# 2

# Arbitrage Theory for Frictionless Markets

## 2.1 Models without Friction

### 2.1.1 DMW Theorem

The classical result by Dalang–Morton–Willinger, usually abbreviated as DMW and sometimes referred to as the Fundamental Theory of Asset (or Arbitrage) Pricing (FTAP) for the discrete finite-time model of a frictionless financial market, says:

> *There is no arbitrage if and only if there is an equivalent martingale measure.*

This formulation is due to Harrison and Pliska, who established it for a model with finite number of states of the nature, i.e., for finite $\Omega$. Retrospectively, one can insinuate that in this case it is mainly a "linguistic" exercise: the result expressed in geometric language was known a long time ago as the Stiemke lemma. This is, to large extent, true. However, a remarkable fact is that, contrarily to its predecessors, exactly this formulation of a no-arbitrage criterion, involving an important probability concept, a martingale measure, opens a way to numerous generalizations of great theoretical and practical value.

Loosely speaking, the result can be viewed as a partial converse to the assertion that one cannot win (in finite time) by betting on a martingale: if one cannot win betting on a process, the latter is a martingale with respect to an equivalent martingale measure.

We start our presentation here with a detailed analysis of the Dalang–Morton–Willinger theorem. The assertion in italics is, in fact, a *grand public* formulation which hides a profound difference between these two results, and the authors of advanced textbooks prefer to give a longer list of *NA* criteria. We follow this tradition.

The model is given by a complete probability space $(\Omega, \mathcal{F}, P)$ with a discrete-time filtration $\mathbf{F} = (\mathcal{F}_t)_{t=0,1,\dots,T}$ and an adapted $d$-dimensional price

Yu. Kabanov, M. Safarian, *Markets with Transaction Costs*,
Springer Finance, DOI 10.1007/978-3-540-68121-2_2,

process $S = (S_t)$ with the constant first component. It is convenient to assume that $\mathcal{F}_0$ is trivial and $\mathcal{F}_T = \mathcal{F}$.

The set of "results" $R_T$ (obtained from zero starting value) consists of the terminal values of discrete-time integrals

$$H \cdot S_T := \sum_{t=1}^{T} H_t \Delta S_t,$$

where $\Delta S_t := S_t - S_{t-1}$, and $H$ runs over the linear space $\mathcal{P}$ of predictable processes, i.e., $H_t \in L^0(\mathbf{R}^d, \mathcal{F}_{t-1})$ (the first component of $S$ plays no role here because $\Delta S_t^1 = 0$).

The common terminology: $H$ is a (portfolio) *strategy*, while $H \cdot S$ is called a *value process*. The larger set $A_T := R_T - L^0_+$ can be interpreted as the set of *hedgeable claims*; it is the set of random variables $H \cdot S_T - h$ where the r.v. $h \geq 0$.

By definition, the *NA property* of the model means that $R_T \cap L^0_+ = \{0\}$ (or, equivalently, $A_T \cap L^0_+ = \{0\}$). We prefer to use from the very beginning these mathematically convenient definitions in terms of intersections of certain sets rather than a popular form like this: the property $H \cdot S_T \geq 0$ implies that $H \cdot S_T = 0$.

**Theorem 2.1.1** *The following properties are equivalent:*

(a) $A_T \cap L^0_+ = \{0\}$ *(NA condition);*
(b) $A_T \cap L^0_+ = \{0\}$ *and* $A_T = \bar{A}_T$ *(closure in* $L^0$*);*
(c) $\bar{A}_T \cap L^0_+ = \{0\}$*;*
(d) *there is a strictly positive process* $\rho \in \mathcal{M}$ *such that* $\rho S \in \mathcal{M}$*;*
(e) *there is a bounded strictly positive process* $\rho \in \mathcal{M}$ *such tat* $\rho S \in \mathcal{M}$*.*

As usual, $\mathcal{M}$ is the space of martingales (if necessary, we shall also use a more complicated notation showing the probability, time range, etc.).

Of course, the last two properties are usually formulated as:

(d′) *there is a probability* $\tilde{P} \sim P$ *such that* $S \in \mathcal{M}(\tilde{P})$;
(e′) *there is a probability* $\tilde{P} \sim P$ *with* $d\tilde{P}/dP \in L^\infty$ *such that* $S \in \mathcal{M}(\tilde{P})$.

However, the chosen versions have more direct analogs in the model with transaction costs. Their equivalences are obvious due to the following elementary fact about martingales with respect to a probability measure $\tilde{P} \ll P$ with density $\rho_T$:

$$S \in \mathcal{M}(\tilde{P}) \text{ if and only if } \rho S \in \mathcal{M}(P) \text{ where } \rho_t = E(\rho_T|\mathcal{F}_t),\ t \leq T.$$

Collecting conditions in the single theorem is useful because one can clearly see that in numerous generalizations and ramifications certain properties remain equivalent (of course, appropriately modified), but others do not. Note also that, in the case of finite $\Omega$, the set $A_T$ is always closed. Indeed, it is the arithmetic sum of a linear space and the polyhedral cone $-L^0_+$ in the finite-dimensional linear space $L^0$. Thus, it is a polyhedral cone. So, we have no difference between the first three properties, while the last two coincide trivially. The situation is completely different for arbitrary $\Omega$. Though the linear space $R_T$ is always closed (we show this later), the set $A_T$ may be not closed even for $T = 1$ and a countable $\Omega$. To see this, let $\mathcal{F}_0$ be trivial and take $\mathcal{F}_1 = \sigma\{\xi\}$ and $\Delta S_1 = \xi$, where $\xi$ is a strictly positive finite random variable such that $P(\xi > \varepsilon) > 0$ whatever is $\varepsilon > 0$. The set $A_1 = \mathbf{R}\xi - L^0_+$ does not contain any strictly positive constant, but each constant $c > 0$ belongs to its closure $\bar{A}_1$ because $(n\xi) \wedge c \to c$ as $n \to \infty$.

To the already long list, one can add several other equivalent conditions:

(f) *there is a strictly positive process* $\rho \in \mathcal{M}$ *such that* $\rho S \in \mathcal{M}_{\text{loc}}$;
(f′) *there is a probability* $\tilde{P} \sim P$ *such that* $S \in \mathcal{M}_{\text{loc}}(\tilde{P})$;
(g) $\{\eta \Delta S_t : \ \eta \in L^0(\mathcal{F}_{t-1})\} \cap L^0_+ = \{0\}$ *for all* $t \le T$ *(NA for one-step models).*

With other conditions already established, the above addendum poses no problems. Indeed, (f′) is obviously implied by (e′). On the other hand, if $S \in \mathcal{M}_{\text{loc}}(\tilde{P})$, then $\tilde{H} \cdot S \in \mathcal{M}(\tilde{P})$ with $\tilde{H}_t := 1/(1 + \tilde{E}(|\Delta S_t| \| \mathcal{F}_{t-1}))$. So, we know that *NA* holds for the process $\tilde{H} \cdot S$; hence, it holds also for $S$ as both processes have the same set of hedgeable claims, i.e., (f′) implies a property from the "main" list of equivalent conditions. Suppose now that the implication (g) $\Rightarrow$ (a) fails. Take the smallest $t \le T$ such that $A_t \cap L^0_+ \ne \{0\}$ (the set of such dates is nonempty: it contains, at least, $T$). We have a strategy $H = (H_s)_{s \le T}$ such that $H \cdot S_t \ge 0$ and $P(H \cdot S_t > 0) > 0$. Due to the choice of $t$, either the set $\Gamma' := \{H \cdot S_{t-1} < 0\}$ is of strictly positive probability (and (g) is violated by $\eta := I_{\Gamma'} H_t$), or the set $\Gamma'' := \{H \cdot S_{t-1} = 0\}$ is of full measure (and (g) is violated by $\eta := I_{\Gamma''} H_t$), a contradiction.

**Remark.** The *NA* property for the class of all strategies, as defined above, is equivalent to the *NA* property in the narrower class of bounded strategies $H$. Indeed, if there is an arbitrage opportunity, then, in virtue of the condition (g), there is an arbitrage opportunity $\eta$ for a certain one-step model. Clearly, when $n$ is sufficiently large, $\eta I_{\{|\eta| \le n\}}$ will be an arbitrage opportunity for this one-step model. Note that the presence of (g) in the list of equivalent conditions is crucial in this reasoning.

Similarly, *NA* is equivalent to the absence of arbitrage in the class of so-called *admissible* strategies for which the value processes are bounded from below by constants (depending on the strategy). Moreover, if $H$ is an arbitrage opportunity generating the value process $V = H \cdot S$, one can find another arbitrage opportunity $\tilde{H}$ such that the value process $\tilde{H} \cdot S \ge 0$. To see this, we consider the sets $\Gamma_t := \{H \cdot S_t < 0\}$ and the last instant $r$ for which the probability of such a set is strictly positive; $0 < r < T$ since $H$ is an arbitrage opportunity. Let us check that the strategy $\tilde{H} := I_{\Gamma_r} I_{]r,T]} H$ has the claimed property. Indeed, the process $\tilde{V} := \tilde{H} \cdot S$ is zero for all $t \le r$ and remains zero outside the set $\Gamma_r$ until $T$. On the set $\Gamma_r$, the increments $\Delta \tilde{V}_t = \Delta V_t$ for

$t \geq r+1$, and hence the trajectories of $\tilde{V}$ are the trajectories of $V$ shifted upwards on the value $-V_r > 0$.

Before the proof of Theorem 2.1.1, we give in the following subsection several elementary results which will be also useful in obtaining NA criteria in models with transaction costs.

### 2.1.2 Auxiliary Results: Measurable Subsequences and the Kreps–Yan Theorem

**Lemma 2.1.2** *Let $\eta^n \in L^0(\mathbf{R}^d)$ be such that $\underline{\eta} := \liminf |\eta^n| < \infty$. Then there are $\tilde{\eta}^k \in L^0(\mathbf{R}^d)$ such that for all $\omega$, the sequence of $\tilde{\eta}^k(\omega)$ is a convergent subsequence of the sequence of $\eta^n(\omega)$.*

*Proof.* Define the random variables $\tau_k := \inf\{n > \tau_{k-1} : \ ||\eta^n| - \underline{\eta}| \leq k^{-1}\}$ starting with $\tau_0 := 0$. Then $\tilde{\eta}_0^k := \eta^{\tau_k}$ is in $L^0(\mathbf{R}^d)$, and $\sup_k |\tilde{\eta}_0^k| < \infty$. Working further with the sequence of $\tilde{\eta}_0^n$, we construct, applying the above procedure to the first component and its lim inf, a sequence of $\tilde{\eta}_1^k$ with convergent first component and such that for all $\omega$, the sequence of $\tilde{\eta}_1^k(\omega)$ is a subsequence of the sequence of $\tilde{\eta}_0^n(\omega)$. Passing on each step to the newly created sequence of random variables and to the next component, we arrive at a sequence with the desired properties. □

**Remark.** The claim can be formulated as follows: there exists a (strictly) increasing sequence of integer-valued random variables $\sigma_k$ such that $\eta^{\sigma_k}$ converges a.s.

**Lemma 2.1.3** *Let $\mathcal{G} = \{\Gamma_\alpha\}$ be a family of measurable sets such that any nonnull set $\Gamma$ has a nonnull intersection with an element of $\mathcal{G}$. Then there is an at most countable subfamily of sets $\{\Gamma_{\alpha_i}\}$ the union of which is of full measure.*

*Proof.* Suppose that $\mathcal{G}$ is closed under countable unions. Then $\sup_\alpha P(\Gamma_\alpha)$ is attained on some $\tilde{\Gamma} \in \mathcal{G}$. The subfamily consisting of a single $\tilde{\Gamma}$ gives the answer. Indeed, $P(\tilde{\Gamma}) = 1$: otherwise we could enlarge the supremum by adding a set from $\mathcal{G}$ having a nonnull intersection with $\tilde{\Gamma}^c$. The general case follows by considering the family formed by countable unions of sets from $\mathcal{G}$. □

The following result is referred to as the Kreps–Yan theorem. It holds for arbitrary $p \in [1, \infty]$, $p^{-1} + q^{-1} = 1$, but the cases $p = 1$ and $p = \infty$ are the most important. Recall that for $p \neq \infty$, the norm closure of a convex set in $L^p$ coincides with the closure in $\sigma\{L^p, L^q\}$.

**Theorem 2.1.4** *Let $\mathcal{C}$ be a convex cone in $L^p$ closed in $\sigma\{L^p, L^q\}$, containing $-L^p_+$ and such that $\mathcal{C} \cap L^p_+ = \{0\}$. Then there is $\tilde{P} \sim P$ with $d\tilde{P}/dP \in L^q$ such that $\tilde{E}\xi \leq 0$ for all $\xi \in \mathcal{C}$.*

*Proof.* By the Hahn–Banach theorem any nonzero $x \in L^p_+ := L^p(\mathbf{R}_+, \mathcal{F})$ can be separated from $\mathcal{C}$: there is a $z_x \in L^q$ such that $Ez_x x > 0$ and $Ez_x \xi \le 0$ for all $\xi \in \mathcal{C}$. Since $\mathcal{C} \supseteq -L^p_+$, the latter property yields that $z_x \ge 0$; we may assume that $\|z_x\|_q = 1$. Let us consider the family $\mathcal{G} := \{z_x > 0\}$. As any nonnull set $\Gamma$ has a nonnull intersection with the set $\{z_x > 0\}$, $x = I_\Gamma$, the family $\mathcal{G}$ contains a countable subfamily of sets (say, corresponding to a sequence $\{x_i\}$) the union of which is of full measure. Thus, $z := \sum 2^{-i} z_{x_i} > 0$, and we can take $\tilde{P} := zP$. □

### 2.1.3 Proof of the DMW Theorem

The implications (b) ⇒ (a), (b) ⇒ (c), and (e) ⇒ (d) are trivial. The implication (d) ⇒ (a) is easy. Indeed, let $\xi \in A_T \cap L^0_+$, i.e., $0 \le \xi \le H \cdot S_T$. Since the conditional expectation with respect to the martingale measure $\tilde{E}(H_t \Delta S_t | \mathcal{F}_{t-1}) = 0$, we obtain by consecutive conditioning that $\tilde{E} H \cdot S_T = 0$. Thus, $\xi = 0$. To complete the proof, it remains to verify that (c) ⇒ (e) and (a) ⇒ (b).

(c) ⇒ (e). Notice that for any random variable $\eta$, there is an equivalent probability $P'$ with bounded density such that $\eta \in L^1(P')$ (e.g., one can take $P' = Ce^{-|\eta|}P$). Property (c) (as well as (a) and (b)) is invariant under equivalent change of probability. This consideration allows us to assume that all $S_t$ are integrable. The convex set $A^1_T := \bar{A}_T \cap L^1$ is closed in $L^1$. Since $A^1_T \cap L^1_+ = \{0\}$, Theorem 2.1.4 ensures the existence of $\tilde{P} \sim P$ with bounded density and such that $\tilde{E}\xi \le 0$ for all $\xi \in A^1_T$, in particular, for $\xi = \pm H_t \Delta S_t$ with bounded and $\mathcal{F}_{t-1}$-measurable $H_t$. Thus, $\tilde{E}(\Delta S_t | \mathcal{F}_{t-1}) = 0$.

(a) ⇒ (b). Lemma 2.1.2 allows us to establish the closedness of $A_T$ by simple recursive arguments even without assuming that the $\sigma$-algebra $\mathcal{F}_0$ is trivial (of course, this does not add any generality but helps to start the induction in the time variable).

Let us consider the case $T = 1$. Let $H^n_1 \Delta S_1 - r^n \to \zeta$ a.s., where $H^n_1$ is $\mathcal{F}_0$-measurable, and $r^n \in L^0_+$. The closedness of $A_1$ means that $\zeta = H_1 \Delta S_1 - r$ for some $\mathcal{F}_0$-measurable $H_1$ and $r \in L^0_+$. To show this, we represent each $H^n_1$ as a column vector and write the whole sequence of these column vectors as the infinite matrix

$$\mathbf{H}_1 := \begin{bmatrix} H^{11}_1 & H^{21}_1 & \dots & \dots & H^{n1}_1 & \dots \\ H^{12}_1 & H^{22}_1 & \dots & \dots & H^{n2}_1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ H^{1d}_1 & H^{2d}_1 & \dots & \dots & H^{nd}_1 & \dots \end{bmatrix}.$$

If the matrix is zero, there is nothing to prove. Suppose that the assertion holds when this (random) matrix has, for each $\omega$, at least $m$ zero lines. We show that the claim holds true also when $\mathbf{H}_1$ has at least $m-1$ zero lines.

It is sufficient to find $\mathcal{F}_0$-measurable random variables $\tilde{H}^k_1$ convergent a.s. and $\tilde{r}^k \in L^0_+$ such that $\tilde{H}^k_1 \Delta S_1 - \tilde{r}^k \to \zeta$ a.s.

Let $\Omega_i \in \mathcal{F}_0$ form a finite partition of $\Omega$. An important (though obvious) observation: we may argue on each $\Omega_i$ separately as on an autonomous measure space (considering the restrictions of random variables and traces of $\sigma$-algebras).

Let $\underline{H}_1 := \liminf |H_1^n|$. On $\Omega_1 := \{\underline{H}_1 < \infty\}$, we take, using Lemma 2.1.2, $\mathcal{F}_0$-measurable $\tilde{H}_1^k$ such that $\tilde{H}_1^k(\omega)$ is a convergent subsequence of $H_1^n(\omega)$ for every $\omega$; $\tilde{r}^k$ are defined correspondingly. Thus, if $\Omega_1$ is of full measure, the goal is achieved.

On $\Omega_2 := \{\underline{H}_1 = \infty\}$, we put $G_1^n := H_1^n/|H_1^n|$ and $h_1^n := r_1^n/|H_1^n|$. Clearly, $G_1^n \Delta S_1 - h_1^n \to 0$ a.s. By Lemma 2.1.2 we find $\mathcal{F}_0$-measurable $\tilde{G}_1^k$ such that $\tilde{G}_1^k(\omega)$ is a convergent subsequence of $G_1^n(\omega)$ for every $\omega$. Denoting the limit by $\tilde{G}_1$, we obtain that $\tilde{G}_1 \Delta S_1 = \tilde{h}_1$ where $\tilde{h}_1$ is nonnegative; hence, in virtue of (a), $\tilde{G}_1 \Delta S_1 = 0$.

As $\tilde{G}_1(\omega) \neq 0$, there exists a partition of $\Omega_2$ into $d$ disjoint subsets $\Omega_2^i \in \mathcal{F}_0$ such that $\tilde{G}_1^i \neq 0$ on $\Omega_2^i$. Define $\bar{H}_1^n := H_1^n - \beta^n \tilde{G}_1$, where $\beta^n := H_1^{ni}/\tilde{G}_1^i$ on $\Omega_2^i$. Then $\bar{H}_1^n \Delta S_1 = H_1^n \Delta S_1$ on $\Omega_2$. The matrix $\bar{\mathbf{H}}_1$ has, for each $\omega \in \Omega_2$, at least $m$ zero lines: our operations did not affect the zero lines of $\mathbf{H}_1$, and a new one has appeared, namely, the $i$th one on $\Omega_2^i$. We conclude by the induction hypothesis.

To establish the induction step in the time variable, we suppose that the claim is true for $(T-1)$-step models. Let $\sum_{t=1}^T H_t^n \Delta S_t - r^n \to \zeta$ a.s., where $H_t^n$ are $\mathcal{F}_{t-1}$-measurable, and $r^n \in L_+^0$. As at the first step, we work with the matrix $\mathbf{H}_1$ using exactly the same reasoning.

On $\Omega_1$ we take an increasing sequence of $\mathcal{F}_0$-random variables $\tau_k$ such that $H^k := H_1^{\tau_k}$ converges to $H_1$. Thus, $\sum_{t=2}^T H_t^{\tau_k} \Delta S_t - r^{\tau_k}$ converges as $k \to \infty$, and we have a reduction to a $(T-1)$-step model.

On $\Omega_2$ we use again the same induction in $m$, the number of zero lines of $\mathbf{H}_1$. The only modification is that the identical operations (passage to subsequences, normalization by $H_1^n$, etc.) should be performed simultaneously over all other matrices $\mathbf{H}_2, \dots, \mathbf{H}_T$.

**Remark 1.** Exactly the same arguments as those used in the proof of the implication (a) $\Rightarrow$ (b) lead to the following assertion referred to as the Stricker lemma:

*The set of results $R_T$ is closed.*

This property holds irrelevantly of the *NA*-condition. Indeed, the latter was used only to check that the nonnegative limit $\tilde{h}_1$ is, in fact, equal to zero. But this holds automatically if we start the arguments with $r_n = 0$.

**Remark 2.** The DMW theorem contains as a corollary the assertion that, in the discrete-time setting with finite horizon, any local martingale is a martingale with respect to a measure $\tilde{P} \sim P$ with bounded density. Moreover, this measure can be chosen in such a way that a given random variable $\xi$ will be $\tilde{P}$-integrable. At the end of this chapter we show that, even in the model with infinite horizon, the local martingale is a martingale with respect to an equivalent probability measure.

### 2.1.4 Fast Proof of the DMW Theorem

Our detailed formulation of the DMW theorem, together with its proof, is intended to prepare the reader to the arguments developed for models with transaction costs. However, a short and elementary proof of the "main" equivalence (a) $\Leftrightarrow$ (e), a proof which can be used in introductory courses for mathematical students, is of separate interest. We give one here combining an optimization approach due to Chris Rogers with Lemma 2.1.2 on measurable subsequences. It is based on the one-step result the first condition of which is just an alternative reformulation of the *NA*-property.

**Proposition 2.1.5** *Let $\xi \in L^0(\mathbf{R}^d)$, and let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$. Then the following conditions are equivalent:*

(i) *for any $\alpha \in L^0(\mathbf{R}^d, \mathcal{G})$, the inequality $\alpha\xi \geq 0$ holds as the equality;*
(ii) *there exists a bounded random variable $\varrho > 0$ such that $E\varrho|\xi| < \infty$ and $E(\varrho\xi|\mathcal{G}) = 0$.*

*Proof.* One needs arguments only for the "difficult" implication (i) $\Rightarrow$ (ii).

First, examine the case where $\mathcal{G}$ is trivial. Let us consider the function $f(a) = Ee^{a\xi - |\xi|^2}$, $a \in \mathbf{R}^d$. If it attains its minimum at some point $a_*$, the problem is solved with $\rho = e^{a_*\xi - |\xi|^2}$, since at this point the derivative of $f$ is zero: $E\xi e^{a_*\xi - |\xi|^2} = 0$. One can check that condition (i) excludes the possibility that the minimum is not attained—we do a verification below.

Let us turn to the general case. A dimension reduction argument allows us to work assuming that the relation $\alpha\xi = 0$ with $\alpha \in L^0(\mathbf{R}^d, \mathcal{G})$ holds only if $\alpha = 0$ (when $\mathcal{G}$ is trivial, this is just the linear independence of the components of $\xi$ as elements of $L^0$). Let $Q(\omega, dx)$ be the regular conditional distribution of $\xi$ with respect to $\mathcal{G}$. Define the function

$$f(\omega, a) := \int e^{ax - |x|^2} Q(\omega, dx)$$

continuous in $a$ and $\mathcal{G}$-measurable in $\omega$. Introduce the $\mathcal{G}$-measurable random variable $f_*(\omega) = \inf_a f(\omega, a)$ and consider, in the product space $\Omega \times \mathbf{R}^d$, the sets $\{(\omega, a) : f(\omega, a) < f_*(\omega) + 1/n\}$ with nonempty open $\omega$-sections $\Gamma_n(\omega)$. Let $\alpha_n$ be a $\mathcal{G}$-measurable random variable with $\alpha_n(\omega) \in \Gamma_n(\omega)$. Such $\alpha_n$ can be constructed easily, without appealing to a measurable selection theorem, e.g., one can take $\alpha_n(\omega) := q_{\theta(n)}$, where

$$\theta(n) := \min\{k : \ f(\omega, q_k) < f_*(\omega) + 1/n\}$$

with an arbitrary countable dense subset $\{q_n\}$ in $\mathbf{R}^d$. Let us consider the set $\Omega_0 := \{\liminf |\alpha_n| < \infty\}$ with its complement $\Omega_1$. Using Lemma 2.1.2, we may assume that on $\Omega_1$ the sequence $\tilde{\alpha}_n := \alpha_n/|\alpha_n|$ converges to some $\beta$ with $|\beta| = 1$ and, by the Fatou lemma,

$$\int e^{\lim |\alpha_n(\omega)|\beta(\omega)x-|x|^2} I_{\{\beta(\omega)x\neq 0\}} Q(\omega,dx)$$
$$\leq \liminf \int e^{\alpha_n(\omega)x-|x|^2} I_{\{\beta(\omega)x\neq 0\}} Q(\omega,dx) \leq f_*(\omega).$$

Necessarily, $Q(\omega,\{x:\ \beta(\omega)x>0\})=0$, implying that $\beta\xi \leq 0$ (a.s.), and, therefore, in virtue of $(i)$, we have that $\beta\xi = 0$. Due to our provision, this equality holds only if $\beta = 0$, and, hence, $\Omega_1$ is a null set which does not matter. Again by Lemma 2.1.2 we may assume that on the set $\Omega_0$ of full measure the sequence $\alpha_n(\omega)$ converges to some $\alpha_*(\omega)$. Clearly, $f(\omega,a)$ attains its minimum at $\alpha_*(\omega)$, and we conclude with $\varrho := e^{\alpha_*\xi-|\xi|^2}/c(\alpha_*)$, where the function $c(a) := \sup_x (1+|x|)e^{ax-|x|^2}$. □

The "difficult" implication (a) ⇒ (e) follows from the above proposition by backward induction. We claim that for each $t = 0,1,\dots,T-1$, there is a bounded random variable $\rho_t^T > 0$ such that $E\rho_t^T|\Delta S_u| < \infty$ and $E\rho_t^T \Delta S_u = 0$ for $u = t+1,\dots,T$. Since (a) implies the $NA$-property for each one-step model, the existence of $\rho_{T-1}^T$ follows from the above proposition with $\xi = \Delta S_T$ and $\mathcal{G} = \mathcal{F}_{T-1}$. Suppose that we have already found $\rho_t^T$. Putting $\xi = E(\rho_t^T|\mathcal{F}_{t-1})\Delta S_{t-1}$ and $\mathcal{G} = \mathcal{F}_{t-2}$, we find bounded $\mathcal{F}_{t-1}$-measurable $\varrho_{t-1} > 0$ such that $E(\varrho_{t-1}E(\rho_t^T|\mathcal{F}_{t-1})|\Delta S_{t-1}|) < \infty$ and $E(\varrho_{t-1}E(\rho_t^T|\mathcal{F}_{t-1})\Delta S_{t-1}) = 0$. It is clear that $\rho_{t-1}^T$ meets the requirements. Property (e) of the DMW theorem holds with $\rho_t := E(\rho_0^T|\mathcal{F}_t)$.

### 2.1.5 *NA* and Conditional Distributions of Price Increments

As shown by Jacod and Shiryaev, the long list of conditions equivalent to the $NA$-property can be completed by the following one involving the regular conditional distributions $Q_t(\omega,dx)$ of the price increments $\Delta S_t$ knowing $\mathcal{F}_{t-1}$:

(h) $0 \in \operatorname{ri}\operatorname{conv}\operatorname{supp} Q_t(\omega,dx)$ *a.s. for all* $t = 1,\dots,T$.

Recall that $Q_t(\omega,\Gamma)$ is an $\mathcal{F}_{t-1}$-measurable random variable in $\omega$ and a measure in $\Gamma$ such that $P(\Delta S_t \in \Gamma|\mathcal{F}_{t-1}) = Q_t(\omega,\Gamma)$ (a.s.) for each Borel set $\Gamma$ in $\mathbf{R}^d$. The topological support of the measure $Q_t(\omega,dx)$ is the intersection of all closed sets the complements of which are null sets for this measure. The abbreviation "ri" denotes the relative interior of a convex set, i.e., the interior in the relative topology of the smallest affine subspace containing it.

Comparing (h) and (g), we see that their equivalence follows from the next one-step result complementing Proposition 2.1.5.

**Proposition 2.1.6** *Let $\xi \in L^0(\mathbf{R}^d)$, and let $Q(\omega, dx)$ be a regular conditional distribution of $\xi$ with respect to a $\sigma$-algebra $\mathcal{G} \subseteq \mathcal{F}$. Then the NA-property (or the equivalent property* (ii)*) holds if and only if the following condition is satisfied:*

(iii) $0 \in \operatorname{ri}\operatorname{conv}\operatorname{supp} Q(\omega, dx)$ *a.s.*

*Proof.* (ii) $\Rightarrow$ (iii). Consider the case where $\mathcal{G}$ is trivial. If the origin does not belong to $A := \operatorname{ri}\operatorname{conv}\operatorname{supp} Q(dx)$, then there exists $a \in \mathbf{R}^d$ such that the latter set lies in the closed half-space $\{x : \ ax \geq 0\}$ but not in the subspace $\{x : \ ax = 0\}$ (to see this, apply the separation theorem in the linear subspace of minimal dimension containing $A$ and extend the separating functional to a functional on the whole $\mathbf{R}^d$ vanishing on the orthogonal complement). So, $Q(x : \ ax > 0) > 0$, and for any strictly positive bounded random variable $\rho$ measurable with respect to $\sigma\{\xi\}$ (i.e., of the form $\rho = r(\xi)$ with a Borel function $r$), we have

$$E\rho a\xi = \int r(x) ax I_{\{x:\ ax>0\}} Q(dx) > 0,$$

in contradiction with (ii).

In the general case we consider the set $\Gamma := \{(\omega, a) : \ Q(\omega, \{ax > 0\}) > 0\}$, which is measurable with respect to the product $\sigma$-algebra $\mathcal{G} \otimes \mathcal{B}^d$; let $\Gamma(\omega)$ be its $\omega$-sections. If (iii) fails, then, as it was just shown, the projection $\operatorname{Pr}_\Omega \Gamma$ of $\Gamma$ on $\Omega$ is nonnull. Due to the measurable selection theorem, there exists a $\mathcal{G}$-measurable $\mathbf{R}^d$-valued random variable $\alpha$ such that $\alpha(\omega) \in \Gamma(\omega)$ for almost all $\omega$ from $\operatorname{Pr}_\Omega \Gamma$. Now, take an arbitrary bounded strictly positive function $r(\omega, x)$ measurable with respect to $\sigma\{\mathcal{G}, \xi\} \otimes \mathcal{B}^d$ and put $\rho(\omega) := r(\omega, \xi(\omega))$. Then

$$E(r\alpha\xi|\mathcal{G}) = \int r(\omega, x)\alpha(\omega) x I_{\{x:\ \alpha(\omega)x>0\}} Q(\omega, dx) > 0 \quad \text{on } \operatorname{Pr}_\Omega \Gamma.$$

It is easy to see that this is a contradiction with (ii).

(iii) $\Rightarrow$ (ii). Again, let us first consider the case of trivial $\mathcal{G}$. Let $L$ be the affine subspace of minimal dimension containing the set $A := \operatorname{ri}\operatorname{conv}\operatorname{supp} Q(dx)$.

The assumption $0 \in A$ implies that the function

$$f(a) := \int e^{ax-|x|^2} Q(dx) = \int e^{ax-|x|^2} I_L(x) Q(dx)$$

attains its minimum at some point $a_*$: otherwise, we could find, as in the proof of Proposition 2.1.5, a vector $\beta$ such $\{x : \ \beta x > 0\} \cap L$ is a $Q$-null set. But this means that the origin is not in the relative interior of the convex hull of $\operatorname{supp} Q(dx)$. In the general case, we can find a $\mathcal{G}$-measurable random variable $\alpha_*$ such that $\alpha_*(\omega)$ is a minimizer of $f(\omega, a)$ and conclude in the same way as in Proposition 2.1.5. □

### 2.1.6 Comment on Absolute Continuous Martingale Measures

One may ask whether the existence of an absolute continuous martingale measure can be related with a certain no-arbitrage property. Indeed, in the case of finite number of states of nature, we have the following criterion:

**Proposition 2.1.7** *Suppose that $\Omega$ is finite. Then the following conditions are equivalent:*

(a) $R_T \cap L^0(\mathbf{R}_+ \setminus \{0\}) = \emptyset$;
(b) *there is a probability measure $\tilde{P} \ll P$ such that $S \in \mathcal{M}(\tilde{P})$.*

Here the implication (b) ⇒ (a) is obvious, while the converse follows easily from the finite-dimensional separation theorem applied to the disjoint convex sets $A_T \setminus \{0\}$ and $L^0(\mathbf{R}_+ \setminus \{0\})$: any separating functional after normalization is a density of probability measure with the needed property. The condition (a) means that there is no "universal" arbitrage strategy $H$, that is, such that $H \cdot S_T > 0$ (a.s.).

Unfortunately, the above proposition cannot be extended to the case of arbitrary $\Omega$.

**Example.** Let us consider a one-period model with two risky assets whose price increments $\Delta S_1^1$ and $\Delta S_1^2$ are random variables defined on a countable probability space $\Omega = \{\omega_i\}_{i\geq 0}$ with all $P(\{\omega_i\}) > 0$. The initial $\sigma$-algebra is trivial. Let $\Delta S_1^1(\omega_0) = 1$, $\Delta S_1^1(\omega_i) = -i$, $i \geq 1$. Let $\Delta S_1^2(\omega_0) = 0$, $\Delta S_1^2(\omega_i) = 1$, $i \geq 1$. Apparently, the equalities $E_Q \Delta S_1^1 = 0$ and $E_Q \Delta S_1^2 = 0$ are incompatible, and, hence, there are no martingale measures. On the other hand, let $(H^1, H^2) \in \mathbf{R}^2$ be a "universal" arbitrage strategy. Then necessarily $H^1 > 0$, and we get a contradiction since in such a case the countable system of inequalities $-iH^1 + H^2 > 0$, $i \geq 1$, is incompatible whatever is $H^2$.

### 2.1.7 Complete Markets and Replicable Contingent Claims

As we observed, the set of results $R_T$ is always closed in $L^0$. It is an easy exercise to deduce from this property that the set $\mathbf{R} + R_T$ is also closed. We use this remark in the proof of the following:

**Proposition 2.1.8** *Suppose that the set $\mathcal{Q}^e$ of equivalent martingale measures is nonempty. Then the following conditions are equivalent:*

(a) *$\mathcal{Q}^e$ is a singleton;*
(b) $\mathbf{R} + R_T = L^0$.

*Proof.* (a) ⇒ (b). We may assume without loss of generality that $P$ is a martingale measure. Suppose that there is $\xi \in L^0$ which is not in the closed subspace $\mathbf{R} + R_T \subseteq L^0$. It follows that the random variables $\xi^n := \xi I_{\{|\xi| \leq n\}}$ are not in this subspace for all $n \geq N$. Applying the separation theorem, one can find $\eta$ with $|\eta| \leq 1/2$ such that $E\eta\zeta = 0$ for the elements $\zeta$ from the closed subspace $(\mathbf{R}+R_T)\cap L^1$ of $L^1$ but $E\eta\xi^N>0$. Put $Q=(1+\eta)P$. Then $E_Q H\cdot S_T=0$ whatever is a bounded predictable process $H$. This means that $Q$ is an equivalent martingale measure different from $P$, contradicting (a).

(b) $\Rightarrow$ (a). Take $\Gamma\in\mathcal{F}_T$. Then $I_\Gamma=c_\Gamma+H_\Gamma\cdot S_T$, where $c_\Gamma$ is a constant. It follows that $Q(\Gamma)=c_\Gamma$ whatever is a martingale measure $Q$, i.e., the latter is unique.

The property (b), in financial literature referred to as the *market completeness*, means that any contingent claim can be replicated, that is, represented as the terminal value of a self-financing portfolio starting from a certain initial endowment. The above statement, asserting that an arbitrage-free market is complete if and only if there is only one equivalent martingale measure, sometimes is called the second fundamental theorem of asset pricing.

The closedness of the subspace $\mathbf{R}+R_T$ leads the next assertion concerning replicable claims on incomplete markets. In its formulation, $Q_l$ and $Q_l^e$ denote the sets of absolutely continuous and equivalent local martingale measures. □

**Proposition 2.1.9** *Suppose that $\mathcal{Q}^e\neq\emptyset$. Let a random variable $\xi\geq 0$ be such that $a=\sup_{Q\in\mathcal{Q}_l^e}E_Q\xi<\infty$ and the supremum is attained on some measure $Q^*$. Then $\xi=a+H\cdot S_T$ for some predictable process (and, hence, the function $Q\mapsto E_Q\xi$ is constant on the set $Q_l^e$).*

*Proof.* Supposing that the statement fails, we apply the Hahn–Banach theorem and separate $\xi$ and the subspace $(\mathbf{R}+R_T)\cap L^1(Q^*)$ in $L^1(Q^*)$, that is, we find $\eta\in L^\infty$ such that $E_{Q^*}\xi\eta>0$ and $E_{Q^*}\eta\zeta=0$ for all $\zeta$ from the subspace. In particular, $E_{Q^*}\eta=0$ and $E_{Q^*}H\cdot S_T\eta=0$ whatever is a predictable process $H$ such that $H\cdot S_T$ is integrable; in particular, the last equality holds for $H=I_{[0,\tau_n]}$, where $\tau_n$ is a localizing sequence for $S$. Normalizing, we may assume that $|\eta|\leq 1/2$. It follows that the measure $\tilde{Q}=(1+\eta)Q^*$ is an element of $\mathcal{Q}_l^e$ and $E_{\tilde{Q}}\xi=a+E_{Q^*}\xi\eta>a$ in an apparent contradiction with the definition of $Q^*$. □

### 2.1.8 DMW Theorem with Restricted Information

Let us consider the following setting, which is only slightly different from the classical one. Namely, assume that we are given a filtration $\mathbf{G}=(\mathcal{G}_t)_{t\leq T}$ with $\mathcal{G}_t\subseteq\mathcal{F}_t$. Suppose that the strategies are now predictable with respect to this smaller filtration (i.e., $H_t\in L^0(\mathcal{G}_{t-1})$), a situation which may happen when the portfolios are revised on the basis of restricted information, e.g., due to a delay. Again, we may define the sets $R_T$ and $A_T$ and give a definition of the arbitrage, which, in these symbols, looks exactly as (a) above, and we can list the corresponding necessary and sufficient conditions.

To this aim, we define the $\mathbf{G}$-optional projection $X^o$ of an integrable process $X$ by putting $X_t^o:=E(X_t|\mathcal{G}_t)$, $t\leq T$.

**Theorem 2.1.10** *The following properties are equivalent:*

(a) $A_T \cap L^0_+ = \{0\}$ *(NA condition);*
(b) $A_T \cap L^0_+ = \{0\}$ *and* $A_T = \bar{A}_T$*;*
(c) $\bar{A}_T \cap L^0_+ = \{0\}$*;*
(d) *there is a strictly positive process* $\rho \in \mathcal{M}$ *with* $(\rho S)^o \in \mathcal{M}(\mathbf{G})$*;*
(e) *there is a bounded strictly positive process* $\rho \in \mathcal{M}$ *with* $(\rho S)^o \in \mathcal{M}(\mathbf{G})$.

The symbol $\mathcal{M}(\mathbf{G})$ stands here for the set of $\mathbf{G}$-martingales, and we presume tacitly in the last two conditions that $E\rho_t|S_t| < \infty$. Clearly, these conditions can be formulated in terms of existence of an equivalent probability $\tilde{P}$ such that $\tilde{E}(S_{t+1}|\mathcal{G}_t) = \tilde{E}(S_t|\mathcal{G}_t)$ for all $t \le T-1$.

We leave to the reader as an (easy) exercise to inspect that the arguments of the previous section go well for this theorem.

**Remark.** Curiously, this result, rather natural and important for practical applications, was established only recently. It happens that all numerous proofs, except one suggested in [131] and reproduced above in Sect. 2.1.3, in their most essential part concerning the construction of equivalent martingale measures given the *NA*-property, are based on the reduction to the one-step case with $T = 1$. Of course, (a) implies (g) (i.e., the *NA*-property for all one-step models). A clever argument in the Dalang–Morton–Willinger paper permits to assemble a required martingale density from martingale densities for one-step models. However, in the model with restricted information, the property (g) drops out from the list of equivalent conditions.

**Example.** Consider the model where $T = 2$, $\mathcal{G}_0 = \mathcal{G}_1 = \{\emptyset, \Omega\}$, but there is $A \in \mathcal{F}_2$ such that $0 < P(A) < 1$. Put

$$\Delta S_1 := I_A - \frac{1}{2} I_{A^c}, \qquad \Delta S_2 := -\frac{1}{2} I_A + I_{A^c}.$$

There is no arbitrage at each of two steps, but the constant process with $H_1 = H_2 = 1$ is an arbitrage strategy for the two-step model.

### 2.1.9 Hedging Theorem for European-Type Options

One of the most fundamental though simple ideas of mathematical finance is the arbitrage pricing of contingent claims.

A *contingent claim* or an *option* is a random variable $\xi$ which can be interpreted as a pay-off of the option seller to the option buyer. For a European-type option, the payment is made at the terminal (maturity) date $T$ and may depend on the whole history up to $T$. What is a "fair" price for such a contract payed at time zero? Apparently, and this is the basic principle, the option price should be such that neither of two parties has arbitrage opportunities, i.e., riskless profits.

Let us define the set

$$\Gamma := \Gamma(\xi) := \{x : \exists H \in \mathcal{P} \text{ such that } x + H \cdot S_T \geq \xi\}.$$

Clearly, if not empty, it is a semi-infinite interval (maybe, coinciding with the whole line). A priori, it can be either of the form $[\bar{x}, \infty[$ or $]\bar{x}, \infty[$.

The theorem below ensures, in particular, that $\bar{x} \in \Gamma$. If the contracted price of the option, say, $x$ is strictly larger than $\bar{x}$, then the seller has a nonrisky profit by pocketing $x - \bar{x}$ and running a self-financing portfolio process in the underlying assets $\bar{x} + H \cdot S$, the terminal value of which dominates the terminal pay-off (so, selling the portfolio at the date $T$ covers the liability).

Similarly, suppose that the right extremity $\underline{x}$ of the semi-infinite interval

$$-\Gamma(-\xi) = \{x : \exists H \in \mathcal{P} \text{ such that } -x + H \cdot S_T \geq -\xi\}$$

belongs to this interval. If $x$ is strictly less than $\underline{x}$, then the option buyer will have an arbitrage opportunity. Indeed, in this case there exists a strategy $H$ such that $-\underline{x} + H \cdot S_T \geq -\xi$. Thus, borrowing $x$ at $t = 0$ to buy the option, the agent runs a portfolio $-x + H \cdot S$, which has a terminal value larger than $\underline{x} - x - \xi$. Therefore, after exercising the option, the agent will have a nonrisky profit $\underline{x} - x$.

These arguments show that "fair" prices lie in the interval $[\underline{x}, \bar{x}]$.

**Remark 1.** Note that it is tacitly assumed that the agent (option seller) may have a short position in option: for the discrete-time model, it is an innocent assumption, but it is questionable for continuous-time models, where the admissibility means that unbounded short positions even in the underlying are not allowed.

In the case where the contingent claim is redundant, that is, of the form $\xi = x + H^{\xi} \cdot S_T$, we necessarily have that $x = \underline{x} = \bar{x}$ is the no-arbitrage price of the option. Indeed, let us consider the hedging portfolio process $\bar{x} + H \cdot S$ for $\xi$. The absence of arbitrage implies that its terminal value must coincide with $\xi$ and, in virtue of the "law of one price" (also due to NA, see the remark below), $x = \bar{x}$ and, by symmetry, $x = \underline{x}$. The same NA arguments show that if $\xi$ is nonredundant, the hedging portfolio starting from $\bar{x}$ is an arbitrage opportunity. Thus, the range of no-arbitrage prices is either a singleton or an open interval $]\underline{x}, \bar{x}[$.

**Remark 2.** The law of one price (L1P) is the property asserting that the equality $x + H \cdot S_T = x' + H' \cdot S_T$ implies the equality $x = x'$. The NA-property is a sufficient condition for L1P that follows from the DMW theorem: the latter ensures that there is a measure under which the process $(H - H') \cdot S$ is a martingale. One may ask what is a necessary and sufficient condition for L1P. The answer is the following:

*L1P holds if and only if there is a bounded martingale $Z$ with $EZ_T = 1$ and $Z_0 > 0$ such that the process $ZS$ is a martingale.*

In this formulation we do not suppose that the $\sigma$-algebra is trivial. Notice that L1P means that $R_T \cap L^0(\mathcal{F}_0) = \{0\}$, where, as already mentioned, the linear space $R_T$ is closed. We hope that with this remark the proof of the nontrivial "only if" part will be an easy exercise for the reader.

Now we present the theorem giving a "dual" description of the set of initial capitals $\Gamma$, from which one can super-replicate (hedge) the contingent claim $\xi$.

**Notation.** Let $\mathcal{Q}$ (resp. $\mathcal{Q}^e$) be the set of all measures $Q \ll P$ (resp. $Q \sim P$) such that $S$ is a martingale with respect to $Q$. We add to this notation the subscript $l$ to denote larger sets of measures $\mathcal{Q}_l$ and $\mathcal{Q}_l^e$ for which $S$ is only a local martingale. We shall denote by $\mathcal{Z}$, $\mathcal{Z}^e$, $\mathcal{Z}_l$, ... the density processes for measures from the corresponding sets.

**Theorem 2.1.11** *Suppose that $\mathcal{Q}^e \neq \emptyset$. Let $\xi$ be a bounded from below random variable such that $E_Q|\xi| < \infty$ for every $Q \in \mathcal{Q}^e$. Then*

$$\Gamma = \big\{x : \ x \geq E\rho_T\xi \text{ for all } \rho \in \mathcal{Z}^e\big\}. \tag{2.1.1}$$

In other words, $\bar{x} = \sup_{Q\in\mathcal{Q}^e} E_Q\xi$ and $\Gamma = [\bar{x}, \infty]$. An obvious corollary of this theorem (applied to the set $\Gamma(-\xi)$) is the assertion that $\underline{x} = \inf_{Q\in\mathcal{Q}^e} E_Q\xi$.

The direct proof of this result is not difficult, but we obtain it from two fundamental facts having their own interest. The first one usually is referred to as the *optional decomposition theorem*, which will be discussed in Sect. 2.1.12.

**Theorem 2.1.12** *Suppose that $\mathcal{Q}^e \neq \emptyset$. Let $X = (X_t)$ be a bounded from below process which is a supermartingale with respect to each probability measure $Q \in \mathcal{Q}^e$. Then there exist a strategy $H$ and an increasing process $A$ such that $X = X_0 + H \cdot S - A$.*

**Proposition 2.1.13** *Suppose that $Q^e \neq \emptyset$. Let $\xi$ be a bounded from below random variable such that $\sup_{Q\in\mathcal{Q}^e} E_Q|\xi| < \infty$. Then the process $X$ with*

$$X_t = \operatorname*{ess\,sup}_{Q\in\mathcal{Q}^e} E_Q(\xi|\mathcal{F}_t)$$

*is a supermartingale with respect to every $Q \in \mathcal{Q}^e$.*

For the proof of this result, we send the reader to Appendix (Proposition 5.3.7).

*Proof of Theorem 2.1.11.* The inclusion $\Gamma \subseteq [\bar{x}, \infty[$ is obvious: if $x + H\cdot S_T \geq \xi$, then $x \geq E_Q\xi$ for every $Q \in \mathcal{Q}^e$. To show the opposite inclusion, we may suppose that $\sup_{Q\in\mathcal{Q}} E_Q|\xi| < \infty$ (otherwise both sets are empty). Applying the optional decomposition theorem, we get that $X = \bar{x} + H \cdot S - A$. Since $\bar{x} + H \cdot S_T \geq X_T = \xi$, the result follows. □

### 2.1.10 Stochastic Discounting Factors

In this subsection we discuss financial aspects of the hedging theorem and give an interpretation of densities of martingale measures as stochastic discounting factors.

Let us consider a "practical example" where the option seller promised to deliver at the expiration date $T$ a "basket" of $d$ assets, namely, $\eta^i$ units of the $i$th asset with positive price process $S^i$. Since the market is frictionless, this is same as to deliver $\eta S_T$ units of the numéraire, i.e., to make a payment $\xi = \eta S_T$. The hedging theorem asserts that the set of initial capitals allowing one to super-replicate $\xi$ can be described in terms of prices. Namely, if the NA-property holds, one can hedge the pay-off from the initial capital $x$ if and only if $x$ dominates the expectation of "stochastically discounted" pay-off $\rho_T \xi = \xi S_T^\rho$ whatever is a martingale density $\rho$. In other words, the comparison should be done not by computing the "value" of the basket using the "true" price process but replacing the latter by a "consistent price system" $S^\rho = \rho S$ obtained by multiplying the "true" price process by the *stochastic discounting factor* $\rho$. The word "consistent" here reflects the fact that $S_t^\rho$ is determined by $S_T^\rho$ via the martingale property: $S_t^\rho = E(S_T^\rho | \mathcal{F}_T)$.

### 2.1.11 Hedging Theorem for American-Type Options

In the American-type option the buyer has the right to exercise at any date before $T$ on the basis of the available information flow, so the exercise date $\tau$ is a stopping time; the buyer gets the amount $Y_\tau$, the value of an adapted process $Y$ at $\tau$. The description of the pay-off process $Y = (Y_t)$ is a clause of the contract (as well as the final maturity date $T$).

By analogy with the case of European options, we define the set of initial capitals starting from which one can run a self-financing portfolio the values of which dominate the eventual pay-off on the considered time-interval:

$$\Gamma := \Gamma(Y) := \{x : \exists H \in \mathcal{P} \text{ such that } x + H \cdot S \geq Y\}.$$

**Theorem 2.1.14** *Suppose that $\mathcal{Q}^e \neq \emptyset$. Let $Y = (Y_t)$ be an adapted process bounded from below and such that $E_Q|Y_t| < \infty$ for all $Q \in \mathcal{Q}^e$ and $t \leq T$. Then*

$$\Gamma = \Big\{x : \; x \geq E\rho_\tau Y_\tau \text{ for all } \rho \in \mathcal{Z}^e \text{ and all stopping times } \tau \leq T\Big\}. \quad (2.1.2)$$

The proof of this result based on application of the optional decomposition is exactly the same as of Theorem 2.1.11. The only difference is that now we take as $X$ the process

$$X_t = \operatorname*{ess\,sup}_{Q \in \mathcal{Q}^e, \tau \in \mathcal{T}_t} E_Q(Y_\tau | \mathcal{F}_t),$$

where $\mathcal{T}_t$ is the set of stopping times with values in the set $\{t, t+1, \ldots, T\}$. Under the assumption $\sup_{Q \in \mathcal{Q}^e} E_Q|Y_t| < \infty$ for each $t$, the process $X$ is a supermartingale with respect to every $Q \in \mathcal{Q}^e$, see Proposition 5.3.8 in Appendix.

### 2.1.12 Optional Decomposition Theorem

We give here a slightly different formulation.

**Theorem 2.1.15** *Suppose that $\mathcal{Q}_l^e \neq \emptyset$. Let $X = (X_t)$ be a process which is a generalized supermartingale with respect to each measure $Q \in \mathcal{Q}_l^e$. Then there are a strategy $H$ and an increasing process $A$ such that $X = X_0 + H \cdot S - A$.*

*Proof.* We start from a one-step version of the result. □

**Lemma 2.1.16** *Let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$, and let $\xi$ and $\eta$ be random variables with values in $\mathbf{R}$ and $\mathbf{R}^d$ and for which $E(|\xi| + |\eta||\mathcal{G}) < \infty$. Assume that $E(\alpha\xi|\mathcal{G}) \leq 0$ whatever is a random variable $\alpha > 0$ with $E(\alpha|\mathcal{G}) = 1$ such that $E(\alpha\eta|\mathcal{G}) = 0$ and $E(\alpha|\xi||\mathcal{G}) < \infty$, $E(\alpha|\eta||\mathcal{G}) < \infty$. Suppose that such $\alpha$ does exist. Then there is $\lambda \in L^0(\mathbf{R}^d, \mathcal{G})$ such that $\xi - \lambda\eta \leq 0$.*

*Proof.* First, we suppose without loss of generality that $\xi$ and $\eta$ are integrable (we may argue with $\tilde{\xi} := \xi/(1+E(|\xi|+|\eta||\mathcal{G}))$ and $\tilde{\eta} := \eta/(1+E(|\xi|+|\eta||\mathcal{G}))$). Define the set $A := \{\lambda\eta : \ \lambda \in L^0(\mathbf{R}^d, \mathcal{G})\} - L^0_+$. By the DMW theorem, it is closed in probability. Thus, the convex set $A^1 := A \cap L^1$ is closed in $L^1$. If the assertion of the lemma fails, $\xi \notin A^1$. Therefore, in virtue of the Hahn–Banach separation theorem, there is $\alpha \in L^\infty$ such that

$$E\alpha\xi > \sup_{\zeta \in A^1} E\alpha\zeta.$$

Necessarily, $\alpha \geq 0$: if not, the right-hand side of the above inequality would be infinite. By the same reason $E\alpha\lambda\eta = 0$ whatever is $\lambda \in L^\infty(\mathbf{R}^d, \mathcal{G})$. Hence, $E(\alpha\eta|\mathcal{G}) = 0$, and the supremum is equal to zero. That is, $E\alpha\xi > 0$. But this is incompatible with the inequality $E(\alpha\xi|\mathcal{G}) \leq 0$ we should have for such $\alpha$. □

With the lemma, the proof of the theorem is easy. Indeed, let $\rho \in \mathcal{Z}_l^e$. Consider the obvious identity $\rho_t = \alpha_1 \dots \alpha_t$, where $\alpha_k := \rho_k/\rho_{k-1}$. The martingale property of $\rho$ means that $E(\alpha_t|\mathcal{F}_{t-1}) = 1$. On the other hand, due to the coincidence of the classes of local and generalized martingales, $\rho \in \mathcal{Z}_l^e$ if and only if $E(\alpha_t|\Delta S_t||\mathcal{F}_{t-1}) < \infty$ and $E(\alpha_t \Delta S_t|\mathcal{F}_{t-1}) = 0$ for all $t \leq T$. Thus, by Lemma 2.1.16, there is $H_t \in L^0(\mathbf{R}^d|\mathcal{F}_{t-1})$ such that $\Delta X_t - H_t\Delta S_t \leq 0$. Denoting the right-hand side by $-\Delta A_t$ and putting $A_0 = 0$, we obtain the desired decomposition.

**Remark.** Let us return to the setting of Lemma 2.1.16, assuming that the $\sigma$-algebra $\mathcal{G}$ is trivial. Consider the maximization problem $E\alpha\xi \to \max$ under two equality constraints $E\alpha = 1$ and $E\alpha\eta = 0$, the constraint $\alpha > 0$ (a.s.), and "admissibility" assumptions on $\alpha$ to ensure the needed integrability. The hypothesis of the lemma says that the value of this problem does not exceed zero. It is not difficult to prove that there is a Lagrange multiplier $\lambda$ "removing" the second equality constraint. For the new maximization problem, we also have that $E\alpha(\xi - \lambda\eta) \le 0$ for all $\alpha$ satisfying the remaining constraints. Clearly, this is possible only if $\xi - \lambda\eta \le 0$.

One may expect that these arguments can be extended for the general case, with conditional expectations. This is still easy for finite or countable $\Omega$. This strategy of proof is feasible for arbitrary $\Omega$, but one needs to look for a $\mathcal{G}$-measurable version of Lagrange multipliers by applying a delicate measurable selection result, requiring, in turn, specific preparations. However, this approach (inspired by the original proof of the DMW) works well also for continuous-time models, see [73]. We use it for an analysis of the structure of the set of equivalent martingale measures in the next subsection.

### 2.1.13 Martingale Measures with Bounded Densities

The following useful result gives, in particular, the positive answer to the question whether the set $\mathcal{Q}^e$ is norm-dense in $\mathcal{Q}^e_l$ (that is whether $\mathcal{Z}^e$ is dense in $\mathcal{Z}^e_l$ in the $L^1$-norm). Indeed, in virtue of the DMW theorem, $\mathcal{Q}^e_l \neq \emptyset$ if and only if $\mathcal{Q}^e \neq \emptyset$. It remains to take as the reference measure an arbitrary element of the latter set and apply the theorem below. This theorem happens to be useful to get a similar property for the discrete-time model with infinite horizon, which will be discussed in the next section.

**Theorem 2.1.17** *Let $P \in \mathcal{Q}^e_l$. Then the set $\{Q \in \mathcal{Q}^e,\ dQ/dP \in L^\infty\}$ is norm-dense in $\mathcal{Q}^e_l$.*

*Proof.* It contains three steps. The first one is a simple lemma on the approximations of positive functions on the probability space $(\mathbf{R}^m, \mathcal{B}^n, \mu)$ by positive functions from $C(\bar{\mathbf{R}}^m)$, where $\bar{\mathbf{R}}^m$ is the one-point compactification of $\mathbf{R}^m$. □

**Lemma 2.1.18** *Let $\phi : \mathbf{R}^m \to \mathbf{R}^l$ be a measurable mapping with $|\phi| \in L^1(\mu)$. Put $U := \{g \in L^1(\mu) :\ g > 0,\ g|\phi| \in L^1(\mu)\}$ and $U_C := U \cap C(\bar{\mathbf{R}}^m)$. Then for any $f \in U$ and $\varepsilon > 0$, there is $f^\varepsilon \in U_C$ such that $\|f - f^\varepsilon\|_{L^1(\mu)} < \varepsilon$ and*

$$E_\mu \phi f = E_\mu \phi f^\varepsilon. \tag{2.1.3}$$

*Proof.* Let $\mathcal{O}_\varepsilon(f)$ be an open ball in $L^1_\mu$ of radius $\varepsilon$ with center at $f$. Define the convex sets $G := U \cap \mathcal{O}_\varepsilon(f)$ and $G_C := U_C \cap \mathcal{O}_\varepsilon(f)$ and consider the affine mapping $\Phi : G \to \mathbf{R}^l$ with $\Phi(g) = E_\mu(f - g)\phi$. We need to show that $0 \in \Phi(G_C)$. Notice that $U_C$ is a dense subset of $U$, and, therefore, $G_C$ is dense in $G$ in $L^1_\mu$. It follows that $\Phi(G_C)$ is dense in $\Phi(G)$. The convexity of these sets implies that $\operatorname{ri}\Phi(G_C) = \operatorname{ri}\Phi(G)$, and to complete the proof, it is sufficient to check that $0 \in \operatorname{ri}\Phi(G)$. To this aim we first observe that without loss of generality we may consider the case where $f\phi^i$, $i = 1, \dots, l$, are linearly independent elements of $L^1_\mu$. Suppose that $0 \notin \operatorname{ri}\Phi(G)$. Let us consider the smallest hyperplane $H$ containing $\Phi(G)$. Since $0 \in \Phi(G)$, it is a subspace. By the separation theorem, there is a nontrivial linear functional $y$ on $H$ such that

$yx \geq 0$ for all $x \in \Phi(G)$. Extending $y$ to a linear functional on the whole $\mathbf{R}^l$, we may rewrite this as $E_\mu(f-g)y\phi \geq 0$ whatever is $g \in G$. Using functions of the form $g = f \pm \delta f I_\Gamma$ where $\Gamma$ is a measurable set and $\delta \in ]0,1[$ is such that $g \in \mathcal{O}_\varepsilon(f)$, we get from here that $E_\mu I_\Gamma f y\phi = 0$ for any $\Gamma$. Hence, $yf\phi = 0$, in contradiction with the assumed linear independence of components. □

With this preparatory result, we can easily prove the claim for the one-period model.

**Lemma 2.1.19** *Let $\mathcal{G}$ be a (complete) sub-$\sigma$-algebra of $\mathcal{F}$, and let $\alpha$ and $\eta$ be random variables taking values, respectively, in $\mathbf{R}_+ \setminus \{0\}$ and $\mathbf{R}^d$ such that $E((1+\alpha)|\eta||\mathcal{G}) < \infty$. Assume that $E(\alpha|\mathcal{G}) = 1$, $E(\eta|\mathcal{G}) = 0$, and $E(\alpha\eta|\mathcal{G}) = 0$. Then there are bounded random variables $\alpha^n > 0$ converging to $\alpha$ a.s. and such that $E(\alpha^n|\mathcal{G}) = 1$, $E(\alpha^n\eta|\mathcal{G}) = 0$.*

*Proof.* Let $\mu(dx,\omega)$ be a regular conditional distribution of the random vector $(\alpha,\eta)$ knowing $\mathcal{G}$. Define on $\mathbf{R}^{d+1}$ the functions $f(x) := x^1$ and $\phi(x) := (1, x^2, \dots, x^{d+1})$. Writing the conditional expectations as the integrals with respect to conditional distribution, we express properties of $\alpha$ as follows: $E_{\mu(.,\omega)} f\phi = e_1$ (the first orth in $\mathbf{R}^{d+1}$) for all $\omega$ except a null-set. The set

$$\Gamma^n := \big\{(\omega, g) \in \Omega \times C\big(\bar{\mathbf{R}}^{d+1}\big) :\ g > 0,\ E_{\mu(.,\omega)} g\phi = e_1,\\ E_{\mu(.,\omega)}|f-g| < 1/n\big\}$$

is $\mathcal{G} \otimes \mathcal{B}(C(\bar{\mathbf{R}}^{d+1}))$-measurable and, according to the previous lemma, has the projection on $\Omega$ of full measure. By the classical measurable selection theorem $\Gamma^n$ admits a $\mathcal{G}$-measurable selector $f^n : \Omega \to C(\bar{\mathbf{R}}^{d+1})$. The function of two variables $f^n(\omega,x)$, being $\mathcal{G}$-measurable in $\omega$ and continuous in $x$, is $\mathcal{G} \otimes \mathcal{B}^{d+1}$-measurable. The random variables $\tilde{\alpha}^n = f^n(\omega,(\alpha(\omega),\eta(\omega)))$ converge to $\alpha$ in $L^1$ and, hence, in probability. Let us define the bounded random variables $\tilde{\alpha}^{n,k}(\omega) := \tilde{f}^{n,k}(\omega,(\alpha(\omega),\eta(\omega)))$, where

$$\tilde{f}^{n,k}(\omega,x) = f^n(\omega,x) I_{\{\|f^n(\omega,.)\| \leq k\}} + I_{\{\|f^n(\omega,.)\| > k\}},$$

and $\|.\|$ is a uniform norm in $x$. Since $E_{\mu(.,\omega)} g\phi = e_1$, we have the equalities $E(\tilde{\alpha}^{n,k}|\mathcal{G}) = 1$, $E(\tilde{\alpha}^{n,k}\eta|\mathcal{G}) = 0$.

Obviously, $\tilde{\alpha}^{n,k}$ converge to $\tilde{\alpha}^n$ in probability. The convergence in probability is a convergence in a metric space, and, therefore, one can take a subsequence $k_n$ such that $\alpha^n := \tilde{\alpha}^{n,k_n}$ converge to $\alpha$ in probability. But then there is a subsequence of $\alpha^n$ convergent to $\alpha$ a.s.

The third, concluding step, is also simple. Note first that we may replace the reference measure by any other from $Q_l^e$ with bounded density. According to the DMW theorem, between such measures, there are measures from $Q^e$, and so we may assume without loss of generality that already $P \in Q^e$.

We again use the multiplicative representation of the density $\rho_T = dQ/dP$, namely, $\rho_T = \alpha_1 \dots \alpha_T$ with $\alpha_t := \rho_t/\rho_{t-1}$. The property $\rho \in \mathcal{Z}_l^e$ holds if and only if $E(\alpha_t|\mathcal{F}_{t-1}) = 1$, $E(\alpha_t|\Delta S_t||\mathcal{F}_{t-1}) < \infty$ and $E(\alpha_t \Delta S_t|\mathcal{F}_{t-1}) = 0$ for all $t \le T$. Applying the preceding lemma, we define the measure $P^n := \rho_T^n P \in Q^e$ with bounded density $\rho_T^n := \alpha_1^n \dots \alpha_T^n$ convergent to $\rho_T$ a.s. But by the Scheffe theorem, here we also have the convergence in $L^1$. $\square$

**Remark.** Theorem 2.1.17 has several obvious corollaries. For example, if $P \in Q_l^e$, then the set of $Q \in \mathcal{Q}_l^e$ with bounded densities $\rho_T = dQ/dP$ and $\rho_T^{-1} = dP/dQ$ is dense in $\mathcal{Q}_l^e$. This fact is easily seen by considering the convex combinations $Q^n = (1 - 1/n)Q + (1/n)P$ and letting $n$ tend to infinity. Noticing that $\mathcal{Q}_l^e$ is dense in the set $\mathcal{Q}_l$ (by the similar consideration), one can further strengthen the claim in another direction, etc.

It is not difficult to check that the set of local martingale measures with finite entropy (i.e., with $E\rho_T \ln \rho_T < \infty$), if nonempty, is also dense in $Q_l^e$. We explain the idea by establishing a more general result, which has applications in portfolio optimization problems.

Let $\varphi : ]0,\infty[ \to \mathbf{R}$ be a measurable function bounded from below, and let

$$\mathcal{Q}_\varphi^e := \big\{ Q \in \mathcal{Q}^e : \; E\varphi(dQ/dP) < \infty \big\}.$$

**Proposition 2.1.20** *If the set $\mathcal{Q}_\varphi^e \neq \emptyset$, then it is dense in $\mathcal{Q}^e$ in the following two cases:*

(a) *for every $c \ge 1$, there exist constants $r_1(c), r_2(c) \ge 0$ such that*

$$\varphi(\lambda y) \le r_1(c)\varphi(y) + r_2(c)(y+1), \quad y \in ]0,\infty[, \; \lambda \in \big[c^{-1}, c\big]; \tag{2.1.4}$$

(b) *$\varphi$ is convex, and $\mathcal{Q}_\varphi^e = \mathcal{Q}_{\varphi_\lambda}^e$ for any $\lambda > 0$, where $\varphi_\lambda(y) := \varphi(\lambda y)$.*

*Proof.* (a) Let $\tilde{P} \in \mathcal{Q}_\varphi^e$. Take an arbitrary measure $Q \in \mathcal{Q}^e$. By the above theorem and the accompanying remark, there exists a sequence $Q^n \in \mathcal{Q}^e$ convergent to $Q$ with the densities $dQ^n/d\tilde{P}$ taking values in intervals $[c_n^{-1}, c_n]$. We have

$$E\varphi\bigg(\frac{dQ^n}{dP}\bigg) = E\varphi\bigg(\frac{dQ^n}{d\tilde{P}}\frac{d\tilde{P}}{dP}\bigg) \le r_1(c_n)E\varphi\bigg(\frac{d\tilde{P}}{dP}\bigg) + 2r_2(c_n) < \infty.$$

Hence, $Q^n \in \mathcal{Q}_\varphi^e$, and the result follows.

(b) We may assume without loss of generality that $\varphi \ge 0$ (by adding a constant) and repeat the same arguments modifying only the last step. Clearly, $dQ^n/d\tilde{P} = \alpha_n c_n^{-1} + (1-\alpha_n)c_n$, where $\alpha_n$ is a random variable taking values in $[0,1]$. By convexity,

$$\begin{aligned} E\varphi\bigg(\frac{dQ^n}{d\tilde{P}}\frac{d\tilde{P}}{dP}\bigg) &\le E\bigg[\alpha_n\varphi\bigg(c_n^{-1}\frac{d\tilde{P}}{dP}\bigg) + (1-\alpha_n)\varphi\bigg(c_n\frac{d\tilde{P}}{dP}\bigg)\bigg] \\ &\le E\varphi\bigg(c_n^{-1}\frac{d\tilde{P}}{dP}\bigg) + E\varphi\bigg(c_n\frac{d\tilde{P}}{dP}\bigg) < \infty \end{aligned}$$

in virtue of assumption, and we conclude as before.

Note that for convex $\varphi$, condition (a) implies (b). The latter hypothesis entangles properties of $\varphi$ and $\mathcal{Q}^e$.

In financial applications, typically, $\varphi(y) = y^p$, $p > 0$, or $\varphi(y) = y \ln y$. In particular, if nonempty, the set $\mathcal{Q}^e_{y \ln y}$ of martingale measures with finite entropy is dense in the set of equivalent martingale measures $\mathcal{Q}^e$.

More generally, let $u : \mathbf{R} \to \mathbf{R}$ be an increasing *concave* differentiable function, and let $u^*$ be its Fenchel dual (which is, by definition, the Fenchel dual of the *convex* function $-u(-.)$), i.e., the convex function

$$u^*(y) = \sup_x \big(u(x) - xy\big).$$

For example, the dual of the exponential utility function $u(x) = 1 - e^{-x}$ is the function $u^*(y) = y \ln y - y + 1$, $y \geq 0$, and $u^*(y) = \infty$, $y < 0$.

Suppose that $u$ has a "reasonable" asymptotic elasticity, i.e.,

$$AE_+(u) := \limsup_{x \to \infty} \frac{xu'(x)}{u(x)} < 1, \qquad AE_-(u) := \liminf_{x \to -\infty} \frac{xu'(x)}{u(x)} > 1.$$

It can be shown that the function $\varphi = u^*$ satisfies the growth condition (a) of Proposition 2.1.20. □

### 2.1.14 Utility Maximization and Convex Duality

In this subsection we explain the importance of the set of equivalent martingale measures in the problem of portfolio optimization. Namely, we consider the simplest model with finite number of states of nature where the investor maximizes the mean value of an exponential utility function of the terminal value of his portfolio. Applying the classical Fenchel theorem, we show that the dual problem involves martingale measures.

So, $\Omega$ is finite, and hence, the space $L^0$ can be identified with a finite-dimensional Euclidean space. As usual, $R_T$ is the set of random variables $H \cdot S_T$, and $\mathcal{Z}^e_T$ (respectively, $\mathcal{Z}^e$) is the set of densities (respectively, density processes) of equivalent martingale measures.

It is *supposed* in the following discussion that $\mathcal{Z}^e_T \neq \emptyset$.

We are interested in the portfolio optimization problem the value of which is

$$J^o := \sup_{\eta \in R_T} E\big(1 - e^{-\eta}\big). \tag{2.1.5}$$

It will be studied jointly with the minimization problem $EZ_T \ln Z_T \to \min$ over the set of equivalent martingale densities $\mathcal{Z}^e_T$; let its value be

$$\underline{J} := \inf_{\xi \in \mathcal{Z}^e_T} E\xi \ln \xi. \tag{2.1.6}$$

The latter problem, by abuse of language, sometimes is referred to as the "dual" one. As we shall see below, this terminology deviates from the standard one of the convex analysis.

The continuous function $\xi \to E\xi \ln \xi$ attains its minimum $\underline{J}$ on the compact set $\mathcal{Z}_T^a$ which is a closure of $\mathcal{Z}_T^e$. Due to strict convexity, the minimizer $\xi^o$ is unique. The derivative of the function $\varphi(x) = x \ln x$ (with $\varphi(0) = 0$) at zero is $-\infty$, and this property implies that $\xi^o$ is strictly positive. Indeed, let us take an arbitrary point $\xi$ of the set $\mathcal{Z}_T^e$ (assumed nonempty) and consider on $[0,T]$ the function $F_t = Ef_t$, where $f_t := \varphi(t\xi + (1-t)\xi^o)$. As $F$ attains its minimum at $t = 0$, the derivative $F_0' \geq 0$. But $F_0' = Ef_0'$. Since $f_0' = \varphi'(0)\xi = -\infty$ on the set $\{\xi^o = 0\}$, the probability of the latter is zero.

The measure $P^o = \xi^o$ is called the *entropy minimal* equivalent martingale measure.

**Proposition 2.1.21** $J^o = 1 - e^{-\underline{J}}$.

This result is a direct consequence of the fundamental Fenchel theorem. We recall the simplest version of its formulation (in its traditional form, for convex functions).

Let $X$ be a Hilbert space, and let $f : X \to \mathbf{R} \cup \{\infty\}$, $g : X \to \mathbf{R} \cup \{\infty\}$ be two convex lower semicontinuous functions not identically equal to infinity, i.e., $\operatorname{dom} f \neq \emptyset$ and $\operatorname{dom} g \neq \emptyset$.

Let us consider two minimization problems, the primal

$$f(\eta) + g(\eta) \to \min \quad \text{on } X \tag{2.1.7}$$

and the dual

$$f^*(-\xi) + g^*(\xi) \to \min \quad \text{on } X^*(=X). \tag{2.1.8}$$

We denote their values $v := \inf_x[f(x) + g(x)]$ and $v_* := \inf_y[f^*(-y) + g^*(y)]$.

Suppose that

$$\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset, \qquad \big(-\operatorname{dom} f^*\big) \cap \operatorname{dom} g^* \neq \emptyset;$$

we rewrite these conditions, to relate them with those in the formulation of theorem below, as

$$0 \in \operatorname{dom} f - \operatorname{dom} g, \qquad 0 \in \operatorname{dom} f^* + \operatorname{dom} g^*.$$

They ensure that $v < \infty$ and $v_* < \infty$.

Note that always $v + v_* \geq 0$ because by the Fenchel inequality

$$f(\eta) + g(\eta) + f^*(-\xi) + g^*(\xi) \geq (-\eta, \xi) + (\eta, \xi) = 0.$$

The following result is a particular case of the Fenchel theorem.

**Theorem 2.1.22** (a) *Let* $0 \in \operatorname{int}(\operatorname{dom} f - \operatorname{dom} g)$. *Then the dual problem (2.1.8) has a solution, and* $v + v_* = 0$.

(b) *Let* $0 \in \operatorname{int}(\operatorname{dom} g^* + \operatorname{dom} f^*)$. *Then the primal problem (2.1.7) has a solution, and* $v + v_* = 0$.

Let us consider the minimization problem

$$f(\eta) + g(\eta) \to \min \quad \text{on } L^0, \tag{2.1.9}$$

where $f(\eta) := E(e^\eta - 1)$, and $g = \delta_{R_T}$, the indicator function (in the sense of convex analysis), which is equal to zero on $R_T$ and infinity on the complement. Clearly, $f^*$ is calculated via the dual to the convex function $e^x - 1$, namely, $f^*(-\xi) = E(\xi \ln \xi - \xi + 1)\delta_{[0,\infty[}(\xi)$, and $g^* = \delta_{R_T^\circ}$. In our case the polar $R_T^\circ$ is just $R_T^\perp$, the subspace orthogonal to $R_T$. The conditions of the Fenchel theorem, part (a), are obviously fulfilled. Thus, $J^o$ coincides with the (attained) value of the dual problem

$$f^*(-\xi) + g^*(\xi) \to \min \quad \text{on } L^0,$$

i.e., $J^o$ is equal to the minimum of $f^*(-\xi)$ on the set $R_T^\perp \cap L_+^0 = \mathbf{R}_+ \mathcal{Z}_T^a$. Since

$$\inf_{\xi \in \mathcal{Z}} \inf_{t \geq 0} E(t\xi \ln t\xi - t\xi + 1) = \inf_{\xi \in \mathcal{Z}_T^a} \left(1 - e^{-E\xi \ln \xi}\right) = 1 - e^{-\underline{J}},$$

we get the result.

**Remark**. If $\mathcal{Z}^e \neq \emptyset$, the hypothesis of the Fenchel theorem, part $(b)$, holds, ensuring the existence in the primal problem. In the case where $\mathcal{Z}^e = \emptyset$, there is an arbitrage strategy $H^a$ with $\eta^a := H^a \cdot S_T \geq 0$ and $\eta^a \neq 0$. Clearly, for any $\eta \in R_T$, the value of the functional in (2.1.5) at $\eta^a + \eta$ is strictly larger than at $\eta$.

**Proposition 2.1.23** *Let $H^o$ be the optimal strategy for the problem of portfolio optimization, Then the random variable*

$$\xi^o := e^{-H^o \cdot S_T} / E e^{-H^o \cdot S_T} \tag{2.1.10}$$

*is the density of the minimal entropy equivalent martingale measure $P^o$.*

*Proof.* The right-hand side of (2.1.10) is the density of a martingale measure. Indeed, for any strategy $H$, the function

$$f_H(\lambda) = 1 - E e^{-H^o \cdot S_T + \lambda H \cdot S_T}$$

attains its maximum at $\lambda = 0$, and, therefore, $f_H'(0) = 0$, i.e.,

$$E(H \cdot S_T) e^{-H^o \cdot S_T} = 0,$$

implying the claimed property. Using it, we can easily verify that

$$1 - e^{-E\xi^o \ln \xi^o} = 1 - E e^{-H^o \cdot S_T}.$$

Accordingly to Proposition 2.1.21, this equality may hold only if $\xi^o$ is the solution of the problem of the entropy minimization. □

## 2.2 Discrete-Time Infinite-Horizon Model

The aim of this section is to present relations between the absence of arbitrage and the existence of an equivalent martingale measure for the model with an $\mathbf{R}^d$-valued price process $S = (S_t)_{t=0,1,\ldots}$ defined on some filtered space $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F})_{t=0,1,\ldots})$. We assume that the initial $\sigma$-algebra is trivial. In the first subsection we discuss some purely probabilistic questions. In particular, we show that if $S$ admits an equivalent local martingale measure, then it admits an equivalent martingale measure. Moreover, the latter can be chosen to ensure the integrability of an arbitrary adapted process fixed in advance. Afterwards we introduce some substitutes for the no-arbitrage property and prove necessary and sufficient conditions for them.

### 2.2.1 Martingale Measures in Infinite-Horizon Model

We consider the discrete-time infinite-horizon model with an $\mathbf{R}^d$-valued process $S = (S_t)_{t\geq 0}$ and introduce, for $p \geq 1$, the set $\mathcal{Q}^{e,p}$ of probability measures $Q \sim P$ such that $S$ is a $Q$-martingale and $S_t \in L^p(Q)$ for all $t \geq 0$. We also use the standard notation $S_t^* := \sup_{s\leq t} |S_s|$.

**Theorem 2.2.1** *Let $S \in \mathcal{M}_{\mathrm{loc}}(P)$. Then there exists a probability measure $\tilde{P} \sim P$ such that $S \in \mathcal{M}(\tilde{P})$.*

In the case of finite time-horizon this assertion is a direct corollary of the DMW theorem (and the measure $\tilde{P}_T \sim P$ even can be chosen with the bounded density $d\tilde{P}_T/dP$). For the infinite-time horizon, we get it from the following much more general assertion.

**Theorem 2.2.2** *Let $S$ be a local martingale, and $Y = (Y_t)$ be an adapted process dominating $S^*$. Let $\varepsilon > 0$. Then there exists a measure $P' \sim P$ such that $S$ is a local $P'$-martingale, $Y_t \in L^1(P')$ for every $t$, and $\|P' - P\| \leq \varepsilon$.*

As an obvious corollary, we have:

**Theorem 2.2.3** *The set $\mathcal{Q}^{e,p}$ is dense in the set $\mathcal{Q}_l^e$.*

Theorem 2.2.2 is a generalization of Theorem 2.1.17. It is interesting that the reference to the latter constitutes the essential ingredients of the proof.

**Lemma 2.2.4** *Let $S = (S_t)_{t\leq T}$ be a local martingale in $\mathbf{R}^d$, and let $\xi \in L^0_+$. Then for any $\varepsilon > 0$, there is a probability measure $P^\varepsilon \sim P$ with density $Z_T^\varepsilon$ such that $S = (S_t)_{t\leq T}$ is a local martingale with respect to $P^\varepsilon$, $E|Z_T^\varepsilon - 1| < \varepsilon$, and $Z_T^\varepsilon(1+\xi)$ is bounded.*

*Proof.* We introduce the probability measure $P^1 = ce^{-\xi}P$. Since the NA-property holds for $P$, it holds for $P^1$. By the DMW theorem there is $P^2 \sim P^1$ with $dP^2/dP^1 \in L^\infty$ such that $S \in \mathcal{M}(P^2)$. Applying Theorem 2.1.17 with $P^2$ as the reference measure, we obtain that there exists a measure $P^\varepsilon \sim P^2$ with $dP^\varepsilon/dP^2 \in L^\infty$ such that $\|P^\varepsilon - P\| < \varepsilon$ and $S \in \mathcal{M}(P^\varepsilon)$. The measure $P^\varepsilon$ meets the requirements. □

*Proof of Theorem 2.2.2.* We may suppose that $\varepsilon < 1$. Take a sequence $\varepsilon_n > 0$ such that $\sum \varepsilon_n < \varepsilon/3$. We define recursively an auxiliary sequence of probability measures $P^n \sim P$ with bounded $\mathcal{F}_n$-measurable densities $dP^n/dP$. Let $\zeta^k$ denote the density process of $P^k$ with respect to $P$, that is, $\zeta_t^k = E(dP^k/dP|\mathcal{F}_t)$. Put

$$Z_t^n := \zeta_t^1 \dots \zeta_t^n, \qquad \bar{Z}_t^n := E\big(Z_\infty^n|\mathcal{F}_t\big).$$

The construction will ensure the following properties:

(a) $\zeta_n^n(1+Y_n) \le c_n$ for some constant $c_n$;
(b) $E(\bar{Z}_t^n S_t|\mathcal{F}_{t-1}) = \bar{Z}_{t-1}^n S_{t-1}$ for all $t$, i.e., $\bar{Z}^n S$ is a local martingale;
(c) $\|P^n - P\| \le \tilde{\varepsilon}_n := \varepsilon_n/(1+c_0 \dots c_{n-1})$.

Using Lemma 2.2.4, we define a probability measure $P^1 \sim P$ with $\mathcal{F}_1$-measurable density $dP^1/dP$ such that $\zeta_1^1(1+Y_1) \le c_1$ for some constant $c_1$, the process $(S_t)_{t\le 1}$ is a $P^1$-martingale, and $\|P^1 - P\| \le \varepsilon_1$. Note that the whole process $S$ remains a local martingale with respect to $P^1$.

Suppose that the measures $P^k$ for $k \le n-1$ are already constructed. Applying Lemma 2.2.4 to the $(d+1)$-dimensional local martingale $(\bar{Z}_t^{n-1} S_t, \bar{Z}_t^{n-1})_{t\le n}$, we find a measure $P^n \sim P$ with $\mathcal{F}_n$-measurable density such that the properties (a) and (c) hold and $(\bar{Z}_t^{n-1} S_t, \bar{Z}_t^{n-1})_{t\le n}$ is a local $P^n$-martingale. The latter property means that $(\bar{Z}_t^{n-1}\zeta_t^n S_t, \bar{Z}_t^{n-1}\zeta_t^n)_{t\le n}$ is a local martingale. The martingale $(\bar{Z}_t^{n-1}\zeta_t^n)$, having at the date $t=n$ the value $Z_n^{n-1}\zeta_n^n = Z_n^n = Z_\infty^n$, coincides with $(\bar{Z}_t^n)$. Thus, $\bar{Z}^n S$ is a local martingale, and the property (b) holds.

By virtue of our construction,

$$E\sum \big|\zeta_\infty^n - 1\big| \le \sum \varepsilon_n < \infty,$$

and hence $\sum|\zeta_\infty^n - 1| < \infty$ a.s. Thus, $Z_\infty^n$ converges almost surely to some finite random variable $Z_\infty > 0$. Moreover, the convergence holds also in $L^1$ because

$$E\big|Z_\infty^n - Z_\infty^{n-1}\big| = EZ_\infty^{n-1}\big|\zeta_\infty^n - 1\big| \le c_0 \dots c_{n-1} E\big|\zeta_\infty^n - 1\big| \le \varepsilon_n$$

and $\sum \varepsilon_n$ is finite. Also,

$$E|Z_\infty - 1| \le \sum E\big|Z_\infty^n - Z_\infty^{n-1}\big| \le \sum \varepsilon_n < \varepsilon/3.$$

It remains to check that the probability measure $P' = (Z_\infty/EZ_\infty)P$ meets the requirements. Since $EZ_\infty \ge 1 - \varepsilon/3 \ge 2/3$, we have that

$$\|P' - P\| = E|Z_\infty/EZ_\infty - 1| \le \frac{2E|Z_\infty - 1|}{EZ_\infty} \le \varepsilon.$$

It is easy to check that, for each fixed $t$, the sequence $Z_\infty^n Y_t, n = t, t+1, \dots,$ is fundamental in $L^1$. Indeed, we again use the properties (a) and (c):

$$E\big|Z_\infty^n Y_t - Z_\infty^{n-1} Y_t\big| \le E\big|\zeta_n^n - 1\big|\zeta_1^1 \dots \zeta_t^t Y_t \zeta_{t+1}^{t+1} \dots \zeta_{n-1}^{n-1}$$
$$\le c_0 \dots c_{n-1} E\big|\zeta_n^n - 1\big| \le \varepsilon_n.$$

It follows that $Z_\infty^n Y_t$ and $Z_\infty^n S_t$ converges in $L^1$ to integrable random variables. Thus, $Z_\infty Y_t \in L^1$, and, in virtue of (b),

$$E(Z_\infty \Delta S_t | \mathcal{F}_{t-1}) = 0,$$

i.e., $P'$ is a martingale measure. □

### 2.2.2 No Free Lunch for Models with Infinite Time Horizon

Infinite-horizon discrete-time market models based on the price process $(S_t)_{t=0,1,\dots}$ pose new interesting mathematical problems related with the so-called doubling strategies or the St.-Petersburg game. It is well known that if $S$ is a symmetric random walk on integers and, hence, a martingale, the strategy $H_t = 2^t I_{\{t \le \tau\}}$ where $\tau := \inf\{t \ge 1 : \ \Delta S_t = 1\}$ looks as an arbitrage opportunity: $H \cdot S_\infty = 1$. This strategy vanishes after the stopping time $\tau$, which is finite but not bounded. So, certain restrictions on strategies are needed to exclude such a one. A satisfactory criterion relating the existence of an equivalent martingale measure with a strengthened no-arbitrage property can be obtained by assuming that there is no trading after some bounded stopping time where the bound depend on the strategy. Using the concepts and notation developed above, we can formalize this easily.

Let $R_\infty$ be the union of all sets $R_T$, $T \in \mathbf{N}$, and let $A_\infty := R_\infty - L_+^0$.

The infinite-horizon model has the *NA-property* if $R_\infty \cap L_+^0 = \{0\}$ (or, equivalently, $A_\infty \cap L_+^0 = \{0\}$). In general, *NA* is weaker than the *EMM*-property claiming the existence of a probability measure $\tilde{P} \sim P$ such that $S$ is a $\tilde{P}$-martingale. The simplest reinforcing of *NA* is the *NFL*-property ("no-free-lunch") suggested by Kreps: $\bar{C}_\infty^w \cap L_+^\infty = \{0\}$, where $\bar{C}_\infty^w$ is the closure of the set $C_\infty := A_\infty \cap L^\infty$ in the topology $\sigma(L^\infty, L^1)$ (i.e., the weak* closure).

**Theorem 2.2.5** *The following properties are equivalent:*

(a) $\bar{C}_\infty^w \cap L_+^\infty = \{0\}$ *(NFL);*
(b) *there exists* $\tilde{P} \sim P$ *such that* $S \in \mathcal{M}_{\mathrm{loc}}(\tilde{P})$*;*
(c) *there exists* $\tilde{P} \sim P$ *such that* $S \in \mathcal{M}(\tilde{P})$*.*

*Proof.* The Kreps–Yan theorem says that the $NFL$-property holds if and only if there exists $P' \sim P$ such that $E'\xi \le 0$ for all $\xi \in \bar{C}_\infty^w$. This $P'$ can be called a *separating measure* since its density is a functional from $L^1$ which separates $\bar{C}_\infty^w$ and $L_+^\infty$. Of course, a local martingale measure $\tilde{P}$ is a separating one. Indeed, if $H \cdot S_T$ is bounded from below, then the process $(H \cdot S_t)_{t \le T}$ is a $\tilde{P}$-martingale. Hence, for any bounded from below random variable $\xi = H \cdot S_T - h$ where $h \in L_+^0$, we have the inequality $\tilde{E}\xi \le 0$. It follows that this inequality holds for any $\xi \in \bar{C}_\infty^w$. This gives us the implication (b) $\Rightarrow$ (a). The more difficult implication (a) $\Rightarrow$ (b) follows from Theorem 2.2.6 below ensuring that, amongst the equivalent separating measures, there is a local martingale measure. The equivalence (b) $\Leftrightarrow$ (c) follows from Theorem 2.2.2. □

**Theorem 2.2.6** *Any neighborhood of a separating measure contains an equivalent probability measure $P'$ under which $S$ is a local martingale.*

*Proof.* We assume without loss of generality that the reference measure $P$ is separating. Fix $\varepsilon > 0$ and a sequence of numbers $\varepsilon_s > 0$ such that $\sum_{s\geq 1} \varepsilon_s < \varepsilon$. The theorem will be proven if, for each $s \geq 1$, we find an $\mathcal{F}_s$-random variable $\alpha_s > 0$ with the following properties:

(i) $E(\alpha_s|\mathcal{F}_{s-1}) = 1$;
(ii) $E(|1-\alpha_s||\mathcal{F}_{s-1}) \leq \varepsilon_s$;
(iii) $E(\alpha_s|\Delta S_s||\mathcal{F}_{s-1}) < \infty$ and $E(\alpha_s \Delta S_s|\mathcal{F}_{s-1}) = 0$.

Indeed, let us consider the process $Z_t := \alpha_1 \dots \alpha_t$, $t \geq 1$, $Z_0 = 1$, which is a martingale in virtue of (i). In virtue of (ii),

$$E|\Delta Z_s| = EZ_{s-1}E\big(|\alpha_s - 1||\mathcal{F}_{s-1}\big) \leq \varepsilon_s.$$

The martingale $Z$, being dominated by an integrable random variable, namely, by $1 + \sum |\Delta Z_s|$, is uniformly integrable. Also, $E\sum|\alpha_s - 1| < \varepsilon$. Therefore, $\sum|\alpha_s - 1| < \infty$ a.s., and the infinite product $Z_\infty > 0$ a.s. Thus, the probability measure $\tilde{P} = Z_\infty P$ is equivalent to $P$. In virtue of (iii), the process $S$ is a generalized martingale under $\tilde{P}$, i.e., belongs to the class coinciding with $\mathcal{M}_{\text{loc}}(\tilde{P})$. Moreover,

$$E|Z_\infty - 1| \leq E\sum_{s\geq 1}|\Delta Z_s| < \varepsilon.$$

Let $H_s \in L^0(\mathbf{R}^d, \mathcal{F}_{s-1})$ be such that the random variable $H_s \Delta S_s$ is bounded from below. Then $(H_s \Delta S_s) \wedge n$, being an element of $C_\infty$, has a negative expectation—we assumed that $P$ is a separating measure. By the Fatou lemma $EH_s\Delta S_s \leq 0$. In the proposition below we show that this ensures the existence of $\alpha_s$ with the required properties. □

So, we need the following one-step result.

**Proposition 2.2.7** *Let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$. Suppose that $\eta \in L^0(\mathbf{R}^d)$ is such that $E\gamma\eta \leq 0$ for any $\gamma \in L^0(\mathbf{R}^d, \mathcal{G})$ for which $\gamma\eta$ is bounded from below. Let $\varepsilon > 0$. Then there is a strictly positive random variable $\alpha$ such that $E(\alpha|\mathcal{G}) = 1$, $E(|1-\alpha||\mathcal{G}) \leq \varepsilon$, $E(\alpha|\eta||\mathcal{G}) < \infty$, and $E(\alpha\eta|\mathcal{G}) = 0$.*

*Proof.* Let $\mu(dx, \omega)$ be a regular conditional distribution of $\eta$ with respect to $\mathcal{G}$. In the space $\Omega \times C(\bar{\mathbf{R}}^d)$ we consider the $\mathcal{G} \otimes \mathcal{B}(C(\bar{\mathbf{R}}^d))$-measurable set $\Gamma$ defined as the intersection of the sets

$$\big\{(\omega, g):\ g > 0,\ E_{\mu(.,\omega)}g = 1,\ E_{\mu(.,\omega)}|1-g| \leq \varepsilon\big\}$$

and

$$\{(\omega, g) : \ E_{\mu(.,\omega)} g|x| < \infty, \ E_{\mu(.,\omega)} gx = 0\}.$$

If the projection of $\Gamma$ on $\Omega$ is of full measure, we apply the measurable selection theorem, take an arbitrary $\mathcal{G}$-measurable selector $f : \Omega \to C(\bar{\mathbf{R}}^d)$, and conclude by putting $\alpha(\omega) = f(\omega, \eta(\omega))$.

Let $\Delta_\omega$ be the image of the convex set

$$\{g \in C(\bar{\mathbf{R}}^d) : \ g > 0, \ E_{\mu(.,\omega)} g = 1, \ E_{\mu(.,\omega)} g|x| < \infty, \ E_{\mu(.,\omega)} |1 - g| \le \varepsilon\}$$

under the linear mapping $\Psi_\omega := g \mapsto E_{\mu(.,\omega)} gx$. The full projection property means that, for almost all $\omega$, the set $\Delta_\omega$ contains the origin.

Let us first consider the case $d = 1$, where $\Delta_\omega$ is just an interval. Define the $\mathcal{G}$-measurable random variables

$$\zeta'(\omega) = \inf\{t : \ \mu(]-\infty, t], \omega) > 0\}, \qquad \zeta''(\omega) = \sup\{t : \ \mu(]-\infty, t], \omega) < 1\}.$$

The random variables $I_A I_{\{-n \le \xi'\}} \eta$, where $A \in \mathcal{G}$, being bounded from below, have negative expectations. Hence, $I_{\{-\infty < \xi'\}} E(\eta|\mathcal{G}) \le 0$. This implies that

$$I_{\{-\infty < \xi'\}} E(\eta^+|\mathcal{G}) \le I_{\{-\infty < \xi'\}} E(\eta^-|\mathcal{G}) < \infty.$$

Therefore, $\Psi(1) \le 0$ on the set $\{-\infty < \xi'\}$ and, by symmetry, $\Psi(1) \ge 0$ on the set $\{\xi'' < \infty\}$ (a.s.). Thus, on the intersection of these sets, $\Psi(1) = 0$. It follows from the elementary lemma below that the interval $\Delta_\omega \supseteq [0, \infty[$ for almost all $\omega \in \{-\infty < \xi', \ \xi'' = \infty\}$. By symmetry, the interval $\Delta_\omega \supseteq ]-\infty, 0]$ for almost all $\omega \in \{-\infty = \xi', \ \xi'' < \infty\}$. □

In the following assertion, $\omega$ is fixed and omitted in notation.

**Lemma 2.2.8** *If $\xi'' = \infty$, then $\Delta$ is unbounded from above.*

*Proof.* Fix $\varepsilon \in ]0, 1]$ and $a > 0$ such that $\mu(\{a\}) = 0$. Consider the subset $W_{\gamma,a}$ formed by the continuous functions $g$ such that $g(a) = 1$, $xg(x) \to 0$ as $x \to \pm\infty$, and $E_\mu g I_{[a,\infty[} = \varepsilon$. Note that $\sup_{g \in W_{\varepsilon,a}} E_\mu xg I_{[a,\infty[} = \infty$. Indeed, as the support of $\mu$ is unbounded, we can find a continuous function $g_0$ with a compact support contained in the interval $]a, \infty[$ such that $E\mu g_0 < \varepsilon$ while the value $E_\mu xg$ is arbitrarily large. Adding to $g_0$ the function $e^{-\lambda|x-a|}$ with an appropriately chosen parameter $\lambda$, we obtain a function $g \in W_{\varepsilon,a}$ with $E_\mu xg I_{[a,\infty[} \ge E_\mu g_0$.

Take $a > 0$ such that $\mu(\{a\}) = 0$ and $\mu(\{x : \ |x| \ge a\}) \le \delta/2$. Take $f = e^{-\lambda|x+a|}$ and choose the parameter $\lambda$ to ensure that

$$\varepsilon := \mu(\{x : \ |x| \ge a\}) - E\mu f I_{]-\infty,a]} > 0.$$

By the above, for any $N > 0$, we can find $f_N \in W_{\varepsilon,a}$ such that $E_\mu x f_N I_{[a,\infty[} \ge N$. The assertion became obvious since, for the function

$$g_N := f I_{]-\infty,-a[} + I_{]-a,a[} + g_N I_{]a,a]},$$

we have $\Psi(g_N) \ge N$. The lemma is proven. □

For the case of $d > 1$ and $\omega$ for which $0 \notin \Delta_\omega$, there is, in virtue of the Hahn–Banach theorem, $l(\omega) \in \mathbf{R}^d$ such that $|l(\omega)| = 1$ and $l(\omega)x > 0$ for all $x \in \Delta_\omega$. Put $l(\omega) = 0$ if $0 \in \Delta_\omega$. Using a measurable version of the Hahn–Banach theorem, one can choose the separating functionals in such a way that the function $\omega \mapsto l(\omega)$ is $\mathcal{G}$-measurable. Applying the above reasoning to the scalar random variable $\eta^l := l\eta$, we find a function $f^l(\omega, y)$ on $\Omega \times \mathbf{R}$ which is $\mathcal{G}$-measurable in $\omega$ and continuous in $x$. Denoting by $\mu^l(dy, .)$ the regular conditional distribution of $\eta^l$ with respect to $\mathcal{G}$, we get, by the change of variable, that

$$l(\omega)\int_{\mathbf{R}^d} x f^l(\omega, l(\omega)x)\mu(dx,\omega) = \int_{\mathbf{R}} y f^l(\omega, y)\mu^l(dy,\omega) = 0.$$

Thus, $l = 0$ (a.s.), and the required property holds.

**Remark.** Let $P$ be a probability measure under which $S$ is a local martingale, and let $H$ be a strategy such that the process $H \cdot S$ is bounded from below. Then this process is a true martingale converging at infinity to a random variable $H \cdot S_\infty$ almost surely. By the Fatou lemma, $H \cdot S_t \geq E(H \cdot S_\infty | \mathcal{F}_t)$. Therefore, for this strategy, $H \cdot S \geq 0$ if and only if $H \cdot S_\infty \geq 0$. These considerations show that there is hope to get conditions for the existence of an equivalent local martingale measure based on strategies of such a type. This is done in the next section.

In the following model the $NA$-property is fulfilled, but there is no equivalent separating measure. Namely, $R_\infty \subset L^\infty$, $R_\infty \cap L^\infty_+ = \{0\}$, but $\bar{C}^w_\infty = L^\infty$!

**Example.** Let $\Omega = \mathbf{N}$, $P(\{2k-1\}) = P(\{2k\}) = 2^{-k-1}$, $k \geq 1$, and let $\mathcal{F}_t := \sigma\{\{1\}, \dots, \{2t\}\}$. Put $S_0 := 0$,

$$\Delta S_k = 2^{5k} I_{\{2k-1\}} + 2^{2k} I_{\{2k\}} - 2^{-k} I_{\{2k+1,\dots\}}.$$

Since $\mathcal{F}_T$ is finite, the random variables $H \cdot S_T$ are bounded, and $R_\infty \subseteq L^\infty$.

Let $0 \leq \xi \leq 1$. Then $S_T \wedge \xi \in C_\infty$, and $S_T \wedge \xi \to \xi$ in probability as $T \to \infty$. Hence, $\xi \in \bar{C}^w_\infty$. It follows that $\bar{C}^w_\infty = L^\infty$.

Let $\eta \neq 0$ be a random variable from $R_\infty \cap L^\infty_+$, i.e., of the form $H \cdot S_T$. Let $k$ be the first integer for which at least one of the values $\eta(2k-1)$ or $\eta(2k)$ is strictly positive. Inspecting sequentially the increments $H_t \Delta S_t$, we deduce that $H_t = 0$ for $t < k$, while the $\mathcal{F}_{k-1}$-measurable random variable $H_k(j)$ is equal to $a > 0$ for $j \geq k-1$. It follows that $H_k(j)\Delta S_k(j) \leq -ae^{-k}$ for $j \geq 2k+1$. The negative values at elementary events $2k+1$ and $2k+2$ can be compensated only if $H_{k+1}(j) \geq a2^{-k}$ for $j \geq 2k+1$. Continuing this inspection, we arrive at the last increment $H_T \Delta S_T$ the negative values of which on elementary events $2T+1, \dots$ cannot be compensated.

More surprisingly, in this example the closure of $R_\infty$ in the $L^1$-norm intersects $L^\infty_+$ only at zero. This can be shown by a similar sequential inspection of $\lim_n H^n_t \Delta S_t$. To ensure the positivity of $\eta$, these random variables should take such large positive values at the elementary events with odd numbers larger than $k$ that the $L^1$-norm of $\eta$ would be infinite in an apparent contradiction.

### 2.2.3 No Free Lunch with Vanishing Risk

The *NFL*-condition can be criticized because the weak* closure has no good financial interpretation.[1] Fortunately, it can be replaced by the more attractive *NFLVR*-property.

To describe the latter we introduce the class of *admissible* strategies $H$ whose value processes $H \cdot S$ are bounded from below (by constants depending on $H$) and converge a.s. to finite limits. Denoting by $R_{\text{ad}}$ the set of random variables $H \cdot S_\infty$, we define the sets $A_{\text{ad}} := R_{\text{ad}} - L^0_+$ and $C_{\text{ad}} := A_{\text{ad}} \cap L^\infty$.

We say that the process $S$ has the *NFLVR-property* (no free lunch with vanishing risk) if $\bar{C}_{\text{ad}} \cap L^\infty = \{0\}$, where $\bar{C}_{\text{ad}}$ is the norm-closure of $C_{\text{ad}}$. "Financial" motivation of the terminology is based on the alternative description: *NFLVR-property* holds if and only if $P$-$\lim \xi_n = 0$ for every sequence $\xi_n \in C_{\text{ad}}$ such that $\|\xi_n^-\|_{L^\infty} \to 0$, see Lemma 2.2.11.

Though the sets $A$ and $A_{\text{ad}}$ may be not related by an inclusion, the property $A_{\text{ad}} \cap L^0_+ = \{0\}$ ensures the property $A_\infty \cap L^0_+ = \{0\}$. Indeed, the former implies that, for any finite $T$, there is no arbitrage in the class of strategies with the value processes $(H \cdot S_t)_{t \le T}$ bounded from below. As we know, this is equivalent to the absence of arbitrage in the class of all strategies and, hence, to the existence of an equivalent martingale measure on $\mathcal{F}_T$. It follows that the property $A_{\text{ad}} \cap L^0_+ = \{0\}$ implies that the bound $H \cdot S_T \le c$ propagates backwards and $C_\infty \subseteq C_{\text{ad}}$.

**Theorem 2.2.9** *NFLVR holds if and only if there is $P' \sim P$ such that $S \in \mathcal{M}_{\text{loc}}(P')$.*

*Proof.* It is easy to see (using the Fatou lemma) that a local martingale measure (and even a separating measure for $R_\infty$) separates $\bar{C}_{\text{ad}}$ and $L^\infty_+$. So, the implication "if" is obvious. On the other hand, the condition $\bar{C}^w_{\text{ad}} \cap L^\infty_+ = \{0\}$, ensuring that $C_\infty \subseteq C_{\text{ad}}$, implies the *NFL*-property, and the needed measure $P'$ does exist in virtue of Theorem 2.2.5. But according to Theorem 2.2.10 below, such a condition holds because under *NFLVR* the set $\bar{C}^w_{\text{ad}}$ coincides with $\bar{C}_{\text{ad}}$. □

**Theorem 2.2.10** *Suppose that $\bar{C}_{\text{ad}} \cap L^\infty = \{0\}$. Then $C_{\text{ad}} = \bar{C}^w_{\text{ad}}$.*

Before the proof we establish some simple facts from functional analysis. Let $\lfloor \eta, \infty \lfloor$ be the set of $\xi \in L^0$ such that $\xi \ge \eta$.

**Lemma 2.2.11** *Let $C$ be a convex cone in $L^\infty$ containing $-L^\infty_+$. Then the following properties are equivalent:*

[1] Of course, the definition of weak* closure involving only halfspaces is even simpler than that of the norm closure. The intuition, though, appeals to the "interior" description, in terms of limits. In general, the weak* sequential closure lies strictly between the norm-closure and weak* closure. To get all points of the latter as limits, one needs to consider convergence along the nets, which is, indeed, not intuitive.

(a) $\bar{C} \cap L^\infty_+ = \{0\}$;
(b) $P$-$\lim \xi_n = 0$ *for every sequence* $\xi_n \in C$ *such that* $\|\xi_n^-\|_{L^\infty} \to 0$;
(c) *the set* $C \cap \lfloor -1, \infty \lfloor$ *is bounded in probability.*

*Proof.* (a) $\Rightarrow$ (b). If the assertion fails, one can find a sequence $\xi_n \in C$ such that $\xi_n \geq -1/n$ and $P(\xi_n > \varepsilon) \geq \varepsilon$ for some $\varepsilon > 0$. Since $\xi_n \wedge 1 \in C$, we may assume that $\xi_n \leq 1$. By the von Weizsäcker theorem there are random variables of the form $\bar{\xi}_k = k^{-1} \sum_{i=1}^k \xi_{n_i}$ (thus, elements of $C$) convergent to a certain random variable $\xi$ a.s. Note that the negative parts of $\bar{\xi}_k$ converge to zero in $L^\infty$. On the other hand, $\xi$ is not zero. Indeed, $\xi$ is also the limit of $k^{-1} \sum_{i=1}^k \tilde{\xi}_{n_i}$, where $\tilde{\xi}_n := \xi_n + 1/n \geq 0$ and $P(\tilde{\xi}_n > \varepsilon) \geq \varepsilon$. It is easy to see that

$$Ee^{-\tilde{\xi}_n} \leq P(\tilde{\xi}_n \leq \varepsilon) + e^{-\varepsilon} P(\tilde{\xi}_n > \varepsilon) \leq 1 - \varepsilon + e^{-\varepsilon}\varepsilon < 1.$$

Due to convexity of the exponential, the same bound holds for the convex combinations of $\tilde{\xi}_n$ and, thus, for the limit $\xi$. So, $\beta := P(\xi > 0) > 0$. By the Egorov theorem, there is a measurable set $\Gamma$ with $P(\Gamma) > 1 - \beta/2$ on which the convergence $\bar{\xi}_n \to \xi$ is uniform. But then the sequence $\bar{\xi}_n^+ I_\Gamma - \bar{\xi}_n^-$ of elements of $C$ converges in $L^\infty$ to a nonzero random variable $\xi I_\Gamma \geq 0$, in contradiction with (a).

(b) $\Rightarrow$ (c). If the set $C \cap \lfloor 1, \infty \lfloor$ is unbounded in probability, then it contains a sequence of random variables $\xi_n^0 \geq 1$ such that $\lim P(\xi_n^0 \geq n) > 0$. But then the sequence $\xi_n := \xi_n^0/n$ violates condition (b).

(c) $\Rightarrow$ (a). If (a) fails to be true, there exist a sequence $\xi_n \in C$ and a nonzero $\xi \in L^\infty_+$ such that $\|\xi - \xi_n\|_{L^\infty} \leq 1/n$. It follows that $\|\xi_n^-\|_{L^\infty} \leq 1/n$. Then the random variables $n\xi_n$ belong to $C \cap \lfloor -1, \infty \lfloor$ and form a sequence divergent to infinity on the set $\{\xi > 0\}$ and, therefore, not bounded in probability. $\square$

The next lemma, comparatively with the previous one, requires a specific structure of the cone $C$. We use the notation $\bar{K}^P$ for the closure of $K$ in $L^0$.

**Lemma 2.2.12** *Let* $C = (K - L^0_+) \cap L^\infty$ *where* $K$ *is a cone,* $K \subseteq \lfloor -1, \infty \lfloor$. *Suppose that* $K$ *is bounded in probability. Let* $\xi_n$ *be a sequence in* $C \cap \lfloor -1, \infty \lfloor$ *convergent to* $\xi$ *a.s. Then the set* $\bar{K}^P \cap \lfloor \xi, \infty \lfloor$ *is nonempty and contains a maximal element* $\eta_0$.

*Proof.* In virtue of the assumed structure of the set $C$, there are $\eta_n \in K$ such that $\eta_n \geq \xi_n$. Applying the von Weizsäcker theorem, we find a subsequence such that $\bar{\eta}_k := k^{-1} \sum_{i=1}^k \eta_{n_i}$ converge a.s. to some $\bar{\eta} \geq \xi$. Since $K$ is bounded in probability, so is the set $\bar{K}^P$. Thus, $\bar{\eta}$ is finite and belongs to $\bar{K}^P \cap \lfloor \xi, \infty \lfloor \neq \emptyset$. It remains to recall that any nonempty closed bounded subset of $L^0$ has a maximal element with respect to the natural partial ordering (each linearly ordered subset $\{\zeta_\alpha\}$ has as a majorant $\operatorname{ess\,sup}_\alpha \zeta_\alpha < \infty$, and the existence of the maximal element holds by the Zorn lemma). $\square$

**Lemma 2.2.13** *Let* $C_{\rm ad} \cap L_+ = \{0\}$. *If* $H$ *is an admissible integrand, then* $H \cdot S_\infty \geq -1$ *if and only if the process* $H \cdot S \geq -1$.

*Proof.* Suppose that $H$ is admissible and $H \cdot S_\infty \ge -1$ but there is $u$ such that $P(\Gamma_u) > 0$, where $\Gamma_u := \{H \cdot S_u < -1\}$. Then the strategy $HI_{[u,\infty[}I_{\Gamma_u}$ is admissible, and the random variable $HI_{]u,\infty[}I_{\Gamma_u} \cdot S_\infty \ge 0$ is strictly positive on $\Gamma_u$. This is a contradiction with the assumption of the lemma. □

*Proof of Theorem 2.2.10.* According to the Krein–Šmulian theorem, a convex set is closed in $\sigma\{L^\infty, L^1\}$ if an only if its intersection with every ball of $L^\infty$ is closed in probability. Obviously, the last condition follows if the set is *Fatou-closed*, that is, if it contains the limit of any bounded from below sequence of its elements convergent almost surely. So, let $\xi_n$ be a sequence in $C_{\text{ad}}$ convergent to $\xi$ a.s. and such that all $\xi_n \ge -c$. It is sufficient to argue with $c = 1$. We apply Lemma 2.2.12 with $K = R_{\text{ad}} \cap \lfloor -1, \infty \lfloor$, which is bounded in probability by virtue of Lemma 2.2.11. The theorem will be proven if we show that a maximal element $\eta_0$ in $\bar{K}^P \cap \lfloor \xi, \infty \lfloor \neq \emptyset$ belongs to $K$. So, we have a sequence $V^n := H^n \cdot S \ge -1$ with $V^n_\infty \to \eta$ a.s. We claim that $\sup_t |V^n_t - V^m_t| \to 0$ in probability as $n, m \to \infty$. If this not true, then $P((\sup_t(V^{i_k}_t - V^{j_k}_t)^+ > \varepsilon) \ge \varepsilon$ with some $\varepsilon > 0$ and $i_k, j_k \to \infty$. For $T_k := \inf\{t : \ V^{i_k}_t - V^{j_k}_t > \alpha\}$, we have $P(T_k < \infty) \ge \varepsilon$. Let us consider the process

$$\tilde{V}^k := \left(I_{[0,T_k]} H^{i_k} + I_{]T_k,T]} H^{j_k}\right) \cdot S,$$

which is an element of $K = R_{\text{ad}} \cap \lfloor -1, \infty \lfloor$. Note that

$$\tilde{V}^k_\infty = V^{i_k}_\infty I_{\{T_k = \infty\}} + V^{j_k}_\infty I_{\{T_k < \infty\}} + \xi_k,$$

where $\xi_k := (V^{i_k}_{T_k} - V^{j_k}_{T_k}) I_{\{T_k < \infty\}} \ge 0$, and $P(\xi_k \ge \varepsilon) \ge \varepsilon$. Using the von Weizsäcker theorem in the same way as in the proof of Lemma 2.2.11, we find a sequence $\bar{V}^k \in K$ such that $\bar{V}^k_\infty \to \eta_0 + \xi$, where $\xi \in L^0$ and $\xi \neq 0$. This contradicts the maximality of $\eta_0$.

Taking a subsequence, we may assume that $\sup_t |V^n_t - V^m_t| \to 0$ a.s. Thus, there is a process $V$ which is a uniform limit of $V^n$ (a.s.). Obviously, $V \ge -1$, and the limit $V_\infty$ exists and is finite. Since $\Delta V^n_t = H^n_t \Delta S_t$ converges to $\Delta V_t$ and $R_T$ is closed, we have $\Delta V_t = H_t \Delta S_t$. □

### 2.2.4 Example: "Retiring" Process

Here we present an example where a martingale measure can be constructed in a rather straightforward way. We shall use the result later, in the study of models with transaction costs.

Let $S = (S_t)_{t \ge 0}$ be an $\mathbf{R}^d$-valued discrete-time adapted process. Put $\xi_t = \Delta S_t$, $\Gamma_t := \{\xi_t = 0\}$.

**Proposition 2.2.14** *Suppose that the following conditions hold:*

(i) *for each finite $T$, the process $(S_t)_{t \le T}$ has the NA-property;*
(ii) $I_{\Gamma_t} \uparrow 1$ *a.s.;*
(iii) $E(I_{\Gamma_t} | \mathcal{F}_{t-1}) > 0$ *a.s. on $\Gamma^c_{t-1}$ for each $t \ge 1$.*

*Then there exists a probability $Q \sim P$ such that $S$ is a $Q$-martingale bounded in $L^2(Q)$ (hence, uniformly integrable with respect to $Q$).*

*Proof.* By the DMW theorem condition (i) is equivalent to the $NA$-property for each one-step model: the relation $\gamma\xi_t \geq 0$ with $\gamma \in L^0(\mathbf{R}^d, \mathcal{F}_{t-1})$ may hold only if $\gamma\xi_t = 0$. The same theorem asserts that each $\xi_t$ admits an equivalent martingale measure which can be chosen to ensure the integrability of any fixed finite random variable, e.g., $|\xi_t|^2$. In terms of densities this means that there are $\mathcal{F}_t$-measurable random variables $\bar{\alpha}_t > 0$ such that $E(\bar{\alpha}_t\xi_t|\mathcal{F}_{t-1}) = 0$ and $c_t := E(\bar{\alpha}_t|\xi_t|^2|\mathcal{F}_{t-1}) < \infty$. Normalizing, to this we can also add the property $E(\bar{\alpha}_t|\mathcal{F}_{t-1}) = 1$.

We define the $\mathcal{F}_t$-measurable random variable $\alpha_t > 0$ by the formula

$$\alpha_t = I_{\Gamma_{t-1}} + \left[\frac{(1-\delta_t)I_{\Gamma_t}}{E(I_{\Gamma_t}|\mathcal{F}_{t-1})} + \frac{\delta_t\bar{\alpha}_t I_{\Gamma_t^c}}{E(\bar{\alpha}_t I_{\Gamma_t^c}|\mathcal{F}_{t-1})}\right] I_{\Gamma_{t-1}^c \cap A_t} + I_{\Gamma_{t-1}^c \cap A_t^c},$$

where $A_t := \{E(\bar{\alpha}_t I_{\Gamma_t^c}|\mathcal{F}_{t-1}) > 0\}$ and $\delta_t := 2^{-t}E(\bar{\alpha}_t I_{\Gamma_t^c}|\mathcal{F}_{t-1})/(1+c_t)$. Clearly, $E(\alpha_t|\mathcal{F}_{t-1}) = 1$.

Noting that $\bar{\alpha}_t I_{\Gamma_t^c} I_{A_t^c} = 0$ (a.s.), we obtain that $E(\alpha_t\xi_t^2|\mathcal{F}_{t-1}) \leq 2^{-t}$ and $E(\alpha_t\xi_t|\mathcal{F}_{t-1}) = 0$.

The process $Z_t := \alpha_1 \dots \alpha_t$ is a martingale which converges (stationarily) a.s. to a random variable $Z_\infty > 0$ with $EZ_\infty \leq 1$. Recalling that $I_{\Gamma_t} \uparrow 1$ (a.s.) and using the identity $Z_\infty I_{\Gamma_t} = Z_t I_{\Gamma_t}$, we obtain that

$$EZ_\infty = E\lim_t Z_\infty I_{\Gamma_t} = \lim_t EZ_\infty I_{\Gamma_t} = \lim_t EZ_t I_{\Gamma_t} = 1 - \lim_t EZ_t I_{\Gamma_t^c}.$$

It follows that $EZ_\infty = 1$ (i.e., $(Z_t)$ is a uniformly integrable martingale). Indeed, $E(\alpha_k I_{\Gamma_k^c}|\mathcal{F}_{k-1} \leq 2^{-k}$, and, hence,

$$EI_{\Gamma_t^c} Z_t = E\prod_{k\leq t}\alpha_k I_{\Gamma_k^c} \leq \prod_{k\leq t} 2^{-k} \to 0.$$

Thus, $Q := Z_\infty P$ is a probability measure under which $S$ is a martingale. At last,

$$E_Q S_t^2 = \sum_{k\leq t} EZ_k\xi_k^2 \leq \sum_{k\leq t} 2^{-k} \leq 1,$$

i.e., $S_t$ belongs to the unit ball of $L^2(Q)$. □

**Remark 1.** Condition (iii) cannot be omitted. Indeed, let $S$ be the symmetric random walk starting from zero and stopped at the moment when it hits unit. It is a martingale, and condition (ii) holds. Since $S_\infty = 1$ a.s., the process $S$ cannot be a uniformly integrable martingale with respect to a measure $Q$ equivalent to $P$.

**Remark 2.** Fix $f : \mathbf{R} \to \mathbf{R}_+$. A minor modification of the arguments leads to a martingale measure $Q$ for which $E_Q \sup_t f(S_t) < \infty$. Indeed, let $(\eta_t)$ be an adapted process with $\eta_t = \eta_t I_{\Gamma_t^c} \geq 0$. As above, we can find $\alpha_t$ with the extra property $E(\alpha_t f(S_t)|\mathcal{F}_{t-1}) \leq 2^{-t}$ implying that $E\sum_t \eta_t < \infty$. It remains to take $\eta_t = f(S_t)I_{\Gamma_t^c}$ and note that $\sup_t f(S_t) \leq \sum_t \eta_t$.

### 2.2.5 The Delbaen–Schachemayer Theory in Continuous Time

This book is addressed to the reader from whom we do not expect the knowledge of stochastic calculus beyond standard textbooks. Luckily, the theory of markets with transaction costs, in current state of art, does not require such a knowledge, in a surprising contrast to the classical continuous-time *NA*-theory initiated by Kreps and largely developed in a series of papers by Delbaen and Schachermayer collected in [57]. However, it seems to be useful to provide a short abstract of the main results of the latter, which will serve as a background for a discussion explaining this difference.

In the classical continuous-time theory we are given a set $\mathcal{X}$ of scalar semimartingales $X$ on a compact interval $[0,T]$ interpreted as value processes; the elements of $R_T := \{X_T : \ X \in \mathcal{X}\}$ are the investor's "results"; the *NA*-property means that $R_T \cap L^0_+ = \{0\}$. Typically, $\mathcal{X}$ is the set of stochastic integrals $H \cdot S$, where $S$ is a fixed $d$-dimensional semimartingale (interpreted as the price processes of risky assets), and $H$ is a $d$-dimensional predictable process for which the integral is defined and is bounded from below by a constant depending on $H$. The condition on $H$ ("admissibility") rules out the doubling strategies. The experience with discrete-time models gives a hint that martingale densities can be obtained by a suitable separation theorem. Put $C_T := (R_T - L^0_+) \cap L^\infty$ (the set of bounded contingent claims hedgeable from zero initial endowment) and introduce the "no-free-lunch condition" (*NFL*): $\bar{C}^w_T \cap L^\infty_+ = \{0\}$, where $\bar{C}^w_T$ is a closure of $C_T$ in the weak* topology, i.e., $\sigma\{L^\infty, L^1\}$. The Kreps–Yan theorem (Theorem 2.1.4) says that *NFL* holds if and only if there exists an equivalent "separating" measure $P' \sim P$ such that $E'\xi \le 0$ for all $\xi$ from $\bar{C}^w_T$ (or $R_T$). It is easy to see that in the model with a bounded (resp., locally bounded) price process $S$, the latter is a martingale (resp. local martingale).

The above result established by Kreps in the context of financial modelling ("FTAP") was completed by Delbaen and Schachermayer by a number of important observations for the model based on the price process $S$. We indicate here only a few.

First, they observed that in the Kreps theorem the condition *NFL* can be replaced by a visibly weaker (but, in fact, equivalent) condition "no-free-lunch condition with vanishing risk" (*NFLVR*): $\bar{C}_T \cap L^\infty_+ = \{0\}$, where $\bar{C}_T$ is the norm-closure of $C_T$ in $L^\infty$. The reason for this is in the following simply formulated (but difficult to prove) result from stochastic calculus:

**Theorem 2.2.15** *Let the NFLVR-condition be fulfilled. Then* $C_T = \bar{C}^w_T$.

This result, which is a generalization of Theorem 2.2.9, can be formulated in a more abstract way for a convex set $\mathcal{X}$ of bounded from below semimartingales which satisfies some closedness and concatenation properties.

Second, they establish that in any neighborhood of a separating measure $P'$, there exists an equivalent probability measure $\tilde{P}$ (also a separating one) such that the semimartingale $S$ with respect to $\tilde{P}$ is a $\sigma$-martingale (i.e., for some predictable integrants $G^i$ with values in $]0,1]$, the processes $G^i \cdot S^i$, $i = 1, \dots, d$, are $\tilde{P}$-martingales). The situation for the continuous time is rather different even with respect to infinite-horizon discrete-time models: one cannot claim the existence of an equivalent local martingale measure! The reason for this is clear: in discrete time there is no difference between local martingales and $\sigma$-martingales (which are just generalized martingales).

As we shall see further, for the model with transaction costs, the portfolio processes are vector-valued, and their dynamics can be described using only the Lebesgue integrals. In the case of zero transaction cost, one can make a reduction to scalar wealth processes $H \cdot S$, but the resulting $H$ are (vector-valued) processes of *bounded variation* and not arbitrary integrands, which is, apparently, an additional complication. In the general case the problem of no-arbitrage criteria has also other particularities arising even in the discrete-time framework.