B + K

3. Stochastic Processes in Discrete Time

3.1 Information and Filtrations

Access to full, accurate, up-to-date information is clearly essential to anyone actively engaged in financial activity or trading. Indeed, information is arguably the most important determinant of success in financial life. Partly for simplicity, partly to reflect the legislation and regulations against insider trading, we shall confine ourselves to the situation where agents take decisions on the basis of information in the public domain, and available to all. We shall further assume that information once known remains known – is not forgotten – and can be accessed in real time.

In reality, of course, matters are more complicated. Information overload is as much of a danger as information scarcity. The ability to retain information, organise it, and access it quickly, is one of the main factors which will discriminate between the abilities of different economic agents to react to changing market conditions. However, we restrict ourselves here to the simplest possible situation and do not differentiate between agents on the basis of their information processing abilities. Thus as time passes, new information becomes available to all agents, who continually update their information. What we need is a mathematical language to model this information flow, unfolding with time. This is provided by the idea of a *filtration*; we outline below the elements of this theory that we shall need.

The Kolmogorov triples $(\Omega, \mathcal{F}, P)$, and the Kolmogorov conditional expectations $E(X|\mathcal{B})$, give us all the machinery we need to handle static situations involving randomness. To handle dynamic situations, involving randomness which unfolds with time, we need further structure.

We may take the initial, or starting, time as $t = 0$. Time may evolve discretely, or continuously. We postpone the continuous case to Chapter 5; in the discrete case, we may suppose time evolves in integer steps, $t = 0, 1, 2, \dots$ (say, stock-market quotations daily, or tick data by the second). There may be a final time T, or time horizon, or we may have an infinite time horizon (in the context of option pricing, the time horizon T is the expiry time).

We wish to model a situation involving randomness unfolding with time. As above, we suppose, for simplicity, that information is never lost (or forgotten): thus, as time increases we learn more. We recall from Chapter 2 that

σ-algebras represent information or knowledge. We thus need a sequence of σ-algebras $\{\mathcal{F}_n : n = 0, 1, 2, \dots\}$, which are increasing:

$$\mathcal{F}_n \subset \mathcal{F}_{n+1} \quad (n = 0, 1, 2, \dots),$$

with $\mathcal{F}_n$ representing the information, or knowledge, available to us at time n. We shall always suppose all σ-algebras to be *complete* (this can be avoided, and is not always appropriate, but it simplifies matters and suffices for our purposes). Thus $\mathcal{F}_0$ represents the initial information (if there is none, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, the trivial σ-algebra). On the other hand,

$$\mathcal{F}_\infty := \lim_{n\to\infty} \mathcal{F}_n$$

represents all we ever will know (the 'Doomsday σ-algebra'). Often, $\mathcal{F}_\infty$ will be $\mathcal{F}$ (the σ-algebra from Chapter 2, representing 'knowing everything'. But this will not always be so; see e.g. [218], §15.8 for an interesting example. Such a family $\{\mathcal{F}_n : n = 0, 1, 2, \dots\}$ is called a *filtration*; a probability space endowed with such a filtration, $\{\Omega, \{\mathcal{F}_n\}, \mathcal{F}, P\}$ is called a *filtered probability space*. These definitions are due to P. A. Meyer of Strasbourg; Meyer and the Strasbourg (and more generally, French) school of probabilists have been responsible for the 'general theory of (stochastic) processes', and for much of the progress in stochastic integration, since the 1960s; see e.g. [59, 60, 163, 164].

For the special case of a finite state space $\Omega = \{\omega_1, \dots, \omega_n\}$ and a given σ-algebra $\mathcal{F}$ on Ω (which in this case is just an algebra) we can always find a unique finite partition $\mathcal{P} = \{A_1, \dots, A_l\}$ of Ω, i.e. the sets A_i are disjoint and $\bigcup_{i=1}^{l} A_i = \Omega$, corresponding to $\mathcal{F}$. A filtration $\mathcal{F}_n$ therefore corresponds to a sequence of finer and finer partitions $\mathcal{P}_n$. At time $t = 0$ the agents only know that some event $\omega \in \Omega$ will happen, at time $T < \infty$ they know which specific event ω^* has happened. During the flow of time the agents learn the specific structure of the (σ-) algebras $\mathcal{F}_n$, which means they learn the corresponding partitions $\mathcal{P}$. Having the information in $\mathcal{F}_n$ revealed is equivalent to knowing in which $A_i^{(n)} \in \mathcal{P}_n$ the event ω^* is. Since the partitions become finer the information on ω^* becomes more detailed with each step.

Unfortunately this nice interpretation breaks down as soon as Ω becomes infinite. It turns out that the concept of filtrations rather than that of partitions is relevant for the more general situations of infinite Ω, infinite T and continuous-time processes.

3.2 Discrete-Parameter Stochastic Processes

The word 'stochastic' (derived from the Greek) is roughly synonymous with 'random'. It is perhaps unfortunate that usage favours 'stochastic process' rather than the simpler 'random process', but as it does, we shall follow it.

We need a framework which can handle dynamic situations, in which time evolves, and in which new information unfolds with time. In particular, we need to be able to speak in terms of 'the information available at time n', or, 'what we know at time n'. Further, we need to be able to increase n – thereby increasing the information available as new information (typically, new price information) comes in, and talk about the information flow over time. One has a clear mental picture of what is meant by this – there is no conceptual difficulty. However, what is needed is a precise mathematical construct, which can be conveniently manipulated - perhaps in quite complicated ways – and yet which bears the above heuristic meaning. Now 'information' is not only an ordinary word, but even a technical term in mathematics – many books have been written on the subject of information theory. However, information theory in this sense is not what we need: for us, the emphasis is on the flow of information, and how to model and describe it. With this by way of motivation, we proceed to give some of the necessary definitions.

A *stochastic process* $X = \{X_t : t \in I\}$ is a family of random variables, defined on some common probability space, indexed by an index-set I. Usually (always in this book), I represents time (sometimes I represents space, and one calls X a spatial process). Here, $I = \{0, 1, 2, \dots, T\}$ (finite horizon) or $I = \{0, 1, 2, \dots\}$ (infinite horizon). The (stochastic) process $X = (X_n)_{n=0}^{\infty}$ is said to be *adapted* to the filtration $(\mathcal{F}_n)_{n=0}^{\infty}$ if

$$X_n \text{ is } \mathcal{F}_n - \text{measurable.}$$

So if X is adapted, we will know the value of X_n at time n. If

$$\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$$

we call $(\mathcal{F}_n)$ the natural filtration of X. Thus a process is always adapted to its natural filtration. A typical situation is that

$$\mathcal{F}_n = \sigma(W_0, W_1, \dots, W_n)$$

is the natural filtration of some process $W = (W_n)$. Then X is adapted to $(\mathcal{F}_n)$, i.e. each X_n is $\mathcal{F}_n$- (or $\sigma(W_0, \cdots, W_n)$-) measurable, iff

$$X_n = f_n(W_0, W_1, \dots, W_n)$$

for some measurable function f_n (non-random) of $n + 1$ variables.

Notation. For a random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$, $X(\omega)$ is the value X takes on ω (ω represents the randomness). For a stochastic process $X = (X_n)$, it is convenient (e.g., if using suffixes, n_i say) to use X_n, $X(n)$ interchangeably, and we shall feel free to do this. With ω displayed, these become $X_n(\omega)$, $X(n, \omega)$, etc.

The concept of a stochastic process is very general – and so very flexible – but it is too general for useful progress to be made without specifying further structure or further restrictions. There are two main types of stochastic process which are both general enough to be sufficiently flexible to model many commonly encountered situations, and sufficiently specific and structured to have a rich and powerful theory. These two types are Markov processes and martingales. A *Markov process* models a situation in which where one is, is all one needs to know when wishing to predict the future – how one got there provides no further information. Such a 'lack of memory' property, though an idealisation of reality, is very useful for modelling purposes. We shall encounter Markov processes more in continuous time (see Chapter 5) than in discrete time, where usage dictates that they are called Markov chains. For an excellent and accessible recent treatment of Markov chains, see e.g. [171]. *Martingales*, on the other hand (see §3.4 below) model fair gambling games – situations where there may be lots of randomness (or unpredictability), but no tendency to drift one way or another: rather, there is a tendency towards stability, in that the chance influences tend to cancel each other out on average.

3.3 Discrete-Parameter Martingales

Excellent accounts of discrete-parameter martingales are Neveu [169] and Williams [218] to which we refer the reader for detailed discussions. We will summarise what we need to use martingales for modelling in finance.

3.3.1 Definition and Simple Properties

Definition 3.3.1. *A process $X = (X_n)$ is called a martingale relative to $((\mathcal{F}_n), \mathbb{P})$ if*

(i) X is adapted (to $(\mathcal{F}_n)$);
(ii) $\mathbb{E}\,|X_n| < \infty$ for all n;
(iii) $\mathbb{E}[X_n|\mathcal{F}_{n-1}] = X_{n-1}$ $\mathbb{P}-a.s.$ $(n \geq 1)$.

X is a supermartingale if in place of (iii)

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] \leq X_{n-1} \quad \mathbb{P}-a.s. \quad (n \geq 1);$$

X is a submartingale if in place of (iii)

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] \geq X_{n-1} \quad \mathbb{P}-a.s. \quad (n \geq 1).$$

Using (iii) we see that the best forecast of unobserved future values of (X_k) based on information at time $\mathcal{F}_n$ is X_n; in more mathematical terms, the $\mathcal{F}_n$ measurable random variable Y which minimises $\mathbb{E}((X_{n+1} - Y)^2|\mathcal{F}_n)$ is X_n.

Martingales also have a useful interpretation in terms of dynamic games: a martingale is 'constant on average', and models a fair game; a supermartingale is 'decreasing on average', and models an unfavourable game; a submartingale is 'increasing on average', and models a favourable game.

Note. 1. Martingales have many connections with harmonic functions in probabilistic potential theory. The terminology in the inequalities above comes from this: supermartingales correspond to superharmonic functions, submartingales to subharmonic functions.
2. X is a submartingale (supermartingale) if and only if $-X$ is a supermartingale (submartingale); X is a martingale if and only if it is both a submartingale and a supermartingale.
3. (X_n) is a martingale if and only if $(X_n - X_0)$ is a martingale. So we may without loss of generality take $X_0 = 0$ when convenient.
4. If X is a martingale, then for $m < n$ using the iterated conditional expectation and the martingale property repeatedly (all equalities are in the a.s.-sense)

$$\mathbb{E}[X_n|\mathcal{F}_m] = \mathbb{E}[\mathbb{E}(X_n|\mathcal{F}_{n-1})|\mathcal{F}_m] = \mathbb{E}[X_{n-1}|\mathcal{F}_m]$$
$$= \ldots = \mathbb{E}[X_m|\mathcal{F}_m] = X_m,$$

and similarly for submartingales, supermartingales.
5. Examples of a martingale include: sums of independent, integrable zero-mean random variables (submartingales: positive mean; supermartingale: negative mean).

From the *Oxford English Dictionary: martingale* (etymology unknown)

1. 1589. An article of harness, to control a horse's head.
2. Naut. A rope for guying down the jib-boom to the dolphin-striker.
3. A system of gambling which consists in doubling the stake when losing in order to recoup oneself (1815).
Thackeray: 'You have not played as yet? Do not do so; above all avoid a martingale if you do.'

Gambling games have been studied since time immemorial – indeed, the Pascal-Fermat correspondence of 1654 which started the subject was on a problem (de Méré's problem) related to gambling. The doubling strategy above has been known at least since 1815.

The term 'martingale' in our sense is due to J. Ville (1939). Martingales were studied by Paul Lévy (1886–1971) from 1934 on (see obituary [155]) and by J.L. Doob (1911–) from 1940 on. The first systematic exposition was [62]. This classic book, though hard going, is still a valuable source of information.

Example. Accumulating data about a random variable ([218], pp. 96, 166–167). If $\xi \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, $M_n := \mathbb{E}(\xi|\mathcal{F}_n)$ (so M_n represents our best estimate of ξ based on knowledge at time n), then using iterated conditional expectations

$$\mathbb{E}[M_n|\mathcal{F}_{n-1}] = \mathbb{E}[\mathbb{E}(\xi|\mathcal{F}_n)|\mathcal{F}_{n-1}] = \mathbb{E}[\xi|\mathcal{F}_{n-1}] = M_{n-1},$$

so (M_n) is a martingale. One has the convergence

$$M_n \to M_\infty := \mathbb{E}[\xi|\mathcal{F}_\infty] \quad a.s. \quad \text{and in } L^1.$$

3.3.2 Martingale Convergence

We turn now to the theorems that make martingales so powerful a tool.

A supermartingale is 'decreasing on average'. Recall that a decreasing sequence (of real numbers) that is bounded below converges (decreases to its greatest lower bound or infimum). This suggests that a supermartingale which is bounded below converges a.s.. This is so (Doob's forward convergence theorem: [218], §§11.5, 11.7).

More is true. Call X L^1-bounded if

$$\sup_n \mathbb{E}\,|X_n| < \infty.$$

Theorem 3.3.1 (Doob). *An L^1-bounded supermartingale is a.s. convergent: there exists X_∞ finite such that*

$$X_n \to X_\infty \quad (n \to \infty) \quad a.s.$$

In particular, we have ([218], §11.5):

Theorem 3.3.2 (Doob's Martingale Convergence Theorem). *An L^1-bounded martingale converges a.s..*

We say that

$$X_n \to X_\infty \quad \text{in } L^1$$

if

$$\mathbb{E}\,|X_n - X_\infty| \to 0 \quad (n \to \infty).$$

For a class of martingales, one gets convergence in L^1 as well as almost surely ([169], IV.2, [218], Chapter 14).

Theorem 3.3.3. *The following are equivalent for martingales $X = (X_n)$:*

(i) X_n converges in L^1;
(ii) X_n is L^1-bounded, and its a.s. limit X_∞ (which exists, by above) satisfies

$$X_n = \mathbb{E}[X_\infty|\mathcal{F}_n];$$

(iii) There exists an integrable random variable X with

$$X_n = \mathbb{E}[X|\mathcal{F}_n].$$

Such martingales are called regular [169] or *uniformly integrable* [218].

3.3.3 Doob Decomposition

Theorem 3.3.4. *Let $X = (X_n)$ be an adapted process with each $X_n \in L^1$. Then X has an (essentially unique) Doob decomposition*

$$X = X_0 + M + A : \quad X_n = X_0 + M_n + A_n \quad \forall n \tag{3.1}$$

with M a martingale null at zero, A a predictable process null at zero. If also X is a submartingale ('increasing on average'), A is increasing: $A_n \le A_{n+1}$ for all n, a.s.

Proof. If X has a Doob decomposition (3.1),

$$\mathbb{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}] = \mathbb{E}[M_n - M_{n-1}|\mathcal{F}_{n-1}] + \mathbb{E}[A_n - A_{n-1}|\mathcal{F}_{n-1}].$$

The first term on the right is zero, as M is a martingale. The second is $A_n - A_{n-1}$, since A_n (and A_{n-1}) is $\mathcal{F}_{n-1}$-measurable by previsibility. So

$$\mathbb{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}] = A_n - A_{n-1}, \tag{3.2}$$

and summation gives

$$A_n = \sum_{k=1}^{n} \mathbb{E}[X_k - X_{k-1}|\mathcal{F}_{k-1}], \quad a.s.$$

We use this formula to *define* (A_n), clearly previsible. We then use (3.1) to *define* (M_n), then a martingale, giving the Doob decomposition (3.1).

If X is a submartingale, the LHS of (3.2) is ≥ 0, so the RHS of (3.2) is ≥ 0, i.e. (A_n) is increasing. □

Although the Doob decomposition is a simple result in discrete time, the analogue in continuous time – the Doob-Meyer decomposition – is deep (see Chapter 5). This illustrates the contrasts that may arise between the theories of stochastic processes in discrete and continuous time.

3.4 Martingale Transforms

Now think of a gambling game, or series of speculative investments, in discrete time. There is no play at time 0; there are plays at times $n = 1, 2, \ldots$, and

$$\Delta X_n := X_n - X_{n-1}$$

represents our net winnings per unit stake at play n. Thus if X_n is a martingale, the game is 'fair on average'.

Call a process $C = (C_n)_{n=1}^{\infty}$ *predictable* (or *previsible*) if

$$C_n \text{ is } \mathcal{F}_{n-1} - \text{measurable for all } n \geq 1.$$

Think of C_n as your stake on play n (C_0 is not defined, as there is no play at time 0). Previsibility says that you have to decide how much to stake on play n based on the history *before* time n (i.e., up to and including play $n-1$). Your winnings on game n are $C_n \Delta X_n = C_n(X_n - X_{n-1})$. Your total (net) winnings up to time n are

$$Y_n = \sum_{k=1}^{n} C_k \Delta X_k = \sum_{k=1}^{n} C_k(X_k - X_{k-1}).$$

We write

$$Y = C \bullet X, \quad Y_n = (C \bullet X)_n, \quad \Delta Y_n = C_n \Delta X_n$$

($(C \bullet X)_0 = 0$ as $\sum_{k=1}^{0}$ is empty), and call $C \bullet X$ the *martingale transform* of X by C.

Theorem 3.4.1. *(i) If C is a bounded non-negative predictable process and X is a supermartingale, $C \bullet X$ is a supermartingale null at zero.*
(ii) If C is bounded and predictable and X is a martingale, $C \bullet X$ is a martingale null at zero.

Proof. With $Y = C \bullet X$ as above,

$$\begin{aligned} \mathbb{E}[Y_n - Y_{n-1}|\mathcal{F}_{n-1}] &= \mathbb{E}[C_n(X_n - X_{n-1})|\mathcal{F}_{n-1}] \\ &= C_n \mathbb{E}[(X_n - X_{n-1})|\mathcal{F}_{n-1}] \end{aligned}$$

(as C_n is bounded, so integrable, and $\mathcal{F}_{n-1}$-measurable, so can be taken out)

$$\leq 0$$

in case (i), as $C \geq 0$ and X is a supermartingale,

$$= 0$$

in case (ii), as X is a martingale. □

Interpretation. You can't beat the system! In the martingale case, previsibility of C means we can't foresee the future (which is realistic and fair). So we expect to gain nothing – as we should.

Note. 1. Martingale transforms were introduced and studied by D.L. Burkholder [33]. For a textbook account, see e.g. [169], VIII.4.
2. Martingale transforms are the discrete analogues of stochastic integrals. They dominate the mathematical theory of finance in discrete time, just as stochastic integrals dominate the theory in continuous time.

Lemma 3.4.1 (Martingale Transform Lemma). *An adapted sequence of real integrable random variables (M_n) is a martingale iff for any bounded previsible sequence (H_n),*

$$\mathbb{E}\left(\sum_{k=1}^{n} H_k \Delta M_k\right) = 0 \quad (n = 1, 2, \ldots).$$

Proof. If (M_n) is a martingale, X defined by $X_0 = 0$,

$$X_n = \sum_{k=1}^{n} H_r \Delta M_r \quad (n \geq 1)$$

is the martingale transform $H \bullet M$, so is a martingale.

Conversely, if the condition of the proposition holds, choose j, and for any $\mathcal{F}_j$-measurable set A write $H_n = 0$ for $n \neq j+1$, $H_{j+1} = I_A$. Then (H_n) is previsible, so the condition of the proposition, $\mathbb{E}(\sum_1^n H_r \Delta M_r) = 0$, becomes

$$\mathbb{E}[I_A(M_{j+1} - M_j)] = 0.$$

Since this holds for every set $A \in \mathcal{F}_j$, the definition of conditional expectation gives

$$\mathbb{E}(M_{j+1}|\mathcal{F}_j) = M_j.$$

Since this holds for every j, (M_n) is a martingale. □

Remark 3.4.1. The proof above is a good example of the value of Kolmogorov's definition of conditional expectation – which reveals itself, not in immediate transparency, but in its ease of handling in proofs. We shall see in Chapter 4 the financial significance of martingale transforms $H \bullet M$.

3.5 Stopping Times and Optional Stopping

A random variable T taking values in $\{0, 1, 2, \ldots; +\infty\}$ is called a *stopping time* (or optional time) if

$$\{T \leq n\} = \{\omega : T(\omega) \leq n\} \in \mathcal{F}_n \quad \forall n \leq \infty.$$

Equivalently,

$$\{T = n\} \in \mathcal{F}_n \quad n \leq \infty, \quad \text{or} \quad \{T \geq n\} \in \mathcal{F}_n, \quad n \leq \infty.$$

Think of T as a time at which you decide to quit a gambling game: whether or not you quit at time n depends only on the history up to and including time n – NOT the future. Thus stopping times model gambling and other situations where there is no foreknowledge, or prescience of the future; in particular, in the financial context, where there is no insider trading. (Elsewhere, T denotes the expiry time of an option. If we mean T to be a stopping time, we will say so.)

The following important classical theorem is discussed in [218], §10.10.

Theorem 3.5.1 (Doob's Optional Stopping Theorem, OST). *Let T be a stopping time, $X = (X_n)$ be a supermartingale, and assume that one of the following holds:*

(i) T is bounded ($T(\omega) \leq K$ for some constant K and all $\omega \in \Omega$);
(ii) $X = (X_n)$ is bounded ($|X_n(\omega)| \leq K$ for some K and all n, ω);
(iii) $ET < \infty$ and $(X_n - X_{n-1})$ is bounded.
Then X_T is integrable, and

$$EX_T \leq EX_0.$$

If X is a martingale, then

$$EX_T = EX_0.$$

The optional stopping theorem is important in many areas, such as sequential analysis in statistics. We turn in the next section to related ideas specific to the gambling/financial context.

Write $X_n^T := X_{n \wedge T}$ for the sequence (X_n) stopped at time T.

Proposition 3.5.1. *(i) If (X_n) is adapted and T is a stopping time, the stopped sequence $(X_{n\wedge T})$ is adapted.*
(ii) If (X_n) is a martingale (supermartingale) and T is a stopping time, (X_n^T) is a martingale (supermartingale).

Proof. If $\phi_j := 1_{\{j \leq T\}}$,

$$X_{T\wedge n} = X_0 + \sum_{j=1}^{n} \phi_j (X_j - X_{j-1})$$

(as the right is $X_0 + \sum_{j=1}^{T\wedge n}(X_j - X_{j-1})$, which telescopes to $X_{T\wedge n}$). Since $\{j \leq T\}$ is the complement of $\{T < j\} = \{T \leq j-1\} \in \mathcal{F}_{j-1}$, (ϕ_n) is predictable. So (X_n^T) is adapted.

If (X_n) is a martingale, so is (X_n^T) as it is the martingale transform of (X_n) by (ϕ_n). Since by predictability of (ϕ_n)

$$\begin{aligned} E(X_{T\wedge n}|\mathcal{F}_{n-1}) &= X_0 + \sum_{j=1}^{n-1} \phi_j (X_j - X_{j-1}) + \phi_n (E[X_n|\mathcal{F}_{n-1}] - X_{n-1}) \\ &= X_{T\wedge(n-1)} + \phi_n (E[X_n|\mathcal{F}_{n-1}] - X_{n-1}), \end{aligned}$$

$\phi_n \geq 0$ shows that if (X_n) is a supermartingale (submartingale), so is $(X_{T\wedge n})$.
□

Note. 1. See e.g. [169], Lemma II-12.14 for a variant on this proof avoiding the language of martingale transforms.
2. Part (ii) of the proposition says that if a game is fair, it remains fair when stopped (by a stopping time, i.e. without prescience); if it is (un-) favourable, it remains (un-) favourable when stopped.

Examples

1. Simple Random Walk. Recall the simple random walk: $S_n := \sum_1^n X_k$, where the X_n are independent tosses of a fair coin, taking values ± 1 with equal probability 1/2. Suppose we decide to bet until our net gain is first +1, then quit. Let T be the time we quit; T is a stopping time. The stopping time T has been analysed in detail; see e.g.[108], §5.3, or Exercise 10.

From this, note:

(i) $T < \infty$ a.s.: the gambler will certainly achieve a net gain of +1 eventually;
(ii) $ET = +\infty$: the mean waiting-time until this happens is infinity.
Hence also:
(iii) No bound can be imposed on the gambler's maximum net loss before his net gain first becomes +1.

At first sight, this looks like a foolproof way to make money out of nothing: just bet until you get ahead (which happens eventually, by (i)), then quit. However, as a gambling strategy, this is hopelessly impractical: because of (ii), you need unlimited time, and because of (iii), you need unlimited capital – neither of which is realistic.

Notice that the optional stopping theorem (Theorem 3.5.1) fails here: we start at zero, so $S_0 = 0$, $ES_0 = 0$; but $S_T = 1$, so $ES_T = 1$. This example shows two things:

(a) The Optional Stopping Theorem does indeed need conditions, as the conclusion may fail otherwise (none of the conditions (i) – (iii) in the OST are satisfied in the example above).
(b) Any practical gambling (or trading) strategy needs to have some integrability or boundedness restrictions to eliminate such theoretically possible but practically ridiculous cases.

2. The Doubling Strategy. The strategy of doubling when losing - *the martingale*, according to the *Oxford English Dictionary* (§3.3) – has similar properties. We play until the time T of our first win. Then T is a stopping time, and is geometrically distributed with parameter $p = 1/2$. If $T = n$, our winnings on the nth play are 2^{n-1} (our previous stake of 1 doubled on each of the previous $n-1$ losses). Our cumulative losses to date are $1 + 2 + \ldots + 2^{n-2} = 2^{n-1} - 1$ (summing the geometric series), giving us a net gain of 1. The mean time of play is $E(T) = 2$ (so doubling strategies accelerate our eventually certain win to give a finite expected waiting time for it). But no bound can be put on the losses one may need to sustain before we win, so again we would need unlimited capital to implement this strategy – which would be suicidal in practice as a result.

3. The Saint Petersburg Game. A single play of the Saint Petersburg game consists of a sequence of coin tosses stopped at the first head; if this is the rth toss, the player receives a prize of \$ 2^r. (Thus the expected gain is $\sum_{r=1}^{\infty} 2^{-r}.2^r = +\infty$, so the random variable is not integrable, and martingale theory does not apply.) Let S_n denote the player's cumulative gain after n plays of the game. The question arises as to what the 'fair price' of a ticket to play the game is. It turns out that fair prices exist (in a suitable sense), but the fair price of the nth play varies with n – surprising, as all the plays are replicas of each other.

Other examples may be constructed of games which are 'fair' in some sense, but in which the player sustains a net loss, tending to $-\infty$, with probability one. For a discussion of such examples, see e.g. [90].

3.6 The Snell Envelope

Definition 3.6.1. *If $Z = (Z_n)_{n=0}^N$ is a sequence adapted to a filtration $(\mathcal{F}_n)$, the sequence $U = (U_n)_{n=0}^N$ defined by*

$$\begin{cases} U_N := Z_N, \\ U_n := \max(Z_n, E(U_{n+1}|\mathcal{F}_n)) \quad (n \le N-1) \end{cases}$$

is called the Snell envelope of Z ([209]).

We shall see in Chapter 4 that the Snell envelope is the tool needed in pricing American options.

Theorem 3.6.1. *The Snell envelope (U_n) of (Z_n) is a supermartingale, and is the smallest supermartingale dominating (Z_n) (that is, with $U_n \ge Z_n$ for all n).*

Proof. First, $U_n \ge E(U_{n+1}|\mathcal{F}_n)$, so U is a supermartingale, and $U_n \ge Z_n$, so U dominates Z.

Next, let $T = (T_n)$ be any other supermartingale dominating Z; we must show T dominates U also. First, since $U_N = Z_N$ and T dominates Z, $T_N \ge U_N$. Assume inductively that $T_n \ge U_n$. Then

$$T_{n-1} \ge E(T_n|\mathcal{F}_{n-1}) \ge E(U_n|\mathcal{F}_{n-1}),$$

and as T dominates Z

$$T_{n-1} \ge Z_{n-1}.$$

Combining,

$$T_{n-1} \ge \max(Z_{n-1}, E(U_n|\mathcal{F}_{n-1})) = U_{n-1}.$$

By repeating this argument (or more formally, by backward induction), $T_n \ge U_n$ for all n, as required. □

Proposition 3.6.1. *$T_0 := \inf\{n \ge 0 : U_n = Z_n\}$ is a stopping time, and the stopped sequence $(U_n^{T_0})$ is a martingale.*

Proof. Since $U_N = Z_N$, $T_0 \in \{0, 1, \dots, N\}$ is well-defined. For $k = 0$, $\{T_0 = 0\} = \{U_0 = Z_0\} \in \mathcal{F}_0$; for $k \ge 1$,

$$\{T_0 = k\} = \{U_0 > Z_0\} \cap \cdots \cap \{U_{k-1} > Z_{k-1}\} \cap \{U_k = Z_k\} \in \mathcal{F}_k.$$

So T_0 is a stopping time.

As in the proof of Proposition 3.5.1,

$$U_n^{T_0} = U_{n \wedge T_0} = U_0 + \sum_{j=1}^{n} \phi_j \Delta U_j,$$

where $\phi_j = \mathbf{1}_{\{T_0 \ge j\}}$ is adapted. For $n \le N-1$,

$$U_{n+1}^{T_0} - U_n^{T_0} = \phi_{n+1}(U_{n+1} - U_n) = \mathbf{1}_{\{n+1 \le T_0\}}(U_{n+1} - U_n).$$

Now $U_n := \max(Z_n, \mathbb{E}(U_{n+1}|\mathcal{F}_n))$, and

$$U_n > Z_n \quad \text{on } \{n+1 \le T_0\}.$$

So from the definition of U_n,

$$U_n = \mathbb{E}(U_{n+1}|\mathcal{F}_n) \quad \text{on } \{n+1 \le T_0\}.$$

We next prove

$$U_{n+1}^{T_0} - U_n^{T_0} = \mathbf{1}_{\{n+1 \le T_0\}}(U_{n+1} - \mathbb{E}(U_{n+1}|\mathcal{F}_n)). \tag{3.3}$$

For, suppose first that $T_0 \ge n+1$. Then the left of (3.3) is $U_{n+1} - U_n$, the right is $U_{n+1} - \mathbb{E}(U_{n+1}|\mathcal{F}_n)$, and these agree on $\{n+1 \le T_0\}$ by above. The other possibility is that $T_0 < n+1$, i.e. $T_0 \le n$. Then the left of (3.3) is $U_{T_0} - U_{T_0} = 0$, while the right is zero because the indicator is zero, completing the proof of (3.3). Now apply $\mathbb{E}(.|\mathcal{F}_n)$ to (3.3): since $\{n+1 \le T_0\} = \{T_0 \le n\}^c \in \mathcal{F}_n$,

$$\begin{aligned} \mathbb{E}[(U_{n+1}^{T_0} - U_n^{T_0})|\mathcal{F}_n] &= \mathbf{1}_{\{n+1 \le T_0\}} \mathbb{E}([U_{n+1} - \mathbb{E}(U_{n+1}|\mathcal{F}_n)]|\mathcal{F}_n) \\ &= \mathbf{1}_{\{n+1 \le T_0)\}} [\mathbb{E}(U_{n+1}|\mathcal{F}_n) - \mathbb{E}(U_{n+1}|\mathcal{F}_n)] = 0. \end{aligned}$$

So $\mathbb{E}(U_{n+1}^{T_0}|\mathcal{F}_n) = U_n^{T_0}$. This says that $U_n^{T_0}$ is a martingale, as required. □

Write $\mathcal{T}_{n,N}$ for the set of stopping times taking values in $\{n, n+1, \dots, N\}$ (a finite set, as Ω is finite). We next see that the Snell envelope solves the optimal stopping problem.

Proposition 3.6.2. *T_0 solves the optimal stopping problem for Z:*

$$U_0 = \mathbb{E}(Z_{T_0}|\mathcal{F}_0) = \sup\{\mathbb{E}(Z_T|\mathcal{F}_0) : T \in \mathcal{T}_{0,N}\}.$$

Proof. To prove the first statement we use that $(U_n^{T_0})$ is a martingale and $U_{T_0} = Z_{T_0}$; then

$$U_0 = U_0^{T_0} = \mathbb{E}(U_N^{T_0}|\mathcal{F}_0) = \mathbb{E}(U_{T_0}|F_0) = \mathbb{E}(Z_{T_0}|\mathcal{F}_0).$$

Now for any stopping time $T \in \mathcal{T}_{0,N}$, since U is a supermartingale (above), so is the stopped process (U_n^T) (see Proposition 3.5.1). Together with the property that (U_n) dominates (Z_n) this yields

$$U_0 = U_0^{T_0} \geq E(U_N^T|\mathcal{F}_0) = E(U_T|\mathcal{F}_0) \geq E(Z_T|\mathcal{F}_0),$$

and this completes the proof. □

The same argument, starting at time n rather than time 0, gives

Corollary 3.6.1. *If* $T_n := \inf\{j \geq n : U_j = Z_j\}$,

$$U_n = \mathbb{E}(Z_{T_n}|\mathcal{F}_n) = \sup\{\mathbb{E}(Z_T|\mathcal{F}_n) : T \in \mathcal{T}_{n,N}\}.$$

As we are attempting to maximise our payoff by stopping $Z = (Z_n)$ at the most advantageous time, the Corollary shows that T_n gives the best stopping time that is realistic: it maximises our expected payoff given only information currently available (it is easy, but irrelevant, to maximise things with hindsight!). We thus call T_0 (or T_n, starting from time n) the *optimal* stopping time for the problem. For textbook accounts of optimal stopping problems, see e.g. [41], [169].

Optimal stopping problems have both an extensive - and quite deep - mathematical theory and applications to areas such as gambling, as well as the more speculative areas of mathematical finance. For a textbook treatment, see e.g. [41] (the Snell envelope is treated in §4.4). The gambling- and game-theoretic side of things is developed in the classic [66], and its recent sequel [159].

There are extensive links between the martingale theory of Chapter 3 and potential theory, classical and probabilistic. The least supermartingale majorant - Snell envelope - in martingale theory corresponds to the least excessive majorant in potential theory. This is called the *réduite* (reduced function) in potential theory. It occurs in the fundamental theorem of gambling ([159], §3.1); for the setting of probabilistic potential theory, see e.g. [163], IX.2.

Exercises

3.1 (i) Show that in general,

$$\mathbb{V}ar\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}ar(X_i) + \sum_{i\neq j} \mathbb{C}ov(X_i, X_j).$$

(ii) Show that if (X_n) is an L^2 martingale difference sequence (that is, $X_n = Z_n - Z_{n-1}$ with (Z_n) an L^2 martingale),

$$\mathbb{V}ar(Z_n) = \mathbb{V}ar\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}ar(X_i).$$

In particular, this holds if the X_n are independent.

3.2 1. Let $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Show that the mean-square error

$$\mathbb{E}\left[(Y - (aX + b))^2\right]$$

is minimized for $a^* = \mathbb{C}ov(X,Y)/\mathbb{V}ar(X)$ and $b^* = \mathbb{E}(Y) - a^*\mathbb{E}(X)$.
2. Now let $Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}$ a σ-algebra with $\mathcal{G} \subseteq \mathcal{F}$. Show that

$$\min_{a,b;X\in L^2(\Omega,\mathcal{G},\mathbb{P})} \mathbb{E}\left[(Y - (aX + b))^2\right] = \mathbb{E}\left[(Y - \mathbb{E}(Y|\mathcal{G}))^2\right].$$

3.3 A number d of balls are distributed between two urns, I and II. At each time $n = 0, 1, 2, \ldots$ a ball is chosen - each with equal probability $1/d$ - and transferred to the other urn.
(i) Show that the number of balls in urn I forms a Markov chain with transition probabilities

$$p_{i,i+1} = (d - i)/d, \quad p_{i,i-1} = i/d, \quad p_{ij} = 0 \text{ otherwise} \quad (i = 0, 1, \ldots, d).$$

(ii) Show that the stationary distribution is (π_i), where

$$\pi_i = 2^{-d}\binom{d}{i}$$

- that is, if the process is started in this distribution, it stays in it. (This is the *Ehrenfest urn*, treated in detail in [43], 129–132, [90], 377–378, etc. It exhibits a strong 'central push' towards the central states, and is a discrete-time analogue of the Ornstein-Uhlenbeck velocity process of §5.7.)

3.4 Consider a gambler who bets a unit stake on a succession of independent plays, each of which he wins with probability p, loses with probability $q := 1 - p$, with the strategy of quitting when first ahead. Write S_n for his net gain after n plays,

$$f_n := \mathbb{P}(S_1 \leq 0, \cdots, S_{n-1} \leq 0, S_n = 1)$$

for the probability that he quits at time n,

$$F(s) := \sum_{n=1}^{\infty} f_n s^n$$

for the generating function of the sequence (f_n). Show that:

(i) $F(s) = (1 - \sqrt{1 - 4pqs^2})/(2qs)$;
(ii) $f_{2n-1} = \frac{(-1)^{n-1}}{2q}\binom{\frac{1}{2}}{n}(4pq)^n$, $f_{2n} = 0$;
(iii) He eventually wins with probability 1 if $p \geq q$, p/q if $p < q$;
(iv) For $p \geq q$ (when he is certain to win eventually), the expected duration of play is $1/(p-q)$ if $p > q$, $+\infty$ if $p = q = \frac{1}{2}$. Thus if the game is fair, the expected waiting time to quitting when first ahead is infinite. (For a detailed account, see e.g. [90], XI.3, [108], §5.3.)

3.5 In the fair game case $p = q = \frac{1}{2}$ of the above:

(i) For each real θ, show that $M_n := (\cosh\theta)^{-n}e^{\theta S_n}$ is a martingale;
(ii) If $T := \inf\{n : S_n = 1\}$ is the duration of play, $\mathbb{P}(T < \infty) = 1$;
(iii) $\mathbb{E}(s^T) = \sum s^n \mathbb{P}(T = n) = (1 - \sqrt{1 - s^2})/s$, $\mathbb{P}(T = 2n - 1) = (-1)^{n-1}\binom{\frac{1}{2}}{n}$.
(iv) $\mathbb{E}(T) = \infty$. (Differentiate $\mathbb{E}(s^T)$ in (iii) and put $s = 1$.)
(This illustrates the power of martingale methods in such problems; for a detailed treatment, see [218], §10.12.)

4. Mathematical Finance in Discrete Time

4.1 The Model

We will study so-called finite markets – i.e. discrete-time models of financial markets in which all relevant quantities take a finite number of values. Following the approach of Harrison and Pliska [115] and Taqqu and Willinger [213], it suffices, to illustrate the ideas, to work with a finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with a finite number $|\Omega|$ of points ω, each with positive probability: $\mathbb{P}(\{\omega\}) > 0$.

We specify a time horizon T, which is the terminal date for all economic activities considered. (For a simple option pricing model the time horizon typically corresponds to the expiry date of the option.)

As before, we use a filtration $\mathbb{F}$ consisting of σ-algebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T$: we take $\mathcal{F}_0 = \{\emptyset, \Omega\}$, the trivial σ-field, $\mathcal{F}_T = \mathcal{F} = \mathcal{P}(\Omega)$ (here $\mathcal{P}(\Omega)$ is the power-set of Ω, the class of all $2^{|\Omega|}$ subsets of Ω: we need every possible subset, as they all – apart from the empty set – carry positive probability).

The financial market contains $d+1$ financial assets. The usual interpretation is to assume one (locally) risk-free asset (bond, bank account) labelled 0, and d risky assets (stocks, say) labelled 1 to d. While the reader may keep this interpretation as a mental picture, we prefer not to use it directly. The prices of the assets at time t are random variables, $S_0(t,\omega), S_1(t,\omega), \ldots, S_d(t,\omega)$ say, non-negative and $\mathcal{F}_t$-measurable (i.e. adapted: at time t, we know the prices $S_i(t)$). We write $S(t) = (S_0(t), S_1(t), \ldots, S_d(t))'$ for the vector of prices at time t. Hereafter we refer to the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the set of trading dates, the price process S and the information structure $\mathbb{F}$, which is typically generated by the price process S, together as a securities market model $\mathcal{M}$.

It will be essential to assume that the price process of at least one asset follows a strictly positive process.

Definition 4.1.1. *A numéraire is a price process* $(X(t))_{t=0}^T$ *(a sequence of random variables), which is strictly positive for all* $t \in \{0, 1, \ldots, T\}$.

For the standard approach the risk-free bank account process is used as numéraire. In some applications, however, it is more convenient to use a security other than the bank account and we therefore just use S_0 without further specification as a numéraire. We furthermore take $S_0(0) = 1$ (that is, we reckon in units of the initial value of our numéraire), and define $\beta(t) := 1/S_0(t)$ as a discount factor. We will occasionally refer to assets $1, \ldots, d$ as the risky assets.

A *trading strategy* (or *dynamic portfolio*) φ is a $\mathbb{R}^{d+1}$ vector stochastic process $\varphi = (\varphi(t))_{t=1}^T = ((\varphi_0(t,\omega), \varphi_1(t,\omega), \ldots, \varphi_d(t,\omega))')_{t=1}^T$ which is predictable (or previsible): each $\varphi_i(t)$ is $\mathcal{F}_{t-1}$-measurable for $t \geq 1$. Here $\varphi_i(t)$ denotes the number of shares of asset i held in the portfolio at time t – to be determined on the basis of information available *before* time t; i.e. the investor selects his time t portfolio after observing the prices $S(t-1)$. However, the portfolio $\varphi(t)$ must be established before, and held until after, announcement of the prices $S(t)$. The components $\varphi_i(t)$ may assume negative as well as positive values, reflecting the fact that we allow short sales and assume that the assets are perfectly divisible.

Definition 4.1.2. *The value of the portfolio at time t is the scalar product*

$$V_\varphi(t) = \varphi(t) \cdot S(t) := \sum_{i=0}^{d} \varphi_i(t) S_i(t), \quad (t = 1, 2, \ldots, T) \quad \text{and } V_\varphi(0) = \varphi(1) \cdot S(0).$$

The process $V_\varphi(t,\omega)$ is called the wealth or value process of the trading strategy φ.

The initial wealth $V_\varphi(0)$ is called the ***initial investment*** or ***endowment*** of the investor.

Now $\varphi(t) \cdot S(t-1)$ reflects the market value of the portfolio just after it has been established at time $t-1$, whereas $\varphi(t) \cdot S(t)$ is the value just after time t prices are observed, but before changes are made in the portfolio. Hence

$$\varphi(t) \cdot (S(t) - S(t-1)) = \varphi(t) \cdot \Delta S(t)$$

is the change in the market value due to changes in security prices which occur between time $t-1$ and t. This motivates:

Definition 4.1.3. *The gains process G_φ of a trading strategy φ is given by*

$$G_\varphi(t) := \sum_{\tau=1}^{t} \varphi(\tau) \cdot (S(\tau) - S(\tau - 1)) = \sum_{\tau=1}^{t} \varphi(\tau) \cdot \Delta S(\tau), \quad (t = 1, 2, \ldots, T).$$

Define $\tilde{S}(t) = (1, \beta(t)S_1(t), \ldots, \beta(t)S_d(t))'$, the vector of discounted prices, and consider the *discounted value process*

$$\tilde{V}_\varphi(t) = \beta(t)(\varphi(t) \cdot S(t)) = \varphi(t) \cdot \tilde{S}(t), \quad (t = 1, 2, \ldots, T)$$

and the *discounted gains process*

$$\tilde{G}_\varphi(t) := \sum_{\tau=1}^{t} \varphi(\tau) \cdot (\tilde{S}(\tau) - \tilde{S}(\tau - 1)) = \sum_{\tau=1}^{t} \varphi(\tau) \cdot \Delta \tilde{S}(\tau), \quad (t = 1, 2, \ldots, T).$$

Observe that the discounted gains process reflects the gains from trading with assets 1 to d only, which in case of the standard model (a bank account and d stocks) are the risky assets.

We will only consider special classes of trading strategies.

Definition 4.1.4. *The strategy φ is self-financing, $\varphi \in \Phi$, if*

$$\varphi(t) \cdot S(t) = \varphi(t+1) \cdot S(t) \quad (t = 1, 2, \ldots, T-1). \tag{4.1}$$

Interpretation. When new prices $S(t)$ are quoted at time t, the investor adjusts his portfolio from $\varphi(t)$ to $\varphi(t+1)$, without bringing in or consuming any wealth. The following result (which is trivial in our current setting, but requires a little argument in continuous time) shows that renormalising security prices (i.e. changing the numéraire) has essentially no economic effects.

Proposition 4.1.1 (Numéraire Invariance). *Let $X(t)$ be a numéraire. A trading strategy φ is self-financing with respect to $S(t)$ if and only if φ is self-financing with respect to $X(t)^{-1}S(t)$.*

Proof. Since $X(t)$ is strictly positive for all $t = 0, 1, \ldots, T$ we have the following equivalence, which implies the claim:

$$\varphi(t) \cdot S(t) = \varphi(t+1) \cdot S(t) \quad (t = 1, 2, \ldots, T-1)$$
$$\Leftrightarrow$$
$$\varphi(t) \cdot X(t)^{-1} S(t) = \varphi(t+1) \cdot X(t)^{-1} S(t) \quad (t = 1, 2, \ldots, T-1).$$

□

Corollary 4.1.1. *A trading strategy φ is self-financing with respect to $S(t)$ if and only if φ is self-financing with respect to $\tilde{S}(t)$.*

We now give a characterisation of self-financing strategies in terms of the discounted processes.

Proposition 4.1.2. *A trading strategy φ belongs to Φ if and only if*

$$\tilde{V}_\varphi(t) = V_\varphi(0) + \tilde{G}_\varphi(t), \quad (t = 0, 1, \ldots, T). \tag{4.2}$$

Proof. Assume $\varphi \in \Phi$. Then using the defining relation (4.1), the numéraire invariance theorem and the fact that $S_0(0) = 1$

$$\begin{aligned} V_\varphi(0) + \tilde{G}_\varphi(t) &= \varphi(1) \cdot S(0) + \sum_{\tau=1}^{t} \varphi(\tau) \cdot (\tilde{S}(\tau) - \tilde{S}(\tau-1)) \\ &= \varphi(1) \cdot \tilde{S}(0) + \varphi(t) \cdot \tilde{S}(t) \\ &\quad - \sum_{\tau=1}^{t-1} (\varphi(\tau) - \varphi(\tau+1)) \cdot \tilde{S}(\tau) - \varphi(1) \cdot \tilde{S}(0) \\ &= \varphi(t) \cdot \tilde{S}(t) = \tilde{V}_\varphi(t). \end{aligned}$$

Assume now that (4.2) holds true. By the numéraire invariance theorem it is enough to show the discounted version of relation (4.1) for each $t = 1, 2, \ldots, T-1$. Summing up to $t = 2$ (4.2) is

$$\varphi(2) \cdot \bar{S}(2) = \varphi(1) \cdot \bar{S}(0) + \varphi(1) \cdot (\bar{S}(1) - \bar{S}(0)) + \varphi(2) \cdot (\bar{S}(2) - \bar{S}(1)).$$

Subtracting $\varphi(2) \cdot \bar{S}(2)$ on both sides gives $\varphi(2) \cdot \bar{S}(1) = \varphi(1) \cdot \bar{S}(1)$. Proceeding similarly – or by induction – we can show $\varphi(t) \cdot \bar{S}(t) = \varphi(t+1) \cdot \bar{S}(t)$ for $t = 2, \ldots, T-1$ as required. □

We are allowed to borrow (so $\varphi_0(t)$ may be negative) and sell short (so $\varphi_i(t)$ may be negative for $i = 1, \ldots, d$). So it is hardly surprising that if we decide what to do about the risky assets, the numéraire will take care of itself, in the following sense.

Proposition 4.1.3. *If $(\varphi_1(t), \ldots, \varphi_d(t))'$ is predictable and V_0 is $\mathcal{F}_0$-measurable, there is a unique predictable process $(\varphi_0(t))_{t=1}^T$ such that $\varphi = (\varphi_0, \varphi_1, \ldots, \varphi_d)'$ is self-financing with initial value of the corresponding portfolio $V_\varphi(0) = V_0$.*

Proof. If φ is self-financing, then by Proposition 4.1.2,

$$\tilde{V}_\varphi(t) = V_0 + \tilde{G}_\varphi(t) = V_0 + \sum_{\tau=1}^{t} (\varphi_1(\tau)\Delta\tilde{S}_1(\tau) + \ldots + \varphi_d(\tau)\Delta\tilde{S}_d(\tau)).$$

On the other hand,

$$\tilde{V}_\varphi(t) = \varphi(t) \cdot \tilde{S}(t) = \varphi_0(t) + \varphi_1(t)\tilde{S}_1(t) + \ldots + \varphi_d(t)\tilde{S}_d(t).$$

Equate these:

$$\varphi_0(t) = V_0 + \sum_{\tau=1}^{t} (\varphi_1(\tau)\Delta\tilde{S}_1(\tau) + \ldots + \varphi_d(\tau)\Delta\tilde{S}_d(\tau)) - (\varphi_1(t)\tilde{S}_1(t) + \ldots + \varphi_d(t)\tilde{S}_d(t)),$$

which defines $\varphi_0(t)$ uniquely. The terms in $\tilde{S}_i(t)$ are

$$\varphi_i(t)\Delta\tilde{S}_i(t) - \varphi_i(t)\tilde{S}_i(t) = -\varphi_i(t)\tilde{S}_i(t-1),$$

which is $\mathcal{F}_{t-1}$-measurable. So

$$\varphi_0(t) = V_0 + \sum_{\tau=1}^{t-1} (\varphi_1(\tau)\Delta\tilde{S}_1(\tau) + \ldots + \varphi_d(\tau)\Delta\tilde{S}_d(\tau)) - (\varphi_1(t)S_1(t-1) + \ldots + \varphi_d(t)\tilde{S}_d(t-1)),$$

where as $\varphi_1, \ldots, \varphi_d$ are predictable, all terms on the right-hand side are $\mathcal{F}_{t-1}$-measurable, so φ_0 is predictable. □

4.2 Existence of Equivalent Martingale Measures

4.2.1 The No-Arbitrage Condition

The central principle in the single period example, §1.4, was the absence of arbitrage opportunities, i.e. the absence of risk-free plans for making profits without any investment. As mentioned there this principle is central for any market model, and we now define the mathematical counterpart of this economic principle in our current setting.

Definition 4.2.1. *Let $\tilde{\Phi} \subset \Phi$ be a set of self-financing strategies. A strategy $\varphi \in \tilde{\Phi}$ is called an arbitrage opportunity or arbitrage strategy with respect to $\tilde{\Phi}$ if $\mathbb{P}\{V_\varphi(0) = 0\} = 1$, and the terminal wealth of φ satisfies*

$$\mathbb{P}\{V_\varphi(T) \geq 0\} = 1 \text{ and } \mathbb{P}\{V_\varphi(T) > 0\} > 0.$$

So an arbitrage opportunity is a self-financing strategy with zero initial value, which produces a non-negative final value with probability one and has a positive probability of a positive final value. Observe that arbitrage opportunities are always defined with respect to a certain class of trading strategies.

Definition 4.2.2. *We say that a security market $\mathcal{M}$ is arbitrage-free if there are no arbitrage opportunities in the class Φ of trading strategies.*

We will allow ourselves to use 'no-arbitrage' in place of 'arbitrage-free' when convenient.

We will use the following mental picture in analysing the the sample paths of the price processes. We observe a realisation $S(t,\omega)$ of the price process $S(t)$. We want to know which sample point $\omega \in \Omega$ – or random outcome – we have. Information about ω is captured in the filtration $\mathcal{F}_t$. In our current setting we can switch to the unique sequence of partitions $\mathcal{P}_t$ corresponding to the filtration $\mathcal{F}_t$ (see Chapter 3). So at time t we know the set $A_t \in \mathcal{P}_t$ with $\omega \in A_t$. Now recall the structure of the subsequent partitions. A set $A \in \mathcal{P}_t$ is the disjoint union of sets $A_1, \ldots, A_K \in \mathcal{P}_{t+1}$. Since $S(u)$ is $\mathcal{F}_u$-measurable $S(t)$ is constant on A and $S(t+1)$ is constant on the A_k, $k = 1, \ldots, K$. So we can think of A as the time 0 state in a single-period model and each A_k corresponds to a state at time 1 in the single-period model. We can therefore think of a multi-period market model as a collection of subsequent single-period markets. What is the effect of a 'global' no-arbitrage condition on the single-period markets?

Lemma 4.2.1. *If the market model contains no arbitrage opportunities, then for all $t \in \{0, 1, \ldots, T-1\}$, for all self-financing trading strategies $\varphi \in \Phi$ and for any $A \in \mathcal{P}_t$, we have*

(i) $\mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \geq 0 | A) = 1 \Rightarrow \mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) = 0 | A) = 1$,

(ii) $\mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \leq 0 | A) = 1 \Rightarrow \mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) = 0 | A) = 1$.

The economic meaning of this result answers the question raised above. No arbitrage 'globally' implies no arbitrage 'locally'. From this the idea of the proof is immediate. Any local trading strategy can be embedded in a global strategy for which we can use the global no-arbitrage condition.

Proof. We only prove (i) ((ii) is shown in a similar fashion). Fix $t \in \{0,\dots,T-1\}$ and $\varphi \in \Phi$. Suppose $\mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \geq 0|A) = 1$ for some $A \in \mathcal{P}_t$ and define a new trading strategy ψ for all times $u = 1,\dots,T$ as follows:

For $u \leq t$: $\psi(u) = 0$ ('do nothing before time t').
For $u = t+1$: $\psi(t+1) = 0$ if $\omega \notin A$, and

$$\psi_k(t+1,\omega) = \begin{cases} \varphi_k(t+1,\omega) & \text{if } \omega \in A \text{ and } k \in \{1,\dots,d\}, \\ \varphi_0(t+1,\omega) - \tilde{V}_\varphi(t,\omega) & \text{if } \omega \in A \text{ and } k = 0. \end{cases}$$

('If ω happens to be in A at time t, follow strategy φ when dealing with the risky assets, but modify the holdings in the numéraire appropriately in order to compensate for doing nothing when $\omega \notin A$.')

For $u > t+1$: $\psi_k(u) = 0$ for $k \in \{1,\dots,d\}$ and

$$\psi_0(u,\omega) = \begin{cases} \tilde{V}_\psi(t+1,\omega) & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases}$$

('Invest the amount $\tilde{V}_\psi(t+1)$ into the numéraire account if ω happens to be in A, otherwise do nothing'.)

The next step now is to show that the strategy ψ is a self-financing trading strategy. By construction ψ is predictable, hence a trading strategy. For $\omega \notin A$ $\psi \equiv 0$, so we only have to consider $\omega \in A$. The relevant point in time is $t+1$. Recall that $\psi(t) = 0$, hence $\psi(t) \cdot \tilde{S}(t) = 0$. Now

$$\begin{aligned} \psi(t+1) \cdot \tilde{S}(t) &= (\varphi_0(t+1) - \tilde{V}_\varphi(t))\tilde{S}_0(t) + \sum_{k=1}^{d} \varphi_k(t+1)\tilde{S}_k(t) \\ &= \sum_{k=0}^{d} \varphi_k(t+1)\tilde{S}_k(t) - \tilde{V}_\varphi(t) \\ &= \varphi(t+1) \cdot \tilde{S}(t) - \tilde{V}_\varphi(t) = \varphi(t) \cdot \tilde{S}(t) - \tilde{V}_\varphi(t) = 0, \end{aligned}$$

using the fact that φ is self-financing. Since $\psi(u) \cdot \tilde{S}(u) = 0$ for $u \leq t$ we have $\psi(u+1) \cdot \tilde{S}(u) = \psi(u) \cdot \tilde{S}(u)$ for all $u \leq t$ (and for all $\omega \in \Omega$). When $u > t+1$ and $\omega \in A$ we only hold the numéraire asset (with constant discounted value equal to 1), so

$$\psi(u+1) \cdot \tilde{S}(u) = \tilde{V}_\psi(t+1) = \psi(u) \cdot \tilde{S}(u).$$

Therefore the strategy ψ is self-financing.

We now analyse the value process of ψ. Using our assumption $\mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \geq 0|A) = 1$ we see that for all $u \geq t+1$ and $\omega \in A$

$$\begin{aligned} \tilde{V}_\psi(u) = \psi(u) \cdot \tilde{S}(u) &= \psi(t+1) \cdot \tilde{S}(t+1) \\ &= (\varphi_0(t+1) - \tilde{V}_\varphi(t))\tilde{S}_0(t) + \sum_{k=1}^{d} \varphi_k(t+1)\tilde{S}_k(t) \\ &= \sum_{k=0}^{d} \varphi_k(t+1)\tilde{S}_k(t) - \tilde{V}_\varphi(t) \\ &= \tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \geq 0. \end{aligned}$$

Since $\tilde{V}_\psi(T) = 0$ on A^c ψ defines a self-financing trading strategy with $\tilde{V}_\psi(0) = 0$ and $\tilde{V}_\psi(T) \geq 0$. The assumption of an arbitrage-free market implies $\tilde{V}_\psi(T) = 0$ or

$$\begin{aligned} 0 = \mathbb{P}(\tilde{V}_\psi(T) > 0) &= \mathbb{P}\left(\{\tilde{V}_\psi(T) > 0\} \cap A\right) \\ &= \mathbb{P}(\tilde{V}_\psi(t+1) - \tilde{V}_\psi(t) > 0|A)\mathbb{P}(A). \end{aligned}$$

Therefore $\mathbb{P}(\tilde{V}_\psi(t+1) - \tilde{V}_\psi(t) = 0|A) = 1$. □

The fundamental insight in §1.4 was the equivalence of the no-arbitrage condition and the existence of risk-neutral probabilities. For the multi-period case we now use the probabilistic machinery of Chapters 2 and 3 to establish the corresponding result.

Definition 4.2.3. *A probability measure $\mathbb{P}^*$ on $(\Omega, \mathcal{F}_T)$ equivalent to $\mathbb{P}$ is called a martingale measure for $\tilde{S}$ if the process $\tilde{S}$ follows a $\mathbb{P}^*$-martingale with respect to the filtration $\mathbb{F}$. We denote by $\mathcal{P}(\tilde{S})$ the class of equivalent martingale measures.*

Proposition 4.2.1. *Let $\mathbb{P}^*$ be an equivalent martingale measure ($\mathbb{P}^* \in \mathcal{P}(\tilde{S})$) and $\varphi \in \Phi$ any self-financing strategy. Then the wealth process $\tilde{V}_\varphi(t)$ is a $\mathbb{P}^*$-martingale with respect to the filtration $\mathbb{F}$.*

Proof. By the self-financing property of φ (compare Proposition 4.1.2, (4.2)), we have

$$\tilde{V}_\varphi(t) = V_\varphi(0) + \tilde{G}_\varphi(t) \quad (t = 0,1,\dots,T).$$

So

$$\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) = \tilde{G}_\varphi(t+1) - \tilde{G}_\varphi(t) = \varphi(t+1) \cdot (\tilde{S}(t+1) - \tilde{S}(t)).$$

So for $\varphi \in \Phi$, $\tilde{V}_t(\varphi)$ is the martingale transform of the $\mathbb{P}^*$ martingale $\tilde{S}$ by φ (see Theorem 3.4.1) and hence a $\mathbb{P}^*$ martingale itself. □

The next result is the key for the further development.

Proposition 4.2.2. *If an equivalent martingale measure exists - that is, if $\mathcal{P}(\tilde{S}) \neq \emptyset$ - then the market $\mathcal{M}$ is arbitrage-free.*

Proof. Assume such a $\mathbb{P}^*$ exists. For any self-financing strategy φ, we have as before

$$\tilde{V}_\varphi(t) = V_\varphi(0) + \sum_{\tau=1}^{t} \varphi(\tau) \cdot \Delta\tilde{S}(\tau).$$

By Proposition 4.2.1, $\tilde{S}(t)$ a (vector) $\mathbb{P}^*$-martingale implies $\tilde{V}_\varphi(t)$ is a P^*-martingale. So the initial and final $\mathbb{P}^*$-expectations are the same,

$$\mathbb{E}^*(\tilde{V}_\varphi(T)) = \mathbb{E}^*(\tilde{V}_\varphi(0)).$$

If the strategy is an arbitrage opportunity its initial value – the right-hand side above – is zero. Therefore the left-hand side $\mathbb{E}^*(\tilde{V}_\varphi(T))$ is zero, but $\tilde{V}_\varphi(T) \geq 0$ (by definition). Also each $\mathbb{P}^*(\{\omega\}) > 0$ (by assumption, each $\mathbb{P}(\{\omega\}) > 0$, so by equivalence each $\mathbb{P}^*(\{\omega\}) > 0$). This and $\tilde{V}_\varphi(T) \geq 0$ force $\tilde{V}_\varphi(T) = 0$. So no arbitrage is possible. □

Proposition 4.2.3. *If the market $\mathcal{M}$ is arbitrage-free, then the class $\mathcal{P}(\tilde{S})$ of equivalent martingale measures is non-empty.*

Because of the fundamental nature of this result we will provide two proofs. The first proof is based on our previous observation that the 'global' no-arbitrage condition implies also no-arbitrage 'locally'. We therefore can combine single-period results to prove the multi-period claim. The second prove uses functional-analytic techniques (as does the corresponding proof in Chapter 1), i.e. a variant of the *Hahn-Banach theorem.*

First proof. From Lemma 4.2.1 we know that each of the underlying single-period market models is free of arbitrage. By the results in §1.4 this implies the existence of risk-neutral probabilities. That is, for each $t \in \{0, 1, \dots, T-1\}$ and each $A \in \mathcal{P}_t$ there exists a probability measure $\mathbb{P}(t, A)$ such that each cell $A_i \subset A$, $i = 1, \dots, K_A$ in the partition $\mathcal{P}_{t+1}$ has a positive probability mass and

$$\sum_{i=1}^{K_A} \mathbb{P}(t, A)(A_i) = 1.$$

Furthermore $\mathbb{E}_{\mathbb{P}(t,A)}(\tilde{S}(t+1)) = \tilde{S}(t)$ (where we restrict ourselves to $\omega \in A$). We can think of the probability measures $\mathbb{P}(t, A)$ as conditional risk-neutral probability measures given the event A occurred at time t. Now we can define a probability measure $\mathbb{P}^*$ on Ω by defining the probabilities of the simple events $\{\omega\}$ (observe that $\mathcal{F}_T = \mathcal{P}(\Omega)$, hence the final partition consists of all simple events). To each such $\{\omega\}$ there exists a single path from 0 to T and $\mathbb{P}^*$ is set equal to the product of the conditional probabilities along the path. By construction

$$\sum_{\omega \in \Omega} \mathbb{P}^*(\{\omega\}) = 1.$$

Since the conditional risk-neutral probabilities are greater than 0, $\mathbb{P}^*(\{\omega\}) > 0$ for each $\omega \in \Omega$ and $\mathbb{P}^*$ is an equivalent measure. The final step is to show that $\mathbb{P}^*$ is a martingale measure. We thus have to show $\mathbb{E}^*(\tilde{S}_k(t+1)|\mathcal{F}_t) = \tilde{S}_k(t)$ for any $k = 1, \dots, d$, $t = 0, \dots, T-1$. Now $\tilde{S}_k(t)$ is $\mathcal{F}_t$-measurable, and since any $A \in \mathcal{F}_t$ can be written as a union of $A' \in \mathcal{P}_t$ the claim follows from

$$\int_{A'} \tilde{S}_k(t+1) d\mathbb{P}^* = \int_{A'} \tilde{S}_k(t) d\mathbb{P}^*,$$

which is true by construction of $\mathbb{P}^*$. (Recall that we have $\mathbb{E}_{\mathbb{P}(A,t)}(\tilde{S}_k(t+1)) = \mathbb{E}_{\mathbb{P}(A,t)}(\tilde{S}_k(t))$.) □

For the second proof we need some auxiliary observations.

Since our market model is finite we can use results from Euclidean geometry. Define $\mathcal{X}^+$ to be the set of non-negative random variables on $(\Omega, \mathcal{F})$ and

$$\Gamma := \{X \in \mathcal{X}^+ : X(\omega) \geq 0 \ \forall \omega \in \Omega \text{ and } \exists\, \omega \in \Omega \text{ such that } X(\omega) > 0\}.$$

Then Γ is a *cone* (closed under vector addition and multiplication by positive scalars). If we assume no-arbitrage with respect to Φ, we have for any self-financing strategy φ,

$$V_\varphi(0) = 0 \quad \Rightarrow \quad \tilde{V}_\varphi(T) \notin \Gamma.$$

By Proposition 4.1.2 it follows that $\tilde{G}_\varphi(T) \notin \Gamma$.

The next lemma shows that we still have $\tilde{G}_\varphi(T) \notin \Gamma$ if we only assume predictability of a vector process $(\varphi_1, \dots, \varphi_d)$ and choose a process φ_0 according to Proposition 4.1.3 in such a way that the strategy $\varphi = (\varphi_0, \varphi_1, \dots, \varphi_d)$ has zero initial value and is self-financing.

Lemma 4.2.2. *In an arbitrage-free market any predictable vector process $\varphi' = (\varphi_1, \dots, \varphi_d)$ satisfies*

$$\tilde{G}_{\varphi'}(T) \notin \Gamma.$$

(Observe the slight abuse of notation: for the value of the discounted gains process the zeroth component of a trading strategy doesn't matter. Hence we use the operator $\tilde{G}$ for d-dimensional vectors as well.)

Proof. By Proposition 4.1.3 there exists a unique predictable process $(\varphi_0(t))$ such that $\varphi = (\varphi_0, \varphi_1, \dots, \varphi_d)$ has zero initial value and is self-financing. Assume $\tilde{G}_{\varphi'}(T) \in \Gamma$. Then using Proposition 4.1.2,

$$V_\varphi(T) = \beta(T)^{-1}\tilde{V}_\varphi(T) = \beta(T)^{-1}(V_\varphi(0) + \tilde{G}_\varphi(T)) = \beta(T)^{-1}\tilde{G}_{\varphi'}(T) > 0.$$

Hence φ is an arbitrage opportunity with respect to Φ by definition of Γ. By Lemma 4.2.1 this contradicts the assumption that the market is arbitrage-free. □

Second proof of Proposition 4.2.3. Denote the vector space of all random variables on Ω by $\mathcal{X}$ (this space can be identified with $\mathbb{R}^{\Omega}$). We now form the set $\mathcal{V}$ of random variables $\tilde{G}_{\varphi'}(T)$, with $\varphi' = (\varphi_1, \ldots, \varphi_d)$ a predictable process:

$$\mathcal{V} := \{X \in \mathcal{X} : X = \tilde{G}_{\varphi'}(T),\ \varphi' \text{ predictable}\}.$$

By linearity of the discounted gains process $\tilde{G}_{\varphi'}$ in φ' this is a vector subspace of the vector space $\mathcal{X}$. By Lemma 4.2.2 this subspace does not meet Γ, i.e.

$$\mathcal{V} \cap \Gamma = \emptyset.$$

So $\mathcal{V}$ does not meet the subset

$$K := \{X \in \Gamma : \sum_{\omega \in \Omega} X(\omega) = 1\}.$$

Now K is a compact convex set. By the separating hyperplane theorem (Theorem C.0.1), there is a vector $\lambda = (\lambda(\omega) : \omega \in \Omega)$ such that for all $X \in K$

$$\lambda \cdot X := \sum_{\omega \in \Omega} \lambda(\omega) X(\omega) > 0,$$

but for all $\tilde{G}_{\varphi'}(T)$ in $\mathcal{V}$,

$$\lambda \cdot \tilde{G}_{\varphi'}(T) = \sum_{\omega \in \Omega} \lambda(\omega) \tilde{G}_{\varphi'}(T)(\omega) = 0. \tag{4.3}$$

Choosing each $\omega \in \Omega$ successively and taking X to be 1 on this ω and zero elsewhere, (4.3) tells us that each $\lambda(\omega) > 0$. So

$$\mathbb{P}^*(\{\omega\}) := \frac{\lambda(\omega)}{\sum_{\omega' \in \Omega} \lambda(\omega')}$$

defines a probability measure equivalent to $\mathbb{P}$ (no non-empty null sets). With $\mathbb{E}^*$ as $\mathbb{P}^*$-expectation, (4.3) says that

$$\mathbb{E}^*\left(\tilde{G}_{\varphi}(T)\right) = 0,$$

i.e.

$$\mathbb{E}^*\left(\sum_{\tau=1}^{T} \varphi(\tau) \cdot \Delta \tilde{S}(\tau)\right) = 0.$$

In particular, choosing for each i to hold only stock i,

$$\mathbb{E}^*\left(\sum_{\tau=1}^{T} \varphi_i(\tau) \Delta \tilde{S}_i(\tau)\right) = 0 \quad (i = 1, \ldots, d).$$

Since this holds for any predictable φ (boundedness holds automatically as Ω is finite), the martingale transform lemma tells us that the discounted price processes $(\tilde{S}_i(t))$ are $\mathbb{P}^*$-martingales. □

Note. Our situation is finite-dimensional, so all we have used here is Euclidean geometry. We have a subspace, and a cone not meeting the subspace except at the origin. Take λ orthogonal to the subspace on the same side of the subspace as the cone. The separating hyperplane theorem holds also in infinite-dimensional situations, where it is a form of the Hahn-Banach theorem of functional analysis. For proofs, variants and background, see e.g. [27, 214].

We now combine Propositions 4.2.2 and 4.2.3 as a first central theorem in this chapter.

Theorem 4.2.1 (No-Arbitrage Theorem). *The market $\mathcal{M}$ is arbitrage-free if and only if there exists a probability measure $\mathbb{P}^*$ equivalent to $\mathbb{P}$ under which the discounted d-dimensional asset price process $\tilde{S}$ is a $\mathbb{P}^*$-martingale.*

4.2.2 Risk-Neutral Pricing

We now turn to the main underlying question of this text, namely the pricing of contingent claims (i.e. financial derivatives). First we have to model these financial instruments in our current framework. This is done in the following fashion.

Definition 4.2.4. *A contingent claim X with maturity date T is an arbitrary non-negative $\mathcal{F}_T$-measurable random variable (which is by the finiteness of the probability space bounded). We denote the class of all contingent claims by $\mathcal{X}^+$.*

A typical example of a contingent claim X is an option on some underlying asset S, then (e.g. for the case of a European call option with maturity date T and strike K) we have a functional relation $X = f(S)$ with some function f (e.g. $X = (S(T) - K)^+$). The general definition allows for more complicated relationships which are captured by the $\mathcal{F}_T$-measurability of X (recall that $\mathcal{F}_T$ is typically generated by the process S).

We say that the claim is *attainable* if there exists a *replicating strategy* $\varphi \in \Phi$ such that

$$V_{\varphi}(T) = X.$$

So the replicating strategy generates the same time T cash-flow as does X. In a highly efficient security market we expect that the law of one price holds true, that is for a specified cash-flow there exists only one price at any time instant. Otherwise arbitrageurs would use the opportunity to cash in a riskless profit (recall the recent case of option mispricing at NatWest Markets as an excellent example of how arbitrageurs exploit mispricing [97]). So the no-arbitrage condition implies that for an attainable contingent claim its time t price must be given by the value of any replicating strategy (we say the claim is uniquely replicated in that case). This is the basic idea of the *arbitrage pricing theory*. Clearly the equivalence of the no-arbitrage condition and the existence of risk-neutral probability measures imply the possibility of using risk-neutral measures for pricing purposes, and we devote this subsection to exploring this relation.

Let us start by investigating the arbitrage pricing approach a bit further. The idea is to replicate a given cash-flow at a given point in time. Using a self-financing trading strategy the investor's wealth may go negative at time $t < T$, but he must be able to cover his debt at the final date. To avoid negative wealth the concept of admissible strategies is introduced. A self-financing trading strategy $\varphi \in \Phi$ is called *admissible* if $V_\varphi(t) \geq 0$ for each $t = 0, 1, \ldots, T$. We write Φ_a for the class of admissible trading strategies. The modelling assumption of admissible strategies reflects the economic fact that the broker should be protected from unbounded short sales. From the mathematical point of view it is not really needed and we use self-financing strategies when addressing the mathematical aspects of the theory. (In fact one can show that a security market which is arbitrage-free with respect to Φ_a is also arbitrage-free with respect to Φ.)

We now return to the main question of the section: given a contingent claim X, i.e. a cash-flow at time T, how can we determine its value (price) at time $t < T$? For an attainable contingent claim this value should be given by the value of any replicating strategy at time t, i.e. there should be a unique value process (say $V_X(t)$) representing the time t value of the simple contingent claim X. The following proposition ensures that the value processes of replicating trading strategies coincide, thus proving the uniqueness of the value process.

Proposition 4.2.4. *Suppose the market $\mathcal{M}$ is arbitrage-free. Then any attainable contingent claim X is uniquely replicated in $\mathcal{M}$.*

Proof. Suppose there is an attainable contingent claim X and strategies φ and ψ such that

$$V_\varphi(T) = V_\psi(T) = X,$$

but there exists a $\tau < T$ such that

$$V_\varphi(u) = V_\psi(u) \quad \text{for every } u < \tau \text{ and } V_\varphi(\tau) \neq V_\psi(\tau).$$

Define $A := \{\omega \in \Omega : V_\varphi(\tau,\omega) > V_\psi(\tau,\omega)\}$, then $A \in \mathcal{F}_\tau$ and $\mathbb{P}(A) > 0$ (otherwise just rename the strategies). Define the $\mathcal{F}_\tau$-measurable random variable $Y := V_\varphi(\tau) - V_\psi(\tau)$ and consider the trading strategy ξ defined by

$$\xi(u) = \begin{cases} \varphi(u) - \psi(u), & u \leq \tau \\ \mathbf{1}_{A^c}(\varphi(u) - \psi(u)) + \mathbf{1}_A(Y\beta(\tau), 0, \ldots, 0), & \tau < u \leq T. \end{cases}$$

Then ξ is predictable and the self-financing condition (4.1) is clearly true for $t \neq \tau$, and for $t = \tau$ we have using that $\varphi, \psi \in \Phi$

$$\xi(\tau) \cdot S(\tau) = (\varphi(\tau) - \psi(\tau)) \cdot S(\tau) = V_\varphi(\tau) - V_\psi(\tau),$$

$$\begin{aligned} \xi(\tau+1) \cdot S(\tau) &= \mathbf{1}_{A^c}(\varphi(\tau+1) - \psi(\tau+1)) \cdot S(\tau) + \mathbf{1}_A Y\beta(\tau) S_0(\tau) \\ &= \mathbf{1}_{A^c}(\varphi(\tau) - \psi(\tau)) \cdot S(\tau) + \mathbf{1}_A (V_\varphi(\tau) - V_\psi(\tau))\beta(\tau)\beta^{-1}(\tau) \\ &= V_\varphi(\tau) - V_\psi(\tau). \end{aligned}$$

Hence ξ is a self-financing strategy with initial value equal to zero. Furthermore

$$\begin{aligned} V_\xi(T) &= \mathbf{1}_{A^c}(\varphi(T) - \psi(T)) \cdot S(T) + \mathbf{1}_A(Y\beta(\tau), 0, \ldots, 0) \cdot S(T) \\ &= \mathbf{1}_A Y\beta(\tau) S_0(T) \geq 0 \end{aligned}$$

and

$$\mathbb{P}\{V_\xi(T) > 0\} = \mathbb{P}\{A\} > 0.$$

Hence the market $\mathcal{M}$ contains an arbitrage opportunity with respect to the class Φ of self-financing strategies. But this contradicts the assumption that the market $\mathcal{M}$ is arbitrage-free. □

This uniqueness property allows us now to define the important concept of an arbitrage price process.

Definition 4.2.5. *Suppose the market $\mathcal{M}$ is arbitrage-free. Let X be any attainable contingent claim with time T maturity. Then the arbitrage price process $\pi_X(t)$, $0 \leq t \leq T$ or simply arbitrage price of X in $\mathcal{M}$ is given by the value process of any replicating strategy φ for X.*

The construction of hedging strategies that replicate the outcome of a contingent claim (for example a European option) is an important problem in both practical and theoretical applications. Hedging is central to the theory of option pricing. The classical arbitrage valuation models, such as the Black-Scholes model ([25]; see §4.6 and Chapter 6), depend on the idea that an option can be perfectly hedged using the underlying asset (in our case the assets of the market model $\mathcal{M}$), so making it possible to create a portfolio that replicates the option exactly. Hedging is also widely used to reduce risk, and the kinds of delta-hedging strategies implicit in the Black-Scholes model are used by participants in option markets. We will come back to hedging problems subsequently (e.g. §4.5 in the finite setting).

Analysing the arbitrage-pricing approach we observe that the derivation of the price of a contingent claim doesn't require any specific preferences of the agents other than nonsatiation, i.e. agents prefer more to less, which rules out arbitrage. So, the pricing formula for any attainable contingent claim must be independent of all preferences that do not admit arbitrage. In particular, an economy of risk-neutral investors must price a contingent claim in the same manner. This fundamental insight, due to Cox and Ross [45] in the case of a simple economy – a riskless asset and one risky asset - and in its general form due to Harrison and Kreps [114], simplifies the pricing formula enormously. In its general form the price of an attainable simple contingent claim is just the expected value of the discounted payoff with respect to an equivalent martingale measure.

Proposition 4.2.5. *The arbitrage price process of any attainable contingent claim X is given by the* **risk-neutral valuation formula**

$$\pi_X(t) = \beta(t)^{-1} \mathbb{E}^* (X\beta(T)|\mathcal{F}_t) \quad \forall t = 0, 1, \ldots, T, \tag{4.4}$$

where $\mathbb{E}^*$ is the expectation operator with respect to an equivalent martingale measure $\mathbb{P}^*$.

Proof. Since we assume the the market is arbitrage-free there exists (at least) an equivalent martingale measure $\mathbb{P}^*$. By Proposition 4.2.1 the discounted value process $\tilde{V}_\varphi$ of any self-financing strategy φ is a $\mathbb{P}^*$-martingale. So for any contingent claim X with maturity T and any replicating trading strategy $\varphi \in \Phi$ we have for each $t = 0, 1, \ldots, T$

$$\begin{aligned}
\pi_X(t) &= V_\varphi(t) = \beta(t)^{-1}\tilde{V}_\varphi(t) \\
&= \beta(t)^{-1} E^*(\tilde{V}_\varphi(T)|\mathcal{F}_t) && \text{(as } \tilde{V}_\varphi(t) \text{ is a } \mathbb{P}^*\text{-martingale)} \\
&= \beta(t)^{-1} E^*(\beta(T)V_\varphi(T)|\mathcal{F}_t) && \text{(undoing the discounting)} \\
&= \beta(t)^{-1} E^*(\beta(T)X|\mathcal{F}_t) && \text{(as } \varphi \text{ is a replicating strategy for } X).
\end{aligned}$$

□

4.3 Complete Markets: Uniqueness of Equivalent Martingale Measures

The last section made clear that attainable contingent claims can be priced using an equivalent martingale measure. In this section we will discuss the question of the circumstances under which all contingent claims are attainable. This would be a very desirable property of the market $\mathcal{M}$, because we would then have solved the pricing question (at least for contingent claims) completely. Since contingent claims are merely non-negative $\mathcal{F}_T$-measurable random variables in our setting, it should be no surprise that we can give a criterion in terms of probability measures. We start with:

Definition 4.3.1. *A market $\mathcal{M}$ is complete if every contingent claim is attainable, i.e. for every non-negative $\mathcal{F}_T$-measurable random variable $X \in \mathcal{X}^+$ there exists a replicating self-financing strategy $\varphi \in \Phi$ such that $V_\varphi(T) = X$.*

In the case of an arbitrage-free market $\mathcal{M}$ one can even insist on replicating contingent claims by an admissible strategy $\varphi \in \Phi_a$. Indeed, if φ is self-financing and $\mathbb{P}^*$ is an equivalent martingale measure under which discounted prices $\tilde{S}$ are $\mathbb{P}^*$-martingales (such $\mathbb{P}^*$ exist since $\mathcal{M}$ is arbitrage-free and we can hence use the no-arbitrage theorem (Theorem 4.2.1), $\tilde{V}_\varphi(t)$ is also a $\mathbb{P}^*$-martingale, being the martingale transform of the martingale $\tilde{S}$ by φ (see Proposition 4.2.1). So

$$\tilde{V}_\varphi(t) = E^*(\tilde{V}_\varphi(T)|\mathcal{F}_t) \quad (t = 0, 1, \ldots, T).$$

If φ replicates X, $V_\varphi(T) = X \geq 0$, so discounting, $\tilde{V}_\varphi(T) \geq 0$, so the above equation gives $\tilde{V}_\varphi(t) \geq 0$ for each t. Thus all the values at each time t are non-negative – not just the final value at time T – so φ is admissible.

Theorem 4.3.1 (Completeness Theorem). *An arbitrage-free market $\mathcal{M}$ is complete if and only if there exists a unique probability measure $\mathbb{P}^*$ equivalent to $\mathbb{P}$ under which discounted asset prices are martingales.*

Proof. '⇒': Assume that the arbitrage-free market $\mathcal{M}$ is complete. Then for any $\mathcal{F}_T$-measurable random variable $X \geq 0$ (contingent claim), there exists an admissible (so self-financing) strategy φ replicating X: $X = V_\varphi(T)$. As φ is self-financing, by Proposition 4.1.2,

$$\beta(T)X = \tilde{V}_\varphi(T) = V_\varphi(0) + \sum_{\tau=1}^{T} \varphi(\tau) \cdot \Delta\tilde{S}(\tau).$$

We know by the no-arbitrage theorem (Theorem 4.2.1) that an equivalent martingale measure $\mathbb{P}^*$ exists; we have to prove uniqueness. So, let $\mathbb{P}_1, \mathbb{P}_2$ be two such equivalent martingale measures. For $i = 1, 2$, $(\tilde{V}_\varphi(t))_{t=0}^T$ is a $\mathbb{P}_i$-martingale. So,

$$\mathbb{E}_i(\tilde{V}_\varphi(T)) = \mathbb{E}_i(\tilde{V}_\varphi(0)) = V_\varphi(0),$$

since the value at time zero is non-random ($\mathcal{F}_0 = \{\emptyset, \Omega\}$) and $\beta(0) = 1$. So

$$\mathbb{E}_1(\beta(T)X) = \mathbb{E}_2(\beta(T)X).$$

Since X is arbitrary, $\mathbb{E}_1, \mathbb{E}_2$ have to agree on integrating all non-negative integrands. Taking negatives and using linearity: they have to agree on non-positive integrands also. Splitting an arbitrary integrand into its positive and negative parts: they have to agree on all integrands. Now $\mathbb{E}_i$ is expectation (i.e. integration) with respect to the measure $\mathbb{P}_i$, and measures that agree on integrating all integrands must coincide. So $\mathbb{P}_1 = \mathbb{P}_2$, giving uniqueness as required.

'⇐': Assume that the arbitrage-free market $\mathcal{M}$ is incomplete: then there exists a non-attainable $\mathcal{F}_T$-measurable random variable $X \geq 0$ (a contingent claim). We may confine attention to self-financing strategies φ (which will in the replicating case be automatically admissible). By Proposition 4.1.3, we may confine attention to the risky assets $S_1, \dots, S_d$, as these suffice to tell us how to handle the numéraire S_0.

Consider the following set of random variables:

$$\tilde{\mathcal{V}} := \left\{ Y \in \mathcal{X} : Y = Y_0 + \sum_{t=1}^{T} \varphi(t) \cdot \Delta\tilde{S}(t),\ Y_0 \in \mathbb{R},\ \varphi \text{ predictable} \right\}.$$

(Recall that Y_0 is $\mathcal{F}_0$-measurable and set $\varphi = ((\varphi_1(t), \dots, \varphi_d(t))')_{t=1}^{T}$ with predictable components.) Then by the above reasoning, the discounted value $\beta(T)X$ does not belong to $\tilde{\mathcal{V}}$, so $\tilde{\mathcal{V}}$ is a *proper* subset of the set $\mathcal{X}$ of all random variables on Ω (which may be identified with $\mathbb{R}^{|\Omega|}$). Let $\mathbb{P}^*$ be a probability measure equivalent to $\mathbb{P}$ under which discounted prices are martingales (such $\mathbb{P}^*$ exist by the no-arbitrage theorem (Theorem 4.2.1). Define the scalar product

$$(Z, Y) \to \mathbb{E}^*(ZY)$$

on random variables on Ω. Since $\tilde{\mathcal{V}}$ is a proper subset, there exists a non-zero random variable Z orthogonal to $\tilde{\mathcal{V}}$ (since Ω is finite, $\mathbb{R}^{|\Omega|}$ is Euclidean: this is just Euclidean geometry). That is,

$$\mathbb{E}^*(ZY) = 0, \quad \forall\, Y \in \tilde{\mathcal{V}}.$$

Choosing the special $Y = 1 \in \tilde{\mathcal{V}}$ given by $\varphi_i(t) = 0, t = 1, 2, \dots, T;\ i = 1, \dots, d$ and $Y_0 = 1$ we find

$$\mathbb{E}^*(Z) = 0.$$

Write $\|X\|_\infty := \sup\{|X(\omega)| : \omega \in \Omega\}$, and define $\mathbb{P}^{**}$ by

$$\mathbb{P}^{**}(\{\omega\}) = \left(1 + \frac{Z(\omega)}{2\|Z\|_\infty}\right) \mathbb{P}^*(\{\omega\}).$$

By construction, $\mathbb{P}^{**}$ is equivalent to $\mathbb{P}^*$ (same null sets - actually, as $\mathbb{P}^* \sim \mathbb{P}$ and $\mathbb{P}$ has no non-empty null sets, neither do $\mathbb{P}^*, \mathbb{P}^{**}$). As Z is non-zero, $\mathbb{P}^{**}$ and $\mathbb{P}^*$ are *different*. Now

$$\begin{aligned}\mathbb{E}^{**}\left(\sum_{t=1}^{T} \varphi(t) \cdot \Delta\tilde{S}(t)\right) &= \sum_{\omega\in\Omega} \mathbb{P}^{**}(\omega) \left(\sum_{t=1}^{T} \varphi(t,\omega) \cdot \Delta\tilde{S}(t,\omega)\right) \\ &= \sum_{\omega\in\Omega} \left(1 + \frac{Z(\omega)}{2\|Z\|_\infty}\right) \mathbb{P}^*(\omega) \left(\sum_{t=1}^{T} \varphi(t,\omega) \cdot \Delta\tilde{S}(t,\omega)\right).\end{aligned}$$

The '1' term on the right gives

$$\mathbb{E}^*\left(\sum_{t=1}^{T} \varphi(t) \cdot \Delta\tilde{S}(t)\right),$$

which is zero since this is a martingale transform of the $\mathbb{P}^*$-martingale $\tilde{S}(t)$ (recall martingale transforms are by definition null at zero). The 'Z' term gives a multiple of the inner product

$$(Z, \sum_{t=1}^{T} \varphi(t) \cdot \Delta\tilde{S}(t)),$$

which is zero as Z is orthogonal to $\tilde{\mathcal{V}}$ and $\sum_{t=1}^{T} \varphi(t) \cdot \Delta\tilde{S}(t) \in \tilde{\mathcal{V}}$. By the martingale transform lemma (Lemma 3.4.1), $\tilde{S}(t)$ is a $\mathbb{P}^{**}$-martingale since φ is an arbitrary predictable process. Thus $\mathbb{P}^{**}$ is a second equivalent martingale measure, different from $\mathbb{P}^*$. So incompleteness implies non-uniqueness of equivalent martingale measures, as required. □

Martingale Representation. To say that every contingent claim can be replicated means that every $\mathbb{P}^*$-martingale (where $\mathbb{P}^*$ is the risk-neutral measure, which is unique) can be written, or *represented*, as a martingale transform (of the discounted prices) by a replicating (perfect-hedge) trading strategy φ. In stochastic-process language, this says that all $\mathbb{P}^*$-martingales can be *represented* as martingale transforms of discounted prices. Such martingale representation theorems hold much more generally, and are very important. For the Brownian case, see §5.8; for background, see [181, 220].

A further characterisation of completeness (due to Harrison and Pliska [115] and Kreps [145]) entirely specific to the finite case, may be obtained along the following lines.

Since we work in a finite probability space we can identify each $\mathcal{F}_t$ with a unique partition $\mathcal{P}_t$ of Ω, and at time t the investors know which cell of this partition contains the true state of the world, but they do not know more than this. The price process S is said to contain a *redundancy* if $\mathbb{P}(\alpha S_{t+1} = 0|A) = 1$ for some non-trivial vector α, some $t < T$, and some $A \in \mathcal{P}_t$. If such a redundancy exists, then there is an event A possible at time t which makes possession of some one security over the coming period completely equivalent to possession of a linear combination of the other securities over that same period. If no such circumstances exist, then we say that the securities are *nonredundant*. For each cell A of $\mathcal{P}_t, t = 0, 1, \dots, T-1$, let $K_t(A)$ be the number of cells of $\mathcal{P}_{t+1}$ which are contained in A. This might be called the *splitting index* of A ([64]; Duffie [70], p.37 calls this number the *spanning number*). Assuming the the market $\mathcal{M}$ is arbitrage-free and that the securities are nonredundant, we must have $K_t(A) \ge d + 1$ for all t and A. Using this we have the following characterisation of completeness for a finite security market.

Proposition 4.3.1. *If the securities are nonredundant, then the model is complete if and only if $K_t(A) = d + 1$ for all $A \in \mathcal{P}_t$ and $t = 0, 1, \dots, T-1$.*

We refer to Kreps [145] for a proof and further discussion (see Harrison and Pliska [115] for an outline). The heuristic message behind Proposition

4.3.1 is that in each circumstance A that may prevail at time t, investors must have available enough linearly independent securities to span the space of contingencies which may prevail at time $t+1$. For a model with many trading dates t and many states ω, completeness depends critically on the way uncertainty resolves itself over time, this being reflected by the splitting indices $K_t(A)$ (again we refer to [64, 72, 114, 115, 145] for further discussion).

4.4 The Fundamental Theorem of Asset Pricing: Risk Neutral Valuation

We summarise what we have achieved so far. We call a measure $\mathbb{P}^*$ under which discounted prices $\tilde{S}(t)$ are $\mathbb{P}^*$-martingales a *martingale measure*. Such a $\mathbb{P}^*$ equivalent to the actual probability measure P is called an *equivalent martingale measure*. Then:

- *No-arbitrage theorem (Theorem 4.2.1):* If the market is *arbitrage-free*, equivalent martingale measures $\mathbb{P}^*$ *exist*.
- *Completeness theorem (Theorem 4.3.1):* If the market is *complete* (all contingent claims can be replicated), equivalent martingale measures are *unique*.

Combining:

Theorem 4.4.1 (Fundamental Theorem of Asset Pricing). *In an arbitrage-free complete market $\mathcal{M}$, there exists a unique equivalent martingale measure $\mathbb{P}^*$.*

The term *fundamental theorem of asset pricing* was introduced by Dybvig and Ross [79]. It is used for theorems establishing the equivalence of an economic modelling condition such as no-arbitrage to the existence of the mathematical modelling condition existence of equivalent martingale measures.

Assume now that $\mathcal{M}$ is an arbitrage-free complete market and let $X \in \mathcal{X}^+$ (≥ 0, $\mathcal{F}_T$-measurable) be any contingent claim, φ an admissible strategy replicating it (which exists by completeness), then:

$$V_\varphi(T) = X.$$

As $\tilde{V}_\varphi(t)$ is the martingale transform of the $\mathbb{P}^*$-martingale $\tilde{S}(t)$ (by $\varphi(t)$), $\tilde{V}_\varphi(t)$ is a $\mathbb{P}^*$-martingale. So $V_\varphi(0)(=\tilde{V}_\varphi(0)) = \mathbb{E}^*(\tilde{V}_\varphi(T)) = \mathbb{E}^*(\beta(T)X)$, giving us the risk-neutral pricing formula

$$V_\varphi(0) = \mathbb{E}^*(\beta(T)X).$$

More generally, the same argument gives $\tilde{V}_\varphi(t) = \beta(t)V_\varphi(t) = \mathbb{E}^*(\beta(T)X|\mathcal{F}_t)$:

$$V_\varphi(t) = \beta(t)^{-1}\mathbb{E}^*(\beta(T)X|\mathcal{F}_t) \quad (t = 0, 1, \ldots, T). \tag{4.5}$$

It is natural to call $V_\varphi(0) = \pi_X(0)$ above the *arbitrage price* (or more exactly, *arbitrage-free price*) of the contingent claim X at time 0, and $V_X(t) = \pi_X(t)$ above the *arbitrage price* (or more exactly, *arbitrage-free price*) of the simple contingent claim X at time t. For, if an investor *sells* the claim X at time t for $V_X(t)$, he can follow strategy φ to replicate X at time T and clear the claim; an investor selling for this value is *perfectly hedged*. To sell the claim for *any other amount* would provide an arbitrage opportunity (as with the argument for put-call parity). We note that, to calculate prices as above, we need to know only:

1. Ω, the set of all possible states,
2. the σ-field $\mathcal{F}$ and the filtration (or information flow) $(\mathcal{F}_t)$,
3. $\mathbb{P}^*$.

We do not need to know the underlying probability measure $\mathbb{P}$ - only its null sets, to know what 'equivalent to $\mathbb{P}$' means (actually, in this finite model, there are no non-empty null-sets, so we do not need to know even this).

Now pricing of contingent claims is our central task, and for pricing purposes $\mathbb{P}^*$ is vital and $\mathbb{P}$ itself irrelevant. We thus may - and shall – focus attention on $\mathbb{P}^*$, which is called the *risk-neutral* probability measure. *Risk-neutrality* is the central concept of the subject and the underlying theme of this text. The concept of risk-neutrality is due in its modern form to Harrison and Pliska [115] in 1981 – though the idea can be traced back to actuarial practice much earlier (see Esscher [87] and also Gerber and Shiu [103]). Harrison and Pliska call $\mathbb{P}^*$ the *reference measure*; Björk [18] calls it the *risk-adjusted* or *martingale measure*; Dothan [64] uses *equilibrium price measure*. The term 'risk-neutral' reflects the $\mathbb{P}^*$-martingale property of the risky assets, since martingales model fair games (one can't win systematically by betting on a martingale).

To summarise, we have:

Theorem 4.4.2 (Risk-Neutral Pricing Formula). *In an arbitrage-free complete market $\mathcal{M}$, arbitrage prices of contingent claims are their discounted expected values under the risk-neutral (equivalent martingale) measure $\mathbb{P}^*$.*

There exist several variants and ramifications of the results we have presented so far.

Finite, Discrete Time; Finite Probability Space (our model)

Like Harrison and Pliska in their seminal paper [115] we used several results from functional analysis. Taqqu and Willinger [213] provide an approach based on probabilistic methods and allowing a geometric interpretation which yields a connection to linear programming. They analyse certain geometric properties of the sample paths of a given vector-valued stochastic process representing the different stock prices through time. They show that under the requirement that no arbitrage opportunities exist, the price increments between two periods can be converted to martingale differences (see Chapter 3) through an equivalent martingale measure. From a probabilistic point of view this provides a converse to the classical notion that 'one cannot win betting on a martingale' by saying 'if one cannot win betting on a process, then it must be a martingale under an equivalent martingale measure'. Furthermore, they give a characterisation of complete markets in terms of an extremal property of a probability measure in the convex set $\bar{\mathcal{P}}(\tilde{S})$ of martingale measures for $\tilde{S}$ (not necessarily equivalent to $\mathbb{P}$):

The market model $\mathcal{M}$ is complete under a measure $\boldsymbol{Q}$ on $(\Omega, \mathcal{F})$ if and only if $\boldsymbol{Q}$ is an extreme point of $\bar{\mathcal{P}}(\tilde{S})$ (i.e. $\boldsymbol{Q}$ cannot be expressed as a strictly convex combination of two distinct probability measures in $\bar{\mathcal{P}}(\tilde{S})$).

They also show that the problem of attainability of a simple contingent claim can be viewed and formulated as the 'dual problem' to finding a certain martingale measure for the price process $\tilde{S}$.

Finite, Discrete Time; General Probability Space

The no-arbitrage condition remains equivalent to the existence of an equivalent martingale measure. The first proof of this was given by Dalang, Morton and Willinger [50] using deep functional analytic methods (such as measurable selection and measure-decomposition theorems). There exist now several more accessible proofs, in particular by Schachermayer [197], using more elementary results from functional analysis (orthogonality arguments in properly chosen spaces, see also Kabanov and Kramkov [135]) and by Rogers [184], using a method which essentially comes down to maximising expected utility of gains from trade over all possible trading strategies.

Discrete Time; Infinite Horizon; General Probability Space

Under this setting the equivalence of no-arbitrage opportunities and existence of an equivalent martingale measure breaks down (see Back and Pliska [7] and Dalang et al. [50] for counterexamples). Introducing a weaker regularity concept than no-arbitrage, namely no free lunch with bounded risk – requiring an absolute bound on the maximal loss occurring in certain basic trading strategies (see [198] for an exact mathematical definition, [144] for related concepts) - Schachermayer [198] established the following beautiful result:

The condition no free lunch with bounded risk is equivalent to the existence of an equivalent martingale measure.

For a recent overview of variants of fundamental asset pricing theorems proved by probabilistic techniques, we refer the reader to [129]. We will not pursue these approaches further, but use our finite discrete-time and finite probability space setting to explore several models which are widely used in practice.

4.5 The Cox-Ross-Rubinstein Model

In this section we consider simple discrete-time financial market models. The development of the risk-neutral pricing formula is particularly clear in this setting since we require only elementary mathematical methods. The link to the fundamental economic principles of the arbitrage pricing method can be obtained equally straightforwardly. Moreover binomial models, by their very construction, give rise to simple and efficient numerical procedures. We start with the paradigm of all binomial models - the celebrated Cox-Ross-Rubinstein model [44].

4.5.1 Model Structure

We take $d = 1$, that is, our model consists of two basic securities. Recall that the essence of the relative pricing theory is to take the price processes of these basic securities as given and price secondary securities in such a way that no arbitrage is possible.

Our time horizon is T and the set of dates in our financial market model is $t = 0, 1, \dots, T$. Assume that the first of our given basic securities is a (riskless) bond or bank account B, with price process

$$B(t) = (1+r)^t, \quad t = 0, 1, \dots, T,$$

implying that the bond yields a riskless rate of return r in each time interval $[t, t+1]$. Furthermore, we have a risky asset (stock) S with price process

$$S(t+1) = \begin{cases} uS(t) & \text{with probability } p, \\ dS(t) & \text{with probability } 1-p, \end{cases} \qquad t = 0, 1, \dots, T-1$$

with $0 < d < u, S_0 \in \mathbb{R}_0^+$ (see Fig. 4.1 below).

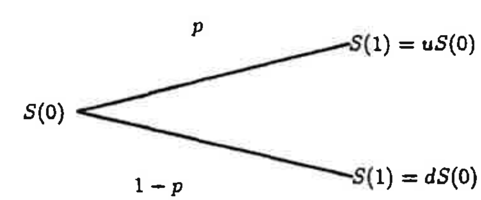

Fig. 4.1. One-step tree diagram

Alternatively we write this as

$$Z(t+1) := \frac{S(t+1)}{S(t)}, \quad t = 0, 1, \dots, T-1.$$

We set up a probabilistic model by considering the $Z(t)$, $t = 1, \dots, T$ as random variables defined on probability spaces $(\tilde{\Omega}_t, \tilde{\mathcal{F}}_t, \tilde{\mathbb{P}}_t)$ with

$$\begin{aligned}\tilde{\Omega}_t &= \tilde{\Omega} = \{d, u\},\\ \tilde{\mathcal{F}}_t &= \tilde{\mathcal{F}} = \mathcal{P}(\tilde{\Omega}) = \{\emptyset, \{d\}, \{u\}, \tilde{\Omega}\},\\ \tilde{\mathbb{P}}_t &= \tilde{\mathbb{P}} \text{ with } \tilde{\mathbb{P}}(\{u\}) = p, \ \tilde{\mathbb{P}}(\{d\}) = 1-p, \ p \in (0,1).\end{aligned}$$

On these probability spaces we define

$$Z(t, u) = u \quad \text{and} \quad Z(t, d) = d, \quad t = 1, 2, \dots, T.$$

Our aim, of course, is to define a probability space on which we can model the basic securities (B, S). Since we can write the stock price as

$$S(t) = S(0) \prod_{\tau=1}^{t} Z(\tau), \quad t = 1, 2, \dots, T,$$

the above definitions suggest using as the underlying probabilistic model of the financial market the *product space* $(\Omega, \mathcal{F}, \mathbb{P})$ (see e.g. [218] ch. 8), i.e.

$$\Omega = \tilde{\Omega}_1 \times \dots \times \tilde{\Omega}_T = \tilde{\Omega}^T = \{d, u\}^T,$$

with each $\omega \in \Omega$ representing the successive values of $Z(t)$, $t = 1, 2, \dots, T$. Hence each $\omega \in \Omega$ is a T-tuple $\omega = (\tilde{\omega}_1, \dots, \tilde{\omega}_T)$ and $\tilde{\omega}_t \in \tilde{\Omega} = \{d, u\}$. For the σ-algebra we use $\mathcal{F} = \mathcal{P}(\Omega)$ and the probability measure is given by

$$\mathbb{P}(\{\omega\}) = \tilde{\mathbb{P}}_1(\{\omega_1\}) \times \dots \times \tilde{\mathbb{P}}_T(\{\omega_T\}) = \tilde{\mathbb{P}}(\{\omega_1\}) \times \dots \times \tilde{\mathbb{P}}(\{\omega_T\}).$$

The role of a product space is to model independent replication of a random experiment. The $Z(t)$ above are two-valued random variables, so can be thought of as tosses of a biased coin; we need to build a probability space on which we can model a succession of such independent tosses.

Now we redefine (with a slight abuse of notation) the $Z(t)$, $t = 1, \dots, T$ as random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ as (the tth projection)

$$Z(t, \omega) = Z(t, \omega_t).$$

Observe that by this definition (and the above construction) $Z_1, \dots, Z_T$ are independent and identically distributed with

$$\mathbb{P}(Z(t) = u) = p = 1 - \mathbb{P}(Z(t) = d).$$

To model the flow of information in the market we use the obvious filtration

$$\begin{aligned}\mathcal{F}_0 &= \{\emptyset, \Omega\} && \text{(trivial } \sigma\text{-field)},\\ \mathcal{F}_t &= \sigma(Z(1), \dots, Z(t)) = \sigma(S(1), \dots, S(t)),\\ \mathcal{F}_T &= \mathcal{F} = \mathcal{P}(\Omega) && \text{(class of all subsets of } \Omega).\end{aligned}$$

This construction emphasises again that a multi-period model can be viewed as a sequence of single-period models. Indeed, in the Cox-Ross-Rubinstein case we use identical and independent single-period models. As we will see in the sequel this will make the construction of equivalent martingale measures relatively easy. Unfortunately we can hardly defend the assumption of independent and identically distributed price movements at each time period in practical applications. This is the reason why we introduce models of Markovian type in §4.7.

Remark 4.5.1. We used this example to show explicitly how to construct the underlying probability space. Having done this in full once, we will from now on feel free to take for granted the existence of an appropriate probability space on which all relevant random variables can be defined.

4.5.2 Risk-Neutral Pricing

We now turn to the pricing of derivative assets in the Cox-Ross-Rubinstein market model. To do so we first have to discuss whether the Cox-Ross-Rubinstein model is arbitrage-free and complete.

To answer these questions we have, according to our fundamental theorems (Theorems 4.2.1 and 4.3.1), to understand the structure of equivalent martingale measures in the Cox-Ross-Rubinstein model. In trying to do this we use (as is quite natural and customary) the bond price process $B(t)$ as numéraire.

Our first task is to find an equivalent martingale measure of the same class as $\mathbb{P}$, i.e. a probability measure $\mathbf{Q}$ defined as a product measure via a measure $\tilde{\mathbf{Q}}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ such that $\tilde{\mathbf{Q}}(\{u\}) = q$ and $\tilde{\mathbf{Q}}(\{d\}) = 1 - q$. We call the set of all measures of that type $\mathcal{P}$. We have:

Proposition 4.5.1. *(i) A martingale measure $\mathbf{Q} \in \mathcal{P}$ for the discounted stock price $\tilde{S}$ exists if and only if*

$$d < 1 + r < u. \tag{4.6}$$

(ii) If equation (4.6) holds true, then there is a unique such measure in $\mathcal{P}$ characterised by

$$q = \frac{1 + r - d}{u - d}. \tag{4.7}$$

Proof. Since $S(t) = \tilde{S}(t)B(t) = \tilde{S}(t)(1+r)^t$, we have $Z(t+1) = S(t+1)/S(t) = (\tilde{S}(t+1)/\tilde{S}(t))(1+r)$. So, the discounted price $(\tilde{S}(t))$ is a $\mathbf{Q}$-martingale if and only if for $t = 0, 1, \dots, T-1$

$$\begin{aligned}&\mathbb{E}^{\mathbf{Q}}[\tilde{S}(t+1)|\mathcal{F}_t] = \tilde{S}(t) \Leftrightarrow \mathbb{E}^{\mathbf{Q}}[(\tilde{S}(t+1)/\tilde{S}(t))|\mathcal{F}_t] = 1\\ &\Leftrightarrow \mathbb{E}^{\mathbf{Q}}[Z(t+1)|\mathcal{F}_t] = 1 + r.\end{aligned}$$

But $Z(1), \dots, Z(T)$ are mutually independent and hence $Z(t+1)$ is independent of $\mathcal{F}_t = \sigma(Z(1), \dots, Z(t))$. So

$$1 + r = \mathbb{E}^{Q}(Z(t+1)|\mathcal{F}_t) = \mathbb{E}^{Q}(Z(t+1)) = uq + d(1-q)$$

is a weighted average of u and d; this can be $1+r$ if and only if $1+r \in [d, u]$. As Q is to be *equivalent* to $\mathbb{P}$ and $\mathbb{P}$ has no non-empty null sets, $r = d-1, u-1$ are excluded and (4.6) is proved.

To prove uniqueness and to find the value of q we simply observe that under (4.6)

$$u \times q + d \times (1-q) = 1 + r$$

has a unique solution. Solving it for q leads to the above formula. □

From now on we assume that (4.6) holds true. Using the above Proposition we immediately get:

Corollary 4.5.1. *The Cox-Ross-Rubinstein model is arbitrage-free.*

Proof. By Proposition 4.5.1 there exists an equivalent martingale measure and this is by the no-arbitrage theorem (Theorem 4.2.1) enough to guarantee that the Cox-Ross-Rubinstein model is free of arbitrage. □

To answer the question of market completeness we have to examine the structure of equivalent martingale measures. We already know that there is a unique martingale measure in the class $\mathcal{P}$, but can we find other equivalent martingale measures outside $\mathcal{P}$? Examining the structure of the underlying measure space (product space) again we can show that every probability measure is a product measure, and hence all probability measures are of class $\mathcal{P}$. Using this result we conclude the following:

Proposition 4.5.2. *The Cox-Ross-Rubinstein model is complete.*

Proof. Since by Proposition 4.5.1 the equivalent martingale measure is unique, the claim follows from the completeness theorem (Theorem 4.3.1). □

We translate the above measure-theoretic argument to economic language. Since every probability measure is a product measure uniqueness of the equivalent martingale measure on the product space is equivalent to uniqueness of the individual martingale measures on the component space. (Observe that this is true for general product space situations – it merely says that for independent random variables, the individual distributions determine the joint distribution.) We thus have the following result.

Corollary 4.5.2. *The multi-period model is complete if and only if every underlying single-period model is complete.*

We can now use the risk-neutral valuation formula to price *every* contingent claim in the Cox-Ross-Rubinstein model.

Proposition 4.5.3. *The arbitrage price process of a contingent claim X in the Cox-Ross-Rubinstein model is given by*

$$\pi_X(t) = B(t)\mathbb{E}^*(X/B(T)|\mathcal{F}_t) \quad \forall t = 0, 1, \dots, T,$$

where $\mathbb{E}^*$ is the expectation operator with respect to the unique equivalent martingale measure $\mathbb{P}^*$ characterised by $p^* = (1+r-d)/(u-d)$.

Proof. This follows directly from Proposition 4.2.4 since the Cox-Ross-Rubinstein model is arbitrage-free and complete. □

As the most prominent example of an contingent claim, we now price the European call option.

Corollary 4.5.3. *Consider a European call option with expiry T and strike price K written on (one share of) the stock S. The arbitrage price process $C(t)$, $t = 0, 1, \dots, T$ of the option is given by (set $\tau = T - t$)*

$$C(t) = (1+r)^{-\tau} \sum_{j=0}^{\tau} \binom{\tau}{j} p^{*j}(1-p^*)^{\tau-j}(S(t)u^j d^{\tau-j} - K)^+. \tag{4.8}$$

Proof. Recall that

$$S(t) = S(0) \prod_{j=1}^{t} Z(j), \quad t = 1, 2, \dots, T.$$

By Proposition 4.5.3 the price $C(t)$ of a call option with strike price K at time t is

$$\begin{aligned} C(t) &= (1+r)^{-(T-t)}\mathbb{E}^*[(S(T)-K)^+|\mathcal{F}_t] \\ &= (1+r)^{-(T-t)}\mathbb{E}^*\left[\left(S(t)\prod_{i=t+1}^{T} Z(i) - K\right)^+ \middle| \mathcal{F}_t\right] \\ &= (1+r)^{-(T-t)}\mathbb{E}^*\left[\left(S(t)\prod_{i=t+1}^{T} Z(i) - K\right)^+\right] \\ &= (1+r)^{-(T-t)}\sum_{j=0}^{T-t}\binom{T-t}{j}p^{*j}(1-p^*)^{T-t-j}(S(t)u^j d^{T-t-j} - K)^+. \end{aligned}$$

We used the extension of property 7 (role of independence) of conditional expectations from §2.6 in the next-to-last equality. It is applicable since $S(t)$ is $\mathcal{F}_t$-measurable, $Z(t+1), \dots, Z(T)$ are independent of $\mathcal{F}_t$ and $(x-K)^+$ is a non-negative function. Here j, $T-t-j$ are the numbers of times $Z(i)$ takes the two possible values d, u. □

For a European put option, we can either argue similarly or use put-call parity (§1.3).

4.5.3 Hedging

Since the Cox-Ross-Rubinstein model is complete we can find unique hedging strategies for replicating contingent claims. Recall that this means we can find a portfolio $\varphi(t) = (\varphi_0(t), \varphi_1(t))$, φ predictable, such that the value process $V_\varphi(t) = \varphi_0(t)B(t) + \varphi_1(t)S(t)$ satisfies

$$\Pi_X(t) = V_\varphi(t), \quad \text{for all } t = 0, 1, \ldots, T.$$

Using the bond as numéraire we get the discounted equation

$$\tilde{\Pi}_X(t) = \tilde{V}_\varphi(t) = \varphi_0(t) + \varphi_1(t)\tilde{S}(t), \quad \text{for all } t = 0, 1, \ldots, T.$$

By the pricing formula, Proposition 4.5.3, we know the arbitrage price process and using the restriction of predictability of φ, this leads to a unique replicating portfolio process φ. We can compute this portfolio process at any point in time as follows. The equation $\tilde{\Pi}_X(t) = \varphi_0(t) + \varphi_1(t)\tilde{S}(t)$ has to be true for each $\omega = (\omega_1, \ldots, \omega_t, \ldots, \omega_T) \in \Omega$ and each $t = 1, \ldots, T$ (observe that $\pi_X(0) = V_\varphi(0) = \varphi_0(1) + \varphi_1(1)S_1(0)$). Given such a t we only can use information up to (and including) time $t-1$ to ensure that φ is predictable. Therefore we know the coordinates $\omega_1, \ldots, \omega_{t-1}$, but ω_t can only have one of the values d or u. This leads to the following system of equations, which can be solved for $\varphi_0(t)$ and $\varphi_1(t)$ uniquely. With a slight abuse of notation we write $\bar{\omega} = (\omega_1, \ldots, \omega_{t-1})$, $\bar{\omega}u = (\omega_1, \ldots, \omega_{t-1}, u)$ and $\bar{\omega}d = (\omega_1, \ldots, \omega_{t-1}, d)$ (recall that since $S(t)$ and $\Pi_X(t)$ are $\mathcal{F}_t$-measurable they are constant on the corresponding partition $\mathcal{P}_t$ of Ω, which in our current setting is just given by the first t coordinates of ω, and the same argument applies for φ). Then

$$\begin{aligned}\tilde{\Pi}_X(t, \bar{\omega}u) &= \varphi_0(t, \bar{\omega}) + \varphi_1(t, \bar{\omega})\tilde{S}(t, \bar{\omega}u),\\ \tilde{\Pi}_X(t, \bar{\omega}d) &= \varphi_0(t, \bar{\omega}) + \varphi_1(t, \bar{\omega})\tilde{S}(t, \bar{\omega}d).\end{aligned}$$

The solution is given by

$$\begin{aligned}\varphi_0(t, \bar{\omega}) &= \frac{\tilde{S}(t, \bar{\omega}u)\tilde{\Pi}_X(t, \bar{\omega}d) - \tilde{S}(t, \bar{\omega}d)\tilde{\Pi}_X(t, \bar{\omega}u)}{\tilde{S}(t, \bar{\omega}u) - \tilde{S}(t, \bar{\omega}d)}\\ &= \frac{u\tilde{\Pi}_X(t, \bar{\omega}d) - d\tilde{\Pi}_X(t, \bar{\omega}u)}{u-d},\\ \varphi_1(t, \bar{\omega}) &= \frac{\tilde{\Pi}_X(t, \bar{\omega}u) - \tilde{\Pi}_X(t, \bar{\omega}d)}{\tilde{S}(t, \bar{\omega}u) - \tilde{S}(t, \bar{\omega}d)}\\ &= \frac{\tilde{\Pi}_X(t, \bar{\omega}u) - \tilde{\Pi}_X(t, \bar{\omega}d)}{\tilde{S}(t-1, \bar{\omega})(u-d)}.\end{aligned}$$

Observe that we only need to know $\bar{\omega}$ to compute $\varphi(t)$, hence φ is predictable. We make this rather abstract construction more transparent by constructing

the hedge portfolio for the European call. We use the following notation. Write

$$C(t, x) := (1+r)^{-(T-t)} \sum_{j=0}^{T-t} \binom{T-t}{j} p^{*j}(1-p^*)^{T-t-j}(xu^j d^{T-t-j} - K)^+.$$

Then $C(t, x)$ is value of the call at time t given that $S(t) = x$.

Proposition 4.5.4. *The perfect hedging strategy $\varphi = (\varphi_0, \varphi_1)$ replicating the European call option with time of expiry T and strike price K is given by*

$$\begin{aligned}\varphi_1(t) &= \frac{C(t, S(t-1)u) - C(t, S(t-1)d)}{S(t-1)(u-d)},\\ \varphi_0(t) &= \frac{uC(t, S(t-1)d) - dC(t, S(t-1)u)}{(1+r)^t(u-d)}.\end{aligned}$$

Proof. $C(t, x)$ must be the value of the portfolio at time t if the strategy $\varphi = (\varphi(t))$ replicates the claim:

$$\varphi_0(t)(1+r)^t + \varphi_1(t)S(t) = C(t, S(t)).$$

Now $S(t) = S(t-1)Z(t) = S(t-1)u$ or $S(t-1)d$, so:

$$\varphi_0(t)(1+r)^t + \varphi_1(t)S(t-1)u = C(t, S(t-1)u),$$

$$\varphi_0(t)(1+r)^t + \varphi_1(t)S(t-1)d = C(t, S(t-1)d).$$

Subtract:

$$\varphi_1(t)S(t-1)(u-d) = C(t, S(t-1)u) - C(t, S(t-1)d).$$

So $\varphi_1(t)$ in fact depends only on $S(t-1)$, thus yielding the predictability of φ, and

$$\varphi_1(t) = \frac{C(t, S(t-1)u) - C(t, S(t-1)d)}{S(t-1)(u-d)}.$$

Using any of the equations in the above system and solving for $\varphi_0(t)$ completes the proof. □

Notice that the numerator in the equation for $\varphi_1(t)$ is the difference of two values of $C(t, x)$, with the larger value of x in the first term (recall $u > d$). When the payoff function $C(t, x)$ is an increasing function of x, as for the European call option considered here, this is non-negative. In this case, the Proposition gives $\varphi_1(t) \geq 0$: the replicating strategy does not involve short-selling. We record this as:

Corollary 4.5.4. *When the payoff function is a non-decreasing function of the asset price $S(t)$, the perfect-hedging strategy replicating the claim does not involve short-selling of the risky asset.*

If we do not use the pricing formula from Proposition 4.5.3 (i.e. the information on the price process), but only the final values of the option (or more generally of a contingent claim) we are still able to compute the arbitrage price and to construct the hedging portfolio by backward induction. In essence this is again only applying the one-period calculations of §1.4 for each time interval and each state of the world. We outline this procedure for the European call starting with the last period $[T-1,T]$. We have to choose a replicating portfolio $\varphi(T) = (\varphi_0(T), \varphi_1(T)$ based on the information available at time $T-1$ (and so $\mathcal{F}_{T-1}$-measurable). So for each $\omega \in \Omega$ the following equation has to hold:

$$C(T,\omega) = \varphi_0(T,\omega)B(T,\omega) + \varphi_1(T,\omega)S(T,\omega).$$

Given the information $\mathcal{F}_{T-1}$ we know all but the last coordinate of ω, and this gives rise to two equations (with the same notation as above $\bar{\omega}$ denotes the first $T-1$ coordinates):

$$C(T,\bar{\omega}u) = \varphi_0(T,\bar{\omega})(1+r)^T + \varphi_1(T,\bar{\omega})S(T,\bar{\omega}u),$$

$$C(T,\bar{\omega}d) = \varphi_0(T,\bar{\omega})(1+r)^T + \varphi_1(T,\bar{\omega})S(T,\bar{\omega}d).$$

Since we know the payoff structure of the call at time T, e.g. $C(T,\bar{\omega}u) = (uS(T-1,\bar{\omega}) - K)^+$ and $C(T,\bar{\omega}d) = (dS(T-1,\bar{\omega}) - K)^+$, we can solve the above system and obtain

$$\varphi_1(T,\bar{\omega}) = \frac{(uS(T-1,\bar{\omega}) - K)^+ - (dS(T-1,\bar{\omega}) - K)^+}{S(T-1,\bar{\omega})(u-d)},$$

$$\varphi_0(T,\bar{\omega}) = \frac{u(dS(T-1,\bar{\omega}) - K)^+ - d(uS(T-1,\bar{\omega}) - K)^+}{(1+r)(u-d)}.$$

Using this portfolio one can compute the arbitrage price of the European call at time $T-1$ in state of the world $\bar{\omega}$ as

$$C(T-1,\bar{\omega}) = \varphi_0(T,\bar{\omega})(1+r)^{T-1} + \varphi_1(T,\bar{\omega})S(T-1,\bar{\omega}).$$

Now the arbitrage prices at time $T-1$ are known and one can repeat the procedure to successively compute the prices at $T-2,\dots,1,0$.

The advantage of our risk-neutral pricing procedure over this approach is that we have a single formula for the price of the contingent claim at all times t at once, and don't have to go to a backwards induction only to compute a price at a special time t.

4.5.4 Comparison With the General Arbitrage Bounds

Using the pricing formula for the European call we see that the price process depends on the current stock price $S(t)$, the strike price K, the expiry date

T, the interest rate r and (hidden in u,d) the volatility of the stock price. We thus have in accordance with §1.3 the determining factors of the European call price explicit in formula (4.8) and can examine their effects explicitly. We use the notation $C(F)$ to mean that we consider the price C of the call as a function of the factor F, keeping all others fixed.

Changes in Option Price Relative to the Underlying Stock Price. If the current stock price S is changed by an amount ΔS we see that

$$\begin{aligned}\Delta C &= C(S+\Delta S) - C(S)\\ &= (1+r)^{-T}\sum_{j=0}^{T}\left(\binom{T}{j}p^{*j}(1-p^*)^{T-j}\right.\\ &\quad \left.\{((S+\Delta S)u^jd^{T-j} - K)^+ - (Su^jd^{T-j} - K)^+\}\right)\\ &\geq 0,\end{aligned}$$

since the function $(x-K)^+$ is increasing in x. Hence $C(S+\Delta S) \geq C(S)$ and the value of the option increases as the underlying stock price increases, in accordance with our findings in §1.3.

The quotient $\Delta C/\Delta S$ measures the change in the option price for an infinitesimal change in the stock price, keeping everything else fixed. If $\Delta S \to 0$ this quotient converges (under conditions of regularity) to $\partial C/\partial S$ and is therefore often referred to as the option's delta.

Looking at Proposition 4.5.4 we see the relation of the delta to the hedging portfolio for the option. The number of stocks to be held in the portfolio during period $(t-1,t]$ is given by the difference quotient

$$\varphi_1(t) = \frac{C(t,S(t-1)u) - C(T,S(t-1)d)}{S(t-1)u - S(t-1)d}$$

and this quotient is called the hedge ratio. This is an expression of the above type and we see that the number of stocks in the hedge portfolio is given approximately by the corresponding derivative, the option's delta.

Changes in Option Price Relative to the Strike Price K. Since the strike price K is only present in the expression $(x-K)^+$ and this function is non-increasing in K, we immediately see that the options value is a non-increasing function of the strike price.

Formula (4.8) can be used to discuss the other determing factors of the option price in a similar (although quite tedious) manner.

4.6 Binomial Approximations

Suppose we observe financial assets during a continuous time period $[0,T]$. To construct a stochastic model of the price processes of these assets (to, e.g. value contingent claims) one basically has two choices: one could model the processes as continuous-time stochastic processes (for which the theory outlined in Chapter 5 is needed and which is done in Chapter 6) or one could construct a sequence of discrete-time models in which the continuous-time price processes are approximated by discrete-time stochastic processes in a suitable sense. We describe the the second approach now by examining the asymptotic properties of a sequence of Cox-Ross-Rubinstein models.

4.6.1 Model Structure

We assume that all random variables subsequently introduced are defined on a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$. As in §4.5 we want to model two assets, a riskless bond B and a risky stock S, which we now observe in a continuous-time interval $[0,T]$. To transfer the continuous-time framework into a binomial structure we make the following adjustments. Looking at the nth Cox-Ross-Rubinstein model in our sequence, there is a prespecified number k_n of trading dates. We set $\Delta_n = T/k_n$ and divide $[0,T]$ in k_n subintervals of length Δ_n, namely $I_j = [j\Delta_n, (j+1)\Delta_n]$, $j = 0, \ldots, k_n - 1$. We suppose that trading occurs only at the equidistant time points $t_{n,j} = j\Delta_n$, $j = 0, \ldots, k_n - 1$. We fix r_n as the riskless interest rate over each interval I_j, and hence the bond process (in the nth model) is given by

$$B(t_{n,j}) = (1+r_n)^j, \quad j = 0, \ldots, k_n.$$

In the continuous-time model we compound continuously with spot rate $r \geq 0$ and hence the bond price process $B(t)$ is given by $B(t) = e^{rt}$. In order to approximate this process in the discrete-time framework, we choose r_n such that

$$1 + r_n = e^{r\Delta_n}. \tag{4.9}$$

With this choice we have for any $j = 0, \ldots, k_n$ that $(1+r_n)^j = \exp(rj\Delta_n) = \exp(rt_{n,j})$. Thus we have approximated the bond process exactly at the time points of the discrete model.

Next we model the one-period returns $S(t_{n,j+1})/S(t_{n,j})$ of the stock by a family of random variables $Z_{n,i}; i = 1, \ldots, k_n$ taking values $\{d_n, u_n\}$ with

$$\mathbb{P}(Z_{n,i} = u_n) = p_n = 1 - \mathbb{P}(Z_{n,i} = d_n)$$

for some $p_n \in (0,1)$ (which we specify later). With these $Z_{n,j}$ we model the stock price process S_n in the nth Cox-Ross-Rubinstein model as

$$S_n(t_{n,j}) = S_n(0) \prod_{i=1}^{j} Z_{n,i}, \quad j = 0, 1, \ldots, k_n.$$

With the specification of the one-period returns we get a complete description of the discrete dynamics of the stock price process in each Cox-Ross-Rubinstein model. We call such a finite sequence $Z_n = (Z_{n,i})_{i=1}^{k_n}$ a

lattice or *tree*. The parameters u_n, d_n, p_n, k_n differ from lattice to lattice, but remain constant throughout a specific lattice. In the triangular array $(Z_{n,i})$, $i = 1, \ldots, k_n$; $n = 1, 2, \ldots$ we assume that the random variables are row-wise independent (but we allow dependence between rows). The approximation of a continuous-time setting by a sequence of lattices is called the lattice approach.

It is important to stress that for each n we get a different discrete stock price process $S_n(t)$ and that in general these processes do not coincide on common time points (and are also different from the price process $S(t)$).

Turning back to a specific Cox-Ross-Rubinstein model, we now have as in §4.5 a discrete-time bond and stock price process. We want arbitrage-free financial market models and therefore have to choose the parameters u_n, d_n, p_n accordingly. An arbitrage-free financial market model is guaranteed by the existence of an equivalent martingale measure, and by Proposition 4.5.1 (i) the (necessary and) sufficient condition for that is

$$d_n < 1 + r_n < u_n.$$

The risk-neutrality approach implies that the expected (under an equivalent martingale measure) one-period return must equal the one-period return of the riskless bond and hence we get (see Proposition 4.5.1(ii))

$$p_n = \frac{(1+r_n) - d_n}{u_n - d_n}. \tag{4.10}$$

So the only parameters to choose freely in the model are u_n and d_n. In the next sections we consider some special choices.

4.6.2 The Black-Scholes Option Pricing Formula

We now choose the parameters in the above lattice approach in a special way. Assuming the risk-free rate of interest r as given, we have by (4.9) $1 + r_n = e^{r\Delta_n}$, and the remaining degrees of freedom are resolved by choosing u_n and d_n. We use the following choice:

$$u_n = e^{\sigma\sqrt{\Delta_n}}, \quad \text{and} \quad d_n = u_n^{-1} = e^{-\sigma\sqrt{\Delta_n}}.$$

By condition (4.10) the risk-neutral probabilities for the corresponding single period models are given by

$$p_n^* = \frac{1 + r_n - d_n}{u_n - d_n} = \frac{e^{r\Delta_n} - e^{-\sigma\sqrt{\Delta_n}}}{e^{\sigma\sqrt{\Delta_n}} - e^{-\sigma\sqrt{\Delta_n}}}.$$

We can now price contingent claims in each Cox-Ross-Rubinstein model using the expectation operator with respect to the (unique) equivalent martingale measure characterised by the probabilities p_n^* (compare §4.5.2). In particular we can compute the price $C(t)$ at time t of a European call on the stock S with strike K and expiry T by formula (4.8) of Corollary 4.5.3. Let us reformulate this formula slightly. We define

$$a_n = \min\{j \in \mathbb{N}_0 | S(0)u_n^j d_n^{k_n-j} > K\}. \tag{4.11}$$

Then we can rewrite the pricing formula (4.8) for $t=0$ in the setting of the nth Cox-Ross-Rubinstein model as

$$\begin{aligned} C(0) &= (1+r_n)^{-k_n}\sum_{j=a_n}^{k_n}\binom{k_n}{j}p_n^{*j}(1-p_n^*)^{k_n-j}(S(0)u^j d^{k_n-j}-K) \\ &= S(0)\left[\sum_{j=a_n}^{k_n}\binom{k_n}{j}\left(\frac{p_n^* u_n}{1+r_n}\right)^j\left(\frac{(1-p_n^*)d_n}{1+r_n}\right)^{k_n-j}\right] \\ &\quad -(1+r_n)^{-k_n}K\left[\sum_{j=a_n}^{k_n}\binom{k_n}{j}p_n^{*j}(1-p_n^*)^{k_n-j}\right]. \end{aligned}$$

Denoting the binomial cumulative distribution function with parameters (n,p) as $B^{n,p}(.)$ we see that the second bracketed expression is just

$$\bar{B}^{k_n,p_n^*}(a_n) = 1 - B^{k_n,p_n^*}(a_n).$$

Also the first bracketed expression is $\bar{B}^{k_n,\hat{p}_n}(a_n)$ with

$$\hat{p}_n = \frac{p_n^* u_n}{1+r_n}.$$

That $\hat{p}_n$ is indeed a probability can be shown straightforwardly. Using this notation we have in the nth Cox-Ross-Rubinstein model for the price of a European call at time $t=0$ the following formula:

$$C_n(0) = S_n(0)\bar{B}^{k_n,\hat{p}_n}(a_n) - K(1+r_n)^{-k_n}\bar{B}^{k_n,p_n^*}(a_n). \tag{4.12}$$

(We stress again that the underlying is $S_n(t)$, dependent on n, but $S_n(0) = S(0)$ for all n.) We now look at the limit of this expression.

Proposition 4.6.1. *We have the following limit relation:*

$$\lim_{n\to\infty} C_n(0) = C_{BS}(0)$$

with $C_{BS}(0)$ given by the Black-Scholes formula (we use $S = S(0)$ to ease the notation)

$$C_{BS}(0) = SN(d_1(S,T)) - Ke^{-rT}N(d_2(S,T)). \tag{4.13}$$

The functions $d_1(s,t)$ and $d_2(s,t)$ are given by

$$d_1(s,t) = \frac{\log(s/K) + (r+\frac{\sigma^2}{2})t}{\sigma\sqrt{t}},$$

$$d_2(s,t) = d_1(s,t) - \sigma\sqrt{t} = \frac{\log(s/K) + (r-\frac{\sigma^2}{2})t}{\sigma\sqrt{t}}$$

and $N(.)$ is the standard normal cumulative distribution function (§2.3).

Proof. Since $S_n(0) = S$ (say) all we have to do to prove the proposition is to show

(i) $\lim_{n\to\infty} \bar{B}^{k_n,\hat{p}_n}(a_n) = N(d_1(S,T))$,

(ii) $\lim_{n\to\infty} \bar{B}^{k_n,p_n^*}(a_n) = N(d_2(S,T))$.

These statements involve the convergence of distribution functions and we use the theory outlined in §§2.7, 2.8, 2.9.

To show (i) we interpret

$$\bar{B}^{k_n,\hat{p}_n}(a_n) = \mathbb{P}(a_n \le Y_n \le k_n)$$

with (Y_n) a sequence of random variables distributed according to the binomial law with parameters $(k_n,\hat{p}_n)$. We normalise Y_n to

$$\tilde{Y}_n = \frac{Y_n - \mathbb{E}(Y_n)}{\sqrt{Var(Y_n)}} = \frac{Y_n - k_n\hat{p}_n}{\sqrt{k_n\hat{p}_n(1-\hat{p}_n)}} = \frac{\sum_{j=1}^{k_n}(B_{j,n}-\hat{p}_n)}{\sqrt{k_n\hat{p}_n(1-\hat{p}_n)}},$$

where $B_{j,n}$, $j = 1,\dots,k_n$; $n = 1,2,\dots$ are row-wise independent Bernoulli random variables with parameter $\hat{p}_n$. Now using the central limit theorem (Theorem 2.9.2) we know that for $\alpha_n \to \alpha$, $\beta_n \to \beta$ we have

$$\lim_{n\to\infty} \mathbb{P}(\alpha_n \le \tilde{Y}_n \le \beta_n) = N(\beta) - N(\alpha).$$

By definition we have

$$\mathbb{P}(a_n \le Y_n \le k_n) = \mathbb{P}\left(\alpha_n \le \tilde{Y}_n \le \beta_n\right)$$

with

$$\alpha_n = \frac{a_n - k_n\hat{p}_n}{\sqrt{k_n\hat{p}_n(1-\hat{p}_n)}} \quad \text{and } \beta_n = \frac{k_n(1-\hat{p}_n)}{\sqrt{k_n\hat{p}_n(1-\hat{p}_n)}}.$$

Using the following limiting relations:

$$\lim_{n\to\infty}\hat{p}_n = \frac{1}{2}, \quad \lim_{n\to\infty} k_n(1-2\hat{p}_n)\sqrt{\Delta_n} = -T\left(\frac{r}{\sigma}+\frac{\sigma}{2}\right),$$

and the defining relation for a_n, formula (4.11), we get

$$\lim_{n\to\infty}\alpha_n = \lim_{n\to\infty}\frac{\frac{\log(K/S)+k_n\sigma\sqrt{\Delta_n}}{2\sigma\sqrt{\Delta_n}} - k_n\hat{p}_n}{\sqrt{k_n\hat{p}_n(1-\hat{p}_n)}}$$

$$= \lim_{n\to\infty}\frac{\log(K/S)+\sigma k_n\sqrt{\Delta_n}(1-2\hat{p}_n)}{2\sigma\sqrt{k_n\Delta_n\hat{p}_n(1-\hat{p}_n)}}$$

$$= \frac{\log(K/S)-(r+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = -d_1(S,T).$$

Furthermore we have

$$\lim_{n\to\infty}\beta_n = \lim_{n\to\infty}\sqrt{k_n\hat{p}_n^{-1}(1-\hat{p}_n)} = +\infty.$$

So $N(\beta_n)\to 1, N(\alpha_n)\to N(-d_1) = 1-N(d_1)$, completing the proof of (i).

To prove (ii) we can argue in very much the same way and arrive at parameters α_n^* and β_n^* with $\hat{p}_n$ replaced by p_n^*. Using the following limiting relations:

$$\lim_{n\to\infty}p_n^* = \frac{1}{2}, \quad \lim_{n\to\infty}k_n(1-2p_n^*)\sqrt{\Delta_n} = T\left(\frac{\sigma}{2}-\frac{r}{\sigma}\right),$$

we get

$$\lim_{n\to\infty}\alpha_n^* = \lim_{n\to\infty}\frac{\log(K/S)+\sigma n\sqrt{\Delta_n}(1-2p_n^*)}{2\sigma\sqrt{n\Delta_n p_n^*(1-p_n^*)}}$$

$$= \frac{\log(K/S)-(r-\frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = -d_2(s,T).$$

For the upper limit we get

$$\lim_{n\to\infty}\beta_n^* = \lim_{n\to\infty}\sqrt{k_n(p_n^*)^{-1}(1-p_n^*)} = +\infty,$$

whence (ii) follows similarly. □

By the above proposition we have derived the classical Black-Scholes European call option valuation formula as an asymptotic limit of option prices in a sequence of Cox-Ross-Rubinstein type models with a special choice of parameters. We will therefore call these models discrete Black-Scholes models. A straightforward analysis of the continuous-time Black-Scholes market model using stochastic calculus is contained in Chapter 6. Let us mention here that in the continuous-time Black-Scholes model the dynamics of the (stochastic) stock price process $S(t)$ are modelled by a geometric Brownian motion (or exponential Wiener process). The sample paths of this stochastic price process are almost all continuous and the probability law of $S(t)$ at any time t is lognormal. In particular the time T distribution of $\log\{S(T)/S(0)\}$ is $N(T\mu, T\sigma^2)$. Looking back at the construction of our sequence of Cox-Ross-Rubinstein models we see that

$$\log\frac{S_n(T)}{S(0)} = \sum_{i=1}^{k_n}\log(Z_{n,i}),$$

with $\log(Z_{n,i})$ Bernoulli random variables with

$$\mathbb{P}(\log(Z_{n,i}) = \sigma\sqrt{\Delta_n}) = p_n = 1-\mathbb{P}(\log(Z_{n,i}) = -\sigma\sqrt{\Delta_n}).$$

By the (triangular array version) of the central limit theorem we know that $\log\frac{S_n(T)}{S(0)}$ properly normalised converges in distribution to a random variable with standard normal distribution. Doing similar calculations as in the above proposition we can compute the normalising constants and get

$$\lim_{n\to\infty}\log\frac{S_n(T)}{S(0)} \sim N(T(r-\sigma^2/2), T\sigma^2),$$

i.e. $\frac{S_n(T)}{S(0)}$ is in the limit lognormally distributed.

Using the terminology of weak convergence we can therefore say that the probability measures $\mathbb{P}^n$ induced by the distributions of $S_n(T)/S(0)$ converge to the probability measure $\boldsymbol{Q}$ induced by $N(T(r-\sigma^2/2), T\sigma^2)$. Now the probability measures $\mathbb{P}^n$ are the risk-neutral equivalent probability measures in the nth Cox-Ross-Rubinstein model, and we will show in Chapter 6 that in continuous-time market models, similarly to discrete-time financial market modelling, contingent claims can be priced using risk-neutral equivalent martingale measures. In particular we will show that in the above Black-Scholes market the (in this case unique) equivalent martingale measure is given by $\boldsymbol{Q}$.

Therefore as a direct consequence of the definition of weak convergence (compare expression (2.2)) we have

Proposition 4.6.2. *Let X be a contingent claim of the form $X = h(S(T))$ with h a bounded, uniformly continuous real function. Denote by Π_X^n resp. Π_X the time $t=0$ price of X in the nth discrete-time resp. the continuous-time Black-Scholes market model. Then*

$$\lim_{n\to\infty}\Pi_X^n = \Pi_X.$$

Proof. Writing the pricing formula for the contingent claim using the expectation operator with respect to the risk-neutral probability measures, we have

$$\Pi_X^n = \mathbb{E}_{\mathbb{P}^n}(h(S_n(T))) = \int h\,d\mathbb{P}^n,$$

resp.

$$\Pi_X = \mathbb{E}_{\boldsymbol{Q}}(h(S(T))) = \int h\,d\boldsymbol{Q}$$

(since the σ-field at $t=0$ is assumed to be trivial, we can use expectation instead of conditional expectation), and the portmanteau theorem gives the result. □

Example. Using $h(x) = \max\{0, (K - x)\}$ we get the above convergence for the European put option, and put-call parity gives the result for the European call option (as above). Observe $g(x) = \max\{0, (x - K)\}$ is unbounded, so we cannot use Proposition 4.6.2 to give another direct proof of Proposition 4.6.1.

The above results relied heavily on the fact that contingent claims considered were of European type, i e, only dependent on the value of $S(T)$. In Chapter 6 we will reconsider the approximation procedure and show that we have in fact weak convergence in functional form, that is, for all $t \in [0, T]$ simultaneously. Thus we get approximation for path-dependent contingent claims also.

We now turn briefly to different choices of u_n and d_n and their effects.

4.6.3 Further Limiting Models

As already mentioned different choices of the sequences (u_n) and (d_n) lead to different asymptotic stock price processes. We briefly discuss two possible choices.

Jump Stock Price Movements

The key to the results in the last section was the weak convergence of the sequence of random variables $\log(\frac{S_n(T)}{S(0)})$. To show this convergence we basically used the De Moivre-Laplace theorem for binomial random variables. We now use another classical limit theorem for binomial random variables - the 'weak law of small numbers' or 'law of rare events', which states that for certain parameters the limiting distribution is a Poisson distribution (compare §2.9). Indeed, if we choose $u_n = u = e^{\zeta}$, $\zeta > 0$ (independent of n) and $d_n = e^{\xi \Delta_n}$ with some $0 < \xi < r$ we have (for large enough n) an arbitrage-free market model with unique risk-neutral probabilities p_n^* given by

$$p_n^* = \frac{\exp(r\Delta_n) - \exp(\xi\Delta_n)}{u - \exp(\xi\Delta_n)} \to 0, \ (n \to \infty).$$

For this lattice approach the step size of an upward move remains constant through all Cox-Ross-Rubinstein models, but the probability it will occur becomes very small. On the other hand, the size of a downward move becomes very small (as $\Delta_n \to 0$, we have $d_n \to 1$), but its probability becomes very close to 1.

Recall that in the sequence of Cox-Ross-Rubinstein models we modelled the stock price at time T as

$$\log \frac{S_n(T)}{S(0)} = \sum_{i=1}^{k_n} \log(Z_{n,i}),$$

with $\log(Z_{n,i})$ Bernoulli random variables. Given the size of the up and down movements and the probabilities p_n^* as above, an application of the law of rare events (see §2.9) shows that the corresponding sequence of equivalent probability measures $\mathbb{P}^n$ of the Cox-Ross-Rubinstein models converges weakly to the probability measure $\boldsymbol{Q}$ induced by a Poisson distribution with parameter $\lambda = \frac{Tu(r-\xi)}{u-1}$.

We can apply the pormanteau theorem again to find the valuation formula of a European put and use put-call parity to get the pricing formula for a European call. We use the following notation: C_n is the time $t = 0$ price of a European call in the nth Cox-Ross-Rubinstein model with parameters as above and

$$\bar{\Psi}^{\mu}(x) = 1 - \Psi^{\mu}(x - 1) = \sum_{i=x}^{\infty} \frac{e^{-\mu}\mu^i}{i!}$$

the complementary Poisson distribution function with parameter μ. With this notation we have the following limiting relation:

$$\lim_{n\to\infty} C_n = S(0)\bar{\Psi}^{\lambda}(x) - Ke^{-rT}\bar{\Psi}^{\frac{\lambda}{u}}(x).$$

The parameter λ is given as above and $x = (\log(K/S(0)) - \xi T)/\log u$.

In the limiting continuous-time model the stock price process has to be modelled in such a way that 'jumps' are possible, i.e. the paths of the stochastic stock price process must allow discontinuities. This is done by using the continuous-time Poisson process (or another point process, see Chapter §5.2). The distribution of the stock price process in the continuous-time model is then log-Poisson. This kind of binomial model was introduced by Cox and Ross in [45]; see also [46], p.365 for a somewhat different textbook treatment.

Constant Elasticity of Variance Diffusion

We now allow the up and down movements of the binomial process to differ predictably from period to period. More explicitly we write (using the notation from above)

$$u_n = u_n(S_n(j\Delta_n), \Delta_n) \quad \text{and} \quad d_n = d_n(S_n(j\Delta_n), \Delta_n).$$

To obtain an arbitrage-free market we have to choose the probabilities in the underlying single-period models according to (4.7), i.e.

$$p_{n,j}^* = p_{n,j}^*(S_n(j\Delta_n)) = \frac{\exp\{r\Delta_n\} - d_n(S_n(j\Delta_n), \Delta_n)}{u_n(S_n(j\Delta_n), \Delta_n) - d_n(S_n(j\Delta_n), \Delta_n)}.$$

This, of course, implies that the equivalent martingale measure for the nth Cox-Ross-Rubinstein model is dependent on the whole family of probabilities $p_{n,0}^*, \ldots, p_{n,k_n-1}^*$.

For instance, if we use the functions

$$u(y,t) = \mu yt + \sigma y^p \sqrt{t} \quad \text{and} \quad d(y,t) = \mu yt - \sigma y^p \sqrt{t}, \quad 0 < p \le 1,$$

and set

$$u_n(S(t),t) = \exp\{u(S(t),t)\} \quad \text{and} \quad d_n(S(t),t) = \exp\{d(S(t),t)\},$$

we have

$$p_{n,j} = \frac{e^{r\Delta_n} - e^{\mu S_n(j\Delta_n)\Delta_n - \sigma S_n^p(j\Delta_n)\sqrt{\Delta_n}}}{e^{\mu S_n(j\Delta_n)\Delta_n + \sigma S_n^p(j\Delta_n)\sqrt{\Delta_n}} - e^{\mu S_n(j\Delta_n)\Delta_n - \sigma S_n^p(j\Delta_n)\sqrt{\Delta_n}}}.$$

With these parameters one can show that the probability measures $\mathbb{P}^n$ converge weakly to a probability measure $\boldsymbol{Q}$ induced by a certain gamma-type distribution. This leads to the constant elasticity of variance option pricing formula for the limit of European call option prices at time 0 in the above sequence of Cox-Ross-Rubinstein models:

$$\lim_{n\to\infty} C_{n,0} = S(0)\sum_{i=1}^{\infty} g(i,x)\bar{G}(i+\lambda,y) - Ke^{-rT}\sum_{i=1}^{\infty} g(i+\lambda,x)\bar{G}(i,y).$$

The function $g(i,u)$ is the gamma density function

$$g(i,u) = \frac{e^{-u}u^{i-1}}{(i-1)!}$$

and the function $\bar{G}(i,z)$ the complementary gamma distribution function

$$\bar{G}(i,z) = \int_z^{\infty} g(i,u)du.$$

The parameters are given as $\lambda = 1/(2(1-p))$, $x = 2\lambda r S(0)^{\frac{1}{\lambda}} e^{rT/\lambda}/(\sigma^2(e^{rT/\lambda} - 1))$ and $y = 2\lambda r K^{\frac{1}{\lambda}}/(\sigma^2(e^{rT/\lambda} - 1))$.

The corresponding continuous-time stock price dynamics are given by

$$dS(t) = \mu S(t)dt + \sigma S(t)^p dW(t)$$

(where $dW(t)$ denotes the stochastic differential with respect to the Wiener process - we treat this in Chapter 5) and the constant elasticity in the (conditional) variance term (in front of $dW(t)$) gives the name to this model.

Remark 4.6.1. The numerics of the above approximations have been subject to investigation for quite some time (see [31, 152] for discussion and references). Such numerical schemes are easy to implement, for instance using Mathematica, and the reader is invited to do so.

4.7 Multifactor Models

We now discuss examples of discrete-time financial market models with more than two underlying assets. Such models are useful for the evaluation of multivariate contingent claims, such as options on multiple assets (options on the maximum of two or more asset prices, dual-strike options, and portfolio or basket options). For the exposition we assume $d+1$ financial assets $S_0, S_1, \dots, S_d$. We assume $S_0 = B$, a risk-free bank account or bond, and use B as numéraire.

4.7.1 Extended Binomial Model

This model, proposed by Boyle, Evnine and Gibbs [28], uses a single binomial tree for each of the underlying d risky assets. So we have 2^d branches per node. We discuss the case $d = 2$ (i.e. the model consists of two risky assets and the bank account) in detail; the generalisation to $d > 2$ is straightforward. To show that this model is arbitrage-free we have to find an equivalent martingale measure and to show that it is complete we have to prove uniqueness of the equivalent martingale measure. A similar argument to that for the Cox-Ross-Rubinstein model shows that the multi-period extended binomial model is arbitrage-free (complete) if and only if the single-period model is (compare §4.5.2). So it is enough to discuss the single-period model with trading dates $t = 0$ and $t = 1 (= T)$. We assume a risk-free rate of return of $r \ge 0$, so $B(0) = 1$ and $B(1) = 1 + r$. Furthermore we have two risky assets, S_1 and S_2. Since both risky assets are modelled by single binomial trees, we have four possible states of the world at time $t = 1$ with values of $(S_1(1), S_2(1))$ given by $(u_1 S_1(0), u_2 S_2(0))$ with probability p_{uu}, $(u_1 S_1(0), d_2 S_2(0))$ with probability p_{ud}, $(d_1 S_1(0), u_2 S_2(0))$ with probability p_{du} and $(d_1 S_1(0), d_2 S_2(0))$ with probability p_{dd}, where we assume $u_i > d_i$, $i = 1, 2$ and positive probabilities. Under the risk-neutral probabilities $p^*_{uu}, p^*_{ud}, p^*_{du}, p^*_{dd}$ the discounted stock price processes $\tilde{S}_i(t) = S_i(t)/B(t)$ have to be martingales. These martingale conditions imply the following two equations:

$$\begin{aligned} \mathbb{E}[\tilde{S}_1(1)] = \tilde{S}_1(0) &\Leftrightarrow (p^*_{uu} + p^*_{ud})u_1 + (p^*_{du} + p^*_{dd})d_1 = (1+r), \\ \mathbb{E}[\tilde{S}_2(1)] = \tilde{S}_2(0) &\Leftrightarrow (p^*_{uu} + p^*_{du})u_2 + (p^*_{ud} + p^*_{dd})d_2 = (1+r). \end{aligned}$$

Furthermore, besides the fact that the p^* have to be positive to generate an equivalent measure, we must have

$$p^*_{uu} + p^*_{ud} + p^*_{du} + p^*_{dd} = 1.$$

So we have three equations for the unknown probabilities $p^*_{uu}, p^*_{ud}, p^*_{du}, p^*_{dd}$ and in general (depending on the parameters u_1, d_1, u_2, d_2, r) we will have several (even infinitely many) solutions of the system of the equations above. This means that the extended binomial model is arbitrage-free, but not complete (in accordance to our rule of thumb (§1.4) that we should have as many financial assets to trade in as states of the world).

4.7.2 Multinomial Models

The extended binomial model shows that while it is tempting to model each asset by a single binomial tree we lose the desirable property of market completeness in doing so. We will therefore now construct a arbitrage-free, complete market model (with $d > 2$ financial assets) following the informal rule of allowing as many different states of the world as we have assets to trade in. Furthermore the stochastic stock price processes in this model can be constructed to be of Markovian nature, that is, rather than the single-period returns being independent unconditionally, they are independent given the present value of the process. This also allows for a more realistic representation of the true prices and is more in line with the most prominent continuous-time model, the Black-Scholes market model, in which the stock price processes are Markovian. We follow an approach which is basically due to He [116]. Again we only discuss the $d = 2$ case (with the risk-free bank account B, with rate of return $r \geq 0$, as numéraire asset and two risky assets S_1, S_2); the case $d > 2$ follows by the same prescription. Let us start with the single-period model. As in the extended binomial case above we assume trading dates $t = 0$ and $t = 1 (= T)$, but now we have only three possible states of the world at time $t = 1$. Indeed we set

$$S_1(1) = S_1(0)Z_1 \quad \text{and} \quad S_2(1) = S_2(0)Z_2,$$

with

$$\mathbb{P}(Z_1 = u_{11}, Z_2 = u_{21}) = p_1; \quad \mathbb{P}(Z_1 = u_{12}, Z_2 = u_{22}) = p_2;$$
$$\mathbb{P}(Z_1 = u_{13}, Z_2 = u_{23}) = p_3.$$

In general Z_1 and Z_2 are not independent, but we still can choose u_{ij} in such a way that they are uncorrelated. Under the risk-neutral probabilities p_1^*, p_2^*, p_3^* the discounted stock price processes $\tilde{S}_i(t) = S_i(t)/B(t)$ have to be martingales. These martingale conditions imply the following two equations:

$$\mathbb{E}[\tilde{S}_1(1)] = \tilde{S}_1(0) \Leftrightarrow u_{11}p_1^* + u_{12}p_2^* + u_{13}p_3^* = (1+r),$$
$$\mathbb{E}[\tilde{S}_2(1)] = \tilde{S}_2(0) \Leftrightarrow u_{21}p_1^* + u_{22}p_2^* + u_{23}p_3^* = (1+r).$$

Furthermore, besides the fact that the p^* have to be positive to generate an equivalent measure, we must have

$$p_1^* + p_2^* + p_3^* = 1.$$

Therefore we have three equations for the three unknown probabilities and in general (given reasonable parameters u_{ij}) we will have a unique solution of the system of the equations above, and hence an arbitrage-free, complete financial market model.

In the multi-period setting with time horizon T and the set of trading dates given by $\{0 = t_0 < t_1 < \ldots < t_n = T\}$ of equidistant time points with

distance Δ_n (observe that we have n time steps), we model the stock price processes by

$$S_i(t_k) = S_i(0) \prod_{j=1}^{k} Z_{ij}, \; k = 0, 1, \ldots, n, \; i = 1, 2,$$

with a sequence of independent random vectors $(Z^{(j)})_{1 \leq j \leq n}$ such that $Z_1^{(j)}$, $Z_2^{(j)}$ are uncorrelated (but possibly dependent) and

$$\mathbb{P}(Z_1^{(j)} = u_{11}^{(j)}, Z_2^{(j)} = u_{21}^{(j)}) = p_1^{(j)};$$
$$\mathbb{P}(Z_1^{(j)} = u_{12}^{(j)}, Z_2^{(j)} = u_{22}^{(j)}) = p_2^{(j)};$$
$$\mathbb{P}(Z_1^{(j)} = u_{13}^{(j)}, Z_2^{(j)} = u_{23}^{(j)}) = p_3^{(j)}.$$

Since for each j the random vector $Z^{(j)}$ can be in one of three possible states, the above argument applies for each 'underlying' single-period market and the multi-period market is arbitrage-free and complete.

The most important case here is $Z_i^{(j+1)} = u_i(S(t_j), t_j, \epsilon^{(j)})$, $i = 1, 2$, $j = 0, \ldots n-1$, with a sequence of independent random vectors $(\epsilon^{(j)})_{j \leq n-1}$ such that $\epsilon_1^{(j)}, \epsilon_2^{(j)}$ are uncorrelated (but possibly dependent) and sufficiently smooth functions u_i. Then $u_i^{(j+1)}$ are a predictable functions of $S(t_j)$ making the discrete stochastic process $S_i(t)$ Markovian. We will construct a financial market model of this type in §6.4.

4.8 Further Contingent Claim Valuation in Discrete Time

4.8.1 American Options

Recall that is never optimal to exercise an American call early.

We now consider how to evaluate an American put option, European and American call options having been treated already. The new feature is the need to distinguish between the *stopping region*, where it is optimal to stop and exercise our American option *early*, and the *continuation region*, where it is optimal to proceed as with the European counterpart and continue. The distinctive feature of American options, and the main difficulty in dealing with them, is that it is not in general possible to *price* the option without at the same time *finding the stopping and continuation regions*. Furthermore, these regions cannot be found 'once and for all' by explicit formulas, but must instead be found by a recursive procedure, case by case. We use the discrete-time Black-Scholes model of §4.6, i.e. two financial assets, a risk-free bank account (used as numéraire) and a risky stock, modelled in a binomial set-up. We now divide the time interval $[0, T]$ into N equal subintervals of length Δ say. Assuming the risk-free rate of interest r (over [0,T]) as given, we have $1+\rho = e^{r\Delta}$ (where we denote the risk-free rate of interest in each subinterval by ρ). The remaining degrees of freedom are resolved by choosing u and d as follows:

$$u = e^{\sigma\sqrt{\Delta}}, \quad \text{and} \quad d = u^{-1} = e^{-\sigma\sqrt{\Delta}}.$$

By condition (4.7) the risk-neutral probabilities for the corresponding single period models are given by

$$p^* = \frac{1+\rho-d}{u-d} = \frac{e^{r\Delta} - e^{-\sigma\sqrt{\Delta}}}{e^{\sigma\sqrt{\Delta}} - e^{-\sigma\sqrt{\Delta}}}.$$

Thus the stock with initial value $S = S(0)$ is worth $Su^i d^j$ after i steps up and j steps down. Consequently, after N steps, there are $N+1$ possible prices, $Su^i d^{N-i}$ ($i = 0, \ldots, N$). There are 2^N possible paths through the tree. It is common to take N of the order of 30, for two reasons:

(i) typical lengths of time to expiry of options are measured in months (9 months, say); this gives a time step around the corresponding number of days,
(ii) 2^{30} paths is about the order of magnitude that can be comfortably handled by computers (recall that $2^{10} = 1,024$, so 2^{30} is somewhat over a billion).

We can now calculate both the value of an American put option and the optimal exercise strategy by working backwards through the tree (this method of backward recursion in time is a form of the dynamic programming (DP) technique, due to Richard Bellman, which is important in many areas of optimisation and Operational Research).

1. Draw a binary tree showing the initial stock value and having the right number, N, of time intervals.
2. Fill in the stock prices: after one time interval, these are Su (upper) and Sd (lower); after two time intervals, Su^2, S and $Sd^2 = S/u^2$; after i time intervals, these are $Su^j d^{i-j} = Su^{2j-i}$ at the node with j 'up' steps and $i-j$ 'down' steps (the '(i,j)' node).
3. Using the strike price K and the prices at the terminal nodes, fill in the payoffs $f_{N,j} = \max\{K - Su^j d^{N-j}, 0\}$ from the option at the terminal nodes underneath the terminal prices.
4. Work back down the tree, from right to left. The values f_{ij}^E of the corresponding European option at the (i,j) node are given in terms of those of its upper and lower right neighbours in the usual way, as discounted expected values under the risk-neutral measure:

$$f_{ij}^E = e^{-r\Delta}[p^* f_{i+1,j+1} + (1-p^*) f_{i+1,j}].$$

The intrinsic (or early-exercise) value of the American put at the (i,j) node – the value there if it is exercised early – is

$$K - Su^j d^{i-j}$$

(when this is non-negative, and so has any value). The value of the American put is the higher of these:

$$\begin{aligned} f_{ij} &= \max\{f_{ij}^E, K - Su^j d^{i-j}\} \\ &= \max\left\{e^{-r\Delta}(p^* f_{i+1,j+1} + (1-p^*) f_{i+1,j}), K - Su^j d^{i-j}\right\}. \end{aligned}$$

5. The initial value of the option is the value f_0 filled in at the root of the tree.
6. At each node, it is optimal to exercise early if the early-exercise value there exceeds the value f_{ij}^E there of the corresponding European option.

Note. The above procedure is simple to describe and understand, and simple to program. It is laborious to implement numerically by hand, on examples big enough to be non-trivial. Numerical examples are worked through in detail in [122], 359-360 and [46], 241-242.

Mathematically, the task remains of describing the *continuation region* – the part of the tree where early exercise is not optimal. This is a classical *optimal stopping problem*, and as we mentioned above, a solution by explicit formulas is not known – indeed, is probably not feasible. It would take us too far afield to pursue such questions here; for a fairly thorough (but quite difficult) treatment, see [207]. We will, however, connect the work above with that of Chapter 3 on the Snell envelope. Consider the pricing of an American put, strike price K, expiry $T = N$, in discrete time, with discount factor $1+\rho$ per unit time as earlier. Let $Z = (Z_n))_{n=0}^N$ be the payoff on exercising at time n. We want to price Z_n, by U_n say (to conform to our earlier notation), so as to avoid arbitrage; again, we work backwards in time. The recursive step is

$$U_{n-1} = \max\left\{Z_{n-1}, \frac{1}{1+r} \mathbb{E}^*[U_n|\mathcal{F}_{n-1}]\right\},$$

the first alternative on the right corresponding to early exercise, the second to the discounted expectation under P^*, as usual. Let $\tilde{U}_n = U_n/(1+r)^n$ be the discounted price of the American option: then

$$\tilde{U}_{n-1} = \max\left\{\tilde{Z}_{n-1}, \mathbb{E}^*[\tilde{U}_n|\mathcal{F}_{n-1}]\right\}.$$

Thus $(\tilde{U}_n)$ is the *Snell envelope* (§3.7) of the discounted payoff process $(\tilde{Z}_n)$. It is thus:

(i) a P^*-supermartingale,
(ii) the smallest supermartingale dominating $(\tilde{Z}_n)$,
(iii) the solution of the optimal stopping problem for Z.

We conclude by showing the equivalence of American and European calls without using arbitrage arguments.

Theorem 4.8.1. *Let $(Z_n)_0^N$ be an adapted sequence, $h := Z_N$; write $C_A(n)$, $C_E(n)$ for the values at time n of the American and European options generated by the payoff function h. Then*

(i) $C_A(n) \geq C_E(n)$,
(ii) If $C_E(n) \geq Z_n$, then $C_A(n) = C_E(n)$.

Proof. (i) We use the supermartingale resp. martingale property of the price processes of the discounted American resp. European call to get

$$\tilde{C}_A(n) \geq \mathbb{E}^* \left(\tilde{C}_A(N)|\mathcal{F}_n\right) = \mathbb{E}^* \left(\tilde{C}_E(N)|\mathcal{F}_n\right)] = \tilde{C}_E(n).$$

(ii) $(\tilde{C}_E(n))$ is a P^*-martingale, so in particular a P^*-supermartingale. Being the Snell envelope of (Z_n), $(\tilde{C}_A(n))$ is the least P^*-supermartingale dominating (Z_n). So if $\tilde{C}_E(n) \geq Z_n$ as in the condition of the theorem, $\tilde{C}_E(n) \geq \tilde{C}_A(n)$, so $\tilde{C}_E(n) = \tilde{C}_A(n)$. □

Corollary 4.8.1. *In the Black-Scholes model with one risky asset, the American call option is equivalent to its European counterpart.*

Proof. Here $Z_n = (S_n - K)_+$. Discounting,

$$\tilde{C}_E(n) = (1+\rho)^{-N} \mathbb{E}^* \left((S_N - K)_+|\mathcal{F}_n\right)$$
$$\geq \mathbb{E}^* \left(\tilde{S}_N - K(1+\rho)^{-N}|\mathcal{F}_n\right) = \tilde{S}_n - K(1+\rho)^{-N},$$

since $\tilde{S}_n$ is a $\mathbb{P}^*$-martingale. Without the discounting, this says

$$C_E(n) \geq S_n - K(1+\rho)^{-(N-n)}.$$

This gives $C_E(n) \geq S_n - K$; also $C_E(n) \geq 0$; so $C_E(n) \geq (S_n - K)_+ = Z_n$, and the result follows from the theorem. □

4.8.2 Barrier Options

Barrier options are options whose payoff depends on whether or not the stock price attains some specified level before expiry. We will be brief here, referring to §6.3.3 for a more extensive discussion of barrier options in continuous time. The simplest case is that of a single, constant barrier at level H. The option may pay ('knock in') or not ('knock out') according as to whether or not level H is attained, from below ('up') or above ('down'). There are thus four possibilities – 'up and in', 'up and out', 'down and in', 'down and out' – for the basic – single, constant barrier – case. In addition, one may have two barriers, with the option knocking in (or out) if the price reaches *either* a lower barrier H_1 *or* an upper barrier H_2. More generally, one may have non-constant – 'moving' – barriers, with the level a function of time.

As always, it pays to be flexible, and to be able to work in discrete or continuous time, as seems more appropriate for the problem in hand. For a full treatment in continuous time, see [224], Chapters 10, 11, or §6.3.3. Now a continuous-time price process model, such as the Black-Scholes model based on geometric Brownian motion (§6.2) may be approximated in various ways by discrete-time models (such as the discrete Black-Scholes model, the Cox-Ross-Rubinstein binomial tree model of §4.5); for the passage from discrete to continuous time, see §4.6 (and more generally, §5.9 below).

When we have a barrier option in discrete time, we price it as with the American options of §6.3.1 by backward induction. Some sample paths hit the barriers, and for these we can fill in the payoff from the boundary conditions that define the barriers; as before, we fill in the payoff at the terminal nodes at expiry. We then proceed backwards in time recursively, at each stage using all current information to fill in, as before, the payoffs at new nodes one time step earlier. When we reach the root, the payoff is the value of the option initially.

Problems may easily be encountered when dealing with barrier options in discrete time if the discretisation process is not chosen and handled with care. A new discretisation process, due to Rogers and Stapleton [187], proceeds by first discretising *space*, by steps $\delta x > 0$, and then discretising *time*, into $\tau_0, \tau_1, \cdots$, where

$$\tau_0 := 0, \quad \tau_{n+1} := \inf\{t > \tau_n : |X(t) - X(\tau_n)| > \delta x\}, \quad n \geq 0,$$

and deal with the resulting random walk (ξ_n), where

$$\xi_n := X(\tau_n).$$

This approximation scheme is accurate, reasonably fast, and very flexible: it is capable of handling a wide variety of problems, with moving as well as fixed barriers. For the theory, and detailed comparison with other available methods, see [187]; another approach is due to AitSahlia and Lai [3]. Techniques useful here include continuity corrections for approximations to normality, Edgeworth expansions, and Richardson extrapolation.

4.8.3 Lookback Options

Lookback – or hindsight – options, which we discuss in more detail in §6.3.4 in continuous time, are options which convey the right to 'buy at the low, sell at the high' – in other words, to eliminate the regret that an investor operating in real time on current, partial knowledge would feel looking back in time with complete knowledge. Again, most of the theory is for continuous time (see e.g. [224], Chapter 12), but a discrete-time framework may be preferred – or needed, if the only prices available are those sampled at certain discrete time-points. Care is obviously needed here, as discretisation of time will miss the extremes of the peaks and troughs giving the highs and lows in continuous time.

Discrete lookback options have been studied from several viewpoints; see e.g. [120, 138, 153]. An interesting approach using duality theory for random walks has recently been given by AitSahlia and Lai [2].

4.8.4 A Three-Period Example

Assume we have two basic securities, a risk-free bond and a risky stock. The one-year risk-free interest rate (continuously compounded) is $r = 0.06$ and the volatility of the stock is 20 %. We price calls and puts in three-period Cox-Ross-Rubinstein model. The up and down movements of the stock price are given by

$$u = e^{\sigma\sqrt{\Delta}} = 1.1224 \quad \text{and} \quad d = u^{-1} = e^{-\sigma\sqrt{\Delta}} = 0.8910$$

with $\sigma = 0.2$ and $\Delta = 1/3$. We obtain risk-neutral probabilities by (4.7)

$$p^* = \frac{e^{r\Delta} - d}{u - d} = 0.5584.$$

We assume that the price of the stock at time $t = 0$ is $S(0) = 100$. To price a European call option with maturity one year ($T = 3$) and strike $K = 10$) we can either use the valuation formula (4.8) or work our way backwards through the tree. Prices of the stock and the call are given in Fig. 4.2 below. It is interesting to compare the approximative Cox-Ross-Rubinstein prices c_n (discrete model with n time steps, see (4.8), (4.12) to the Black-Scholes price $C_{BS} = 10.9895$ (according to (4.13)). We have

n	c(n)
5	11.33
10	10.79
50	10.95
100	10.97
200	10.98
500	10.99

To price a European put, with price process denoted by $p(t)$, and an American put, $P(t)$, (maturity $T = 3$, strike 100), we can for the European put either use the put-call parity (1.1), the risk-neutral pricing formula, or work backwards through the tree. For the prices of the American put we use the technique outlined in §4.8.1. Prices of the two puts are given in Fig. 4.3. We indicate the early exercise times of the American put in bold type. Recall that the discrete-time rule is to exercise if the intrinsic value $K - S(t)$ is larger than the value of the corresponding European put.

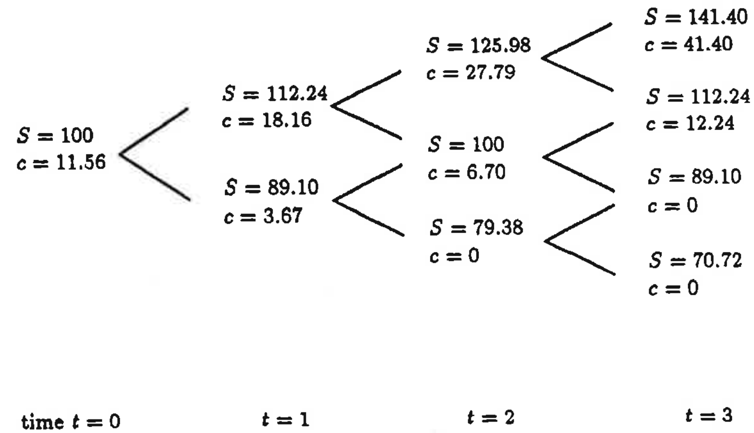

Fig. 4.2. Stock and European call prices

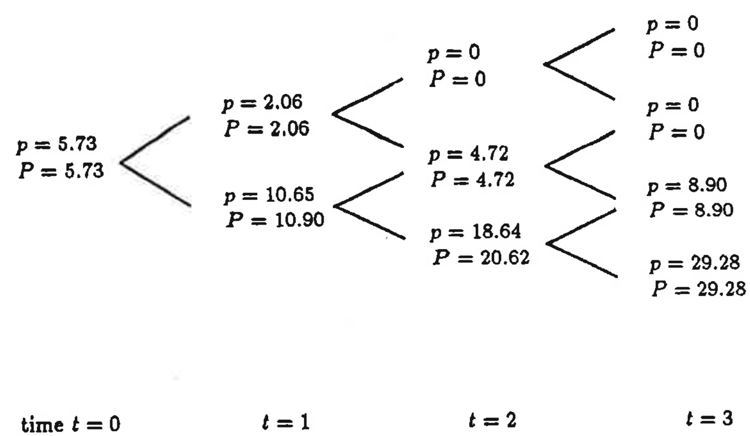

Fig. 4.3. European $p(.)$ and American $P(.)$ put prices

Exercises

4.1 Construct hedging strategies for the European call and put in the setting of the example in §4.8.4.

4.2 Compare the Black-Scholes price with Cox-Ross-Rubinstein price approximations. Is the convergence of Cox-Ross-Rubinstein prices to the Black-Scholes price 'smooth' or 'oscillating'? (See [152] for details.)

4.3 Consider a European call option, written on a stock S, with strike price 100 which matures in one year. Assume the continuously compounded riskfree interest rate is 5%, the current price of the stock is 90 and its volatility is $\sigma = 0.2$.

1. Set up a three-period binomial (Cox-Ross-Rubinstein) model for the stock price movements.
2. Compute the risk-neutral probabilities and find the value of the call at each node.
3. Construct a hedging portfolio for the call.

4.4 Consider put options, written on a stock S, with strike price 100 which mature in one year. Assume the continuously compounded riskfree interest rate is 6%, the current price of the stock is 100 and its volatility is $\sigma = 0.25$.

1. Set up a three-period binomial (Cox-Ross-Rubinstein) model for the stock price movements.
2. Compute the risk-neutral probabilities and find the value of a European put at each node.
3. Construct a hedging portfolio for the European put.
4. Now compute the values of a corresponding American put at each node and set up a hedging portfolio. Compare with the hedging portfolio in 3.

4.5 Consider a European powered call option, written on a stock S, with expiry T and strike K. The payoff is ($p > 1$):

$$C_p(T) = \begin{cases} (S(T) - K)^p, & S(T) \geq K; \\ 0 & S(T) < K. \end{cases}$$

Assume that $T = 1$ year, $S(0) = 90, \sigma = 0.3, K = 100$. Consider a two-period binomial model.

1. Price C_p using the risk-neutral valuation formula.
2. Construct a hedge portfolio and compute arbitrage prices (which of course will agree with the risk-neutral prices) using the hedging portfolio.
3. Compare the hedge portfolio with a hedge portfolio for a usual European call. What are the implications for the risk-management of powered call options?

4.6 In *static hedging* of exotic options one tries to construct a portfolio of standard options – with varying strikes and maturities but fixed weights that will not require any further adjustment – that will exactly replicate the value of the given target option for a chosen range of future times and market levels.

We will construct a static hedge for a barrier option in a binomial five-period model. Consider a zero interest-rate world with a stock worth 100 today. The stock price can move up and down 10 with probability 0.5 at the end of a fixed period.

Our target for replication is a five-period up-and-out European-style call with a strike of 70 and a barrier of 120. This option has natural boundaries both at expiration in five periods and on the knockout barrier at 120.

Create a portfolio of ordinary options that collectively have the same pay-off as the up-and-out call on the boundaries. To create such a portfolio following the steps:

1. Start with an ordinary call struck at 70. It has the same payoff if the barrier is never reached.
2. Add a short position in 10 five-period calls with strike 120 to the portfolio to make the portfolio value 0 at the time 4 boundary point.
3. Add a long position in 5 three-period calls struck at 120 to complete the portfolio.

For each portfolio compute the value-process at every node and compare it with the value of the barrier option.