and for the diffusion,

$$a(t) = \begin{bmatrix} a_{II}(t) & a_{IJ}(t) \\ a_{JI}(t) & a_{JJ}(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & + \end{bmatrix},$$

and

$$\alpha_1(t) = \begin{bmatrix} + & 0 & * \\ & 0 & 0 \\ & & + \end{bmatrix}, \quad \alpha_2(t) = \begin{bmatrix} 0 & 0 & 0 \\ & + & * \\ & & + \end{bmatrix}, \quad \alpha_3(t) = 0.$$

Here, $+$ denotes a non-negative real function, and $*$ denotes any real function, such that a and α_i, $i = 1, 2, 3$ are symmetric and positive semi-definite for all t.

Proof. Assume that $\mathbf{X}$ is affine on $\mathbb{R}^m_+ \times \mathbb{R}^n$. Then by Theorem 2.6, $a(t, x)$ and $b(t, x)$ are affine of the form (2.9). Since $a(t, x)$ must be symmetric and positive semi-definite, for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^m_+ \times \mathbb{R}^n$, we require $a(t)$ and $\alpha_i(t)$ to be symmetric and positive semi-definite for $i \in I$, and $\alpha_j(t) = 0$ for $j \in J$.

With Remark 2.11, one obtains for $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$ with $x_k = 0$, for $k \in I$,

$$b_k(t) + \sum_{i \in I \setminus \{k\}} \beta_{i,k}(t) x_i + \sum_{j \in J} \beta_{j,k}(t) x_j \geq 0,$$
$$a_{kl}(t) + \sum_{i \in I \setminus \{k\}} \alpha_{i,kl}(t) x_i = 0,$$

for all $l \in \{1, \ldots, d\}$ and $t \in \mathbb{R}_+$. The conditions on $b(t)$ and $\mathcal{B}$ in (2.20) are seen from the following;

- Inserting $x = 0$, one realises that $b(t) \in \mathbb{R}^m_+ \times \mathbb{R}^n$.
- Since $x_J \in \mathbb{R}^m$ has both positive and/or negative entries, $\beta_{j,I}(t) = 0$ for $j \in J$, i.e. $\mathcal{B}_{IJ}(t) = 0$.
- Since $x_i \in \mathbb{R}_+$ for $i \in I \setminus \{k\}$ then $\beta_{i,k}(t) \geq 0$, for $i \in I \setminus \{k\}$, $k \in I$.

The remaining conditions on $a(t)$ and $(\alpha_i(t))_{i \in I}$ are seen from the following;

- Inserting x with $x_I = 0$ we have $a_{kl}(t) = 0$ for $k \in I$ and $l \in \{1, \ldots, d\}$, thus $a_{II}(t) = 0$ and $a_{IJ}(t) = 0$.
- Inserting x such that x_i, $i \in I$ is the only non-zero entry, we obtain $\alpha_{i,kl}(t) = 0$ for $k \in I \setminus \{i\}$ and $l \in \{1, \ldots, d\}$. By symmetry, also $\alpha_{i,lk}(t) = 0$.

Thus the necessary conditions (2.20) on the parameters are shown.

Now assume that $a(t, x)$ and $b(t, x)$ are affine of the form (2.9) and satisfy the conditions (2.20). Using (2.20), for $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$, with $x_k = 0$ for $k \in I$, we see that

$$b_k(t, x) = b_k(t) + \sum_{i \in I \setminus \{k\}} \beta_{i,k}(t) x_i \geq 0,$$

and for $l \in \{1, \ldots, d\}$,

$$a_{kl}(t,x) = \sum_{i \in I \setminus \{k\}} \alpha_{i,kl}(t) x_i = 0,$$

and by Remark 2.11, $\mathbf{X}$ has state space $\mathbb{R}_+^m \times \mathbb{R}^n$.

It remains to show that $\mathbf{X}$ is affine. By Theorem 2.6 it is sufficient to show that there exists a solution $(\phi(t,T,u), \psi(t,T,u))$ to the Riccati equations satisfying (2.11) for all $0 \le t \le T$, $u \in \mathrm{i}\mathbb{R}^d$ and $x \in \mathbb{R}_+^m \times \mathbb{R}^n$. We will for the more general case $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ show that there exists a solution in $(\phi, \psi) \in \mathbb{C}_- \times \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ implying that (2.11) holds.

First, realise that inserting the parameters from (2.20) into the Riccati equations (2.10) yields the Riccati equations of the form (2.21).

Consider the differential equation for $\psi_J(t,T,u)$. Since, for $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$, $\Re \frac{\partial}{\partial t} \psi_J(t,T,u) = 0$ for $t = T$, we see that $\Re \psi_J(t,T,u) = 0$ is a solution for $t < T$. By Theorem A.3 (a) it is unique. Thus, $\psi_J(t,T,u) \in \mathrm{i}\mathbb{R}$ on the domain where the solution exists. We now show that the solution exists on $[0,T]$. An application of Lemma A.9 to $\psi_J(t,T,u)$ yields

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \|\psi_J(t,T,u)\|^2 &= -2\Re \left(\overline{\psi_J(t,T,u)}^\top \mathcal{B}_{JJ}(t)^\top \psi_J(t,T,u) \right) \\ &= -2 \sum_{j \in J} \Re \left(\overline{\psi_j(t,T,u)} \mathcal{B}_{Jj}(t)^\top \psi_J(t,T,u) \right). \end{aligned}$$

Since $\mathcal{B}(t)$ is continuous, each entry has a maximum and minimum on the compact interval $[0,T]$, thus there exists $0 \le K < \infty$ such that $|\mathcal{B}_{kl}(t)| \le K$ for all $k, l \in \{1, \ldots, d\}, t \in [0,T]$. Now see, for $v_J \in \mathrm{i}\mathbb{R}^n$,

$$\begin{aligned} \left| 2\Re \left(\overline{v}_j \mathcal{B}_{Jj}(t)^\top v_J \right) \right| &= 2 \left| \Im(v_j) \mathcal{B}_{Jj}(t)^\top \Im(v_J) \right| \\ &= 2 \left| \sum_{k \in J} \Im(v_j) \mathcal{B}_{kj}(t) \Im(v_k) \right| \\ &\le 2K \sum_{k \in J} |\Im(v_j) \Im(v_k)| \\ &\le 2Kn \|v_J\|^2, \end{aligned}$$

using (A.5). Combining yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\psi_J(t,T,u)\|^2 \le 2Kn^2 \|\psi_J(t,T,u)\|^2,$$

and integrating and applying Grönwall's inequality, Lemma A.2,

$$\begin{aligned}\|\psi_J(t,T,u)\|^2 &\le \|u_J\|^2 + \left|\int_T^t 2Kn^2\|\psi_J(s,T,u)\|^2\,\mathrm{d}s\right| \\ &\le \|u_J\|^2 + \left|\int_T^t \|u_J\|^2 2Kn^2 e^{2Kn^2|t-s|}\,\mathrm{d}s\right| \\ &< \infty.\end{aligned}$$

By Theorem A.3 (b), a solution $\psi_J(\cdot,T,u)$ exists on $[0,T]$, for $T>0$ and $u \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$.

Consider now the differential equation for $\psi_i(t,T,u), i \in I$. Let R_i denote the right hand side, $\frac{\partial}{\partial t}\psi(t,T,u) = R(t,\psi(t,T,u))$, and see, for $v \in \mathbb{C}^m \times \mathrm{i}\mathbb{R}^n$, that

$$\begin{aligned}\Re R_i(t,v) &= \Re\left(-\frac{1}{2}v^\top \alpha_i(t) v - \beta_i(t)^\top v\right) \\ &= -\frac{1}{2}\Re\left[(\Re v + i\Im v)^\top \alpha_i(t)\,(\Re v + i\Im v)\right] - \beta_i(t)^\top \Re v \\ &= -\frac{1}{2}\Re v^\top \alpha_i(t)\Re v + \frac{1}{2}\Im v^\top \alpha_i(t) \Im v - \beta_{i,I}(t)^\top \Re v_I \\ &= -\frac{1}{2}\alpha_{i,ii}(t)(\Re v_i)^2 + \frac{1}{2}\alpha_{i,ii}(t)(\Im v_i)^2 - \beta_{i,I}(t)^\top \Re v_I.\end{aligned}$$

The differential equation system for $\Re\psi$ can be written as $\mathrm{d}\Re\psi(t) = b(t,\Re\psi(t))\,\mathrm{d}t$, $\psi(T,T,u) = u$ where

$$b_i(t,v) = -\frac{1}{2}\alpha_{i,ii}(t)(v_i)^2 + \frac{1}{2}\alpha_{i,ii}(t)(\Im\psi_i(t,T,u))^2 - \beta_{i,I}(t)^\top v_I,$$

for $i \in I$, and $b_j(t,v) = 0$ for $j \in J$. Consider now $\Re\psi(\cdot,T,u)$ on the interval where it exists. Corollary 2.10 can be applied here, with $u_i = -e_i$ for $i \in I$ giving $H = \mathbb{R}^m_- \times \mathbb{R}^n$; By positive semi-definiteness, $\alpha_{i,ii}(t) \ge 0$, thus $-b_i(t,v) \le 0$ for $v \in H$, $v_i = 0$ and all t. We conclude that $\Re\psi(t,T,u) \in H$. Seeing that also $\Re\psi_j = 0$, we have $\psi(t,T,u) \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$ for $u \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$.

It remains to show that $\psi_I(\cdot,T,u)$ exists on $[0,T]$. As above, Lemma A.9 is applied, and

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\psi_I(t,T,u)\|^2 = 2\sum_{i\in I}\Re\left(\overline{\psi_i(t,T,u)}R_i(t,\psi(t,T,u))\right).$$

By continuity on $[0,T]$, a constant $K<\infty$ can be chosen such that

$$\forall t \in [0,T]\ \forall l,k,m \in \{1,\dots,d\} : |\alpha_{l,km}(t)| < K \text{ and } |\mathcal{B}_{km}(t)| < K.$$

See that, for $v \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$,

$$\begin{aligned}
-\Re\left(\overline{v}_i R_i(t,v)\right) &= \frac{1}{2}\Re\left(\overline{v}_i v^\top \alpha_i(t) v\right) + \Re\left(\overline{v}_i \beta_i(t)^\top v\right) \\
&= \frac{1}{2}\Re\Big(\overline{v}_i \alpha_{i,ii}(t) v_i v_i + \overline{v}_i v_i \alpha_{i,iJ}(t) v_J \\
&\quad + \overline{v}_i v_J^\top \alpha_{i,Ji}(t) v_i + \overline{v}_i v_J^\top \alpha_{i,JJ}(t) v_J\Big) \\
&\quad + \Re\left(\overline{v}_i \beta_i(t)^\top v\right) \\
&\leq \frac{1}{2}\alpha_{i,ii}(t)|v_i|^2 \Re v_i + K|v_i|^2 \sum_{j\in J}|v_j| + \frac{1}{2}\left|\overline{v}_i v_J^\top \alpha_{i,JJ}(t) v_J\right| \\
&\quad + \left|\overline{v}_i \beta_{i,I}(t)^\top v_I + \overline{v}_i \beta_{i,J}(t)^\top v_J\right| \\
&\leq 0 + K\|v_I\|^2 n\left(1+\|v_J\|^2\right) + \frac{1}{2}(1+\|v_I\|^2) K n^2 \|v_J\|^2 \\
&\quad + K|v_i|\sum_{k\in I}|v_k| + K|v_i|\sum_{l\in J}|v_l| \\
&\leq \frac{1}{2}K\left(1+\|v_I\|^2\right)\left(3n^2\left(1+\|v_J\|^2\right)\right) \\
&\quad + Km\|v_I\|^2 + K(1+\|v_I\|^2)m(1+\|v_J\|^2) \\
&\leq \frac{\widetilde{K}}{2m}\left(1+\|v_I\|^2\right)\left(1+\|v_J\|^2\right),
\end{aligned}$$

where $\widetilde{K} \in (0,\infty)$ is independent of v. Here, the inequalities (A.4)–(A.6) were used.

We obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|\psi_I(t,T,u)\|^2 &= 2\sum_{i\in I}\Re\left(\overline{\psi_i(t,T,u)} R_i(t,\psi(t,T,u))\right) \\
&\geq -\widetilde{K}\left(1+\|\psi_I(t,T,u)\|^2\right)\left(1+\|\psi_J(t,T,u)\|^2\right),
\end{aligned}$$

and have $\psi_I(T,T,u) = u_I$, implying, with $t \leq T$,

$$\begin{aligned}
\|\psi_I(t,T,u)\|^2 &= \|u_I\|^2 - \int_t^T \frac{\mathrm{d}}{\mathrm{d}t}\|\psi_I(s,T,u)\|^2 \,\mathrm{d}s \\
&\leq \|u_I\|^2 + \int_t^T \widetilde{K}\left(1+\|\psi_I(s,T,u)\|^2\right)\left(1+\|\psi_J(s,T,u)\|^2\right)\mathrm{d}s.
\end{aligned}$$

Applying Grönwall's inequality, Lemma A.2, to the function $t \mapsto 1 + \|\psi_I(t,T,u)\|^2$

yields

$$
\begin{aligned}
&1+\|\psi_I(t,T,u)\|^2 \leq \\
&\qquad 1+\|u_I\|^2+\int_t^T (1+\|u_I\|^2)\widetilde{K}\left(1+\|\psi_J(s,T,u)\|^2\right) e^{\int_t^s \widetilde{K}(1+\|\psi_J(\tau,T,u)\|^2)\,\mathrm{d}\tau}\,\mathrm{d}s,
\end{aligned}
$$

and we see that $\|\psi_I(t,T,u)\|^2 < \infty$ on $[0,T]$. By Theorem A.3 (b), $\psi_I(\cdot,T,u)$ exists on $[0,T]$, $T \geq 0$, $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$.

By continuity of $\psi(\cdot,T,u)$, $a(\cdot,x)$ and $b(\cdot,x)$ we have that $\phi(\cdot,T,u)$, given by Remark 2.7, exists. By the conditions (2.20) and the fact that $\psi(t,T,u) \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^m$, we have $b(t)^\top\psi(t,T,u) \in \mathbb{C}_-$ and

$$
\begin{aligned}
\frac{1}{2}\psi(t,T,u)^\top a(t)\psi(t,T,u) &= \frac{1}{2}\psi_J(t,T,u)^\top a_{JJ}(t)\psi_J(t,T,u) \\
&= -\frac{1}{2}\Im\psi_J(t,T,u)^\top a_{JJ}(t)\Im\psi_J(t,T,u) \\
&< 0,
\end{aligned}
$$

using that $a_{JJ}(t)$ is positive semi-definite. Thus $\phi(t,T,u) \in \mathbb{C}_-$. □

We have obtained a characterisation of the affine processes in the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$, in form of a set of parameter functions. An interesting observation about the theorem is, that existense of a unique solution to the Riccati equations (2.21) is shown, not only for $u \in \mathrm{i}\mathbb{R}^d$, but for $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$. This will be exploited in the next section.

2.5 The Moment Generating Function

So far we have considered imaginary vectors u in the affine transformation formula, (2.2), which is then the (conditional) characteristic function. However, for applications, one is also interested in real vectors u, and hence the moment generating function. For example, if $\mathbf{X}$ represents mortality or interest rates, then an essential quantity to be studied is,

$$
\mathrm{E}\left[e^{-\int_t^T \gamma^\top X(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t)\right],
$$

for a real vector γ. If we let $\mathrm{d}Y(t) = \gamma^\top X(t)\,\mathrm{d}t$, this is a special case of the affine transformation formula for the affine process $(\mathbf{X},\mathbf{Y})$ and a real u, namely $u = (0,\ldots,0,-1)$.

Existence of the left side of the affine transformation formula (2.2),

$$\mathrm{E}\left[e^{u^\top X(T)} \middle| \mathcal{F}(t) \right] \tag{2.22}$$

is in general only guaranteed when $u \in \mathrm{i}\mathbb{R}^d$, which is why this is the assumption in Section 2. This allow us to classify the affine processes with Theorem 2.12. The moment generating function, i.e. (2.22) with real valued u, is not a useful tool for this task, since there exist affine processes where the moment generating function does not exist.

In order to apply the results of Section 2 in life insurance, integrability of the moment generating function has to be verified. Thus, the case $u \in \mathbb{R}^d$ is addressed.

Lemma 2.13. *Let* $\mathbf{X}$ *be an affine process and let* $u \in \mathbb{R}^d$, $T > 0$. *If there exists a solution to the Riccati equations* (2.21) *in* u *on* $[0, T]$ *and either of the following holds,*

1. $e^{\phi(t,T,u)+\psi(t,T,u)^\top X(t)}$ *is uniformly bounded on* $[0, T]$,
2. $\mathrm{E}\left[\int_0^T e^{2\phi(t,T,u)+2\psi(t,T,u)^\top X(t)} \psi(t,T,u)^\top a(t, X(t)) \psi(t,T,u)\,\mathrm{d}t\right] < \infty$,

then the affine transformation formula (2.2) *holds for* u, *for all* $t \in [0, T]$.

Proof. With (ϕ, ψ) being a solution to the Riccati equations (2.10), the process

$$M(t) = e^{\phi(t,T,u)+\psi(t,T,u)^\top X(t)}$$

is a local martingale. If we can show it is a martingale, the affine transformation formula holds, and $\mathbf{X}$ is affine.

For the case 1, the result is obtained by Remark A.14.

For the case 2, recall that, if

$$\mathrm{E}\langle M(T)\rangle < \infty, \tag{2.23}$$

then $\mathbf{M}$ is a a martingale for $t \leq T$. This is for example shown in [11], Lemma 8.2. We show that the condition (2.23) is the same as the assumption, 2. In the proof of Theorem 2.6, the dynamics of M is found, and we find

$$\begin{aligned}
\langle M(T)\rangle &= \left\langle \int_0^T M(t)\psi(t,T,u)^\top \rho(t, X(t))\,\mathrm{d}W(t) \right\rangle \\
&= \left\langle \int_0^T M(t) \sum_{i=1}^d \sum_{j=1}^d \psi_j(t,T,u)^\top \rho_{ji}(t, X(t))\,\mathrm{d}W_i(t) \right\rangle
\end{aligned}$$

$$= \sum_{i=1}^{d} \int_0^T M^2(t) \left(\sum_{j=1}^{d} \psi_j(t,T,u)^\top \rho_{ji}(t,X(t)) \right)^2 \mathrm{d}\langle W_i(t) \rangle$$
$$= \int_0^T M^2(t) \psi(t,T,u)^\top a(t,X(t)) \psi(t,T,u) \,\mathrm{d}t.$$

For the last equality, recall that $a(t,X(t)) = \rho(t,X(t))\rho(t,X(t))^\top$. □

The case 2 is not a handy tool for checking that the process **M** is a martingale, since it is not easily checked. For the case 1, things are more managable, and a large class of affine processes are naturally uniformly bounded, as we will see now.

Let $X(t) \in \mathbb{R}^d_+$, i.e. $m = d, n = 0$, and let $u \in \mathbb{R}^d_-$. Then $u^\top X(t) \leq 0$. This is a special case of Theorem 2.12 where $u \in \mathbb{R}^m_- \times \{0\}^n$. Now, $\Im u = 0$, and the functions with $\Im \psi(t,T,u) = 0$ are solutions to the imaginary part of the Riccati equations (2.21). Thus the unique solution (ϕ, ψ) satisfies

$$(\phi(t,T,u), \psi(t,T,u)) \in \mathbb{R}_- \times \mathbb{R}^m_- \times \{0\}^n.$$

In other words, $M(t) \leq 1$ for all $t \leq T$, and the following corollary is proven.

Corollary 2.14. *Let* **X** *be an affine process. For* $u \in \mathbb{R}^m_- \times \{0\}^n$ *there exists a unique solution to the Riccati equations* (2.21) *and the affine transformation formula* (2.2) *holds.*

For the general case of $u \in \mathbb{C}^d$, the author believes that the expectation $\mathrm{E}\left[e^{u^\top X(T)}\right]$ is finite if and only if there exists a solution to the Riccati equations. In [7] it is proved for the time-homogeneous case with a long technical proof. With time and courage, the proof could probably be extended to cover the time-inhomogeneous setup here. This is sated in the following conjecture.

Conjecture 2.15. *Let* **X** *be an affine process, and let* $u \in \mathbb{R}^d$. *For* $T \geq 0$,

$$\mathrm{E}\left[e^{u^\top X(T)}\right] < \infty,$$

if and only if there exists a solution to the Riccati equations on $[0,T]$. *In that case, the affine transformation formula* (2.2) *holds for* $t \in [0,T]$.

3 Affine Processes in Life Insurance

3.1 Survival Probabilities and Discount Factors

For applications in life insurance, one is interested in expressions of the form $\mathrm{E}\left[e^{-\int_t^n \mu(s)\,\mathrm{d}s}\middle|\,\mathcal{F}(t)\right]$, where μ is a stochastic process, e.g. a mortality rate. In general, let $\mathbf{X}$ be an affine d-dimensional process, and define

$$\mathrm{d}Y(t) = \big(c(t) + \gamma(t)^\top X(t)\big)\ \mathrm{d}t \tag{3.1}$$

for continuous functions $c : \mathbb{R}_+ \to \mathbb{R}$ and $\gamma : \mathbb{R}_+ \to \mathbb{R}^d$. We show that the extended process $(\mathbf{X}, \mathbf{Y})$ is an affine process, which allows us to consider expressions of the form

$$\mathrm{E}\left[e^{\int_0^T \left(c(s)+\gamma(s)^\top X(s)\right)\mathrm{d}s}\middle|\,\mathcal{F}(t)\right].$$

The corollary also addresses the special case $m = d$, i.e. the case where the affine process is non-negative, where it is known that the affine transformation formula holds. The corollary is inspired by Theorem 10.4 in [7], but the Corollary presented here does not rely on Conjecture 2.15, and hence the existence is only guaranteed for the case $m = d$.

Corollary 3.1. *Let* $\mathbf{X}$ *be an affine process satisfying* (2.1) *in the canonical state space* $\mathbb{R}_+^m \times \mathbb{R}^n$ *with* $m + n = d > 0$ *and* $m, n \geq 0$. *Let the process* $\mathbf{Y}$ *satisfy* (3.1). *Then the* $(d+1)$*-dimensional extended process* $(\mathbf{X}, \mathbf{Y})$ *is affine, with Riccati equations*

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(t,T,u) &= -\frac{1}{2}\psi_J(t,T,u)^\top a_{JJ}(t)\psi_J(t,T,u) - b(t)^\top\psi_{\{1,\ldots,d\}}(t,T,u) - c(t)u_{d+1},\\
\phi(T,T,u) &= 0,\\
\frac{\partial}{\partial t}\psi_i(t,T,u) &= -\frac{1}{2}\psi_{\{1,\ldots,d\}}(t,T,u)^\top\alpha_i(t)\psi_{\{1,\ldots,d\}}(t,T,u)\\
&\quad - \beta_i(t)^\top\psi_{\{1,\ldots,d\}}(t,T,u) - \gamma_i(t)u_{d+1},\\
\frac{\partial}{\partial t}\psi_J(t,T,u) &= -\mathcal{B}_{JJ}(t)^\top\psi_J(t,T,u) - \gamma_J(t)u_{d+1},\\
\psi_{d+1}(t,T,u) &= u_{d+1},\\
\psi(T,T,u) &= u,
\end{aligned} \tag{3.2}$$

for $i \in I$ *and* $u \in \mathbb{C}^{d+1}$*, where* $I = \{1, \dots, m\}$ *and* $J = \{m+1, \dots, d\}$*.*

Moreover, if $n = 0$ *and for all* $t \geq 0$*,*

$$\gamma(t) \in \mathbb{R}_+^m,$$

then

$$\mathrm{E}\left[e^{-\int_t^T \left(c(s) + \gamma(s)^\top X(s) \right) \mathrm{d}s} \,\middle|\, \mathcal{F}(t) \right] = e^{\Phi(t,T) + \Psi(t,T)^\top X(t)} \tag{3.3}$$

exists and (Φ, Ψ) *is the unique solution to the Riccati equations*

$$\begin{aligned} \frac{\partial}{\partial t}\Phi(t,T) &= -b(t)^\top \Psi(t,T) + c(t), \\ \frac{\partial}{\partial t}\Psi_i(t,T) &= -\frac{1}{2}\Psi(t,T)^\top \alpha_i(t)\Psi(t,T) - \beta_i(t)^\top \Psi(t,T) + \gamma_i(t), \\ \Phi(T,T) &= 0, \\ \Psi(T,T) &= 0, \end{aligned} \tag{3.4}$$

for $i \in I$*.*

Proof. The extended process $\widetilde{\mathbf{X}} = (\mathbf{X}, \mathbf{Y})$ is a diffusion process satisfying (2.1) with drift $\widetilde{b}(t,x) = \widetilde{b}(t) + \widetilde{\mathcal{B}}(t)x$ and diffusion $\widetilde{a}(t,x) = \widetilde{a}(t) + \sum_{i=1}^d \widetilde{\alpha}_i(t)x_i$, where

$$\widetilde{b}(t) = \begin{bmatrix} b(t) \\ c(t) \end{bmatrix}, \quad \widetilde{\mathcal{B}}(t) = \begin{bmatrix} \mathcal{B}(t) & 0 \\ \gamma(t)^\top & 0 \end{bmatrix},$$

and

$$\widetilde{a}(t) = \begin{bmatrix} a(t) & 0 \\ 0 & 0 \end{bmatrix}, \quad \widetilde{\alpha}_i(t) = \begin{bmatrix} \alpha_i(t) & 0 \\ 0 & 0 \end{bmatrix},$$

for $i \in I$. The process $\widetilde{\mathbf{X}}$ satisfies Assumption 2.1 since $c(t) + \gamma(t)^\top X(t)$ is continuous and therefore integrable. Then, by Theorem 2.12, the process is affine in $\mathbb{R}_+^m \times \mathbb{R}^n \times \mathbb{R}$. Inserting into the Riccati equations (2.21) one especially obtains $\frac{\mathrm{d}}{\mathrm{d}t}\psi_{d+1}(t,T,u) = 0$, and therefore $\psi_{d+1}(t,T,u) = u_{d+1}$. Then, the Riccati equations (3.2) are obtained by insertion into (2.21).

Assume now that $n = 0$ and $\gamma(t) \in \mathbb{R}_+^m$. Let $c(t) = 0$ for all $t \geq 0$, i.e. we consider the process $\mathrm{d}Y(t) = \gamma(t)^\top X(t)\,\mathrm{d}t$. By Theorem 2.12 the state space is $\mathbb{R}_+^{m+1}$, and with $u = (0, \dots, 0, -1)$, where u is $(m+1)$-dimensional, a unique solution to the Riccati equations is ensured. Inserting into the Riccati equations (3.2), one obtains

$$\mathrm{E}\left[e^{-Y(T)}\middle|\,\mathcal{F}(t)\right] = e^{\phi(t,T,u)+\psi_I(t,T,u)^\top X(t)-Y(t)}.$$

Multiplying the equation with $e^{-\int_t^T c(s)\,\mathrm{d}s+Y(t)}$ and defining

$$\Psi(t,T) = \psi_I(t,T,u), \quad \Phi(t,T) = \phi(t,T,u) - \int_t^T c(s)\,\mathrm{d}s$$

the results (3.3) and (3.4) are obtained. □

Corollary 3.1 and Theorem 2.12 are the main results so far, providing the basis for applications of affine processes in life insurance. Theorem 2.12 is used to determine if a process is affine, and given such a process, Corollary 3.1 provides the formula (3.3) with a system of Riccati equations that has a guaranteed solution.

3.2 Generalised Forward Rates

This section contains a generalisation and application of the concept of forward rates in life insurance mathematics. The author believes that this section contains new results not present in existing litterature.

A popular quantity in life insurance mathematics is

$$\mathrm{E}\left[e^{-\int_t^T X(s)\,\mathrm{d}s}X(T)\middle|\,\mathcal{F}(t)\right],$$

for a 1-dimensional process $\mathbf{X}$, e.g a mortality rate. Treating the quantity as a derivative, as in Lemma 1.12, it is seen that

$$\mathrm{E}\left[e^{-\int_t^T X(s)\,\mathrm{d}s}X(T)\middle|\,\mathcal{F}(t)\right] = e^{\Phi(t,T)+\Psi(t,T)X(t)}\left(-\frac{\partial}{\partial T}\Phi(t,T) - X(t)\frac{\partial}{\partial T}\Psi(t,T)\right).$$

In this section we consider a 2-dimensional process $\mathbf{X}$ and the quantity,

$$\mathrm{E}\left[e^{-\int_t^T (X_1(s)+X_2(s))\,\mathrm{d}s}X_1(T)\middle|\,\mathcal{F}(t)\right].$$

To this end, define the “indicator” function

$$\mathcal{I} : \mathbb{R}\times\mathbb{R}_{++}\times\mathbb{R}_+ \to \mathbb{R}_+$$

such that it is continuous and for $\varepsilon > 0$ satisfies

$$\mathcal{I}(t, S, \varepsilon) = \begin{cases} 1, & t \leq S - \varepsilon, \\ 0, & t > S. \end{cases}$$

Also, for $\varepsilon > 0$ we assume that $\mathcal{I}$ is $\mathcal{C}^1$ in S and ε and increasing in S and decreasing in ε. For $\varepsilon = 0$, let

$$\mathcal{I}(t, S, 0) = 1.$$

Then, in particular, $\mathcal{I}(t, S, \varepsilon) \nearrow 1_{\{t<S\}}$ for $\varepsilon \searrow 0$, and if $t < S$, we have $\mathcal{I}(t, S, \varepsilon) \nearrow \mathcal{I}(t, S, 0)$. It is assumed that $\mathcal{I}$ exists.

To give an intuition about the contents of this section, we begin with a heuristic presentation of the calculations. Let $t < S < T$,

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T (X_1(s) + \mathcal{I}(s,S,\varepsilon) X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t) \right] \\
&= \mathrm{E}\left[-\frac{\partial}{\partial T} e^{-\int_t^T X_1(s)\,\mathrm{d}s - \int_t^S \mathcal{I}(s,S,\varepsilon) X_2(s)\,\mathrm{d}s} \middle| \mathcal{F}(t) \right] \\
&= -\frac{\partial}{\partial T} \mathrm{E}\left[e^{-\int_t^T (X_1(s) + \mathcal{I}(s,S,\varepsilon) X_2(s))\,\mathrm{d}s} \middle| \mathcal{F}(t) \right] \\
&= -\frac{\partial}{\partial T} e^{\Phi(t,T,S,\varepsilon) + \Psi(t,T,S,\varepsilon)^\top X(t)} \\
&= e^{\Phi(t,T,S,\varepsilon) + \Psi(t,T,S,\varepsilon)^\top X(t)} \frac{\partial}{\partial T} \int_t^T \left(\frac{\partial}{\partial t} \Phi(s,T,S,\varepsilon) + X(t)^\top \frac{\partial}{\partial t} \Psi(s,T,S,\varepsilon) \right) \mathrm{d}s \\
&= e^{\Phi(t,T,S,\varepsilon) + \Psi(t,T,S,\varepsilon)^\top X(t)} \Bigg(\frac{\partial}{\partial t} \Phi(T,T,S,\varepsilon) + X(t)^\top \frac{\partial}{\partial t} \Psi(T,T,S,\varepsilon) \\
&\qquad + \int_t^T \left(\frac{\partial}{\partial T} \frac{\partial}{\partial t} \Phi(s,T,S,\varepsilon) + X(t)^\top \frac{\partial}{\partial T} \frac{\partial}{\partial t} \Psi(s,T,S,\varepsilon) \right) \mathrm{d}s \Bigg) \\
&= e^{\Phi(t,T,S,\varepsilon) + \Psi(t,T,S,\varepsilon)^\top X(t)} \left(W(t,T,S,\varepsilon) + V(t,T,S,\varepsilon) X(t)^\top \right).
\end{aligned}$$

Here, W and V solve a linear differential equation system, namely the system of differential equations that Φ and Ψ solve, but differentiated with respect to T. The boundary conditions are $W(T,T,S,\varepsilon) = \frac{\partial}{\partial t}\Phi(T,T,S,\varepsilon)$ and $V(T,T,S,\varepsilon) = \frac{\partial}{\partial t}\Psi(T,T,S,\varepsilon)$.

To reach the result presented in Theorem 3.4, take limits as $S \nearrow T$ and $\varepsilon \searrow 0$. The arguments are carried out in full detail below, split into two lemmas and a theorem.

Lemma 3.2. *Let* $\mathbf{X}$ *be a 2-dimensional non-negative affine process satisfying* (2.1), *and let* $t < T$. *If there exists an open interval* A *containing* T *and a stochastic bound*

$$Z \geq \sup_{t \in A} X_1(t),$$

such that $\mathrm{E}\left[Z \middle| \mathcal{F}(t)\right] < \infty$, *then*

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_t^T (X_1(s)+\mathcal{I}(s,\lambda)X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t)\right] \\ &= e^{\Phi(t,T,\lambda)+\Psi(t,T,\lambda)^\top X(t)} \left(W(t,T,\lambda) + V(t,T,\lambda)^\top X(t)\right), \end{aligned} \tag{3.5}$$

for $\lambda = (S, \varepsilon) \in (t, T) \times \mathbb{R}_{++}$. *The functions* Φ *and* Ψ *satisfy the system of differential equations*

$$\begin{aligned} \frac{\partial}{\partial t}\Psi_1(t,T,\lambda) &= -\frac{1}{2}\Psi(t,T,\lambda)^\top \alpha_1(t)\Psi(t,T,\lambda) - \beta_1(t)^\top \Psi(t,T,\lambda) + 1, \\ \frac{\partial}{\partial t}\Psi_2(t,T,\lambda) &= -\frac{1}{2}\Psi(t,T,\lambda)^\top \alpha_2(t)\Psi(t,T,\lambda) - \beta_2(t)^\top \Psi(t,T,\lambda) + \mathcal{I}(t,\lambda), \\ \frac{\partial}{\partial t}\Phi(t,T,\lambda) &= -b(t)^\top \Psi(t,T,\lambda), \\ \Phi(T,T,\lambda) &= 0, \\ \Psi(T,T,\lambda) &= 0, \end{aligned} \tag{3.6}$$

and W *and* V *satisfy the system of differential equations*

$$\begin{aligned} \frac{\partial}{\partial t}V_i(t,T,\lambda) &= -\Psi(t,T,\lambda)^\top \alpha_i(t) V(t,T,\lambda) - \beta_i(t)^\top V(t,T,\lambda), \\ V_1(T,T,\lambda) &= 1, \\ V_2(T,T,\lambda) &= 0, \\ \frac{\partial}{\partial t}W(t,T,\lambda) &= -b(t)^\top V(t,T,\lambda), \\ W(T,T,\lambda) &= 0. \end{aligned} \tag{3.7}$$

Proof. Let $\lambda = (S, \varepsilon) \in (t, T) \times \mathbb{R}_{++}$. Considering Corollary 3.1, let $\gamma(t) = (1, \mathcal{I}(t, S, \varepsilon))^\top$. Then (3.3) holds,

$$\mathrm{E}\left[e^{-\int_t^T (X_1(s)+\mathcal{I}(s,\lambda)X_2(s))\,\mathrm{d}s} \middle| \mathcal{F}(t)\right] = e^{\Phi(t,T,\lambda)+\Psi(t,T,\lambda)^\top X(t)},$$

where (Φ, Ψ) satisfies (3.4) with $c(t) = 0$, which is (3.6).

Differentiating both sides with respect to T, and multiplying with -1 yields,

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T (X_1(s)+\mathcal{I}(s,\lambda)X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t)\right] \\
&= e^{\Phi(t,T,\lambda)+\Psi(t,T,\lambda)^\top X(t)} \left(-\frac{\partial}{\partial T}\Phi(t,T,\lambda) - \frac{\partial}{\partial T}\Psi(t,T,\lambda)^\top X(t) \right),
\end{aligned}$$

where we from Theorem A.5 know that (Φ, Ψ) is differentiable in T. Here, we interchanged expectation and differentiation, and by the assumptions of the lemma, it holds for all $T \in A$.

From Theorem A.5 a system of differential equations for $t \mapsto \frac{\partial}{\partial T}(\Psi, \Phi)(t, T, \lambda)$ is also obtained. With $(x, y) = (\Psi, \Phi)$, let $R_i(t, (x, y), \lambda)$ for $i = 1, 2$ denote the right hand side of $\frac{\partial}{\partial t}\Psi_i(t, T, \lambda)$ in (3.6) and let $R_3(t, (x, y), \lambda)$ denote the right hand side of $\frac{\partial}{\partial t}\Phi(t, T, \lambda)$. Then, for $j = 1, 2$,

$$\begin{aligned}
\frac{\partial}{\partial x_j} R_i(t,(x,y),\lambda) &= -x^\top \alpha_i(t) e_j - \beta_i(t)^\top e_j, \\
\frac{\partial}{\partial y} R_i(t,(x,y),\lambda) &= 0, \\
\frac{\partial}{\partial x_j} R_3(t,(x,y),\lambda) &= -b_j(t), \\
\frac{\partial}{\partial y} R_3(t,(x,y),\lambda) &= 0,
\end{aligned}$$

and multiplying with a 3-dimensional vector $(v^\top, w) = (v_1, v_2, w)$ yields

$$\begin{aligned}
\frac{\partial}{\partial (x,y)} R_i(t,(x,y),\lambda)^\top (v_1, v_2, w)^\top &= -x^\top \alpha_i(t) v - \beta_i(t)^\top v, \\
\frac{\partial}{\partial (x,y)} R_3(t,(x,y),\lambda)^\top (v_1, v_2, w)^\top &= -b(t)^\top v.
\end{aligned}$$

This defines a system of differential equations for the function

$$t \mapsto \frac{\partial}{\partial T}(\Psi, \Phi)(t, T, \lambda),$$

with boundary condition

$$-R(T, (\Psi, \Phi)(T, T, \lambda), \lambda) = (-1, 0, 0)^\top.$$

Now, let

$$\begin{aligned}
W(t,T,\lambda) &= -\frac{\partial}{\partial T}\Phi(t,T,\lambda), \\
V(t,T,\lambda) &= -\frac{\partial}{\partial T}\Psi(t,T,\lambda),
\end{aligned}$$

then the system of differential equations (3.7) is obtained. □

It is immediate to take limits on the left hand side of (3.5). Also, the limit for $S \nearrow T$ can be taken on the right hand side.

Lemma 3.3. *Under the conditions of Lemma 3.2,*

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_t^T (X_1(s)+X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t) \right] \\ &= \lim_{\varepsilon \searrow 0} e^{\Phi(t,T,\varepsilon)+\Psi(t,T,\varepsilon)^\top X(t)} \left(W(t,T,\varepsilon) + V(t,T,\varepsilon)^\top X(t) \right), \end{aligned} \tag{3.8}$$

where Ψ, Φ, V *and* W *satisfy the system of differential equations* (3.6) *and* (3.7) *for* $\lambda = (T, \varepsilon)$.

Proof. First, we prove that

$$\begin{aligned} &\lim_{\varepsilon \searrow 0} \lim_{S \nearrow T} \mathrm{E}\left[e^{-\int_t^T (X_1(s)+\mathcal{I}(t,S,\varepsilon)X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t) \right] \\ &= \mathrm{E}\left[e^{-\int_t^T (X_1(s)+X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t) \right]. \end{aligned}$$

Since $e^{-\int_t^T (X_1(s)+\mathcal{I}(t,S,\varepsilon)X_2(s))\,\mathrm{d}s} X_1(T) \leq X_1(T)$, which is assumed to have finite expectation, dominated convergence can be applied to obtain

$$\begin{aligned} &\lim_{\varepsilon \searrow 0} \lim_{S \nearrow T} \mathrm{E}\left[e^{-\int_t^T (X_1(s)+\mathcal{I}(t,S,\varepsilon)X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t) \right] \\ &= \mathrm{E}\left[\lim_{\varepsilon \searrow 0} \lim_{S \nearrow T} e^{-\int_t^T (X_1(s)+\mathcal{I}(t,S,\varepsilon)X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t) \right]. \end{aligned}$$

The first part of the lemma follows by continuity of $x \mapsto e^{-x}$ and monotone convergence, which can be applied since $\mathcal{I}(t, S, \varepsilon)$ is increasing in S and decreasing in ε.

By Theorem A.4, the solution (Ψ, Φ, V, W) of the combined system of differential equations (3.6) and (3.7) is continuous in $(S, \varepsilon) \in \mathbb{R}_{++} \times \mathbb{R}_{++}$. In particular, it is continuous in (T, ε) for all $\varepsilon \in \mathbb{R}_{++}$, which is the second part. □

Note that, since the left hand side of (3.8) exists, the limit on the right hand side exists. It remains to be verified that the solutions are continuous in $\varepsilon = 0$ and that the limit functions solve the particular system of differential equations where $\varepsilon = 0$.

Theorem A.4 cannot be applied for the limit as $\varepsilon \searrow 0$, since $\varepsilon \mapsto \mathcal{I}(t, T, \varepsilon)$ is discontinuous in 0 when $t \geq T$ is considered. However, we have proved above that the limit exists, and the following theorem shows that the limiting functions solve the particular system of differential equations where $(S, \varepsilon) = (T, 0)$.

Theorem 3.4. *Let* $\mathbf{X}$ *be a 2-dimensional non-negative affine process satisfying* (2.1)*, and let* $t < T$*. If there exists an open interval* A *containing* T *and a stochastic bound*

$$Z \geq \sup_{t \in A} X_1(t),$$

such that $\mathrm{E}\left[Z \middle| \mathcal{F}(t)\right] < \infty$*, then*

$$\mathrm{E}\left[e^{-\int_t^T (X_1(s)+X_2(s))\,\mathrm{d}s} X_1(T) \middle| \mathcal{F}(t)\right] = e^{\Phi(t,T)+\Psi(t,T)X(t)} \left(W(t,T) + V(t,T)^\top X(t)\right),$$

where Φ, Ψ *solve the system of differential equations,*

$$\begin{aligned}
\frac{\partial}{\partial t}\Psi_i(t,T) &= -\frac{1}{2}\Psi(t,T)^\top \alpha_i(t)\Psi(t,T) - \beta_i(t)^\top \Psi(t,T) + 1,\\
\frac{\partial}{\partial t}\Phi(t,T) &= -b(t)^\top \Psi(t,T),\\
\Phi(T,T) &= 0,\\
\Psi(T,T) &= 0,
\end{aligned}$$

and V, W *solve the system of differential equations,*

$$\begin{aligned}
\frac{\partial}{\partial t}V_i(t,T) &= -\Psi(t,T)^\top \alpha_i(t)V(t,T) - \beta_i(t)^\top V(t,T),\\
V_1(T,T) &= 1,\\
V_2(T,T) &= 0,\\
\frac{\partial}{\partial t}W(t,T) &= -b(t)^\top V(t,T),\\
W(T,T) &= 0,
\end{aligned}$$

for $i = 1, 2$*.*

Proof. Consider first Ψ as defined by (3.6). By Lemma 3.3 we set $S = T$ and suppress the notation of S so that we instead of λ simply write ε.

In Corollary 3.1 with $\gamma(t) = (1, \mathcal{I}(t,T,\varepsilon))^\top$, it is stated that Ψ exists, in particular for all $t \in [0,T]$ and $\varepsilon \in [0,1]$. A uniform bound $K > 1$ then exists,

$$\sup_{(t,\varepsilon)\in[0,T]\times[0,1]} \|\Psi(t,T,\varepsilon)\|^2 \leq K.$$

Split the solution into two intervals at the point $T-\varepsilon$, and denote by $\Psi^{(1)}$ and $\Psi^{(2)}$ these two solutions,

$$\Psi(t,T,\varepsilon) = \begin{cases} \Psi^{(1)}(t,T-\varepsilon), & t \in [0,T-\varepsilon), \\ \Psi^{(2)}(t,T,\varepsilon), & t \in [T-\varepsilon,T]. \end{cases} \tag{3.9}$$

For $\Psi^{(1)}$ the dependence of ε is supressed, since (as we will see) the differential equation system only depends on ε through the boundary condition.

Let $R_i(t,x,\varepsilon)$ denote the right side of $\frac{\partial}{\partial t}\Psi_i$ in (3.6),

$$R_i(t,x,\varepsilon) = -\frac{1}{2}x^\top \alpha_i(t)x - \beta_i(t)^\top x + \mathcal{I}(t,\varepsilon 1_{\{i=2\}}),$$

where we have suppressed the T in $\mathcal{I}(t,\varepsilon) = \mathcal{I}(t,T,\varepsilon)$. Recall that $\mathcal{I}(t,0) = 1$. Now, $\Psi^{(1)}$ and $\Psi^{(2)}$ solve the systems of differential equations, respectively

$$\begin{aligned} \frac{\partial}{\partial t}\Psi^{(2)}(t,T,\varepsilon) &= R(t,\Psi^{(2)}(t,T,\varepsilon),\varepsilon), \\ \Psi^{(2)}(T,T,\varepsilon) &= 0, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial t}\Psi^{(1)}(t,T-\varepsilon) &= R(t,\Psi^{(1)}(t,T-\varepsilon),0), \\ \Psi^{(1)}(T-\varepsilon,T-\varepsilon) &= \Psi^{(2)}(T-\varepsilon,T,\varepsilon). \end{aligned}$$

Now see, for $t \in [T-\varepsilon,T]$, first using that $\Psi^{(2)}(T,T,\varepsilon) = \Psi^{(2)}(T,T,0)$ and next using the mean value theorem to find vectors $\xi = (\xi_1,\xi_2)^\top$ and $\zeta = (\zeta_1,\zeta_2)^\top$,

$$\begin{aligned} &\|\Psi(t,T,\varepsilon) - \Psi(t,T,0)\|_\infty \\ &\le \sum_{i=1}^{2} \left|\Psi_i^{(2)}(t,T,\varepsilon) - \Psi_i^{(2)}(t,T,0)\right| \\ &\le \sum_{i=1}^{2} \left(\left|\Psi_i^{(2)}(t,T,\varepsilon) - \Psi_i^{(2)}(T,T,\varepsilon)\right| + \left|\Psi_i^{(2)}(T,T,0) - \Psi_i^{(2)}(t,T,0)\right| \right) \\ &= \sum_{i=1}^{2} \left(\left|R_i(\xi_i,\Psi^{(2)}(\xi_i,T,\varepsilon),\varepsilon)\right|(T-t) + \left|R_i(\zeta_i,\Psi^{(2)}(\zeta_i,T,0),0)\right|(T-t) \right) \\ &\le (T-t)4C. \end{aligned}$$

We saw above that $\Psi^{(2)}$ is uniformly bounded by K, so the above holds for C satisfying

$$\sup_{(t,x,\varepsilon)\in[0,T]\times[-K,K]^2\times[0,1]} |R(t,x,\varepsilon)| \leq C.$$

Such a C exists since R is continuous on the compact set of the supremum. We already know by Theorem A.4 that Ψ (and therefore also $\Psi^{(2)}$) is locally Lipschitz continuous in ε for $\varepsilon > 0$, and by the inequalities above, we conclude that $\Psi^{(2)}$ is locally Lipschitz continuous also for $\varepsilon = 0$.

Consider now the system of differential equations for $\Psi^{(1)}$, and realise that $R(t,x,0)$ is continuous in t and $\mathcal{C}^1$ in x (and in particular, it is locally Lipschitz continuous). Also, note that the system of differential equations is parameter independent. For a $\delta > 0$, let $J = (-\delta, T+\delta)$, and let $R(t,x,0)$ be defined for $t \in J$ (possibly by an extension like $R((t \wedge T) \vee 0, x, 0)$). Then J is an open interval, and by Theorem A.4 the solution $\Psi^{(1)}$ is locally Lipschitz continuous in the boundary condition, $(T-\varepsilon, \Psi^{(2)}(T-\varepsilon, T, \varepsilon))$. Above, we saw that the boundary condition itself is $\mathcal{C}^1$ in ε, so considering $\Psi^{(1)}$ as a function of the boundary condition, we conclude that $\Psi^{(1)}$ is locally Lipschitz continuous in ε for $\varepsilon \in [0,1]$. In total, $\Psi(t,T,\varepsilon)$ is locally Lipschitz continuous in $\varepsilon \in [0,1]$, by (3.9). In particular, for $t \in [0,T)$, choose $\varepsilon \leq 1$ such that $t < T - \varepsilon$. Then,

$$\Psi(t,T,\varepsilon) = \Psi^{(1)}(t, T-\varepsilon) \to \Psi(t,T,0), \qquad \text{for } \varepsilon \searrow 0.$$

For the functions $\Phi(t,T,\varepsilon)$, $V(t,T,\varepsilon)$ and $W(t,T,\varepsilon)$, see in (3.6) and (3.7) that they depend on ε only through $\Psi(t,T,\varepsilon)$: Denote by $Q = (Q^{\Phi}, Q^{V_1}, Q^{V_2}, Q^{W})^{\top}$ the right hand side of the system of differential equations for (Φ, V, W). Then $Q(t,x,\varepsilon)$ depend on ε only through Ψ. Thus, Q is locally Lipschitz continuous in (x,ε).

To apply Theorem A.4, ε must belong to an open interval, containing 0. Extending $\mathcal{I}$ to $\varepsilon < 0$ is trivial however; Set $\mathcal{I}(t,\varepsilon) = \mathcal{I}(t,|\varepsilon|)$, and apply Theorem A.4 for $\varepsilon \in (-1,1)$: We conclude that $(\Phi, V, W)(t,T,\varepsilon)$ is continuous in $\varepsilon = 0$. □

Remark 3.5. None of the calculations above are critical to the fact that we are considering a 2-dimensional process, and it is also straightforward to replace $X_i(t)$ by $c_i(t) + \gamma_i(t) X_i(t)$ for deterministic functions c_i and $\gamma_i \geq 0$. With a d-dimensional non-negative affine process $\mathbf{X}$ and $N \subset \{1, \dots, d\}$, for $d \geq 1$, one can extend the result to expressions of the form

$$\mathrm{E}\left[e^{-\int_t^T \sum_{i=1}^d (c_i(s)+\gamma_i(s)X_i(s))\,\mathrm{d}s} \sum_{j\in N} (c_j(T) + \gamma_j(T) X_j(T)) \middle| \mathcal{F}(t) \right].$$

Then the system of differential equations would read,

$$\begin{aligned}
\frac{\partial}{\partial t}\Psi_i(t,T) &= -\frac{1}{2}\Psi(t,T)^\top\alpha_i(t)\Psi(t,T) - \beta_i(t)^\top\Psi(t,T) + \gamma_i(t),\\
\frac{\partial}{\partial t}\Phi(t,T) &= -b(t)^\top\Psi(t,T) + \sum_{i=1}^{d} c_i(t),\\
\Phi(T,T) &= 0,\\
\Psi(T,T) &= 0,\\
\frac{\partial}{\partial t}V_i(t,T) &= -\Psi(t,T)^\top\alpha_i(t)V(t,T) - \beta_i(t)^\top V(t,T)\\
V_i(T,T) &= \gamma_i(T)1_N(i),\\
\frac{\partial}{\partial t}W(t,T) &= -b(t)^\top V(t,T),\\
W(T,T) &= \sum_{i\in N} c_i(T),
\end{aligned}$$

for $i = 1,\dots,d$. $\diamond$

Although the proof of the remark is not written down, we will assume that it holds for the rest of the thesis.

The result is applicable when one wants to model dependent interest and transition rates. An example is a market value reserve for a life insurance, where, when modelling with a surrender option, the market value is of the form

$$\begin{aligned}
&\mathrm{E}\left[\int_t^n e^{-\int_t^s (r(\tau)+\mu(\tau)+\nu(\tau))\,\mathrm{d}\tau}\left(\mu(s)S + \nu(s)G(s)\right)\,\mathrm{d}s\,\middle|\,\mathcal{F}(t)\right]\\
&= \int_t^n \Bigg(\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu(\tau)+\nu(\tau))\,\mathrm{d}\tau}\mu(s)\,\middle|\,\mathcal{F}(t)\right] S\\
&\qquad + \mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu(\tau)+\nu(\tau))\,\mathrm{d}\tau}\nu(s)\,\middle|\,\mathcal{F}(t)\right] G(s)\Bigg)\,\mathrm{d}s.
\end{aligned} \tag{3.10}$$

Here, n is the end of the contract period, μ is the mortality rate, ν is the surrender rate, S is the sum paid upon death and $G(s)$ is the amount paid upon surrender at time s. The result allows one to model r, μ and ν as dependent affine processes.

It is often convenient to express the solution of the differential equations in the form of forward rates and for this purpose some new notation is introduced. The forward rates are labeled by a topscript, such that

$$\mathrm{E}\left[e^{-\int_t^T A(s)\,\mathrm{d}s}B(T)\,\middle|\,\mathcal{F}(t)\right] = e^{-\int_t^T f_t^{A:A}(s)\,\mathrm{d}s} f_t^{B:A}(T).$$

This is made precise in the following, which is an extension of Definition 1.11, and these forward rates are termed generalised forward rates.

Definition 3.6. *Let*

$$Y_i(t) = c_i(t) + \gamma_i(t)X_i(t)$$

be interest and/or transition rates. For $N \subset \{1,\dots,d\}$, the generalised forward rates *are defined by*

$$f_t^{\left(\sum_{j\in N} Y_j\right):\left(\sum_{i=1}^d Y_i\right)}(T) = \begin{cases} W(t,T) + V(t,T)^\top X(t), & t \le T, \\ \sum_{j\in N} Y_j(T), & t > T, \end{cases}$$

where V and W are defined in Remark 3.5. If $N = \{1,\dots,d\}$, we suppress the notation after ":" and write

$$f_t^{\sum_{i=1}^d Y_i}(T) = f_t^{\left(\sum_{j\in N} Y_j\right):\left(\sum_{i=1}^d Y_i\right)}(T).$$

Note that the generalised forward rates are only defined for stochastic interest and transition rates that are *affine*. The definition makes sense, since for $N = \{1,\dots,d\}$,

$$\Phi(t,T) + \Psi(t,T)^\top X(t) = -\int_t^T \left(W(t,s) + V(t,s)^\top X(t)\right) \,\mathrm{d}s,$$

and hence

$$e^{\Phi(t,T)+\Psi(t,T)^\top X(t)} = e^{-\int_t^T f_t^{\sum_{i=1}^d Y_i}(s)\,\mathrm{d}s}.$$

With the notation we can consider (3.10). With e.g. $Y_1 = r$, $Y_2 = \mu$ and $Y_3 = \nu$, it is written as

$$\int_t^n e^{-\int_t^s f_t^{(r+\mu+\nu)}(\tau)\,\mathrm{d}\tau} \left(f_t^{\mu:(r+\mu+\nu)}(s)S + f_t^{\nu:(r+\mu+\nu)}(s)G(s)\right) \,\mathrm{d}s.$$

We make two observations about the new notation. First, an interesting result is obtained if one lets $S = G(s)$ above.

Lemma 3.7. *The forward rates satisfy, for $N \subset \{1,\dots,d\}$ and $k \notin N$,*

$$f_t^{\left(\sum_{j\in N} Y_j\right):\left(\sum_{i=1}^d Y_i\right)}(T) + f_t^{Y_k:\left(\sum_{i=1}^d Y_i\right)}(T) = f_t^{\left(Y_k+\sum_{j\in N} Y_j\right):\left(\sum_{i=1}^d Y_i\right)}(T).$$

Proof. Note that

$$
\begin{aligned}
&e^{-\int_t^T f_t^{\sum_{i=1}^d Y_i}(s)\,\mathrm{d}s} f_t^{\left(Y_k+\sum_{j\in N} Y_j\right):\left(\sum_{i=1}^d Y_i\right)}(T) \\
&= \mathrm{E}\left[e^{-\int_t^T \sum_{i=1}^d Y_i(s)\,\mathrm{d}s} \left(Y_k(T) + \sum_{j\in N} Y_j(T) \right) \middle| \mathcal{F}(t) \right] \\
&= \mathrm{E}\left[e^{-\int_t^T \sum_{i=1}^d Y_i(s)\,\mathrm{d}s} Y_k(T) \middle| \mathcal{F}(t) \right] + \mathrm{E}\left[e^{-\int_t^T \sum_{i=1}^d Y_i(s)\,\mathrm{d}s} \sum_{j\in N} Y_j(T) \middle| \mathcal{F}(t) \right] \\
&= e^{-\int_t^T f_t^{\sum_{i=1}^d Y_i}(s)\,\mathrm{d}s} f_t^{Y_k:\left(\sum_{i=1}^d Y_i\right)}(T) + e^{-\int_t^T f_t^{\sum_{i=1}^d Y_i}(s)\,\mathrm{d}s} f_t^{\left(\sum_{j\in N} Y_j\right):\left(\sum_{i=1}^d Y_i\right)}(T),
\end{aligned}
$$

which yields the result when multiplied with $e^{\int_t^T f_t^{\sum_{i=1}^d Y_i}(s)\,\mathrm{d}s}$. □

Second, one can ask if the second part of the notation, after the ":", is necessary. In other terms, are the forward rates for μ in

$$
\mathrm{E}\left[e^{-\int_t^T (r(s)+\mu(s))\,\mathrm{d}s} \mu(T) \middle| \mathcal{F}(t) \right]
$$

and

$$
\mathrm{E}\left[e^{-\int_t^T \mu(s)\,\mathrm{d}s} \mu(T) \middle| \mathcal{F}(t) \right]
$$

identical, i.e. does it hold that

$$
f_t^{\mu:(r+\mu)}(T) \stackrel{?}{=} f_t^{\mu:\mu}(T). \tag{3.11}
$$

The general answer is **no**. It holds if r and μ are independent, but considering the differential equations describing the two forward rates, one can realise that they are not equal in general.

3.3 Comparison With Current Forward Rate Definitions

The *forward rate* was originally used in financial mathematics, and only for the interest rate. Later the term was proposed for use in actuarial mathematics for the mortality rate and, to the authors knowledge, it was first seen in Milevsky and Promislow (2001) [15]. The interest and mortality rate do, for simple life insurance contracts, play the same role, and the adoption of the term seems natural.

Most of the definitions and uses of the forward mortality rate have been for independent interest and mortality rates, however Miltersen and Persson (2005) [16] attempt to define the forward mortality rate for dependent interest and mortality rates. Norberg (2010) [18] examines the concept of forward mortality rates and criticises the concept for its apparent lack of desirable properties and the difficulty of generalising the forward rates to multi-state life insurance models. It is the belief that the generalised forward rates proposed in this thesis solve some of the problems pointed out in [18] for dependent interest and mortality rates.

The approach in [16] is examined in the following section in light of the generalised forward rates presented in this thesis. Since the generalised forward rates are only defined for affine processes, we work under the assumption that interest and mortality rates are affine processes.

3.3.1 Forward Rates Proposed in the Literature

Let $r(t)$ and $\mu(t)$ be the stochastic interest and mortality rate, respectively. The approach in [16] is as follows. First, the *forward interest rate*, g_t, is defined as usual, thus satisfying

$$\mathrm{E}\left[e^{-\int_t^T r(s)\,\mathrm{d}s}\,\middle|\,\mathcal{F}(t)\right] = e^{-\int_t^T g_t(s)\,\mathrm{d}s}. \tag{3.12}$$

Second, the *forward mortality spread rate*, h_t, is defined as the function satisfying

$$\mathrm{E}\left[e^{-\int_t^T (r(s)+\mu(s))\,\mathrm{d}s}\,\middle|\,\mathcal{F}(t)\right] = e^{-\int_t^T (g_t(s)+h_t(s))\,\mathrm{d}s}. \tag{3.13}$$

Third, the *forward mortality rate for the term insurance*, h_t^{ti}, is defined as the function satisfying

$$\mathrm{E}\left[e^{-\int_t^T (r(s)+\mu(s))\,\mathrm{d}s}\mu(T)\,\middle|\,\mathcal{F}(t)\right] = e^{-\int_t^T \left(g_t(s)+h_t^{\mathrm{ti}}(s)\right)\mathrm{d}s}h_t^{\mathrm{ti}}(T). \tag{3.14}$$

The forward rates are named g and h instead of f to avoid confusion with the generalised forward rates defined above. The term *spread rate* is motivated by the forward credit default spread rates from credit risk, where they play a similar role in connection with the forward interest rate. Simple calculations yield the following expression for the forward mortality spread rate,

$$h_t(T) = \frac{-\frac{\partial}{\partial T}\,\mathrm{E}\left[e^{-\int_t^T (r(s)+\mu(s))\,\mathrm{d}s}\,\middle|\,\mathcal{F}(t)\right]}{\mathrm{E}\left[e^{-\int_t^T (r(s)+\mu(s))\,\mathrm{d}s}\,\middle|\,\mathcal{F}(t)\right]} - g_t(T), \tag{3.15}$$

assuming that differentiation and integration may be interchanged. The expression for h_t is valuable: it tells us that h_t is well-defined. With the notation from Section 3.2, it is immediate that $g_t(s) = f_t^{r:r}(s)$. From the definition (3.13) we obtain

$$e^{-\int_t^T (g_t(s)+h_t(s))\,\mathrm{d}s} = \mathrm{E}\left[e^{-\int_t^T (r(s)+\mu(s))\,\mathrm{d}s}\middle|\,\mathcal{F}(t)\right] = e^{-\int_t^T \left(f_t^{r:(r+\mu)}(s)+f_t^{\mu:(r+\mu)}(s)\right)\mathrm{d}s},$$

thus

$$h_t(s) = f_t^{r:(r+\mu)}(s) - g_t(s) + f_t^{\mu:(r+\mu)}(s), \tag{3.16}$$

which is identical to (3.15). From the discussion about (3.11) we know that the equality $f_t^{r:(r+\mu)}(T) = f_t^{r:r}(T)$ does **not** hold in general. We are interested in the general case, thus we assume they are unequal. Hence, using that $f_t^{r:r}(T) = g_t(T)$, we obtain

$$h_t(T) \neq f_t^{\mu:(r+\mu)}(T),$$

which is also (3.11). The generalised forward rates thus differ from the already proposed ones.

In [16], there is a discussion of the two mortality rates and the relation between them. It is immediate to see, that if the interest and mortality rate are independent, the two mortality rates are equal. However, when they are dependent, it is not straightforward. Norberg (2010) [18] discusses the forward rate concept, and examines in Section 4 the difference between the forward mortality spread rate and the forward mortality rate for term insurance proposed in [16]. By a counterexample, he concludes that they are not equal in general.

Before such a study can be carried out though, it is of interest to consider the definitions first. Are they well defined? For the forward mortality spread rate h_t from (3.13), we found that it is well defined by (3.15), but such an expression is not immediate for the forward mortality rate for the term insurance, h_t^{ti}. As we will see, the answer to the question raised is in fact no; the forward rate for the term insurance is not well-defined.

From the equation (3.14) we obtain

$$\begin{aligned} e^{-\int_t^T \left(g_t(s)+h_t^{\mathrm{ti}}(s)\right)\mathrm{d}s} h_t^{\mathrm{ti}}(T) &= \mathrm{E}\left[e^{-\int_t^T (r(s)+\mu(s))\,\mathrm{d}s}\mu(T)\middle|\,\mathcal{F}(t)\right] \\ &= e^{-\int_t^T \left(f_t^{r:(r+\mu)}(s)+f_t^{\mu:(r+\mu)}(s)\right)\mathrm{d}s} f_t^{\mu:(r+\mu)}(T), \end{aligned}$$

For the expressions to be natural, one must expect that the exponentiated part, and the non-exponentiated part respectively, are identical on both side of the equation, i.e.

$$-\int_t^T \left(g_t(s) + h_t^{\mathrm{ti}}(s)\right) \mathrm{d}s = -\int_t^T \left(f_t^{r:(r+\mu)}(s) + f_t^{\mu:(r+\mu)}(s)\right) \mathrm{d}s,$$

and

$$h_t^{\mathrm{ti}}(T) = f_t^{\mu:(r+\mu)}(T).$$

If these two equalities hold, not only at the timepoints stated, but for the functions in general (i.e. $h_t^{\mathrm{ti}} = f_t^{\mu:(r+\mu)}$), there is a contradiction, since $g_t \neq f_t^{r:(r+\mu)}$. The conclusion is that h_t^{ti} is not well-defined: *In general it does not exist.*

This indicates that there is something wrong. The generalised forward rates proposed in this thesis do not have the same problem, and a forward mortality rate $f_t^{\mu:(r+\mu)}(s)$ exists, which can be applied for both the term insurance and the pure endowment.

The problem with the proposed forward mortality rates from [16] is that they insist on not changing the forward interest rate. In the setup proposed here, it is accepted that $f_t^{r:(r+\mu)} \neq f_t^{r:r}$, where the latter is the usual definition. This can both be seen as natural, since if the equality sign holds, one could insist that it should also hold for the respective forward mortality rate. In that case, the formula from the independent case holds, and there is a contradiction. On the other hand, it can be seen as troubling that the definition of the forward interest rate has changed, as the intuitive interpretation is not that clear. By applications of the generalised forward rates, it should come to light if the μ-dependent forward interest rate, $f_t^{r:(r+\mu)}$, has its own meaningful and intuitive interpretation.

4 Examples

4.1 Stochastic Mortality

We consider a simple example with a stochastic mortality rate, where the stochastic part is driven by a CIR process. This is inspired by the setup considered in [6], where the mortality is of the form $\mu_x^\circ(t)X_x(t)$ for a time-inhomogeneous CIR process $\mathbf{X}_x$, where the parameters depend on the initial age x. There, $\mu_x^\circ(t)$ is the mortality without expected mortality improvements, and $\mathbf{X}_x$ describes the expected mortality improvements for an x-year old.

Note that the x in the notation $\mathbf{X}_x$ is the age, and not the starting value of the process, $X(0) = \tilde{x}$, that was denoted $\mathbf{X}^{\tilde{x}}$ in Section 2.

Example 4.1. Consider the 1-dimensional CIR process $\mathbf{X}$ satisfying the SDE,

$$\mathrm{d}X(t) = (b + \beta X(t))\, \mathrm{d}t + \sigma\sqrt{X(t)}\, \mathrm{d}W(t), \tag{4.1}$$

for parameters $b, \sigma \in \mathbb{R}_{++}$ and $\beta \in \mathbb{R} \setminus \{0\}$. It is immediately seen from Theorem 2.12 that $\mathbf{X}$ is affine, with state space $\mathcal{X} = \mathbb{R}_+$.

Before defining the mortality rate, we show that $X(T) \mid \mathcal{F}(t)$ has a rescaled noncentral χ^2-distribution. To be specific, we show that, conditioning on $\mathcal{F}(t)$, the stochastic variable $\frac{X(T)}{C(T-t)}$, with $C(\tau) = \frac{\sigma^2}{4\beta}\left(e^{\beta\tau} - 1\right)$, has a noncentral χ^2-distribution with $\delta = \frac{4b}{\sigma^2}$ degrees of freedom and noncentrality parameter $\zeta = \frac{e^{\beta(T-t)}X(t)}{C(T-t)}$.

We find the conditional characteristic function

$$\mathrm{E}\left[e^{u\frac{X(T)}{C(T-t)}} \middle| \mathcal{F}(t)\right] = \mathrm{E}\left[e^{v_{T-t}X(T)} \middle| \mathcal{F}(t)\right],$$

with $v_\tau = \frac{u}{C(\tau)}$ and $u \in \mathrm{i}\mathbb{R}$. For general $v \in \mathrm{i}\mathbb{R}$, the Riccati equations (2.21) become,

$$\begin{aligned}
\frac{\partial}{\partial t}\psi(t,T,v) &= -\frac{1}{2}\sigma^2\psi(t,T,v)^2 - \beta\psi(t,T,v), \\
\frac{\partial}{\partial t}\phi(t,T,v) &= -b\psi(t,T,v),
\end{aligned}$$

with $\psi(T,T,v) = v$ and $\phi(T,T,v) = 0$. An application of Lemma A.6 yields analytic solutions, where $\theta = \sqrt{\beta^2} = |\beta|$,

$$\begin{aligned}
\psi(t,T,v) &= \frac{\left(|\beta|\left(e^{|\beta|(T-t)}+1\right)+\beta\left(e^{|\beta|(T-t)}-1\right)\right)v}{|\beta|\left(e^{|\beta|(T-t)}+1\right)-\beta\left(e^{|\beta|(T-t)}-1\right)-\sigma^2\left(e^{|\beta|(T-t)}-1\right)v} \\
&= \begin{cases} \frac{2\beta e^{\beta(T-t)}v}{2\beta-\sigma^2 v\left(e^{\beta(T-t)}-1\right)}, & \beta > 0, \\ \frac{-2\beta v}{-2\beta e^{-\beta(T-t)}-\sigma^2 v\left(e^{-\beta(T-t)}-1\right)}, & \beta < 0, \end{cases} \\
&= \frac{2\beta e^{\beta(T-t)}v}{2\beta-\sigma^2 v\left(e^{\beta(T-t)}-1\right)}, \\
\phi(t,T,v) &= \frac{2b}{\sigma^2}\log\left(\frac{2\,|\beta|\,e^{\frac{1}{2}(|\beta|-\beta)(T-t)}}{|\beta|\left(e^{|\beta|(T-t)}+1\right)-\beta\left(e^{|\beta|(T-t)}-1\right)-\sigma^2\left(e^{|\beta|(T-t)}-1\right)v}\right) \\
&= \begin{cases} \frac{2b}{\sigma^2}\log\left(\frac{2\beta}{2\beta-\sigma^2 v\left(e^{\beta(T-t)}-1\right)}\right), & \beta > 0, \\ \frac{2b}{\sigma^2}\log\left(\frac{-2\beta e^{-\beta(T-t)}}{-2\beta e^{-\beta(T-t)}-\sigma^2 v\left(e^{-\beta(T-t)}-1\right)}\right), & \beta < 0. \end{cases} \\
&= \frac{2b}{\sigma^2}\log\left(\frac{2\beta}{2\beta-\sigma^2 v\left(e^{\beta(T-t)}-1\right)}\right).
\end{aligned}$$

Inserting $v = v_{T-t}$ and performing straightforward calculations, we obtain,

$$\begin{aligned}
\psi(t,T,v_{T-t}) &= \frac{e^{\beta(T-t)}u}{C(T-t)(1-2u)}, \\
\phi(t,T,v_{T-t}) &= \log\left[\frac{1}{(1-2u)^{\frac{2b}{\sigma^2}}}\right],
\end{aligned}$$

and the conditional characteristic function becomes,

$$\mathrm{E}\left[e^{u\frac{X(T)}{C(T-t)}} \mid \mathcal{F}(t)\right] = \frac{e^{\frac{e^{\beta(T-t)}}{C(T-t)(1-2u)}X(t)}}{(1-2u)^{\frac{2b}{\sigma^2}}} = \frac{e^{\frac{\zeta u}{1-2u}}}{(1-2u)^{\frac{\delta}{2}}}.$$

By Lemma A.8, the distributional result is shown. Especially, $\frac{X(t)}{C(t)}$ itself has a noncentral χ^2-distribution, with δ degrees of freedom and noncentrality parameter $\zeta_t = \frac{e^{\beta t}}{C(t)}X(0)$.

Let $c_x : \mathbb{R}_+ \to \mathbb{R}$ and $\gamma_x : \mathbb{R}_+ \to \mathbb{R}_+$ be continuous functions, and define the mortality rate

$$\mu_x(t) = c_x(t) + \gamma_x(t)X_x(t),$$

for an x-year old at time 0. By the notation X_x we underline that the parameters of $\mathbf{X}_x$ are allowed to depend on the age at time 0, x, thus we write

$$\mathrm{d}X_x(t) = (b_x + \beta_x X_x(t))\,\mathrm{d}t + \sigma_x \sqrt{X_x(t)}\,\mathrm{d}W(t).$$

An application of Corollary 3.1 yields the conditional survival probabilities,

$$_{n-t}p^t_{x+t} = \mathrm{E}\left[e^{-\int_t^n \mu_x(s)\,\mathrm{d}s} \middle| \mathcal{F}(t) \right] = e^{\Phi_x(t,n)+\Psi_x(t,n)X_x(t)},$$

where Ψ_x and Φ_x solve the Riccati equations (3.4),

$$\begin{aligned}
\frac{\partial}{\partial t}\Psi_x(t,n) &= -\frac{1}{2}\sigma_x^2 \Psi_x(t,n)^2 - \beta_x \Psi_x(t,n) + \gamma_x(t),\\
\frac{\partial}{\partial t}\Phi_x(t,n) &= -b_x \Psi_x(t,n) + c_x(t),\\
\Phi_x(n,n) &= \Psi_x(n,n) = 0.
\end{aligned}$$

If γ_x is constant in the time t, then by Lemma A.6, there exists an analytic solution Ψ_x,

$$\Psi_x(t,n) = \frac{-2\gamma_x \left(e^{\theta_x(n-t)} - 1 \right)}{(\theta_x - \beta_x)\left(e^{\theta_x(n-t)} - 1 \right) + 2\theta_x},$$

with $\theta_x = \sqrt{\beta_x^2 + 2\gamma_x \sigma_x^2}$, and in that case,

$$\Phi_x(t,n) = \frac{2b_x}{\sigma_x^2} \log\left(\frac{2\theta_x e^{\frac{1}{2}(\theta_x - \beta_x)(n-t)}}{(\theta_x - \beta_x)\left(e^{\theta_x(n-t)} - 1 \right) + 2\theta_x} \right) - \int_t^n c_x(s)\,\mathrm{d}s.$$

The above model is advantageous in the sense that the distribution of $_{n-t}p^t_{x+t}$ is a (more or less analytic, depending on the functions c_x and γ_x) transformation of a noncentral χ^2-distribution, from which there is easy access to both quantiles and simulation. If γ_x isn't constant, or if it is and c_x isn't analytically integrable, then the functions Ψ_x and Φ_x can be solved numerically first, and then there is access to easy simulation of, and quantiles in, the distribution of $_{n-t}p^t_{x+t}$.

When considering term insurance contracts, the survival probability itself is not sufficient for valuation of the contract. There, one also needs to consider, using the setup in Section 1.1,

$$\begin{aligned}
\mathrm{E}\left[\int_t^n \mathrm{d}N_{01}(s) \middle| I(t) = 1, \mathcal{F}^Y(t) \right] &= \mathrm{E}\left[\int_t^n I(s)\mu_x(s)\,\mathrm{d}s \middle| I(t) = 1, \mathcal{F}^Y(t) \right]\\
&= \int_t^n \mathrm{E}\left[e^{-\int_t^s \mu_x(\tau)\,\mathrm{d}\tau} \mu_x(s) \middle| \mathcal{F}^Y(t) \right] \mathrm{d}s,
\end{aligned}$$

where $\mathbf{N}_{01}$ is the counting process counting the number of deaths, and $\mathbf{I}\mu_x$ is the intensity of the predictable compensator. To this end, the forward mortality rates are useful. We again consider the general case where $\gamma_x(t)$ is a function of t. The differential equations describing the forward rate, which satisfies

$$e^{-\int_t^n f^{\mu}_{x,t}(s)\,\mathrm{d}s} = e^{\Phi_x(t,n)+\Psi_x(t,n)X_x(t)},$$

can be found. The forward rate is labeled by the age at time 0, x, as in Section 1.3. By Remark 3.5, in the simple case $d = 1$ and $N = \{1\}$, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t}V_x(t,n) &= -\left(\beta_x + \sigma_x^2\Psi_x(t,n)\right)V_x(t,n),\\
V_x(n,n) &= \gamma(n),\\
\frac{\partial}{\partial t}W_x(t,n) &= -b_xV_x(t,n),\\
W_x(n,n) &= c_x(n),
\end{aligned}$$

where $f^{\mu}_{x,t}(n) = W_x(t,n) + V_x(t,n)X(t)$. The linear differential equation for V_x can be solved,

$$\begin{aligned}
V_x(t,n) &= \gamma(n)e^{\int_t^n\left(b_x+\sigma_x^2\Psi_x(s,n)\right)\mathrm{d}s},\\
W_x(t,n) &= c_x(n) + b_x\int_t^n V_x(s,n)\,\mathrm{d}s,
\end{aligned}$$

thus

$$f^{\mu}_{x,t}(n) = c_x(n) + b_x\int_s^n \gamma(n)e^{\int_t^n\left(b_x+\sigma_x^2\Psi_x(\tau,n)\right)\mathrm{d}\tau}\,\mathrm{d}s + \gamma(n)e^{\int_t^n\left(b_x+\sigma_x^2\Psi_x(s,n)\right)\mathrm{d}s}X(t).$$

With the forward mortality rate, the survival probability can be rewritten as

$${}_{n-t}p^t_{x+t} = e^{-\int_t^n f^{\mu}_{x,t}(s)\,\mathrm{d}s} = e^{-\int_t^n W_x(s,n)\,\mathrm{d}s - X(t)\int_t^n V_x(s,n)\,\mathrm{d}s}.$$

Now, the essential quantity, for dealing with term insurances, has the expression

$$\mathrm{E}\left[e^{-\int_t^n \mu_x(s)\,\mathrm{d}s}\mu_x(n)\middle|\,\mathcal{F}^Y(t)\right] = e^{-\int_t^n f^{\mu}_{x,t}(s)\,\mathrm{d}s}f^{\mu}_{x,t}(n),$$

obtained by Remark 3.5, with $d = 1$ and $N = \{1\}$ (or alternatively, Lemma 1.12).

The stochastic part of the forward rate $f^{\mu}_{x,t}(s)$ is $X(t)$, thus independent of s. As with the survival probability, once the functions W_x and V_x are found, there is easy access to simulation of, and quantiles in, the distribution of

$$\mathrm{E}\left[\int_t^n \mathrm{d}N_{01}(s)\middle|\,I(t) = 1, \mathcal{F}^Y(t)\right].$$

○

For the example to be applied, the parameters must be chosen, and there are different sensible choices. One possible choice of parameters is to have $c_x(t) = 0$ and let $\gamma_x(t) = \mu_x^\circ(t)$ be the basic mortality intensity *including* expected mortality improvements. The process $\mathbf{X}_x$ can be the deviations from the expected future mortality and one could choose $X_x(0) = 1$ and let $b_x = -\beta_x$, such that the process is mean reverting to 1, possibly with a slow mean reversion rate.

Another possible choice of parameters is to let $c_x(t) = 0$, and $\gamma_x(t) = \mu_x^\circ(t)$ be the basic mortality intensity *excluding* expected mortality improvements. The process $\mathbf{X}_x$ then plays the part of the longevity effect, and should start in 1 and converge slowly towards zero, or a positive value close to 0. This could be obtained by setting $\frac{b_x}{-\beta_x} = \varepsilon$ for a small ε, and find β_x such that $\mathbf{X}_x$ tends toward $\frac{b_x}{-\beta_x}$ at a suitable rate.

The system of differential equations for the forward rate $f^\mu_{x,t}(T)$ allows one to find the solution for all $t \leq T$, where T is fixed. In practice, when applied to a term insurance or life annuity, the solution is needed for fixed t and all $T \geq t$. Thus, the system of differential equations must be solved for each value of T where the solution is needed, which is not ideal in practice.

4.2 The Distribution of Particular Time-Inhomogeneous CIR Processes

In Example 4.1 it was seen that the CIR process has a noncentral χ^2-distribution. The result can be generalized to time-inhomogeneous CIR processes,

$$\mathrm{d}X(t) = (b(t) + \beta(t)X(t))\,\mathrm{d}t + \sigma(t)\sqrt{X(t)}\,\mathrm{d}W(t), \tag{4.2}$$

where

$$\delta(t) = \frac{4b(t)}{\sigma^2(t)}$$

is constant in t. Maghsoodi (1996) [14] proved this when $\delta(t) \in \mathbb{N}$ by representing the process as the squared norm of sums of Ornstein-Uhlenbeck processes, but it actually holds for all constant $\delta(t) \in \mathbb{R}_+$.

Theorem 4.2. *Let* $\mathbf{X}$ *be a time-inhomogeneous CIR process satisfying* (4.2) *where* b, β *and* σ *are real-valued and continuous functions with* $b, \sigma > 0$*. Assume*

$$\delta(t) = \frac{4b(t)}{\sigma^2(t)}$$

is constant. Then $\frac{X(T)}{\mathcal{C}(t,T)} \mid \mathcal{F}(t)$ *has a noncentral* χ^2*-distribution with* δ *degrees of freedom and noncentrality parameter*

$$\zeta(t,T) = \frac{e^{\int_t^T \beta(s)\,\mathrm{d}s} X(t)}{\mathcal{C}(t,T)},$$

where

$$\mathcal{C}(t,T) = \frac{1}{4}\int_t^T \sigma^2(s) e^{\int_s^T \beta(\tau)\,\mathrm{d}\tau}\,\mathrm{d}s.$$

Proof. By Lemma A.8 we need to show that

$$\mathrm{E}\left[e^{u\frac{X(T)}{\mathcal{C}(t,T)}} \mid \mathcal{F}(t)\right] = e^{-\frac{\delta(t)}{2}\log(1-2u)} e^{\frac{u}{1-2u}\frac{e^{\int_t^T \beta(\tau)\,\mathrm{d}\tau}}{\mathcal{C}(t,T)} X(t)}, \tag{4.3}$$

for $u \in \mathrm{i}\mathbb{R}$. From Theorem 2.6 we know that

$$\mathrm{E}\left[e^{vX(T)} \mid \mathcal{F}(t)\right] = e^{\phi(t,T,v)+\psi(t,T,v)X(t)},$$

where ϕ and ψ solve the Riccati equations,

$$\begin{aligned} \frac{\partial}{\partial t}\phi(t,T,v) &= -b(t)\psi(t,T,v), \\ \frac{\partial}{\partial t}\psi(t,t,v) &= -\frac{1}{2}\sigma^2\psi(t,T,v)^2 - \beta(t)\psi(t,T,v), \end{aligned} \tag{4.4}$$

with $\phi(T,T,v) = 0$ and $\psi(T,T,v) = v$ for $v \in \mathrm{i}\mathbb{R}$. Define the functions

$$\begin{aligned} \tilde{\phi}(t,T,v) &= -\frac{\delta(t)}{2}\log\left(1 - 2\mathcal{C}(t,T)v\right) \\ \tilde{\psi}(t,T,v) &= \frac{v e^{\int_t^T \beta(\tau)\,\mathrm{d}\tau}}{1 - 2\mathcal{C}(t,T)v}. \end{aligned}$$

Then the right side of (4.3) can be rewritten as

$$e^{\tilde{\phi}(t,T,v_{t,T})+\tilde{\psi}(t,T,v_{t,T})X(t)},$$

with $v_{t,T} = \frac{u}{\mathcal{C}(t,T)}$. If $\tilde{\phi}$ and $\tilde{\psi}$ solve the Riccati equations (4.4), we may conclude that (4.3) holds.

We see that $\tilde{\phi}(T,T,v) = 0$ and $\tilde{\psi}(T,T,v) = v$, and differentiating $\tilde{\phi}(t,T,v)$, remembering that $\delta(t)$ is constant,

$$\begin{aligned} \frac{\partial}{\partial t}\tilde{\phi}(t,T,v) &= -\delta(t)\frac{1}{1-2\mathcal{C}(t,T)v} v \frac{1}{4}\sigma^2(t) e^{\int_t^T \beta(\tau)\,\mathrm{d}\tau} \\ &= -b(t)\tilde{\psi}(t,T,v), \end{aligned}$$

and differentiating $\tilde{\psi}(t,T,v)$,

$$\begin{aligned}\frac{\partial}{\partial t}\tilde{\psi}(t,T,v) &= -\beta(t)\tilde{\psi}(t,T,v) - \frac{v e^{\int_t^T \beta(\tau)\,\mathrm{d}\tau}\frac{1}{2}\sigma^2(t) e^{\int_t^T \beta(\tau)\,\mathrm{d}\tau} v}{\left(1-\frac{1}{2}\int_t^T \sigma^2(s) e^{\int_s^T \beta(\tau)\,\mathrm{d}\tau} v\right)^2} \\ &= -\beta(t)\tilde{\psi}(t,T,v) - \frac{1}{2}\sigma^2(t)\tilde{\psi}(t,T,v)^2,\end{aligned}$$

the proof is concluded. □

Inserting constant parameters it is seen that the time-homogeneous distributional result is obtained.

In Example 4.1 a stochastic mortality rate of the form $\mu_x(t) = c_x(t) + \gamma_x(t)X_x(t)$ was considered, where $\mathbf{X}_x$ is a time-homogeneous CIR process. The application of the forward rate in the example is well known, and the formulas presented allow us to consider an endowment insurance, possibly combined with a life annuity. The interesting part of the example is the form of the mortality, where the stochastic part is a CIR process, thus having a known distribution, namely a rescaled non-central χ^2 distribution. This makes it particularly easy to perform simulations and find quantiles. The example can be extended to certain time-inhomogeneous CIR processes by use of Theorem 4.2.

4.3 Generalised Forward Rates

The generalised forward rates allow us to apply dependent interest and transition rates for valuation of certain life insurance contracts. To exemplify this in a simple setting, a survival model is considered with dependent interest and mortality rate. Throughout the section the setup from Section 1.1 is adopted.

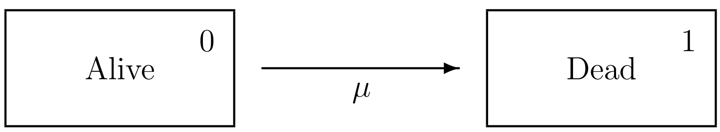

Figure 3: *State space for the survival model.*

Example 4.3. Consider a Markov model with state space $\mathcal{J} = \{0,1\}$ with states *alive* and *dead*, shown in Figure 3. State 1 is absorbing, and the only transition intensity is μ, from state *alive* to *dead*. We define the payment process $\mathbf{B}$ by the dynamics

$$\mathrm{d}B(t) = b_0(t)I(t)\,\mathrm{d}t + \Delta B(t)I(t)\,\mathrm{d}\varepsilon_n(t) + b_{01}(t)\,\mathrm{d}N(t),$$

for a piecewise continuous function $b_0(t)$, representing continuous payments (premiums or benefits) while alive, a single payment, $\Delta B(n)$ at time n, if alive, and a piecewise continuous function $b_{01}(t)$ representing the payment upon death at time t. We assume that there are no payments after time n, that is, $b_0(t) = b_{01}(t) = 0$ for $t > n$. The interest rate $r(t)$ and the mortality rate $\mu_x(t)$ are modelled as affine processes as follows. Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)^\top$ be a 2-dimensional affine process with parameters $b(t)$, $\mathcal{B}(t) = (\beta_1(t), \beta_2(t))$, $a(t)$ and $\alpha_1(t), \alpha_2(t)$, satisfying Theorem 2.12 such that the process is non-negative. Then, let

$$\begin{aligned} r(t) &= c^r(t) + \gamma^r(t)X_1(t), \\ \mu_x(t) &= c^\mu_x(t) + \gamma^\mu_x(t)X_2(t). \end{aligned}$$

Here, c^r and c^μ are real-valued functions, and γ^r and γ^μ are positive real-valued functions. It is, for notational simplicity, assumed that the affine process $\mathbf{X}$ is age-independent, i.e. the parameters are age-independent. As will be seen, this is in essence not much of a simplification, since $\gamma^\mu_x(t)$ is age-dependent.

Using Remark 3.5, a system of differential equations can be found for the forward rates. We have $d = 2$, and the forward interest rate can be found setting $N = \{1\}$, while the forward mortality rate can be found setting $N = \{2\}$. This yields two differential equation systems. However, Φ and Ψ are independent of N, and they solve the system of differential equations,

$$\begin{aligned} \frac{\partial}{\partial t}\Psi_1(t,T) &= -\frac{1}{2}\Psi(t,T)^\top\alpha_1(t)\Psi(t,T) - \beta_1(t)^\top\Psi(t,T) + \gamma^r(t), \\ \frac{\partial}{\partial t}\Psi_2(t,T) &= -\frac{1}{2}\Psi(t,T)^\top\alpha_2(t)\Psi(t,T) - \beta_2(t)^\top\Psi(t,T) + \gamma^\mu_x(t), \\ \Psi(T,T) &= 0, \\ \frac{\partial}{\partial t}\Phi(t,T) &= -b(t)^\top\Psi(t,T) + c^r(t) + c^\mu_x(t), \\ \Phi(T,T) &= 0, \end{aligned} \tag{4.5}$$

where $\Psi = (\Psi_1, \Psi_2)^\top$. The forward interest rate,

$$f^{r:(r+\mu)}_{x,t}(T) = W^r(t,T) + V^r(t,T)X(t),$$

satisfies the system of differential equations,

$$\begin{aligned}
\frac{\partial}{\partial t} V_1^r(t,T) &= -\Psi(t,T)^\top \alpha_1(t) V^r(t,T) - \beta_1(t)^\top V^r(t,T),\\
\frac{\partial}{\partial t} V_2^r(t,T) &= -\Psi(t,T)^\top \alpha_2(t) V^r(t,T) - \beta_2(t)^\top V^r(t,T),\\
V_1^r(T,T) &= \gamma^r(T),\\
V_2^r(T,T) &= 0,\\
\frac{\partial}{\partial t} W^r(t,T) &= -b(t)^\top V^r(t,T),\\
W^r(T,T) &= c^r(T),
\end{aligned} \tag{4.6}$$

where $V^r = (V_1^r, V_2^r)^\top$, and the forward mortality rate

$$f_{x,t}^{\mu:(r+\mu)}(T) = W^\mu(t,T) + V^\mu(t,T)X(t),$$

satisfies the system of differential equations,

$$\begin{aligned}
\frac{\partial}{\partial t} V_1^\mu(t,T) &= -\Psi(t,T)^\top \alpha_1(t) V^\mu(t,T) - \beta_1(t)^\top V^\mu(t,T),\\
\frac{\partial}{\partial t} V_2^\mu(t,T) &= -\Psi(t,T)^\top \alpha_2(t) V^\mu(t,T) - \beta_2(t)^\top V^\mu(t,T),\\
V_1^\mu(T,T) &= 0,\\
V_2^\mu(T,T) &= \gamma_x^\mu(T),\\
\frac{\partial}{\partial t} W^\mu(t,T) &= -b(t)^\top V^\mu(t,T),\\
W^\mu(T,T) &= c_x^\mu(T),
\end{aligned} \tag{4.7}$$

where $V^\mu = (V_1^\mu, V_2^\mu)^\top$. We note that the solutions all depend on the age x, since the age x appears in $\gamma_x^\mu(t)$ in the diffential equation for Ψ_2, and Ψ_2 appears in all the other differential equations. Thus Ψ, Φ, V^r, W^r, V^μ, and W^μ depend on x. This also implies that the advantage of age-independence of the affine process $(\mathbf{X}_1, \mathbf{X}_2)$ is limited. In the following, we omit the dependence in the notation of the solutions, and keep it in the notation of the forward rates.

An interesting observation is, that if one adds the differential equation systems for (V^r, W^r) and (V^μ, W^μ), which is the system of differential equations describing $f^{r:(r+\mu)} + f^{\mu:(r+\mu)}$, a system of differential equations for $(V^r + V^\mu, W^r + W^\mu)$ is obtained. This is seen to be identical to the one obtained by letting $N = \{1, 2\}$, which is the system of differential equations for the forward rate written $f^{(r+\mu):(r+\mu)}$. That is exactly the content of Lemma 3.7.

The present value for the future payments in the life insurance contract is, for $t < n$,

$$\begin{aligned} PV(t) &= \int_t^n e^{-\int_t^s r(\tau)\,\mathrm{d}\tau}\,\mathrm{d}B(s) \\ &= \int_t^n e^{-\int_t^s r(\tau)\,\mathrm{d}\tau}\left(I(s)b_0(s)\,\mathrm{d}s + b_{01}(s)\,\mathrm{d}N(s)\right) + I(n)\Delta B(n)e^{-\int_t^n r(\tau)\,\mathrm{d}\tau}, \end{aligned}$$

and conditioning on $I(t) = 1$ and $\mathcal{F}^Y(\infty)$ the usual results for deterministic rates can be used to obtain

$$\begin{aligned} &\mathrm{E}\left[PV(t)\,\middle|\,I(t) = 1, \mathcal{F}^Y(\infty)\right] \\ &= \int_t^n e^{-\int_t^s (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\left(b_0(s) + b_{01}(s)\mu_x(s)\right)\,\mathrm{d}s + \Delta B(n)e^{-\int_t^n (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}, \end{aligned}$$

which is well known: It is the prospective reserve when $r(t)$ and $\mu_x(t)$ are deterministic. To find the prospective reserve at time t, which we denote by $R(t)$, we condition on $\mathcal{F}^Y(u)$ and find

$$R(t) = \int_t^n e^{-\int_t^s f_{x,t}^{r+\mu}(\tau)\,\mathrm{d}\tau}\left(b_0(s) + b_{01}(s)f_{x,t}^{\mu:(r+\mu)}(s)\right)\,\mathrm{d}s + \Delta B(n)e^{-\int_t^n f_{x,t}^{r+\mu}(\tau)\,\mathrm{d}\tau},$$

where Remark 3.5 is applied. We remember from Definition 3.6 and Lemma 3.7 that $f_{x,t}^{r+\mu} = f_{x,t}^{(r+\mu):(r+\mu)} = f_{x,t}^{r:(r+\mu)} + f_{x,t}^{\mu:(r+\mu)}$.

The expression is similar to the one with independent interest and mortality. It is not straightforward to put up Thiele's differential equation, since the forward rates now depend on t. However, with a little trick proposed in [17], a differential equation can be found. We introduce the auxiliary function,

$$R^\circ(u;t) = \int_u^n e^{-\int_u^s f_{x,t}^{r+\mu}(\tau)\,\mathrm{d}\tau}\left(b_0(s) + b_{01}(s)f_{x,t}^{\mu:(r+\mu)}(s)\right)\,\mathrm{d}s + \Delta B(n)e^{-\int_u^n f_{x,t}^{r+\mu}(\tau)\,\mathrm{d}\tau},$$

for $t, u \in [0, n]$, which is the reserve at time u, calculated with the forward rates at time t. It is straightforward to find a differential equation for this function, and by differentiation we obtain

$$\begin{aligned} \frac{\partial}{\partial u}R^\circ(u;t) &= f_{x,t}^{r+\mu}(u)R^\circ(u;t) - b_0(u) - f_{x,t}^{\mu:(r+\mu)}(u)b_{01}(u) \\ &= f_{x,t}^{r:(r+\mu)}(u)R^\circ(u;t) - b_0(u) - f_{x,t}^{\mu:(r+\mu)}(u)\left(b_{01}(u) - R^\circ(u;t)\right), \end{aligned}$$

with boundary condition $R^\circ(n-;t) = \Delta B(n)$. Solving on the interval $[t, n]$ yields $R^\circ(t;t) = R(t)$, i.e. the prospective reserve at time t. ○

The example shows that in a simple alive-dead insurance contract, we can apply dependent affine interest and mortality rates, with forward rates determined by a differential equation system. We also find a Thiele differential equation for the prospective reserve at time t. The results build on Remark 3.5, which is not restricted to only two processes. Therefore, it is also immediate to extend the example to include a surrender state with a surrender rate dependent on the interest and mortality rate, if the value paid upon surrender at time s is independent of the interest and transition rate process.

To examine the possible applications of the generalised forward rates in more advanced models, a 3-state disability model is considered in the following example. The model is chosen such that there, in the case of deterministic interest and transition rates, exist closed form expressions. It is then examined whether results can be obtained when (some of) the interest and transition rates are dependent affine processes.

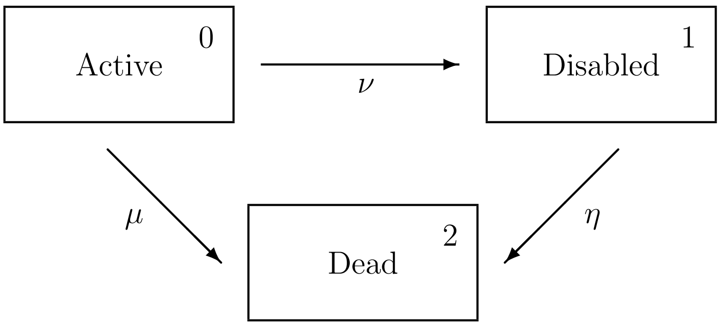

Figure 4: *State space for the disability model.*

Example 4.4. Let $\mathcal{J} = \{0, 1, 2\}$ be the state space corresponding to the states *active*, *disabled* and *dead*, shown in Figure 4. We assume that no recovery from the state *disabled* is possible. Thus the non-zero transition rates are,

- the disability rate, ν_x, from state 0 to 1,
- the mortality rate while active, μ_x, from state 0 to 2,
- the mortality rate while disabled η_x, from state 1 to 2.

The interest rate, r, and the transition rates ν_x and μ_x are modelled by a 3-dimensional affine process, $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)^\top$, similar to Example 4.3. The transition rate η_x is assumed to be deterministic. Let the parameters $b(t)$, $\mathcal{B}(t) =$

$(\beta_1(t), \beta_2(t), \beta_3(t))$, $a(t)$ and $\alpha_1(t), \alpha_2(t), \alpha_3(t)$ satisfy Theorem 2.12 such that the process is non-negative, and let

$$\begin{aligned} r(t) &= c^r(t) + \gamma^r(t) X_1(t), \\ \nu_x(t) &= c_x^\nu(t) + \gamma_x^\nu(t) X_2(t), \\ \mu_x(t) &= c_x^\mu(t) + \gamma_x^\mu(t) X_3(t). \end{aligned}$$

Again, for notational simplicity, it is assumed that the affine process does not depend on the age, x.

Define the payment process $\mathbf{B}$ by the dynamics,

$$\begin{aligned} \mathrm{d}B(t) = 1_{\{Z(t)=0\}} b_0(t)\,\mathrm{d}t + 1_{\{Z(t)=1\}} b_1(t)\,\mathrm{d}t \\ + b_{01}(t)\,\mathrm{d}N_{01}(t) + b_{02}(t)\,\mathrm{d}N_{02}(t) + b_{12}(t)\,\mathrm{d}N_{12}(t). \end{aligned}$$

Here, the functions $b_i(t)$ and $b_{ij}(t)$ are piecewise continuous functions, equal to 0 for $t > n$, i.e. the contract ends at time n.

The present value at time t is found,

$$\begin{aligned} PV(t) &= \int_t^n e^{-\int_t^s r(\tau)\,\mathrm{d}\tau} dB(s) \\ &= \int_t^n e^{-\int_t^s r(\tau)\,\mathrm{d}\tau} \Big(\big(1_{\{Z(s)=0\}} b_0(s) + 1_{\{Z(s)=1\}} b_1(s)\big)\,\mathrm{d}s \\ &\quad + b_{01}(s)\,\mathrm{d}N_{01}(s) + b_{02}(s)\,\mathrm{d}N_{02}(s) + b_{12}(s)\,\mathrm{d}N_{12}(s)\Big). \end{aligned}$$

Conditioning on $Z(t) = 1$ and $\mathcal{F}^Y(\infty)$, we find

$$\mathrm{E}\left[PV(t)\,\middle|Z(t) = 1, \mathcal{F}^Y(\infty)\right] = \int_t^n e^{-\int_t^s (r(\tau)+\eta_x(\tau))\,\mathrm{d}\tau} \left(b_1(s) + \eta_x(s) b_{12}(s)\right)\,\mathrm{d}s,$$

which is the prospective reserve for the state *disabled*, conditioned on knowing the stochastic interest and transition rates. Conditioning on $\mathcal{F}^Y(t)$, the unconditioned prospective reserve, which we denote $R_1(t)$, is obtained

$$R_1(t) = \int_t^n e^{-\int_t^s \left(f_{x,t}^r(\tau)+\eta_x(\tau)\right)\mathrm{d}\tau} \left(b_1(s) + \eta_x(s) b_{12}(s)\right)\,\mathrm{d}s.$$

If one sets $c_x^\nu = c_x^\mu = \gamma_x^\nu = \gamma_x^\mu = 0$, then a system of differential equations for the forward rate can be found using Remark 3.5 with $d = 3$. The set N should be such that $1 \in N$. (It makes no difference if $2 \in N$ or $3 \in N$ when the functions mentioned equal 0.)

For the state *active*, we also condition on $Z(t)=0$ and $\mathcal{F}^Y(\infty)$. In that case, we can perform the calculations for deterministic interest and transition rates, thus

$$\begin{aligned}
&\mathrm{E}\left[PV(t)\,\middle|\,Z(t)=0,\mathcal{F}^Y(\infty)\right]\\
&=\int_t^n e^{-\int_t^s (r(\tau)+\mu_x(\tau)+\nu_x(\tau))\,\mathrm{d}\tau}\Big(b_0(s)+\mu_x(s)b_{01}(s)\\
&\quad+\nu_x(s)\,\mathrm{E}\left[PV(s)\,\middle|\,Z(s)=1,\mathcal{F}^Y(\infty)\right]\Big)\,\mathrm{d}s\\
&=\int_t^n e^{-\int_t^s (r(\tau)+\mu_x(\tau)+\nu_x(\tau))\,\mathrm{d}\tau}\left(b_0(s)+\mu_x(s)b_{01}(s)\right.\\
&\quad\left.+\nu_x(s)\int_s^n e^{-\int_s^u (r(\tau)+\eta_x(\tau))\,\mathrm{d}\tau}\left(b_1(u)+\eta_x(u)b_{12}(u)\right)\,\mathrm{d}u\right)\mathrm{d}s.
\end{aligned}$$

To find the prospective reserve, we condition on $\mathcal{F}^Y(t)$. Then we need to find the quantities,

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu_x(\tau)+\nu_x(\tau))\,\mathrm{d}\tau}\,\middle|\,\mathcal{F}^Y(t)\right],\\
&\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu_x(\tau)+\nu_x(\tau))\,\mathrm{d}\tau}\mu_x(s)\,\middle|\,\mathcal{F}^Y(t)\right],\\
&\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu_x(\tau)+\nu_x(\tau))\,\mathrm{d}\tau}\nu_x(s)e^{-\int_s^u r(\tau)\,\mathrm{d}\tau}\,\middle|\,\mathcal{F}^Y(t)\right].
\end{aligned}\tag{4.8}$$

The first two terms are straightforward, and equal

$$e^{-\int_t^s f_{x,t}^{r+\mu+\nu}(\tau)\,\mathrm{d}\tau}\quad\text{and}\quad e^{-\int_t^s f_{x,t}^{r+\mu+\nu}(\tau)\,\mathrm{d}\tau}f_{x,t}^{\mu:(r+\mu+\nu)}(s),$$

respectively. Differential equations for the forward rates can be found using Remark 3.5, with the actual c and γ functions and $N=\{1\}$. It may be noted, that the system of differential equations for Φ and Ψ (thus also for the W and V functions) obtained now is different from the one for the forward rate $f_{x,t}^r$ above.

Remark 3.5 does not give a solution to the third quantity. Hence the prospective reserve is not immediate. It is, however, the belief of the author that a trick similar to the one that led to Theorem 3.4 could result in a new system of differential equations describing the third line of (4.8); The idea is to consider the quantity

$$\mathrm{E}\left[e^{-\int_t^u \left(r(\tau)+\mathcal{I}(\tau,S,\varepsilon)\mu_x(\tau)+1_{(t,s)}(\tau)\nu_x(\tau)\right)\mathrm{d}\tau}\,\middle|\,\mathcal{F}^Y(t)\right],$$

where $S\in(t,s)$ and $\mathcal{I}$ is the continuous "indicator" function defined in Section 3.2. Differentiating the expression with respect to s yields an expression similar to the one we seek, i.e. the third line in (4.8). It will converge as desired for $S \nearrow s$ and $\varepsilon \searrow 0$. The task is then to show that the system of differential equations and the solution also converge. This is similar to the problem studied throughout Lemma 3.2 and 3.3 and Theorem 3.4.

If the mortality rate for the state *disabled* was modelled as an affine stochastic process, dependent on the other rates, the quantities to be found instead of the third line of (4.8) would have been

$$\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu_x(\tau)+\nu_x(\tau))\,\mathrm{d}\tau} \nu_x(s) e^{-\int_s^u (r(\tau)+\eta_x(\tau))\,\mathrm{d}\tau} \middle| \mathcal{F}^Y(t) \right],$$
$$\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu_x(\tau)+\nu_x(\tau))\,\mathrm{d}\tau} \nu_x(s) e^{-\int_s^u (r(\tau)+\eta_x(\tau))\,\mathrm{d}\tau} \eta_x(u) \middle| \mathcal{F}^Y(t) \right].$$

Applying similar methods, and differentiating both with respect to s and u, the author believes it should be possible to reach a set of differential equations for these quantities as well. ○

In the example, it was attempted to apply the results to a more complicated life insurance contract that, in the case of deterministic rates, has a closed form expression for the prospective reserve. The results are not directly applicable, but it is the belief of the author that the results of this thesis can be extended, such that they can be applied to this example.

In multistate contracts, where closed form expressions for the prospective reserve do not exist, it is unknown whether the generalised forward rates are applicable.

4.4 Solvency II Capital Requirement

In Section 1 we discussed how one can determine the capital requirement, x_{SCR}, which is a quantile in the distribution of the loss after 1 year, with the loss defined in Definition 1.1. This is the loss realised after one year, associated with the systematic risk, while the unsystematic risk is not considered.

In Example 1.7 we realised that a mortality rate where the stochastic part was ignored for the mortality during the first year had significant advantages. For a mortality rate based on an affine process $\mathbf{X}$, this could be formalised as

$$\mu_x(t) = c_x(t) + \gamma_x(t) X(t \cdot 1_{\{t>1\}}).$$

Considering a stochastic interest rate as well, the same idea can be applied, thus the interest rate for the first year is deterministic. In practice one can use a forward interest rate for the first year, and the stochastic interest rate afterwards.

We consider the setup from Example 4.3, and find the loss at time 1.

Example 4.5. Let the payment process $\mathbf{B}$ be defined as in Example 4.3, and let the interest and transition rates be defined by,

$$r(t) = \begin{cases} c^r(t) + \gamma^r(t)X_1(0), & t \leq 1, \\ c^r(t) + \gamma^r(t)X_1(t), & t > 1, \end{cases}$$

$$\mu_x(t) = \begin{cases} c^\mu_x(t) + \gamma^\mu_x(t)X_2(0), & t \leq 1, \\ c^\mu_x(t) + \gamma^\mu_x(t)X_2(t), & t > 1. \end{cases}$$

with c^r, γ^r, c^μ_x and γ^μ_x being continuous functions where γ^r_x and γ^μ_x are positive. Then $r(t)$ and $\mu_x(t)$ are deterministic for $t \leq 1$, and affine processes for $t > 1$.

By Definition 1.1 we consider the loss

$$\begin{aligned} L(t) &= \mathrm{E}\left[PV(0)\,\middle|\mathcal{F}^Z(0) \vee \mathcal{F}^Y(t)\right] - \mathrm{E}\left[PV(0)\,\middle|\mathcal{F}^Z(0) \vee \mathcal{F}^Y(0)\right] \\ &= \mathrm{E}\left[PV(0)\,\middle|\mathcal{F}^Y(t)\right] - R(0). \end{aligned}$$

Proceding along the lines of Example 4.3 with the aim of finding the loss at time 1, we consider

$$\mathrm{E}\left[e^{-\int_0^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\middle|\, \mathcal{F}^Y(u)\right] = e^{-\int_0^1 (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\, \mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\middle|\, \mathcal{F}^Y(u)\right],$$

and

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_0^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\mu_x(t)\middle|\, \mathcal{F}^Y(u)\right] \\ &\qquad = e^{-\int_0^1 (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\, \mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\mu_x(t)\middle|\, \mathcal{F}^Y(u)\right], \end{aligned}$$

for $u \in \{0, 1\}$. If $t \leq 1$, the interest and transition rates are deterministic, thus the case $t > 1$ is considered. If $u = 1$, the solution is given by Remark 3.5

$$\begin{aligned} \mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\middle|\, \mathcal{F}^Y(1)\right] &= e^{-\int_1^t f^{r+\mu}_{x,1}(\tau)\,\mathrm{d}\tau}, \\ \mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau}\mu_x(t)\middle|\, \mathcal{F}^Y(1)\right] &= e^{-\int_1^t f^{r+\mu}_{x,1}(\tau)\,\mathrm{d}\tau} f^{r+\mu}_{x,1}(t), \end{aligned}$$

where the system of differential equations for the forward rates written is those from Example 4.3, i.e. (4.5)–(4.7), remembering that $f^{r+\mu}_{x,1} = f^{r:(r+\mu)}_{x,1} + f^{\mu:(r+\mu)}_{x,1}$.

If $u = 0$ things are more complicated, since Remark 3.5 cannot be applied directly. Now define the functions

$$\tilde{r}(t) = 1_{\{t>1\}} r(t),$$
$$\tilde{\mu}_x(t) = 1_{\{t>1\}} \mu_x(t),$$

and see that we can write

$$\mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau} \middle| \mathcal{F}^Y(u) \right] = \mathrm{E}\left[e^{-\int_0^t (\tilde{r}(\tau)+\tilde{\mu}_x(\tau))\,\mathrm{d}\tau} \middle| \mathcal{F}^Y(u) \right],$$

and

$$\mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau} \mu_x(t) \middle| \mathcal{F}^Y(u) \right] = \mathrm{E}\left[e^{-\int_0^t (\tilde{r}(\tau)+\tilde{\mu}_x(\tau))\,\mathrm{d}\tau} \tilde{\mu}_x(t) \middle| \mathcal{F}^Y(u) \right].$$

If $\tilde{r}$ and $\tilde{\mu}_x$ are affine processes, then Remark 3.5 can be applied to find a system of differential equations for appropriate forward rates. However, $\tilde{r}$ and $\tilde{\mu}_x$ have discontinuous parameter functions. To see this, note that we can write

$$\tilde{r}(t) = 1_{\{t>1\}} c^r(t) + 1_{\{t>1\}} \gamma^r(t) X_1(t) = \tilde{c}^r(t) + \tilde{\gamma}^r(t) X_1(t),$$
$$\tilde{\mu}_x(t) = 1_{\{t>1\}} c^\mu_x(t) + 1_{\{t>1\}} \gamma^\mu_x(t) X_2(t) = \tilde{c}^\mu_x(t) + \tilde{\gamma}^\mu_x(t) X_2(t).$$

Believing that Remark 3.5 holds for these discontinuous parameter functions, one finds that a solution is given by

$$\mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau} \middle| \mathcal{F}^Y(0) \right] = e^{-\int_0^t \tilde{f}^{r+\mu}_{x,0}(\tau)\,\mathrm{d}\tau},$$
$$\mathrm{E}\left[e^{-\int_1^t (r(\tau)+\mu_x(\tau))\,\mathrm{d}\tau} \mu_x(t) \middle| \mathcal{F}^Y(0) \right] = e^{-\int_0^t \tilde{f}^{r+\mu}_{x,0}(\tau)\,\mathrm{d}\tau} \tilde{f}^{\mu:(r+\mu)}_{x,0}(t),$$

with differential equations (4.5)–(4.7), where the functions c^r, γ^r, c^μ_x and γ^μ_x are replaced by the functions $\tilde{c}^r$, $\tilde{\gamma}^r$, $\tilde{c}^\mu_x$ and $\tilde{\gamma}^\mu_x$, respectively. It is the belief of the author that the result is true: It is almost straightforward to see that $\int_0^t \tilde{r}(s)\,\mathrm{d}s$ and $\int_0^t \tilde{\mu}_x(s)\,\mathrm{d}s$ are affine processes, and the author believes that the results of Section 3.2 could be extended to cover the case of $\tilde{r}$ and $\tilde{\mu}_x$.

Assuming that the result holds, we see that the functions c^r, γ^r, c^μ_x and γ^μ_x equal the functions $\tilde{c}^r$, $\tilde{\gamma}^r$, $\tilde{c}^\mu_x$ and $\tilde{\gamma}^\mu_x$ for $t > 1$. Thus, by inspection of the system of differential equations, we see that

$$\tilde{f}^{r:(r+\mu)}_{x,1}(t) = f^{r:(r+\mu)}_{x,1}(t),$$
$$\tilde{f}^{\mu:(r+\mu)}_{x,1}(t) = f^{\mu:(r+\mu)}_{x,1}(t),$$

and we only have one set of forward rates to consider. Note also that by the definition of the forward rates, for $t \leq 1$

$$\tilde{f}^{r:(r+\mu)}_{x,1}(t) = r(t),$$
$$\tilde{f}^{\mu:(r+\mu)}_{x,1}(t) = \mu_x(t).$$

We find,

$$\begin{aligned} &\mathrm{E}\left[PV(0)\,\middle|\mathcal{F}^Z(0) \vee \mathcal{F}^Y(1)\right] \\ &\qquad = \int_0^n e^{-\int_0^s \tilde{f}^{r+\mu}_{x,1}(\tau)\,\mathrm{d}\tau} \left(b_0(s) + b_{01}(s)\tilde{f}^{\mu:(r+\mu)}_{x,1}(s)\right)\,\mathrm{d}s + \Delta B(n) e^{-\int_0^n \tilde{f}^{r+\mu}_{x,1}(\tau)\,\mathrm{d}\tau}, \end{aligned}$$

and

$$R(0) = \int_0^n e^{-\int_0^s \tilde{f}^{r+\mu}_{x,0}(\tau)\,\mathrm{d}\tau} \left(b_0(s) + b_{01}(s)\tilde{f}^{\mu:(r+\mu)}_{x,0}(s)\right)\,\mathrm{d}s + \Delta B(n) e^{-\int_0^n \tilde{f}^{r+\mu}_{x,0}(\tau)\,\mathrm{d}\tau}.$$

The forward rates at time 0, $f^{(\cdot)}_{x,0}(t)$ are deterministic, thus $R(0)$ is deterministic. The forward rates at time 1 contain the stochastic variable $X(1) = (X_1(1), X_2(1))$. With simulation of $X(1)$, one finds the empirical distribution of $L(1)$ which is the solvency capital requirement, x_{SCR}. ○

References

[1] Herbert Amann. *Ordinary Differential Equations, An Introduction to Nonlinear Analysis.* de Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin - New York 1983, 1990.

[2] Matthias Börger. Deterministic Shock vs. Stochastic Value-at-Risk – An Analysis of the Solvency II Standard Model Approach to Longevity Risk. *Blätter der DGVFM*, 31 (2):225–259, 2010.

[3] Marcus C. Christiansen. A Sensitivity Analysis Concept for Life Insurance with Respect to a Valuation Basis of Infinite Dimension. *Insurance: Mathematics & Economics*, 42:680–690, 2008.

[4] Marcus C. Christiansen. Biometric Worst Case Scenarios for Multi-State Life Insurance Policies. *Insurance: Mathematics & Economics*, 47:190–197, 2010.

[5] Marcus C. Christiansen and Mogens Steffensen. Safe-Side Scenarios for Financial and Biometrical Risk. (Working Paper), Institut für Versicherungswissenschaften, Universität Ulm and Department of Mathematical Sciences, University of Copenhagen, January 2011.

[6] Mikkel Dahl and Thomas Møller. Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk. *Insurance: Mathematics & Economics*, 39 (2):193–217, 2006.

[7] Damir Filipovic. *Term Structure Models.* Spring Finance. Springer, 2009.

[8] Damir Filipovic and Eberhard Mayerhofer. Affine Diffusion Processes: Theory and Applications. *Radon Series Comp. Appl. Math*, 8:1–40, 2009.

[9] Paul Glasserman. *Monte Carlo Methods in Financial Engineering.* Applications of Mathematics. Springer, 2004.

[10] Ernst Hansen. *Measure Theory.* Department of Mathematical Sciences, University of Copenhagen, Fourth edition, 2009.

[11] Martin Jacobsen. Stokastiske Integraler. (Lecture Notes), Department of Mathematical Sciences, University of Copenhagen, 2008.

[12] Vladimir Kalashnikov and Ragnar Norberg. On the Sensitivity of Premiums and Reserves to Changes in Valuation Elements. *Scandinavian Actuarial Journal*, 3:238–256, 2003.

[13] George James Lidstone. Changes in Pure Premium Values Consequent upon Variations in the Rate of Interest or the Rate of Mortality. *J. Institute of Actuaries*, 3, 1905.

[14] Yoosef Maghsoodi. Solution of the Extended CIR Term Structure and Bond Option Valuation. *Mathematical Finance*, 6:89–109, 1996.

[15] Moshe A. Milevsky and David Promislow. Mortality Derivatives and the Option to Annuitise. *Insurance: Mathematics & Economics*, 29 (3):299–318, 2001.

[16] Kristian R. Miltersen and Svein-Arne Persson. Is Mortality Dead? Stochastic Forward Force of Mortality Rate Determined by No Arbitrage. (Working Paper), Norwegian School of Economics and Business Administration, 2005.

[17] Thomas Møller and Mogens Steffensen. *Market-Valuation Methods in Life and Pension Insurance*. International Series on Actuarial Science. Cambridge University Press, 2007.

[18] Ragnar Norberg. Forward mortality and other vital rates - Are they the way forward? *Insurance: Mathematics & Economics*, 47 (2):105–112, 2010.

[19] Jesper Lund Pedersen. Stochastic Processes in Life Insurance. (Lecture Notes), Department of Mathematical Sciences, University of Copenhagen, 2010.

[20] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.

A Mathematical Appendix

This appendix contains various auxiliary results used in the thesis.

Lemma A.1. *Let A be a $d \times d$ symmetric positive semi-definite matrix. If the ith diagonal entry is zero, $A_{ii} = 0$, then $A_{ik} = A_{ki} = 0$ for $k = 1, \ldots, d$.*

Proof. A is symmetric and can be diagonalized by an ortonormal matrix Q and a diagonal matrix D with $D_{ii} = \lambda_i$,

$$A = QDQ^\top.$$

See that A has a square root, $A^{\frac{1}{2}} = QD^{\frac{1}{2}}Q^\top$. Assume now that $A_{ii} = 0$. Then,

$$0 = e_i^\top A e_i = e_i^\top A^{\frac{1}{2}} A^{\frac{1}{2}} e_i = \|A^{\frac{1}{2}} e_i\|^2,$$

thus $A^{\frac{1}{2}} e_i = 0$. But then $Ae_i = A^{\frac{1}{2}} A^{\frac{1}{2}} e_i = 0$, and

$$A_{ki} = e_k^\top A e_i = 0$$

for $k = 1, \ldots, d$. By symmetry $A_{ik} = A_{ki} = 0$. □

A.1 Differential Equations

Grönwall's[3] inequality is a classic result which is essential when dealing with differential equations. It is stated here without proof. It is proved in [1], Lemma (6.1).

Lemma A.2. *(Grönwall's inequality) Let $J \subset \mathbb{R}$ be an interval, $t_0 \in J$ and let $a, \beta, u : J \to \mathbb{R}_+$ be continuous functions. If*

$$u(t) \leq a(t) + \left| \int_{t_0}^{t} \beta(s) u(s) \, \mathrm{d}s \right|, \quad \forall t \in J,$$

then

$$u(t) \leq a(t) + \left| \int_{t_0}^{t} a(s)\beta(s) e^{\left| \int_s^t \beta(\tau) \, \mathrm{d}\tau \right|} \, \mathrm{d}s \right|, \quad \forall t \in J.$$

[3] Thomas Hakon Grönwall (1877-1932) was a Swedish mathematician. After emigrating to the United States he spelled his name Gronwall, and the result is often referred to as Gronwall's inequality.

We now present a theorem containing some classic results about existence and uniqueness of solutions of differential equations.

Theorem A.3. *Let $K \in \{\mathbb{R}, \mathbb{C}\}$. Consider the system of ordinary differential equations*

$$\begin{aligned} \frac{\partial}{\partial t} f(t,u) &= R(t, f(t,u)), \\ f(t_0, u) &= u, \end{aligned} \tag{A.1}$$

where $R : \mathbb{R} \times K^d \to K^d$ and $(t_0, u) \in \mathbb{R} \times K^d$. If $R(t,v)$ is continuous in t and locally Lipschitz continuous in v, then the following holds;

(a) *For every $u \in K^d$, there exist life times*

$$t_-(u), t_+(u) \in [-\infty, \infty], \quad t_-(u) < t_0 < t_+(u)$$

such that there exists a unique solution $f(\cdot, u) : (t_-(u), t_+(u)) \to K^d$ of (A.1).

(b) *The domain*

$$\mathcal{D}_K = \left\{ (t,u) \in \mathbb{R} \times K^d \mid t_-(u) < t < t_+(u) \right\}$$

is open in $\mathbb{R} \times K^d$ and maximal in the sense, that either $t_+(u) = \infty$ or

$$\lim_{t \nearrow t_+(u)} \|f(t,u)\| = \infty,$$

and either $t_-(u) = -\infty$ or

$$\lim_{t \searrow t_-(u)} \|f(t,u)\| = \infty.$$

(c) *For every $t \in \mathbb{R}$, the t-section*

$$\mathcal{D}_K(t) = \left\{ u \in K^d \mid (t,u) \in \mathcal{D}_K \right\}$$

is open in K^d, and non-expanding in t in the following sense,

$$K^d = \mathcal{D}_K(t_0) \supset \mathcal{D}_K(t_1) \supset \mathcal{D}_K(t_2),$$

for $t_0 < t_1 < t_2$ and for $t_2 < t_1 < t_0$.

Proof. Part (a) is proved in [1], Theorem (7.6), which also states that D_K is maximal. In [1], Theorem (8.3), it is proved that D_K is open.

For part (c) assume (a) and (b) holds. $D_K(t)$ is open since D_K is. See that $\mathcal{D}_K(t) = \left\{u \in K^d \mid t_-(u) < t < t_+(u)\right\}$. For $u \in K^d$ we always have $t_-(u) < t_0 < t_+(u)$, and therefore $K^d = \mathcal{D}_K(t_0)$.

Let $t_0 \leq t_1 < t_2$ and $u \in \mathcal{D}_K(t_2)$. Then $t_+(u) > t_2 > t_1$, and then $u \in \mathcal{D}_K(t_1)$, i.e. $\mathcal{D}_K(t_2) \subset \mathcal{D}_K(t_1)$. The argument for the case $t_2 < t_1 < t_0$ is analogous. □

The theorem is useful when considering the Riccati equations (2.10). See for example that with $u = 0$, we trivially have $\frac{\partial}{\partial t}\psi(t, T, u) = 0$, i.e. $t_+(0) = +\infty$. Thus, by Remark 2.7, ϕ and ψ are well defined for $u = 0$, i.e. $0 \in \mathcal{D}_{\mathbb{C}}(t)$ for all $t \geq 0$.

We now consider continuity and differentiability of solutions of parameter dependent differential equation problems. In the following let $J \subset \mathbb{R}$ be an open interval, let $K \in \{\mathbb{R}, \mathbb{C}\}$ and let $D \subset K^d$ be open. Also, let $\Lambda \subset \mathbb{R}^k$ be a locally compact[4] convex set, denoting the parameter space. We then consider the function $f : J \to K^d$ given by the differential equation,

$$\begin{aligned} \frac{\partial}{\partial t} f(t, t_0, u, \lambda) &= R(t, f(t, t_0, u, \lambda), \lambda), \\ f(t_0, t_0, u, \lambda) &= u, \end{aligned} \tag{A.2}$$

for $(t_0, u, \lambda) \in J \times D \times \Lambda$. We assume that $R(t, x, \lambda)$ is continuous in t and locally Lipschitz continuous in x. Remember that a $\mathcal{C}^1$ function is locally Lipschitz continuous.

By Theorem A.3 there exists a unique solution to the differential equation system (A.2) in the interval $J(t_0, u, \lambda) = (t_-(u, \lambda), t_+(u, \lambda))$ for all $\lambda \in \Lambda$. We denote the domain of the function f by,

$$\mathcal{D}(f, \Lambda) = \{(t, t_0, u, \lambda) \in J \times J \times D \times \Lambda \mid t \in J(t_0, u, \lambda)\}.$$

Theorem A.4. (Continuity) *Assume that Λ is open, and that $R(t, x, \lambda)$ is locally Lipschitz continuous in (x, λ). Then the solution $f : \mathcal{D}(f, \Lambda) \to K^d$ of (A.2) is locally Lipschitz continuous.*

Proof. This is proven in [1], Theorem (8.4). □

[4] A subset of $\mathbb{R}^d$ or $\mathbb{C}^d$ is *locally compact* if every point of the set has a compact neighborhood.

The following theorem assumes slighty more about R, but in turn guarantees differentiability of the solution.

Theorem A.5. (Differentiability) *Assume that* Λ *is open, and that* $R(t,x,\lambda)$ *is* $\mathcal{C}^1$ *in* (x,λ). *Then the solution* $f:\mathcal{D}(f,\Lambda)\to K^d$ *of* (A.2) *is* $\mathcal{C}^1$. *Moreover, the partial derivative* $a(t)=\frac{\partial f}{\partial t_0}(t,t_0,u,\lambda)$ *satisfies the differential equation system,*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}a(t) &= \frac{\partial}{\partial x}R(t,f(t,t_0,u,\lambda),\lambda)a(t),\\ a(t_0) &= -R(t_0,u,\lambda).\end{aligned} \tag{A.3}$$

Proof. This is proven in [1], Theorem (9.2). □

A Riccati (differential) equation is a differential equation that is quadratic in the unknown function, i.e. of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}f(t)=a(t)f(t)^2+b(t)f(t)+c(t),\quad f(t_0)=u.$$

They are named after Count Jacopo Francesco Riccati (1676-1754). Riccati differential equations are fundamental when studying affine process (see Theorem 2.6).

The Riccati equations can be solved explicitly when the coefficients are time-independent, as stated in the following lemma. Also, as is evident from the proof of Theorem 4.2, the Riccati equations has an explicit solution in certain cases of time-inhomogeneous coefficients. This is however not stated in the following lemma.

Lemma A.6. *Consider the Riccati differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}f(t)=af(t)^2+bf(t)+c,\quad f(t_0)=u.$$

where $a,b,c,u\in\mathbb{R}$ *such that* $a\neq 0$ *and* $b^2-4ac>0$, *and with* $t_0\in\mathbb{R}_+$. *Define* $\theta=\sqrt{b^2-4ac}$. *Then a unique solution on the maximal interval of existence* $(t_-(u),t_+(u))$ *is given by*

$$f(t)=\frac{-2c\left(e^{\theta(t_0-t)}-1\right)+\left(\theta\left(e^{\theta(t_0-t)}+1\right)-b\left(e^{\theta(t_0-t)}-1\right)\right)u}{\theta\left(e^{\theta(t_0-t)}+1\right)+b\left(e^{\theta(t_0-t)}-1\right)+2a\left(e^{\theta(t_0-t)}-1\right)u}.$$

Also,

$$\int_t^{t_0}f(s)\,\mathrm{d}s=\frac{1}{a}\log\left(\frac{2\theta e^{\frac{1}{2}(\theta+b)(t_0-t)}}{\theta\left(e^{\theta(t_0-t)}+1\right)+b\left(e^{\theta(t_0-t)}-1\right)+2a\left(e^{\theta(t_0-t)}-1\right)u}\right).$$

Moreover,

- *if $a > 0$, $c \leq 0$ and $u \in \mathbb{R}_-$, then $t_+(u) = +\infty$ and $f(t) \in \mathbb{R}_-$ for all $t \geq t_0$,*
- *if $a < 0$, $c \geq 0$ and $u \in \mathbb{R}_-$, then $t_-(u) = -\infty$ and $f(t) \in \mathbb{R}_-$ for all $t \leq t_0$.*

Proof. For the analytic expression of f and $\int_t^{t_0} f(s)\,\mathrm{d}s$, let $G(t) = f(t_0 - t)$ and see [7], Lemma 10.12.

For the second part, let $a > 0$, $c \leq 0$ and $u \in \mathbb{R}_-$. Then see, for $f(t) = 0$, that

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = c \leq 0.$$

By Lemma 2.8, $f(t) \in \mathbb{R}_-$ for $t \geq t_0$. Then,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} f(t)^2 = f(t)\frac{\mathrm{d}}{\mathrm{d}t} f(t) = af(t)^3 + bf(t)^2 + cf(t) \leq (|b| + |c|)f(t)^2 + |c|\,.$$

An application of Grönwall's inequality yields, for $t \geq t_0$,

$$\begin{aligned} f(t)^2 &= 2|c|(t - t_0) + \int_{t_0}^{t} 2\left(|b| + |c|\right) f(s)^2\,\mathrm{d}s \\ &\leq 2|c|(t - t_0) + \int_{t_0}^{t} |c|(s - t_0)2\left(|b| + |c|\right) e^{2(|b|+|c|)(t-s)}\,\mathrm{d}s < \infty, \end{aligned}$$

and by Theorem A.3 (b), we have $t_+(u) = +\infty$.

Assume now that $a < 0$, $c \geq 0$ and $u \in \mathbb{R}_-$, and let $g(t) = f(t_0 - t)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} g(t) = -ag(t)^2 - bg(t) - c.$$

Analogue to above, we have $g(t) \in \mathbb{R}_-$ for $t \geq 0$, and

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} g(t)^2 \leq (|b| + |c|)g(t)^2 + |c|,$$

and Grönwalls inequality can be applied to conclude that $g(t)^2$ is finite for all $t \geq 0$. Thus, for $t \leq t_0$, we have $f(t) \in \mathbb{R}_-$ and $t_-(u) = -\infty$. □

A.2 The χ^2 Distribution

The noncentral χ^2-distribution appears in the case of certain (extended) CIR processes. It is defined here.

Definition A.7. *The noncentral χ^2-distribution with $\delta > 0$ degrees of freedom and noncentrality parameter $\zeta > 0$ has density*

$$f(x) = \frac{1}{2}e^{-\frac{x+\zeta}{2}} \left(\frac{x}{\zeta}\right)^{\frac{\delta-2}{4}} I_{\frac{\delta-2}{2}}\left(\sqrt{\zeta x}\right), \quad x \geq 0,$$

where $I_\nu(x) = \sum_{j\geq 0} \frac{1}{j!\Gamma(j+\nu+1)} \left(\frac{x}{2}\right)^{2j+\nu}$ is the modified Bessel function of the first kind of order $\nu > -1$.

The characteristic function can be found.

Lemma A.8. *Let X be noncentral χ^2 distributed with δ degrees of freedom and noncentrality parameter ζ. Then, for $u \in \mathbb{C}_-$,*

$$\mathrm{E}\left[e^{uX}\right] = \frac{e^{\frac{\zeta u}{1-2u}}}{(1-2u)^{\frac{\delta}{2}}}.$$

Proof. See [7], Lemma 10.4.

For simulation from the noncentral χ^2 distribution, see [9].

A.3 Complex Numbers

A complex function $f : \mathbb{C} \to \mathbb{R}$ can be written as $f = g + \mathrm{i}h$ for real valued functions g, h. If f, h and g is differentiable, differentiation is straightforward, $f'(z) = g'(z) + \mathrm{i}h'(z)$. Differentiating the exponential function can then be carried out,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}z}e^{f(z)} &= \frac{\mathrm{d}}{\mathrm{d}z}\left(e^{g(z)}\left(\cos h(z) + \mathrm{i}\sin h(z)\right)\right) \\
&= g'(z)e^{f(z)} + e^{g(z)}\left(-h'(z)\sin h(z) + ih'(z)\cos h(z)\right) \\
&= g'(z)e^{f(z)} + ih'(z)e^{g(z)}\left(i\sin h(z) + \cos h(z)\right) \\
&= e^{f(z)}f'(z),
\end{aligned}$$

and one sees that the usual differentiation holds true.

We now find some basic inequalities for complex vectors and matrices. Let A be a d-dimensional real matrix, $A \in [-K, K]^{d^2}$ and let $v \in \mathbb{C}^d$. Then it holds that

$$|v_i| \leq \max\left\{1, |v_i|^2\right\} \leq 1 + \|v\|^2, \tag{A.4}$$

implying that $\sum_{i=1}^{d} |v_i| \le d(1+\|v\|^2)$. Also

$$|v_i| \sum_{j=1}^{d} |v_j| \le \sum_{j=1}^{d} \max\left\{|v_i|^2, |v_j|^2\right\} \le d\|v\|^2, \tag{A.5}$$

and

$$\begin{aligned} \left|v^\top A v\right| &\le K \sum_{k=1}^{d} (|v_k|, \dots, |v_k|)\,(|v_1|, |v_2|, \dots, |v_d|)^\top \\ &\le K \sum_{k=1}^{d}\sum_{l=1}^{d} \max\left\{|v_k|^2, |v_l|^2\right\} \\ &\le K d^2 \|v\|^2, \end{aligned} \tag{A.6}$$

holds.

A useful lemma when dealing with differential equations is the following.

Lemma A.9. *For an interval J and an integer $d \ge 1$, assume $f : J \to \mathbb{C}^d$ is differentiable. Then,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\|f(t)\|^2 = 2\Re\left(\overline{f(t)}^\top \frac{\mathrm{d}}{\mathrm{d}t} f(t)\right).$$

Proof. Let $f(t) = g(t) + \mathrm{i}h(t)$ for real functions $g, h : J \to \mathbb{R}^d$. The result is obtained with straightforward calculations,

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\|f(t)\|^2 \\ &= \sum_{i=1}^{d} \frac{\mathrm{d}}{\mathrm{d}t}\left(\overline{f_i(t)} f_i(t)\right) \\ &= \sum_{i=1}^{d} \Big((g_i'(t) - \mathrm{i}h_i'(t))\,(g_i(t) + \mathrm{i}h_i(t)) + (g_i(t) - \mathrm{i}h_i(t))\,(g_i'(t) + \mathrm{i}h_i'(t)) \Big) \\ &= \sum_{i=1}^{d} \Big(2g_i(t)g_i'(t) + 2h_i(t)h_i'(t) \\ &\qquad + \mathrm{i}\big(g_i'(t)h_i(t) - h_i'(t)g_i(t) + g_i(t)h_i'(t) - h_i(t)g_i'(t)\big)\Big) \\ &= 2\sum_{i=1}^{d} (g_i(t)g_i'(t) + h_i(t)h_i'(t)) \end{aligned}$$

$$= 2\sum_{i=1}^{d} \Re\big((g_i(t) - \mathrm{i}h_i(t)) \, (g_i'(t) + \mathrm{i}h_i'(t)) \big)$$
$$= 2\Re\left(\overline{f(t)}^{\top} \frac{\mathrm{d}}{\mathrm{d}t} f(t)\right).$$

□

We now turn to probability theory and define the complex martingale.

Definition A.10. *An $(\mathcal{F}(t))_{t\in[0,\infty)}$ adapted stochastic process $\mathbf{M} = (M(t))_{t\in[0,\infty)}$, $M(t) \in \mathbb{C}$ is a* martingale *if $\mathrm{E}\,|M(t)| < \infty$ for all $t \in [0,\infty)$, and for all $0 \le s \le t$,*

$$\mathrm{E}\left[M(t) \mid \mathcal{F}(s)\right] = M(s).$$

Lemma A.11. *Let μ be a positive measure on a measure space $(E, \mathcal{F})$. For $f : E \to \mathbb{C}$,*

$$\left|\int f \,\mathrm{d}\mu\right| \le \int |f| \,\mathrm{d}\mu,$$

where $|a + \mathrm{i}b| = \sqrt{a^2 + b^2}$.

Proof. Any complex number $x \in \mathbb{C}$ has representation $x = e^{a+\mathrm{i}b}$, for $a \in [-\infty, \infty)$ and $b \in [0, 2\pi)$. Then

$$|x| = |e^a (\cos b + \mathrm{i} \sin b)| = e^a = e^{-\mathrm{i}b} x.$$

Using this, see that there exists $\theta \in [0, 2\pi)$ such that $\left|\int f \,\mathrm{d}\mu\right| = e^{\mathrm{i}\theta} \int f \,\mathrm{d}\mu$, and with $f = e^{g+\mathrm{i}h}$ for real functions g and h on E, we obtain

$$\left|\int f \,\mathrm{d}\mu\right| = \int e^{g+\mathrm{i}(h+\theta)} \,\mathrm{d}\mu = \int \Re e^{g+\mathrm{i}(h+\theta)} \,\mathrm{d}\mu + \mathrm{i} \underbrace{\int \Im e^{g+\mathrm{i}(h+\theta)} \,\mathrm{d}\mu}_{=0}$$
$$= \int e^g \cos(h+\theta) \,\mathrm{d}\mu \le \int e^g \,\mathrm{d}\mu = \int |f| \,\mathrm{d}\mu.$$

□

Remark A.12. Using lemma A.11 one sees that

$$\left|\mathrm{E}\left[e^{\mathrm{i}X}\right]\right| \le \mathrm{E}\left|e^{\mathrm{i}X}\right| = 1,$$

i.e. that any characteristic function is bounded by 1. ◇

Lemma A.13. *Let* $\mathbf{M}$ *be a complex local martingale. If for all* $t \in \mathbb{R}_+$ *there exists* $Y(t)$*, such that* $\mathrm{E}\,|Y(t)| < \infty$ *and*

$$\forall s \in [0,t] : |M(s)| \leq Y(t),$$

then $\mathbf{M}$ *is a complex martingale.*

Proof. By the definition of local martingales, there exists a sequence of stopping times (τ_n) such that $\mathbf{M}^{\tau_n} = (M(\tau_n \wedge t))_{t \in \mathbb{R}_+}$ is a martingale. Let $A \in \mathcal{F}(s)$, and see that for $t \geq s$,

$$\begin{aligned}\mathrm{E}\left[1_A M^{\tau_n}(s)\right] &= \mathrm{E}\left[1_A \,\mathrm{E}\left(M^{\tau_n}(t) \mid \mathcal{F}(s)\right)\right] \\ &= \mathrm{E}\left[\mathrm{E}\left(1_A M^{\tau_n}(t) \mid \mathcal{F}(s)\right)\right] = \mathrm{E}\left[1_A M^{\tau_n}(t)\right].\end{aligned}$$

We have $M^{\tau_n}(u) \overset{\text{a.s.}}{\to} M(u)$ for $n \to \infty$ for all $u \geq 0$, and by dominated convergence, since $M(s)$ and $M(t)$ are bounded by $Y(t)$,

$$\mathrm{E}\left[1_A M(s)\right] = \mathrm{E}\left[1_A M(t)\right], \quad A \in \mathcal{F}(s).$$

By the definition of conditional expectations, $M(s) = \mathrm{E}\left[M(t) \mid \mathcal{F}(s)\right]$. □

Remark A.14. The conditions of lemma A.13 are obviously satisfied if the complex local martingale $\mathbf{M}$ is uniformly bounded. ◇

A.4 Quadratic Variation

The following are results stated in [11].

$\langle \mathbf{M}, \mathbf{M} \rangle = \langle \mathbf{M} \rangle$ is the quadratic variation of $\mathbf{M}$, and $\langle \mathbf{M}, \mathbf{N} \rangle$ denotes the quadratic variation between $\mathbf{M}$ and $\mathbf{N}$.

For any locally L^2-bounded martingales $\mathbf{M}$ and $\mathbf{N}$, and any predictable and locally bounded functions $\mathbf{H}$ and $\mathbf{K}$,

$$\left\langle \int_0^t H(s)\,\mathrm{d}M(s), \int_0^t K(s)\,\mathrm{d}N(s) \right\rangle = \int_0^t H(s)K(s)\,\mathrm{d}\left\langle M(s), N(s)\right\rangle,$$

and in particular

$$\left\langle \int_0^t H(s)\,\mathrm{d}M(s) \right\rangle = \int_0^t H^2(s)\,\mathrm{d}\left\langle M(s)\right\rangle.$$

For a continuous semimartingale $\mathbf{X} = \mathbf{V} + \mathbf{M}$ where $\mathbf{M}$ is a locally L^2-bounded càdlàg martingale and $\mathbf{V}$ is an adapted càdlàg process of finite variation, the quadratic variation is defined as

$$\langle \mathbf{X} \rangle = \langle \mathbf{M} \rangle .$$

Thus, for a 1-dimensional diffusion process

$$\mathrm{d}X(t) = \alpha(X(t))\,\mathrm{d}t + \sigma(X(T))\,\mathrm{d}W(t), \tag{A.7}$$

the quadratic variation is the quadratic variation of the diffusion term,

$$\langle X(t) \rangle = \left\langle \int_0^t \sigma(X(s))\,\mathrm{d}W(s) \right\rangle = \int_0^t \sigma^2(X(s))\,\mathrm{d}\,\langle W(s) \rangle .$$

For a d-dimensional Brownian motion $\mathbf{W}$, we have

$$\langle W_i(t), W_j(t) \rangle = \delta_{ij} t,$$

where δ_{ij} is the Kronecker delta. Thus, for a general d-dimensional diffusion process $\mathbf{X}$ from (A.7), we get the classic result

$$\mathrm{d}\,\langle X_i(t), X_j(t) \rangle = \sum_{k=1}^{d} \sigma_{ik}(X(t))\sigma_{jk}(X(t))\,\mathrm{d}t,$$

which in the 1-dimensional case reduces to $\mathrm{d}\,\langle X(t) \rangle = \sigma^2(X(t))\,\mathrm{d}t$.