

AN ELEMENTARY INTRODUCTION TO CHANGES OF NUMERAIRE

SIMON ELLERSGAARD NIELSEN

1. AN OVERVIEW

1.1. The Philosophy. One of the murkiest parts of any introductory level course on stochastic calculus for finance is indubitably the process of changing the probability measure in the pricing process of financial derivatives (otherwise known as the change in numeraire). Whilst it might appear natural to people well versed in a measure theoretic probability theory, the governing ideas underpinning this procedure are mostly too opaque to the general reader to establish any kind of genuine comprehension. Thus, the intention of the following document is to facilitate greater understanding through an informal overview of some integral results of the numeraire. Accordingly, any sense of mathematical rigor will be fleeting (or illusory), considering that a plethora of more technical exegeses of the subject already exists.

1.2. Reminder. Our first elementary encounter with the risk neutral probability measure \mathbb{Q} in continuous time finance usually comes from considering what it would take to transform the drift term of a traded asset to the risk free rate. The answer turns out to be a mapping of the real world Wiener process $dW_t^{\mathbb{P}}$ into $dW_t^{\mathbb{Q}} - \gamma dt$ where γ is the Sharpe ratio, $(\mu_t - r_t)/\sigma_t$. Importantly, this probabilistic transformation transforms all *discounted* traded asset into \mathbb{Q} martingales. Formally, if χ_s is the time s value of our contingent claim, and $B_s^{(0)}$ is the money account $\exp(\int_0^s r_u du)$, then $d(\chi_t/B_t^{(0)})$ is driftless, or identically

$$(1) \quad \frac{\chi_t}{B_t^{(0)}} = \mathbb{E}_t^{\mathbb{Q}} \left[\frac{\chi_T}{B_T^{(0)}} \right].$$

A simple re-arrangement of this equation gives us the all important Feynman-Kac formula for derivative pricing

$$(2) \quad \chi_t = \mathbb{E}_t^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_u du \right) \chi_T \right],$$

which inter-alia is employed in the derivation of the Black-Scholes formula. Rather remarkably, it can be shown that postulating the existence of such a measure \mathbb{Q} (under which discounted traded assets become martingales) is equivalent to a no arbitrage condition on the market. This result is so important that the literature has labelled it *The First Fundamental Theorem*. Alas, it is also non-trivial to prove rigorously wherefore we shall accept it at face value.

2. A GENERALIZED NUMERAIRE

2.1. Beyond the Money Account. One must inevitably ponder whether it is merely incidental that we can price our derivatives by deflating them by the money account. Could one conceivably obtain a pricing equation like (1) using some other numeraire, that is to say some other reference asset qua which other assets are priced? The answer is a resounding yes. Not only can we deploy any **non-dividend paying traded asset** in our pricing equation; there is also the chance that it will reduce

the computational complexity considerably if aptly chosen vis-a-vis the contingent claim we wish to price. Explicitly, let $A^{(i)}$ be any one such asset, then any $A^{(i)}$ -deflated asset is a $\mathbb{Q}^{(i)}$ martingale and thus the contingent claim χ can be priced as

$$(3) \quad \frac{\chi_t}{A_t^{(i)}} = \mathbb{E}_t^{\mathbb{Q}^{(i)}} \left[\frac{\chi_T}{A_T^{(i)}} \right]$$

(here $\mathbb{Q}^{(i)}$ is the probability measure associated with $A^{(i)}$). Of course, recalling that the dynamics of the underlying usually is expressed in terms of more familiar measures such as \mathbb{P} or \mathbb{Q} this immediately raises the question how the expected value in (3) is computed in practice. Clearly, we should be able to find a function which maps our initial Wiener process into a random variate under the desired measure. Fortunately, *Girsanow's Theorem* provides a straightforward answer to this question: if dW_t^\odot captures the dynamics of our initial contingent claim (where \odot typically will be \mathbb{P} or \mathbb{Q}), then $dW_t^{\mathbb{Q}^{(i)}}$ can be expressed as a linear function thereof. In particular,

$$(4) \quad dW_t^\odot = \varphi_t dt + dW_t^{\mathbb{Q}^{(i)}}$$

where φ is a function which can be determined from the ratio of the probability distributions of the two measures. Nonetheless, for our present purposes the linearity property will suffice for computational reasons, as we are only interested in making $\chi_t/A_t^{(i)}$ a martingale under $\mathbb{Q}^{(i)}$. And (4) is precisely telling us that if we have the stochastic differential equation, $d(\chi_t/A_t^{(i)})$, under \odot then all we have to do is to cross out the drift term and replace \odot with $\mathbb{Q}^{(i)}$ in the diffusion. A few examples should make this considerably clearer:

2.1.1. The Exchange Option. An exchange option gives the holder the right, but not the obligation, to exchange one unit of the held stock $S^{(1)}$ for one unit of some other stock $S^{(2)}$ at time T . At maturity the pay-off is clearly $\chi_T^{\text{exch}} = \max\{S_T^{(2)} - S_T^{(1)}, 0\}$. Assume that the stocks follow Geometric Brownian Motion $dS_t^{(i)} = rS_t^{(i)}dt + \sigma_i S_t^{(i)} dW_{i,t}^\mathbb{Q}$, where $r, \sigma \in \mathbb{R}^+$ and $dW_{1,t}^\mathbb{Q} \cdot dW_{2,t}^\mathbb{Q} = \rho dt$. If we deflate our contingent claim with $S^{(1)}$ rather than the money account we immediately obtain the pricing equation

$$(5) \quad \chi_t^{\text{exch}} = S_t^{(1)} \mathbb{E}_t^{\mathbb{Q}^{(1)}} [\max\{Z_T - 1, 0\}]$$

where $Z_t \equiv S_t^{(2)}/S_t^{(1)}$. To compute this we must know the dynamics of Z_t under $\mathbb{Q}^{(1)}$. By Itô's formula, the \mathbb{Q} dynamics is

$$(6) \quad \begin{aligned} dZ_t &= \frac{\partial Z_t}{\partial t} dt + \frac{\partial Z_t}{\partial S^{(1)}} dS^{(1)} + \frac{\partial Z_t}{\partial S^{(2)}} dS^{(2)} \\ &+ \frac{1}{2} \left(\frac{\partial^2 Z_t}{\partial (S^{(1)})^2} (dS^{(1)})^2 + 2 \frac{\partial^2 Z_t}{\partial S^{(1)} \partial S^{(2)}} dS^{(1)} dS^{(2)} + \frac{\partial^2 Z_t}{\partial (S^{(2)})^2} (dS^{(2)})^2 \right) \\ &= (\text{non-zero drift}) + Z_t \sigma_2 dW_{2,t}^\mathbb{Q} - Z_t \sigma_1 dW_{1,t}^\mathbb{Q} \\ &= (\text{non-zero drift}) + Z_t (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^{1/2} dW_{1,t}^\mathbb{Q}. \end{aligned}$$

where, in the last line, we have used the standard result that $\text{Var}[X - Y] = \text{Var}[X] + \text{Var}[Y] - 2\text{Cov}[X, Y]$. Now, from Girsanow's Theorem, it follows that

$$(7) \quad dZ_t = Z_t \tilde{\sigma} dW_{1,t}^{\mathbb{Q}^{(1)}}$$

where we have defined $\tilde{\sigma} = (\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^{1/2}$. By Itô the solution to this equation is

$$(8) \quad \ln Z_T = \ln Z_t - \frac{1}{2}\tilde{\sigma}^2(T-t) + \int_t^T \tilde{\sigma} dW_u^{\mathbb{Q}^{(i)}}$$

which is normally distributed with mean $\ln Z_t - \frac{1}{2}\tilde{\sigma}^2(T-t)$ and variance $\tilde{\sigma}^2(T-t)$. Hence,

$$(9) \quad \begin{aligned} \chi_t^{\text{exch}} &= S_t^{(1)} \mathbb{E}_t^{\mathbb{Q}^{(1)}} [\max\{Z_T - 1, 0\}] \\ &= S_t^{(1)} \mathbb{E}_t^{\mathbb{Q}^{(1)}} [\max\{e^{\ln Z_T} - 1, 0\}] \\ &= S_t^{(1)} (\mathbb{E}_t^{\mathbb{Q}^{(1)}} [Z_T] \Phi(d_1) - \Phi(d_2)) \\ &= S_t^{(1)} (Z_t \Phi(d_1) - \Phi(d_2)) \\ &= S_t^{(2)} \Phi(d_1) - S_t^{(1)} \Phi(d_2) \end{aligned}$$

where we have defined

$$(10) \quad \begin{aligned} d_1 &\equiv \frac{1}{\tilde{\sigma}\sqrt{T-t}} \ln \left[\frac{S_t^{(2)}}{S_t^{(1)}} \right] + \frac{1}{2}\tilde{\sigma}\sqrt{T-t}, \\ d_2 &\equiv d_1 - \tilde{\sigma}\sqrt{T-t}. \end{aligned}$$

The third line in (9) makes use of the standard result that if $Y \sim N(m, s^2)$ then $\mathbb{E}[\max\{e^Y - K, 0\}] = \mathbb{E}[e^Y] \Phi(d_1) - K \Phi(d_2)$, where $\mathbb{E}[e^Y] = e^{m+0.5s^2}$, $d_1 = \ln(\mathbb{E}[e^Y]/K)/s + 0.5s$ and $d_2 = d_1 - s$. This completes the derivation of the exchange option price.

2.1.2. Call on Zero Coupon Bonds. Suppose we have a call option with strike K and maturity T on a zero coupon bond which matures (pays \$1) at $S > T$. We imagine that the bond yield is a function of a Merton governed short rate $dr_t = \alpha dt + \beta dW_t^{\mathbb{Q}}$. In particular, we shall assume the exponential affine form $P(t, T) = \exp(-A(T-t) - B(T-t)r_t)$, where $A(T-t)$ and $B(T-t)$ are functions of the time to maturity, $\tau = T-t$, which must be chosen such that the PDE for Merton governed bonds is satisfied¹. I'll leave it as an exercise for the reader to show that

$$(11) \quad B(T-t) = T-t, \quad A(T-t) = \frac{1}{2}\alpha(T-t)^2 - \frac{1}{6}\beta^2(T-t)^3.$$

Clearly, when the call option expires the pay-off will be $\chi_T^{\text{zcb}} = \max\{P(T, S) - K, 0\}$. Now, if we choose to price our claim under \mathbb{Q} then inevitably we will have to face a stochastic r_t both in the discount factor and the ZCB itself, which would give rise to a grueling covariance of factors. Rather, we might get the idea of deflating our option using a ZCB of maturity T

$$(12) \quad \chi_t^{\text{zcb}} = P(t, T) \mathbb{E}_t^{\mathbb{Q}^T} [\max\{F_T(T, S) - K, 0\}]$$

where we have defined the "forward price" $F_t(T, S) \equiv P(t, S)/P(t, T) = \exp(-[A(t, S) - A(t, T)] - [B(t, S) - B(t, T)]r_t)$. (Notice that the strike K remains unchanged: that's because $P(T, T) = 1$, i.e. we are basically dividing by 1 under the expectation operator). Using Itô's Lemma on the latter we get

$$(13) \quad \begin{aligned} dF_t(T, S) &= (\text{non-zero drift}) - F_t(T, S)[B(t, S) - B(t, T)]\beta dW_t^{\mathbb{Q}} \\ &= (\text{non-zero drift}) - F_t(T, S)[S - T]\beta dW_t^{\mathbb{Q}} \end{aligned}$$

¹I.e. $P_t + \alpha P_r + 0.5\beta^2 P_{rr} = rP$ with terminal condition $P(T, T) = 1$.

so by Girsanow, under \mathbb{Q}^T ,

$$(14) \quad dF_t(T, S) = -F_t(T, S)[S - T]\beta dW_t^{\mathbb{Q}^T}.$$

This looks like something we've seen before. Indeed it is now trivial to show that $\ln(F_t(T, S))$ is normally distributed with variance $\nu^2(t, T, S) \equiv \beta^2(S - T)^2(T - t)$ and mean $\ln(F_t(T, S)) - \frac{1}{2}\nu(t, T, S)^2$. So using a derivation analogous to (9) we get

$$(15) \quad \begin{aligned} \chi_t^{\text{exch}} &= P(t, T)\mathbb{E}_t^{\mathbb{Q}^T}[\max\{F_T(T, S) - 1, 0\}] \\ &= P(t, S)\Phi(d_1) - KP(t, T)\Phi(d_2) \end{aligned}$$

where, in this case, we have defined

$$(16) \quad \begin{aligned} d_1 &\equiv \frac{1}{\nu(t, T, S)} \ln \left[\frac{P(t, S)}{P(t, T)} \right] + \frac{1}{2}\nu(t, T, S), \\ d_2 &\equiv d_1 - \nu(t, T, S). \end{aligned}$$

2.2. Preliminary Conclusion. From the above, we conclude that the pricing of financial derivatives can be simplified considerably by adopting a clever choice of numeraire. In particular, it is important to realize that this is not merely an obscure form a mathematical trickery with limited applicability besides the examples already highlighted. Other derivatives where a change of numeraire can be employed include convertible bonds and options where the strike is in a currency different from the stock price (Björk et al. 2001) and dimensional reduction in discretely sampled Asian and lookback options (Andreasen 1998).

3. A CLOSER LOOK

3.1. The Radon-Nikodym Derivative. Of course, we haven't done anything so far to emphasize what exactly permits us to jump between different numeraires (or rather, probability measures) when we price our derivatives. In particular, assuming the correctness of (1) what justifies the transition to (3)? The answer is, of course, a clever mapping of probabilities, but it is not immediately obvious how this should be carried out formally.

First of all we shall require that our probability measures on the space (Ω, \mathcal{F}) (let's label them \mathbb{Q} and $\mathbb{Q}^{(i)}$) between which we transform are equivalent. In other words, our measures agree on which events in our sigma algebra have probability zero, or identically which elements have probability one. If this is the case, then a measure theoretic result known as the *Radon-Nikodym Theorem* guarantees the existence of some non-negative function $L : \Omega \mapsto \mathbb{R}^+$ such that

$$(17) \quad \int_{\mathcal{A}} d\mathbb{Q}^{(i)}(\omega) = \int_{\mathcal{A}} L(\omega) d\mathbb{Q}(\omega)$$

for any event $\mathcal{A} \in \mathcal{F}$. Here the $d\mathbb{Q}$ s should be interpreted as the associated probability density functions under those measures. In particular, (17) basically tells us that there is a function which, when multiplying one of the pdfs, can get our two measure to agree in terms of probabilities assigned to any event. Notice that this can be identically expressed infinitesimally, as the existence of a non-negative function L such that

$$(18) \quad L(\omega) = \frac{d\mathbb{Q}^{(i)}}{d\mathbb{Q}}$$

which explains the nomenclature the "Radon-Nikodym derivative".

3.2. Properties.

3.2.1. *Positivity.* As stated above the Radon-Nikodym derivative, L , is non-negative.

3.2.2. *Unit Mean.* Letting \mathcal{A} be the universal set, Ω , and remembering the properties of density functions we get $1 = \int_{\Omega} L(\omega) d\mathbb{Q}(\omega)$. I.e, by definition of expectation values,

$$(19) \quad \mathbb{E}^{\mathbb{Q}}[L] = 1.$$

Furthermore, if $X : \Omega \mapsto \mathbb{R}$ is a random variable on $(\Omega, \mathcal{F}, \mathbb{Q})$, then it follows from (17) that

$$(20) \quad \mathbb{E}^{\mathbb{Q}^{(i)}}[X] = \mathbb{E}^{\mathbb{Q}}[L \cdot X].$$

3.2.3. *Martingale.* Let \mathcal{H} be a sub-sigma algebra of \mathcal{F} , i.e. $\mathcal{H} \subseteq \mathcal{F}$. Then we have two Radon-Nikodym derivatives $L_{\mathcal{H}}$ and $L_{\mathcal{F}}$, which generally will be unequal as $L_{\mathcal{F}}$ will not be \mathcal{H} -measurable. However, it is easy to show that they will be related through the transformation

$$(21) \quad L_{\mathcal{H}} = \mathbb{E}^{\mathbb{Q}}[L_{\mathcal{F}} | \mathcal{H}].$$

In particular, consider the filtered space $(\Omega, \mathcal{F}, \mathbb{Q}, \{\mathcal{F}_t\}_{t \geq 0})$. Then an immediate corollary of (21) is that $L = d\mathbb{Q}^{(i)}/d\mathbb{Q}$ is a \mathbb{Q} martingale, i.e.

$$(22) \quad L_s = \mathbb{E}^{\mathbb{Q}}[L_t], \quad t \geq s.$$

3.3. Transforming the Pricing Equation. Armed with these properties, let us return to the question of how we transform the standard pricing equation (1) into something along the lines of (3). Basically, we need to find an $L = d\mathbb{Q}^{(i)}/d\mathbb{Q}$ through which we can *define* the density function $d\mathbb{Q}^{(i)}$ such that equation (3) holds, i.e. such that the $A^{(i)}$ deflated asset is indeed a $\mathbb{Q}^{(i)}$ martingale.

To get a feeling for what the relevant L is, it might help to set the conditional expectation in our two equations equal to time zero ($\mathcal{F}_0 = \emptyset$). Obviously, we require that the contingent claims priced under different measures must have equal values, hence equating χ_0 under \mathbb{Q} to χ_0 under $\mathbb{Q}^{(i)}$, we obtain $B_0^{(0)} \mathbb{E}^{\mathbb{Q}}[\chi_T / B_T^{(0)}] = A_0^{(i)} \mathbb{E}^{\mathbb{Q}^{(i)}}[\chi_T / A_T^{(i)}]$, or identically,

$$(23) \quad \mathbb{E}^{\mathbb{Q}} \left[\frac{B_0^{(0)}}{A_0^{(i)}} \frac{A_T^{(i)}}{B_T^{(0)}} \chi_T \right] = \mathbb{E}^{\mathbb{Q}^{(i)}}[\chi_T].$$

Comparing this expression with (20) this prompts us to postulate that by defining the measure $\mathbb{Q}^{(i)}$ under the Radon-Nikodym derivative

$$(24) \quad L_t \equiv \frac{d\mathbb{Q}^{(i)}}{d\mathbb{Q}}(t) = \frac{B_0^{(0)}}{A_0^{(i)}} \frac{A_t^{(i)}}{B_t^{(0)}}$$

then the deflated process $\chi_t / A_t^{(i)}$ will be a martingale under $\mathbb{Q}^{(i)}$. To see this, first notice that the definition of L_t makes sense as Radon-Nikodym derivative insofar as it satisfies the properties of unit unconditional mean, $\mathbb{E}^{\mathbb{Q}}[L_t] = 1$ and being a \mathbb{Q} martingale $L_s = \mathbb{E}^{\mathbb{Q}}[L_t]$ where $t \geq s$. Assuming the validity of (1) we can then demonstrate equation (3), with $\mathbb{Q}^{(i)}$ defined as above, through a simple application of the Abstract Bayes' Theorem²:

²Assume X is a random variable on $(\Omega, \mathcal{F}, \mathbb{Q})$, and let $\mathbb{Q}^{(i)}$ be another measure on (Ω, \mathcal{F}) , where $L = d\mathbb{Q}^{(i)}/d\mathbb{Q}$ is the Radon-Nikodym derivative. Then, if $\mathcal{H} \subseteq \mathcal{F}$ the Abstract Bayes' Theorem holds (a.s. under $\mathbb{Q}^{(i)}$):

$$(25) \quad \mathbb{E}^{\mathbb{Q}^{(i)}}[X | \mathcal{H}] = \frac{\mathbb{E}^{\mathbb{Q}}[L \cdot X | \mathcal{H}]}{\mathbb{E}^{\mathbb{Q}}[L | \mathcal{H}]}.$$

$$\begin{aligned}
(26) \quad \mathbb{E}_t^{\mathbb{Q}^{(i)}} \left[\frac{\chi_T}{A_T^{(i)}} \right] &= \frac{\mathbb{E}_t^{\mathbb{Q}} \left[\frac{d\mathbb{Q}^{(i)}(T)}{d\mathbb{Q}} \cdot \frac{\chi_T}{A_T^{(i)}} \right]}{\mathbb{E}_t^{\mathbb{Q}} \left[\frac{d\mathbb{Q}^{(i)}(T)}{d\mathbb{Q}} \right]} = \frac{\mathbb{E}_t^{\mathbb{Q}} \left[\frac{B_0^{(0)}}{A_0^{(i)}} \frac{A_T^{(i)}}{B_T^{(0)}} \frac{\chi_T}{A_T^{(i)}} \right]}{\frac{d\mathbb{Q}^{(i)}(t)}{d\mathbb{Q}}} \\
&= \frac{\frac{B_0^{(0)}}{A_0^{(i)}} \mathbb{E}_t^{\mathbb{Q}} \left[\frac{\chi_T}{B_T^{(0)}} \right]}{\frac{B_0^{(0)}}{A_0^{(i)}} \frac{A_t^{(i)}}{B_t^{(0)}}} = \frac{\frac{\chi_t}{B_t^{(0)}}}{\frac{A_t^{(i)}}{B_t^{(0)}}} = \frac{\chi_t}{A_t^{(i)}}
\end{aligned}$$

This established the generalized pricing equation.

3.4. The Girsanov Theorem. Now that we know how we formally transform between probability measures when we price under different numeraires, the only important question that remains (for our purposes) is how our Wiener process transforms in the process. As already suggested, the answer is contained in Girsanov's Theorem, which we rudimentarily described as a linear relationship between the Wiener increments. Qua the Radon-Nikodym derivative we can now make this statement considerably more exact. In particular, let $W^{\mathbb{Q}}$ be a Wiener process on the filtered space $(\Omega, \mathcal{F}, \mathbb{Q}, \{\mathcal{F}_t\}_{t \geq 0})$ and suppose we have defined a new measure through $L_t = d\mathbb{Q}^{(i)}/d\mathbb{Q}(t)$, where L_t conforms with the properties listed above. Then, from the martingale condition under \mathbb{Q} , it follows that dL_t can be written as

$$(27) \quad dL_t = \varphi_t L_t dW_t^{\mathbb{Q}}.$$

where φ is some function (in practice computed by applying Itô to our specified L_t). Girsanov's Theorem then states that the Wiener process transforms as

$$(28) \quad dW_t^{\mathbb{Q}} = \varphi_t dt + dW_t^{\mathbb{Q}^{(i)}}.$$

However, as already suggested, for computational reasons it might be less relevant to know the exact form of φ when pricing contingent claims, so this result is merely stated for completeness.

3.5. Conclusion. The importance of numeraire changes within derivative pricing is inexorable. The most outstanding example of this would be the "risk forward measure", \mathbb{Q}^T , used when pricing bond options, but we are by no means restricted to this alone. This paper has been but a cursory exploration of the vast richness the subject contains, with a pseudo-rigorous overview of the underlying probability theory. Hopefully, it has been enough to keep the reader entertained, and will serve as a springboard into analyzing the subject in greater detail.