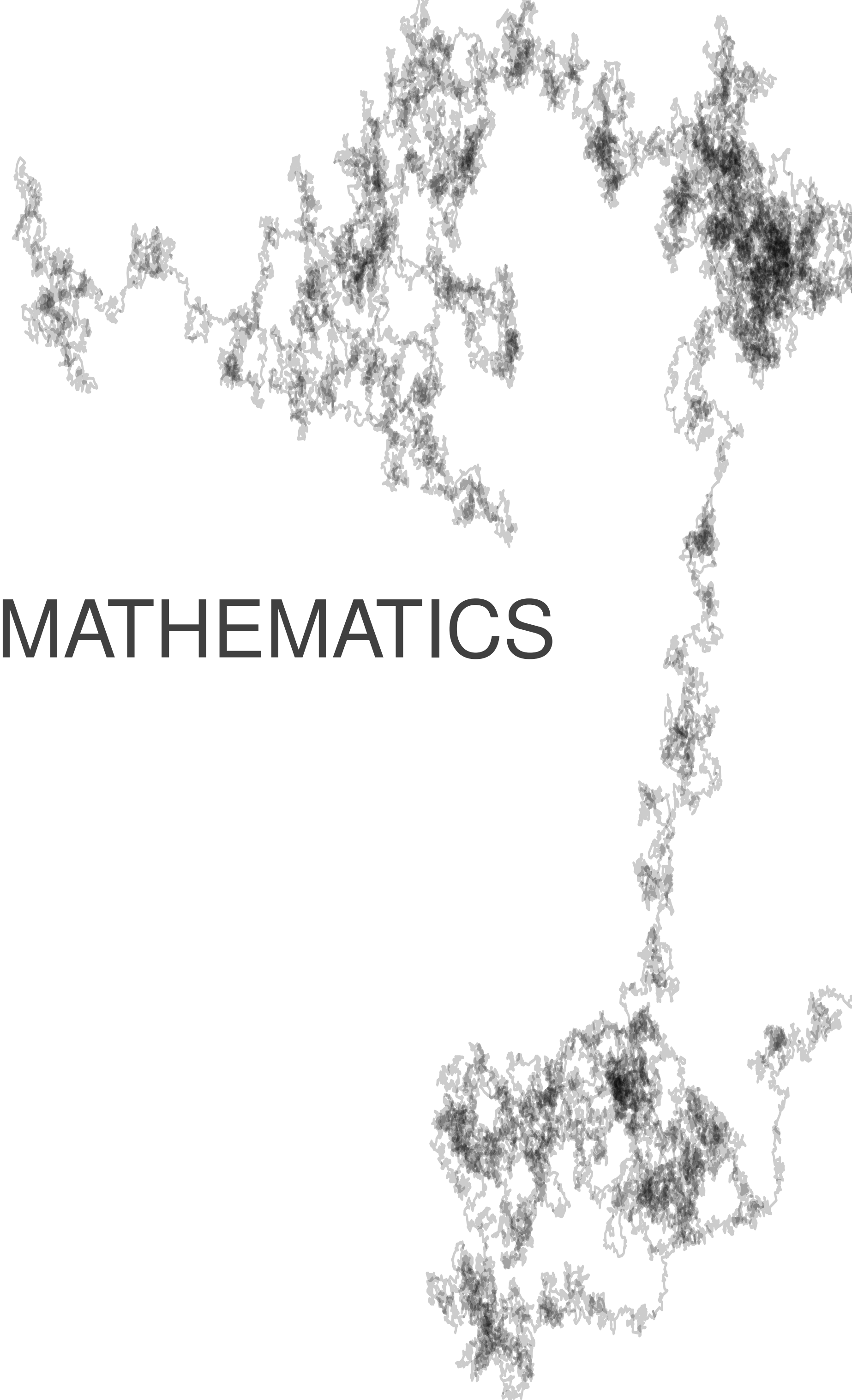
THORSTEN SCHMIDT



FINANCIAL MATHEMATICS

UNIVERSITÄT FREIBURG

Contents

I
Discrete time

Historically, financial mathematics originated in continuous time - like in the famous works from Black & Scholes (1973) and Bachelier (1900) and the main driving force was a Brownian motion. It seems very plausible, that a large number of traders who act independently can be approximated by a Gaussian distribution through the central limit theorem, such that this is a very appealing setup.

However, this requires the full power of stochastic integration and the technical details are quite subtle. It is remarkable, that the main concepts of financial markets, like absence of arbitrage, the first and second fundamental theorem can be proven in discrete time without the need to dive into the technicalities while providing similarly deep insights. I therefore believe, it is a good start to spend some time on discrete time.

1 A discrete-time financial market

An excellent introduction to the field is Föllmer & Schied (2016). We follow this book for the introduction and directly start in a multi-period financial market. The advantage of this approach - as we will soon see - is that a multi-period market essentially can be reduced to a one-period market.

To this end we fix a probability space $(\Omega, \mathscr{F}, P)$. A financial market consists of one primary risk-free asset S^0 which is assumed to be strictly positive. Furthermore, we have d risky assets $S = (S^1, \dots, S^d)$ which are assumed to be non-negative. All assets are described as stochastic processes on the time interval $\mathbb{T} = \{0, \dots, T\}$.

The information flow is described by the filtration $\mathbb{F} = (\mathscr{F}_t)_{t \in \mathbb{T}}$. We denote by $\bar{S} = (S^0, S)$ the $d+1$-dimensional stochastic process including the risk-free account. We assume that $\bar{S}$ is adapted to the filtration $\mathbb{F}$.

Definition 1. A *trading strategy* $\bar{H}$ is a predictable, $d+1$-dimensional stochastic process. The trading strategy is *self-financing*, if

$$\bar{H}_t \cdot \bar{S}_t = \bar{H}_{t+1} \cdot \bar{S}_t, \qquad t = 1, \dots, T-1.$$

Intuitively, a self-financing trading strategy does not require external funds while rebalancing at time t, neither does it produce a consumable profit.

Let us denote the increments of S by

$$\Delta S_t = S_t - S_{t-1}, \qquad t = 1, \dots, T.$$

It is a remarkable result, that for a self-financing trading strategy, the position at time t can be decomposed in the initial value plus the gains from trading. The gains itself can be written as a (discrete) stochastic integral.

Lemma 2. *For a self-financing trading strategy $\bar{H}$, we obtain that*

$$\bar{H}_t \cdot \bar{S}_t = \bar{H}_1 \cdot \bar{S}_0 + \sum_{k=1}^{t} \bar{H}_k \cdot \Delta \bar{S}_k, \qquad 1 \leq t \leq T.$$

Proof. This follows in two steps:

$$\begin{aligned}
\bar{H}_t \cdot \bar{S}_t &= \bar{H}_t \cdot \bar{S}_t + \bar{H}_{t-1} \cdot \bar{S}_{t-1} - \bar{H}_{t-1} \cdot \bar{S}_{t-1} \\
&= \bar{H}_t \cdot \bar{S}_t + \bar{H}_{t-1} \cdot \bar{S}_{t-1} - \bar{H}_t \cdot \bar{S}_{t-1} \\
&= \bar{H}_t \cdot (\bar{S}_t - \bar{S}_{t-1}) + \bar{H}_{t-1} \cdot \bar{S}_{t-1} \\
&= \sum_{k=2}^{t} \bar{H}_k \cdot (\bar{S}_k - \bar{S}_{k-1}) + \bar{H}_1 \cdot \bar{S}_1
\end{aligned}$$

where we used that $\bar{H}$ is self-financing. For the last time step we obtain,

$$\begin{aligned}
\bar{H}_1 \cdot \bar{S}_1 &= \bar{H}_1 \bar{S}_1 + \bar{H}_1 \cdot \bar{S}_0 - \bar{H}_1 \cdot \bar{S}_0 \\
&= \bar{H}_1 (\bar{S}_1 - \bar{S}_0) + \bar{H}_1 \cdot \bar{S}_0
\end{aligned}$$

and the claim follows. □

Example 3 (Bank account)**.** A typical example for S^0 is the bank account. The bank account starts at $S_0^0 = 1$ and offers the interest rate r_t from $t-1$ to tt. Note that r_t is of course already known at $t-1$ and hence predictable. Hence,

$$S_t^0 = \prod_{s=1}^{t} (1 + r_s).$$

We always require $r_t > -1$. But often one additionally assumes that $r_t \geq 0$.

1.1 Moving to discounted quantities

An important step - economically, and mathematically - is to move to discounted quantities. While this simplifies that setup drastically, it also has a number of subtle consequences (in particular in continuous time).

We introduce the *discounted price process*

$$X_t^i := \frac{S_t^i}{S_t^0}, \quad t = 0, \ldots, T, \quad i = 0, \ldots, d.$$

Note that $X^0 \equiv 1$ and in particular $\Delta X_t^0 = 0$. As previously, we use the notation $\bar{H} = (H^0, H)$.

Definition 4. The *discounted value process* $V = V^{\bar{H}}$ of the trading strategy $\bar{H}$ is given by

$$V_t := \bar{H}_t \cdot \bar{X}_t, \quad t = 1, \ldots, T$$

with $V_0 := \bar{H}_1 \cdot \bar{X}_0$. The *discounted gains process* $G = G^{\bar{H}}$ is

$$G_t := \sum_{k=1}^{t} H_k \cdot \Delta X_k, \quad t = 1, \ldots, T$$

with $G_0 = 0$.

Of course,

$$G_t = \sum_{k=1}^{t} \bar{H}_k \cdot \Delta \bar{X}_k,$$

which explains why we can switch from $\bar{H}$ to H on discounted quantities.

Proposition 5. *Consider the trading strategy $\bar{H}$. Then the following are equivalent:*

(i) $\bar{H}$ *is self-financing,*

(ii) $\bar{H}_t \cdot \bar{X}_t = \bar{H}_{t+1} \cdot \bar{X}_t, \quad t = 1, \ldots, T-1,$

(iii) $V_t = V_0 + G_t$ *for* $0 \leqslant t \leqslant T$.

Proof. By definition, self-financing is equivalent to

$$\begin{aligned} & \bar{H}_t \cdot \bar{S}_t = \bar{H}_{t+1} \cdot \bar{S}_t & & t = 0, \ldots, T-1, \\ \Leftrightarrow \quad & \bar{H}_t \cdot \frac{\bar{S}_t}{S_t^0} = \bar{H}_{t+1} \cdot \frac{\bar{S}_t}{S_t^0}, & & t = 0, \ldots, T-1, \end{aligned}$$

since $S^0 > 0$. This yields equivalence of (i) and (ii). For the next step we compute the increments of the value process. By (ii),

$$V_t - V_{t-1} = \bar{H}_{t+1} \cdot \bar{X}_{t+1} - \bar{H}_t \cdot \bar{X}_t = \bar{H}_{t+1} \cdot (\bar{X}_{t+1} - \bar{X}_t) = H_{t+1} \cdot (X_{t+1} - X_t).$$

Hence,

$$V_t - V_0 = \sum_{s=1}^{t} H_s \cdot (X_s - X_{s-1}), \quad t = 1, \ldots, T$$

and the conclusion follows. □

Remark 6 (Trading strategies). If we start with a d-dimensional trading strategy H, we can determine the associated self-financing strategy $\bar{H}$ as follows: choose H^0 according to

$$H_{t+1}^0 - H_t^0 = -(H_{t+1} - H_t) \cdot X_t, \qquad t = 0, \ldots, T-1,$$

and $H_1^0 = V_0 - H_1 \cdot X_0$. In the following, if we speak of a self-financing trading strategy H, we mean equivalently this associated strategy $\bar{H}$.

1.2 *Arbitrage and martingales*

The central concept for our analysis of financial markets is the concepts of arbitrage.

Definition 7. An *arbitrage* is a self-financing trading strategy H, such that the associated discounted value process V satisfies

(i) $V_0 \leqslant 0$,

(ii) $V_T \geqslant 0$, and

(iii) $P(V_T > 0) > 0$.

If there are no arbitrages on a financial market, we call it arbitrage-free.

If we recall our introduction, we realize that a financial market consists of the triplet $(\mathbb{F}, S, P)$. It can be easily shown that the above conditions are equivalent to the same conditions for the undiscounted value process $(\bar{H}_t \cdot \bar{S}_t)$.

Proposition 8. *A financial market is free of arbitrage if and only if every one-period financial market* (S_t, S_{t+1})*,* $t = 0, \dots, T-1$ *is free of arbitrage.*

Proof. The idea of the proof is to show the equivalence of the negotiations: there exists an arbitrage if and only if for a $t \in \{1, \dots, T\}$ there exists a $\mathscr{F}_{t-1}$-measurable random $\xi \in \mathbb{R}^d$, such that $\xi \cdot \Delta X_t \geq 0$ P-a.s. and $P(\xi \cdot \Delta X_t > 0) > 0$.

For the first part we start with an arbitrage and show that there exists a single period with an arbitrage: let $\bar{H}$ be an arbitrage with value process V. Let

$$t := \min\left\{s \in \{1, \dots, T\} : V_s \geqslant 0 \text{ und } P(V_s > 0) > 0\right\}$$

be a deterministic time with the convention that $\min \varnothing = \infty$. Since $\bar{H}$ is an arbitrage, we obtain that $t \leqslant T$. There are two possibilities to be taken into account: either $V_{t-1} = 0$, or $P(V_{t-1} < 0) > 0$. In the first case, we are ready, since

$$H_t \cdot (X_t - X_{t-1}) = V_t - V_{t-1} = V_t,$$

so $\xi = H_t$ does the job.

For the second case we choose $\xi := H_t \mathbb{1}_{\{V_{t-1}<0\}}$. Then ξ is $\mathscr{F}_{t-1}$-measurable and

$$\xi \cdot (X_t - X_{t-1}) = (V_t - V_{t-1})\mathbb{1}_{\{V_{t-1}<0\}} \geqslant -V_{t-1}\mathbb{1}_{\{V_{t-1}<0\}} \geqslant 0.$$

Now observe that the r.h.s. is positive with positive probability and the first part is finished.

The other direction is straightforward: set $H_s = \xi \mathbb{1}_{\{s=t\}}$ and construct the associated self-financing trading strategy $\bar{H}$, which is an arbitrage. □

Remark 9. It is interesting to see that the property to deduce absence of arbitrage from one time period only breaks down if one allows for two filtrations, see Kabanov & Stricker (2006). A fundamental theorem for two markets with two filtrations is still an open research question.

Definition 10. A probability measure Q on $(\Omega, \mathscr{F}_T)$ is called *martingale measure*, if X is a Q-martingale.

We note that a martingale measure refers to the *discounted* price process being a martingale. It is a bit surprising that an artificial probability measure takes up such a prominent role in no-arbitrage theory. We will see later, why. Recall that Q is called absolutely continuous with respect to P ($Q \ll P$) if $P(F) = 0$ for an $F \in \mathscr{F}_T$ implies that $Q(F) = 0$. Q is called equivalent to P ($P \sim Q$), if $Q \ll P$ and $P \ll Q$.

1.3 *Martingale measures*

We denote the set of *equivalent martingale measures* by $\mathscr{M}_e(\mathbb{F}) = \mathscr{M}_e$. If we assume that the initial filtration is trivial, the following result can already be obtained. It is remarkable that the integrability condition of a martingale can be obtained from no-arbitrage (equivalently the existence of a martingale measure as we will see later) and a substantially weakened integrability.

Satz 11. *Let $\mathscr{F}_0 = \{\emptyset, \Omega\}$. Then, the following are equivalent:*

(i) $Q \in \mathscr{M}_E(\mathbb{F})$

(ii) For any bounded self-financing trading strategy $\bar{H}$, $V^{\bar{H}}$ is a Q-martingale.

(iii) For any self-financing trading strategy $\bar{H}$ such that $E_Q[(V_T^{\bar{H}})^-] < \infty$ $V^{\bar{H}}$ is a Q- martingale.

(iv) For any self-financing trading strategy $\bar{H}$ with $V_T = V_T^{\bar{H}} \geqslant 0$ it holds that

$$V_0 = E_Q[V_T].$$

Proof. i) ⇒ ii): we start with a bounded self-financing strategy $\bar{H}$, i.e. $|H_t^i| \leqslant c$, for $i = 0, \dots, d$, $t = 0, \dots, T$. Then, integrability of $V = V^{\bar{H}}$ follows from the integrability of X since

$$|V_t| \leqslant |V_0| + \sum_{k=1}^{t} c \cdot \sum_{i=1}^{d} \Big(|X_k^i| + |X_{k-1}^i| \Big).$$

Moreover, we have that

$$\begin{aligned} E_Q[V_t|\mathscr{F}_{t-1}] &= E_Q[V_{t-1} + \bar{H}_t \cdot (\bar{X}_t - \bar{X}_{t-1})|\mathscr{F}_{t-1}] \\ &= V_{t-1} + \bar{H}_t \cdot (E_Q[\bar{X}_t|\mathscr{F}_{t-1}] - \bar{X}_{t-1}) = V_{t-1}. \end{aligned}$$

ii) ⇒ iii): We want to show the martingale property under a minimal integrability assumption. We start with the assumption that $E_Q[V_T^-] < \infty$. Then $E_Q[V_T]$ and, similarly, $E_Q[V_T|\mathscr{F}_{T-1}]$ is well-defined (though possibly not finite).

Consider $a > 0$. Then,

$$\begin{aligned} E_Q[V_T|\mathscr{F}_{T-1}]\mathbb{1}_{\{|\bar{H}_T| \leqslant a\}} &= E_Q[\bar{H}_T \cdot \bar{X}_T \mathbb{1}_{\{|\bar{H}_T| \leqslant a\}}|\mathscr{F}_{T-1}] \\ &= E_Q[\bar{H}_T(\bar{X}_T - \bar{X}_{T-1})\mathbb{1}_{\{|\bar{H}_T| \leqslant a\}}|\mathscr{F}_{T-1}] + V_{T-1}\mathbb{1}_{\{|\bar{H}_T| \leqslant a\}} \\ &= V_{T-1}\mathbb{1}_{\{|\bar{H}_T| \leqslant a\}}. \end{aligned} \tag{12}$$

Now with $a \to \infty$, $\{|\bar{H}_T| \leqslant a\} \to \Omega$, since $\bar{H}_T$ is a $\mathbb{R}^{d+1}$-value random variable. Hence, $E_Q[V_T|\mathscr{F}_{T-1}] = V_{T-1}$ Q-a.s. The next step is to show $E_Q[V_{T-1}^-] < \infty$ with Jensens' inequality:

this follows since

$$E_Q[V_{T-1}^-] = E_Q[E_Q[V_T|\mathscr{F}_{T-1}]^-] \leqslant E_Q[E_Q[V_T^-|\mathscr{F}_{T-1}]] = E_Q[V_T^-] < \infty.$$

Proceeding iteratively we obtain $E_Q[V_t^-] < \infty$ as well as $E_Q[V_t|\ \mathscr{F}_{t-1}] = V_{t-1}$.

Then,

$$V_0 = E_Q[V_t].$$

Since $E_Q[V_t^-]$, the expectation is well-defined. Moreover, $V_0 \in \mathbb{R}$, so $E_Q[V_t^+] < \infty$, and hence $E_Q[|V_t|] < \infty$. We obtain integrability and hence V is a Q-martingale.

iii) $\Rightarrow$ iv): clear.

iv) $\Rightarrow$ i): First, we show integrability of X_t^i. This can be achieved using $X_t^i = X_0^i + \sum_{s=1}^t \Delta X_s^i$. We choose accordingly $H_s^j = \mathbb{1}_{\{s\leq t\}}\mathbb{1}_{\{j=i\}}$, such that $G_T = \sum_{s=1}^t \Delta X_s^i$ and, (by Proposition 1.1 (iii)), $V_0 = X_0^i$. Moreover, $V_T = X_t^i \geq 0$. Then we can apply (iv), such that

$$\infty > X_0^i = V_0 = E_Q[V_T] = E_Q[X_t^i] = E_Q[|X_t^i|]. \tag{13}$$

The next goal is to show

$$E_Q[X_{t+1}^i 1_F] = E_Q[X_t^i 1_F] \qquad \forall F \in \mathscr{F}_t$$

and for $1 \leq t \leq T$. For this, we aim at a similar strategy, but $H_s = \mathbb{1}_{\{s\leq t\}}\mathbb{1}_F - \mathbb{1}_{\{s\leq t-1\}}\mathbb{1}_F$ is not possible since $\mathbb{1}_F$ is only $\mathscr{F}_t$-measurable. Instead we search for a strategy which achieves $V_T = X_t^i\mathbb{1}_F + X_{t+1}^i\mathbb{1}_{F^C} \geq 0$. Note that $V_T = X_t^i + \Delta X_{t+1}^i\mathbb{1}_{F^C}$ which can be achieved by the trading strategy

$$H_s^i = \mathbb{1}_{\{s\leq t\}} + \mathbb{1}_{F^C}1_{\{s=t+1\}},$$

and $H^j = 0$ for $j \neq i$ together with $V_0 = X_0^i$, as above. Again, by (iv),

$$X_0^i = V_0 = E_Q[V_T] = E_Q[\mathbb{1}_F X_t^i + \mathbb{1}_{F^C}X_{t+1}^i].$$

Together with (13),

$$E_Q[X_{t+1}^i] = X_0^i = E_Q[\mathbb{1}_F X_t^i + \mathbb{1}_{F^C}X_{t+1}^i],$$

hence $E_Q[1_F X_t^i] = E_Q[1_F X_{t+1}^i]$, and the claim follows. □

2 *The fundamental theorem*

The following fundamental theorem relates absence of arbitrage with a very simple criterion - the existence of a martingale measure. This measure can be used for pricing in a very simple manner, which explains the immense success of this approach.

Theorem 14 (FTAP). *A financial market is free of arbitrage, if and only if $\mathscr{M}_e(\mathbb{F}) \neq \varnothing$. In this case there exists a $Q \in \mathscr{M}_e$ with bounded density dQ/dP.*

This theorem is proved in several steps. The first step, which is the most important step in applications, is surprisingly easy.

Proposition 15. *If $\mathscr{M}_e \neq \varnothing$, then there is no arbitrage.*

Proof. We use proposition 11 (iii). Assume that H is an arbitrage with discounted value process V and chose a $Q \in \mathscr{M}_e$.

Note that with $V_0 \leqslant 0$ P-a.s. it holds that $V_0 \leqslant 0$ Q-a.s. In the same manner, we obtain that $V_T \geq 0$ Q-a.s. and hence $E_Q[V_T^-] = 0 < \infty$.

Since H is an arbitrage, $P(V_T > 0) > 0$, and hence $Q(V_T > 0) > 0$. This implies $E_Q[V_T] > 0$. Since Proposition 11 (iii) yields that V is a Q-martingale, i.e.

$$V_0 = E_Q[V_T] > 0,$$

we obtain a contradiction to $V_0 \leqslant 0$. □

2.1 Examples

Let us visit shortly some examples which illustrate the importance to the application of Proposition 15. Note that if we add a Q-martingale as additional coordinate to the price process X the market remains arbitrage-free. It is therefore natural to use the *risk-neutral pricing rule* for pricing additional contingent claims.

Consider n $\mathscr{F}_T$-measurable contingent claim with (discounted) payoff $C_T \geq 0$ and let

$$X_t^{d+1} = E_Q[C_T | \mathscr{F}_t], \qquad t \in \mathbb{T}.$$

Then, the extended market $\widetilde{X} = (X^1, \ldots, X^{d+1})$ is free of arbitrage.

Example 16 (Black-Scholes Model)**.** The famous Black-Scholes modelgives the stock price under P as a geometric Brownian motion, precisely:

$$dS_t = S_t \mu dt + S_t \sigma dW_t$$

with a Brownian motion W and initial value $\mathbb{R} \ni S_0 > 0$. The unique strong solution of this SDE is given by

$$S_t = S_0 \exp\Big((\mu - \sigma^2/2)t + \sigma W_t \Big), \qquad t \geq 0.$$

A typical derivative is an *European call* which offers the pay-off

$$(S_T - K)^+.$$

If we additionally assume that the Bank account is e^{rt}, then the Girsanov theorem shows that the measure $dQ = L_T dP$ with

$$dL_t = -L_t \lambda dW_t,$$

and $\lambda = {r-\mu}/{\sigma}$ is a martingale measure. Under Q, $\widetilde{W}_t = W_t + \lambda t,\ t \geq 0$ is a Brownian motion and so

$$dS_t = S_t r dt + S_t \sigma d\widetilde{W}_t,$$

and – of course – $dX_t = S_t \sigma d\widetilde{W}_t$. Let us compute shortly the price of the call option,

$$E_Q[C_T] = S_0 \Phi(d_1) - K e^{-rT} \Phi(d_2), \tag{17}$$

$$d_{1/2} = \frac{\log \frac{S_0}{K e^{-rT}} \pm \frac{\sigma^2 T}{2}}{\sigma \sqrt{T}} \tag{18}$$

While this model is in continuous time, the discrete analogue $S = (S_{t_i})_{i=0,\dots,n}$ with $t_i = \delta i$ is of course in discrete time. It is remarkable that the continuous time model has only one equivalent martingale measure, while the discrete time model has many.

Example 19 (The binomial model)**.** An example in discrete time which is very illustrative, is the Binomial model (also called the Cox-Ross-Rubinstein model). Here, we assume that

$$S_t = S_0 \prod_{i=1}^{t} \xi_i, \qquad t = 1, \dots, T$$

with $\xi_i \in \{1+u, 1+d\}$, $-1 \le d \le u$. The bank account is $S_t^0 = (1+r)^t$, $r > -1$. This model has no-arbitrage if and only if

$$d < r < u.$$

Indeed, for a martingale measure we need

$$1 = E_Q[(1+r)^{-1} \cdot \xi_t | \mathscr{F}_{t-1}] = (1+r)^{-1}\Big(q_t \cdot (1+u) + (1-q_t) \cdot (1+d)\Big)$$

which is equivalent to

$$q_t = \frac{r-d}{u-d}.$$

It is remarkable to realise that under Q, (ξ_t) are i.i.d., while this assumption is of course not necessary under P. We also may see that the martingale measure is actually used to characterise the convex hull of $(1+d, 1+u)$ and so a geometric interpretation of existence of a martingale measure is the property that the return of the bank account $1+r$ lies in the interior of the convex span of the return of the stock.

It is also interesting to see that for a contingent claim with price C_{t-1} at $t-1$ and values $\{C_t^+, C_t^-\}$ at t, the replicating strategy is determined by

$$H_t = \frac{C_t^+ - C_t^-}{S_{t-1} \cdot (u-d)}.$$

Note that this strategy is unique !

2.2 *Complete markets*

We call a contingent claim C_T *replicable*, if there exists a self-financing trading strategy, such that $V_T = C_T$. We call an arbitrage-free market *complete*, if every contingent claim is replicable.

Proposition 20. *Assume that $\mathscr{F}_0 = \{\emptyset, \mathscr{F}\}$. Then, an arbitrage-free market is complete, if and only if*

$$\mathscr{M}_e = \{Q\}.$$

Proof. Assume that the market is complete. We first note that a replicable claim has a unique price: we first note that by Proposition 11 the value processes of replicable trading strategies are martingales, since $V_T = C_T \ge 0$. Hence, $E_Q[V_T] = V_0 < \infty$ for all $Q \in \mathscr{M}_e$. This implies that the price of a *replicable* contingent claim C_T is unique, since for any replicable trading strategy

$$V_0 = E_Q[V_T] = E_Q[C_T]$$

and the right hand side does not depend on V_T.

Now we show that for $Q, Q' \in \mathscr{M}_e$ it holds that $Q = Q'$. Indeed, consider $C_T = \mathbb{1}_F$ for any $F \in \mathscr{F}_T$. Then

$$Q(F) = E_Q[\mathbb{1}_F] = E_{Q'}[\mathbb{1}_F] = Q'(F).$$

We postpone the second assertion - since it needs a little bit more technique. □

2.3 *The second part of the proof of the FTAP*

Now we return to the fundamental theorem. The main goal is to show the existence of an equivalent martingale measure. Since this relates to linear functionals, our proof will reside on the Hahn-Banach theorem

Since we already know that it is sufficient to consider one period only, we study the gains which can be obtained in a self-financing manner with zero initial investment. Fix $0 < t \leq T$. Then these gains are given by

$$K = \{H \cdot \Delta X_t : \ H \in L^0(P, \mathscr{F}_{t-1}, \mathbb{R}^d)\}.$$

Then, absence of arbitrage (NA) is equivalent to

$$K \cap L^0_+(\mathscr{F}_t, P, \mathbb{R}) = \{0\}.$$

Since, t is fixed we shortly write $L^0_+(\mathscr{F}_t, P, \mathbb{R}) = L^0_+$. One central and very useful observation in the following theorem is that the set K can be replaced by the claims which can be *super-replicated* which is given by the set $K - L^0_+$.

Satz 21. *Consider the one period-market from $t-1$ to t. Then, the following are equivalent:*

(i) $K \cap L^0_+ = \{0\}$,

(ii) $(K - L^0_+) \cap L^0_+ = \{0\}$,

(iii) there exists an equivalent martingale measure with bounded density,

(iv) there exists an equivalent martingale measure.

Proof. We show (iv) $\Rightarrow$ (i) $\Leftrightarrow$ (ii) and (iii) $\Rightarrow$ (iv). The part (ii) $\Rightarrow$ (iii) is the most difficult part and will be treated separately.

(iv) $\Rightarrow$ (i): we aim at a contradiction. Consider $Q \in \mathscr{M}_e$ and assume there exists $H \in L^0(P, \mathscr{F}_{t-1}, \mathbb{R}^d)$ such that $H \cdot (X_t - X_{t-1}) \geqslant 0$ while $P(H \cdot (X_t - X_{t-1}) > 0) > 0$. Note that this implies $Q(H \cdot (X_t - X_{t-1})) > 0$.

This cannot hold if H is bounded. We therefore consider $H^c := H \mathbb{1}_{\{H| \leqslant c\}}$ for $c > 0$. Since $\{H| \leqslant c\} \uparrow \Omega$ for $c \to \infty$, we can exploit σ-continuity of the probability measure Q. Hence $Q(H^{c^*}(X_t - X_{t-1}) > 0) \to Q(H(X_t - X_{t-1}) > 0) > 0$. Then there exists a c^* such that $Q(H^{c^*}(X_t - X_{t-1}) > 0) > 0$.

But,

$$E^Q[H^{c^*} \cdot (X_t - X_{t-1}) | \mathscr{F}_{t-1}] = H^{c^*} E^Q[X_t - X_{t-1} | \mathscr{F}_{t-1}] = 0,$$

which contradicts with $H^{c^*} \cdot (X_t - X_{t-1}) \geqslant 0$ and $Q(H^{c^*}(X_t - X_{t-1}) > 0) > 0$.

(i) $\Rightarrow$ (ii): Consider $Z \in (K - L^0_+) \cap L^0_+$. With a $\mathscr{F}_{t-1}$-measurable H and $U \in L^0_+$,

$$Z = H \cdot (X_t - X_{t-1}) - U \geqslant 0.$$

Hence $H \cdot (X_t - X_{t-1}) \geqslant U \geqslant 0$, such that $H \cdot (X_t - X_{t-1}) \in K \cap L^0_+$. By (i),t $H \cdot (X_t - X_{t-1}) = 0$, also $U = 0$ and so $Z = 0$.

The missing (ii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (iv) are immediate. □

For the remaining step (ii) $\Rightarrow$ (iii) will be achieved through a number of steps:

(i) We first show that integrability can always achieved through a change of measure (Lemma 22). This allows us to consider the convex cone C as a subset of L^1, see (23).

(ii) Then we apply the Hahn-Banach theorem: consider $F \in L^1_+$ together with $B = \{F\}$ and C. The difficulty is to show that C is closed (which we postpone for a moment). This gives us a strictly separating continuous functional.

(iii) Using duality we obtain a density (Lemma 28 (i)). Since C is countably convex we can even select a positive density, hence an equivalent martingale measure (Lemma 28 (ii)).

(iv)

(v)

The following step shows that we can always achieve integrability by an equivalent measure change with a bounded density.

Lemma 22. *There exists* $\widetilde{P} \sim P$ *such that* $\widetilde{E}[|X_t|] < \infty$ *and* $\widetilde{E}[|X_{t-1}|] < \infty$.

Proof. Consider $c > 0$, let

$$Z := \frac{c}{1 + |X_t| + |X_{t-1}|} \leqslant c$$

and $d\widetilde{P} = Z dP$. Obviously, $\widetilde{E}[|X_t|] < \infty$ and $\widetilde{E}[|X_{t-1}|] < \infty$. □

Since (ii) only depends on the nullsets of P, it holds if and only it if holds with respect to $\widetilde{P}$ when $\widetilde{P} \sim P$. The same holds for boundedness of the density and we therefore can assume without loss of generality that $E[|X_t|] < \infty$ und $E[|X_{t-1}|] < \infty$.

Define the convex cone

$$C = (K - L^0_+) \cap L^1. \tag{23}$$

Example 24 (C not closed)**.** It is remarkable that NA actually implies closedness of C: indeed, the following example (where an arbitrage exists) shows that C is not always closed: consider $\Omega = [0,1]$, the Borel σ-field $\mathscr{F}_1 =$, trivial $\mathscr{F}_0 = \{\varnothing, \Omega\}$, and $\Delta X(\omega) = \omega$ (clearly, we have arbitrages here).

Note that C is a true subset of L^1, since for $F \geqslant 1$ (for all $\omega \in \Omega$), $F \notin C$. Define

$$F_n = (F^+ \wedge n)\mathbb{1}_{[1/n,1]} - F^- \quad \text{for } F \in L^1,$$

such that $F_n \xrightarrow{L^1} F$. Moreover, $F_n \in C$: note that

$$(F^+ \wedge n)\mathbb{1}_{[1/n,1]} \leqslant \begin{cases} n & \omega \in [1/n, 1], \\ 0 & \omega \in [0, 1/n). \end{cases}$$

Hence,

$$(F^+ \wedge n)\mathbb{1}_{[1/n,1]} \leqslant n \cdot n\Delta X = n^2 \Delta X,$$

such that $(F^+ \wedge n)\mathbb{1}_{[1/n,1]} = n^2\Delta X - U \in C$.

2.4 *The Hahn-Banach theorem*

Recall that a topological vector space is a vector space with a topology, such that addition and scalar multiplication are continuous. It is called *locally convex* if its topology is generated by convex sets. Note that any Banach space is locally convex, since we can use as base the ε balls for each element. However, the space of all random variable L^0 with the topology of convergence in probability is not convex if $(\Omega, \mathscr{F}, P)$ has no atoms.

Theorem 25 (Hahn-Banach). *Consider two non-empty sets B and C of the locally convex space E and assume that*

(i) $B \cap C = \varnothing$,

(ii) B, C *are convex,*

(iii) B *is compact and* C *is closed.*

Then there exists a continuous linear functional $\ell : E \to \mathbb{R}$, *such that*

$$\sup_{x \in C} \ell(x) < \inf_{x \in B} \ell(x).$$

By a duality argument the linear functional can be represented by a Z satisfying the following property (27). We first show that this is sufficient to obtain a martingale measure.

Lemma 26. *Let* $c \geqslant 0$ *and* $Z \in L^\infty(P, \mathscr{F}_t)$, *such that*

$$E[ZW] \leqslant c \qquad \text{for all } W \in C. \tag{27}$$

'Then

(i) $E[ZW] \leqslant 0$ *for all* $W \in C$,

(ii) $Z \geqslant 0$ *P-a.s.,*

(iii) If $P(Z > 0) > 0$, *then*

$$\frac{dQ}{dP} := Z$$

defines a martingale measure and $Q \ll P$.

Proof. (i) Since C is a cone, it follows for any $W \in C$ and $\alpha > 0$, that

$$E[ZW] = \alpha \cdot E[Z\alpha^{-1}W] \leqslant \alpha c,$$

and the equation holds already for $c = 0$.

(ii) Choose $W = -\mathbb{1}_{\{Z<0\}} \in C$. Then,

$$E[Z_-] = E[ZW] \leqslant 0,$$

such that $Z_- = 0$ and hence $Z \geq 0$ P-a.s.

(iii) Choose H in $L^\infty(P, \mathscr{F}_{t-1}, \mathbb{R}^d)$, $\alpha \in \mathbb{R}$ and $Y = (X_t - X_{t-1})$. Then, $Y \in C$ and, since H and Z are bounded,

$$E[ZHY] \leqslant c \quad \text{and} \quad E[\alpha ZHY] \leqslant c.$$

As above, $E[ZHY] \leqslant 0$. Außerdem ist für alle $\alpha \in \mathbb{R}$

$$\alpha E[ZHY] \leqslant 0,$$

also $E[ZHY] = 0 = E[H(X_t - X_{t-1})]$. Wir erhalten $E^Q[\mathbb{1}_F(X_t^i - X_{t-1}^i)] = 0$ für alle $F \in \mathscr{F}_{t-1}$, also ist X Q-Martingal. □

So, by Lemma 26, the existence of an equivalent martingale measure is equivalent to find an element of the following set:

$$\mathcal{Z} := \{Z \in L^\infty,\ 0 \leqslant Z \leqslant 1,\ P(Z > 0) > 0,\ E[ZW] \leqslant 0\ \forall W \in C\}.$$

Lemma 28. *Assume that C is closed in L^1 and that $C \cap L^1_+ = \{0\}$. Then,*

(i) for all $F \in L^1_+ \backslash \{0\}$ there exists an $Z \in \mathcal{Z}$, such that $E[FZ] > 0$, and

(ii) there exists $Z^* \in \mathcal{Z}$, such that $Z^* > 0$.

Proof. For the first part, consider $B = \{F\}$. Then $B \cap C = \emptyset$, $C \neq \emptyset$; both sets are convex, B is compact and C is closed by assumption.

Then, we can apply the theorem of Hahn-Banach which gives us a continuous linear functional ℓ, such that

$$\sup_{W \in C} \ell(W) < \ell(F). \tag{29}$$

The dual space of L^1 can be identified with L^∞ by the Riesz theorem, such the linear function ℓ can be represented by a $Z \in L^\infty$ such that

$$\ell(F') = E[Z \cdot F'], \qquad F' \in L^1.$$

Without loss of generality we may assume that $\|Z\|_\infty = 1$.

By Equation (29), we obtain that $\ell(W) = E[ZW] < \ell(F) = E[ZF]$ for all $W \in C$. Hence, Z satisfies (27) and we can apply Lemma 26. This yields that $Z \in \mathcal{Z}$. Since $0 \in C$ we obtain that $E[FZ] > 0$.

For the second part, we start by showing that $\mathcal{Z}$ is countably convex. To this choose $\alpha_k \in [0, 1]$, $k \in \mathbb{N}$ with $\sum_{k=1}^\infty \alpha_k = 1$, and $(Z_k)_{k \in \mathbb{N}} \subset \mathcal{Z}$ and consider

$$Z := \sum_{k=1}^\infty \alpha_k Z_k.$$

For $W \in C$,

$$\sum_{k=1}^\infty |\alpha_k Z_k W| \leqslant |W| \sum_{k=1}^\infty |\alpha_k| = |W| \in L^1.$$

such that by dominated convergende

$$E[ZW] = \sum_{k=1}^\infty \alpha_k E[Z_k W] \leqslant 0$$

and hence $Z \in \mathcal{Z}$.

Now set

$$c := \sup\{P(Z > 0) :\ Z \in \mathcal{Z}\}.$$

and choose a sequence $(Z_n) \in \mathcal{Z}$, such that $P(Z_n > 0) \to c$. Then,

$$Z^* := \sum_{k=1}^{\infty} \frac{1}{2^k} Z_k \in \mathcal{Z}.$$

Now we show that $P(Z^* > 0) = 1$. Indeed, $\{Z^* > 0\} = \bigcup_{k=1}^{\infty} \{Z_k > 0\}$, such that

$$P(Z^* > 0) \geqslant \sup_{k \in \mathbb{N}} P(Z_k > 0) = c.$$

If we would have on the contrary that $P(Z^* = 0) > 0$, then $F := 1_{\{Z^*=0\}} \neq 0$ and $F \in L^1_+$. By Lemma 28 there exists $Z' \in \mathcal{Z}$, s.t.

$$0 < E[FZ'] = E[1_{\{Z^*=0\}} Z'].$$

Hence $P(\{Z' > 0\} \cap \{Z^* = 0\}) > 0$. This implies that the convex combination achieves

$$P\Big(\frac{1}{2}(Z' + Z^*) > 0\Big) > P(Z^* > 0),$$

a contradiction to the maximality of c and the claim follows. □

The following Lemma generalizes the theorem of *Bolzano-Weierstraß* to infinite dimensional spaces. Boundedness is not sufficient in infinite dimensional spaces, such that we require existence of an accumulation point instead.

Lemma 30. *Consider a sequence (H_n) of d-dimensional random variables and assume that*

$$\lambda := \liminf_n \parallel H_n \parallel < \infty.$$

Then there exists $H \in L^0(\mathbb{R}^d)$ and a strictly increasing sequence (σ_m) such that

$$H_{\sigma_m(\omega)}(\omega) \to H(\omega)$$

for P-almost all $\omega \in \Omega$.

The idea is to proceed pointwise, such that we can rely on the classical Bolzano-Weierstraß.

Proof. Define $\sigma_m = m$ on $\{\lambda = \infty\}$. On $\{\lambda < \infty\}$ let $\sigma_1^0 := 1$ and

$$\sigma_m^0(\omega) := \inf\Big\{ n > \sigma_{m-1}^0(\omega) : \parallel H_n(\omega) \parallel - |\lambda(\omega)| \leqslant \frac{1}{m} \Big\} \qquad m = 2, 3, \ldots$$

Now we proceed inductively through the coordinates. We denote for the sequence (σ^{i-1}), $H^i := \liminf_{m \to \infty} H^i_{\sigma_m^{i-1}}$ and construct (σ^i) as follows: let $\sigma_1^i = 1$ and

$$\sigma_m^i(\omega) := \inf\Big\{ \sigma_n^{i-1}(\omega) : \ \sigma_n^{i-1}(\omega) > \sigma_{m-1}^i(\omega) \quad \text{und} \quad |H^i_{\sigma_n^{i-1}(\omega)}(\omega) - H^i(\omega)| \leqslant \frac{1}{n} \Big\}.$$

Then, $\sigma_m := \sigma_m^d$ on $\{\lambda < \infty\}$ does the job. □

We are almost ready, but two portfolios can lead to the same payoff, which creates problems. Or, equivalently, it could happen that

$$H(X_t - X_{t-1}) = 0$$

holds even if $H \neq 0$. Using orthogonal projections we create a subset where this can not happen.

We consider the not locally convex space L^0 with the topology of convergence in probability, which is generated by the semi-metric $E[|X-Y| \wedge 1]$.

Lemma 31. *Define:*

$$N = \{H \in L^0(\Omega, \mathscr{F}_{t-1}, P; \mathbb{R}^d) : \; H(X_t - X_{t-1}) = 0 \quad P - \textit{f.s.}\}$$
$$N^{\perp} = \{G \in L^0(\Omega, \mathscr{F}_{t-1}, P; \mathbb{R}^d) : \; G \cdot H = 0 \textit{ für alle } H \in N\}$$

Then it holds that

(i) $N, N^{\perp}$ are closed in L^0. Moreover, for $g \in L^0(\Omega, \mathscr{F}_{t-1}, P; \mathbb{R})$ it holds that

$$gH \in N \textit{ if } H \in N \quad \textit{ and } gG \in N^{\perp} \textit{ if } G \in N^{\perp}.$$

(ii) $N \cap N^{\perp} = \{0\}$.
(iii) Every $G \in L^0(\Omega, \mathscr{F}_{t-1}, P, \mathbb{R}^d)$ has the following unique decomposition

$$G = H + G^{\perp}, \quad H \in N, \; G^{\perp} \in N^{\perp}.$$

Proof. (i) Consider a sequence $H_n \xrightarrow{P} H$. Then we have an a.s. converging subsequence (H_{σ_m}). This implies

$$H_{\sigma_m}(\omega) \cdot (X_t(\omega) - X_{t-1})(\omega) \to H(\omega)(X_t(\omega) - X_{t-1}(\omega)) \quad \text{for } P\text{-almost all } \omega \in \Omega. \tag{32}$$

Similarly, if we now consider a subsequence $(H_n) \subseteq N$ with $H_n \to H$ a.s., then the left hand side of (32) is equal to 0 for all n and so is the limit. Hence, $H \in N$.

Similarly, for $(G_k) \subseteq N^{\perp}$ with $G_n \to^{a.s.} G$, we obtain $G \in N^{\perp}$.

The additional property is immediate.

(ii) Since for $G \in N \cap N^{\perp}$, it holds by definition that

$$0 = G \cdot G = |G|$$

which is equivalent to $G = 0$ a.s.

(iii) For $\xi \in \mathbb{R}^d$ we write $\xi = \xi^1 e_1 + \cdots + \xi^d e_d$ with respect to a basis $\{e_1, \ldots, e_d\}$.

First, assume that $e_i = n_i + e_i^{\perp}$ with $n_i \in N$ and $e_i^{\perp} \in N^{\perp}$, $i = 1, \ldots, d$. Then

$$\xi = \underbrace{\sum_{i=1}^{d} \xi n_i}_{\in N} + \underbrace{\sum_{i=1}^{d} \xi_i e_i^{\perp}}_{\in N^{\perp}}.$$

The decomposition is unique since $N \cap N^{\perp} = \{0\}$.

Now we show $e_i = n_i + e_i^{\perp}$. To this end consider the Hilbert space $L^2 = L^2(\Omega, \mathscr{F}_t, P; \mathbb{R}^d)$ with scalar product $\langle X, Y \rangle = E[XY]$. Both $N \cap L^2$ and $N^{\perp} \cap L^2$ are closed subspaces of L^2, since convergence in probability implies L^2-convergence and we already showed closeness of N and $N^{\perp}$. We define the orthogonal projections

$$\pi : L^2 \to N \cap L^2, \quad \pi^{\perp} : L^2 \to N^{\perp} \cap L^2$$

and set $n_i = \pi(e_i)$, $e_i^\perp = \pi^\perp(e_i)$.

Now consider $\xi := e_i - \pi(e_i)$. Since $\pi(e_i)$ is the orthogonal projection, it holds that

$$\langle \xi, n \rangle = 0 \tag{33}$$

for all $n \in N \cap L^2$. Note that $e_i \in L^2$ and so is $\pi(e_i)$, such that $\xi \in L^2$. We show that $\xi \in N^\perp$: assume, $\xi \notin N^\perp \cap L^2$. Then there exists $H \in N$ with $P(\xi \cdot H > 0) > 0$. Set

$$\widetilde{H} := H \mathbb{1}_{\{\xi \cdot H > 0\ ,\ |H| \leqslant c\}} \in N \cap L^2.$$

If c is large enough,

$$0 < E[\widetilde{H} \cdot \xi] = \langle \widetilde{H}, \xi \rangle,$$

a contradiction to (33).

□

The final step is done in the following lemma: it shows, that already $K \cap L^0_+ = \{0\}$ implies the required closeness.

Lemma 34. *If $K \cap L^0_+ = \{0\}$, then $K - L^0_+$ is closed in L^0.*

Proof. Consider a sequence (W_n) of elements of $K - L^0_+$ converging in L^0 (hence in probability) to W. By changing to a subsequence we can assume that the convergence is even almost surely. Then we have the representation

$$W_n = \widetilde{H}_n \cdot \Delta X - U_n \overset{L.31}{=} \widetilde{\widetilde{H}}_n \Delta X + H_n^\perp \Delta X - U_n = H_n^\perp \Delta X - U_n =: H_n \Delta X - U_n,$$

with $H_n \in N^\perp$, since $\widetilde{\widetilde{H}}_n \Delta X = 0$.

First, we assume that $\liminf |H_n| < \infty$ P-a.s. Then, Lemma 30 implies that we find a subsequence for which $H_{\sigma_n} \to H$ P-a.s. Moreover,

$$0 \leqslant U_{\sigma_n} = H_{\sigma_n} \Delta X - W_{\sigma_n} \to H \Delta X - W =: U \qquad P - \text{f.s.}$$

with some $U \geqslant 0$, such that $W \in K - L^0_+$ and closeness holds.

The proof is finished when we can show that $\liminf |H_n| < \infty$ P-a.s. To this end, consider the trading strategy $\xi_n = \frac{H_n}{|H_n|}$ and $A = \{\omega \in \Omega \ : \ \liminf |H_n| = \infty\}$. We apply 30 to $\xi_n = \frac{H_n}{|H_n|}$. This yields a subsequence (τ_n), such that $\xi_{\tau_n} \to \xi$ P-a.s. Now it holds that

$$0 \leqslant \mathbb{1}_A \frac{U_{\tau_n}}{|H_{\tau_n}|} = \mathbb{1}_A \left(\frac{H_{\tau_n}}{|H_{\tau_n}|} \cdot \Delta X - \frac{W_{\tau_n}}{|H_{\tau_n}|} \right) \to \mathbb{1}_A \xi \Delta X \quad P - \text{a.s.},$$

since $\frac{W_{\tau_n}}{|H_{\tau_n}|} \to 0$. This yield that $\mathbb{1}_A \xi \Delta X \in K \cap L^0_+$, such that by our assumption $\mathbb{1}_A \xi \Delta X = 0$.

Note that for $\eta \in N$,

$$\xi_{\tau_n} \cdot \eta = \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_n = k\}} \frac{1}{|H_k|} H_k \cdot \eta = 0,$$

since $H_k \in N^\perp$. Hence, $\xi_{\tau_n} \in N^\perp$ and so is $\mathbb{1}_A \xi_{\tau_n}$. Since $N^\perp$ is closed, $\mathbb{1}_A \xi \in N^\perp$. But we also showed that $\mathbb{1}_A \xi \in N$. This is, by (ii) of Lemma 30, only possible if $\mathbb{1}_A \xi = 0$. But $|\xi| = 1$, such that $P(A) = 0$ follows. □

3 The 2nd fundamental theorem

In this section we study the second fundamental theorem of asset pricing. A proof for this theorem in discrete time under trivial initial conditions can be found in Föllmer & Schied (2016). We follow Niemann & Schmidt (2024) and show a full proof of the conditional version using non-linear expectations.

We start with a super short introduction to non-linear expectations.

3.1 Dynamic non-linear expectations

As before, we consider a measurable space $(\Omega, \mathscr{F})$ with a filtration $\mathbb{F} = (\mathscr{F}_t)_{t \in \{0,...,T\}}$. For this section, we assume that $\mathscr{F}_T = \mathscr{F}$ and that $\mathscr{F}_0$ is trivial.

Consider a set of probability measures $\mathscr{P}$ on $(\Omega, \mathscr{F})$. A P-null set $A \subseteq \Omega$ is a possibly not measurable set being a subset of a measurable set $A' \in \mathscr{F}$ with $P(A') = 0$. A set $A \subseteq \Omega$ is called a *$\mathscr{P}$-polar set*, if A is a P-null set for every $P \in \mathscr{P}$. We denote the collection of $\mathscr{P}$-polar sets by $\mathrm{Pol}(\mathscr{P})$. We say a property holds $\mathscr{P}$-quasi surely, in short $\mathscr{P}$-q.s., if it holds outside a $\mathscr{P}$-polar set. If $\mathscr{P} = \{P\}$, we write short $\mathrm{Pol}(P)$ instead of $\mathrm{Pol}(\{P\})$.

For two subsets of probability measures $\mathscr{P}$ and $\mathscr{Q}$, we call $\mathscr{Q}$ *absolutely continuous* with respect to $\mathscr{P}$, denoted by $\mathscr{Q} \ll \mathscr{P}$, if $\mathrm{Pol}(\mathscr{P}) \subseteq \mathrm{Pol}(\mathscr{Q})$. We write $\mathscr{Q} \sim \mathscr{P}$, if $\mathscr{Q} \ll \mathscr{P}$ and $\mathscr{P} \ll \mathscr{Q}$.

On $\mathscr{L}^0(\Omega, \mathscr{F}) = \{X : \Omega \to \mathbb{R} \colon X \ \mathscr{F}\text{-measurable}\}$ we introduce the equivalence relation $\sim_{\mathscr{P}}$ by $X \sim_{\mathscr{P}} Y$ if and only if $X = Y$ $\mathscr{P}$-q.s.. Then, we set

$$\begin{aligned} L^0(\Omega, \mathscr{F}, \mathscr{P}) &:= \mathscr{L}^0(\Omega, \mathscr{F}) / \mathscr{P} \\ L^p(\Omega, \mathscr{F}, \mathscr{P}) &:= \{X \in L^0(\Omega, \mathscr{F}, \mathscr{P}) : \sup_{P \in \mathscr{P}} E_P[|X|^p] < \infty\} \\ L^\infty(\Omega, \mathscr{F}, \mathscr{P}) &:= \{X \in L^0(\Omega, \mathscr{F}, \mathscr{P}) : \exists C > 0 : |X| \le C \ \mathscr{P}\text{-q.s.}\} \end{aligned}$$

Then, Proposition 14 in Denis et al. (2011) shows that for each $p \in [1, \infty]$, $L^p(\Omega, \mathscr{F}, \mathscr{P})$ is a Banach space. The space $L^0(\Omega, \mathscr{F}, \mathscr{P})$ can be equipped with the metric d given by

$$d(X, Y) := \sup_{P \in \mathscr{P}} E_P[|X - Y| \wedge 1]$$

inducing uniform convergence in probability.

We consider a set $\mathscr{H} \subseteq L^0(\Omega, \mathscr{F}, \mathscr{P})$ containing all constants and set, for $t \in \{0, ..., T\}$,

$$\mathscr{H}_t := \mathscr{H} \cap L^0(\Omega, \mathscr{F}_t, \mathscr{P}).$$

The following definition introduces the notion of a conditional nonlinear expectation and the associated notion of a dynamic nonlinear expectation which is a set of conditional nonlinear expectations.

Definition 35. We call a mapping $\mathcal{E}_t : \mathscr{H} \to \mathscr{H}_t$ an *$\mathscr{F}_t$-conditional nonlinear expectation*, if

(i) $\mathcal{E}_t$ is *monotone*: for $X, Y \in \mathscr{H}$ the condition $X \le Y$ implies $\mathcal{E}_t(X) \le \mathcal{E}_t(Y)$,

(ii) $\mathcal{E}_t$ *preserves measurable functions*: for $X_t \in \mathscr{H}_t$ we have $\mathcal{E}_t(X_t) = X_t$.

We call $\mathcal{E} = (\mathcal{E}_t)_{t \in \{0,...,T\}}$ a *dynamic nonlinear expectation*, if for every $t \in \{0, ..., T\}$ the mapping $\mathcal{E}_t : \mathscr{H} \to \mathscr{H}_t$ is an $\mathscr{F}_t$-conditional nonlinear expectation.

We introduce further properties which will be of interest in the context of dynamic nonlinear expectations. First, we introduce some well-known properties regarding the set $\mathscr{H}$, all in an appropriate conditional formulation. Denote $\mathscr{H}_t^+ := \{X \in \mathscr{H}_t : X \ge 0\}$.

Definition 36. We call the set $\mathscr{H}$

(i) *symmetric*, if $-\mathscr{H} = \mathscr{H}$.

(ii) *additive*, if $\mathscr{H} + \mathscr{H} \subseteq \mathscr{H}$.

(iii) *$\mathscr{F}_t$-translation-invariant*, if $\mathscr{H} + \mathscr{H}_t \subseteq \mathscr{H}$.

(iv) *$\mathscr{F}_t$-convex*, if for $\lambda_t \in \mathscr{H}_t$ with $0 \leq \lambda_t \leq 1$ we have

$$\lambda_t \mathscr{H} + (1 - \lambda_t)\mathscr{H} \subseteq \mathscr{H}.$$

(v) *$\mathscr{F}_t$-positively homogeneous*, if $\mathscr{H}_t^+ \cdot \mathscr{H} \subseteq \mathscr{H}$.

(vi) *$\mathscr{F}_t$-local*, if $\mathbb{1}_A \mathscr{H} \subseteq \mathscr{H}$ for every $A \in \mathscr{F}_t$.

We simply call $\mathscr{H}$ *translation-invariant*, if it is $\mathscr{F}_t$-translation-invariant for every $t \in \{0, ..., T\}$ and do so in a similar fashion for all the other properties.

Next, we introduce well-known properties of nonlinear conditional expectations, all in an appropriate conditional formulation which are frequently used for example in the context of risk measures.

Definition 37. An $\mathscr{F}_t$-conditional expectation $\mathcal{E}_t$ is called

(i) *subadditive*, if $\mathscr{H}$ is additive

$$\mathcal{E}_t(X + Y) \leq \mathcal{E}_t(X) + \mathcal{E}_t(Y), \qquad X,\ Y \in \mathscr{H}.$$

(ii) *$\mathscr{F}_t$-translation-invariant*, if $\mathscr{H}$ is $\mathscr{F}_t$-translation-invariant and

$$\mathcal{E}_t(X + X_t) = \mathcal{E}_t(X) + X_t, \qquad X \in \mathscr{H},\ \ X_t \in \mathscr{H}_t$$

(iii) *$\mathscr{F}_t$-convex*, if $\mathscr{H}$ is $\mathscr{F}_t$-convex and

$$\mathcal{E}_t\left(\lambda_t X + (1 - \lambda_t)Y\right) \leq \lambda_t \mathcal{E}_t(X) + (1 - \lambda_t)\mathcal{E}_t(Y), \quad 0 \leq \lambda_t \leq 1,\ \lambda_t \in \mathscr{H}_t,\ X, Y \in \mathscr{H}.$$

(iv) *$\mathscr{F}_t$-positively homogeneous*, if $\mathscr{H}$ is $\mathscr{F}_t$-positively homogeneous and

$$\mathcal{E}_t(X_t X) = X_t \mathcal{E}_t(X), \qquad X \in \mathscr{H},\ X_t \in \mathscr{H}_t^+.$$

(v) *$\mathscr{F}_t$-sublinear*, if it is subadditive and $\mathscr{F}_t$-positively homogeneous.

(vi) *$\mathscr{F}_t$-local*, if $\mathscr{H}$ is $\mathscr{F}_t$ local and

$$\mathcal{E}_t(\mathbb{1}_A X) = \mathbb{1}_A \mathcal{E}_t(X), \qquad X \in \mathscr{H},\ A \in \mathscr{F}_t.$$

Moreover, we call a dynamic expectation $\mathcal{E} = (\mathcal{E}_t)_{t \in \{0,...,T\}}$ translation-invariant, (subadditive, convex or positively homogeneous) if for every $t \in \{0, ..., T\}$ the $\mathscr{F}_t$-conditional expectation $\mathcal{E}_t$ has the corresponding property.

3.2 Sensitivity and time consistency

In contrast to a classical expectation, a nonlinear expectation might contain only little information on underlying random variables. Sensitivity is a property which allows at least to separate zero from positive random variables. It should be noted that such sensitivity on a suitable set of random variables is implied by no-arbitrage.

Definition 38. We call an $\mathscr{F}_t$-conditional nonlinear expectation $\mathcal{E}_t$ *sensitive*, if for every $X \in \mathscr{H}$ with $X \geq 0$ and $\mathcal{E}_t(X) = 0$ we have $X = 0$.

Similarly, we call the dynamic nonlinear expectation $\mathcal{E}$ *sensitive*, if all $\mathcal{E}_t$, $t = 0, \dots, T$ are sensitive.

Time consistency is an important property in the context of dynamic risk measures, which has been intensively studied. It transports the tower-property to the non-linear setting.

Definition 39. We call a dynamic expectation $\mathcal{E}$ *time-consistent*, if

$$\mathcal{E}_s = \mathcal{E}_s \circ \mathcal{E}_t, \qquad 0 \le s \le t \le T.$$

Since $\mathscr{F} = \mathscr{F}_T$, $\mathcal{E}_T$ is the identity and hence, every expectation is time-consistent between $T-1$ and T, i.e.,

$$\mathcal{E}_{T-1} \circ \mathcal{E}_T = \mathcal{E}_{T-1}.$$

Remark 40 (Extension of time-consistency to stopping times)**.** For simplicity, we restrict our definition of time consistency to deterministic times $s, t \in \{0, \dots, T\}$. This can easily be generalized when $\mathcal{E}$ is translation-invariant and local: indeed, let τ be a stopping time with values in $\{0, \dots, T\}$. Given $(\mathcal{E}_t)_t$, we define

$$\mathcal{E}_\tau(H) := \sum_s \mathbb{1}_{\{\tau = s\}} \mathcal{E}_s(H) .$$

If $(\mathcal{E}_t)_t$ is time-consistent, then for any two such stopping times σ, τ with $\sigma \le \tau$, the equality

$$\mathcal{E}_\sigma \circ \mathcal{E}_\tau = \mathcal{E}_\sigma$$

holds whenever $\mathcal{E}$ is translation-invariant and local.

The remarkable connection between sensitivity and time consistency can already be seen from the simple observation that a time-consistent dynamic expectation is already sensitive, if $\mathcal{E}_0$ is sensitive.

Remark 41. Let $\mathscr{P}$ and $\mathscr{Q}$ be two sets of probability measures on $(\Omega, \mathscr{F})$, and let $\mathcal{E}_t : L^\infty(\Omega, \mathscr{F}, \mathscr{Q}) \to L^\infty(\Omega, \mathscr{F}_t, \mathscr{Q})$ be a conditional nonlinear expectation. Then, $\mathcal{E}_t$ is well-defined on $L^\infty(\Omega, \mathscr{F}, \mathscr{P})$ if and only if $\mathscr{Q} \ll \mathscr{P}$. However, for $H \in L^\infty(\Omega, \mathscr{F}, \mathscr{P})$ the evaluation $\mathcal{E}_t(H)$ is a priori only an element of $L^\infty(\Omega, \mathscr{F}_t, \mathscr{Q})$. For it to be well-defined in $L^\infty(\Omega, \mathscr{F}, \mathscr{P})$ we require $\mathscr{Q} \sim \mathscr{P}$ on $\mathscr{F}_t$. Hence, if $\mathscr{Q} \ll \mathscr{P}$ and $\mathscr{Q} \sim \mathscr{P}$ on $\mathscr{F}_t$, the conditional nonlinear expectation $\mathcal{E}_t$ induces a nonlinear expectation $\bar{\mathcal{E}}_t : L^\infty(\Omega, \mathscr{F}, \mathscr{P}) \to L^\infty(\Omega, \mathscr{F}_t, \mathscr{P})$. In case $\mathcal{E}_t$ is sensitive, sensitivity of $\bar{\mathcal{E}}_t$ is equivalent to $\mathscr{P} \sim \mathscr{Q}$.

Lemma 42 below generalizes the well-known result that the acceptance sets of time-consistent expectations are decreasing: if $\mathcal{E}$ is time-consistent, then

$$\{\mathcal{E}_s \le 0\} \supseteq \{\mathcal{E}_t \le 0\}$$

for $s \le t$.

Lemma 42. *Let $\mathcal{E}$ be a time-consistent, local dynamic nonlinear expectation, fix $t \in \{0, \dots, T\}$ and consider $H \in \mathscr{H}$. If $\mathcal{E}_t(H) \le 0$, then*

$$\mathcal{E}_s(\mathbb{1}_A H) \le 0 \qquad \text{for all } A \in \mathscr{F}_t \text{ and all } s \le t.$$

If $\mathcal{E}_0$ is sensitive, then $\mathcal{E}_s(\mathbb{1}_A H) \le 0$ for all $A \in \mathscr{F}_t$ and some $s \le t$ implies that $\mathcal{E}_t(H) \le 0$.

Proof. Since $\mathcal{E}$ is local, $\mathcal{E}_t(H \mathbb{1}_A) = \mathbb{1}_A \mathcal{E}_t(H) \le 0$. Together with monotonicity we obtain

$$\mathcal{E}_s(\mathbb{1}_A H) = \mathcal{E}_s(\mathbb{1}_A \mathcal{E}_t(H)) \le 0.$$

Now, suppose $\mathcal{E}_0$ is sensitive. If $s = t$, the result is clear with $A = \Omega$. Let $s < t$ and note that $\mathcal{E}_s$ is sensitive. To show that $\mathcal{E}_t(H) \leq 0$, it thus suffices to show

$$\mathcal{E}_s(\mathbb{1}_A \mathcal{E}_t(H)) = 0$$

for $A := \{\mathcal{E}_t(H) \geq 0\} \in \mathscr{F}_t$. However, as above,

$$\mathcal{E}_s(\mathbb{1}_A \mathcal{E}_t(H)) = \mathcal{E}_s(\mathbb{1}_A H)$$

and the latter vanishes by assumption. □

Let $\mathcal{E}_0^*$ be a $\mathscr{F}_0$-conditional expectation. A *dynamic extension* of $\mathcal{E}_0^*$ is a dynamic expectation $\mathcal{E}$ such that $\mathcal{E}_0 = \mathcal{E}_0^*$.

Note that for any collection $\mathscr{P}$ of probability measures, the associated nonlinear expectation $\sup_{P \in \mathscr{P}} E_P[\cdot]$ is sensitive; see Remark 41. We call $\mathscr{P}$ *dominated* if there exists a probability measure P on $(\Omega, \mathscr{F})$ with $\mathscr{P} \ll \{P\}$, i.e., every P-null set is $\mathscr{P}$-polar. In this case, the Halmos-Savage Lemma guarantees the existence of a countable collection $\{P_n \colon n \in \boldsymbol{N}\} \subseteq \mathscr{P}$ with $\{P_n \colon n \in \boldsymbol{N}\} \sim \mathscr{P}$. In particular, there exists a measure P^* (not necessarily contained in $\mathscr{P}$) such that $\mathscr{P} \sim P^*$. Consequently, for any set of random variables $M \subseteq L^0(\Omega, \mathscr{F}, \mathscr{P}) = L^0(\Omega, \mathscr{F}, P^*)$, there exists a random variable called the $\mathscr{P}$-essential infimum and denoted by $\mathscr{P} - \operatorname{ess\,inf} M$ such that

(i) $\mathscr{P} - \operatorname{ess\,inf} M \leq Y$ $\mathscr{P}$-q.s. for every $Y \in M$,

(ii) $\mathscr{P} - \operatorname{ess\,inf} M \geq Z$ $\mathscr{P}$-q.s. for every random variable Z satisfying $Z \leq Y$ $\mathscr{P}$-q.s. for every $Y \in M$.

If $\mathscr{P}$ is not dominated, the $\mathscr{P}$-essential infimum might not exist, and it has in general no countable representation. In light of the financial applications we have in mind, we will assume in the next lemma that $\mathscr{P}$ is dominated.

Lemma 43. *Assume that $\mathscr{P}$ is dominated. Then, every sensitive $\mathscr{F}_0$-conditional expectation $\mathcal{E}_0$ on a symmetric set $\mathscr{H}$ has at most one translation-invariant, local, time-consistent dynamic extension $\mathcal{E}$. If it exists, it is given by*

$$\mathcal{E}_t(H) = \mathscr{P} - \operatorname{ess\,inf}\{H_t \in \mathscr{H}_t : H - H_t \in A_t\},$$

where

$$A_t := \{H \in \mathscr{H} : \mathcal{E}_0(\mathbb{1}_A H) \leq 0,\ \forall A \in \mathscr{F}_t\}.$$

Proof. Lemma 42 characterizes for $t \geq 1$ the acceptance set $A_t := \{H \in \mathscr{H} \colon \mathcal{E}_t \leq 0\}$ solely in terms of $\mathcal{E}_0$ and $\boldsymbol{F}$. Indeed, it yields that

$$A_t = \{H \in \mathscr{H} : \mathcal{E}_0(\mathbb{1}_A H) \leq 0\ \forall A \in \mathscr{F}_t\}.$$

This allows to recover every translation-invariant nonlinear expectation on a symmetric set from its acceptance set through the representation

$$\begin{aligned} \mathcal{E}_t(H) &= \mathscr{P} - \operatorname{ess\,inf}\{H_t \in \mathscr{H}_t : H_t \geq \mathcal{E}_t(H)\} \\ &= \mathscr{P} - \operatorname{ess\,inf}\{H_t \in \mathscr{H}_t : H - H_t \in A_t\}. \end{aligned}$$

Summarizing, we have obtained an explicit expression of the extension. □

Next, we verify that translation-invariance implies locality if $\mathscr{H} \subseteq L^\infty(\Omega, \mathscr{F}, \mathscr{P})$. This implies that every conditional risk measure on $L^\infty(\Omega, \mathscr{F}, \mathscr{P})$ is local. In particular, for every probability measure P, every dynamic risk-measure on $L^\infty(\Omega, \mathscr{F}, P)$ has at most one time-consistent extension. Moreover, one can show that not every coherent risk measure has a time-consistent extension.

Proposition 44. *Every translation-invariant expectation on a local set $\mathscr{H} \subseteq L^\infty(\Omega, \mathscr{F}, \mathscr{P})$ is local.*

Proof. Let $\mathcal{E}_t$ be translation-invariant and $A \in \mathscr{F}_t$. Further, let $H \in \mathscr{H}$. The inequality

$$\mathbb{1}_A H - \mathbb{1}_{A^c} \|H\|_\infty \leq H \leq \mathbb{1}_A H + \mathbb{1}_{A^c} \|H\|_\infty ,$$

yields

$$\mathcal{E}_t(H) \geq \mathcal{E}_t(\mathbb{1}_A H - \mathbb{1}_{A^c} \|H\|_\infty),$$

and additionally

$$\mathcal{E}_t(H) \leq \mathcal{E}_t(\mathbb{1}_A H + \mathbb{1}_{A^c} \|H\|_\infty).$$

Multiplying both inequalities with $\mathbb{1}_A$ gives, exploiting translation invariance,

$$\mathbb{1}_A \mathcal{E}_t(H) = \mathbb{1}_A \mathcal{E}_t(\mathbb{1}_A H),$$

and thus

$$\begin{aligned}\mathcal{E}_t(\mathbb{1}_A H) &= \mathbb{1}_A \mathcal{E}_t(\mathbb{1}_A H) + \mathbb{1}_{A^c} \mathcal{E}_t(\mathbb{1}_A H) \\ &= \mathbb{1}_A \mathcal{E}_t(H) + \mathbb{1}_A \mathbb{1}_{A^c} \mathcal{E}_t(\mathbb{1}_A H) \\ &= \mathbb{1}_A \mathcal{E}_t(H).\end{aligned}$$

□

3.3 *Super- and Sub-hedging*

Now we turn back to a financial market. Recall that we worked on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, P)$ with a fixed probability measure P. Our aim is to study the upper bound of the set of no-arbitrage prices in more detail. It is given by

$$\bar{\mathcal{E}}_t(C_T) := \text{esssup}\{E_Q[C_T \mid \mathscr{F}_t] : Q \in \mathscr{M}_e\}, \tag{45}$$

for $H \in L^\infty(P)$. We are interested in its relation to the smallest super-hedging price given by

$$\mathcal{E}_t(C_T) := \text{ess}\inf\{C_t \in L^0_t : \exists\, H \in \text{Pred} : C_t + G_t(H) \geq C_T\}\,, \tag{46}$$

where the gains process for the predictable (and hence self-financing) strategy H is given by $G_t(H) := (H \cdot X)_T - (H \cdot X)_t$. We denote $\mathcal{E}(C)$ for the process $(\mathcal{E}_t(C))_{t \in \{0,t\dots,T\}}$ and use a similar notation for the other dynamic non-linear expectations.

Lemma 47. *Assume that NA holds. Then, the super-hedging price $\mathcal{E}_t$ defined in* (46) *is a sensitive and $\mathscr{F}_t$-sub-linear, $\mathscr{F}_t$-conditional nonlinear expectation on L^∞ for all* $0 \leq t \leq T$.

Proof. First, we show that $\mathcal{E}_t(C_T)$ is bounded for $C_T \in L^\infty(P)$. The inequality $\mathcal{E}_t(C_T) \leq \|C_T\|$ follows by definition. If $C_t \in L^0_t$ is a superhedging price, choose $H \in \text{Pred}$ with $C_t + G_t(H) \geq C_T$.

Consider the set $A := \{C_t < -\|C_T\|\} \in \mathscr{F}_t$. Then there exists $H \in \text{Pred}$ with $G_t(C_T) \geq \mathbb{1}_A(C_T - C_t) \geq 0$ and therefore $P(A) = 0$ by absence of arbitrage. We conclude that $-\|C_T\| \leq \mathcal{E}_t(C_T) \leq \|C_T\|$.

Second, one easily checks the properties of a sublinear expectation. The sensitivity of $\mathcal{E}_t$ follows from the no-arbitrage assumption. □

Up to now we treated only bounded random variables, which excludes for example European calls. Typically one would consider the space $L^0_+(\mathscr{F}_T)$, which is of course not symmetric. The extension to $L^0(\mathscr{F}_T)$ is done by establishing continuity from below of the superheding price. The main argument resides of course on monotone convergence.

3.4 The superhedging duality

At time $t < T$, an $\mathscr{F}_t$ - measurable random variable π_t is a superhedging price for the European claim C_T due at time T, if there is a self-financing trading strategy which provides always a terminal wealth greater than C_T, i.e.there exists a predictable process H, such that

$$\pi_t + G_t(H) \geq C_T.$$

Remark 48. For every $C_T \in L^0(\mathscr{F}_T)$, and every predictable H one has the orthogonality

$$\mathcal{E}_t(C_T + G_t(H)) = \mathcal{E}_t(C_T). \tag{49}$$

For the following lemma we recall that on a upwards directed set, i.e. for $X, Y \in M$ there exists $Z \in M$ such that $Z \geq X \vee Y$, the essential supremum can be approximated by a sequence, see for example (Föllmer & Schied 2016, Theorem A.37). The lemma shows that the essential infimum of all superhedging prices is itself a superhedging price.

Lemma 50. *Assume that NA holds. For every $C_T \in L^0(\mathscr{F}_T)$, $\mathcal{E}_t(C_T)$ is a superhedging price for C_T.*

Proof. The set $M := \{C_t \in L^0_t : \exists H \in \text{Pred} : C_t + G_t(H) \geq C_T\}$ of superhedging prices is directed downwards. Hence, by (Föllmer & Schied 2016, Theorem A.37) there exists a sequence $(C^n_t)_n \subseteq M$ with $C^n_t \downarrow \mathcal{E}_t(C)$ a.s. By construction, we may write for each $n \in \mathbb{N}$,

$$C_T = C^n_t + G_t(H^n) - U^n$$

for some $H^n \in \text{Pred}$ and $U^n \in L^0_+(\mathscr{F}_T)$.

As the cone $\{G_t(H) - U : H \in \text{Pred},\ U \in L^0_+(\mathscr{F}_T)\}$ is closed by Lemma 34, the claim follows. □

The following result shows that the superhedging prices are actually time-consistent.

Theorem 51. *Assume that NA holds. Then, the dynamic non-linear expectation $\mathcal{E}$ is time-consistent on L^∞.*

Proof. Applying Lemma 50 to the European contingent claims $\mathcal{E}_{t+1}(C_T)$ and C_T allows to choose strategies $H, H' \in \text{Pred}$ such that

$$\mathcal{E}_t(\mathcal{E}_{t+1}(C_T)) + G_t(H) \geq \mathcal{E}_{t+1}(C_T)$$

and

$$\mathcal{E}_{t+1}(C_T) + G_{t+1}(H') \geq C_T.$$

Combining both inequalities yields

$$\mathcal{E}_t(\mathcal{E}_{t+1}(C_T)) + G_t(H) + G_{t+1}(H') \geq C_T.$$

Hence, the claim C_T can be super-replicated at time t at price $\mathcal{E}_t(\mathcal{E}_{t+1}(C_T))$. As $\mathcal{E}_t(C_T)$ is by definition the smallest super-hedging price for claim C_T at time t, we obtain

$$\mathcal{E}_t(\mathcal{E}_{t+1}(C_T)) \geq \mathcal{E}_t(C_T).$$

Next, we obtain from Lemma 50 the existence of $H'' \in \text{Pred}$, such that

$$\mathcal{E}_t(C_T) + G_t(H'') \geq C_T.$$

Applying $\mathcal{E}_{t+1}$ to this inequality gives

$$\mathcal{E}_t(C_T) + \mathcal{E}_{t+1}(G_t(H'')) \geq \mathcal{E}_{t+1}(C_T).$$

By Equation (49),

$$\mathcal{E}_{t+1}(G_t(H'')) = H''_{t+1}\Delta X_{t+1} + \mathcal{E}_{t+1}(G_{t+1}(H'')) = H''_{t+1}\Delta X_{t+1}$$

and hence

$$\mathcal{E}_t(C_T) + H''_{t+1}\Delta X_{t+1} \geq \mathcal{E}_{t+1}(C_T).$$

Again, by Equation (49),

$$\mathcal{E}_t(\mathcal{E}_{t+1}(C_T)) \leq \mathcal{E}_t(C_T)$$

and the result is proven. □

Remark 52 (Time-consistency of $\bar{\mathcal{E}}$)**.** Consistency of $\bar{\mathcal{E}}$ is related to the stability of $\mathscr{M}_e$. With a little bit of work we obtain from (Föllmer & Schied 2016, Theorem 11.22) that the expectation $\bar{\mathcal{E}}$ is time-consistent.

The following result is the famous super-hedging duality.

Corollary 53 (Superhedging duality on L^∞)**.** *Assume (NA) holds. Then, for every $0 \leq t \leq T$ and every $C_T \in L^\infty(P)$, the superhedging-duality*

$$\mathcal{E}_t(C_T) = \bar{\mathcal{E}}_t(C_T) \tag{54}$$

holds.

Proof. By Theorem 51 and Remark 52, both $\mathcal{E}$ and $\bar{\mathcal{E}}$ are time-consistent. Moreover, they are translation invariant and hence local by Proposition 44. We leave the claim that

$$\mathcal{E}_0 = \bar{\mathcal{E}}_0$$

to the reader. Then, Lemma 43 implies the claim. □

By some monotone convergence arguments this can be extended to the space of claims (i.e. non-negative random variables).

Proposition 55 (Superhedging duality for claims)**.** *Assume that (NA) holds. The superhedging-duality* (54)*, and consistency of $\mathcal{E}$ extends to $L^0_+(\mathscr{F}_T)$.*

For the proof we refer to Proposition 2.16 in Niemann & Schmidt (2024).

Next, we prove a version of the optional decomposition directly by relying on the superheding-duality. For the stochastic integral until t we use the following notation

$$(H \cdot X)_t = \sum_{s=1}^{t} H_s \, \Delta X_s.$$

Theorem 56 (Optional decomposition). *Assume that (NA) holds and let V be a non-negative $\mathscr{M}_e$-supermartingale. Then there exists an adapted increasing process C with $C_0 = 0$, and a predictable process H such that*

$$V_t = V_0 + (H \cdot X)_t - C_t \,.$$

Proof. By assumption,

$$E_Q[V_t \mid \mathscr{F}_{t-1}] \leq V_{t-1}$$

for every $0 \leq t \leq T$ and $Q \in \mathscr{M}_e$. This is equivalent to $\bar{\mathcal{E}}_{t-1}(V_t) \leq V_{t-1}$ and hence $\bar{\mathcal{E}}_{t-1}(\Delta V_t) \leq 0$. By Proposition 55,

$$\mathcal{E}_{t-1}(\Delta V_t) \leq 0.$$

Hence, for $t \in \{1, \dots, T\}$ there exists a strategy $H = H^{(t)} \in \text{Pred}$ such that

$$\Delta V_t \leq G_{t-1}(H) = \sum_{s=t}^{T} H_s \Delta X_s.$$

an application of $\mathcal{E}_t$ on both sides yields, by Equation (49),

$$\mathcal{E}_t(\Delta V_t) = \Delta V_t \leq \mathcal{E}_t(H_t \Delta X_t + G_t(H)) = H_t \Delta X_t.$$

Summing over $t \in \{0, \dots, T\}$, we obtain a predictable H' such that $(H' \cdot X) - V$ is increasing. □

3.5 *Structure of arbitrage-free prices*

The main goal in computing arbitrage-free prices relying on the fundamental theorem of asset pricing is to obtain a price process for a new security which can be added to the market without violating absence of arbitrage.

In this spirit, an $\mathscr{F}_t$-measurable random variable π_t is called *arbitrage-free price* (at time t) of a European contingent claim C_T if there exists an adapted process X^{d+1} such that $X_t^{d+1} = \pi_t$, $X_T^{d+1} = H$ and the market (X, X^{d+1}) extended with X^{d+1} is free of arbitrage. Note that everything is formulated in discounted terms here. Denote by $\Pi_t(C_T)$ the collection of arbitrage free prices at time t.

Denote the upper and the lower bound of the no-arbitrage set at time t by

$$\pi_t^{\sup}(C_T) := \operatorname{esssup} \Pi_t(C_T), \quad \text{and} \quad \pi_t^{\inf}(C_T) := \operatorname{ess\,inf} \Pi_t(C_T).$$

To achieve countable convexity of the set of equivalent martingale measures we exploit nonnegativity of the price process and triviality of the initial σ-algebra $\mathscr{F}_0$ in the following lemma.

Lemma 57. *$\mathscr{M}_e$ is countably convex.*

Proof. Let $(Q^n) \subseteq \mathscr{M}_e$ and $(\lambda^n)_n \subseteq \mathbb{R}_+$ with $\sum_n \lambda^n = 1$. Set $Q^* := \sum_n \lambda^n Q^n$. Obviously, $Q^* \sim P$. For every $t \in \{1, \dots, T\}$ we have by monotone convergence

$$E_{Q^*}[X_t] = \sum_n \lambda^n E_{Q^n}[X_t] = \sum_n \lambda^n X_0 = X_0 < \infty$$

and hence $X_T \in L^1(Q^*)$. Similarly, for $A \in \mathscr{F}_{t-1}$,

$$E_{Q^*}[X_t \mathbb{1}_A] = \sum_n \lambda^n E_{Q^n}[X_t \mathbb{1}_A] = \sum_n \lambda^n E_{Q^n}[X_{t-1} \mathbb{1}_A] = E_{Q^*}[X_{t-1} \mathbb{1}_A]$$

and therefore $E^*[X_t \mid \mathscr{F}_{t-1}] = X_{t-1}$. □

It is important to acknowledge that, in the notation of Lemma 57, we typically do not have

$$E_{Q^*}[H \mid \mathscr{F}_t] = \sum_n \lambda^n E_{Q^n}[H \mid \mathscr{F}_t]$$

for bounded H at $t > 0$, while this holds, as just shown, for X^i_T, $i = 1, \dots, d$.

Proposition 58. *For every $t \in \{0, \dots, T\}$, and for every $C_T \in L^0_+(\mathscr{F}_T)$ the set*

$$\left\{ E_Q[C_T \mid \mathscr{F}_t] : Q \in \mathscr{M}_e \right\} \tag{59}$$

is $\mathscr{F}_t$-countably convex.

Proof. Let $(Q^n) \subseteq \mathscr{M}_e$. By pasting we may assume that all Q^n agree on $\mathscr{F}_t$.

Set $Q^* := \sum_n 2^{-n} Q^n$. By Lemma 57, $Q^* \in \mathscr{M}_e$. Denote by $Z^n := dQ^n/dQ^*$ the associated densities. As $Q^* = Q^n$ on $\mathscr{F}_t$ for each $n \in \mathbf{N}$, we have

$$Z^n_t = E_{Q^*}[Z^n \mid \mathscr{F}_t] = 1.$$

Since $C_T \geq 0$, monotone convergence implies for a sequence $(\lambda^n_t) \in L^0_+(\mathscr{F}_t)$ with $\sum_n \lambda_t = 1$, that

$$\sum_n \lambda^n_t E_{Q^n}[C_T \mid \mathscr{F}_t] = E_{Q^*}\Big[C_T \sum_n \lambda^n_t Z^n \mid \mathscr{F}_t\Big].$$

Set $Z := \sum_n \lambda^n_t Z^n > 0$. Note that

$$E_{Q^*}\Big[\sum_n \lambda^n_t Z^n \mid \mathscr{F}_t\Big] = \sum_n \lambda^n_t = 1$$

and we may therefore define the measure Q by

$$dQ/dQ^* := Z.$$

Then,

$$E_{Q^*}\Big[C_T \sum_n \lambda^n_t Z^n \mid \mathscr{F}_t\Big] = E_Q[C_T \mid \mathscr{F}_t].$$

It remains to verify that Q is indeed a martingale measure (after t). As the price process is nonnegative, its conditional expectation is well-defined, and we obtain by monotone convergence for $s \geq t$

$$\begin{aligned} E_Q[X_{s+1} \mid \mathscr{F}_s] &= E_{Q^*}\Big[\frac{Z}{Z_s} X_{s+1} \mid \mathscr{F}_s\Big] \\ &= \frac{1}{Z_s} \sum_n \lambda^n_t E_{Q^*}[Z^n X_{s+1} \mid \mathscr{F}_s] \\ &= \frac{1}{Z_s} \sum_n \lambda^n_t Z^n_s E_{Q^n}[X_{s+1} \mid \mathscr{F}_s] = X_s \end{aligned}$$

such that $Q \in \mathscr{M}_e$. □

Note that, due to the integrability conditions, $\Pi_t(C_T)$ is not necessarily $\mathscr{F}_t$-countably convex. Even in the unconditional case this fails. It is an easy consequence that integrability is the only difference between the set of risk-neutral expectations in (59) and $\Pi_t(C_T)$.

Lemma 60. *Consider $Q \in \mathscr{M}_e$. If $E_Q[C_T \mid \mathscr{F}_t]$ is finite-valued, then it is an arbitrage-free price.*

Proof. If $\xi := E_Q[C_T \mid \mathscr{F}_t]$ is finite, it is an element of $L^0_+(\mathscr{F}_T)$. We recall that we always may achieve integrability for ξ under an equivalent martingale measure: indeed, we can always find a P' such that $\xi \in L^1(P)$. Then we can choose a martingale measure with a bounded density.

Hence, there exists $\widetilde{Q} \in \mathscr{M}_e$ such that $E_Q[C_T \mid \mathscr{F}_t]$ is integrable with respect to $\widetilde{Q}$. We now paste Q and $\widetilde{Q}$, which is again a martingale measure. By construction, we even have $\widetilde{Q} \odot_t Q \in \mathscr{M}_e^{C_T}$. Moreover, it follows that

$$E_{\widetilde{Q}\odot_t Q}[C_T \mid \mathscr{F}_t] = E_Q[C_T \mid \mathscr{F}_t].$$

Since the associated price process is a martingale, this an arbitrage-free price by the fundamental theorem 15. □

Corollary 61. *Consider a claim $C_T \in L^0_+(\mathscr{F}_T)$, let $(Q^n) \subseteq \mathscr{M}_e^H$ and $(\lambda_t^n) \subseteq L^0_+(\mathscr{F}_t)$ with $\sum_n \lambda_t^n = 1$. If $\sum_n \lambda_t^n E_{Q^n}[C_T \mid \mathscr{F}_t]$ is finite-valued, it is contained in $\Pi_t(C_T)$.*

Proof. Due to Proposition 58, there exists $Q \in \mathscr{M}_e$ with

$$\sum_n \lambda_t^n E_{Q^n}[C_T \mid \mathscr{F}_t] = E_Q[C_T \mid \mathscr{F}_t]\,.$$

Now the claim follows by Lemma 60. □

Now we collect some properties of the non-linear expectation Π_t.

Corollary 62. *Consider $t \in \{0, \dots, T\}$. Then*

(i) $\Pi_t(H)$ is $\mathscr{F}_t$-convex for every claim $H \in L^0_+(\mathscr{F}_T)$,

(ii) $\Pi_t(H)$ is directed upwards for every claim $H \in L^0_+(\mathscr{F}_T)$,

(iii) $\Pi_t(H)$ is $\mathscr{F}_t$-countably convex for every bounded claim $H \in L^\infty(P)$, and,

(iv) for $H \in L^0_+(\mathscr{F}_T)$, any partition $(A^n) \subseteq \mathscr{F}_t$, and any sequence $(Q^n) \subseteq \mathscr{M}_e^H(P)$,

$$\sum_n \mathbb{1}_{A^n} E_{Q^n}[H \mid \mathscr{F}_t] \in \Pi_t(H)\,.$$

The next step is to show that Π_t is also local.

Lemma 63. *For $t \in \{0, \dots, T\}$, $A \in \mathscr{F}_t$ and $H \in L^0_+(\mathscr{F}_T)$, it holds that*

$$\Pi_t(\mathbb{1}_A H) = \mathbb{1}_A \Pi_t(H).$$

Proof. Let $Q \in \mathscr{M}_e$ such that $H\mathbb{1}_A$ is integrable with respect to Q. By construction $E_Q[H \mid \mathscr{F}_t]\mathbb{1}_A + E_{\widetilde{Q}}[H \mid \mathscr{F}_t]\mathbb{1}_{A^c}$ is finite, and by Corollary 62 and Lemma 60 there exists $Q^* \in \mathscr{M}_e^H$ with

$$E_{Q^*}[H \mid \mathscr{F}_t] = E_Q[H \mid \mathscr{F}_t]\mathbb{1}_A + E_{\widetilde{Q}}[H \mid \mathscr{F}_t]\mathbb{1}_{A^c}$$

and therefore

$$E_{Q^*}[H \mid \mathscr{F}_t]\mathbb{1}_A = E_Q[H\mathbb{1}_A \mid \mathscr{F}_t]$$

which finishes the proof. □

The next Proposition shows that, for every claim H, the non-linear expectation

$$\bar{\mathcal{E}}(H) = \operatorname{esssup}\{E_Q[H \mid \mathscr{F}_t] : Q \in \mathscr{M}_e\}$$

can be computed by considering a subset of $\mathscr{M}_e$ only: one can restrict to the set of martingale measure $\mathscr{M}_e^H(P)$ under which H is integrable. In particular, for every claim H, the non-linear expectation $\bar{\mathcal{E}}(H)$ agrees with the upper bound of the no-arbitrage interval $\pi_t^{\sup}(H)$. This links the superhedging-duality Proposition 55 with the pricing in financial markets.

Proposition 64. *For every $H \in L^0_+(\mathscr{F}_T)$ we have the equalities*

$$\operatorname{esssup}\{E_Q[H \mid \mathscr{F}_t] : Q \in \mathscr{M}_e\} = \operatorname{esssup}\{E_Q[H \mid \mathscr{F}_t] : Q \in \mathscr{M}_e^H\}$$

and

$$\operatorname{ess\,inf}\{E_Q[H \mid \mathscr{F}_t] : Q \in \mathscr{M}_e\} = \operatorname{ess\,inf}\{E_Q[H \mid \mathscr{F}_t] : Q \in \mathscr{M}_e^H\}$$

Proof. We only show the first equality, the other one is left as exercise. Using Lemma 63 and Lemma 60, it suffices to show the following: if there exits $Q \in \mathscr{M}_e$ with $E_Q[H \mid \mathscr{F}_t] = +\infty$, then $\pi_t^{\sup}(H) = +\infty$.

In this regard, consider $Q \in \mathscr{M}_e$ with $E_Q[H \mid \mathscr{F}_t] = +\infty$ and some $Y_t \in L^0_t$. Then, there exists $n \in \mathbb{N}$ such that $\{Y_t \le E_Q[H \wedge n \mid \mathscr{F}_t]\}$ has positive probability. By the fundamental theorem of asset pricing we find $\pi_t \in \Pi_t(H)$ such that $\{Y_t \le \pi_t\}$ has positive probability.

Since Y_t was arbitrary, it follows that $\pi_t^{\sup} = +\infty$ with positive probability. Now set $A := \{\pi_t^{\sup} < +\infty\}$. Using Lemma 63, and arguing as above for the claim $H\mathbb{1}_A$, we deduce that $P(A) = 0$. □

For the next lemma, recall that the smallest superhedging price $\mathcal{E}_t$ was defined in (46).

Lemma 65. *Consider the claim $C_T \in L^0_+(\mathscr{F}_T)$. Then, C_T is symmetric w.r.t. $\mathcal{E}_t$ if and only if C_T is attainable at time t.*

Proof. We start with some observations. Symmetry requires to consider $\mathcal{E}_t^*(\cdot) = -\mathcal{E}_t(-\cdot)$. This is the smallest subhedging price, and as a consequence of the superhedging-duality, Corollary 53,

$$\mathcal{E}_t^*(H) = \operatorname{esssup}\{C_t \in L^0_t : \exists\, \underline{H} \in \text{Pred} : H_t + G_t(\underline{H}) \le C_T\}$$

is the largest sub-hedging price. Due to Lemma 50, $\mathcal{E}_t^*(H)$ is itself a sub-hedging price.

Now, suppose that C_T is attainable, i.e. $C_T = C_t + G_t(H)$ for some predictable process $H \in \text{Pred}$. Then,

$$\begin{aligned}\mathcal{E}_t^*(C_T) &= -\mathcal{E}_t(-C_T) = -\mathcal{E}_t(-C_t - G_t(H)) \\ &= \mathcal{E}_t^*(C_t) = C_t = \mathcal{E}_t(C_t + G_t(H)) = \mathcal{E}_t(C_T).\end{aligned}$$

On the contrary, if C_T is symmetric, $\mathcal{E}_t(C_T) = \mathcal{E}_t^*(C_T)$ is by definition finite. Hence there is a super- and a subhedging strategy such that

$$\mathcal{E}_t^*(C_T) + G_t(\underline{H}) \le C_T \le \mathcal{E}_t(C_T) + G_t(H)\,. \tag{66}$$

This implies

$$0 \le H - \mathcal{E}_t(C_T) - G_t(H) \le G_t(\underline{H} - H)\,.$$

By no-arbitrage, $G_t(H) = G_t(\underline{H})$ and so $C_T = \mathcal{E}_t(H) + G_t(H)$. □

Theorem 67 (2nd fundamental theorem)**.** *The market is complete at time t if and only if every European contingent claim $C_T \in L^0_+(\mathscr{F}_T)$ has a unique price at time t.*

Proof. Assume that the market is complete. Then, by Lemma 65 and the superhedging duality, Proposition 55, every contingent claim has a unique arbitrage-free price.

On the contrary, if a contingent claim has a unique arbitrage-free price, the superhedging duality implies that C_T is symmetric and hence the claim is attainable. □

One can additionally show a number of things: for example completeness is equivalent to the pasting property $\mathscr{M}_e \subset \mathscr{M}_e \odot_t Q$ with some $Q \in \mathscr{M}_e$. Moreover, if the market

For details, we again refer to Niemann & Schmidt (2024).

4 *A fundamental theorem under uncertainty*

II
Continuous-time Finance

We start with a gentle introduction to semimartingale theory, relying on the scriptum on stochastic processes from last semester.

1 *Semimartingale theory*

Let us first visit some important examples for semimartingales. We recall that a process is called càdlàg, if it is a process which has almost surely left limits and is almost surely continuous from the right (RCLL - right continuous with left limits).

We will consider a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, P)$ satisfying the usual conditions, i.e. the filtration is right-continuous and $\mathscr{F}$ is complete (subsets of null-sets are $\mathscr{F}_0$-measurable).

1.1 *The Poisson process*

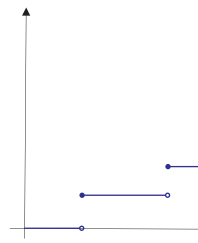


Figure II.1: Path of a

Definition 1 (Poissonprocess)**.** An adapted, càdlàg process X taking values in $\mathbb{N}$ is called *extended Poisson process*, if

(i) $X_0 = 0$,

(ii) $\Delta X_t \in \{0, 1\}$,

(iii) $E[X_t] < \infty$ for all $t \geq 0$,

(iv) $X_t - X_s$ is independent of $\mathscr{F}_s$, for $0 \leq s \leq t$.

We define the *cumulated intensity* Λ of X by

$$\Lambda(t) = E[X_t], \qquad t \geq 0.$$

Note that this is again an increasing, right-continuous process. X is called *Poisson-Prozess* with intensity $\lambda > 0$, if $\Lambda(t) = \lambda t,\ t \geq 0$.

If the cumulated intensity is absolutely continuous, i.e.

$$\Lambda(t) = \int_0^t \lambda(s)\, ds, \qquad t \geq 0$$

then the function λ is called the intensity of X. X is called standard Poissonprocess if $\lambda = 1$.

We note that we also may look at the time-transformed Poisson-process

$$X_{T_t}, \qquad t \geq 0,$$

Whenever T is increasing. If T is continuous and independent of X, we obtain that X is a Poisson process *conditional* on the filtration generated by the time-transformation T. If T has jumps we can no longer guarantee $\Delta X_t \in \{0, 1\}$.

The Poisson process corresponds one-to-one to a point process. Indeed if $T_n = \inf\{t \geq t : N_t \geq n\}$ denotes the n-th jumping time of N, then $(T_n)_{n\geq 1}$ are an increasing sequence of stopping time, so a *point process*.

If we additionally have a sequence $(Z_n)_{n\geq 1}$ of random variable on a Polish space E, then the double sequence $(T_n, Z_n)_{n\geq 1}$ constitutes a *marked point process*.

If we want to construct integrals over the marked point process we would be interested in expressions like

$$\sum_{n\geq 1} H(T_n, Z_n) = \int H(s, x)\mu(ds, dx)$$

where we can introduce the associated random measure

$$\mu(\omega; ds, dx) = \sum_{n\geq 1} \delta_{(T_n, Z_n)}(ds, dx),$$

where δ_a is the Dirac measure in point a.

1.2 *Survival processes*

In many applications, the first jump of the Poisson process is the most important one: mortality, default, insurance, etc. and it is therefore interesting to study this process in more generality.

Hence, consider a càdlàg process H with $H_0 = 0$ and a single jump of size 1. Then, this process is increasing, and hence by the Doob-Meyer decomposition there exists a unique compensator H^p which is a predictable process such that

$$H - H^p$$

is a local martingale. H^p takes over the role of a generalised intensity: indeed, in the Poisson example above, $H^p = \Lambda$. For a deeper study and applications to credit risk we refer to Gehmlich & Schmidt (2018).

1.3 *Brownian motion*

Definition 2. A continuous, adapted process W is called *Brownian motion* if

(i) $W_0 = 0$,

(ii) $E[W_t] = 0$ and $\operatorname{Var}(W_t) < \infty$, for all $t \geq 0$,

(iii) $W_t - W_s$ is independent of $\mathscr{F}_s$.

One can show that

$$W_t - W_s \sim \mathcal{N}(0, t - s),$$

i.e. the increments are even normally distributed. If we choose a time-change T appropriately, we can even construct a Poisson process as time-changed Brownian motion.

1.4 Classes of stochastic processes

To a stochastic process X we associate the mappgin

$$\widehat{X}\colon \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{R}^d.$$

This allows us to consider the stochastic process as a simple random variable on the product space $\Omega \times \mathbb{R}_{\geq 0}$. In particular, measurability can be considered with respect to $\mathscr{F} \otimes \mathscr{B}(\mathbb{R}_{\geq 0})$, which however lacks the link to the filtration.

Definition 3. (i) X is called *progressivly measurable*, if for all $t \geq 0$ the mapping

$$\begin{aligned} \Omega \times [0,t] &\to \mathbb{R}^d \\ (\omega, s) &\mapsto X_s(\omega) \end{aligned}$$

is $\mathscr{F}_t \otimes \mathscr{B}\big([0,t]\big)$-$\mathscr{B}(\mathbb{R}^d)$-measurable.

(ii) The *optional* σ-algebra $\mathscr{O}$ is the σ-algebra on $\Omega \times \mathbb{R}_{\geq 0}$, generated by adapted, càdlàg-processes. X is called *optional*, if $\widehat{X}$ is $\mathscr{O}$-measurable.

(iii) The *predictable* σ-algebra $\mathscr{P}$ is the σ-algebra on $\Omega \times \mathbb{R}_{\geq 0}$, generated by adapted, càg-processes. X is called *predictable*, if $\widehat{X}$ is $\mathscr{P}$-measurable.

In particular we obtain the following inclusions: predictable $\Rightarrow$ optional $\Rightarrow$ progressive $\Rightarrow$ adapted. A classical example is that for a progressive process X, $X^* = \sup_{s \leq \cdot} X_s$ is optional. Moreover, if we denote by X^T the process stopped at the stopping time T, then the following properties are kept while stopping: adapted, predictable, optional, progressive.

For the reverse consider an adapted process X. If X is càd, then X is progressive. If it is càg, then it is optional. If it is càdlàg, then X_- and $\Delta X = X - X_-$ are optional. For the following result, see the almost sure blog (see `https://almostsuremath.com/2016/11/15/optional-processes/`.)

Lemma 4. *Consider an adapted process X which is làd. Assume that X is càd everywhere except of a countable set $S \subset \mathbb{R}_{\geq 0}$. Then X is optional.*

We will often study random intervals, defined for two random times S and T by

$$[\![S,T]\!] := \{(\omega, t) \in \Omega \times \mathbb{R}_{\geq 0} : S(\omega) \leq t \leq T(\omega)\}.$$

As above we can call the interval optional or predictable if it is $\mathscr{O}$ resp. $\mathscr{P}$-measurable.

1.5 Localization

If we have a property $\mathscr{C}$ of a class of properties, then we introduce the localised class $\mathscr{C}_{loc}$ by all those processes X for which it holds there exists a sequence of stopping times $T_n \to \infty$ such that $X^{T_n} \in \mathscr{C}$ for all n. The sequence (T_n) is called localising sequence.

Definition 5. (i) A martingale X is uniformly integrable, if the family $(X_t)_{t \geq 0}$ is *uniformly integrable*. We denote by $\mathscr{M}$ the class of all uniformly integrable martingales.

(ii) A martingale X is called *square integrable*, if $\sup_{t \geq 0} E[X_t^2] < \infty$. This class is denoted by $\mathscr{H}^2$.

(iii) A process in $\mathscr{M}_{loc}$ is called *local martingale* and a process in $\mathscr{H}^2_{loc}$ *locally square integrable*.

Definition 6. A function $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ has *locally finite variation*, if

$$\operatorname{Var}(f)_t := \sup_{0\leq t_0 \dots t_n \leq t} \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| < \infty$$

for all $t \geq 0$. A process is called of locally finite variation, if it has paths of locally finite variation.

Let $\mathscr{V}^+ :=$ denotes the increasing càdlàg process A, with $A_0 = 0$,

$$\begin{aligned}
\mathscr{V} &:= \mathscr{V}^+ - \mathscr{V}^+, \\
\mathscr{A}^+ &:= \{A \in \mathscr{V}^+ \colon E[A_\infty] < \infty\}, \\
\mathscr{A} &:= \mathscr{A}^+ - \mathscr{A}^+ = \{A \in \mathscr{V} \colon E[\operatorname{Var}(A)_\infty] < \infty\}.
\end{aligned}$$

Then $\mathscr{V}$ is the set of all adapted processes of locally finite variation. For each $A \in \mathscr{V}$ we can associate $t \mapsto A_t(\omega)$ with a signed measure, denoted by $dA_t(\omega)$. Then we can define (pathwise) for optional processes H,

$$\left(H \cdot A\right)_t(\omega) = \begin{cases} \int_0^t H_s dA_s & \text{falls } \int_0^t |H_s| d\operatorname{Var}(A)_s < \infty \\ \infty & \text{sonst.} \end{cases}$$

We obtained the following theorem.

Theorem 7 (Integral of finite variation processes)**.** *Let $A \in \mathscr{V}(\mathscr{V}^+)$ and $H \geq 0$ be optional, such that $B = H \cdot A < \infty$. Then $B \in \mathscr{V}(\mathscr{V}^+)$. If H and A are predictable, so is B.*

We also obtained the important result that the only predictable local martingale with finite variation is $M = 0$ (Recall that the Brownian motion is of course predictable).

Theorem 8 (Dual predictable projection)**.** *Consider $A \in \mathscr{A}^+_{loc}$. Then there is a unique predictable process $A^p \in \mathscr{A}^+_{loc}$, satisfying one of the following, equivalent properties*

(i) $A - A^p \in \mathscr{M}_{loc}$,

(ii) $E[A^p_T] = E[A_T]$ *for all stopping times T,*

(iii) $E\big[(H \cdot A)_\infty)\big] = E\big[(H \cdot A^p)_\infty)\big]$ *for all predictable $H \geq 0$.*

One can also directly project on the predictable σ-algebra, but here we have a more versatile tool, the *dual* predictable projection. It helps us to generate local martingales, which is of course very importance to classify absence of arbitrage.

As an example, you might want to check for an extended Poisson process X, $X^p = \Lambda$.

1.6 Semimartingales

We can now define the set of all square integrable martingales by

$$\mathscr{H}^2 = \{M \in \mathscr{M} \colon \sup_{t\geq 0} E[X_t^2] < \infty\}$$

.

Proposition 9 (Predictable covariation). *Let $M, N \in \mathscr{H}^2_{loc}$. Then there exists a unique predictable process $\langle M, N\rangle \in \mathscr{V}$, s.t.*

$$MN - \langle M, N\rangle \in \mathscr{M}_{loc}\,.$$

If $M, N \in \mathscr{H}^2$, then $\langle M, N\rangle \in \mathscr{A}$ and $MN - \langle M, N\rangle \in \mathscr{M}$.

The process $\langle M, N\rangle$ is called predictable covariation and $\langle M\rangle = \langle M, M\rangle$ (predictable) quadratic variation.

Now you could show that a Wiener process with $\sigma^2(t) = \mathrm{Var}(W_t)$ is a continuous, square-integrable martingale with $\langle W\rangle = \sigma^2(t)$.

Mit $\mathscr{L}$ bezeichnen wir die Teilmenge von $\mathscr{M}_{loc}$ für die $M_0 = 0$ gilt

Definition 10. (i) If the process X can be decomposed as

$$X = X_0 + M + A \tag{11}$$

with $M \in \mathscr{L}$ and $A \in \mathscr{V}$, then X is called a *semimartingale*. By $\mathscr{S}$ we denote the space of semimartingales.

(ii) If A is predictable, the decomposition in (11) is unique and we call X special. The space of special semimartingales is denoted by $\mathscr{S}_p$.

If a semimartingale is continuous, it is special and M and A in its decomposition are continuous. If a semimartingale has bounded jumps, it is also special. So the issue of not being a special semimartingale arises from the large jumps.

We even can show a little bit more: for every semimartingale, there exists the decomposition

$$X = X_0 + X^c + M + A$$

with a continuous local martingale X^c and a purely discontinuous local martingale $M \in \mathscr{H}^2_{loc}$ and $A \in \mathscr{V}$.

1.7 *The stochastic integral*

We call H simple, if

$$H = Y\mathbb{1}_{[\![0]\!]} \quad \text{oder} \quad H = Y\mathbb{1}_{]\!]S,T]\!]}$$

with stopping times S und T and bounded, $\mathscr{F}_S$-measurable Y. These are the prototypes of simple processes, where it is clear how to integrate them. Indeed, let us define for simple H its stochastic integral $H \cdot X$ with respect to a stochastic process X by

$$(H \cdot X)_t := \begin{cases} 0 & \text{if } H = Y\mathbb{1}_{[\![0]\!]} \\ Y \cdot (X_{T\wedge t} - X_{S\wedge t}) & \text{otherwise.} \end{cases} \tag{12}$$

By $\mathscr{E}$ we denote the space of simple (elementary) processes.

Theorem 13 (The stochastic integral)**.** *Let X be a semimartingale. The mapping $H \mapsto H \cdot X$ has an extension from $\mathscr{E}$ to the space of locally bounded, predictable processes, such that*

(i) $H \cdot X$ *is adapted and càdlàg,*

(ii) $H \mapsto H \cdot X$ *is linear,*

(iii) if predictable (H^n) converge pointwise to H, and is $|H^n| \leq K$ for a locally bounded, predictable process K, then

$$(H^n \cdot X)_t \xrightarrow{P} (H \cdot X)_t \qquad \forall\, t > 0.$$

We obtained the following properties:

(i) $H \cdot X$ is again a semimartingale.

(ii) If X is a local martingale, so is $H \cdot X$.

(iii) If $X \in \mathscr{V}$, then $H \cdot X$ is the Lebesgue-Stieltjes integral.

(iv) $(H \cdot X)_0 = 0$ and $H \cdot (X - X_0) = H \cdot X$.

(v) $K \cdot (H \cdot X) = (KH) \cdot X$.

(vi) $\Delta(H \cdot X) = H \cdot \Delta X$.

(vii) Is T predictable and Y $\mathscr{F}_T$-messbar, then

$$(Y \mathbb{1}_{[\![T]\!]}) \cdot X = Y \cdot \Delta X_T \mathbb{1}_{[\![T,\infty[\![}$$

If X is even locally square integrable we can allow a larger class of integrands.

Theorem 14. *Let $X \in \mathscr{H}^2_{loc}$. Then, $H \mapsto H \cdot X$ has an extension from $\mathscr{E}$ to L^2_{loc} such that*

(i) $H \cdot X \in \mathscr{H}^2_{loc}$

(ii) $H \in L^2(X) \iff H \cdot X \in \mathscr{H}^2$

(iii) For $X, Y \in \mathscr{H}^2_{loc}$ and predictable $K, M \in L^2_{loc}(X)$,

$$\langle H \cdot X, K \cdot Y \rangle = HK \cdot \langle X, Y \rangle .$$

For two semimartingales $X, Y \in \mathscr{S}$ we can define the quadratic covariation of X and Y by

$$[X, Y] = XY - X_0 Y_0 - X_- \cdot Y - Y_- \cdot X . \tag{15}$$

And we showed a number of properties: Consider $X, X' \in \mathscr{S}$ and $Y \in \mathscr{V}$. Then

(i) $[X, X'] \in \mathscr{V}$ and $[X] \in \mathscr{V}^+$,

(ii) $[X, Y] = \Delta X \cdot Y$,

(iii) if Y is predictable, then $[X, Y] = \Delta Y \cdot X$,

(iv) if either X or Y is continuous, then $[X, Y] = 0$.

(v) $[X, X']_t = \langle X, X' \rangle_t + \sum_{s \leq t} \Delta X_s \Delta X'_s$.

1.8 The Itô-formula

One major result was the following result on the semimartingale property of twice differentiable functions of semimartingales.

Theorem 16 (Itô-Formula)**.** *Consider a d-dimensional semimartingale* $X = (X^1, \dots, X^d)$ *and* $f \in \mathscr{C}^2(\mathbb{R}^d, \mathbb{R})$*. Then,* $f(X) \in \mathscr{S}$ *and*

$$\begin{aligned} f(X) = f(X_0) &+ \sum_{i \le d} D_i f(X_-) \cdot X^i \\ &+ \frac{1}{2} \sum_{i,j \le d} D_{ij} f(X_-) \cdot \langle X^{i,c}, X^{j,c} \rangle \\ &+ \sum_{0 \le s \le \cdot} \Big(f(X_s) - f(X_{s-}) - \sum_{i \le d} D_i f(X_{s-}) \Delta X_s^i \Big). \end{aligned} \tag{17}$$

As a first application we considered stochastic exponentials. Here Y was called a *sotchastic exponential*, if $X \in \mathscr{S}$ and

$$Y = 1 + Y_- \cdot X. \tag{18}$$

We denote the solution of (18) by $Y = \mathscr{E}(X)$.

As an important example we have met the geometric Brownian motion, $\mathscr{E}(W)$. If W is a standard Brownian motion, then

$$\mathscr{E}(W)_t = \exp\Big(W_t - \frac{1}{2}t\Big), \qquad t \ge 0.$$

Theorem 19. *Consier* $X \in \mathscr{S}$*. Then there exists a unique solution of (18) given by*

$$\mathscr{E}(X)_t := Y_t = \prod_{0 < s \le t} (1 + \Delta X_s) e^{-\Delta X_s} \cdot \exp\Big(X_t - X_0 - \frac{1}{2}\langle X^c \rangle_t\Big), \qquad t \ge 0.$$

1.9 Girsanovs theorem

We have already seen that measure changes are of prime importance in financial mathematics. The key tool here is Girsanovs theorem. Define for a stopping time T

$$P_T := P|_{\mathscr{F}_T}.$$

P' is called *locally absolutely continuous* w.r.t. P, if

$$P'_t \ll P_t, \qquad \forall t \ge 0;$$

which we denote by $P' \overset{loc}{\ll} P$. This is even equivalent to our localisation procedure (there exist stopping times $T_n \to \infty$ for which $P'_{T_n} \ll P_{T_n}, \qquad \forall n$.

Theorem 20. *Let* $P' \overset{loc}{\ll} P$*. Then there exists a unique P-martingale* $Z \ge 0$*, such that*

$$Z_t = \frac{dP'_t}{dP_t}, \quad t \ge 0. \tag{21}$$

Z is called density of P' w.r.t. P and $E[Z_t] = 1$ for all $t \geq 0$. If $P' \ll P$, then Z is uniformly integrable and

$$Z_\infty = \frac{dP'}{dP}.$$

A typical example is a geometric Brownian motion

$$Z_t = e^{aW_t - \frac{a^2 t}{2}}, \qquad t \geq 0,$$

which however is *not* uniformly integrable!

Theorem 22 (Girsanov)**.** *Let $P' \overset{loc}{\ll} P$ with density Z. Consider $M \in \mathscr{M}_{loc}(P)$ with $M_0 = 0$. Then*

$$M' = M - \frac{1}{Z} \cdot [M, Z]$$

is P'-almost surely well-defined and a P'-local martingale. If $[M, Z] \in \mathscr{A}_{loc}$, then

$$M'' = M - \frac{1}{Z_-} \langle M, Z \rangle$$

is a P'-local martingale.

As a typical application we consider

$$Z_t = \exp\Big(\theta W_t - \frac{\theta^2 t}{2}\Big), \qquad 0 \leq t \leq T,$$

hence $Z_t = Z_0 + Z_- \cdot \theta W_t$. Then

$$\begin{aligned} M'_t &= W_t - \frac{1}{Z} \cdot [W, Z]_t \\ &= W_t - \frac{1}{Z} \cdot \theta Z d\langle W \rangle_t = W_t - \theta t \end{aligned}$$

is a local martingale. Since quadratic variation is not changed by the measure change, $(W_t - \theta t)_{0 \leq t \leq T}$ is a Brownian motion under P'.

1.10 Semimartingale characteristics

With the above tools we have a good access to jump-diffusion, i.e. processes of the type

$$dX_t = \mu_t dt + \sigma_t dW_t + \kappa_t dJ_t.$$

We note that the distributional characteristics depend on the drift μ, the volatility σ and the jump term κ. Semimartingale characteristics generalise this notion to the full generality of semimartingale processes.

To describe the jumps of the semimartingale in a precise way, we utilize the concept of random measure. A *random measure* on $\mathbb{R}_{\geq 0} \times \mathbb{R}^d$ is a family $\mu = (\mu(\omega; dt, dx) \colon \omega \in \Omega)$ of non-negative measures on $(\mathbb{R}_{\geq 0} \times \mathbb{R}^d, \mathscr{B}(\mathbb{R}_{\geq 0}) \otimes \mathscr{B}(\mathbb{R}^d))$ such that $\mu(\omega; \{0\} \times \mathbb{R}^d) = 0$ for all $\omega \in \Omega$.

To this end, we introduce

$$\widetilde{\Omega} = \Omega \times \mathbb{R}_{\geq 0} \times \mathbb{R}^d \tag{23}$$

with σ-fields $\widetilde{\mathscr{O}} = \mathscr{O} \otimes \mathscr{B}(\mathbb{R}^d)$ and $\widetilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{B}(\mathbb{R}^d)$. In principle, one can replace $\mathbb{R}^d$ be a general Polish space, but we will not need this level of generality here.

A function W on $\widetilde{\Omega}$ is called optional (predictable) when it is $\widetilde{\mathscr{O}}$ ($\widetilde{\mathscr{P}}$)-measurable. Now we define the integral

$$W * \mu_t = \int_{[0,t]\times E} W(\omega, s, x)\mu(\omega; ds, dx)$$

whenever the integral is finite and $+\infty$ otherwise.

The random measure μ is called optional (predictable) if $W * \mu$ is optional (predictable) for every optional (predictable) function W.

The Doob-Meyer decomposition can also be extended to this processes by the following tool: for every optional $\mathscr{P}$-σ-finite random measure μ there exists a unique, predictable random measure ν^p such that for every $\widetilde{\mathscr{P}}$-measureable W such that $|W| * \mu \in \mathscr{A}^+_{loc}$, $|W| * \mu^p \in \mathscr{A}^+_{loc}$

$$W * (\mu - \mu^p) \qquad \text{is a local martingale.}$$

In that case there exists a predictable $A \in \mathscr{A}^+$ and a kernel $K(\omega, t; dx)$ such that

$$\mu^p(\omega; dt, tx) = K(\omega, t; dx)\, dA_t(\omega).$$

We call μ^p the (predictable) compensator of μ or the dual predictable projection of μ (compare Theorem 8).

To an adapted càdlàg process X we associate the integer-valued random measure

$$\mu^X(\omega; dt, dx) = \sum_s \mathbb{1}_{\{\Delta X_s(\omega)\neq 0\}} \delta_{(s, \Delta X_s(\omega))}(dt, dx).$$

As for semimartingales one can construct a stochastic integral with respect to the compensated random measure $\mu^X - (\mu^X)^p$.

We call h a truncation function if it bounded and satisfies $h(x) = x$ in a neighbourhood of 0. For a semimartingale X we introduce

$$\begin{aligned} \check{X}(h) &= \sum_{s\leq \cdot} \big(\Delta X_s - h(\Delta X_s)\big), \\ X(h) &= X - \check{X}(h) \end{aligned} \tag{24}$$

the process $X(h)$ with truncated, in particular bounded, jumps. Then $X(h)$ is special and hence it may be decomposed, see Equation (11) uniquely into

$$X(h) = X_0 + B(h) + M(H). \tag{25}$$

Definition 26. The triplet (B, C, ν) is called *characteristics* of the semimartingale X for the truncation function h where

(i) $B = B(h)$ in decomposition (25),
(ii) $C = C^{ij}$ with $C^{ij} = \langle X^{i,c}, X^{j,c}\rangle$,
(iii) ν is the compensator of μ^X.

Proposition 27. *One can find a version of the characteristics of X such that*

$$\begin{aligned} B^i &= b^i \cdot A, \\ C^{ij} &= c^{ij} \cdot A, \\ \nu(\omega; dt, dx) &= K(\omega, t; dx) dA_t(\omega) \end{aligned}$$

where

(i) A is predictable and $A \in \mathscr{A}^+_{loc}$ *which may be chosen continuous of X is quasi-left-continuous;*
(ii) $b = (b^i)$ *is d-dimensional and predictable,*
(iii) $c = (c^{ij})$ *is predictable with values in the set of symmetric nonnegative matrices,*
(iv) K is a transition kernel, such that

$$\begin{aligned} &K(\{0\}) = 0, \quad \int (|x|^2 \wedge 1) K(\omega, t; dx) < \infty; \\ &\Delta A_t(\omega) > 0 \;\Rightarrow\; b_t(\omega) = \int h(x) K(\omega, t; dx) \\ &\Delta A_t(\omega) K(\omega, t; \mathbb{R}^d) \leq 1. \end{aligned} \tag{28}$$

We then have a nice version of the Itô-formula: for each bounded $f \in C^2$, the following process is a local martingale

$$\begin{aligned} f(X) - f(X_0) - \sum D_j f(X_-) \cdot B^j - \frac{1}{2} \sum D_{jk} f(X_-) \cdot C^{jk} \\ - \big(f(X_- + x) - f(X_-) - \sum_j D_j f(X_-) h^j(x) \big) * \nu. \end{aligned} \tag{29}$$

This is is also sufficient for X being a semimartingale with characteristics (B, C, ν).

2 *Affine semimartingales*

We recall from our lecture on stochastic processes that there are a number of interesting affine processes, in particular those satisfying

$$dX_t = b_0 + b_1 X_t dt + \sqrt{a_0} dW_t$$
$$dX_t = b_0 + b_1 X_t dt + \sqrt{a_1 X_t} dW_t.$$

For those processes we showed that

$$E[e^{iuX_t} | X_0 = x] = \exp\Big(\phi(t,u) + \psi(t,u) \cdot x\Big)$$

with functions ϕ and ψ solving some Riccatti equations. Our aim is now to extend this to the semimartingale level.

Note that it is clear, that there are also affine processes in discrete time, which certainly do not have a jump-compensator of dt-type. We hence should expect that dt can be replaced by a more general dA_t - but what kind of properties does this process have ? Moreover, we can now longer expect a homogenous affine process (where the functions ϕ and ψ do only depend on the length of the considered interval and not on the place of the interval). We consider a state space $D \subset \mathbb{R}^d$ which is a closed convex cone of dimension d together with a filtered probability space, as previously, satisfying the usual conditions. Moreover, we define the *complex dual cone* of the state space D by

$$\mathcal{U} := \{u \in \mathbb{C}^d : \langle \Re u, x\rangle \le 0 \text{ for all } x \in D\}. \tag{30}$$

An important example is the set $\mathbb{R}^m_{\ge 0} \times \mathbb{R}^n$ with $m + n = d$, which we call s the 'canonical state-space'.

Definition 31. Let X be a càdlàg semimartingale, taking values in D. The process X is called an *affine semimartingale*, if there exist $\mathbb{C}$ and $\mathbb{C}^d$-valued deterministic functions $\phi_s(t,u)$ and $\psi_s(t,u)$, continuous in $u \in \mathcal{U}$ and with $\phi_s(t,0) = 0$ and $\psi_s(t,0) = 0$, such that

$$E\big[e^{\langle u, X_t\rangle} | \mathscr{F}_s\big] = \exp\big(\phi_s(t,u) + \langle \psi_s(t,u), X_s\rangle\big) \tag{32}$$

for all $0 \le s \le t$ and $u \in \mathcal{U}$. Moreover, X is called *time-homogeneous*, if $\phi_s(t,u) = \phi_0(t-s,u)$ and $\psi_s(t,u) = \psi_0(t-s,u)$, again for all $0 \le s \le t$ and $u \in \mathcal{U}$.

Condition 33. We say that an affine semimartingale X has *support of full convex span*, if

$$\operatorname{conv}(\operatorname{supp}(X_t)) = D, \qquad \text{for all } t > 0.$$

Under Condition 33, ϕ and ψ are uniquely specified:

Lemma 34. *Let X be an affine semimartingale satisfying the support condition 33. Then $\phi_s(t,u)$ and $\psi_s(t,u)$ are uniquely specified by (32) for all $0 < s \le t$ and $u \in \mathcal{U}$.*

Proof. Fix $0 < s \le t$ and suppose that $\widetilde{\phi}_s(t,u)$ and $\widetilde{\psi}_s(t,u)$ are also continuous in $u \in \mathcal{U}$ and satisfy (32). Write $p_s(t,u) := \widetilde{\phi}_s(t,u) - \phi_s(t,u)$ and $q_s(t,u) := \widetilde{\phi}_s(t,u) - \phi_s(t,u)$. Due to (32) it must hold that

$$p_s(t,u) + \langle q_s(t,u), X_s\rangle \quad \text{takes values in} \quad \{2\pi i k : k \in \mathbb{N}\} \quad a.s.\ \forall\, u \in \mathcal{U}.$$

However, the set $\mathcal{U}$ is simply connected, and hence its image under a continuous function must also be simply connected. It follows that $u \mapsto p_s(t,u) + \langle q_s(t,u), X_s\rangle$ is constant on $\mathcal{U}$ and therefore equal to $p_s(t,0) + \langle q_s(t,0), X_s\rangle = 0$. Hence,

$$p_s(t,u) + \langle q_s(t,u), x\rangle = 0,$$

for all $x \in \operatorname{supp}(X_s)$ and $u \in \mathcal{U}$. Taking convex combinations, the equality can be extended for $x \in D$. Since D has full linear span, we conclude that $p_s(t,u) = 0$ and $q_s(t,u) = 0$ for all $u \in \mathcal{U}$, completing the proof. □

Lemma 35. *Let X be an affine semimartingale satisfying the support condition 33. Then,*

(i) the function $u \mapsto \phi_s(t,u)$ maps $\mathcal{U}$ to $\mathbb{C}_{\leq 0}$ and $u \mapsto \psi(t,u)$ maps $\mathcal{U}$ to $\mathcal{U}$, for all $0 < s \leq t$,

(ii) ϕ and ψ satisfy the semi-flow property, *i.e. for all $0 < s \leq r \leq t$ and $u \in \mathcal{U}$,*

$$\begin{aligned} \phi_s(t,u) &= \phi_r(t,u) + \phi_s(r,\psi_r(t,u)), \qquad & \phi_t(t,u) &= 0 \\ \psi_s(t,u) &= \psi_s(r,\psi_r(t,u)), & \psi_t(t,u) &= u. \end{aligned} \tag{36}$$

Proof. To show the first property, recall that by Equation (32) we have

$$E\left[e^{\langle u,X_t\rangle}|\mathcal{F}_s\right] = \exp\left(\phi_s(t,u) + \langle\psi_s(t,u),X_s\rangle\right) \tag{37}$$

for all $u \in \mathcal{U}$ and $0 \leq s \leq t$. Since $\langle \Re u, X_t\rangle \leq 0$, a.s., the left hand side is bounded by one in absolute value. Thus, also

$$\Re\phi_s(t,u) + \langle \Re\psi_s(t,u), X_s\rangle \leq 0, \ a.s.$$

and consequently

$$\Re\phi_s(t,u) + \langle \Re\psi_s(t,u), x\rangle \leq 0, \quad \text{for all } x \in \operatorname{supp}(X_s).$$

Taking arbitrary convex combinations of these inequalities and using that $\operatorname{conv}(\operatorname{supp}(X_s)) = D$ by Condition 33, we obtain that the inequality must in fact hold for all $x \in D$. Since D is a cone this implies that $\Re\phi_s(t,u) \leq 0$ and $\psi_s(t,u) \in \mathcal{U}$, proving (i).

To show the semi-flow equations we apply iterated expectations to the left hand side of (37), yielding

$$\begin{aligned} E\left[E\left[e^{\langle u,X_t\rangle}|\mathcal{F}_r\right]|\mathcal{F}_s\right] &= E\left[\exp\left(\phi_r(t,u) + \langle\psi_r(t,u),X_r\rangle\right)|\mathcal{F}_s\right] = \\ &= \exp\left(\phi_s(r,u) + \phi_s(r,\psi_r(t,u)) + \langle\psi_s(r,\psi_r(t,u)),X_s\rangle\right). \end{aligned}$$

Note that the exponent on the right hand side is continuous in u and that the same holds true for (37). By the same argument as in the proof of Lemma 34 we conclude that

$$\phi_s(t,u) + \langle\psi_s(t,u),x\rangle = \phi_s(r,u) + \phi_s(r,\psi_r(t,u)) + \langle\psi_s(r,\psi_r(t,u)),x\rangle\,,$$

for all $x \in D$. Since the linear hull of D is $\mathbb{R}^d$ the semi-flow equations (36) follow. Note that the terminal conditions $\psi_t(t,u) = u$ and $\phi_t(t,u) = 0$ are a simple consequence of $E\left[\exp(\langle u,X_t\rangle)|\,\mathcal{F}_t\right] = \exp(\langle u,X_t\rangle)$ and the uniqueness property from Lemma 34. □

Definition 38. Let A be a non-decreasing càdlàg function with continuous part A^c and jump points $J^A := \{t \geq 0 | \Delta A_t > 0\}$. Let $(\gamma,\beta,\alpha,\mu) = (\gamma_i,\beta_i,\alpha_i,\mu_i)_{i\in\{0,\dots,d\}}$ be functions such that $\gamma_0\colon \mathbb{R}_{\geq 0} \times \mathcal{U} \to \mathbb{C}$, $\bar{\gamma}\colon \mathbb{R}_{\geq 0} \times \mathcal{U} \to \mathbb{C}^d$, $\beta_i\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^d$, $\alpha_i\colon \mathbb{R}_{\geq 0} \to \mathcal{S}^d$ and $(\mu_i(t,\cdot))_{t\geq 0}$ are families of (possibly signed) Borel measures on $D \setminus \{0\}$ with $\int_{D\setminus\{0\}}(1 + \|x\|^2)\mu_i(t,dx) < \infty$. We call $(A,\gamma,\beta,\alpha,\mu)$ a *good parameter set* if for all $i \in \{0,\dots,d\}$,

(i) α_i and β_i are locally integrable w.r.t. A^c,

(ii) for all compact sets $K \subset D \setminus \{0\}$, $\mu(\cdot,K)$ is locally A^c-integrable.

(iii) $\gamma(t,u) = 0$ for all $(t,u) \in (\mathbb{R}_{\geq 0} \setminus J^A) \times \mathcal{U}$.

Definition 39. An affine semimartingale is called *quasi-regular*, if the following holds:

(i) The functions ϕ and ψ are of finite variation in s and càdlàg in both s and t. More precisely, we assume that for all $(t,u) \in \mathbb{R}_{\geq 0} \times \mathcal{U}$

$$s \mapsto \phi_s(t,u) \quad \text{and} \quad s \mapsto \psi_s(t,u)$$

are càdlàg functions of finite variation on $[0,t]$, and for all $(s,u) \in \mathbb{R}_{\geq 0} \times \mathcal{U}$

$$t \mapsto \phi_s(t,u) \quad \text{and} \quad t \mapsto \psi_s(t,u)$$

are càdlàg functions on $[s,\infty)$.

(ii) For all $0 < s \leq t$, the functions

$$u \mapsto \phi_{s-}(t,u) \quad \text{and} \quad u \mapsto \psi_{s-}(t,u)$$

are continuous on $\mathcal{U}$.

Theorem 40. *Let X be a quasi-regular affine semimartingale satisfying the support condition 33. Then there exists a good parameter set $(A,\gamma,\beta,\alpha,\mu)$ such that the semimartingale characteristics (B,C,ν) of X w.r.t. the truncation function h satisfy, $\mathbb{P}$-a.s. for any $t>0$,*

$$B_t^c(\omega) = \int_0^t \big(\beta_0(s) + \sum_{i=1}^d X_{s-}^i(\omega)\beta_i(s)\big) dA_s^c \tag{41a}$$

$$C_t(\omega) = \int_0^t \big(\alpha_0(s) + \sum_{i=1}^d X_{s-}^i(\omega)\alpha_i(s)\big) dA_s^c \tag{41b}$$

$$\nu^c(\omega,ds,dx) = \big(\mu_0(s,dx) + \sum_{i=1}^d X_{s-}^i(\omega)\mu_i(s,dx)\big) dA_s^c \tag{41c}$$

$$\int_D \left(e^{\langle u,\xi\rangle} - 1\right) \nu(\omega,\{t\},d\xi) = \left(\exp\left(\gamma_0(t,u) + \sum_{i=1}^d \langle X_{t-}^i(\omega), \bar{\gamma}_i(t,u)\rangle\right) - 1\right). \tag{41d}$$

Moreover, for all $(T,u) \in (0,\infty) \times \mathcal{U}$, the functions ϕ and ψ are absolutely continuous w.r.t A and solve the following generalized measure Riccati equations: their continuous parts satisfy

$$\frac{d\phi_t^c(T,u)}{dA_t^c} = -F(t,\psi_t(T,u)), \tag{42}$$

$$\frac{d\psi_t^c(T,u)}{dA_t^c} = -R(t,\psi_t(T,u)), \tag{43}$$

dA^c-a.e., where

$$\begin{aligned} F(s,u) &= \langle \beta_0(s), u\rangle + \frac{1}{2}\langle u, \alpha_0(s)u\rangle + \int_D \left(e^{\langle x,u\rangle} - 1 - \langle h(x),u\rangle\right)\mu_0(s,dx) \\ R_i(s,u) &= \langle \beta_i(s), u\rangle + \frac{1}{2}\langle u, \alpha_i(s)u\rangle + \int_D \left(e^{\langle x,u\rangle} - 1 - \langle h(x),u\rangle\right)\mu_i(s,dx), \end{aligned} \tag{44}$$

while their jumps are given by

$$\begin{aligned} \Delta\phi_t(T,u) &= -\gamma_0(t,\psi_t(T,u)) \\ \Delta\psi_t(T,u) &= -\bar{\gamma}(t,\psi_t(T,u)), \end{aligned} \tag{45}$$

and their terminal conditions are

$$\phi_T(T,u) = 0 \quad \text{and} \quad \psi_T(T,u) = u. \tag{46}$$

Before we start with the proof, we visit some examples:

Example 47. Consider the following discrete-time variant of the (time-inhomogeneous) Poisson process: let $X_0 = x \in \mathbb{N}$. Furthermore, assume that X is constant except for $t \in \{1, 2, \dots\}$ and assume that $\Delta X_n \in \{0, 1\}$, $n \in \{1, 2, \dots\}$ are independent with $P(\Delta X_n = 1) = p_n \in (0, 1)$. Then X is an affine semimartingale because for $0 \le s \le t$,

$$E[e^{uX_t} | \mathcal{F}_s] = \exp\Big(uX_s + \sum_{s<n\le t, n\in\mathbb{N}} \phi_n(u)\Big)$$

where

$$\phi_n(u) = E[e^{u\Delta X_n}] = e^u(p_n + e^{-u}(1-p_n)) = \exp(u + \log(p_n + e^{-u}(1-p_n))).$$

Clearly, it may happen that $\Delta X_n = 0$ while $\phi(u, n, t) - \phi(u, n-, t) = \phi_n(u) \neq 0$. Stochastic discontinuity is reflected by having jumps at $t \in \{1, 2, \dots\}$ with positive probability. The considered process falls in the class of point processes whose associated jump measure is an *extended* Poisson measure. In contrast to Poisson processes, X is not quasi-left continuous. In summary, X is a process with independent increments, but not a time-inhomogeneous Lévy process. ◇

When we consider processes in discrete time we emphasize this by using a hat $\widehat{X} = (\widehat{X}_n)_{n\in\mathbb{N}}$.

Example 48 (AR(1))**.** A (time-inhomogeneous) autoregressive time series of order (1) is given by

$$\widehat{X}_n = \alpha(n)\widehat{X}_{n-1} + \varepsilon_n$$

where we assume that (ε_n) are independent (not necessarily identically nor normally distributed). Then, $\widehat{X}$ is affine, as

$$E[e^{uX_n} | \widehat{\mathcal{F}}_{n-1}] = E[e^{u\varepsilon_n}]e^{\alpha(n)X_{n-1}}$$

with $\widehat{\mathcal{F}}_{n-1} = \sigma(\widehat{X}_0, \dots, X_{n-1})$. The generalization to higher order requires an extension of the state space. So an AR(p) series gives an affine process $(\widehat{X}_n, \dots, \widehat{X}_{n-p})_{n\ge p}$. ◇

Analogously we obtain the class of discrete affine processes (Exercise). Are GARCH time series affine ?

2.1 *Proof of Theorem 40*

We start with some easy observations

Lemma 49. *Let X be a quasi-regular affine semimartingale. Then,*

$$E\big[e^{\langle u, X_{t-}\rangle} | \mathcal{F}_s\big] = \exp\big(\phi_s(t-, u) + \langle \psi_s(t-, u), X_s\rangle\big), \quad \forall\, 0 \le s < t, u \in \mathcal{U}. \tag{50}$$

$$E\big[e^{\langle u, X_t\rangle} | \mathcal{F}_{s-}\big] = \exp\big(\phi_{s-}(t, u) + \langle \psi_{s-}(t, u), X_{s-}\rangle\big), \quad \forall\, 0 < s \le t, u \in \mathcal{U}. \tag{51}$$

If in addition X has full support, it also holds that

$$E\big[e^{\langle u, \Delta X_t\rangle} | \mathcal{F}_{t-}\big] = \exp\big(-\Delta\phi_t(t, u) - \langle \Delta\psi_t(t, u), X_{t-}\rangle\big), \quad \forall\, (t, u) \in \mathbb{R}_{\ge 0} \times \mathcal{U}. \tag{52}$$

Lemma 53. *Let X be a quasi-regular affine semimartingale of full support and with characteristics (B, C, ν). For any $(t, u) \in (0, \infty) \times \mathcal{U}$,*

$$\int_D \Big(e^{\langle u, \xi\rangle} - 1\Big)\, \nu(\omega; \{t\}, d\xi) = \exp\Big(-\Delta\phi_t(t, u) - \langle \Delta\psi_t(t, u), X_{t-}\rangle\Big) - 1. \tag{54}$$

Proof. By definition, $\nu(\{t\}, d\xi)$ is the dual predictable projection of $\delta_{\Delta X_t}(d\xi)$ such that

$$\int_D \left(e^{\langle u,\xi\rangle} - 1\right) \nu(\omega; \{t\}, d\xi) = E\left[\left(e^{\langle u,\Delta X_t\rangle} - 1\right)\Big| \mathcal{F}_{t-}\right].$$

Combining with (52), the claim follows. □

Next, we consider the continuous parts of the semimartingale characteristics, and make the following definition:

$$G(dt,\omega,T,u) := \langle \psi_t, dB_t^c(\omega)\rangle + \frac{1}{2}\langle \psi_t, dC_t(\omega)\psi_t\rangle + \\ + \int_D \left(e^{\langle \psi_t,\xi\rangle} - 1 - \langle \psi_t, h(\xi)\rangle\right) \nu^c(\omega, dt, d\xi), \tag{55}$$

where we write $\psi_t := \psi_t(T,u)$ for short.

Lemma 56. *Let X be a quasi-regular affine semimartingale with a good version of its characteristics (B,C,ν), let $(T,u) \in (0,\infty) \times \mathcal{U}$ and let $G(dt,\omega,T,u)$ be the complex-valued random measure defined in* (55)*. It holds that*

$$G(dt;\omega,T,u) + d\phi_t^c(T,u) + \langle X_t(\omega), d\psi_t^c(T,u)\rangle = 0, \qquad \mathbb{P} - a.s, \tag{57}$$

as identity between measures on $[0,T]$.

Proof. For $(T,u) \in (0,\infty) \times \mathcal{U}$ consider the process

$$M_t^{u,T} := \mathbb{E}\left[e^{\langle u,X_T\rangle} \middle| \mathcal{F}_t\right] = \exp\left(\phi_t\left(T,u\right) + \langle \psi_t\left(T,u\right), X_t\rangle\right) \quad t \in [0,T),$$

which is a càdlàg martingale with the terminal value $M_T^{u,T} = \exp\left(\langle u, X_T\rangle\right)$. To alleviate notation we consider (T,u) fixed and write

$$M_t = M_t^{u,T} = \exp\left(\phi_t + \langle \psi_t, X_t\rangle\right),$$

with $\phi_t := \phi_t(T,u)$ and $\psi(t) := \psi_t(T,u)$. Applying the Itô-formula to M we obtain a decomposition

$$M_t = L_t + F_t,$$

where L is a local martingale and F is the predictable finite variation process

$$F_t := \int_0^t M_{s-}\left\{d\phi_s^c + \langle X_{s-}, d\psi_s^c\rangle + \langle \psi_{s-}, dB_s\rangle + \frac{1}{2}\langle \psi_{s-}, dC_s\psi_{s-}\rangle \right. \\ \left. + \int_D \left(e^{\Delta\phi_s + \langle \psi_s, X_{s-}+\xi\rangle - \langle \psi_{s-}, X_{s-}\rangle} - 1 - \langle \psi_{s-}, h(\xi)\rangle\right) \nu(\omega, ds, d\xi)\right\}. \tag{58}$$

The jump part ΔF vanishes due to Lemma 53, and we are left with the continuous part

$$F_t = F_t^c = \int_0^t M_{s-}\left\{d\phi_s^c + \langle X_{s-}, d\psi_s^c\rangle + \langle \psi_{s-}, dB_s^c\rangle + \frac{1}{2}\langle \psi_{s-}, dC_s\psi_{s-}\rangle \right. \\ \left. + \int_D \left(e^{\langle \psi_{s-},\xi\rangle} - 1 - \langle \psi_{s-}, h(\xi)\rangle\right) \nu^c(\omega, ds, d\xi)\right\}.$$

Recall that M is a martingale, and hence $M \equiv L$ and $F \equiv 0$ on $[0,T]$, $\mathbb{P}$-a.s. With (55), F can be rewritten as

$$F_t = \int_0^t M_{s-}\left\{d\phi_s^c + \langle X_{s-}, d\psi_s^c\rangle + G(ds;\omega,T,u)\right\}.$$

Since none of the measures appearing above charges points, the left limits X_{s-}, ψ_{s-} can be substituted by right limits X_s, ψ_s. Moreover, M_{s-} is nonzero everywhere and (57) follows. □

In order to make efficient use of the full-support condition (Definition 33), we introduce the following convention: Given an affine semimartingale X, a tuple $\mathbf{X} = (X^0, \dots, X^d)$ represents $d+1$ *stochastically independent* copies of X. Formally, the tuple $\mathbf{X}$ can be realized on the product space $(\Omega^{(d+1)}, \mathcal{F}^{\otimes(d+1)}, (\mathcal{F}_t^{\otimes(d+1)})_{t\geq 0})$ equipped with the associated product measure. Moreover, for any points $\xi_0, \dots, \xi_d$ in $\mathbb{R}^d$, we define the $(d+1) \times (d+1)$-matrix

$$H(\xi_0, \dots, \xi_n) := \begin{pmatrix} 1 & \xi_0^\top \\ \vdots & \vdots \\ 1 & \xi_n^\top \end{pmatrix}. \tag{59}$$

The matrix-valued process Θ_t is formed by inserting $\mathbf{X} = (X^0, \dots, X^d)$ into H, i.e. we set

$$\Theta_t(\omega) = H(X^0, \dots, X^d) = \begin{pmatrix} 1 & X_t^0(\omega)^\top \\ \vdots & \vdots \\ 1 & X_t^d(\omega)^\top \end{pmatrix}. \tag{60}$$

Lemma 61. *Let $s > 0$ and let X be an affine semimartingale with full support. Then there exists $\varepsilon > 0$ and a set $E \in \mathcal{F}_s$ with $P(E) > 0$, such that the matrices $\Theta_t(\omega)$ and $\Theta_{t-}(\omega)$ are regular for all $(t, \omega) \in (s, s+\varepsilon) \times E$.*

Proof. Define the first hitting time

$$\tau := \inf\{t > s : \Theta_t \text{ singular, or } \Theta_{t-} \text{ singular}\}.$$

Since the set of singular matrices is a closed subset of the vector space of $\mathbb{R}^{(d+1)\times(d+1)}$-matrices, τ is a stopping time. Moreover, by monotone convergence, we have

$$\lim_{n\to\infty} P\Big(\Theta_t \text{ and } \Theta_{t-} \text{ regular for all } t \in (s, s+1/n)\Big) = \lim_{n\to\infty} P(\tau \geq s + 1/n) = P(\tau > s).$$

If we can show that $P(\tau > s) > 0$, then the claim follows by choosing N large enough and setting $\varepsilon = 1/N$ and $E = \{\tau \geq s + 1/N\}$. But by right-continuity of X, the set $\{\omega : \tau(\omega) > s\}$ is equal to $\{\omega : \Theta_s(\omega) \text{ is regular}\}$ and it remains to show that Θ_s is regular with strictly positive probability. To this end, we use the full support condition $\mathrm{conv}(\mathrm{supp}(X_s)) = D$ to find $d+1$ convex independent points $\xi^0, \dots, \xi^d$ in $\mathrm{supp}(X_s)$. Recalling the definition of H in (59), it follows that $H(\xi^0, \dots, \xi^d)$ is regular. Since the set of regular matrices is open we find $\delta > 0$ such that even $H(y_0, \dots, y_d)$ is regular for all $y_i \in U_\delta(\xi_i), i \in \{0, \dots, d\}$, where $U_\delta(\xi_i)$ is the open ball of radius δ centered at ξ_i. Now, by independence of $X^0, \dots, X^d$, it follows that

$$\begin{aligned} P\left(\Theta_s \text{ is regular}\right) &\geq P\Big(X_s^i \in U_\delta(\xi_i) \quad \forall i \in \{0, \dots, d\}\Big) \\ &= \prod_{i=0}^d P\left(X_s \in U_\delta(\xi_i)\right). \end{aligned}$$

Since for each $i \in \{0, \dots, d\}$ the intersection of $U_\delta(\xi_i)$ with the support of X_s is non-empty, all probabilities are strictly positive, and the proof is complete. □

Similar to the $\mathbb{R}^{(d+1)\times(d+1)}$-valued process process $(\Theta_t)_{t\geq 0}$ defined in (60), we define $d+1$ independent copies of the complex-valued random measure $G(dt, \omega, T, u)$ from equation (55) and denote them by $G_0, \dots, G_d$, respectively. With this notation and for any $(T, u) \in \mathbb{R}_{\geq 0} \times \mathcal{U}$, the $d+1$ corresponding equations (57) can be written in matrix-vector form as

$$\Theta_t(\omega) \cdot \begin{pmatrix} d\phi_t^c(T, u) \\ d\psi_t^{c,1}(T, u) \\ \vdots \\ d\psi_t^{c,d}(T, u) \end{pmatrix} = - \begin{pmatrix} G_0(dt; \omega, T, u) \\ \vdots \\ G_d(dt; \omega, T, u) \end{pmatrix} \tag{62}$$

which holds $\mathbb{P}$-a.s. as an identity between complex-valued measures on $[0, T]$. The next Lemma gives a 'local' version of the continuous part of Theorem 40.

Lemma 63. *Let X be a quasi-regular affine semimartingale with full support and let $\tau \in (0, \infty)$ be a deterministic timepoint. Then there exists an interval $I_\tau := (\tau, \tau + \varepsilon)$, where $\varepsilon = \varepsilon(\tau) > 0$, and good parameters $(A^c, \beta, \alpha, \mu)$ on I_τ. With respect to these parameters, and with F and R as in* (44)*, the measure Riccati equations* (80) *and* (81) *hold true for each $(T, u) \in \mathbb{R}_{\geq 0} \times \mathcal{U}$ and $t \in I_\tau \cap [0, T]$.*

Remark 64. We emphasize that in this lemma the parameters $(A^c, \beta, \alpha, \mu)$ as well as the functions F and R may depend on τ.

Recall that for a semimartingale X there exists a càdlàg, increasing, predictable, $\mathbb{R}_{\geq 0}$-valued process $\widetilde{A}$ starting in 0 and with continuous part A^c, such that the semimartingale characteristics of X can be 'disintegrated' with respect to $\widetilde{A}$. For the continuous parts (B^c, C, ν^c) of the characteristics, this implies the representation

$$\begin{aligned} B_t^c &= \int_0^t b_s d\widetilde{A}_s^c \\ C_t &= \int_0^t c_s d\widetilde{A}_s^c \\ \nu^c(\omega, dt, dx) &= K_{\omega,t}(dx) d\widetilde{A}_t^c(\omega), \end{aligned} \tag{65}$$

where b and c are predictable processes and $K_{\omega,t}(dx)$ a transition kernel from $\Omega \times \mathbb{R}_{\geq 0}$, endowed with the predictable σ-algebra, to $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Proof. Let $X^0, \ldots, X^d$ be $d + 1$ stochastically independent copies of X. Denote the semimartingale characteristics of X^i by (B^i, C^i, ν^i) and define $G_i(\omega; t, T, u)$ as in (55), $i = 0, \ldots, d$. The semimartingale characteristics (B^i, C^i, ν^i) can be disintegrated as in (65). Since we consider only a finite collection of semimartingales, we may assume that the process $\widetilde{A}_s^c(\omega)$ is the same for each X^i.

By Lemma 61, there exists an interval $I_\tau = (\tau, \tau + \varepsilon)$, $\varepsilon > 0$, and a set $E \in \mathcal{F}$ with $P(E) > 0$ and such that $\Theta_t(\omega)$ is invertible for all $(t, \omega) \in I_\tau \times E$. Multiplying (62) from the left with the inverse of this matrix yields

$$\begin{pmatrix} d\phi_t^c(T, u) \\ d\psi_t^{c,1}(T, u) \\ \vdots \\ d\psi_t^{c,d}(T, u) \end{pmatrix} = -\Theta_t(\omega)^{-1} \cdot \begin{pmatrix} G_0(dt; \omega, T, u) \\ \vdots \\ G_d(dt; \omega, T, u) \end{pmatrix}, \tag{66}$$

as an identity between complex-valued measures on I_τ for all $\omega \in E$. Since $P(E) > 0$, we can choose some particular $\omega_* \in E$ where (66) holds. Setting

$$A_t^c := \widetilde{A}_t^c(\omega_*), \qquad t \in I_\tau$$

we observe that $G_i(dt; \omega_*, T, u) \ll dA_t^c$ for each $i \in \{0, \ldots, d\}$ and conclude that also the left hand side of (66) is absolutely continuous with respect to A^c on I_τ. Denote by (b^i, c^i, K^i) the disintegrated semi-martingale characteristics of X^i, as in (65). Note that the random measures $G_i(dt; \omega, T, u)$ depend linearly on (b^i, c^i, K^i), which in light of (66) suggests to apply the linear transformation $\Theta_t(\omega)^{-1}$ directly to the disintegrated semimartingale characteristics. Evaluating at ω_*, we hence define the *deterministic* functions $(\beta^i, \alpha^i, \mu^i)_{i\in\{0,\ldots,d\}}$ on I_τ by setting

$$\begin{aligned}
\left(\beta^0, \beta^1, \ldots, \beta^d\right)_t^\top &:= \Theta_{t-}(\omega_*)^{-1} \cdot \left(b^0, b^1, \ldots, b^d\right)_t^\top(\omega_*) \\
\left(\alpha^0_{kl}, \alpha^1_{kl}, \ldots, \alpha^d_{kl}\right)_t^\top &:= \Theta_{t-}(\omega_*)^{-1} \cdot \left(c^0_{kl}, c^1_{kl}, \ldots, c^d_{kl}\right)_t^\top(\omega_*), \quad k,l \in \{1,\ldots,d\} \\
\left(\mu^0, \mu^1, \ldots, \mu^d\right)_t^\top &:= \Theta_{t-}(\omega_*)^{-1} \cdot \left(K^0, K^1, \ldots, K^d\right)_t^\top(\omega_*).
\end{aligned}$$

Using these parameters, the functions F, R can be defined on I_τ as in (44). In combination with (66) it follows that

$$\begin{pmatrix} d\phi^c_t(T,u) \\ d\psi^{c,1}_t(T,u) \\ \vdots \\ d\psi^{c,d}_t(T,u) \end{pmatrix} = -\Theta_t(\omega_*)^{-1} \cdot \begin{pmatrix} G_0(dt;\omega_*,T,u) \\ \vdots \\ G_d(dt;\omega_*,T,u) \end{pmatrix} = -\begin{pmatrix} F(t,\psi_t(T,u)) \\ R^1(t,\psi_t(T,u)) \\ \vdots \\ R^d(t,\psi_t(T,u)) \end{pmatrix} dA^c_t \tag{67}$$

for $t \in I_\tau \cap [0,T]$, which yields validity of the Riccati equations (80) and (81) on I_τ. □

Proof of Thm. 40. We consider first the continuous parts of the Riccati equations, and thereafter treat their jumps. Applying Lemma 63 to each $\tau \in (0,\infty)$ we obtain a family of intervals I_τ, each with non-empty interior I°_τ, such that $(I^\circ_\tau)_{\tau\in(0,\infty)}$ is an open cover of the positive half-line $(0,\infty)$. Since $\mathbb{R}_{\geq 0}$ can be exhausted by compact sets such a cover has a countable subcover $\mathcal{S}$.

To each interval $I \in \mathcal{S}$, Lemma 63 associates good parameters $(A^{c,I}, \beta^I, \alpha^I, \nu^I)$. By countability of $\mathcal{S}$ there exists a continuous common dominating function $A^c : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $A^{c,I} \ll A^c$ for all $I \in \mathcal{S}$.

We remark that passing from $A^{c,I}$ to A^c has merely the effect of multiplying all parameters with the Radon-Nikodym derivative $\frac{dA^{c,I}}{dA^c}$. Hence, we may assume without loss of generality that $A^{c,I} = A^c$ for each $I \in \mathcal{S}$.

Let now I and $\widetilde{I}$ be two intervals with non-empty intersection, taken from the countable subcover $\mathcal{S}$. Denote by $(A^c, \beta, \alpha, \mu)$ and $(A^c, \widetilde{\beta}, \widetilde{\alpha}, \widetilde{\mu})$ the respective parameter sets obtained for these intervals by application of Lemma 63 and by (F, R) and $(\widetilde{F}, \widetilde{R})$ the corresponding functions defined by (44). We say that these two parameter sets are *compatible* if they agree (up to a dA^c_t-nullset) on the intersection $I \cap \widetilde{I}$. Once we have shown compatibility for arbitrary intervals I and $\widetilde{I}$ it is clear that we can find a single good parameter set (A, β, α, μ), defined on the whole real half-line $\mathbb{R}_{\geq 0}$, such that the Riccati equations (80) and (81) hold true. To condense notation, we introduce the vectors

$$d\Psi^c_t(T,u) := \begin{pmatrix} d\phi^c_t(T,u) \\ d\psi^{c,1}_t(T,u) \\ \vdots \\ d\psi^{c,d}_t(T,u) \end{pmatrix}, \quad \mathcal{R}(t,u) := \begin{pmatrix} F(t,u) \\ R^1(t,u) \\ \vdots \\ R^d(t,u) \end{pmatrix}, \qquad \widetilde{\mathcal{R}}(t,u) := \begin{pmatrix} \widetilde{F}(t,u) \\ \widetilde{R}^1(t,u) \\ \vdots \\ \widetilde{R}^d(t,u) \end{pmatrix}.$$

Applying equation (67) once on the interval I and once on $\widetilde{I}$ yields

$$\mathcal{R}(t,\psi_t(T,u))dA^c_t = d\Psi^c_t(T,u) = \widetilde{\mathcal{R}}(t,\psi_t(T,u))dA^c_t, \quad t \in I \cap \widetilde{I} \cap [0,T]. \tag{68}$$

Let now $\mathcal{T} \times \mathcal{E}$ be a countable dense subset of $\mathbb{R}_{\geq 0} \times \mathcal{U}$. Taking the union over the countable set $\mathcal{T} \times \mathcal{E}$ we obtain from (68) that

$$\mathcal{R}(t,\psi_t(T,u)) = \widetilde{\mathcal{R}}(t,\psi_t(T,u)) \quad \text{for all } (T,u) \in \mathcal{T} \times \mathcal{E} \text{ and } t \in (I \cap \widetilde{I} \cap [0,T]) \setminus N, \tag{69}$$

where N is a dA^c_t-nullset, independent of (T,u).

The next step is to 'evaluate' (69) at $T = t$ and to use that $\psi_t(t, u) = u$ by taking limits in the countable set $\mathcal{T}$. Observe that as functions of Lévy-Khintchine-form (cf. (44)) both F and R are continuous in u. By denseness of $\mathcal{T}$ in $\mathbb{R}_{\geq 0}$ we can find a sequence $(T_n) \subseteq \mathcal{T}$ such that $T_n \downarrow t$ as $n \to \infty$.

Together with the right-continuity of $\psi_t(T, u)$ in T this yields

$$\mathcal{R}(t,u) = \lim_{n\to\infty} \mathcal{R}(t, \psi_t(T_n, u)) = \lim_{n\to\infty} \widetilde{\mathcal{R}}(t, \psi_t(T_n, u)) = \widetilde{\mathcal{R}}(t,u), \tag{70}$$

for all $u \in \mathcal{E}$. Using continuity of F and R in u, Equation (70) can be extended from the dense subset $\mathcal{E}$ to all of $\mathcal{U}$. It is well-known that a function of Lévy-Khintchine-form determines its parameter triplet uniquely, cf. (**?**, Thm. 8.1). Hence, we may conclude that

$$\beta_t^i = \widetilde{\beta}_t^i, \quad \alpha_t^i = \widetilde{\alpha}_t^i, \quad \mu_t^i = \widetilde{\mu}_t^i,$$

for each $i \in \{0, \dots, d\}$ and $t \in I \cap \widetilde{I}$ with exception of the dA_t^c-nullset N. This is the desired compatibility property and shows the existence of good parameters $(A^c, \beta, \alpha, \nu)$.

We now turn to the continuous parts of the semimartingale characteristics (B, C, ν) and show (41a), (41b) and (41c). To this end, fix $(T, u) \in \mathbb{R}_{\geq 0} \times \mathcal{U}$ and let (b, c, K) be the continuous semimartingale characteristics of X, disintegrated with respect to the increasing predictable process $\widetilde{A}_t^c(\omega)$, as in (65). For each $\omega \in \Omega$, write

$$\widetilde{A}_t^c(\omega) = \int_0^t a_s(\omega) dA_t^c + S_t(\omega)$$

for the Lebesgue decomposition of $\widetilde{A}_t^c(\omega)$ with respect to A_t^c. Note that our argument does not require measurability of $\omega \mapsto a_s(\omega)$ or $\omega \mapsto S_t(\omega)$. Furthermore, define

$$\begin{aligned} g(\omega, t, T, u) := \langle \psi_t, b_t(\omega) \rangle + \frac{1}{2} \langle \psi_t, c_t(\omega) \psi_t \rangle + \\ + \int_D \left(e^{\langle \psi_t, \xi \rangle} - 1 - \langle \psi_t, h(\xi) \rangle \right) K_t(\omega, d\xi), \end{aligned} \tag{71}$$

which can be considered as the disintegrated analogue of (55). Combining (62) with the Riccati equations, we obtain that

$$\Theta_t(\omega; x) \cdot \mathcal{R}(t, \psi_t(T, u)) dA_t^c = g(\omega, t, u, T) a_t(\omega) dA_t^c + g(\omega, t, u, t) d\mathcal{S}_t(\omega) \tag{72}$$

for all $(T, u) \in \mathbb{R}_{\geq 0} \times \mathcal{U}$ and $t \in [0, T]$. By the uniqueness of the Lebesgue decomposition we conclude that

$$\begin{cases} a_t(\omega) g(\omega, t, T, u) = \Theta_t(\omega) \cdot \mathcal{R}(t, \psi_t(T, u)), & dA_t^c - a.e \\ \qquad g(\omega, t, T, u) = 0, & dS_t(\omega) - a.e. \end{cases} \tag{73}$$

As in the first part of the proof, we consider a countable dense subset $\mathcal{T} \times \mathcal{E}$ of $\mathbb{R}_{\geq 0} \times \mathcal{U}$. Taking the union over all (T, u) in $\mathcal{T} \times \mathcal{E}$ and repeating the density arguments of (70) we find an dA_t^c-nullset N_1 and a $d\mathcal{S}_t(\omega)$-nullset N_2, such that

$$\begin{cases} a_t(\omega) g(\omega, t, t, u) = \Theta_t(\omega) \cdot \mathcal{R}(t, u), & \text{for all } t \in \mathbb{R}_{\geq 0} \setminus N_1, u \in \mathcal{E} \\ \qquad g(\omega, t, t, u) = 0, & \text{for all } t \in \mathbb{R}_{\geq 0} \setminus N_2, u \in \mathcal{E}. \end{cases} \tag{74}$$

As functions of u, both sides are of Lévy-Khintchine-form. In addition, $\mathcal{E}$ is dense in $\mathcal{U}$, which allows us to conclude from the first equation that

$$\begin{aligned} a_t(\omega) b_t(\omega) &= \Theta_t(\omega) \cdot (\beta_t^0, \dots, \beta_t^d) \\ a_t(\omega) c_t(\omega) &= \Theta_t(\omega) \cdot (\alpha_t^0, \dots, \alpha_t^d) \\ a_t(\omega) K_t(\omega) &= \Theta_t(\omega) \cdot (\mu_t^{c,0}, \dots, \mu_t^{c,d}) \end{aligned}$$

for all $t \in \mathbb{R}_{\geq 0} \setminus N_1$ and from the second equation that

$$b_t(\omega) = 0, \quad c_t(\omega) = 0, \quad K_t(\omega) = 0, \qquad d\mathcal{S}_t(\omega) - a.e.$$

Integrating with respect to $\widetilde{A}^c_t(\omega)$ and adding up yields(41).

To conclude the proof, we finally turn to the discontinuous part. Note that Lemma 53 already provides us with parameters γ, a set J^ν and (not shown here) the validity of (41c) and (45). Taking the continuous increasing function A^c from the first part of the proof and inserting jumps of strictly positive hight at each time $t \in J^\nu$ we obtain an increasing function A with continuous part A^c and jump set $J^A = J^\nu$. Note that the heights of the jumps are arbitrary; for example the values of the summable series $(2^{-n})_{n\in\mathbb{N}}$ can be taken. Together, $(A, \gamma, \alpha, \beta, \mu)$ is now a good parameter set in the sense of Definition 38 and all parts of Theorem 40 have been shown. □

2.2 *Examples*

Now it is the time to study some of the important examples of affine processes - we focus on applications in climate and financial mathematics.

- Lévy processes - processes with independent and stationary increments (also called PIIS) are a classical example.
- Stationarity is of course not needed - such that processes with independent increments are affine (PII). But they are not always semimartingales.
- We already know the class of Ornstein-Uhlenbeck processes and the class of CIR-processes (Feller processes).

Example 75 (The Heston model)**.** In 1993, Heston propose a model for a stock S with stochastic volatility, extending the Black-Scholes framework. The model is given by $X = \log S$ and

$$\begin{aligned} dX_t &= (\mu + \delta Y_t)dt + \sqrt{Y_t}dW_t \\ dY_t &= (\kappa - \lambda Y_t)dt + \sigma\sqrt{Y_t}dZ_t \end{aligned}$$

with correlated Brownian motions (W, Z). This is an affine model with a lot of success, in particular it can model the *smile*.

Example 76 (Constructing stochastically discontinuous affine processes from stochastically continuous ones)**.** Consider an affine semimartingale X which is stochastically continuous. We assume that D denotes the state space of the affine semimartingale and that ϕ and ψ are the characteristics of X as in (32).

Let $\{t_1 < \cdots < t_N\} \subset \mathbb{R}_{\geq 0}$ be some time points and $a_i \in \mathbb{R}^d$, $b_i \in \mathbb{R}^{d\times d}$ such that $a_i + b_i \cdot x \in D$ for all $x \in D$, $i = 1, \ldots, N$. Then

$$\widetilde{X}_t := \sum_{i=1}^N \mathbb{1}_{\{t \geq t_i\}}(a_i + b_i \cdot X_t), \quad t \geq 0 \tag{77}$$

is an affine semimartingale in the sense of Definition 31. Note that $\widetilde{X}$ is in general not stochastically continuous, as it jumps with positive probability at the time points t_i, $i = 1, \ldots, N$.

Indeed, by the affine property of X and using iterated conditional expectations, we obtain for $t_k \le t < t_{k+1}$,

$$E\Big[e^{\langle u,\widetilde{X}_t\rangle}|\mathcal{F}_{t_k}\Big] = E\Big[\exp\Big(\langle u, \sum_{i=1}^k (a_i + b_i \cdot X_t)\rangle\Big)|\mathcal{F}_{t_k}\Big]$$
$$= e^{\sum_{i=1}^k \langle u,a_i\rangle} E\Big[\exp\Big(\langle \sum_{i=1}^k u b_i^\top, X_t\rangle\Big)|\mathcal{F}_{t_k}\Big]$$
$$= \exp\Big(\sum_{i=1}^k \langle u, a_i\rangle + \phi_{t_k}(t,u') + \langle \psi_{t_k}(t,u'), X_{t_k}\rangle\Big), \tag{78}$$

since X is affine; here we set $u' := \sum_{i=1}^k u b_i^\top$. The affine characteristics of $\widetilde{X}$ are directly obtained from Equation (78). ◇

Example 79 (Benth & Benth weather modeling)**.** The authors suggest to use the following model for the time evolution of temperatures:

$$dT_t = ds(t) + \kappa(T(t) - s(t))dt + \sigma dL_t, \tag{80}$$

with a Lévy process L. This variant of the OU - process has the following explicit solution:

$$T_t = s(t) + (T(0) - s(0))e^{\kappa t} + \int_0^t \sigma(u)\, e^{\kappa(u-t)}\, dL_u.$$

Example 81 (Affine term structure models)**.** One very prominent example of affine models arise in the interest-rate markets: the so-called affine term structure models. The bond market is a high-dimensional market, where to each bond we associated its maturity. The price of a bond with maturity T and discounting rate r is given by

$$P(t,T) = E_Q\Big[e^{-\int_t^T r_s ds}\mathscr{F}_s\Big]. \tag{82}$$

If the driving process X is affine and

$$r_t = a + b^\top X_t,$$

then $(X, \int_0^\cdot r_s ds)$ is again affine and hence

$$P(t,T) = \exp\Big(A(t,T) + B(t,T)^\top X_t\Big), \qquad 0 \le t \le T. \tag{83}$$

We will meet this markets a bit later in the lecture.

One of the main feature of affine processes is that the Fourier transform is of exponential-affine form, i.e. in a tractable form. It turns out that this allows, by variants of Plancherel's theorem, one is able to solve the pricing problem in an efficient manner, i.e. to compute

$$E_Q[F(X_T)] = \int \widehat{F}(u)\, E[\exp(iuX_t)]du, \tag{84}$$

where $\widehat{F}$ is the Fourier transform of the function F.

3 *Polynomial processes*

A further, even more flexible and surprising class of processes are *polynomial processes*. The property which characterize this class is that expectations of polynomials of the process are given as polynomials of the initial value. Let us formalize this a little bit and discuss recent results in this class.

We introduce the set $\mathscr{P}_k$ of polynomials over $\mathbb{R}^d$ with real coefficients and of most degree k,

$$\mathscr{P}_k = \Big\{ x \mapsto \sum_{|\mathbf{l}|=0}^{k} \alpha_{\mathbf{l}} x^{\mathbf{l}} \mid \alpha_{\mathbf{l}} \in \mathbb{R} \Big\}.$$

In this definition we use the so-called multi-index notation. Of course all polynomials of order smaller than k in, say dimension 2, are given by

$$\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_{12} x_1 x_2 + \dots$$

So if $\mathbf{l} = (l_1, \dots l_k) \in \mathbb{N}^k$, we denote

$$\alpha_{\mathbf{l}}\, x^{\mathbf{l}} = \alpha_{l_1,\dots,l_k} x_1^{l_1} \cdots x_k^{l_k}.$$

Definition 85. The semimartingale X is a *k-polynomial process,* if for all $l \in \{0, \dots, k\}$, $f \in \mathscr{P}_l$ we have $E[\| X_t \|_1^k] < \infty \; \forall\, t \geq 0$ and for any $0 \leq s \leq t$ there exists a $q^f_{s,t} \in \mathscr{P}_l$ such that we have

$$E[f(X_t) \mid \mathscr{F}_s] = q^f_{s,t}(X_s). \tag{86}$$

Finally, if X is has the k-polynomial property for all $k \in \mathbb{N}$, we say it has the polynomial property.

Up to now it is unclear how this definition can be applied in the semimartingale setting without further assumptions. The main problem seems to be to bring the notion to the level of characteristics. Such arguments typically can be done via a suitable generalization of the Markov property.

Under suitable assumptions one can show that a k-polynomial semimartingale satisfies the following properties: Assume there is a deterministic A which dominates the semimartingale characteristics and denote the local characteristics by (b, c, K).

(i) $b_{t,i} = b_i(t, X_{t-})$ for some $b_i \in \mathscr{P}_1$,
(ii) $c_{t,ij} = c_{ij}(t, X_{t-})$ for some measurable function c_{ij},
(iii) $K(\omega, t; d\xi) = K(t, X_{t-}, d\xi)$, where K is some transition kernel K
(iv) the functions c_{ij} and K satisfy for all $i, j \in \{1, \dots, d\}$ and $x \in \mathbb{R}^d$,

$$c_{ij}(u, x) + \int_{\mathbb{R}^d} \xi_i \xi_j K(u, x; d\xi) = a_{ij}(u, x),$$

where all $a_{ij} \in \mathscr{P}_2$,

As example we consider diffusion processes only. In the one-dimensional case, the polynomial processes extend over the affine processes in the form that the square of the volatility can be a polynomial of degree 2 !

Example 87 (Jacobi process)**.** Consider the SDE

$$dX_t = b_0 + b_1 X_t dt + \sigma \sqrt{X_t \cdot (1 - X_t)} dW_t. \tag{88}$$

This SDE has a unique solution and lives on the state space $[0, 1]$.

4 Continuous-time financial markets

For the following we consider a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, P)$ satisfying the usual conditions. As in discrete time, we consider $d+1$ traded assets with price process $S = (S^0, \dots, S^d)$. The numéraire is assumed to be strictly positive, $S^0 > 0$.

We directly introduce the discounted price process

$$X^i = \frac{S^i}{S^0} \tag{89}$$

and (only now!) assume that $X = (X^1, \dots, X^d)$ is a semimartingale.

Example 90. Construct a price process which is not a semimartingale (but X still being a semimartingale): we do not need to assume that price processes are semimartingales, since we will only work on the discounted assets.

In view of Proposition 1.1, we call a X-integrable process H a *self-financing trading strategy*. It is called admissible if there exists $\lambda \geq 0$ such that $(H \cdot X) \geq -\lambda$, i.e. when the (discounted) gains process is bounded from below.

Definition 91 (Simple trading strategies)**.** A simple trading strategy is of the form

$$H = \sum_{i=1}^{n} h_i \mathbb{1}_{]\!]\tau_{i-1}, \tau_i]\!]}$$

with stopping times $0 = \tau_0 \leq \dots \leq \tau_n < \infty$ and $\mathscr{F}_{\tau_{i-1}}$-measurable h_i, $i = 1, \dots, n$.

We call the simple strategy *admissible* if S^{τ_n} and $h_1, \dots, h_n$ are uniformly bounded.

By induction, one can show that if H is simple and an arbitrage strategy, that then there exists already a simple buy-and-hold strategy $K = h\mathbb{1}_{]\!]\sigma_1, \sigma_2]\!]}$ which is admissible and an arbitrage.

Of course, existence of an equivalent local martingale measure implies that there are no simple arbitrages. Show that absence of simple arbitrages does not imply the existence of an equivalent martingale measure by constructing an appropriate example.

4.1 No Free Lunch

The first fundamental theorem of asset pricing dates back the David Kreps and Jia-An Yan. Assume that X is locally bounded. We begin by extending Lemma 11 to the continuous-time setting.

Lemma 92. *Let $Q \ll P$. The locally bounded process X is a local Q-martingale iff*

$$E_Q[(H \cdot X)_\infty] = 0 \tag{93}$$

for each admissible simple trading strategy H.

Proof. Since X is locally bounded, there exists a sequence $(T_n)_{n \geq 1}$ of stopping times such that each X^{T_n} is bounded.

It is straightforward to show that S being a local Q-martingale implies (93) for admissible simple trading strategies.

For the converse, it is sufficient to show that each X^{T_n} is a Q-martingale, i.e. that there exist stopping times $0 \leq S_1 \leq S_2 \leq T_n$ such that

$$E_Q[X_{S_2} | \mathscr{F}_{S_1}] = X_{S_1}.$$

But this is equivalent to the requirement that for each $\mathbb{R}^d$-valued $\mathscr{F}_{S_1}$-measurable bounded function h

$$E_Q[h \cdot (X_{S_2} - X_{S_1})] = 0,$$

which holds since $h\mathbb{1}_{]\!]S_1,S_2]\!]}$ is a simple strategy. □

We consider simple trading strategies and introduce

$$K^{\text{simple}} = \{(H \cdot X)_\infty \in L^\infty(\Omega, \mathscr{F}, P) : H \text{ simple and admissible}\}$$

the vector space of claims replicable by simple strategies at zero initial cost. Moreover, we define

$$C^{\text{simple}} = K^{\text{simple}} - L^\infty_+(\Omega, \mathscr{F}, P)$$

the convex cone of contingent claims superreplicable at zero initial cost. By $\overline{C}^*$ we denote the closure of C^{simple} with respect to the weak-star topology of $L^\infty(\Omega, \mathscr{F}, P)$.

In this section we denote by $\mathscr{M}^e(X) = \mathscr{M}^e$ the set of equivalent *local* martingale measures, i.e. the set of those equivalent measures under which X is a local martingale.

Definition 94. The financial market S satisfies the condition of *no free lunch* (NFL) if

$$\overline{C}^* \cap L^\infty_+(\Omega, \mathscr{F}, P) = \{0\}. \tag{95}$$

We refer to the appendix on details on the weak*-topology.

The NFL condition is tailor-made such that the following fundamental theorem holds.

Theorem 96 (Kreps-Yan)**.** *A locally bounded process X satisfies (NFL), iff there exists an equivalent local martingale measure (ELMM):*

$$(NFL) \Longleftrightarrow \mathscr{M}^e(X) \neq \varnothing$$

Proof. We first prove that $(ELMM) \Rightarrow (NFL)$. For any $f \in C^{\text{simple}}$ and any EMM Q we have by Lemma 92 that $E_Q[f] \leq 0$. Since $f \mapsto E_Q[f]$ is weak-star continuous this also extends to $\overline{C}^*$. If (NFL) would not hold, then we would find a $Q \in \mathscr{M}^e(X)$ and $f \in \overline{C}^*$, $f \geq 0$ not vanishing almost surely, which would imply $E_Q[f] > 0$, a contradiction.

The converse $(NFL) \Rightarrow (ELMM)$ follows the lines of the proof in discrete time with some modifications. The first step is the *Hahn-Banach argument*: we claim that for $f \in L^\infty_+$, $f \neq 0$ there exists $g \in L^1_+$ which, viewed as a linear functional on L^∞ satisfies $g\overline{C}^* \leq 0$ and $(f, g) > 0$.

To see this, apply the separation theorem to the weak*-closed convex set $\overline{C}^*$ and the compact set $\{f\}$ to obtain $g \in L^1$ and $\alpha < \beta$ such that $g\overline{C}^* \leq \alpha$ and $(f, g) \ > \beta$.

Since $0 \in C$, it follows that $\alpha \geq 0$. As $\overline{C}^*$ is a cone and g is a linear functional, we have that g is zero or negative on $\overline{C}^*$ and, in particular, non-negative on L^∞_+ such that we obtain $g \in L^1_+$. Since $\beta > 0$, we have proved the above claim and step 1 is finished.

The second step is an *exhaustion argument*: denote by G the set of all $g \in L^1_+$ such that $gC \leq 0$. Since $0 \in G$, G is non-empty.

Let $\mathscr{S}$ be the family of (equivalence classes of) subsets of Ω formed by the supports $\{g > 0\}$ of the elements $g \in G$. Note that $\mathscr{S}$ is closed under countable unions, as for a sequence $(g_n) \in G$ we may find strictly positive scalars (α_n), such that $\sum_n \alpha_n g_n \in G$. Hence there is $g_0 \in G$ such that, for $\{g_0 > 0\}$ we have

$$P(g_0 > 0) = \sup\{P(g > 0) : g \in G\}.$$

We now claim that $P(g_0 > 0) = 1$. Indeed, if $P(g_0 > 0) < 1$, then we could apply step 1 to $f = \mathbb{1}_{\{g_0 = 0\}}$ to find $g_1 \in G$ with

$$E[fg_1] = \langle f, g_1 \rangle = \int_{g_0=0} g_s dP > 0.$$

Hence, $g_0 + g_1$ would be an element of G with support bigger that $P(g_0 > 0)$, a contradiction.

Finally, we normalise g_0 such that $\| g_0 \| = 1$ and let Q be the measure with Radon-Nikodym derivative $\frac{dQ}{dP} = g_0$.

By definition of g_0, we have $E_Q[(H \cdot X)_\infty] \leq 0$ for all simple and admissible strategies H. But, since $-H$ is also admissible, we obtain that $E_Q[(H \cdot X)_\infty] = 0$

We conclude from Lemma 92 that Q is a local martingale measure for X and the proof is finished. □

As a next step, we drop the assumption of locally boundedness of S, to elaborate the difficulties in this more general setup. Note that local boundedness is typically not the case: simply consider the case of discrete time - here no stopping will imply boundedness.

We will need to replace ELMM by a weaker condition: we say that X satisfies (ESM), if there exists and equivalent separating measure, i.e. there exists an equivalent measure $Q \sim P$, such that

$$E_Q[(H \cdot X)_\infty] \leq 0$$

for all H that are X-integrable and there exist some $\lambda > 0$ such that $(H \cdot X) \geq -\lambda$.

We then obtain the following theorem.

Theorem 97. *Fix $p \in [1, \infty)$ and q such that $p^{-1} + q^{-1} = 1$. Suppose $C \subset L^p$ is a convex cone with $C \supseteq -L^p_+$ and*

$$C \cap L^p_+ = \{0\}. \tag{98}$$

If C is closed in $\sigma(L^p, L^q)$, then there exists $Q \sim P$ with $\frac{dQ}{dP} \in L^q(P)$ and $E_Q[Y] \leq 0$ for all $Y \in C$.

Proof. As for the Kreps-Yan theorem we may utilize the Hahn-Banach argument: chose $f \in L^p_+ \backslash \{0\}$, which is disjoint from C by (98). We obtain some $g \in L^q$ separating $\{f\}$ from C, i.e. $E[gY] \leq 0$ for all $Y \in C$. Since $C \supset -L^p_+$, we obtain $g \geq 0$. Moreover, since we have strict separation $g \not\geq 0$, such that we can normalize to $E[g] = 1$.

The exhaustion argument works identically, such that we find a $Z > 0$ almost surely, $Z \in L^q$, $E[ZY] \leq 0$ for all $Y \in C$. Through normalization we obtain $E[Z] = 1$, such that $dQ := ZdP$ does the job. □

4.2 *The general case*

There are a number of no-arbitrage principles around and we beging by recalling or introducing them. We introduce the trading strategies with bounded risk as

$$K^1 = \{(H \cdot X)_\infty : H \text{ 1-admissible}\},$$

the gains from trade by admissible trading

$$K = \{(H \cdot X)_\infty : H \text{ admissible}\},$$

and set

$$C = (K - L^0_+) \cap L^\infty.$$

Then we are able to introduce the following definitions

(i) The process X satisfies no arbitrage (NA), if

$$(K - L^0_+(\Omega, \mathscr{F}, P)) \cap L^0_+ = \{0\}.$$

This is equivalent to

$$((K - L^0_+) \cap L^\infty) \cap L^\infty_+ = C \cap L^\infty_+ = \{0\}.$$

(ii) The process X satisfies no free lunch with vanishing risk (NFLVR), if

$$\overline{C} \cap L^\infty_+ = \{0\},$$

where $\bar{C}$ denotes the norm closure of C in L^∞.

(NFL) The process X satisfies NFL, if

$$\overline{C}^* \cap L^\infty_+ = \{0\}.$$

(iii) The process X satisfies no unbounded profit with bounded risk (NUPBR), if K^1 is a bounded subset of L^0.

Some remarks are due. First, coming from (NFL), we replaced the technical weak-* closure by the norm closure[1] in L^∞ for (NFLVR). It is very intuitive, to understand what a free lunch with fanishing risk is. By our definition, $f \in L^\infty_+ \backslash \{0\}$ is a FLVR, if there exists a sequence $(f_n) = (H_n \cdot X)_\infty$ (lying in K) and a sequence of $g_n \leq f_n$, such that

$$\lim \| f - g_n \|_\infty = 0.$$

[1] Compare Equation ? $\| f \|_\infty := \text{esssup}_{\omega \in \Omega}$ | variable in L^∞ is a ran which is bounded (al

In particular the negative parts (f_n^-) and (g_n^-) have to converge to zero, which explains the notion *vanishing risk*.

Second, an unbounded profit with bounded risk is a sequence $(f_n) = (H_n \cdot X)_\infty$ (lying in K^1) if (f_n) are *unbounded in probability*, i.e.

$$\lim_{m\to\infty} \sup_n P(f_n > m) > 0.$$

Of course we have

$$\text{(NFL)} \Rightarrow \text{(NFLVR)} \Rightarrow \text{(NA)}.$$

It can additionally be shown that

$$\text{NFLVR} \Leftrightarrow \text{(NA)} + \text{(NUPBR)}.$$

and it is a deep insight of Delbaen and Schachermayer that under (NFLVR), $C = \overline{C}^*$, i.e. the cone C is already weak-* closed (and hence, NFL holds) and we obtain the fundamental theorem of asset pricing.

Theorem 99. *Under (NFLVR), the cone C is weak-* closed, hence*

$$\textit{(NFLVR)} \Leftrightarrow \textit{(ESM)}.$$

For a detailed proof we refer to Delbaen & Schachermayer (2006) and Cuchiero & Teichmann (2015).

5 *Markets with transaction costs*

In this section we extend the approach to incorporate a central feature of financial markets: *transaction costs* and *bid-ask-spreads*. If we want to buy a stock it turns out that the price (the *ask*-price) is actually higher than the price for which I can sell the stock (the *bid*-price). Also transaction costs could be incorporated in such a way.

Hence in our market we face to price processes: the bid price, S and the ask price $\bar{S}$ where we assume

$$S^i_t \leq \bar{S}^i_t$$

for all $t \geq 0$ and $i = 0, \dots, d$. As usual, $S^0 > 0$ and $S^i \geq 0$ is assumed.

As previously, trading strategies are given by predictable processes, but now we assume – for simplicity – that they are of finite variation. We decompose the trading strategy $H^i = H^{i,\uparrow} - H^{i,\downarrow}$ into an increasing process $H^{i,\uparrow}$ and a decreasing process $-H^{i,\downarrow}$. For intuition, we consider the discrete-time setting.

Example 100 (Discrete-time market)**.** Self-financing is ensured when gains and expenses from rebalancing at any time sum up to 0, i.e.,

$$\sum_{i=0}^{d} \left(\bar{S}^i_{t-1} \Delta H^{i,\uparrow}_t - S^i_{t-1} \Delta H^{i,\downarrow}_t \right) = 0$$

for all $t = 1, \dots, T$. If we sum these terms up we obtain of course stochastic integrals, such that this is equivalent to

$$\bar{S}_- \cdot H^{\uparrow} - S_- \cdot H^{\downarrow} = 0. \tag{9.19}$$

However, the continuous-time counterpart raises difficulties since the stochastic integral is only defined for càdlàg processes. But note that

$$\begin{aligned}
(\bar{S}^i_- \cdot H^{i,\uparrow})_t &= \sum_{s=1}^{t} \bar{S}^i_{s-1} \left(H^{i,\uparrow}_s - H^{i,\uparrow}_{s-1} \right) \\
&= \sum_{s=1}^{t} \left(H^{i,\uparrow}_s \bar{S}^i_s - H^{i,\uparrow}_{s-1} \bar{S}^i_{s-1} - H^{i,\uparrow}_s (\bar{S}^i_s - \bar{S}^i_{s-1}) \right) \\
&= H^{i,\uparrow}_t \bar{S}^i_t - H^{i,\uparrow}_0 \bar{S}^i_0 - H^{\uparrow} \cdot \bar{S}.
\end{aligned}$$

Definition 101. A predictable process H of finite variation is a *self-financing trading strategy*, if

$$\bar{S}_- \cdot H^{\uparrow} - S_- \cdot H^{\downarrow} = 0, \tag{9.20}$$

where the integrals are defined as

$$\begin{aligned}
(\bar{S}_- \cdot H^{\uparrow})_t &:= \langle H^{\uparrow}_t, \bar{S}_t \rangle - \langle H^{\uparrow}_0, \bar{S}_0 \rangle - (H^{\uparrow} \cdot \bar{S})_t, \\
(S_- \cdot H^{\downarrow})_t &:= \langle H^{\downarrow}_t, S_t \rangle - \langle H^{\downarrow}_0, S_0 \rangle - (H^{\downarrow} \cdot S)_t.
\end{aligned}$$

Observe that this equation reduces to the usual self-financing condition if $S = \bar{S}$, i.e., if bid and ask prices coincide. Moreover, note that we no longer have a direct value of a portfolio and we consider the *liquidation value* instead. It is obtained by selling the positive part of the portfolio at the bid price and buying the negative part at the ask price, i.e.

$$V_t = V^H_t := H^+_t S_t - H^-_t \bar{S}_t = \sum_{i=0}^{d} \left(H^{i,+}_t S^i - H^{i,-}_t \bar{S}^i_t \right). \tag{102}$$

Definition 103. An *arbitrage* is a self-financing trading strategy starting with initial value $V_0 = 0$ with liquidation value $V_T \geq 0$ and $P(V_T > 0) > 0$.

The idea of the FTAP is to construct a fictitious process X which is directly between S and $\bar{S}$. In this regard we call C a *consistent price process* if C is adapted and

$$S^i \leq C^i \leq \bar{S}^i \qquad i = 0, 1 \dots, d.$$

An investment in C is at least as attractive as investing in the true market with transaction costs since the investor buys at lower prices and sells at higher prices.

Theorem 104 (FTAP umder bid-/ask-spreads)**.** *The following statements are equivalent:*

(i) There are no arbitrage opportunities.

(ii) There exists a consistent price system such that there are no arbitrage opportunities for this price process.

Proof. We roughly sketch the idea of the proof. Start with a market with no arbitrage opportunities and define

$$K := \{H_T - \xi \colon H \text{ self-financing with } V_0 = 0 \text{ and } \xi \in L^0_+(\mathbb{R}^d)\}. \tag{105}$$

K is the cone of attainable terminal portfolio holdings with zero initial investments when we may remove assets at T. We define the *solvency cone* as those final positions which can be liquidated at less then zero cost, i.e.

$$M := \Big\{H \in L^0(\mathbb{R}^d) : v^H := H_T^+ S_T - H^- \bar{S}_T \geq 0 \text{ and } E_P[v^H] = 1\Big\}.$$

Absence of arbitrage now implies that

$$K \cap M = \varnothing.$$

Let us assume that we are able to apply the Hahn-Banach separation theorem, yielding a random vector g such that $E[g\xi] \leq 0$ for any $\xi \in K$ and $E[g\xi] > 0$ for any $\xi \in M$.

Now consider the unit vector $e_0 = (1, 0, \dots, 0) \in \mathbb{R}^{1+d}$. Note that

$$\frac{e_0 \mathbb{1}_F}{E[S_T^0 \mathbb{1}_C]} \in M$$

for any $F \in \mathscr{F}$, since a positive position always liquidates at a positive value. Hence, $E[g^0 \mathbb{1}_F] > 0$ for F with $P(F) > 0$, i.e. $g^0 > 0$ almost surely. Normalizing, we may even assume $E[g^0] = 1$. Then we can define the equivalent probability measure $Q \sim P$ by $dQ = g^0 dP$.

The density processes $Z_t^i = E[g^i | \mathscr{F}_t]$, $i = 0, \dots, d$ allow us to define the candidate for the consistent price system C by

$$C_t^i := \frac{S_t^0 Z_t^i}{Z_t^0}$$

for $0 \leq t \leq T$ and $i = 0, \dots, d$. The corresponding discounted price process is denoted by $X = C(S^0)^{-1}$.

Observe that $Z^0 X^i = Z^i$ is by definition a martingale, such that X is a Q-martingale, i.e. the frictionless market with price process X satisfies (NA).

Finally, we need to verify that C is a consistent price system. To this end, fix $i \in \{1, \dots, d\}$, $t \in [0, T]$ and $F \in \mathscr{F}_t$. We consider the self-financing strategy H starting at 0 which buys S^i at t on F. This gives

$$H_T = \left(-\mathbb{1}_C \frac{\bar{S}_t^i}{S_t^0}, 0, \dots, 0, \mathbb{1}_F, 0, \dots, 0 \right).$$

By definition $H_T \in K$, and hence

$$E\Big[\mathbb{1}_F \Big(Z_T^i - Z_T^0 \frac{\bar{S}_t^i}{S_t^0} \Big)\Big] \leq 0.$$

Since Z is a martingale,

$$E\Big[\mathbb{1}_F \Big(Z_t^i - Z_t^0 \frac{\bar{S}_t^i}{S_t^0} \Big)\Big] \leq 0.$$

such that $Z_t^i - Z_t^0 \frac{\bar{S}_t^i}{S_t^0} \leq 0$ since F was arbitrary. Since Z^0 and S^0 are non-negative, this is equivalent to

$$C_t^i = \frac{S_t^0 Z_t^i}{Z_t^0} \leq \bar{S}_t^i.$$

Considering a selling strategy instead shows $S_t^i \leq C_t^i$ and the first part of the proof is finished.

For the reverse part, consider a self-financing strategy H for the market with bid ask-spreads. We can construct an associated self-financing strategy G on the corresponding price system C by $G^i = H^i$ for $i = 1, \ldots, d$ and adjusting G^0 such that G is self-financing. The difference of $G^0 - H^0$ is of course increasing, such that

$$V_T^G \geq V_T^H.$$

Now suppose $V_T^H \geq 0$. Then we also have $V_T^G \geq 0$ and by absence of arbitrage for the consistent price system $V^G = V^H = 0$. □

III
Term-structure modelling

In this chapter we study interest rate modelling and credit risk. References in this field are Filipović (2009) and McNeil et al. (2015) and towards semimartingale modelling Gehmlich & Schmidt (2018), Fontana & Schmidt (2018), Fontana et al. (2020).

1 Interest rate markets - the classical theory

In this section we meet different type of interest rates and the associated market conventions. The basic observation in interest rate theory is that 1 EUR today has a different value than 1 EUR at a future timepoints, say in 1 year. To make this tradeable we consider the following simple product.

Definition 1. A *zero-coupon bond* with nominal N and maturity T promises the owner the payment of N units of currency at time T.

For simplicity we consider always zero-coupon bonds with nominal $N = 1$. The price of the zero-coupon bond at time $t \leq T$ is denoted by $P(t, T)$. The bond with maturity T will also be called T-bond. We throughout assume that $P(t, T) > 0$ (which means no credit risk) for all $t \leq T$ and that $T \mapsto P(t, T)$ is differentiable for all $t \geq 0$.

Example 2. Assume a bank offers you for your investment of 1 today the payment of 1.04 in 2 years. Then the associated yield (annualized) equals $(1.04 - 1)/2$. This is a so-called simple yield. Can we fix the yield today for a future period, say from S to T? The answer is yes: we mimic the following payment scheme:

- Invest 1 EUR at S
- Get x EUR at $T > S$

by the following trading strategy:

- Sell 1 S-bond at t.
- Buy $P(t, S)/P(t, T)$ T-bonds at t.

This trading strategy has zero cost at t, the cash-flow of -1 at S and $P(t, S)/P(t, T)$ at T.

The associated simple rates to the above example lead to the following definition.

Definition 3. The *simple forward rate* for the time interval $[S, T]$ at time $t \leq S$ is given by

$$F(t, S, T) := \frac{1}{T - S}\left(\frac{P(t, S)}{P(t, T)} - 1\right).$$

The *spot forward rate* is

$$F(t, T) := F(t, t, T).$$

Example 4. In comparison to yearly compounding one could also consider monthly, daily or even finer compounding. In principle we observe that getting finer and finer leads to

$$\Big(1+\frac{R}{n}\Big)^{nT} \to e^{RT}$$

which is called *continuous compounding*.

Definition 5. The *continuously compounded forward rate* for the time interval $[S,T]$ at time $t \le S$ is given by

$$R(t,S,T) := \frac{1}{T-S}\log\frac{P(t,S)}{P(t,T)}$$

and the *continuously compounded spot rate* is

$$R(t,T) := R(t,t,T).$$

Finally, we define instantaneous rates which make the time variation of the rates best visible by letting $T \downarrow S$. We compute

$$\begin{aligned} f(t,T) &:= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\Big(\log P(t,T) - \log P(t,T+\varepsilon)\Big) \\ &\to -\partial_T \log P(t,T). \end{aligned}$$

By integration we obtain that

$$\int_t^T f(t,s)ds = -\big(\log P(t,T) - \log P(t,t)\big) = -\log P(t,T)$$

such that all bond-prices can be represented by forward rates in the following way

$$P(t,T) = \exp\Big(-\int_t^T f(t,s)ds\Big). \tag{6}$$

This leads us to the following definition.

Definition 7. Assume that

$$P(t,T) = e^{-\int_t^T f(t,u)du}, \qquad 0 \le t \le T,$$

then $f(t,T)$ is called the (instantaneous) *forward rate* for maturity T at time $t \le T$ and the (instantaneous) *spot rate* is given by

$$r_t := f(t,t).$$

Exercise 8. Compute $F(t,S,T)$ and $R(t,S,T)$ from the forward rates.

1.1 The money market account

Seen in an idealized way, the money market account B is obtained by continuously investing in the short rate,

$$dB_t = B_t r_t dt \qquad B(0) = 1.$$

The solution is of course $B_t = \exp\int_0^t r(s)ds$. This may be proxied by a roll-over strategy,

$$\begin{aligned} B_t^n &= \prod_{i=1}^n \frac{1}{P(t_{i-1}^n, t_i^n)} = \prod_{i=1}^n \exp\Big(\int_{t_{i-1}^n}^{t_i^n} f(t_{i-1}^n, s)ds\Big) \\ &\to \exp\Big(\int_0^t r_s ds\Big). \end{aligned}$$

1.2 Interest rate products

On the market we find fixed coupon bonds, floating range notes and swaps which we will discuss now.

Definition 9. A *fixed coupon bond* with tenor structure $T_1 < \cdots < T_n = T$, coupon sizes $c_1, \ldots, c_n$ and nominal N guarantees the payment of c_i at T_i and N at T.

For any payment stream A given by $A_t = \sum_{i=1}^n c_i \mathbb{1}_{\{t \geq T_i\}}$ we obtain the pricing formula

$$\pi_t^A = \int_t^{T_n} P(t,s) dA(s) = \sum_{i=1}^n \mathbb{1}_{\{T_i \leq t\}} P(t, T_i) c_i.$$

Consequently the price of the fixed coupon bond is

$$P_t^c = \sum_{i=1}^n c_i \mathbb{1}_{\{t \leq T_i\}} P(t, T_i) + \mathbb{1}_{\{t \leq T\}} P(t, T) N.$$

Example 10. Can we also price random payments? Consider the investment of 1 in a T_1-bond, yielding $\xi_1 := P(0, T_1)^{-1}$ at 1. Pay out $\xi_1 - 1$ at T_1 and invest 1 in a T_2-bond and so on. This gives the payment of $\mathscr{F}_{T_{i-1}}$-measurable but random payments at T_i.

Definition 11. A *floating range note* (FRN) with tenor structure $T_0 < T_1 < \cdots < T_n = T$ and notional N offers the payment of $N(T_i - T_{i-1}) F(T_{i-1}, T_i)$ at $T_i, i = 1, \ldots, n$ and the payment of N at T.

T_0 is called reset date. As in the above example we obtain that the price of the floating range note equals

$$p_t^{FRN} = P(t, T_0) N.$$

Exercise 12. Consider a deterministic term structure, i.e. deterministic forward rates $f(t, T)$ and prove that the value of the payment stream given by

$$A_t = \sum_{i=1}^n (T_i - T_{i-1}) F(T_{i-1}, T_i) \mathbb{1}_{\{t \geq T_i\}} + \mathbb{1}_{\{t \geq T\}}$$

coincides with p_t^{FRN}.

1.3 Interest rate swaps

The holder of a payer swap pays fixed rates in exchange for floating rates, and vice versa for a receiver swap.

Definition 13. The holder of a payer *interest rate swap* with tenor structure $T_0 < T_1 < \cdots < T_n = T$ and notional N

- pays $N(T_i - T_{i-1})\kappa$
- receives $N(T_i - T_{i-1}) F(T_i, T_{i-1})$

at times T_i, $i = 1, \ldots, n$.

Note that no notional is exchanged in a swap agreement. We compute the price of the payer swap

$$\Pi_t^p = N\Big[\sum_{i=1}^n P(t,T_i)(T_i - T_{i-1})\Big(F(T_i,T_{i-1}) - \sum_{i=1}^n \kappa\Big)\Big]$$
$$= N\Big[P(t,T_0) - P(t,T) - \kappa\sum_{i=1}^n P(t,T_i)(T_i - T_{i-1})\Big].$$

The value of the receiver swap equals

$$\Pi_t^r = -\Pi_t^p.$$

Definition 14. The *swap rate* for the tenor structure $T_0 < T_1 < \cdots < T_n = T$ is the rate κ leading to zero value of the associated swap.

We compute the swap rate to

$$R_t^{swap} = \frac{P(t,T_0) - P(t,T_n)}{\sum_{i=1}^n (T_i - T_{i-1})P(t,T_i)}, \qquad t \le T_0.$$

Exercise 15. The swap rate can be seen as a weighted sum of the forward rates. Compute the weights.

1.4 *Duration and convexity*

Duration and convexity measure the first and second order sensitivity of bonds with respect to parallel shifts of the yield curve. In this regard define the yield $y_i = R(t,T_i)$ such that the value of a fixed coupon bond equals

$$P_t^c = \sum_{i=1}^n \mathbb{1}_{\{T_i \ge t\}} c_i e^{-y_i(T_i - t)}.$$

Define the associated duration by

$$D_t := \frac{1}{P_t^c}\sum_{i=1}^n \mathbb{1}_{\{T_i \ge t\}} c_i T_i e^{-y_i(T_i - t)}.$$

Then for $t \le T_1$

$$\partial_s \sum_{i=1}^n c_i e^{-(y_i+s)(T_i-t)}\Big|_{s=0} = -DP_t^c.$$

The second derivative is called convexity

$$C_t := \frac{1}{P_t^c}\sum_{i=1}^n \mathbb{1}_{\{T_i \ge t\}} c_i (T_i - t)^2 e^{-y_i(T_i - t)}.$$

And we obtain the important second-order expansion by the Taylor formula:

$$\Delta P_t^c \approx -D_t P_t^c \Delta y + \frac{1}{2} C_t P_t^c (\Delta y)^2.$$

2 *What we need ...*

We shortly reframe our knowledge on no-arbitrage theory in the setting considered here.

2.1 *From no-arbitrage theory*

As is customary in financial literature we will mainly use the claim that the existence of an equivalent martingale measure (EMM) is sufficient for absence of arbitrage. Note that we consider typically an infinite-dimensional market and the right concept is no asymptotic free lunch with vanishing risk (NAFLVR), see for example Klein et al. (2016).

To obtain arbitrage-free prices of contingent claim we proceed as follows: specify an EMM Q and price a T-claim X (a $\mathscr{F}_T$-measurable random variable such that the following expectation exist) by

$$\pi_t^X := S_t^0 E_Q\Big[\frac{X}{S_T^0} \mid \mathscr{F}_t\Big].$$

Then the market enlarged by π^X, $(S_0, \dots, S_d, \pi^X)$ is free of arbitrage. Recall Bayes' theorem:

Theorem 16. *Consider a measure $\widetilde{P}$ such that $P \ll \widetilde{P}$. Let $\Lambda := dP/d\widetilde{P}$. For any random variable X with $E|X| < \infty$ and any σ-field $\mathscr{G} \subset \mathscr{F}$*

$$E[X|\mathscr{G}] = \frac{\widetilde{E}[X\Lambda|\mathscr{G}]}{\widetilde{E}[\Lambda|\mathscr{G}]} \qquad P\text{-a.s.} \tag{17}$$

As a next step, define

$$\pi_t := \frac{1}{S_t^0} E_Q\Big[\frac{dQ}{dP}|\mathscr{F}_t\Big].$$

Then Bayes' rule implies that

$$\begin{aligned}\pi_t^X = S_t^0 E_Q\Big[\frac{X}{S_T^0}|\mathscr{F}_t\Big] &= S_t^0 \frac{E\Big[\frac{X}{S_T^0}\frac{dQ}{dP}|\mathscr{F}_t\Big]}{E\Big[\frac{dQ}{dP}|\mathscr{F}_t\Big]} \\ &= \frac{E_Q\Big[X\pi_T|\mathscr{F}_t\Big]}{\pi_t}.\end{aligned}$$

We also find that π is a martingale, i.e.

$$\pi_t := E_Q\Big[\frac{1}{S_T^0}\frac{dQ}{dP}|\mathscr{F}_t\Big].$$

It plays a role of the Girsanov density in our financial context and is often called *state-price density process*.

Example 18. We determine the T-bond price by the state price density. Note

$$P(t,T) = E\Big[\frac{1}{}\frac{\pi_T}{\pi_t}|\mathscr{F}_t\Big] = E_Q\Big[\frac{S_t^0}{S_T^0}|\mathscr{F}_t\Big].$$

If the numèraire S^0 is the bank account $S_t^0 = \exp\Big(\int_0^t r_s ds\Big)$ we obtain

$$P(t,T) = E^Q\Big[\exp\Big(-\int_t^T r_s ds\Big)|\mathscr{F}_t\Big]. \tag{19}$$

2.2 *The stochastic Fubini Theorem*

It will be a central tool in the following Heath-Jarrow-Morton approach to use

$$\int_t^T \int_0^t b(s,u) dW_s \, du = \int_0^t \int_t^T b(s,u) du \, dW_s,$$

a problem which requires a stochastic version of the Fubini theorem. It is quite surprising that very strong assumptions are needed for the stochastic Fubini theorem to hold (in contrast to the classical Fubini theorem).

Let us first recall the monotone class theorem from the lecture on stochastic processes (Theorem I.3.2 there)

Theorem 20 (Monotone class theorem). *Let $\mathscr{H}$ be a set of bounded functions $S \to \mathbb{R}$, s.t.*

(i) $\mathscr{H}$ is a $\mathbb{R}$-vector space,

(ii) $1 \in \mathscr{H}$,

(iii) for a sequence $(f_n) \geq 0$ in $\mathscr{H}$, with $f_n \uparrow f$ where f is bounded it holds that $f \in \mathscr{H}$.

If a set $\mathscr{C} \subseteq \mathscr{H}$ of bounded functions $S \to \mathbb{R}$, is closed under multiplication, then $\mathscr{H}$ contains all bounded $\sigma(\mathscr{C})$-measurable functions.

Recall that the monotone class theorem is an important tool when we want to deal with classes of stochastic processes, for example predictable ones, adapted ones, etc.

The integrand (b,u) in the above equation is allowed to depend on an additional parameter. We consider a general measurable space $A, \mathscr{A}$ where this parameter can lie in. Recall that we denoted by $\mathscr{P}$ the σ-algebra of predictable processes, i.e. the σ-algebra generated by all adapted processes which are càg (left continuous).

We further introduce a convergence which is slightly weaker than convergence in the Emery topology.

Definition 21. A sequence of processes X^n and a process X for which $\sup_{[0,t]} |X^n - X|$ is measurable for all n is said to *converge uniformly on compacts in probability* (ucp), if

$$P\Big(\sup_{[0,t]} |X^n - X| > K\Big) \to 0$$

for all $K, t > 0$.

Note that for càdlàg processes the supremum is measurable, but in general it might not be. This convergence is metrizable and the space of càdlàg (continuous) processes is complete under this topology. As an exercise, give an example of processes which converge in u.c.p. but not in the semimartingale (Emergy) topology. A further exercise would be to deduce ucp convergence from semimartingale convergence.

We have the following theorem on dominated convergence of stochastic integrals (see Theorem 2.32 in **?**).

Theorem 22. *Let X be a semimartingale and (H^n) a sequence of predictable processes converging a.s. to H. If there exists $G \in L(X)$, such that $|H^n| \leq G$ for all n, then $H^n, H \in L(X)$ and*

$$H^n \cdot X \xrightarrow{ucp} X.$$

We are now able to prove our first step towards the Fubini theorem.

Proposition 23. *Let X be a semimartingale with $X_0 = 0$ and let $H(a,t,\omega) = H_t^a(\omega)$ be $\mathscr{A} \otimes \mathscr{P}$-measurable and bounded. Then there is a function $Z \in \mathscr{A} \otimes \mathscr{B}(\mathbb{R}_+) \otimes \mathscr{F}$ such that for each $a \in A$, $Z^a = Z(a,.,.)$ is a càdlàg, adapted version of $H^a \cdot X$.*

Proof. We use the monotone class theorem. In this regard, let $\mathscr{H}$ be the class of bounded $\mathscr{A} \otimes \mathscr{P}$-measurable functions such that the claim holds.

Consider K bounded predictable, f bounded, $\mathscr{A}$-measurable and $H = Kf$, then

$$\int_0^t H(a,s,\cdot)dX_s = \int_0^t f(a)\,K(s,\cdot)dX_s = f(a)\int_0^t K(s,\cdot)dX_s, \tag{24}$$

and $H \in \mathscr{H}$. Moreover, $\mathscr{H}$ is a vector space and $1 \in \mathscr{H}$. Theorem 22 also shows that it is a monotone class. Hence, the monotone class theorem yields the result. □

Theorem 25. *Let X be a semimartingale, $H(a,t,\omega) = H_t^a(\omega)$ be $\mathscr{A} \otimes \mathscr{P}$-measurable and bounded and let μ be a finite measure on $(A, \mathscr{A})$. Let Z^a be the càdlàg version of $H^a \cdot X$ from Proposition 23.*

Then,

$$Y = \int_A Z^a \mu(da)$$

is a càdlàg version of $H \cdot H$ where $H_t = \int_A H_t^a \mu(da)$.

Proof. First we stop and therefore may assume that $X \in \mathscr{H}^2$. The finite variation part of the semimartingale X satisfies the claim by the classical Fubini theorem, such that we may assume X is a martingale with $E[X]_\infty < \infty$.

Consider $H_t^a(\omega) = f(a)K(t,\omega)$, where K is bounded and predictable, and f is bounded and $\mathscr{A}$-measurable. Then $K \in L(X)$ and $\int |f| d\mu < \infty$. Hence $Z^a = fK \cdot X$ and

$$\begin{aligned}\int_A Z_t^a \mu(da) &= \int_A f(a)\,K \cdot X \mu(da)\\ &= \Big(\int_A f(a)\mu(da)K\Big)\cdot X = H \cdot X.\end{aligned}$$

By linearity this holds for the vector space generated by all processes of the form Kf.

For the monotone class theorem, consider $H_n \to H$ and let $Z_{n,t}^a = H_n^a \cdot X$ be the càdlàg version according to Proposition 23. Then, by Jensens' inequality applied to the probability measure $\mu(A)^{-1}\mu$,

$$\begin{aligned}\frac{1}{\mu(A)}\Big(E\Big[\int_A \sup_{\mathbb{R}} |Z_{n,t}^a - Z_t^a|\;\mu(da)\Big]\Big)^2 &\le E\Big[\int_A \sup_{\mathbb{R}} (Z_{n,t}^a - Z_t^a)^2\;\mu(da)\Big]\\ &= \int_A E\Big[\sup_{\mathbb{R}}(Z_{n,t}^a - Z_t^a)^2\Big]\;\mu(da)\\ &\le 4\int_A E\Big[(Z_{n,\infty}^a - Z_\infty^a)^2\Big]\\ &= 4\int_A E\Big[[Z_n^a - Z^a]_\infty\Big]\;\mu(da),\end{aligned}$$

using Doobs maximal inequality and that for a martinale M, $M^2 - [M]$ is again a martingale.

Since $Z^a = H^a \cdot X$ and $Z_n^a = H_n^a \cdot X$,

$$= 4\int_A E\Big[\int_0^\infty (H_{n,s}^a - H_s^a)^2\,d[X]_s\Big]\mu(da) \to 0$$

by appropriate and repeated use of dominated convergence. □

3 *The Heath-Jarrow-Morton approach*

In the late 1980's Heath et al. (1992) started looking at dynamic models for the whole term structure. The starting point are the forward rates. We assume that they are Itô-processes of the form

$$f(t,T) = f(0,T) + \int_0^t a(s,T)ds + \int_0^t b(s,T)^\top dW_s.$$

We remark that W is a d-dimensional Brownian motion, such that if we integrate with a d-dimensional process b, we need to write $b^\top dW$ to obtain a one-dimensional integral. We will continue this kind of notation in this section.

We will need the following assumptions

$$f(0,\cdot) \text{ is } \mathscr{B}\text{-measurable}, \tag{26}$$

$$a, b \text{ are } \mathrm{Prog} \otimes \mathscr{B}\text{-measurable}, \tag{27}$$

$$\int_0^T \int_0^T |a(s,t)| ds\, dt < \infty \quad \text{for all } T \geq 0, \tag{28}$$

$$\sup_{s,t \leq T} \| b(s,t) \| < \infty \quad \text{for all } T \geq 0. \tag{29}$$

To study absence of arbitrage[1] we study equivalent measures of the form

$$dQ = \mathcal{E}_\infty(\gamma^\top \cdot W)\, dP$$

for some $\gamma \in L(W)$. Then $W^* := W - \int_0^\cdot \gamma_s ds$ is a Q-Brownian motion. We obtain the following, sufficient condition for absence of arbitrage:

$$(B_t^{-1} P(t,T))_{0 \leq t \leq T} \quad \text{is } Q\text{-local martingale for all } T > 0. \tag{30}$$

We are interested in the Q-dynamics of $f(t,T)$ where we use the representation

$$f(t,T) = f(0,T) + \int_0^t a^*(s,T) ds + \int_0^t b(s,T)^\top dW^*(s).$$

Theorem 31. *Assume that* (26)–(29) *hold. Then, the no-arbitrage condition* (30) *holds if and only if*

$$-\int_t^T a(t,u) du + \frac{1}{2} \| B(t,T) \|^2 = -B(t,T)^\top \gamma(t) \tag{32}$$

for all T, $dP \otimes dt$-a.s. with

$$B(t,T) = -\int_t^T b(t,u) du.$$

In this case

$$a^*(t,T) = b(s,T) \int_s^T b(s,u)^\top du. \tag{33}$$

The equation (33) is called HJM-drift condition. We will also show that discounted bond-prices are stochastic exponentials, i.e.

$$B_t^{-1} P(t,T) = P(0,T) \mathcal{E}_t\big(B(\cdot,T)^\top \cdot W^*\big).$$

The proof will be in several steps.

Lemma 34. *The zero coupon bond is an Itô-process satisfying*

$$\frac{dP(t,T)}{P(t,T)} = \big(r_t + \alpha(t,T)\big) dt + B(t,T) dW_t$$

where

$$\alpha(t,T) := -\int_t^T a(t,u) du + \frac{1}{2} \| B(t,T) \|^2 .$$

[1] The arbitrage-theory in general is more sub example Cuchiero et a Klein et al. (2016).

Proof. We first consider

$$\begin{aligned}
\log P(t,T) &= -\int_t^T f(t,u)du \\
&= -\int_t^T \left[f(0,u) + \int_0^t a(s,u)ds + \int_0^t b(s,u)^\top dW_s\right] du \\
&= -\int_t^T f(0,u)du - \int_0^t \int_t^T a(s,u)du\,ds - \int_0^t \int_t^T b(s,u)^\top du\,dW_s \\
&= -\int_0^T f(0,u)du - \int_0^t \int_s^T a(s,u)du\,ds - \int_0^t \int_s^T b(s,u)^\top du\,dW_s \\
&+ \int_0^t f(0,u)du + \int_0^t \int_s^t a(s,u)du\,ds + \int_0^t \int_s^t b(s,u)^\top du\,dW_s
\end{aligned}$$

where we apply the stochastic Fubini theorem, Theorem 25. Moreover note that by the classical Fubini theorem,

$$\int_0^t \int_s^t a(s,u)du\,ds = \int_0^t \int_0^u a(s,u)ds\,du$$

and similarly for the other integrals in the last line. Summarizing, the last line equals $\int_0^t r(u)du$. Hence,

$$\log P(t,T) = -\int_0^T f(0,u)du - \int_0^t \int_s^T a(s,u)du\,ds + \int_0^t v(s,T)dW_s + \int_0^t r(u)du.$$

Applying the Itô-formula yields that

$$dP(t,T) = P(t,T)\Big[\Big(r(t) - \int_t^T a(t,u)du + \frac{1}{2} \parallel v(t,T) \parallel^2 \Big)dt + v(t,T)dW_t\Big]$$

and we conclude. □

We are ready to prove the main result.

Proof. We derive the dynamics of the discounted bond prices by the Itô-formula and Lemma 34,

$$d\big(B_t^{-1}P(t,T)\big) = P(t,T)\Big(\alpha(t,T)dt + B(t,T)^\top dW_t\Big).$$

Girsanov's theorem yields that $W^* = W - \int_0^t \gamma^\top(s)ds$ is a Q-Brownian motion. Hence,

$$d\big(B_t^{-1}P(t,T)\big) = P(t,T)\Big((\alpha(t,T) + B(t,T)^\top\gamma(t))dt + B(t,T)^\top dW_t^*\Big).$$

This process is a local martingale if and only $\alpha(t,T) + B(t,T)^\top\gamma(t))$ vanishes (being equivalent to (32)) and $B(\cdot,T) \in L(W)$ (which is implied by (29)).

To obtain the HJM-drift condition we differentiate once more: note that $\alpha(t,T) + B(t,T)^\top\gamma(t))$ is differential (w.r.t. T) such that we obtain by differentiation

$$\begin{aligned}
0 &= \partial_T\alpha(t,T) + \partial_T B(t,T)\gamma^\top(t) \\
&= -a(t,T) + b(t,T)^\top \int_t^T b(t,u)du - b^\top(t,T)\gamma(t) \\
&= -a(t,T) + b(t,T)^\top \int_t^T b(t,u)du,
\end{aligned}$$

and we obtain the drift condition. On the converse it is straightforward to check that the drift condition implies absence of arbitrage. □

4 *Credit risk modelling*

In this section we start with static modelling of credit risk with a fixed time horizon. If there is frequently updated information on the underlying credits, an intensity based approach as studied in the following chapter seems more appropriate. This is the case in credit derivatives pricing and hedging. On the other side, if one studies a large portfolio of creditors with infrequent information updates, the static viewpoint can be useful as it is the case in typical credit risk management situations. It also turns out to be much simpler.

4.1 *Intoduction to credit risk*

In his landmark paper Merton (1974) applied the framework of Black & Scholes (1973) to the pricing of a corporate bond. A corporate bond promises the repayment F at maturity T. Since the issuing company might not be able to pay the full amount of money back, the payoff is subject to *default risk*.

Let V_t denote the firrm's value at time t. If, at time T, the firm's value V_T is below F, the company is not able to make the promised repayment so that a *default event* occurs. In Merton's model it is assumed that there are no bankruptcy costs and that the bond holder receives the remaining V_T, thus facing a financial loss. If we consider the payoff of the corporate bond in this model, we see that it is equal to F in the case of no default ($V_T \geq F$) and V_T otherwise, i.e.,

$$\mathbb{1}_{\{V_T \geq F\}} F + V_T \mathbb{1}_{\{V_T < F\}} = F - (F - V_T)^+.$$

If we split the liabilities into smaller bonds with face value 1, then we can replicate the payoff of this bond by a portfolio of a riskless bond $p^0(t,T)$ with face value 1 (long) and $1/F$ puts with strike F (short).

Let $(\Omega, \mathscr{F}, \mathbb{F}, P)$ be a filtered probability space where $\mathbb{F}$ satisfies the usual conditions. We consider a setup similar to the Black-Scholes model: the short rate r is constant, such that the bank account equals $B(t) = e^{rt}$ and risk-neutral bond prices are given by

$$P(t,T) = e^{-r(T-t)}.$$

Moreover, we assume that V is the unique strong solution of the stochastic differential euqation

$$dV_t = V_t\big(\mu dt + \sigma dW_t\big), \quad V(0) = v_0$$

with $\sigma, v_0 > 0$. Note that V is *not* a traded asset. There exist many equivalent martingale measures (in fact every equivalent measure is also a martingale measure) and we describe them by the means of the Girsanov theorem. Fix a finite time horizon T. Every equivalent measure Q admits a density of the form

$$Z_T = \mathcal{E}_T(a \cdot W)$$

with some progressively measurable process a such that $\mathbb{E}[Z_T] = 1$. A sufficient criterion for this is Novikov's criterion. Under the measure Q

$$W_t^* := W_t - \int_0^t a_s ds, \quad 0 \leq t \leq T$$

is a Brownian motion. We obtain that V satisfies

$$dV_t = V_t\big((\mu + \sigma a_t)dt + \sigma dW_t^*\big), \quad V_0 = v_0$$

which is notably more complicated than the dynamics under P. One key insight of arbitrage-free pricing is that, changing to an equivalent martingale measure implies vanishing drift. Unfortunately only for all traded and discounted assets, which is not the case for V. We therefore make the following assumption:

(A1) Under Q, the drift of V under Q equals a constant

$$\mu^* := \mu + \sigma a_t, \quad 0 \le t \le T$$

P-a.s.

The Girsanov density is then given by $a_t \equiv (\mu^* - \mu)/\sigma$ which turns out to be a constant as well. The additional yield over the risk-free rate is an important quantity, which is called credit spread.

Definition 35. For a defaultable T-bond with price process $p(\cdot, T)$ the process $s = (s(t,T))_{0 \le t \le T}$ given by

$$s(t,T) = -\frac{1}{T-t} \log \frac{p(t,T)}{P(t,T)}$$

is called *credit spread*.

Recall that $P(t,T)$ was the price of a risk-free bond with maturity T.

Proposition 36. *Under (A1), the price of the defaultable bond at time $t \le T$ equals*

$$p(t,T) = e^{-r(T-t)}\Phi(d_2) + \frac{V_t}{F} e^{(\mu^*-r)(T-t)}\Phi(-d_1)$$

with

$$d_2 = \frac{(\mu^* - \sigma^2/2)(T-t) + \ln(V_t/F)}{\sigma\sqrt{(T-t)}},$$

and

$$d_1 = \frac{(\mu^* + \sigma^2/2)(T-t) + \ln(V_t/F)}{\sigma\sqrt{(T-t)}}.$$

Moreover,

$$s(t,T) = -\frac{1}{T-t} \log\Big(\Phi(d_2) + \frac{V_t}{Fe^{-\mu^*(T-t)}}\Phi(-d_1)\Big).$$

Here Φ denotes the cumulative distribution function of the standard normal distribution. The defaultable bond is therefore cheaper than the default-free bond because it carries additional risk.

A main criticism of the Merton model is that $s(t,T) \to 0$ as $t \to T$ which is easy to see. However, its practical implementation in the KMV approach became famous.

A further problem is that defaults can only happen at T. This was improved by so-called first-passage time approaches, as pioneered in Black & Cox (1976). Including jumps also allows for a default coming unforeseen and we discuss this approach as proposed in Zhou (2001).

Because change of measure for processes with jumps is not part of this course, we assume that we directly work under the risk-neutral measure Q.

Consider i.i.d. $U_1, U_2, \dots$ and a Poisson process N with intensity λ and a Brownian motion W^*, all being independent. The firm value is the unique strong solution of

$$dV_t = V_{t-}\big(\mu^* dt + \sigma dW_t^* + dJ_t\big), \quad t \ge 0 \tag{37}$$

where $J(t) = \sum_{i=1}^{N_t} U_i$ is a compound Poisson process. First, observe that

$$\Delta V_t = V_{t-}\Delta J_t.$$

We denote the jumping times of N by $s_1, s_2, \dots$ and obtain that

$$V_{s_i} = V_{s_i-}(U_i - 1).$$

Consequently, V has multiplicative jumps, and $(U-1)$ denote the size of the jump factor and the SDE (37) has the solution

$$V(t) = V(0)\prod_{i=1}^{N_t}(U_i - 1)\ \exp\Big((\mu^* - \frac{\sigma^2}{2})t + \sigma W_t^*\Big), \quad t \geq 0.$$

As in the Merton model we consider default at T when $V_T < K$. The bond prices are computed in the following proposition.

Proposition 38. *Assume $U_i = 1 + e^{\eta_i}$ with i.i.d. and normally distributed $\eta_1, \eta_2, \ldots$. Assume the bond-holder receives zero at default. Then*

$$p(t,T) = e^{-(r+\lambda)(T-t)}\sum_{k=0}^{\infty} F(k, T-t)$$

where for $k = 0, 1, 2, \ldots$ the function F is given by

$$F(k,t) = \frac{(\lambda t)^k}{k!}\Phi\Big(\frac{(\mu^* - \sigma^2/2)t + k\log(E[U_1]-1) - k\sigma_\eta^2/2 - \log(F/V_t)}{\sqrt{\sigma^2 t + k\sigma_\eta^2}}\Big).$$

Proof. We compute the bond-price for $t = 0$, the general expression following from the independent and stationary increments of W and J. As $N(T)$ is Poisson distributed with parameter λT, we have that

$$Q(N_T = k) = e^{-\lambda T}\frac{(\lambda T)^k}{k!}.$$

Consequently, by the risk-neutral pricing rule

$$p(0,T) = E_Q\Big[e^{-rT}\mathbb{1}_{\{V_T \geq F\}}\Big] = e^{-rT}\sum_{k\geq 0} e^{-\lambda T}\frac{(\lambda T)^k}{k!}\cdot D(k)$$

with

$$\begin{aligned} D(k) &= Q\Big(V_0\prod_{i=1}^{k}(U_i - 1)e^{(\mu^* - \sigma^2/2)T + \sigma W_T} \geq F\Big) \\ &= Q\Big(\sum_{i=1}^{k}\eta_i + (\mu^* - \sigma^2/2)T + \sigma W_T \geq \log(F/V_0)\Big). \end{aligned}$$

The term on the l.h.s. is normally distributed with variance $k\sigma_\eta^2 + \sigma^2 T$ and mean $kE[\eta_1] + (\mu^* - \sigma^2/2)T$. As $U_1 = 1 + \exp(\eta_1)$ we obtain, using the Laplace transform of the normal distribution that

$$E[U_1] = 1 - E[e^{\eta_1}] = 1 + e^{E[\eta_1] + 1/2\sigma_\eta^2},$$

such that

$$E[\eta_1] = \log(E[U_1] - 1) - \frac{\sigma_\eta^2}{2}.$$

Note that $Q(a + b\xi \geq x) = \Phi((a-x)/b)$ for $\xi \sim \mathcal{N}(0,1)$ and, summarizing

$$D(k) = \Phi\Big(\frac{(\mu^* - \sigma^2/2)T - \log(F/V_0) + k\log(\mathbb{E}U_1 - 1) - k\sigma_\eta^2/2}{\sqrt{k\sigma_\eta^2 + \sigma^2 T}}\Big).$$

□

A first-passage time model is of the form

$$\tau = \inf\{t \geq 0 : \Lambda(t) \geq \Theta\}.$$

Exercise 39. Find the corresponding expression for the Merton model.

4.2 *Reduced-form credit modelling*

An interesting alternative model approach to credit is the so-called reduced-form modelling of credit. In contrast to the Merton model, where an explicit mechanism leading to default is specified, here a probabilistic model is used which is then calibrated to data.

A *doubly-stochastic* model for credit risk is of the form

$$\tau = \inf\{t \geq 0 : \Lambda_t \geq \Theta\},$$

here Λ and Θ are independent (therefore doubly stochastic), and Λ is increasing.

Example 40 (Intensity-based approach)**.** This approach leads to the simplest setting. Given a non-negative, progressively measurable process λ, we set

$$\Lambda_t := \int_0^t \lambda_s ds,$$

and assume that Θ is independent of Λ and standard exponential.

Then,

$$\begin{aligned} P(\tau > T | \tau > t, \mathscr{F}_T^\Lambda) &= \mathbb{1}_{\{\tau > t\}} P(\tau > T | \tau > t, \mathscr{F}_T^\Lambda) \\ &= \mathbb{1}_{\{\tau > t\}} P(\Theta > \Lambda_T | \, \Theta > \Lambda_t, \mathscr{F}_T^\Lambda) \\ &= \mathbb{1}_{\{\tau > t\}} P(\Theta' > \Lambda_T - \Lambda_t | \, \mathscr{F}_T^\Lambda) \\ &= \mathbb{1}_{\{\tau > t\}} e^{\Lambda_T - \Lambda_t}. \end{aligned}$$

Now we continue with a more general setting. Introduce the default indicator process

$$H_t = \mathbb{1}_{\{t \leq \tau\}}$$

since this is an increasing process, hence a submartingale, its Doob-Meyer decomposition exists and we denote it by

$$H_t = M_t - H_t^p$$

with a martingale M and an increasing, predictable process H^p.

The goal is now to provide a sufficiently large class of predictable processes for H^p which captures stylized facts. In particular we would be interested in H^p having predictable jumps.

Definition 41. We call a random time U *announced* if there exists an $\mathbb{F}$-stopping time S with $S < U$ almost surely and U is $\mathscr{F}_S$-measurable.

Any announced time is predictable. Announced times are a highly tractable class of predictable times.

We say that H is in the *intensity-based setting*, if

$$H_t^p = \int_0^{t \wedge \tau} h_s ds.$$

Example 42 (Beyond stochastic continuity)**.** The intensity-based setting is a case very often considered. It neglects however, that many jumps occur at predictable times. This restriction was relaxed in Gehmlich & Schmidt (2018), where it was assumed that

$$H_t^p = \int_0^{t\wedge\tau} h_s ds + \sum_{U_i \le t\wedge\tau} \Gamma_i,$$

where U_i is announced at S_i and Γ_i is $\mathscr{F}_{S_i}$-measurable.

We extend the HJM-setting to the defaultable setting under *zero recovery* by starting from the following assumption:

$$P(t,T) = \mathbb{1}_{\{\tau>t\}} \exp\Big(- \int_t^T f(t,u)du \Big), \quad 0 \le t \le T \le T^*. \tag{43}$$

As previously, we assume that the processes f are Itô processes of the form

$$f(t,T) = f(0,T) + \int_0^t a(s,T)ds + \int_0^t b(s,T)\cdot dW_s, \tag{44}$$

We will need the following technical assumptions. Let us assume we work on a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, Q^*)$, where Q^* is a candidate for a risk-neutral measure, i.e. equivalent to the objective measure.

(B1) The process h is non-negative, predictable and integrable on $[0,T^*]$:

$$\int_0^{T^*} |h_s| ds < \infty, \quad Q^*\text{-a.s.},$$

(B2) the initial forward curves $f(\omega,0,t)$ are $\mathscr{F}_0 \otimes \mathscr{B}$-measurable, and integrable on $[0,T^*]$:

$$\int_0^{T^*} |f(0,u)| du < \infty, \qquad Q^*\text{-a.s.},$$

(B3) the *drift parameters* $a(\omega,s,t)$ are $\mathbb{R}$-valued, and $\mathscr{O}\otimes\mathscr{B}$-measurable. The parameter a is integrable on $[0,T^*]$:

$$\int_0^{T^*}\int_0^{T^*} |a(s,t)| ds\, dt < \infty, \quad Q^*\text{-a.s.}.$$

(B4) the *volatility parameter* $b(\omega,s,t)$ is $\mathbb{R}^n$-valued, $\mathscr{O}\otimes\mathscr{B}$-measurable, and bounded on $[0,T^*]$:

$$\sup_{s,t\le T^*} \| b(s,t) \| < \infty, \quad Q^*\text{-a.s.}$$

The following result gives the desired drift condition rendering the measure Q^* an equivalent local martingale measure. Set

$$\bar a(t,T) = \int_t^T a(t,u)du, \qquad \bar b(t,T) = \int_t^T b(t,u)du.$$

Theorem 45. *Assume that (B1) - (B4) hold. Then Q^* is an ELMM if and only if the following two conditions hold:*

$$f = r + h, \tag{46}$$

$$\bar a(t,T) = \frac{1}{2} \| \bar b(t,T) \|^2, \tag{47}$$

$0 \le t \le T \le T^*$, *$dQ^* \otimes dt$-almost surely on* $\{t < \tau\}$.

In principle this means that the HJM condition (compare (30)) needs to hold and the risky short rate $f(s,s)$ is the sum of risk-free short-rate r and default intensity h.

The proof bases on the Ito-formula and the Fubini theorem, as previously. The main novelty is the following step:

Proof of Theorem 45. Set $F(t,T) := \exp\Big(-\int_t^T f(t,u)du \Big)$ and $E(t) := \mathbb{1}_{\{\tau>t\}}$ such that

$$P(t,T) = E_t \, F(t,T).$$

Then, by integration by parts,

$$\begin{aligned} dP(t,T) &= F(t,T)\, dE_t + E_{t-} dF(t,T) + d[E, F(.,T)]_t \\ &=: (1'') + (2'') + (3'') \end{aligned} \tag{48}$$

and we compute the terms next. Regarding $(1'')$, we obtain

$$E_t + \int_0^{t\wedge\tau} h_s ds =: M_t^1 \tag{49}$$

is a martingale. Regarding $(2'')$, we have that

$$dF(t,T) = F(t,T)\Big(f(t,t) - \bar a(t,T) + \frac{1}{2} \parallel \bar b(t,T) \parallel^2 \Big) dt - F(t,T) \bar b(t,T) dW_t \tag{50}$$

by Theorem 31.

Inserting (49) and (50) into (48), we obtain that on $\{t < \tau\}$,

$$\frac{dP(t,T)}{P(t-,T)} = \Big(-h(t) + f(t,t) + \frac{1}{2} \parallel \bar b(t,T) \parallel^2 - \bar a(t,T) \Big) dt + dM_t^2$$

with a local martingale M^2. The discounted price process $(X_t^{-1} P(t,T))_{0\le t\le T}$ is a local martingale if and only if the predictable part in the semimartingale decomposition vanishes. Letting $t = T$ one recovers

$$0 = \int_0^t (f(s,s) - h(s) - r_s) ds$$

for $0 \le t \le T^*$, on $\{t < \tau\}$, which is equivalent to $f(s,s) = h(s) + r_s$, such that first (46) and then (47) follow. The converse is easy to see. □

4.3 *An extension of the HJM-approach*

Now we are in the position to extend the HJM approach in an appropriate way to obtain arbitrage-free defaultbale term structure models under weak assumptions. Consider a measure $Q^* \sim P$. Our intention is to find conditions which render Q^* an equivalent local martingale measure. From now on, only occasionally the measure P will be used, such that all appearing terms (like martingales, almost sure properties, etc.) are to be considered with respect to Q^* if not stated otherwise.

We return to the general setting of Example 42: recall that $H_t = \mathbb{1}_{\{\tau>t\}}$ was the default indicator process and H^p its compensator. To keep the arising technical difficulties at a minimum, we assume that H^p can be decomposed in an absolutely continuous and a (predictable) pure-jump part, such that

$$H_t^p = \int_0^{t\wedge\tau} h_s ds + \int_0^{t\wedge\tau} \int_{\mathbb{R}} x \Gamma(ds,dx), \quad t \ge 0, \tag{51}$$

with a non-negative, predictable process h and with a predictable integer-valued random measure.

We recall definition 41 of announced times. The intuition behind this definition is as follows: at the announcement time S the market receives new information about a future date U (i.e. $S < U$) at which default may happen with positive probability. For example, at time S the market realises that a country has difficulties to pay some of its obligations which are due at the coupon payment date U. Note that any deterministic, positive time is announced and that an announced time is always predictable.

To ensure that the subsequent analysis is meaningful, we make the following technical assumptions. We assume that (B1) holds and introduce

(C1) the random measure $\Gamma(ds, dx)$ is given by

$$\Gamma([0,t], dx) = \sum_{i=1}^{N} \mathbb{1}_{\{U_i \le t\}} \delta_{\Gamma_i}(dx),$$

where each risky time U_i is announced, say by S_i, and $\Gamma_i : \Omega \to (0,1)$ is $\mathcal{F}_{S_i}$-measurable, $1 \le i \le N$.

Assumption (C1) implies that the set $\mathcal{U}$ of (default) risky times is finite. This is a reasonable assumption while working on a finite time interval. If $\Gamma_i = 1$, default happens with probability one at time U_i, a case which we exclude for simplicity of exposition.

Regarding defaultable bond prices we will start from a forward-rate framework and allow for discontinuities in the term structure at risky times. Consider current time $t \in [0, T^*)$ and a bond with maturity $T \in (t, T^*]$. If the risky time U_i was announced before time t, investors will obtain an additional premium for the event $\{\tau = U_i\}$ only when $T \ge U_i$. For $T < U_i$ the investors are not exposed to this risk and hence will not receive an additional premium. This naturally leads to a discontinuity in the term structure $T \mapsto P(., T)$ at U_i. Motivated by this, we consider a family of random measures $(\mu_t)_{t \ge 0}$, defined by

$$\mu_t(du) := \sum_{S_i \le t} \delta_{U_i}(du)$$

and assume that defaultable bond prices are given by

$$P(t,T) = \mathbb{1}_{\{\tau > t\}} \exp\Big(- \int_t^T f(t,u)du - \int_t^T g(t,u)\mu_t(du)\Big), \quad 0 \le t \le T \le T^*. \tag{52}$$

The processes f and g are assumed to be Itô processes of the form

$$f(t,T) = f(0,T) + \int_0^t a(s,T)ds + \int_0^t b(s,T) \cdot dW_s, \tag{53}$$

$$g(t,T) = g(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \beta(s,T) \cdot dW_s, \tag{54}$$

with an n-dimensional Q^*-Brownian motion W. By $\mathscr{B}$ we denote the Borel σ-algebra generated by the open sets in $\mathbb{R}_{\ge 0}$ and by $\mathscr{O}$ we denote the optional σ-algebra generated by all $\mathbb{F}$-adapted càdlàg processes. We will need the following technical assumptions.

(C2) the initial forward curves $f(\omega, 0, t)$ and $g(\omega, 0, t)$ are $\mathscr{F}_0 \otimes \mathscr{B}$-measurable, and integrable on $[0, T^*]$:

$$\int_0^{T^*} |f(0,u)| + |g(0,u)| du < \infty, \qquad Q^*\text{-a.s.},$$

(C3) the *drift parameters* $a(\omega, s, t)$ and $\alpha(\omega, s, t)$ are $\mathbb{R}$-valued, and $\mathscr{O} \otimes \mathscr{B}$-measurable. The parameter a is integrable on $[0, T^*]$:

$$\int_0^{T^*} \int_0^{T^*} |a(s,t)| ds\, dt < \infty, \quad Q^*\text{-a.s.},$$

while α is bounded on $[0, T^*]$:

$$\sup_{s,t \le T^*} |\alpha(s,t)| < \infty, \quad Q^*\text{-a.s.},$$

(C4) the *volatility parameter* $b(\omega, s, t)$ is $\mathbb{R}^n$-valued, $\mathscr{O} \otimes \mathscr{B}$-measurable, and bounded on $[0, T^*]$:

$$\sup_{s,t \le T^*} \| b(s,t) \| < \infty, \quad Q^*\text{-a.s.},$$

while $\beta(\omega, s, t)$ is $\mathbb{R}^n$-valued, $\mathscr{O} \otimes \mathscr{B}$-measurable, and square integrable on $[0, T^*]$:

$$\int_0^{T^*} \int_0^{T^*} \| \beta(s,t) \|^2 \, ds \, dt < \infty, \quad Q^*\text{-a.s.},$$

(C5) we assume that the dual predictable projection ν of the integer-valued random measure $\mu(dt, du) = \sum_{i=1}^n \delta_{(S_i, U_i)}(dt, du)$ satisfies $\nu(dt, du) = \nu(t, du)dt$ with a kernel $\nu(\omega, t, du)$, and

$$\int_0^{T^*} \int_0^{T^*} |e^{-g(t,u)} - 1| \, \nu(t, du) dt < \infty, \quad Q^*\text{-a.s.}$$

Moreover $Q^*(\tau = S_i) = 0$ for all $i \ge 1$.

Remark 55. Assumption (C5) requires that announcing times are totally inaccessible, i.e. come as a surprise. Moreover, there is no default by news, i.e. τ does not coincide with an announcing time. Both assumptions have been made to simplify the exposition but could be relaxed without big difficulties at the cost of lengthier formulas.

The following result gives the desired drift condition rendering the considered measure Q^* an equivalent local martingale measure. Set

$$\begin{aligned}
\bar{a}(t,T) &= \int_t^T a(t,u)du, \\
\bar{b}(t,T) &= \int_t^T b(t,u)du, \\
\bar{\alpha}(t,T) &= \int_t^T \alpha(t,u)\mu_t(du), \\
\bar{\beta}(t,T) &= \int_t^T \beta(t,u)\mu_t(du).
\end{aligned}$$

Theorem 56. *Assume that* (B1) *and* (C1)*-*(C5) *hold. Then Q^* is an ELMM if and only if the following two conditions hold:*

$$\int_0^t f(s,s)ds + \sum_{U_i \le t} g(U_i, U_i) = \int_0^t (r_s + h_s)ds - \sum_{U_i \le t} \log(1 - \Gamma_i), \tag{57}$$

$$\bar{a}(t,T) + \bar{\alpha}(t,T) = \frac{1}{2} \| \bar{b}(t,T) + \bar{\beta}(t,T) \|^2 + \int_t^T \left(e^{-g(t,u)} - 1 \right) \nu(t, du), \tag{58}$$

$0 \le t \le T \le T^*$, *$dQ^* \otimes dt$-almost surely on $\{t < \tau\}$.*

In comparison to the classical HJM drift condition in the default-risk free case, $\bar{a}(t,T) = \frac{1}{2} \| \bar{b}(t,T) \|^2$, a number of additional terms appear here. First, Equation (57) under $g(.,.) = 0$ and $\mathcal{U} = \varnothing$ is the condition in intensity-based dynamic term structure models, compare Equation (46). The additional terms incorporate additional returns due to the extra default risk at risky times. It turns out, that if $\Delta H^p \neq 0$, then a classical HJM-approach with $g(.,.) = 0$ allows for arbitrage profits.

The additional term in (58), $\int_t^T \left(e^{-g(t,u)} - 1\right) \nu(t, du)$, appears as compensation for jumps in the term structure at news news arrival times $S_1, S_2, \dots$ and can be linked to similar expressions in classical HJM-Models with jumps.

The following simple example illustrates the extension of our approach over intensity-based models and builds up intuition on condition (57).

Example 59. Consider a non-negative integrable and progressive process λ, constants $0 < u_1 < \dots < u_N$, positive random variables $\lambda'_1, \dots, \lambda'_N$, with λ'_i being $\mathcal{F}_{u_i}$-measurable, and set

$$\Lambda_t = \int_0^t \lambda(s)ds + \sum_{u_i \le t} \lambda'_i.$$

Let ζ be a standard exponential random variable, independent from Λ, and set

$$\tau = \inf\{t \ge 0 : \Lambda_t \ge \zeta\}.$$

Here, we have $\Delta H^p_{u_i} > 0$ because u_i is a risky time: by the memoryless-property of exponential random variables,

$$Q^*(\tau = u_i | \tau \ge u_i) = Q^*(\lambda'_i \ge \zeta) = E^*[1 - \exp(-\lambda'_i)]. \tag{60}$$

If Λ is deterministic and the short-rate vanishes, we obtain the following term-structure

$$P(t,T) = \mathbb{1}_{\{\tau > t\}} Q^*(\tau > T | \ \tau > t) = \mathbb{1}_{\{\tau > t\}} \exp\Big(- \int_t^T \lambda(s)ds - \sum_{u_i \in (t,T]} \lambda'_i \Big),$$

which clearly falls into the class of models considered here. A simple computation yields

$$H^p_t = \int_0^{t\wedge\tau} \lambda(s)ds + \sum_{i: u_i \le (t \wedge \tau)} (1 - e^{-\lambda'_i}) \tag{61}$$

and it is easily checked that the drift conditions (57)-(58) hold. ◇

The proof of Theorem 56 will make use of the following lemma in which we derive the canonical decomposition of the second integral in (52), denoted by

$$I(t,T) := \int_t^T g(t,u)\mu_t(du), \qquad 0 \le t \le T. \tag{62}$$

Lemma 63. *Assume that (A1), (A2), and (B1), (B2) hold. Then, for each $T \in [0, T^*]$ the process $(I(t,T))_{0 \le t \le T}$ is a special semimartingale and*

$$I(t,T) = \int_0^t \bar{\alpha}(s,T)ds + \int_0^t \bar{\beta}(s,T) \cdot dW_s + \int_0^t \int_0^T g(s,u) \mathbb{1}_{\{s<u\}} \mu(ds,du) - \int_0^t g(s,s)\mu^U(ds)$$

with $\mu^U(ds) = \sum_{i=1}^n \delta_{U_i}(ds)$.

Proof. We start with the observation that, by the definition of μ_t,

$$\begin{aligned} I(t,T) &= \int_0^t \int_t^T g(t,u)\mu(ds,du) \\ &= \int_0^t \int_0^T \mathbb{1}_{\{u>t\}} g(t,u)\mu(ds,du) \\ &= \int_0^t \int_0^T \mathbb{1}_{[0,u)}(t) g(t,u)\mu(ds,du). \end{aligned} \tag{64}$$

The semimartingales $(\mathbb{1}_{[0,u)}(t)g(t,u))$ have the following canonical decompositions,

$$\begin{aligned}\mathbb{1}_{[0,u)}(t)g(t,u) &= g(0,u) + \int_0^t \mathbb{1}_{[0,u]}(v)\, dg(v,u) + \int_0^t g(v,u)d(\mathbb{1}_{[0,u)}(v)) \\ &= g(0,u) + \int_0^t \mathbb{1}_{[0,u]}(v)\alpha(v,u)dv + \int_0^t \mathbb{1}_{[0,u]}(v)\beta(v,u)\cdot dW_v - g(u,u)\mathbb{1}_{\{u\le t\}} \quad (65)\end{aligned}$$

and we obtain that

$$\begin{aligned}(64) &= \int_0^t \int_0^T g(0,u)\mu(ds,du) + \int_0^t \int_0^T \int_0^t \mathbb{1}_{[0,u]}(v)\alpha(v,u)dv\, \mu(ds,du) \\ &\quad + \int_0^t \int_0^T \int_0^t \mathbb{1}_{[0,u]}(v)\beta(v,u)\cdot dW_v\, \mu(ds,du) - \int_0^t \int_0^T g(u,u)\mathbb{1}_{\{u\le t\}}\mu(ds,du) \\ &=: (1') + (2') + (3') + (4').\end{aligned}$$

With (C3) it is possible to interchange the appearing integrals as the integral with respect to μ is a finite sum. Hence,

$$\begin{aligned}(2') &= \int_0^t \int_0^t \int_0^T \mathbb{1}_{[0,u]}(v)\alpha(v,u)\mu(ds,du)dv \\ &= \int_0^t \int_0^v \int_0^T \mathbb{1}_{[0,u]}(v)\alpha(v,u)\mu(ds,du)dv + \int_0^t \int_v^t \int_0^T \mathbb{1}_{[0,u]}(v)\alpha(v,u)\mu(ds,du)dv \\ &= \int_0^t \int_0^v \int_0^T \mathbb{1}_{[0,u]}(v)\alpha(v,u)\mu(ds,du)dv + \int_0^t \int_0^T \int_0^s \mathbb{1}_{[0,u]}(v)\alpha(v,u)dv\, \mu(ds,du)\end{aligned}$$

with an analogous expression for $(3')$. Note that the first term in the last line equals $\int_0^t \bar{\alpha}(v,T)dv$. By (65),

$$\int_0^s \mathbb{1}_{[0,u]}(v)\alpha(v,u)dv + \int_0^s \mathbb{1}_{[0,u]}(v)\beta(v,u)\cdot dW_v = \mathbb{1}_{[0,u)}(s)g(s,u) - g(0,u) + g(u,u)\mathbb{1}_{\{u\le s\}}$$

such that (64) is equal to

$$\int_0^t \bar{\alpha}(v,T)dv + \int_0^t \bar{\beta}(v,T)\cdot dW_v + \int_0^t \int_0^T \mathbb{1}_{[0,u)}(s)g(s,u)\mu(ds,du) - \int_0^t \int_0^T \mathbb{1}_{[s,t]}(u)g(u,u)\mu(ds,du).$$

By Assumption (C1),

$$\int_0^t \int_0^T \mathbb{1}_{[s,t]}(u)g(u,u)\mu(ds,du) = \sum_{U_i\le t} g(U_i,U_i) = \int_0^t g(s,s)\mu^U(ds)$$

which is a special semimartingale and we conclude. □

The previous lemma allows us to obtain the semimartingale representation of

$$G(t,T) := \exp(-I(t,T)), \qquad 0\le t\le T.$$

Proposition 66. *Assume that the above assumptions hold. Then,*

$$\begin{aligned}\frac{dG(t,T)}{G(t-,T)} &= \Big(-\bar{\alpha}(t,T) + \frac{1}{2} \parallel \bar{\beta}(t,T) \parallel^2 + \int_t^T \left(e^{-g(t,u)} - 1\right)\nu(t,du)\Big)dt \\ &\quad - \bar{\beta}(t,T)\cdot dW_t + \left(e^{g(t,t)} - 1\right)\mu^U(dt) + dM_t^1,\end{aligned}$$

with a local martingale M^1.

Proof. The Itô-formula together the representation of G given in Lemma 63 yields that

$$\begin{aligned}G(t,T) &= G(0,T) + \int_0^t G(s-,T)\Big(-\bar{\alpha}(s,T) + \frac{1}{2} \parallel \bar{\beta}(s,t) \parallel^2 \Big)ds - \int_0^t G(s-,T)\bar{\beta}(s,T)\cdot dW_s \\ &\quad + \int_0^t \int_0^T G(s-,T)\left(e^{-g(s,u)\mathbb{1}_{\{u>s\}}} - 1\right)\mu(ds,du) + \int_0^t G(s-,T)\left(e^{g(s,s)} - 1\right)\mu^U(ds). \quad (67)\end{aligned}$$

Using Assumption (C5), we compensate $\mu(ds,du)$ by $\nu(s,du)ds$ and obtain the result. □

Proof of Theorem 56. Set $F(t,T) := \exp\Big(-\int_t^T f(t,u)du\Big)$, $E(t) := \mathbb{1}_{\{\tau>t\}}$ such that

$$P(t,T) = E(t)F(t,T)G(t,T).$$

Then, by integration by parts,

$$\begin{aligned} dP(t,T) &= F(t,T)G(t-,T)dE(t) + E(t-)d(F(t,T)G(t,T)) + d[E, F(.,T)G(.,T)]_t \\ &=: (1'') + (2'') + (3'') \end{aligned} \tag{68}$$

and we compute the according terms in the following. Regarding $(1'')$, we obtain from (51), that

$$E(t) + \int_0^{t\wedge\tau} h_s ds + \int_0^{t\wedge\tau}\int_{\mathbb{R}} x\,\Gamma(ds,dx) =: M_t^2 \tag{69}$$

is a martingale. Regarding $(2'')$, we have that

$$d(F(t,T)G(t,T)) = G(t-,T)dF(t,T) + F(t,T)dG(t,T) + d\,\langle G^c(.,T), F^c(.,T)\rangle_t,$$

where $F^c(.,T)$ and $G^c(.,T)$ are the continuous local martingale parts of $F(.,T)$ and $G(.,T)$, respectively. Computing the dynamics of $F(t,T)$ gives

$$dF(t,T) = F(t,T)\Big(f(t,t) - \bar{a}(t,T) + \frac{1}{2} \parallel \bar{b}(t,T) \parallel^2\Big)dt - F(t,T)\bar{b}(t,T)dW_t. \tag{70}$$

Together with Proposition 66 this leads to

$$\begin{aligned} \frac{d(F(t,T)G(t,T))}{F(t,T)G(t-,T)} &= M_t^3 + \Big(e^{g(t,t)} - 1\Big)\mu^U(dt) \\ &+ \Big(f(t,t) - \bar{a}(t,T) + \frac{1}{2} \parallel \bar{b}(t,T) + \bar{\beta}(t,T) \parallel^2 - \bar{\alpha}(t,T)\Big)dt \\ &- (\bar{b}(t,T) + \bar{\beta}(t,T))\cdot dW_t + \int_t^T \Big(e^{-g(t,u)} - 1\Big)\nu(t,du)dt, \end{aligned} \tag{71}$$

where we used that $\parallel \bar{b}(t,T) \parallel^2 + \parallel \bar{\beta}(t,T) \parallel^2 + 2\bar{b}(t,T)\cdot\bar{\beta}(t,T)^\top = \parallel \bar{b}(t,T) + \bar{\beta}(t,T) \parallel^2$ and a local martingale M^3. In view of $(3'')$, we obtain from (67) that

$$\frac{\Delta G(t,T)}{G(t-,T)} = \int_t^T (e^{-g(t,u)} - 1)\mu(\{t\},du) + (e^{g(t,t)} - 1)\mu^U(\{t\}).$$

By Assumption (B4), $\Delta E(t)\mu(\{t\},\mathbb{R}) = -\sum_{i\geq 1}\mathbb{1}_{\{\tau=t\}}\mathbb{1}_{\{S_i=t\}} = 0$. Hence, using (69),

$$\begin{aligned} \sum_{0<s\leq t}\Delta E(s)\Delta G(s,T) &= \int_0^t G(s-,T)(e^{g(s,s)} - 1)\mu^U(\{s\})dE(s) \\ &= \int_0^t G(s-,T)(e^{g(s,s)} - 1)\mu^U(\{s\})dM_s^2 \\ &- \int_0^{t\wedge\tau}\int_{\mathbb{R}} G(s-,T)(e^{g(s,s)} - 1)\mu^U(\{s\})x\,\Gamma(ds,dx); \end{aligned} \tag{72}$$

where we used that for an integrable function $f:\mathbb{R}\to\mathbb{R}$, $\int f(s)\mu_T(\{s\})ds = 0$ as μ is concentrated on a finite set. Note that $(e^{g(t,t)}\mu^U(\{t\}))_{t\geq 0}$ is predictable due to Assumption (A2) and $\mu^U(\{s\})\Gamma(ds,dx) = \Gamma(ds,dx)$.

Inserting (69), (71) and (72) into (68), we arrive that on $\{t < \tau\}$,

$$\begin{aligned}\frac{dP(t,T)}{P(t-,T)} &= -h(t)dt - \int_{\mathbb{R}} x\,\Gamma(dt,dx) \\ &+ \Big(f(t,t) + \frac{1}{2} \parallel \bar{b}(t,T) + \bar{\beta}(t,T) \parallel^2 - \bar{a}(t,T) - \bar{\alpha}(t,T)\Big)dt \\ &+ \int_{\mathbb{R}} (e^{g(t,t)} - 1)\mu^U(dt) \\ &+ \int_t^T \left(e^{-g(t,u)} - 1\right) \nu(t,du)dt \\ &- \int_{\mathbb{R}} (e^{g(t,t)} - 1)x\,\Gamma(dt,dx) + dM_t^4\end{aligned}$$

with a local martingale M^4. The process $(X_t^{-1}P(t,T))_{0\le t\le T}$ is a local martingale if and only if the predictable part in the semimartingale decomposition vanishes. Letting $t = T$ one recovers

$$0 = \int_0^t (f(s,s) - h(s) - r_s)ds + \sum_{i:U_i\le t} \left(e^{g(U_i,U_i)} - 1 - \Gamma_i e^{g(U_i,U_i)}\right)$$

for $0 \le t \le T^*$, on $\{t < \tau\}$, which is equivalent to $f(s,s) = h(s) + r_s$ and

$$1 - e^{-g(U_i,U_i)} = \Gamma_i$$

on $\{U_i \le T^* \wedge \tau\}$ such that (57) and (58) follow. The converse is easy to see. □

Example 73 (Announced random times)**.** Consider a Poisson process with intensity 1 whose first N jumping times $S_1 < S_2 < \cdots < S_N$ denote the arrival times of news. There is a independent sequence $(\sigma_i)_{i\ge 1}$ of positive random variables with distribution function F_σ and set

$$U_i := S_i + \sigma_i.$$

Then, U_i are announced by S_i and we are just in a setting suggested by **(**B4). Assume for simplicity that $F_\sigma(x) = 1 - e^{-x}$, i.e. σ_1 is standard exponentially distributed and let

$$\tau = \inf\{t \ge 0 : t + \sum_{U_i\le t} 1 \ge \Theta\}$$

with a standard exponential random variable Θ, independent of all other appearing random variables. Then there is no deterministic risky time, i.e. $Q^*(\tau = t) = 0$ for all $t \ge 0$. However, each U_i is a risky time because

$$Q^*(\tau = U_i | S_i, \sigma_i, \tau \ge U_i) = 1 - e^{-1},$$

similar to Equation (60).

4.4 *Affine models*

The aim of this section is to give an affine specification of our credit risky setting. In this case, however, the risky times need to be deterministic, by Theorem 40.

We assume that $\mathcal{U} = \{u_1, \dots, u_N\}$ is the set of deterministic jump times, and – for simplicity – a vanishing short rate $r_t = 0$. The idea is to consider an affine process X and study arbitrage-free doubly stochastic term structure models where the compensator H^p of the default indicator process $H = \mathbb{1}_{\{\cdot \le \tau\}}$ is given by

$$H_t^p = \int_0^t \Big(\phi_0(s) + \psi_0(s)^\top \cdot X_s\Big) ds + \sum_{i=1}^n \mathbb{1}_{\{t \ge u_i\}} \Big(1 - e^{-\phi_i - \psi_i^\top \cdot X_{u_i}}\Big). \tag{74}$$

To ensure that H^p is non-decreasing we will require that $\phi_0(s) + \psi_0(s)^\top \cdot X_s \ge 0$ for all $s \ge 0$ and $\phi_i + \psi_i^\top \cdot X_{u_i} \ge 0$ for all $i = 1, \dots, N$.

Consider a state space in canonical form $\mathcal{X} = \mathbb{R}^m_{\ge 0} \times \mathbb{R}^n$ for integers $m, n \ge 0$ with $m + n = d$ and a d-dimensional Brownian motion W. Let μ and σ be defined on $\mathcal{X}$ by

$$\mu(x) = \mu_0 + \sum_{i=1}^d x_i \mu_i, \tag{75}$$

$$\frac{1}{2}\sigma(x)^\top \sigma(x) = \sigma_0 + \sum_{i=1}^d x_i \sigma_i, \tag{76}$$

where $\mu_0, \mu_i \in \mathbb{R}^d$, $\sigma_0, \sigma_i \in \mathbb{R}^{d \times d}$, for all $i \in \{1, \dots, d\}$. We assume that the parameters μ^i, σ^i, $i = 0, \dots, d$ are admissible in the sense of Theorem 10.2 in Filipović (2009). Then the continuous, unique strong solution of the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x, \tag{77}$$

is an *affine* process X on the state space $\mathcal{X}$ by Theorem 41.

We call a bond-price model *affine* if there exist functions $A : \mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}$, $B : \mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}^d$ such that

$$P^M(t, T) = \mathbb{1}_{\{\tau > t\}} e^{-A(t,T) - B(t,T)^\top \cdot X_t}, \tag{78}$$

for $0 \le t \le T \le T^*$. The following proposition gives sufficient conditions such that the affine model is arbitrage-free.

Proposition 79. *Assume that $\phi_0 : \mathbb{R}_{\ge 0} \to \mathbb{R}$, $\psi_0 : \mathbb{R}_{\ge 0} \to \mathbb{R}^d$ are continuous, $\psi_0(s) + \psi_0(s)^\top \cdot x \ge 0$ for all $s \ge 0$ and $x \in \mathcal{X}$ and the constants $\phi_1, \dots, \phi_n \in \mathbb{R}$ and $\psi_1, \dots, \psi_n \in \mathbb{R}^d$ satisfy $\phi_i + \psi_i^\top \cdot x \ge 0$ for all $1 \le i \le n$ and $x \in \mathcal{X}$. Moreover, let the functions $A : \mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}$ and $B : \mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}^d$ be the unique solutions of*

$$\begin{aligned} A(T, T) &= 0 \\ A(u_i, T) &= A(u_i-, T) - \phi_i \\ -\partial_t^+ A(t, T) &= \phi_0(t) + \mu_0^\top \cdot B(t, T) - \frac{1}{2} B(t, T)^\top \cdot \sigma_0 \cdot B(t, T), \end{aligned} \tag{80}$$

and

$$\begin{aligned} B(T, T) &= 0 \\ B_k(u_i, T) &= B_k(u_i-, T) - \psi_{i,k} \\ -\partial_t^+ B_k(t, T) &= \psi_{0,k}(t) + \mu_k^\top \cdot B(t, T) - \frac{1}{2} B(t, T)^\top \cdot \sigma_k \cdot B(t, T), \end{aligned} \tag{81}$$

for $0 \le t \le T$. Then, the doubly-stochastic affine model given by (78) is a family of local martingales, i.e. the market is free of arbitrages.

Proof. By construction,

$$A(t,T) = \int_t^T a'(t,u)du + \sum_{i:u_i \in (t,T]} \phi_i$$
$$B(t,T) = \int_t^T b'(t,u)du + \sum_{i:u_i \in (t,T]} \psi_i$$

with suitable functions a' and b' and $a'(t,t) = \phi_0(t)$ as well as $b'(t,t) = \psi_0(t)$.

We therefore arrive at a model of the structure

$$P(t,T) = \mathbb{1}_{\{\tau>t\}} \exp\Big(- \int_t^T f(t,u)\mu^M(du)\Big), \tag{82}$$

where $\mu^M(du) = du + \sum_{i=1}^n \delta_{u_i}(du)$. In comparison with the HJM model, (78), yields the following: on the one hand, for $T = u_i \in \mathcal{U}$, we obtain

$$f(t,u_i) = \phi_i + \psi_i^\top \cdot X_t. \tag{83}$$

Hence, the coefficients $a(t,T)$ and $b(t,T)$ in (53) for $T = u_i \in \mathcal{U}$ compute to $a(t,u_i) = \psi_i^\top \cdot \mu(X_t)$ and $b(t,u_i) = \psi_i^\top \cdot \sigma(X_t)$.

On the other hand, for $T \notin \mathcal{U}$ we obtain that $f(t,T) = a'(t,T) + b'(t,T)^\top \cdot X_t$. Then, the coefficients $a(t,T)$ and $b(t,T)$ in the dynamics of $f(t,T)$, see Equation (53), can be computed by Itô's formula as follows:

$$\begin{aligned} a(t,T) &= \partial_t a'(t,T) + \partial_t b'(t,T)^\top \cdot X_t + b'(t,T)^\top \cdot \mu(X_t) \\ b(t,T) &= b'(t,T)^\top \cdot \sigma(X_t). \end{aligned} \tag{84}$$

For the following, set $\bar{a}'(t,T) = \int_t^T a'(t,u)du$ and $\bar{b}'(t,T) = \int_t^T b'(t,u)du$ and note that,

$$\int_t^T \partial_t a'(t,u)du = \partial_t \bar{a}'(t,T) + a'(t,t).$$

As $\partial_t^+ A(t,T) = \partial_t \bar{a}'(t,T)$, and $\partial_t^+ B(t,T) = \partial_t \bar{b}'(t,T)$, we obtain from (84) that

$$\begin{aligned} \bar{a}(t,T) &= \int_t^T a(t,u)\mu^M(du) = \int_t^T a(t,u)du + \sum_{u_i \in (t,T]} \psi_i^\top \cdot \mu(X_t) \\ &= \partial_t^+ A(t,T) + a'(t,t) + \big(\partial_t^+ B(t,T) + b'(t,t)\big)^\top \cdot X_t + B(t,T)^\top \cdot \mu(X_t), \\ \bar{b}(t,T) &= \int_t^T b(t,u)\mu^M(du) = \int_t^T b(t,u)du + \sum_{u_i \in (t,T]} \psi_i^\top \cdot \sigma(X_t) \\ &= B(t,T)^\top \cdot \sigma(X_t) \end{aligned}$$

for $0 \le t \le T \le T^*$. We now show that under our assumptions, the drift conditions hold. We first concentrate on (57), which in our setting reads

$$\begin{aligned} \int_0^t f(s,s)\mu^M(ds) &= \int_0^t f(s,s)ds + \sum_{u_i \le t} f(u_i,u_i) \\ &= \int_0^t h_s ds - \sum_{u_i \le t} \log(1-\Gamma_i) \\ &= \int_0^t (\phi_0(s) + \psi_0(s)^\top \cdot X_s)ds + \sum_{u_i \le t} \phi_i + \psi_i^\top \cdot X_{u_i} \end{aligned}$$

Since $f(s,s) = a'(s,s) + b'(s,s)^\top \cdot X_s$ and $a'(s,s) = \phi_0(s)$ as well as $b'(s,s) = \psi_0(s)$, this condition is clearly satisfied. Next, we observe that for the second drift condition, (58), in our setting

$$\int_t^T a(t,u)du + \int_t^T a(t,u)du\mu_t(du) = \int_t^T a(t,u)\mu^M(du) = \bar{a}(t,T). \tag{85}$$

Then the second drift condition follows by (80), (81), and the affine specification (75), and (76). □

Example 86. In the one-dimensional case we consider X, given as solution of

$$dX_t = (\mu_0 + \mu_1 X_t)dt + \sigma\sqrt{X_t}dW_t, \qquad t \geq 0.$$

We assume for simplicity that $u_1 = 1$ and $N = 1$ such that there is a single risky time, 1, and choose $\mu^M(du) = \delta_1(du)$. Moreover, let $\phi_0 = 0$, $\psi_0 = 1$ as well as $\phi_1 = 0$ and $\psi_1 \geq 0$, such that

$$H^p = \int_0^t X_s ds + \mathbb{1}_{\{t\geq 1\}}(1 - e^{-\psi_1 X_1}).$$

Hence the probability of having no default at time 1 just prior to 1 is given by $e^{-\psi_1 X_1}$, compare Example 59.

An arbitrage-free model can be obtained by choosing A and B according to Proposition 79 which can be immediately achieved using Lemma 10.12 from Filipović (2009) (see in particular Section 10.3.2.2 on the CIR short-rate model): denote $\theta = \sqrt{\mu_1^2 + 2\sigma^2}$ and

$$\begin{aligned}
L_1(t) &= 2(e^{\theta t} - 1),\\
L_2(t) &= \theta(e^{\theta t} + 1) + \mu_1(e^{\theta t} - 1),\\
L_3(t) &= \theta(e^{\theta t} + 1) - \mu_1(e^{\theta t} - 1),\\
L_4(t) &= \sigma^2(e^{\theta t} - 1).
\end{aligned}$$

Then

$$A_0(s) = \frac{2\mu_0}{\sigma^2}\log\Big(\frac{2\theta e^{\frac{(\sigma-\mu_1)t}{2}}}{L_3(t)}\Big), \quad B_0(s) = -\frac{L_1(t)}{L_3(t)}$$

are the unique solutions of the Riccati equations $B_0' = \sigma^2 B_0^2 - \mu_1 B_0$ with boundary condition $B_0(0) = 0$ and $A_0' = -\mu_0 B_0$ with boundary condition $A_0(0) = 0$. Note that with $A(t,T) = A_0(T-t)$ and $B(t,T) = B_0(T-t)$ for $0 \leq t \leq T < 1$, the conditions of Proposition 79 hold. Similarly, for $1 \leq t \leq T$, choosing $A(t,T) = A_0(T-t)$ and $B(t,T) = B_0(T-t)$ implies again the validity of (80) and (81). On the other hand, for $0 \leq t < 1$ and $T \geq 1$ we set $u(T) = B(1,T) + \psi_1 = B_0(T-1) + \psi_1$, according to (81), and let

$$\begin{aligned}
A(t,T) &= \frac{2\mu_0}{\sigma^2}\log\Big(\frac{2\theta e^{\frac{(\sigma-\mu_1)(1-t)}{2}}}{L_3(1-t) - L_4(1-t)u(T)}\Big)\\
B(t,T) &= -\frac{L_1(1-t) - L_2(1-t)u(T)}{L_3(1-t) - L_4(1-t)u(T)}.
\end{aligned}$$

It is easy to see that (80) and (81) are also satisfied in this case, in particular $\Delta A(1,T) = -\phi_1 = 0$ and $\Delta B(1,T) = -\psi_1$. Note that, while X is continuous, the bond prices are not even stochastically continuous because they jump almost surely at $u_1 = 1$. We conclude by Proposition 79 that this affine model is arbitrage-free. ◇

IV
Insurance-Finance Arbitrage

In this chapter we analyse the relation of insurance products to financial markets based on a fundamental theorem developed in Artzner et al. (2024). The goal is to value finance-related insurance claims like equity-linked insurance contracts or variable annuities. The key difference to financial approaches is that insurance contracts are (typically) not traded and the related information flow is not publicly available.

We shortly recall Bayes' rule for the equivalent probability measures Q and P on a measurable space $(\Omega, \mathscr{G})$.

Proposition 1 (Bayes)**.** *Assume that $dQ = LdP$ with $E_P[L|\mathscr{F}] > 0$. Then, for every random variable $X \geq 0$, it holds that*

$$E_Q[X|\mathscr{F}] = \frac{E_P[LX|\ \mathscr{F}]}{E_P[L|\mathscr{F}]}, \qquad P\text{-}f.s. \tag{2}$$

Our first goal is to provide a technical tool for valuing insurance products. This rule should be arbitrage-free for financial products and should also include statistical information in an appropriate way. For this we rely on the statistical measure P on $(\Omega, \mathscr{G})$. Financial information is given by a sub-σ-field $\mathscr{F} \subset \mathscr{G}$ and we assume that there exists an equivalent martingale measure Q on $(\Omega, \mathscr{F}, P|_{\mathscr{F}})$.

Proposition 3. *Let $Q \sim P|_{\mathscr{F}}$. Then, there exists a unique measure $\mathrm{Q}\odot\mathrm{P}$ on $(\Omega, \mathscr{G})$, such that*

(i) $\mathrm{Q}\odot\mathrm{P} = Q$ *on* $\mathscr{F}$ *and*

(ii) for alll $G \in \mathscr{G}$ it holds that $\mathrm{Q}\odot\mathrm{P}(G|\mathscr{F}) = P(G|\mathscr{F})$ (P-almost surely).

The first property states that – as wanted – $\mathrm{Q}\odot\mathrm{P}$ coincides with Q on the financial information. Hence, all financial products are automatically valued in an arbitrage-free sense. The second property is a little bit more subtle. Conditional on *all* the financial information the measure $\mathrm{Q}\odot\mathrm{P}$ coincides with the statistical measure. In particular, for those claims which are independent from the financial markets, the rule provides the classical insurance valuation by conditional expectation.

We now generalise this to the conditional level.

Proposition 4. *Assume $Q \sim P|_{\mathscr{F}}$ and let $X \geq 0$. Then,*

(i) for every σ-field $\mathscr{H} \subset \mathscr{F}$ it holds that:

$$E_{\mathrm{Q}\odot\mathrm{P}}[X|\mathscr{H}] = E_Q\big[E_P[X|\mathscr{F}]|\mathscr{H}\big],$$

(ii) and for every σ-field $\mathscr{H}$ with $\mathscr{F} \subset \mathscr{H} \subset \mathscr{G}$,

$$E_{\mathrm{Q}\odot\mathrm{P}}[X|\mathscr{H}] = E_P\big[X|\mathscr{H}\big].$$

Both propositions. We start with existence: Since $Q \sim P|_{\mathscr{F}}$ on $(\Omega, \mathscr{F})$ the Radon-Nikodym theorem yields the existence of a $\mathscr{F}$-measurable density L such that $dQ = L\,dP$. Define $\mathrm{Q}\odot\mathrm{P}$ by

$$d(\mathrm{Q}\odot\mathrm{P}) = L\,dP.$$

We obtain that $\mathrm{Q}\odot\mathrm{P}(G) = E_P[\mathbb{1}_G L]$. Then, by construction, (i) of Proposition 3 holds:

$$\mathrm{Q}\odot\mathrm{P}(F) = \int_F L dP = \int_F dQ = Q(F)$$

for all $F \in \mathscr{F}$.

For $G \in \mathscr{G}$ we have that

$$\begin{aligned}\int_F \mathbb{1}_G d(\mathrm{Q}\odot\mathrm{P}) &= \int_F \mathbb{1}_G L dP = \int_F LP(G|\mathscr{F})dP \\ &= \int_F P(G|\mathscr{F})d(\mathrm{Q}\odot\mathrm{P}), \qquad F \in \mathscr{F},\end{aligned} \tag{5}$$

such that $\mathrm{Q}\odot\mathrm{P}(G|\mathscr{F}) = P(G|\mathscr{F})$ ($\mathrm{Q}\odot\mathrm{P}$-a.s. and hence Q- as well as P-almost surely since all these measures have the same nullsets) and hence also (ii) of Proposition 3 holds.

For uniqueness consider R, such that (i) and (ii) of Proposition 3 hold. Then, for all $G \in \mathscr{G}$,

$$R(G) = \int R(G|\mathscr{F})dR \overset{(ii)}{=} \int P(G|\mathscr{F})\,dR \overset{(i)}{=} \int P(G|\mathscr{F})dQ = \mathrm{Q}\odot\mathrm{P}(G) \tag{6}$$

and hence $\mathrm{Q}\odot\mathrm{P} = R$.

Now consider $H \in \mathscr{H} \subset \mathscr{F}$, such that $H \in \mathscr{F}$. Analogous to (5), $\mathrm{Q}\odot\mathrm{P}(G|\mathscr{H}) = P(G|\mathscr{H})$. With $H = \Omega$,

$$\int \mathbb{1}_G d(\mathrm{Q}\odot\mathrm{P}) = \int LP(G|\mathscr{H})dP = \int P(G|\mathscr{H})dQ \tag{7}$$

and hence (i) follows by approximation with elementary functions.

Finally, from $\mathscr{F} \subset \mathscr{H} \subset \mathscr{G}$ it follows by Bayes rule that

$$E_{\mathrm{Q}\odot\mathrm{P}}[X|\mathscr{H}] = \frac{E_P[L_T X|\,\mathscr{H}]}{E_P[L_T|\,\mathscr{H}]} = E_P[X|\mathscr{H}],$$

since L_T is $\mathscr{F}$-measurable ($E_P[L_T|\,\mathscr{H}] > 0$ holds since $P|_{\mathscr{F}}$ and Q are equivalent). □

In a typical application we will have two filtrations $\mathbb{G} = (\mathscr{G}_t)_{t=0,\dots,T}$ and $\mathbb{F} = (\mathscr{F}_t)_{t=0,\dots,T}$ such that

$$\mathscr{F}_t \subseteq \mathscr{G}_t, \qquad t = 0, \dots, T.$$

Here $\mathbb{F}$ is publicly available information and $\mathbb{G}$ is the information available to the insurance company, which contains a lot of private information. Possibly it could also allow for arbitrage, which we will take into account in the following.

The QP-rule

The above result allows the valuation of an insurance product by the so-called *QP-rule*. It is similar to the risk-neutral valuation rule and prices the insurance product by computing the conditional expectation of the cumulated and discounted payuments under the measure $Q\odot P$. Note that by proposition 4 for $X \geq 0$

$$E_{Q\odot P}[X|\mathscr{F}_t] = E_Q\Big[\, E_P[X|\mathscr{F}_T]\,|\mathscr{F}_t\Big]. \tag{8}$$

This explains the name QP-rule: first we condition on the evolution of the financial market until T and compute the expected value under the statistical measure. The result is of course $\mathscr{F}_T$-measurable and is valued like a classical European option by the risk-neutral pricing rule, i.e. by taking expectations under Q.

Example 9 (Stochastic mortality)**.** Inspired by our knowledge on credit risk, we can now introduce a quite flexible framework for the valuation of life-insurance. As a typical example think of a pension fund or a life insurance, which promises the (discounted) payment X_T at the maturity time T, if the insured is still alive. We introduce its survival time τ as a $\mathbb{G}$-stopping time (typically not a $\mathbb{F}$-stopping time). A classical model for this is the so-called doubly stochastic approach:

Let $E \sim \mathrm{Exp}(1)$ be a $\mathscr{G}_T$-measurable standard exponential random variable which is independent of $\mathscr{F}_T$. Moreover, let $\Lambda = (\Lambda_t)_{t=0,\dots,T}$ be increasing and, $\mathbb{F}$-adapted such that $\Lambda_0 = 0$. Set

$$\tau = \inf\{t \geq 0 : \Lambda_t \geq E\}, \tag{10}$$

with the convention $\inf \varnothing = \infty$. Then the first-entry time τ is not a $\mathbb{F}$-stopping time, but $\mathscr{G}_T$-measurable. (Currently τ might not even be a $\mathbb{G}$-stopping time. This, however, can be achieved for example by considering $\mathscr{G}_t = \sigma(\mathscr{F}_T, \tau \wedge t)$.)

We compute

$$\begin{aligned} P(\tau > T|\, \mathscr{F}_T) &= E_P\big[\mathbb{1}_{\{\tau > T\}}|\, \mathscr{F}_T\big] \\ &= E_P\big[\mathbb{1}_{\{\Lambda_T < E\}}|\, \mathscr{F}_T\big] \\ &= e^{-\Lambda_T}, \end{aligned}$$

since Λ_T is $\mathscr{F}_T$-measurable and E independent of $\mathscr{F}_T$. The we obtain

$$\begin{aligned} E_{Q\odot P}\Big[X_T \mathbb{1}_{\{\tau > T\}}|\mathscr{F}_t\Big] &= E_Q\Big[E_P[X_T \mathbb{1}_{\{\tau > T\}}|\mathscr{F}_T]|\, \mathscr{F}_t\Big] \\ &= E_Q\Big[e^{-\Lambda_T}|\, \mathscr{F}_t\Big] \end{aligned}$$

as price of the insurance

Often τ is modelled in continuous time with an intensity or mortality rate. This is the $\mathbb{F}$-adapted, non-negative process λ which satisfies

$$\Lambda_t := \int_0^t \lambda_s ds, \qquad t \geq 0.$$

1 *The fundamental theorem of insurance-finance markets*

Up to now, the QP-rule was just a rule without a foundational explanation. In this section we will analyse its role from a more economic viewpoint and introduce the trading strategies of the insurance. If the insurance is able to make a risk-less profit, an insurance-finance arbitrage exists, which of course should be avoided.

The key is to use an appropriate version of the strong law of large number. The insurance can build larger and larger portfolios of insurance clients and by this reduce its risk. If the risk is (in the limit) vanishing, we obtain a very precise rule in the following fundamental theorem. This result will show, that - under the appropriate assumptions - the QP-rule is actually a rule which guarantees absence of insurance-finance arbitrage.

To this we consider a time point $t \in \{0, \dots, T-1\}$. At this time point an insurance can be bought which offers a payment (wlog) at maturity T. We call the discounted payment the *insurance benefit*

$$B_{t,T} \geq 0.$$

Of course $B_{t,T}$ is a $\mathscr{G}_T$-measurable random variable. On the other side the insured pays a premium, which we denote by $p_t \geq 0$. This premium is $\mathscr{G}_t$-measurable.

For the insurance portfolio we consider insurance clients, numbered $1, 2, \dots$ which are willing to contract insurances at arbitrary fractions. The payment of course differ from client to client (in contrast to financial contracts) and we denote the benefit of client i by

$$B^i_{t,T}.$$

Again $B^i_{t,T}$ is $\mathscr{G}_T$-measurable.

Example 11 (Stochastic mortality II)**.** Consider a portfolio of insurance clients with stochastic mortality. To this, let $E_1, E_2, \dots$ independent, standard exponential random variables which are independent of $\mathscr{F}_T$ and let Λ be $\mathbb{F}$-adapted and increasing with $\Lambda_0 = 0$. Set

$$\tau_i := \inf\{t \geq 0 : \Lambda_t \geq E_i\}, \quad i \geq 1.$$

The payoff for the i-th insured is

$$B^i_{t,T} = X_T \mathbb{1}_{\{\tau_i > T\}}$$

which can be evaluated with the QP-rule. ⋄.

Define

$$\mathscr{G}_{t,T} = \mathscr{G}_t \vee \mathscr{F}_T, \tag{12}$$

where $\mathscr{G}_t \vee \mathscr{F}_T$ is the from $\mathscr{G}_t$ and $\mathscr{F}_T$ generated σ-algebra. Foundational will be the following assumption:

Assumption 13. *Assume that the following holds:*

(i) The random variables $B^i_{t,T}$, $i = 1, 2, \dots$ are independent conditional on $\mathscr{G}_{t,T}$,

(ii) $E[B^i_{t,T} | \mathscr{G}_{t,T}] = E[B^1_{t,T} | \mathscr{G}_{t,T}]$, $i = 1, 2, \dots$, *and*

(iii) $\mathrm{Var}[B^i_{t,T} | \mathscr{G}_{t,T}] = \mathrm{Var}[B^1_{t,T} | \mathscr{G}_{t,T}]$, $i = 1, 2, \dots$.

The insurance portfolio

To reduce risk, the insurance forms a portfolio in a suitable way, which we call an *allocation*. A allocation at time t is a *sequence* of random variables $\psi_t = (\psi_t^1, \psi_t^2, \dots)$. Each allocation has only finitely many non-zero entries (like for example $1, 2, 3, 0, 0, \dots$). Here, $psi_t^i = 1$ means that the insurance sells on contract to the insurance client i. Summarising, an allocation is associated to the following payments:

$$\sum_{i \geq 1} \psi_t^i \, B_{t,T}^i.$$

On the other side, those insurance contracts are sold for a *premium*. Even if the benefits all differ (since they depend on client-specific information) we consider a homogenous group, meaning they all come at the same price, p_t. Hence the sum of the premia is given by

$$\sum_{i \geq 1} \psi_t^i p_t = \Big(\sum_{i \geq 1} \psi_t^i \Big) p_t.$$

And we arrive at the profit and loss of the allocation $\psi = (\psi_t)_{t=0,\dots,T}$ given by

$$V_T(\psi) := \sum_{t=0}^{T-1} \sum_{i \geq 1} \psi_t^i \big(p_t - B_{t,T}^i \big). \tag{14}$$

Inspired by large financial markets we think of an insurance strategy as a sequence of allocations. The intuition is that the insurance company can take more and more clients into the portfolio and hence reduce the risk by diversification. We allow appropriate bounded portfolios, but possibly with infinitely many clients.

To this end, we call a *insurance strategy* a sequence $(\psi^n)_{n \geq 1}$ of allocations. The strategy is called *admissible*, if the following conditions hold:

(i) *Uniform boundedness*: there exists $C > 0$, such that

$$\parallel \psi_t^n \parallel := \sum_{i \geq 1} \psi_t^{n,i} \leq C \tag{15}$$

for all $n \geq 1$ and $0 \leq t < T$,

(ii) *convergence of the total mass*: there exists $0 < \gamma_t \in \mathscr{F}_t$, such that

$$\parallel \psi_t^n \parallel \to \gamma_t \quad \text{a.s. for all } t < T, \tag{16}$$

(iii) *convergence of the limit*: there exists a random variable $V = V^\psi$, such that

$$\lim_{n \to \infty} V_T^I(\psi^n) = V,$$

P-a.s.

In addition to the insurance part of the strategy, the insurance can also trade on the financial market. We assume that this is done in a classical, discrete-time manner (although this can easily be relaxed).

We assume that the financial market consists of $d + 1$ assets $S = (S^0, \dots, S^d)$ with numeraire $S^0 > 0$. The discounted price process is denoted by X. A self-financing trading strategy is given by a d-dimensional, $\mathbb{F}$-predictable process H. The associated gains process is given by

$$G_T(H) := \sum_{t=1}^{T-1} H_t \cdot \Delta X_t = \sum_{t=1}^{T-1} \sum_{i=1}^{d} H_t^i \cdot \Delta X_t^i,$$

where $\Delta X_t = X_{t+1} - X_t$. We assume that the financial markets is free of arbitrage. This is, by the fundamental theorem, equivalent to the existence of a martingale measure which is equivalent to $P|_{\mathscr{F}_T}$, in short

$$\mathcal{M}_e(\mathbb{F}) \neq \varnothing. \tag{17}$$

The insurance-finance market (B, p, S) therefore consists of three parts: the benefits B, the premia p and the assets S.

Definition 18. In the insurance-finance market (B, p, S) there exists an *arbitrage*, if there exists an admissible insurance strategy $(\psi^n)_{n \geq 1}$ and a self-financing trading strategy H such that

$$\lim_{n \to \infty} V_T(\psi^n) + G_T(H) \in L_0^+ \backslash \{0\}. \tag{19}$$

Otherwise we say that the insurance-finance market is free of arbitrage.

Recall that for a family of random variables Ξ,

$$\operatorname{esssup}_{\mathscr{F}} \Xi := \operatorname{ess\,inf}\{q \mid q \text{ is } \mathscr{F}\text{-measurable and } q \geq \xi \text{ for all } \xi \in \Xi\}. \tag{20}$$

We denote by

$$p_t^{\uparrow} = \operatorname{esssup}_{\mathscr{F}_t} p_t, \quad \text{and} \quad p_t^{\downarrow} = \operatorname*{ess\,inf}_{\mathscr{F}_t} p_t. \tag{21}$$

Theorem 22. *Consider the insurance-finance market (B, p, S) and assume that Assumption 13 holds.*

(i) If there exists $Q \in \mathcal{M}_e(\mathbb{F})$, such that for all $t < T$

$$p_t \leq E_{Q \odot P}[X_{t,T} | \mathscr{F}_t], \qquad P\text{-a.s.}, \tag{23}$$

then there is no insurance-finance arbitrage.

(ii) If there exists $t < T$, such that

$$P\Big(\bigcap_{Q \in \mathcal{M}_e(\mathbb{F})} \big\{ p_t^{\downarrow} > E_{Q \odot P}\big[X_{t,T} | \mathscr{F}_t\big] \big\} \Big) > 0 \tag{24}$$

then there exists and insurance-finance arbitrage.

Intuitively, the first assertion shows that - if you use the QP rule for pricing, there is never an insurance-finance arbitrage. For (ii), note that if you are above a QP-price, there actually might be Q' and the associated $Q'P$-price could be higher and there would be again no arbitrage. So, we have to make sure that for all martingale measures Q the probability that the price is higher than the associated QP price has a positive probability. Then, (ii) shows that under this minimal assumption, there is an insurance-finance arbitrage. Note that this is not an if-and-only-if result, which to the best of my knowledge is unknown today.

We start with a technical result, showing that under our assumptions the insurance strategies converge.

Satz 25. *Assume that 13, (15) and (16) hold. Then for every admissible insurance strategy* (ψ^n) *and for every* $Q \in \mathcal{M}_e(\mathbb{F})$ *it holds that*

$$E_{Q\odot P}\Big[\lim_{n\to\infty}\sum_{i\geq 1}\psi_t^{n,i}\, p_t\Big] = E_{Q\odot P}[\gamma_t\, p_t], \qquad \text{for all} t < T \qquad \text{and} \tag{26}$$

$$E_{Q\odot P}\Big[\lim_{n\to\infty}\sum_{t=0}^{T-1}\sum_{i\geq 1}\psi_t^{n,i}B_{t,T}^i\Big] = \sum_{t<T} E_{Q\odot P}\big[\gamma_t B_{t,T}\big]. \tag{27}$$

Proof. For (26) we use dominated convergence: by our assumption on uniform boundedness, Equation (15),

$$\sum_{i\geq 1}\psi_t^{n,i}\, p_t \leq C p_t.$$

This allows us to interchange limit and summation in the following expression. Together with our assumption on convergence of the total mass, Equation (16), it follows for $0 \leq t < T$, that

$$\lim_{n\to\infty}\sum_{i\geq 1}\psi_t^{n,i}\, p_t = \gamma_t\, p_t.$$

Recall that $V_T^I(\psi^n) = \sum_{t<T}\sum_{i\geq 1}\psi_t^{n,i}B_{t,T}^i$. As in Proposition 3, we denote the Radon-Nikodym density of Q with respect to $P|_{\mathscr{F}_T}$ by L_T. Then,

$$\begin{aligned} E_{Q\odot P}\Big[\lim_{n\to\infty} V_T^I(\psi^n)\Big] &= E_P\Big[L_T \lim_{n\to\infty} V_T^I(\psi^n)\Big] \\ &= E_P\Big[L_T E_P\big[\lim_{n\to\infty} V_T^I(\psi^n)|\mathscr{G}_{t,T}\big]\Big]. \end{aligned}$$

For any $G \in \mathscr{G}_{t,T}$, we have

$$\begin{aligned} \int_G E_P\big[\lim_{n\to\infty} V_T^I(\psi^n)|\mathscr{G}_{t,T}\big]dP &= \int_G \lim_{n\to\infty} V_T^I(\psi^n)dP \\ &= \int \liminf_{n\to\infty} \mathbb{1}_G V_T^I(\psi^n)dP \\ &\leq \liminf_{n\to\infty}\int_G V_T^I(\psi^n)dP = \liminf_{n\to\infty}\int_G E_P\big[V_T^I(\psi^n)|\mathscr{G}_{t,T}\big]dP. \end{aligned}$$

Furthermore,

$$\begin{aligned} E_P\big[V_T^I(\psi^n)|\mathscr{G}_{t,T}\big] &= \sum_{t<T}\sum_{i\geq 1}\psi_t^{n,i}E_P\big[X_{t,T}^i|\mathscr{G}_{t,T}\big] = \sum_{t<T}\sum_{i\geq 1}\psi_t^{n,i}E_P\big[X_{t,T}|\mathscr{G}_{t,T}\big] \\ &= \sum_{t<T} E_P\big[X_{t,T}|\mathscr{G}_{t,T}\big] \parallel \psi_t^n \parallel. \end{aligned}$$

Hence,

$$\begin{aligned} \liminf_{n\to\infty}\int_G E_P\big[V_T^I(\psi^n)|\mathscr{G}_{t,T}\big]dP &= \liminf_{n\to\infty}\int_G \sum_{t<T} E_P\big[X_{t,T}|\mathscr{G}_{t,T}\big] \parallel \psi_t^n \parallel dP \\ &= \int_G E_P\big[\sum_{t<T}\gamma_t X_{t,T}|\mathscr{G}_{t,T}\big]dP, \end{aligned}$$

where we used uniform boundedness, (16), and dominated convergence for the last equality. We obtain that

$$E_{Q\odot P}\Big[\lim_{n\to\infty} V_T^I(\psi^n)\Big] \leq \sum_{t<T} E_{Q\odot P}\big[\gamma_t X_{t,T}\big]. \tag{28}$$

On the other side, for any $G \in \mathscr{G}_{t,T}$, we have

$$\begin{aligned}\int_G E_P\big[\lim_{n\to\infty} V_T^I(\psi^n)|\mathscr{G}_{t,T}\big]dP &= \int_G \lim_{n\to\infty} V_T^I(\psi^n)dP \\ &= \int \limsup_{n\to\infty} \mathbb{1}_G V_T^I(\psi^n)dP \\ &\geq \limsup_{n\to\infty} \int_G V_T^I(\psi^n)dP = \limsup_{n\to\infty} \int_G E_P\big[V_T^I(\psi^n)|\mathscr{G}_{t,T}\big]dP.\end{aligned}$$

Hence, we obtain as above

$$E_{Q\odot P}\Big[\lim_{n\to\infty} V_T^I(\psi^n)\Big] \geq \sum_{t<T} E_{Q\odot P}\big[\gamma_t X_{t,T}\big] \tag{29}$$

and the claim follows. □

Proof of Theorem 22. (i) Assume that (23) holds and that we had an insurance-finance arbitrage, i.e.

$$\lim_{n\to\infty} V_T(\psi^n) + G_T(H) \in L_0^+ \backslash \{0\}$$

for some insurance portfolio strategy $(\psi^n)_{n\geq 1} = (\psi_t^n)_{n\geq 1, t<T}$ and some financial strategy $(\xi_t)_{t\leq T}$. First, for any $Q \in \mathcal{M}_e(\mathbb{F})$, $E_{Q\odot P}[G_T(H)] = E_Q[G_T(H)] = 0$ by Proposition 3 (i). Then,

$$\begin{aligned}E_{Q\odot P}\Big[\lim_{n\to\infty} V_T(\psi^n) + G_T(H)\Big] &= E_{Q\odot P}\Big[\lim_{n\to\infty} V_T(\psi^n)\Big] \\ &= E_{Q\odot P}\Big[\lim_{n\to\infty} \sum_{t=0}^{T-1}\sum_{i\geq 1} \psi_t^i (p_t - X_{t,T}^i)\Big].\end{aligned} \tag{30}$$

With the equations (26), (27) we obtain that

$$(30) = \sum_{t=0}^{T-1} E_{Q\odot P}\Big[\ \gamma_t \left(p_t - E_{Q\odot P}[X_{t,T}|\mathscr{F}_T]\right)\Big].$$

We consider the specific Q from (23). Then, $p_t \leq E_{Q\odot P}[X_{t,T}|\mathscr{F}_t]$, P-a.s., hence

$$(30) \leq \sum_{t=0}^{T-1} E_{Q\odot P}\Big[\ \gamma_t \Big(E_{Q\odot P}[X_{t,T}|\mathscr{F}_T] - E_{Q\odot P}[X_{t,T}|\mathscr{F}_T]\Big)\Big] = 0,$$

a contradiction to the assumption of an insurance-finance arbitrage.

(ii) Assume that (24) holds. We set,

$$A_t := \bigcap_{Q\in\mathcal{M}_e(\mathbb{F})} \Big\{p_t^{\downarrow} > E_{Q\odot P}\big[X_{t,T}\big|\mathscr{F}_t\big]\Big\} \in \mathscr{F}_t$$

and, by assumption, $P(A_t) > 0$.

This implies that

$$\mathbb{1}_{A_t} p_t \geq \mathbb{1}_{A_t} p_t^{\downarrow} \geq \mathbb{1}_{A_t} \operatorname*{esssup}_{Q\in\mathscr{M}_e(\mathbb{F})} E_{Q\odot P}\big[X_{t,T}\big|\mathscr{F}_t\big] = \mathbb{1}_{A_t} \operatorname*{esssup}_{Q\in\mathscr{M}_e(\mathbb{F})} E_Q\Big[E_P\big[X_{t,T}\big|\mathscr{G}_{t,T}\big]\Big|\mathscr{F}_t\Big] =: \pi_t. \tag{31}$$

At t we take the uniform allocation $\psi_t^n = \mathbb{1}_{A_t}(n^{-1},...,n^{-1},0,...)$ over the first n insurance seekers, restricted to the set A_t and $\psi_s^n = 0$ for all $s \neq t < T$. This is an admissible strategy and since

$$\sum_{i\geq 1} \frac{1}{i^2}\operatorname{Var}(X_{t,T}^i|\mathscr{G}_{t,T}) = \operatorname{Var}(X_{t,T}|\mathscr{G}_{t,T})\sum_{i\geq 1}\frac{1}{i^2} < \infty,$$

we are entitled to apply the conditional strong law of large numbers given in Theorem 3.5 in Majerek et al. (2005). Hence with Assumption 13 and the fact that $\gamma_t = \sum_{i\geq 1} \psi_t^{n,i} = \mathbb{1}_{A_t} \in \mathscr{F}_t$, we get

$$\sum_{i\geq 1} \psi^{n,i} \; X^i_{t,T} \to \mathbb{1}_{A_t} E_P[X_{t,T}|\mathscr{G}_{t,T}] =: H, \tag{32}$$

P-almost surely as $n \to \infty$. Therefore

$$\lim_{n\to\infty} V_T^I(\psi^n) = \mathbb{1}_{A_t} \; \big(p_t - E_P[X_{t,T}|\mathscr{G}_{t,T}]\big) \geq \mathbb{1}_{A_t} \; \big(p_t^{\downarrow} - H\big).$$

As π_t is the conditional superhedging price of H, we obtain from Theorem 7.2 in Föllmer & Schied (2004)[1] that there is a superhedging strategy $\xi = \mathbb{1}_{A_t} \xi$ such that

$$\mathbb{1}_{A_t}\big(\pi_t + \sum_{s=t}^{T-1} \xi_s \cdot \Delta S_s\big) \geq H. \tag{33}$$

Using this financial trading strategy ξ, we find from (31) that

$$\lim_{n\to\infty} V_T^I(\psi^n) + V_T^F(\xi) \geq \mathbb{1}_{A_t} \; \big(p_t^{\downarrow} - H + \sum_{s=t}^{T-1} \xi_s \cdot \Delta S_s\big) \tag{34}$$

almost surely. Now, using (31) and (33), we obtain

$$(34) \geq \mathbb{1}_{A_t}\big(\pi_t - \pi_t\big) = 0. \tag{35}$$

For the final step we have to distinguish if the claim H is replicable or not. For the first case let

$$B_t := \big\{ \operatorname*{ess\,sup}_{Q\in\mathscr{M}_e(\mathbb{F})} E_Q\big[H\big|\mathscr{F}_t\big] = \operatorname*{ess\,inf}_{Q\in\mathscr{M}_e(\mathbb{F})} E_Q\big[H\big|\mathscr{F}_t\big]\big\}$$

and assume $P(A_t \cap B_t) > 0$. By assumption (24), we have $p_t^{\downarrow} > \pi_t$ on a set of positive probability. This allows to drop equality in (35): indeed, since $p_t^{\downarrow} > \pi_t$ with positive probability, we obtain from (34) that

$$(34) \geq \mathbb{1}_{A_t\cap B_t} \; \big(p_t^{\downarrow} - H + \sum_{s=t}^{T-1} \xi_s \cdot \Delta S_s\big) > \mathbb{1}_{A_t\cap B_t}\big(\pi_t - \pi_t\big) = 0, \tag{36}$$

with positive probability. Hence, $\lim_{n\to\infty} V_T^I(\psi^n) + V_T^F(\xi) \neq 0$ and therefore

$$\lim_{n\to\infty} V_T^I(\psi^n) + V_T^F(\xi) \in L^0_+\backslash\{0\}.$$

For the second case let

$$B'_t := A_t \cap \big\{ \operatorname*{ess\,sup}_{Q\in\mathscr{M}_e(\mathbb{F})} E_Q\big[H\big|\mathscr{F}_t\big] > \operatorname*{ess\,inf}_{Q\in\mathscr{M}_e(\mathbb{F})} E_Q\big[H\big|\mathscr{F}_t\big]\big\}$$

and assume $P(A_t \cap B'_t) > 0$. Again, we can drop equality in (35): indeed, we obtain analogously that

$$(34) \geq \mathbb{1}_{A_t\cap B'_t} \; \big(p_t^{\downarrow} - H + \sum_{s=t}^{T-1} \xi_s \cdot \Delta S_s\big) \geq \mathbb{1}_{A_t\cap B'_t} \; \big(\pi_t - H + \sum_{s=t}^{T-1} \xi_s \cdot \Delta S_s\big) \geq 0. \tag{37}$$

Since on B'_t, the no-arbitrage interval of the European claim H is a true interval, the upper bound of the conditional no-arbitrage interval, π_t, already yields a (financial) arbitrage (on B'_t). Hence, $\mathbb{1}_{A_t\cap B'_t} \; \big(\pi_t - H + \xi\Delta S\big) \in L^0_+\backslash\{0\}$,and, in addition,

$$\lim_{n\to\infty} V_T^I(\psi^n) + V_T^F(\xi) \in L^0_+\backslash\{0\}.$$

The existence of an insurance-finance arbitrage is proved. □

[1] Compare Theorem 2
& Schachermayer (20
however requires a fi

Bibliography

Artzner, P., Eisele, K.-T. & Schmidt, T. (2024), 'Insurance - finance arbitrage', *Mathematical Finance, doi:10.1111/mafi.12412* pp. 1–35.

Bachelier, L. (1900), Théorie de la spéculation, *in* 'Annales scientifiques de l'École normale supérieure', Vol. 17, pp. 21–86.

Black, F. & Cox, J. C. (1976), 'Valuing corporate securities: some effects of bond indenture provisions', **31**, 351–367.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–654.

Cuchiero, C., Klein, I. & Teichmann, J. (2020), 'A fundamental theorem of asset pricing for continuous time large financial markets in a two filtration setting', *Theory of Probability & Its Applications* **65**(3), 388–404.

Cuchiero, C. & Teichmann, J. (2015), 'A convergence result for the Emery topology and a variant of the proof of the fundamental theorem of asset pricing', *Finance and Stochastics* **19**, 743–761.

Delbaen, F. & Schachermayer, W. (2006), *The Mathematics of Arbitrage*, Springer Verlag, Berlin.

Denis, L., Hu, M. & Peng, S. (2011), 'Function spaces and capacity related to a sublinear expectation: application to g-brownian motion paths', *Potential analysis* **34**, 139–161.

Filipović, D. (2009), *Term Structure Models: A Graduate Course*, Springer Verlag, Berlin.

Föllmer, H. & Schied, A. (2004), *Stochastic Finance*, 2nd edn, Walter de Gruyter, Berlin.

Föllmer, H. & Schied, A. (2016), *Stochastic Finance*, 4th edn, Walter de Gruyter, Berlin.

Fontana, C., Grbac, Z., Gümbel, S. & Schmidt, T. (2020), 'Term structure modelling for multiple curves with stochastic discontinuities', *Finance and Stochastics* **24**, 465–511.

Fontana, C. & Schmidt, T. (2018), 'General dynamic term structures under default risk', *Stochastic Processes and their Applications* **128**(10), 3353 – 3386.

Gehmlich, F. & Schmidt, T. (2018), 'Dynamic defaultable term structure modelling beyond the intensity paradigm', *Mathematical Finance* **28**(1), 211–239.

Heath, D., Jarrow, R. A. & Morton, A. J. (1992), 'Bond pricing and the term structure of interest rates', *Econometrica* **60**, 77–105.

Kabanov, Y. & Stricker, C. (2006), The Dalang–Morton–Willinger theorem under delayed and restricted information, *in* 'In Memoriam Paul-André Meyer', Springer, pp. 209–213.

Klein, I., Schmidt, T. & Teichmann, J. (2016), 'No arbitrage theory for bond markets', *Advanced Methods in Mathematical Finance* .

Majerek, D., Nowak, W. & Zieba, W. (2005), 'Conditional strong law of large number', *International Journal of Pure and Applied Mathematics* **20**(2), 143 – 156.

McNeil, A., Frey, R. & Embrechts, P. (2015), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press.

Merton, R. (1974), 'On the pricing of corporate debt: the risk structure of interest rates', **29**, 449–470.

Niemann, L. & Schmidt, T. (2024), 'A conditional version of the second fundamental theorem of asset pricing in discrete time', *Frontiers of Mathematical Finance* **3**(2), 239–269.

Zhou, C. (2001), 'The term structure of credit spreads with jump risk', *Journal of Banking and Finance* **25**, 2015–2040.