Chapter 4

No-Arbitrage Pricing

In these notes we try to explain what arbitrage is and how to price financial derivatives in order to avoid arbitrage. The argument runs from a two-state model to the fundamental theorem in general form, and then asks what the theorem needs: completeness, which is what makes a price a number rather than a range. Jumps break it, so we work out what a jump model hands over instead — a split of the risk into the part that can be hedged continuously and the part that cannot — and close by deriving, rather than assuming, why every model in these notes is a diffusion.

4.1 Binomial Model

To get an intuition for the concept of arbitrage and risk neutral pricing, we digress a bit from the continuous time stochastic process model and work for a while on the simpler discrete time binomial model, where in each time step, given the value at initial time, the process can take one of two possible values with certain probabilities at the final time.

Definition 4.1 (Forward Contract)**.** A forward contract is an agreement between two parties to buy or sell a financial tradable at a decided date in the future at a price determined at the inception of the contract. The set price is known as the forward price, and the decided date in future is known as the expiration date. Usually there is no initial exchange of cashflows before the expiration.

Let's start with a simple example. Suppose we have a stock trading in the market for \$100 and from the historical probability we know that in a year the stock may go to \$120 with probability 0.5 and may go to \$90 with probability 0.5. Assume that the bank borrows and lends money at an interest of 2%. Suppose that there is a forward contract trading in the market for the stock which lets you buy or sell the stock a year later. With the given historic probabilities, one may be inclined to price the forward contract at

$$\mathbb{E}^{\mathbb{P}^{\text{historic}}}[S_{1\text{year}}] = 0.5 \cdot \$120 + 0.5 \cdot \$90 = \$105$$

Suppose that a bank is actually trading the forward for \$105. Can you figure out a way to make profit with probability 1?

We can sell the forward derivative for \$105, take a loan from the bank for \$100, use this money to buy the stock right now, give this stock to the buyer of the forward for \$105 a year from now, and repay the \$102 to the bank for the loan we took. We start with no money and at the end of the year we have a profit of \$3 with probability 1. This is arbitrage.

Definition 4.2 (Arbitrage)**.** A portfolio is an arbitrage if we have

$$\pi_0 = 0 \qquad \mathbb{P}^{\text{historic}}(\pi_T \geq 0) = 1 \qquad \mathbb{P}^{\text{historic}}(\pi_T > 0) > 0 \tag{4.1}$$

where π_t is the value of the portfolio at time t. The portfolio must be self-financing meaning that we do not put more cash into the portfolio during the execution of the strategy.

The two conditions on probability mean that almost surely we incur no loss from executing our strategy, and that there is a non-zero chance of gaining profit.

Definition 4.3 (Option contract)**.** An option contract is an agreement between two parties where the buyer of the contract has the option, but not the obligation, to buy or sell a financial tradable at a decided date in the future at a price determined at the inception of the contract. An option contract with the option to buy (sell) is called a call (put) option, the set price is known as the strike, and the decided date in future is known as the option expiration date.

Definition 4.4 (Bond)**.** A bond is a contract between two parties that pays \$1 to the buyer of the bond at the expiration date.

The riskless bond price is governed primarily by the interest rates.

Example 4.1 (Binomial Option Pricing)**.** Let's see how we can price an option in a binomial model so as to avoid arbitrage. Take the same stock in the previous example. Suppose a client wants to buy a call option on the stock with strike as \$105 and expiration in a year. In a year if the stock price goes up to \$120, the client will want to exercise the option and get the stock for \$105, and if the stock price goes down to \$90 the client is better off not exercising the option and simply buying the stock from the market. Note that if we simply buy the stock now to give it to the client in future we are protected against the risk of stock prices rising if the client chooses to exercise, however if the stock prices go down and the client does not exercise we have the risk of losing money on the stock. So we try to think of a replication portfolio made up of stocks and bonds such that the final payoff of the portfolio is always equal to the payoff of the option at expiration. And if we can buy this replication portfolio now then no matter what turn the stock price takes we have zero risk. The final payoff

$$\begin{aligned}
&= \begin{cases} S_{1\text{year}} - \$105 & \text{if} \quad S_{1\text{year}} = \$120 \\ 0 & \text{if} \quad S_{1\text{year}} = \$90 \end{cases} \\
&= \begin{cases} \frac{1}{2}S_{1\text{year}} - \$45 & \text{if} \quad S_{1\text{year}} = \$120 \\ \frac{1}{2}S_{1\text{year}} - \$45 & \text{if} \quad S_{1\text{year}} = \$90 \end{cases} \\
&= \frac{1}{2}S_{1\text{year}} - \$45 \\
&= \frac{1}{2}S_{1\text{year}} - 45B_{1\text{year}}
\end{aligned}$$

So the option payoff can indeed be replicated by a portfolio of stocks and bonds. If we buy 1/2 stock and sell 45 bonds then we have completely hedged our risk. The price of this replication portfolio, and hence the call option, is

$$\frac{1}{2} \cdot S_{\text{now}} - 45 \cdot B_{\text{now}} = \frac{1}{2} \cdot \$100 - 45 \cdot \$\frac{1}{1.02} \sim \$6$$

If the price of the call option were any different from the price of the replication portfolio, one could trade in option and replication portfolio and make unlimited profits.

Exercise (Binomial Forward Price)**.** Show that the forward price of the forward contract in the first example should be equal to $S_{\text{now}}/B_{\text{now}}$ to not have arbitrage.

Because of the discrete binary branching in the binomial model it is always possible to form a replicating portfolio with just stocks and bond. We simply have to solve a system of two linear equations

$$\begin{aligned} a \cdot S_u + b \cdot B &= D_u \\ a \cdot S_d + b \cdot B &= D_d \end{aligned}$$

where S_u and S_d are the two possible stock price a year from now, and D_u and D_d are the corresponding payoff of the stock derivative. B is a bond with 1 payoff a year from now. The current price of the derivative is then

$$D_{\text{now}} = a \cdot S_{\text{now}} + b \cdot B_{\text{now}} \tag{4.2}$$

Solving for a and b we get

$$\begin{aligned} a &= \frac{D_u - D_d}{S_u - S_d} \\ b &= \frac{D_d \cdot S_u - D_u \cdot S_d}{S_u - S_d} \\ D_{\text{now}} &= \frac{D_u - D_d}{S_u - S_d} \cdot S_{\text{now}} + \frac{D_d \cdot S_u - D_u \cdot S_d}{S_u - S_d} \cdot B_{\text{now}} \\ &= \frac{S_{\text{now}} - S_d B_{\text{now}}}{S_u - S_d} \cdot D_u + \frac{-S_{\text{now}} + S_u B_{\text{now}}}{S_u - S_d} \cdot D_d \\ &\equiv q \cdot B_{\text{now}} D_u + (1-q) \cdot B_{\text{now}} D_d \end{aligned}$$

Definition 4.5 (Risk Neutral Probability)**.** In the final representation of the no-arbitrage derivative pricing formula above, q is sometimes interpreted as a probability measure and is known as the risk neutral probability. The formula can then be interpreted as an expectation under this probability.

Note that the real world historic probabilities don't show up in the no-arbitrage pricing formula at all. Before getting back into the continuous time stochastic model, note an identity that translates meaningfully into the continuous time version as well.

Remark (Return on Stock in Risk Neutral Measure)**.**

$$\mathbb{E}^{\mathbb{P}-\text{risk-neutral}}[S_{1\text{year}}] = q \cdot S_u + (1-q) \cdot S_d = S_{\text{now}}/B_{\text{now}} \tag{4.3}$$

i.e. expected return on a stock in the risk neutral probability measure is the same as the return on a riskless bond.

4.2 Continuous Time Itô Diffusion Model

Let's get back to continuous time processes now.

Consider a portfolio of stocks and money market. The value of portfolio at any time is

$$\begin{aligned}\pi_t &= \Delta_t S_t + M_t \\ \pi_0 &= 0\end{aligned}$$

where Δ_t is the amount of stocks of price S_t held at time t and M_t is the amount invested in money market which gives an instantaneous return of rdt where r is the riskfree interest rate.

$$d_{\text{market}}\pi_t = \Delta_t dS_t + rM_t dt$$

i.e. the change in portfolio value due to market movement. After any time step we neither inject nor withdraw cash from the portfolio, we just rebalance the positions giving us

$$\begin{aligned}d\pi_t &= d_{\text{market}}\pi_t \\ &= \Delta_t dS_t + rM_t dt \\ (d\Delta_t)S_t + \Delta_t dS_t + (d\Delta_t)(dS_t) + dM_t &= \Delta_t dS_t + rM_t dt && \text{(using Itô's Lemma)} \\ dM_t &= rM_t dt - (d\Delta_t)(S_t + dS_t)\end{aligned}$$

Take the ansatz $M_t = e^{\int_0^t rds}\tilde{M}_t$ and $S_t = e^{\int_0^t rds}\tilde{S}_t$. We then have

$$\begin{aligned}e^{\int_0^t rds} d\tilde{M}_t + re^{\int_0^t rds}\tilde{M}_t dt &= re^{\int_0^t rds}\tilde{M}_t dt \\ &\quad - (d\Delta_t)\, e^{\int_0^t rds}\big(\tilde{S}_t + d\tilde{S}_t + r\tilde{S}_t dt\big) \\ d\tilde{M}_t &= -(d\Delta_t)((1 + rdt)\tilde{S}_t + d\tilde{S}_t) \\ &= -(d\Delta_t)(\tilde{S}_t + d\tilde{S}_t) && (\because (d\Delta_t)(dt) = 0)\end{aligned}$$

Therefore, with the ansatz $\pi_t = e^{\int_0^t rds}\tilde{\pi}_t$, we have

$$\begin{aligned}\tilde{\pi}_t &= \Delta_t \tilde{S}_t + \tilde{M}_t \\ d\tilde{\pi}_t &= d(\Delta_t \tilde{S}_t) + d\tilde{M}_t \\ &= d\Delta_t d\tilde{S}_t + (d\Delta_t)\tilde{S}_t + \Delta_t d\tilde{S}_t + d\tilde{M}_t \\ &= \Delta_t d\tilde{S}_t\end{aligned}$$

Denoting $e^{\int_0^t rds}$ by A_t, also known as the accumulation factor, we have

$$\frac{\pi_T}{A_T} = \int_0^T \Delta_t d\frac{S_t}{A_t}$$

We can take any price of any tradable asset, including derivatives of stocks, as S_t and the analysis still holds. $D_t = 1/A_t$ is also known as discount factor and $D_t S_t = S_t/A_t$ is known as the discounted asset price.

Definition 4.6 (Equivalent Measures)**.** Two probability measures $\mathbb{P}^a$ and $\mathbb{P}^b$ are equivalent if for any event A of the σ-algebra $\mathbb{P}^a(A) = 0$ if and only if $\mathbb{P}^b(A) = 0$.

Definition 4.7 (Admissible strategy)**.** A self-financing strategy is admissible if there is a constant $a \geq 0$ such that its discounted wealth satisfies $\pi_t/A_t \geq -a$ for all t, almost surely. The restriction is not a technicality: §4.5 exhibits a strategy whose wealth is unbounded below along the way, and would otherwise have to be called risk-free.

Lemma 4.8 (Admissibility buys a supermartingale)**.** *Let S/A be a continuous local martingale and Δ a strategy whose discounted wealth $\pi_t/A_t = \int_0^t \Delta \, d(S/A)$ satisfies $\pi_t/A_t \geq -a$ for a constant a. Then π/A is a supermartingale, and in particular $\mathbb{E}[\pi_T/A_T] \leq 0$.*

Proof. A stochastic integral against a continuous local martingale is a continuous local martingale, so $\pi/A + a$ is one too, and it is non-negative. A non-negative local martingale is a supermartingale — the lemma of chapter 1, which is conditional Fatou and nothing more — so $\pi/A + a$ is a supermartingale, and subtracting the constant leaves one. Its expectation is therefore at most its initial value, which is zero. □

Lemma 4.9 (Sufficient condition for no-arbitrage)**.** *If there is a measure $\mathbb{P}$ equivalent to the real world historic measure $\mathbb{P}^{historic}$ such that the discounted tradable asset process S_t/A_t is a local martingale under $\mathbb{P}$, then the market $\{S_t, M_t\}$ has no arbitrage.*

Proof. We show the two conditions defining an arbitrage,

$$\mathbb{P}^{\text{historic}}(\pi_T \geq 0) = 1 \qquad \mathbb{P}^{\text{historic}}(\pi_T > 0) > 0,$$

cannot hold together. The first step is the only place equivalence of the measures is used, and it is used twice over: equivalent measures agree on which events have probability zero, so both conditions may be rewritten under $\mathbb{P}$,

$$\mathbb{P}(\pi_T \geq 0) = 1 \qquad \mathbb{P}(\pi_T > 0) > 0.$$

This is why the theorem asks for an equivalent measure rather than merely some measure. A measure that assigned zero probability to a state could price away an arbitrage that occurs only in that state.

Now the discounted wealth is $\pi_T/A_T = \int_0^T \Delta_t \, d(S_t/A_t)$, a stochastic integral against a local martingale, and if the strategy is admissible then by Lemma 4.8 it is a supermartingale, so

$$\mathbb{E}^{\mathbb{P}}[\pi_T/A_T] \leq \pi_0/A_0 = 0.$$

An inequality is all that is available — a local martingale need not have constant expectation, and the doubling strategy later in this chapter is the case where it does not — and an inequality is all that is needed. A non-negative variable with non-positive expectation is zero almost surely, and $A_T > 0$, so $\mathbb{P}(\pi_T = 0) = 1$. That contradicts the second condition. □

Two things the proof did not need should be named, because their absence is what makes the direction easy. It did not need the market to be complete, and it did not need the measure to be unique: any one equivalent local martingale measure rules out arbitrage. Completeness is what decides whether the measure is unique, and §4.6 is about the case where it is not.

Together with its converse this gives the Fundamental Theorem of Asset Pricing.

Theorem 4.10 (First Fundamental Theorem of Asset Pricing)**.** *The market $\{S_t, M_t\}$ has no arbitrage if and only if there is a measure $\mathbb{P}$ equivalent to the real world historic measure $\mathbb{P}^{historic}$ such that the discounted tradable asset process S_t/A_t is a local martingale under $\mathbb{P}$.*

Where the two directions are proved. The statement as written is true in discrete settings and *false* in continuous time without repair, so it pays to be exact about what is proved where before proceeding.

The direction just established — measure implies no arbitrage — holds as stated, in every setting, for admissible strategies. It is the half that gets used, since a modelling exercise always starts by writing the measure down.

The converse is proved completely below in two finite settings, one period in Theorem 4.12 and a finite tree in Theorem 4.13, and the geometry there is the whole idea. In continuous time it fails: §4.5 exhibits a market with no arbitrage in the sense defined above and no equivalent martingale measure. The repair is to strengthen the hypothesis from no arbitrage to no free lunch with vanishing risk, and the resulting statement is the Delbaen-Schachermayer theorem of that section. So the honest reading of this theorem is that it is exactly right in finite dimensions and is the finite-dimensional shadow of a harder statement in infinite ones. □

4.3 From No Arbitrage to a Martingale Measure

What we proved above is the half of the theorem we actually use: hand me a martingale measure and I will tell you the market is safe. The other half says that the measure is not a lucky accident of the models we happen to write down — if a market admits no arbitrage at all, such a measure must exist. Without it, "find an equivalent martingale measure" would be a modelling convention. With it, it is the only thing one can do.

This direction is genuinely harder. So we will first prove it completely in a simpler setting — one period, finitely many states — then extend it to many periods, and only then look at what breaks in continuous time. The finite case is not a toy. It contains the whole idea, and the idea is geometric: the strategies available to you sweep out a flat subspace of payoffs, arbitrage means that subspace touches the positive orthant, and a subspace that misses the positive orthant can be separated from it by a plane. The coefficients of that plane, normalised, are the risk neutral probabilities. That is the entire proof.

Lemma 4.11 (Separating a subspace from the simplex)**.** *Let $K \subseteq \mathbb{R}^n$ be a linear subspace and let*

$$C = \left\{ x \in \mathbb{R}^n : x_i \geq 0 \text{ for all } i, \sum_{i=1}^{n} x_i = 1 \right\}$$

be the unit simplex. If $K \cap C = \emptyset$, then there exists $\lambda \in \mathbb{R}^n$ with $\lambda_i > 0$ for every i, such that $\lambda \cdot k = 0$ for every $k \in K$.

Proof. Consider the set of differences

$$D = C - K = \{c - k : c \in C, k \in K\}.$$

D is convex, being a sum of convex sets. It is also closed: if $c_m - k_m \to z$ with $c_m \in C$ and $k_m \in K$, then since C is compact we may pass to a subsequence along which $c_m \to c \in C$, whence

$k_m \to c - z$, and K is closed (every linear subspace of $\mathbb{R}^n$ is), so $c - z \in K$ and $z = c - (c - z) \in D$. Finally $0 \notin D$, since $0 = c - k$ would mean $c = k \in K \cap C$.

Let λ be the point of D closest to the origin; it exists because D is closed and non-empty, and $\lambda \neq 0$ because $0 \notin D$. We claim $\lambda \cdot z \geq |\lambda|^2$ for every $z \in D$. Indeed, by convexity $\lambda + t(z - \lambda) \in D$ for $t \in [0, 1]$, so by the choice of λ

$$\begin{aligned} |\lambda + t(z-\lambda)|^2 &\geq |\lambda|^2 \\ 2t\lambda \cdot (z-\lambda) + t^2|z-\lambda|^2 &\geq 0 \\ 2\lambda \cdot (z-\lambda) + t|z-\lambda|^2 &\geq 0 \end{aligned}$$

for every $t \in (0, 1]$, and letting $t \downarrow 0$ gives $\lambda \cdot z \geq \lambda \cdot \lambda = |\lambda|^2$.

Written out, this says

$$\lambda \cdot c - \lambda \cdot k \geq |\lambda|^2 > 0 \qquad \text{for all } c \in C,\ k \in K.$$

Now fix any $k \in K$. Since K is a subspace, $sk \in K$ for every real s, so $\lambda \cdot c - s(\lambda \cdot k) \geq |\lambda|^2$ for every s. A linear function of s that is bounded below must be constant, so $\lambda \cdot k = 0$. This holds for every $k \in K$, which is the second claim.

We are left with $\lambda \cdot c \geq |\lambda|^2 > 0$ for every $c \in C$. Taking c to be the i-th vertex of the simplex, the vector with a 1 in position i and zeros elsewhere, gives $\lambda_i > 0$. $\square$

Remark (Why the proof is arranged this way)**.** Forming $D = C - K$ looks like a trick until one asks what the alternative would be. Separating two sets requires the pair to be disjoint and at least one of them to be compact — K is a subspace and is neither bounded nor, on its own, easy to argue about. Subtracting collapses the two-set problem to a one-set problem: the single question "is the origin in D?", which is the same question as $K \cap C = \emptyset$ and much easier to attack, since the nearest point of a closed convex set to a point outside it always exists and is unique.

That is the entire idea. Everything after it in the proof is checking that D really is closed and convex, and then reading off consequences of λ being nearest.

Now the market. Take one period, from time 0 to time 1, and a sample space $\Omega = \{\omega_1, \ldots, \omega_n\}$ with finitely many states, each of which the historic measure gives positive probability — a state of probability zero can simply be deleted from the list. There are d risky tradables and a money market account. As before we work with discounted prices, writing $\tilde{S}_0 \in \mathbb{R}^d$ for today's discounted prices and $\tilde{S}_1(\omega) \in \mathbb{R}^d$ for tomorrow's; discounting makes the money market account worth 1 in every state, at both dates.

A portfolio is a vector $\theta \in \mathbb{R}^d$ of holdings in the risky assets, with whatever cash is needed to fund it borrowed or lent in the money market. Its discounted value at time 1, having started from $\pi_0 = 0$, is

$$\tilde{\pi}_1 = \theta \cdot (\tilde{S}_1 - \tilde{S}_0),$$

which is the one-period version of the identity $\pi_T/A_T = \int_0^T \Delta_t \, d(S_t/A_t)$ we derived for the continuous time model: a self-financing portfolio starting from nothing ends up holding exactly its gains from trading, measured in the numeraire.

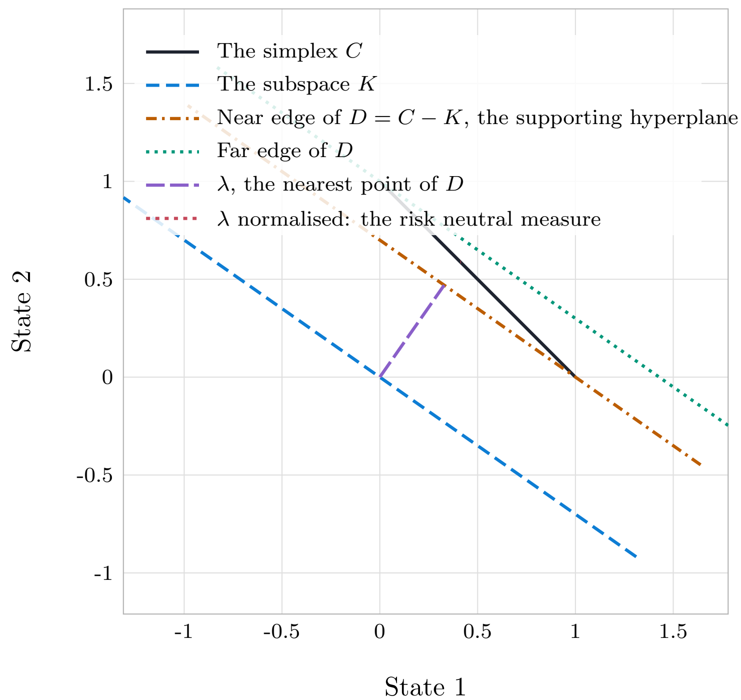

Figure 4.1: The lemma's proof in two dimensions, where K is a line and C is the segment joining the two axis vectors. The difference set $D = C - K$ is that segment swept along the line, which makes it an infinite strip parallel to K, and λ is its nearest point to the origin. Three of the proof's steps become things to look at rather than verify. λ is perpendicular to K because the nearest point of a strip is reached by walking straight at it. The supporting hyperplane is not a new object at all — it is the near edge of the strip, which is why the proof never has to build one. And λ falls in the positive quadrant, which is what makes it a measure rather than merely a separating direction. Rotate K into the first quadrant and the hypothesis fails in front of you: the line cuts the simplex, the strip swallows the origin, and the nearest point of D to the origin becomes the origin itself, leaving nothing to normalise.

Drawn by `separation`.

Theorem 4.12 (First Fundamental Theorem, one period and finitely many states). *The market $\{\tilde{S}_0, \tilde{S}_1\}$ admits no arbitrage if and only if there is a probability measure $\mathbb{Q}$ on Ω giving every state positive probability, such that*

$$\mathbb{E}^{\mathbb{Q}}[\tilde{S}_1] = \tilde{S}_0.$$

Proof. One direction is the lemma we already proved, and in this setting it is a single line: if such a $\mathbb{Q}$ exists then $\mathbb{E}^{\mathbb{Q}}[\tilde{\pi}_1] = \theta \cdot (\mathbb{E}^{\mathbb{Q}}[\tilde{S}_1] - \tilde{S}_0) = 0$ for every θ, and a random variable that is non-negative everywhere with mean zero under a measure that charges every state must be zero in every state.

For the converse, read a payoff as a vector in $\mathbb{R}^n$ by listing its value in each state, and let

$$K = \left\{ \left(\theta \cdot (\tilde{S}_1(\omega_1) - \tilde{S}_0), \ldots, \theta \cdot (\tilde{S}_1(\omega_n) - \tilde{S}_0)\right) : \theta \in \mathbb{R}^d \right\} \subseteq \mathbb{R}^n$$

be the set of discounted payoffs reachable from zero capital. It is the image of $\mathbb{R}^d$ under a linear map, so it is a linear subspace of $\mathbb{R}^n$ — this is the only thing we need about it, and it is why the geometry is so simple.

An arbitrage is precisely an element of K that is non-negative in every coordinate and non-zero. Suppose $x \in K$ were such an element. Then $\sum_i x_i > 0$, and since K is a subspace it contains $x / \sum_i x_i$, which lies in the simplex C. So an arbitrage exists if and only if $K \cap C \neq \emptyset$.

Assume then that there is no arbitrage, so $K \cap C = \emptyset$, and let λ be the vector supplied by Lemma 4.11. Set

$$q_i = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j},$$

which is well defined and strictly positive because every $\lambda_i > 0$, and sums to 1 by construction, so it defines a probability measure $\mathbb{Q}$ charging every state. That λ annihilates K says that for every $\theta \in \mathbb{R}^d$

$$\sum_{i=1}^n q_i \, \theta \cdot \left(\tilde{S}_1(\omega_i) - \tilde{S}_0\right) = 0$$
$$\theta \cdot \left(\mathbb{E}^{\mathbb{Q}}[\tilde{S}_1] - \tilde{S}_0\right) = 0,$$

and a vector orthogonal to every $\theta \in \mathbb{R}^d$ is the zero vector. □

With two states the whole argument fits on a page of graph paper.

Let the axes be the payoff in each of the two states. The attainable payoffs K are the multiples of a single vector, so they are a line through the origin. The simplex C is the segment joining $(1, 0)$ to $(0, 1)$. Arbitrage is a point of K with both coordinates non-negative and not both zero — so, after scaling, a point where the line crosses the segment.

The separating vector is not an abstraction either. In two dimensions the orthogonal complement of a line is a line, so there is exactly one candidate for λ, and the only question the lemma settles is whether it points into the positive quadrant. It does exactly when the price sits inside the range, and normalising it to sum to one gives

$$q = \left(\frac{S_{\text{now}} - S_d}{S_u - S_d}, \frac{S_u - S_{\text{now}}}{S_u - S_d} \right),$$

which is the q the binomial calculation produced at the start of the chapter. The geometry and the algebra are the same object.

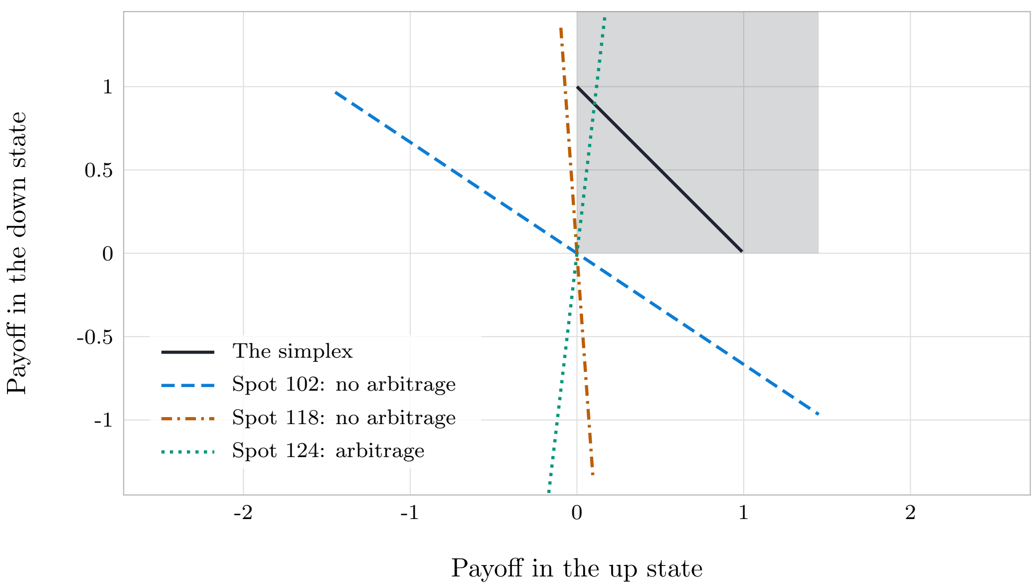

Figure 4.2: The whole of the theorem in one picture, drawn with the axes at equal scale so that the angles are the angles. The shaded region is the payoffs that are negative in neither state — money for nothing — and the segment across it is the simplex, which is that region normalised. Each line is the set of payoffs attainable from a zero-cost portfolio at one value of today's price, and every one of them passes through the origin, because doing nothing is always attainable. At a price inside the range of tomorrow's outcomes the line runs through the second and fourth quadrants and misses the shading entirely: there is no arbitrage, and Lemma 4.11's separating vector is the risk neutral measure. Push the price above every outcome and the line tips into the first and third, cutting the shaded region — and that intersection is the arbitrage, which is to say selling an asset that cannot rise.

Drawn by `TwoState::has_arbitrage` and `TwoState::risk_neutral`.

Remark. Return to the binomial example at the start of this chapter with this in hand. There $n = 2$ and $d = 1$, K is a line in the plane, and we found q by solving two linear equations for the replicating portfolio. Nothing in that calculation explained why q came out between 0 and 1; we simply observed that it had. The separation argument is what that observation was: q is positive exactly because the line K misses the simplex, and it misses the simplex exactly because there is no arbitrage. Had we chosen a forward price outside the range the market allows, K would have cut the positive quadrant, the "probability" would have come out negative, and the strategy realising the arbitrage would have been the point where it did so.

Exercise (Uniqueness and completeness)**.** In the setting of Theorem 4.12, show that $\mathbb{Q}$ is unique if and only if K, together with the constant payoffs, is all of $\mathbb{R}^n$ — that is, if and only if every payoff can be replicated. (Hint: the set of measures with $\mathbb{E}^{\mathbb{Q}}[\tilde{S}_1] = \tilde{S}_0$ is the intersection of the simplex with the orthogonal complement of K, and an affine set is a single point exactly when the subspace it is parallel to is trivial.) This is the Second Fundamental Theorem: a market is complete if and only if the martingale measure is unique.

4.4 Many Periods

Extending this to a finite tree costs surprisingly little, because a multi-period market is nothing more than a collection of one-period markets glued together, one at each node.

Let the market trade at times $t = 0, 1, \dots, T$, let Ω still be finite, and let $\mathcal{F}_t$ be the information available at time t. Since Ω is finite, $\mathcal{F}_t$ is generated by a partition of Ω into finitely many cells, and we call each cell a node at time t. Sitting at a node ν at time t, the market you face over the next period is exactly a one-period market: today's discounted prices $\tilde{S}_t(\nu)$ are known, and tomorrow's take one of finitely many values, one for each of the successor nodes of ν.

Theorem 4.13 (First Fundamental Theorem, finite tree)**.** *A finite multi-period market admits no arbitrage if and only if there is a measure $\mathbb{Q}$ equivalent to $\mathbb{P}^{historic}$ under which the discounted price process $\tilde{S}_t$ is a martingale.*

Proof. Again one direction is the lemma of the previous section. For the converse, suppose the market admits no arbitrage.

First, no one-period market at any node admits an arbitrage either. For suppose the market at node ν at time t did, realised by holdings θ. Consider the strategy: hold nothing until time t; if the state at time t is not in ν, continue to hold nothing; if it is, take the position θ for one period and then put whatever it yields into the money market and leave it there. This is self-financing, costs nothing to set up, and its discounted value at T is the payoff of the node arbitrage — non-negative always, and strictly positive on a set of states which has positive probability, since ν itself is reached with positive probability. That is an arbitrage in the multi-period market, which we assumed there is none of.

So Theorem 4.12 applies at every node, giving for each node ν at each time t a strictly positive probability $q(\mu \mid \nu)$ on the successors μ of ν, with

$$\sum_{\mu} q(\mu \mid \nu)\, \tilde{S}_{t+1}(\mu) = \tilde{S}_t(\nu).$$

Define $\mathbb{Q}$ by multiplying these together along each path: for $\omega \in \Omega$, let $\nu_0, \nu_1, \ldots, \nu_T$ be the nodes it passes through and set

$$\mathbb{Q}(\omega) = \prod_{t=0}^{T-1} q(\nu_{t+1} \mid \nu_t).$$

This is a probability measure, being a product of conditional probabilities down a tree, and it is strictly positive on every state because each factor is, so it is equivalent to $\mathbb{P}^{\text{historic}}$. And by construction the conditional distribution of ν_{t+1} given ν_t under $\mathbb{Q}$ is $q(\cdot \mid \nu_t)$, so the displayed identity above is exactly

$$\mathbb{E}^{\mathbb{Q}}[\tilde{S}_{t+1} \mid \mathcal{F}_t] = \tilde{S}_t,$$

which is the martingale property. □

Remark. The restriction to finitely many states is not needed. With finitely many trading dates but arbitrarily many states, no arbitrage still implies the existence of an equivalent martingale measure; this is the Dalang-Morton-Willinger theorem. The geometry is the same and the work is in replacing "closed because it is a subspace of $\mathbb{R}^n$" with a genuine closedness argument, which is the whole difficulty. That should be a warning about what is coming next.

4.5 What Changes in Continuous Time

The proof above rested on two facts, and in continuous time both of them fail.

The first is that the set of attainable payoffs was closed. In $\mathbb{R}^n$ it was closed for free, being a subspace. With continuously many trading dates the attainable payoffs form a subspace of an infinite dimensional space, and such a subspace need not be closed at all. The separating hyperplane theorem needs closedness, so the argument stops applying.

The second, and the more interesting one, is that the notion of arbitrage we have been using is too weak once trading is continuous. To see this, forget the market for a moment and consider a single Brownian motion W_t on $[0, 1)$, which is a martingale, so on the face of it nothing should be extractable from it.

Example 4.2 (A doubling strategy)**.** Fix a target of \$1 and bet on W until the winnings first reach it. The obstacle is that the first passage

$$\tau = \inf\{t : W_t = 1\}$$

is finite almost surely but has no bound — its mean is infinite — so holding one unit does not finish by time 1 with certainty. The repair is to compress the clock. Take the deterministic position

$$\Delta_t = \frac{1}{1-t}, \qquad 0 \le t < 1, \tag{4.4}$$

and let $\pi_t = \int_0^t \Delta_s \, dW_s$. This is a continuous local martingale, and the quadratic variation of a stochastic integral is the integral of the squared integrand, so

$$\langle \pi \rangle_t = \int_0^t \frac{ds}{(1-s)^2} = \frac{t}{1-t},$$

which runs to infinity as $t \uparrow 1$. By the Dambis-Dubins-Schwarz theorem of chapter 2 there is then a Brownian motion B with

$$\pi_t = B_{t/(1-t)}.$$

The wealth is a Brownian motion on a clock that reaches infinity before calendar time 1. Nothing has been assumed about the size of the position beyond (4.4): the compression is a consequence of the quadratic variation, and the position is the square root of the clock's speed.

Now B attains the level 1 at some almost surely finite clock time τ_B, and the calendar time corresponding to it, $t^* = \tau_B/(1+\tau_B)$, is almost surely strictly less than 1. So stop there:

$$\Delta_t = \frac{1}{1-t}\,\mathbf{1}_{\{t<t^*\}}.$$

This is adapted, since t^* is a stopping time, and the integral is defined on the whole of $[0,1]$ because the integrand is zero after t^*. The wealth satisfies $\pi_0 = 0$ and $\pi_1 = 1$ almost surely.[1] It is an arbitrage, in the exact sense of our definition, in a market whose only asset is a martingale.

Nothing is wrong with the mathematics; what is wrong is the definition. The strategy works by being prepared to lose an unbounded amount along the way, which no counterparty would fund and no exchange would clear — and "unbounded" can be made precise.

Calculation 4.14 (What the strategy has to be willing to lose)**.** Write M for the lowest the wealth reaches before the target is struck. By the gambler's ruin identity for Brownian motion, the chance of falling to $-a$ before rising to $+1$ is

$$\mathbb{P}\big(M \le -a\big) = \frac{1}{1+a}, \tag{4.5}$$

a tail decaying like $1/a$.[2] The drawdown $-M$ is non-negative, and for such a variable the mean is the area under the tail,

$$\mathbb{E}[X] = \mathbb{E}\left[\int_0^\infty \mathbf{1}_{\{a<X\}}\,da\right] = \int_0^\infty \mathbb{P}(X>a)\,da,$$

exchanging the two integrals by Tonelli, which is available because the integrand is non-negative. It is the same expectation as $\int x\,dF(x)$, integrated by parts, and it is the convenient form here because the barrier law is a tail. So

$$\mathbb{E}\big[-M\big] = \int_0^\infty \mathbb{P}\big(-M>a\big)\,da = \int_0^\infty \frac{da}{1+a} = \infty.$$

The divergence is logarithmic, which is the useful form of the statement: the float needed to survive a drawdown of a is about $\ln(1+a)$, so tolerating ten times the drawdown costs only a little over twice the capital, and no finite capital suffices.[3]

---

[1] `simulated`: the fraction reaching the target matches the reflection principle's $2(1-\Phi(1/\sqrt{U}))$ at clock horizon U, which is what confirms the wealth really is a Brownian motion on the stretched clock, and the calendar times at which it does are inside the interval.

[2] `checked by simulation`. The grid has to be fine: a discretised path overshoots the barrier it crosses, which biases the probability upwards by a few per cent at these depths.

[3] `the capped mean` $\mathbb{E}[\min(-M,a)] = \ln(1+a)$`, simulated`. The uncapped mean cannot be simulated at all: every run stops somewhere, stopping censors the tail that carries the divergence, and the sample mean then converges to a finite number and converges to the wrong one.

Equation (4.5) is why the repair is a bound rather than a moment condition. Requiring finite expected loss would not exclude the strategy, since one could stop it early and lose only a little on average; requiring the wealth to stay above a fixed level does exclude it, and excludes it for every level, however large. It is the continuous time version of doubling your stake at roulette until you win, and it is why the wealth process must be required to stay above some fixed level: Definition 4.7 already named the restriction, for exactly this reason, and Lemma 4.8 supplied the one fact the sufficiency argument needed from it: an admissible strategy's discounted wealth is a supermartingale, so its expectation can only fall from zero.

That is all the sufficiency argument ever needed, and it is exactly what the doubling strategy does not have: its wealth is bounded below by no constant, so there is no a to add and Fatou has nothing to work on. When the discounted price has jumps the first sentence of the proof needs more care — a stochastic integral against a local martingale need not be a local martingale — and the statement that repairs it, that a bounded-below integral against a local martingale is a local martingale, is the Ansel-Stricker theorem. The conclusion is the same.

Even after restricting to admissible strategies, no arbitrage is still strictly weaker than the existence of an equivalent martingale measure. One can construct markets with no arbitrage in which no such measure exists: the reason is again closedness, and the failure shows up as a sequence of strategies whose payoffs are not arbitrages but converge to one, the losses required shrinking to zero along the way. Ruling this out is the right condition.

Definition 4.15 (No Free Lunch with Vanishing Risk)**.** A market satisfies NFLVR if there is no sequence of discounted payoffs f_n of admissible strategies starting from zero, and no sequence of constants $\epsilon_n \to 0$ with $f_n \geq -\epsilon_n$, such that f_n converges uniformly to some f with $f \geq 0$ and $\mathbb{P}(f > 0) > 0$.

Read it as: not only can you not make something out of nothing, you cannot get arbitrarily close to doing so with an arbitrarily small stake at risk. It implies no arbitrage — take $f_n = f$ and $\epsilon_n = 0$ — and is strictly stronger.

Theorem 4.16 (Delbaen and Schachermayer)**.** *Let S be a locally bounded semimartingale. The market satisfies NFLVR if and only if there is a measure equivalent to $\mathbb{P}^{historic}$ under which the discounted price process is a local martingale.*

How the proof is organised. The theorem splits into two steps of very unequal difficulty, and only one of them is deep.

- *Step one.* NFLVR implies that the cone of attainable claims is closed in the right topology.
- *Step two.* Given that closedness, a separation argument produces the measure.

Step two is the continuous time version of the separating hyperplane already used, it needs nothing beyond the Hahn-Banach theorem, and it is carried out in full below as Theorem 4.17. Step one is the deep part; it is not proved here, and the remark following the separation argument says what its proof consists of and which of its ingredients have no finite-dimensional counterpart. □

Theorem 4.17 (Kreps and Yan: closedness gives the measure)**.** *Let $K \subset L^\infty$ be the cone of discounted payoffs of admissible strategies starting from zero, and let*

$$C = \left(K - L^\infty_+\right) \cap L^\infty$$

be the claims that can be superreplicated from nothing. Suppose C is weak- closed and contains no non-negative claim other than zero, that is $C \cap L^\infty_+ = \{0\}$. Then there is a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ with $\mathbb{E}^{\mathbb{Q}}[f] \le 0$ for every $f \in C$.*

Proof. Fix any event B with $\mathbb{P}(B) > 0$ and consider the indicator $\mathbb{I}_B$, which lies in L^∞_+ and, by hypothesis, not in C. Since C is weak-* closed and convex and $\{\mathbb{I}_B\}$ is compact and convex, the Hahn-Banach separation theorem applied in the duality between L^∞ and L^1 gives a $g_B \in L^1$, not identically zero, and a real c with

$$\mathbb{E}\big[g_B f\big] \;\le\; c \;<\; \mathbb{E}\big[g_B \mathbb{I}_B\big] \qquad \text{for every } f \in C.$$

Two properties of g_B follow from the shape of C rather than from the separation. Because C is a *cone*, $f \in C$ implies $\lambda f \in C$ for every $\lambda > 0$, so the left inequality forces $\mathbb{E}[g_B f] \le 0$ and c may be taken to be zero. And because C contains $-L^\infty_+$, taking $f = -h$ for arbitrary $h \ge 0$ gives $\mathbb{E}[g_B h] \ge 0$, so $g_B \ge 0$ almost surely.

So each B supplies a non-negative g_B that is strictly positive somewhere on B and integrates non-positively against everything attainable. What remains is to assemble one density that is strictly positive *everywhere*, and that is a counting argument. Let $\mathcal{G}$ be the family of such g, normalised to have $\mathbb{E}[g] = 1$, and let β be the supremum of $\mathbb{P}(g > 0)$ over $\mathcal{G}$. Choose g_n approaching the supremum and set

$$g = \sum_{n \ge 1} 2^{-n} g_n,$$

which converges in L^1, lies in $\mathcal{G}$ since $\mathcal{G}$ is convex and closed under such combinations, and has $\{g > 0\} = \bigcup_n \{g_n > 0\}$, so $\mathbb{P}(g > 0) = \beta$. If $\beta < 1$ then applying the first part to $B = \{g = 0\}$ produces a g_B strictly positive somewhere on B, and $\frac{1}{2}(g + g_B)$ would have $\mathbb{P}(\cdot > 0) > \beta$, contradicting the supremum. So $\beta = 1$, $g > 0$ almost surely, and $d\mathbb{Q} = g\, d\mathbb{P}$ is the measure required. $\square$

Remark (What the hard step actually is, and why)**.** Theorem 4.17 assumed the closedness. Supplying it is the whole difficulty.

The set K is a set of *stochastic integrals*, indexed by the strategies producing them. To show it is closed one takes a sequence of payoffs $f_n = \int \Delta_n\, dS$ converging to some f, and must produce a single strategy Δ with $\int \Delta\, dS = f$. There is no reason for the integrands Δ_n to converge to anything — they live in a space with no useful compactness — so the limit strategy has to be constructed rather than extracted, and constructing it is what Delbaen and Schachermayer's proof does. It uses NFLVR to get a uniform bound on the strategies, then a compactness theorem for the resulting integrals.

The finite dimensional case had none of this because a linear subspace of $\mathbb{R}^n$ is closed automatically. That single sentence is the entire difference in difficulty between the two halves of this chapter.

It helps to know roughly what the construction costs, both to see why it is not reproduced here and to see that nothing conceptually new enters. Three ingredients do the work. NFLVR is converted into a bound: the terminal values of admissible strategies are bounded in probability, which is where the vanishing-risk sequences are ruled out and where the strengthened hypothesis earns its place. That bound feeds a compactness argument of Komlós type, which extracts from a bounded sequence not a convergent subsequence — there is none — but a sequence of convex combinations whose Cesàro averages converge almost surely; this is the step with no finite-dimensional analogue, because in $\mathbb{R}^n$ boundedness already gives convergent subsequences. The limit is then identified as a stochastic integral, which needs a result on when a limit of integrals is an integral, and Ansel and Stricker's theorem on bounded-below integrals against local martingales — the same one that repaired the jump case in Lemma 4.8 — is what closes it.

One further honesty. If S is not locally bounded the conclusion weakens: what NFLVR delivers is an equivalent *sigma-martingale* measure, and a sigma-martingale is a stochastic integral against a martingale which need not itself be a local martingale. For the diffusions in these notes the distinction never arises, and it is mentioned only so that the statement above is not read as more general than it is.

Structure (The same three moves, twice)**.** The continuous case is not a different argument but the same one carried out in a harder space.

	Finite	**Continuous**
the space	$\mathbb{R}^n$	L^∞, paired with L^1
closedness	free: a subspace	the theorem, and it is hard
separation	separating hyperplane	Hahn-Banach, Theorem 4.17
positivity of the normal	from the cone $\mathbb{R}^n_+$	from $C \supset -L^\infty_+$
strict positivity	automatic in finite dimension	the exhaustion argument

Every row but the second is a translation. The second is a theorem. The summary is that the fundamental theorem is a separation argument wearing a hard closedness result as a hat, and a reader who has understood the finite case has understood the shape of the general one — while a reader who wants the general one in full needs a monograph.

Two features of the statement explain the shape of everything we do afterwards.

The first is that the conclusion is a *local* martingale, not a martingale. This is not a technicality to be waved away: local martingales that are not martingales appear in perfectly reasonable models, and when one does, prices computed as expectations can fail to be the prices of the tradables they are supposed to be. Dropping local boundedness weakens the conclusion further, to a σ-martingale.

The second is that this is why the notes are organised the way they are. From here on we will always work in the direction the easy half licenses: we will posit a model, find the measure in which discounted tradables are martingales, and price by taking expectations. The theorem is what tells us that in doing so we are not choosing one pricing convention among many, but following the only one a market without arbitrage permits.

4.6 Splitting the Risk in Two

A jump model does something more useful than price the jump, and it answers the question a desk actually asks: of the risk in this trade, how much can I hedge by trading continuously, and how much must I hedge with an instrument that pays on the event itself?

The answer is precise. Under the pricing measure the claim's value V_t is a martingale, and in a market driven by a Brownian motion and a jump measure every martingale splits into a piece driven by each:

$$dV_t = \xi_t \, dW_t + \int \psi_t(x) \, \tilde{N}(dt, dx). \tag{4.6}$$

Here ξ is the claim's sensitivity to continuous movement and $\psi(x)$ its sensitivity to a jump of size x — the amount the value changes if that jump happens now.

Suppose the only thing we can trade is the underlying S, which has both kinds of movement itself. That is one instrument against two sources of risk, so we cannot expect to remove both. What we can do is remove as much as possible, and the statement of how much is the Galtchouk-Kunita-Watanabe decomposition:

$$V_T = V_0 + \int_0^T \phi_t \, dS_t + L_T, \qquad \langle L, S \rangle = 0. \tag{4.7}$$

The integral is everything trading S can deliver. The residual L is orthogonal to S — no position in S has any effect on it — and is exactly the part that continuous trading cannot touch. Its size $\mathbb{E}[L_T^2]$ is the variance of the best possible hedging error, a number that can be computed and reserved against as chapter 24 describes.

Calculation 4.18 (The hedge ratio, derived)**.** Take a price with both kinds of motion,

$$\frac{dS}{S} = \mu \, dt + \sigma \, dW + \int \big(e^x - 1\big) \, N(dt, dx),$$

and hold ϕ units of it against a claim worth $V(S)$. Over an instant the hedged position moves by

$$d\Pi = \underbrace{\Big(\frac{\partial V}{\partial S} - \phi\Big)\sigma S \, dW}_{\text{diffusive}} + \underbrace{\int \Big[V(Se^x) - V(S) - \phi S(e^x - 1)\Big] N(dt, dx)}_{\text{on a jump}} + (\ldots)\, dt. \tag{4.8}$$

Two features of (4.8) decide everything that follows. The diffusive term can be set to zero exactly, by one choice of ϕ, and that choice is the delta. The jump term cannot be set to zero by any choice, because it must vanish for *every* x at once and ϕ is one number. So perfect hedging is unavailable and the question changes from eliminating the error to minimising it.

Minimise its variance. The two sources are independent, so the variance per unit time adds:

$$\mathrm{Var}\big[d\Pi\big]/dt = \Big(\frac{\partial V}{\partial S} - \phi\Big)^2 \sigma^2 S^2 + \int \Big[V(Se^x) - V(S) - \phi S(e^x - 1)\Big]^2 \nu(dx),$$

using that a Poisson random measure of intensity ν contributes $\int f(x)^2 \nu(dx)$ to the variance rate of $\int f \, dN$. This is a quadratic in ϕ with a positive coefficient, so it has one minimum, found by differentiating:

$$-2\Big(\frac{\partial V}{\partial S} - \phi\Big)\sigma^2 S^2 - 2\int \Big[V(Se^x) - V(S) - \phi S(e^x - 1)\Big] S(e^x - 1)\, \nu(dx) = 0.$$

Collecting the terms in ϕ on one side,

$$\phi \left[\sigma^2 S^2 + \int S^2 (e^x - 1)^2 \nu(dx)\right] = \sigma^2 S^2 \frac{\partial V}{\partial S} + \int S(e^x - 1)\big[V(Se^x) - V(S)\big]\nu(dx).$$

The bracket on the left is the total variance rate of S, and the right hand side is the covariance rate of V with S — so the minimising ϕ is $d\langle V, S \rangle / d\langle S \rangle$, and no separate argument is needed to identify it with the projection of (4.7).

Structure (A hedge ratio is a regression coefficient)**.** Written as $\phi = d\langle V, S\rangle / d\langle S\rangle$ the answer is a familiar object in unfamiliar clothing: it is the slope of a least squares regression of the claim's moves on the hedge instrument's moves. Covariance over variance is what a regression coefficient is.

That is the right way to hold the whole subject. In a complete market the regression happens to be exact — the residual is zero, the coefficient is a derivative, and calling it "delta" hides the fact that it was ever a projection. Once the market is incomplete the residual is not zero and the coefficient stops being a derivative, but it is the same construction throughout. Nothing new appears with jumps; what disappears is the coincidence that made the projection look like differentiation.

Reading the ratio as a regression also says immediately what it optimises and what it does not. It minimises variance, which is a symmetric penalty: it treats an unexpected gain as costly as an unexpected loss, and a desk short a jump does not feel that way. A hedge chosen to minimise a one-sided measure — chapter 24's expected shortfall, say — is a different number, and the difference is largest exactly where the payoff is most asymmetric across the jump.

Remark (The hedge ratio is not the delta)**.** Reading 4.18's answer term by term,

$$\phi = \frac{\sigma^2 S^2 \dfrac{\partial V}{\partial S} + \displaystyle\int S(e^x - 1)\big[V(Se^x) - V(S)\big]\,\nu(dx)}{\sigma^2 S^2 + \displaystyle\int S^2 (e^x - 1)^2\,\nu(dx)}.$$

the numerator's first term is the ordinary delta weighted by diffusive variance; the second is what best offsets the jumps on average, weighted by jump variance. The denominator is the total variance of S from both sources. So ϕ is a variance-weighted average of two ratios, and the weights are the shares of S's variance that each source contributes.

The second ratio deserves a look on its own. For a single jump size x it collapses to

$$\frac{V(Se^x) - V(S)}{S(e^x - 1)},$$

which is the *chord* of V between the pre-jump and post-jump levels. So the jump-hedging ratio is a secant where the delta is a tangent, and the two agree only where V is affine over the interval the jump crosses. For a linear product they coincide and none of this matters; for anything with curvature they do not, and the gap grows with the size of the jump and with the gamma.

So ϕ is a compromise, and it is neither the delta nor a jump hedge. Setting the position to $\partial V/\partial S$ in a market that jumps is not a conservative approximation — it is the wrong ratio, and it is wrong in the direction of ignoring the jumps entirely, since that is the limit it comes from.

Now add an instrument that pays on the event. Whether the residual disappears depends on something we already have a name for: whether the jump is one size or many.

Default. With a known recovery there is a single jump, of a single size, at an unpredictable time. So (4.6) has one ψ rather than a function of x, and there are two risks and — once a credit default swap on the same name is available — two instruments. The system is square. Solving it gives an exact hedge: a position in the underlying for the diffusive part and a determined notional of protection for the default. The residual L vanishes and the market is complete again.

This is why a credit desk reports *jump to default* as a risk in its own right. It is ψ for the default jump: the profit and loss that would arrive if the name defaulted this instant, after the continuous hedge has done what it can. It sits beside the delta on the risk report rather than inside it, because the decomposition says they are different risks requiring different instruments.

A scheduled event, in general. The date is known but the move is not, and in general it can be any size. Then $\psi(\cdot)$ is a genuine function on a continuum and spanning it would need an instrument for every size. Two or three cannot do it.

What they can do is span the first few moments. A future dated across the event hedges the expected move; a straddle expiring just after it hedges the convexity of the outcome; a strangle reaches further into the tails. Each instrument removes another moment of ψ and shrinks L without eliminating it. That is why hedging an event requires instruments expiring on the far side of it, and why each additional strike buys one more moment of the distribution.

And what a policy meeting actually is. The general case overstates the difficulty for the most important scheduled event in rates, and the reason repays attention: it moves the problem into territory we have already solved.

A central bank does not move its policy rate by an arbitrary amount. It moves in multiples of twenty-five basis points, and at the great majority of meetings the outcome is one of

$$\{\, -25\text{bp}, \quad 0, \quad +25\text{bp}\,\},$$

with fifty and occasionally seventy-five appearing only at turning points. So ψ is not a function on a continuum at all. It is a short list of numbers, and the state space of the event is finite and small.

That is this chapter's setting exactly. With n outcomes the market is complete precisely when the traded payoffs span $\mathbb{R}^n$, and n here is three or four rather than infinite — so a handful of instruments expiring across the meeting is not an approximation to spanning the event, it is spanning the event. The separating hyperplane picture drawn above is the right picture, in three dimensions rather than two.

It also means the risk neutral probabilities are not merely assumed to exist. They are quoted.

Example 4.3 (Reading the hike probability off the futures)**.** A fed funds future settles on the *average* effective rate over its delivery month, which is what makes the arithmetic work. Take a target rate of 4.00%, a thirty day month, and a meeting on the tenth. If the rate is r_1 after the meeting, the average over the month is

$$\frac{10 \times 4.00 + 20 \times r_1}{30},$$

so a future implying an average of 4.15% implies

$$r_1 = \frac{30 \times 4.15 - 10 \times 4.00}{20} = 4.225\%.$$

With only two outcomes on the table — hold at 4.00 or hike to 4.25 — the implied rate is a weighted average of the two, so

$$q_{\text{hike}} = \frac{4.225 - 4.00}{0.25} = 90\%.$$

Compare this with the binomial calculation that opens this chapter. There we solved for q from two possible values and a traded price, and observed that the historic probabilities never appeared.

Here the market has done the solve for us and publishes the answer daily. The number quoted as "the probability of a hike" is not anyone's forecast; it is the q derived above, extracted from a price, and it will differ from the true probability by whatever risk premium the market demands.

Where the continuum comes back. None of which means the general case was wasted, because it applies as soon as the exposure is to something other than the policy rate itself.

What moves on a meeting day is not only the target. The statement, the projections and the press conference revise the expected path of *all* future meetings, and a ten year swap rate aggregates those revisions into a number that can take any value. So a position in fed funds futures faces three outcomes and is exactly hedgeable; a Bermudan swaption faces a continuum on the same afternoon and is not.

Which case one is in is decided, as usual, by the product rather than by the event.

Remark (Which hedge, though)**.** One honesty is owed here. In an incomplete market the pricing measure is not unique, and ϕ in (4.7) depends on which one was chosen and on what "best" was taken to mean. Minimising the variance of the terminal error, minimising it period by period, and super-replicating the payoff give three different hedges, and they can differ materially.

So a jump model does not hand over a hedge the way the Black-Scholes model does. It hands over a decomposition — this much is hedgeable, that much is not, and here is the instrument that would span the remainder — together with a choice about the residual that has to be made on other grounds. That is more useful than it sounds, and it is considerably more honest than a model that reports the unhedgeable part as zero because it cannot represent it.

It is a real restriction, and it deserves to be stated clearly. It is not that a jumping asset is beyond reach — credit, as above, is modelled with the same tools — but that the specific chain of reasoning in these notes, in which no arbitrage pins down a single price, needs the continuous case. Where risk arrives suddenly rather than by accumulation, a continuous model can be calibrated to the right prices and will still misstate the hedge, because it is hedging a move it does not believe can happen.

4.7 Why Every Model Here Is a Diffusion

There is one loose end, and it should be tied off before the modelling starts, because otherwise it will look like a habit rather than a result.

Every model in these notes writes the price as

$$dS_t = \mu_t \, dt + \sigma_t \, dW_t,$$

and a reader is entitled to ask what licenses that. It is a very specific shape. Why should a price take it?

The answer assembles from two things we now have, and it turns out that almost nothing is being assumed.

First, the class is forced, not chosen. A price process that is not a semimartingale admits a free lunch with vanishing risk. This is a theorem, not a convention — it follows from the Bichteler-Dellacherie theorem, and is part of what Delbaen and Schachermayer established. So a model that is arbitrage-free has no option: its price must be a semimartingale. Nothing has been assumed yet; this is the market's constraint speaking.

Second, a continuous semimartingale has only one possible shape. Every semimartingale decomposes into a local martingale plus a process of finite variation, and if the price path is continuous both pieces are continuous:

$$S_t = S_0 + \underbrace{M_t}_{\text{local martingale}} + \underbrace{A_t}_{\text{finite variation}}.$$

Chapter 2 dealt with the first piece. By Lévy's characterization and its consequence there, a continuous local martingale accumulating variance at a rate σ_t^2 is *necessarily* $M_t = \int_0^t \sigma_s dW_s$ for some Brownian motion W — there is no other continuous local martingale it could be. And a continuous finite-variation process that accumulates at a rate is $A_t = \int_0^t \mu_s ds$. Adding them,

$$S_t = S_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s,$$

which is the Itô diffusion, arrived at rather than posited.

Remark (What is actually being assumed)**.** Reading the argument backwards is the useful exercise, because it isolates the one place a choice was made.

No arbitrage was not a choice — it is the premise of the subject. The decomposition was not a choice — it is a theorem. Lévy's characterization was not a choice. The only assumption in the chain is *continuity of the path*, and the only thing it excludes is a jump.

So the entire content of "we model prices as diffusions" is the claim that the price does not jump. That is an empirical claim about an asset. It is reasonable for a liquid index over a quiet afternoon. It is plainly false for a bond that can default, for a stock through an earnings announcement, for a currency with a peg that may break, and for anything over a horizon long enough to contain a crisis.

Remark (What lies outside)**.** When the assumption fails, the honest response is to enlarge the model rather than to reinterpret the volatility, and the shape of the enlargement is dictated by the same decomposition: what a general semimartingale has and a continuous one does not is a jump term. Adding it gives

$$dS_t = \mu_t \, dt + \sigma_t \, dW_t + dJ_t,$$

with J built from a Poisson random measure. Merton's jump diffusion and the variance gamma model are two such, and the general theory of Lévy processes is the study of what J may be.

One consequence deserves flagging now, because it changes the character of everything that follows. With jumps the market is generically *incomplete*: the underlying and the money market can no longer replicate an arbitrary payoff, because a jump of random size cannot be hedged by a position chosen before it happens. By the exercise on uniqueness above, the martingale measure is then no longer unique — there is an interval of arbitrage-free prices rather than a single one, and choosing within it is a modelling decision that no amount of no-arbitrage reasoning will make for you.

Incompleteness is relative to the instruments available, though, and markets respond to it by trading the missing risk directly — chapter 2's remark on credit default swaps makes the point. So the interval of prices narrows as the traded set grows, and the practical question is never "is this model complete" but "which instrument spans the risk I cannot hedge".

That is the honest reason these notes stay with continuous models. Not that jumps are unimportant — they are how credit is modelled, and how any asset with a scheduled announcement behaves — but that continuity is what buys the uniqueness that makes this particular line of argument as sharp as it is.

References

- Harrison, J. M., & Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3), 381–408.
- Harrison, J. M., & Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3), 215–260.
- Dalang, R. C., Morton, A., & Willinger, W. (1990). Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics and Stochastic Reports*, 29(2), 185–201.
- Delbaen, F., & Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1), 463–520.
- Delbaen, F., & Schachermayer, W. (1998). The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen*, 312(2), 215–250.
- Bichteler, K. (1981). Stochastic integration and L^p-theory of semimartingales. *Annals of Probability*, 9(1), 49–89.