FinKont2: Hand-In 1

Gustav Dyhr / msh606

February 2025

I have created a full GitHub of my code used for this Hand-In at `https://github.com/guxel/HandIn-1`. The main code is in HandIn1.Py, while the rest is supporting objects. Additionally, short code snippets have been inserted where relevant.

The Hand-In references Tomas Björk's book *Arbitrage Theory in Continuous Time* (4th edition), simply refereed to as Björk.

1 Model Moments

This part will focus on determining the distributions of the two processes typically used for interest-rate modelling.

1.1 Vasicek Distribution

We have been given the following Vasicek SDE:

$$dX_t = \kappa(\theta - X_t)dt + \sigma dW_t \tag{1.1}$$

We must show that the process has the distribution:

$$X_{t+u}|\,\mathcal{F}_t \sim N\left(X_t e^{-\kappa u} + \theta\left(1 - e^{-\kappa u}\right), \frac{\sigma^2\left(1 - e^{-2\kappa u}\right)}{2\kappa}\right). \tag{1.2}$$

To determine the distribution, we can use Björk's Lemma 4.18. The Lemma requires a process defined solely through a brownian motion and a deterministic σ. Therefore, we first create a new related process with the formula:

$$Z_t = f(X_t, t) = e^{\kappa t} X_t \tag{1.3}$$

Using Itô's Lemma (Björk's Theorem 4.11), we can determine the dynamics of the process:

$$dZ_t = df(X_t, t) = \left(f_t + f_x \mu_t + \frac{1}{2}\sigma_t^2 f_{xx}\right)dt + \sigma_t f_x dW_t \tag{1.4}$$

With the following differentials:

$$f_t = \kappa e^{\kappa t} X_t = \kappa Z_t$$
$$f_x = e^{\kappa t} = \frac{Z_t}{X_t}$$
$$f_{xx} = 0$$

And the Vasicek definitions:

$$\mu_t = \kappa\left(\theta - X_t\right)$$
$$\sigma_t = \sigma$$

Inserting these formulas yields:

$$dZ_t = \left(\kappa Z_t + \frac{Z_t}{X_t}\kappa(\theta - X_t) + \frac{1}{2}\sigma^2 \cdot 0\right) dt + \sigma \frac{Z_t}{X_t} dW_t$$

Simplifying the process yields:

$$dZ_t = \left(\kappa Z_t + \frac{Z_t}{X_t}\kappa\theta - \kappa Z_t\right) dt + \sigma e^{\kappa t} dW_t$$

$$dZ_t = e^{\kappa t}(\kappa\theta dt + \sigma) dW_t$$

We now have a simple stochastic process with deterministic drifts and volatility. Looking only at the stochastic parts (imagine subtracting a deterministic part with the drift) then we have the following process:

$$dY_t = dZ_t - e^{\kappa t}\kappa\theta dt = e^{\kappa t}\sigma dW_t$$

Using Björk's Lemma 4.18 we can now determine the distribution of this process. It says that the process defined as a deterministic function multiplied with a Standard Brownian Motion (SBM) satisfies the following:

$$\mathrm{E}_t[Y_{t+s}] = Y_t + 0 = Y_t$$

$$\mathrm{Var}_t[Y_{t+u}] = \int_t^{t+u} \sigma_s^2 ds = \int_t^{t+u} (e^{\kappa s}\sigma)^2 ds = \int_t^{t+u} e^{2\kappa s}\sigma^2 ds = e^{2\kappa t}\frac{\sigma^2(e^{2\kappa u} - 1)}{2\kappa}$$

$$Y_{t+u}|\mathcal{F}_t \sim N\left(Y_t, e^{2\kappa t}\frac{\sigma^2(e^{2\kappa u} - 1)}{2\kappa}\right)$$

As Z_t is Y_t plus something deterministic, then it is clear that:

$$Z_{t+u} = Y_{t+u} + \int_t^{t+u} e^{\kappa s}\kappa\theta ds = Y_{t+u} + e^{kt}\theta(e^{\kappa u} - 1)$$

$$Z_{t+u}|\mathcal{F}_t \sim N\left(Z_t + e^{kt}\theta(e^{\kappa u} - 1), e^{2\kappa t}\frac{\sigma^2(e^{2\kappa u} - 1)}{2\kappa}\right)$$

From equation (1.3) we see that:

$$X_t = e^{-\kappa t} Z_t$$

Which yields using basic stochastic algebra:

$$\begin{aligned}
\mathrm{E}_t[X_{t+u}] =& \mathrm{E}_t[e^{-\kappa(t+u)} Z_{t+u}] = e^{-\kappa(t+u)}\mathrm{E}_t[Z_{t+u}] \\
=& e^{-\kappa(t+u)} Z_t + e^{-\kappa(t+u)} e^{kt}\theta(e^{\kappa u} - 1) \\
=& e^{-\kappa(t+u)} e^{kt} X_t + e^{-\kappa u}\theta(e^{\kappa u} - 1) = X_t e^{-\kappa u} + \theta(1 - e^{-\kappa u})
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}_t[X_{t+u}] =& \mathrm{Var}_t[e^{-\kappa(t+u)} Z_{t+u}] = e^{-2\kappa(t+u)}\mathrm{Var}_t[Z_{t+u}] \\
=& e^{-2\kappa(t+u)} e^{2\kappa t}\frac{\sigma^2(e^{2\kappa u} - 1)}{2\kappa}) = e^{-2\kappa u}\frac{\sigma^2(e^{2\kappa u} - 1)}{2\kappa} = \frac{\sigma^2(1 - e^{-2\kappa u})}{2\kappa}
\end{aligned}$$

Now we have the final distribution for $X_{t+u}|\mathcal{F}_t$:

$$X_{t+u}|\mathcal{F}_t \sim N\left(X_t e^{-\kappa u} + \theta(1 - e^{-\kappa u}), \frac{\sigma^2(1 - e^{-2\kappa u})}{2\kappa}\right)$$

This matches the distribution defined in (1.2). We can also determine the long term distribution of the process:

$$\lim_{u\to\infty} X_{t+u}|\mathcal{F}_t \sim N\left(\theta, \frac{\sigma^2}{2\kappa}\right)$$

Which highlights the mean reversion aspect of the model.

1.2 CIR Expectation

We have been given the following CIR SDE:

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 = x \tag{1.5}$$

With $m_t = \mathrm{E}[X_t]$, we must show that

$$m_t = X_0 e^{-\kappa t} + \theta\left(1 - e^{-\kappa t}\right) \tag{1.6}$$

We start by writing the SDE on integral form. To accomplish this, we define a new process identical to in the last question.

$$Z_t = f(X_t, t) = e^{\kappa t} X_t \tag{1.7}$$

The dynamics can then by calculated in the same way using (1.4): With the same differentials and drift:

$$f_t = \kappa e^{\kappa t} X_t = \kappa Z_t$$

$$f_x = e^{\kappa t} = \frac{Z_t}{X_t}$$

$$f_{xx} = 0$$

$$\mu_t = \kappa(\theta - X(t))$$

But the CIR diffusion term is more complicated:

$$\sigma_t = \sigma\sqrt{X_t}$$

Inserting these definitions yields:

$$dZ_t = (\kappa Z_t + \frac{Z_t}{X_t}\kappa(\theta - X_t) + \frac{1}{2}(\sigma\sqrt{X_t})^2 \cdot 0)dt + \sigma\sqrt{X_t}\frac{Z_t}{X_t}dW_t$$

$$dZ_t = e^{\kappa t}\kappa\theta dt + e^{\kappa t}\sigma\sqrt{X_t}dW_t \tag{1.8}$$

We now have a simple process that we can write on integral form:

$$Z_t = Z_0 + \int_0^t e^{\kappa s}\kappa\theta ds + \int_0^t e^{\kappa s}\sigma\sqrt{X_s}dW_s = Z_0 + \theta(e^{\kappa t} - 1) + \sigma\int_0^t e^{\kappa s}\sqrt{X_s}dW_s \tag{1.9}$$

This is not a perfect closed form solution (no such solution as available for the CIR model), but it is sufficient for our task of analysing the expectations. The expected value of the process is:

$$\mathrm{E}[Z_t] = \mathrm{E}[Z_0 + \theta(e^{\kappa t} - 1) + \sigma\int_0^t e^{\kappa s}\sqrt{X_s}dW_s] = Z_0 + \theta(e^{\kappa t} - 1) + 0 = Z_0 + \theta(e^{\kappa t} - 1)$$

Where we used the linearity of expectations. Tthe first term is deterministic, and the expectation of an Itô integral is 0 under some reasonable constraints (Björk's proposition 4.5). We can now compute the expectation of X_t based on this:

$$\begin{aligned} m_t = \mathrm{E}[X_t] =& \mathrm{E}[e^{-\kappa t}Z_t] = e^{-\kappa t}\mathrm{E}[Z_t] \\ =& e^{-\kappa t}Z_0 + e^{-\kappa t}\theta(e^{\kappa t} - 1) = e^{-\kappa t}e^{-\kappa 0}X_0 + \theta(1 - e^{-\kappa t}) = X_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) \end{aligned} \tag{1.10}$$

Yielding exactly the expression for m_t that we wished to derive, thereby completing the question. Interestingly, the expectation is identical to the expectation of the Vasicek model, which is logical considering their identical drifts. The long-term expectation is also the same, defined as the mean-reversion level θ.

We accidentally ended up not following the hint, but we arrived at the correct conclusion, using the hint would yield the following formulas: The integral form of the CIR model:

$$X_t = X_0 + \int_0^t \kappa(\theta - X_s)ds + \int_0^t \sigma\sqrt{X_s}dW_s \tag{1.11}$$

Take mean:

$$\mathrm{E}[X_t] = E\left[X_0 + \int_0^t \kappa(\theta - X_s)ds + \int_0^t \sigma\sqrt{X_s}dW_s\right] = X_0 + \int_0^t \kappa(\theta - \mathrm{E}[X_s])ds + 0 \tag{1.12}$$

Using the linearity of expectations and the idea that "integrals are simply sums" to move our expectations inside the integrals. Again we use the fact that the expectation of stochastic integrals is 0. Putting $m_t = \mathrm{E}[X_t]$ and solving the integral:

$$m_t = m_0 + \int_0^t \kappa(\theta - m_s)ds = m_0 + \kappa\int_0^t \theta ds - \kappa\int_0^t m_s ds = m_0 + \kappa\theta t - \kappa\int_0^t m_s ds \tag{1.13}$$

We can now differentiate with regards to t to gain our ODE:

$$\frac{d}{dt}m_t = \kappa\theta - \kappa m_t. \tag{1.14}$$

Where we used the following rule for the derivative of a bounded integral:

$$\frac{d}{dx}\int_a^x f(b)db = f(x) \tag{1.15}$$

Solving the ODE:

$$m_t = \theta + ce^{-\kappa t} = \theta + (m_0 - \theta)e^{-\kappa t} = X_0e^{-\kappa t} + \theta(1 - e^{-\kappa t}) \tag{1.16}$$

Where we used the boundary condition $m_0 = X_0$. We have hereby completed the task the "correct" way.

1.3 CIR Variance

First we must derive an ODE for the second moment of the CIR model.

$$h_t = \mathrm{E}[X_t^2] \tag{1.17}$$

Next we must prove that:

$$\mathrm{Var}[X_t] = X_0\frac{\sigma^2}{\kappa}\left(e^{-\kappa t} - e^{-2\kappa t}\right) + \theta\frac{\sigma^2}{2\kappa}\left(1 - e^{-\kappa t}\right)^2 \tag{1.18}$$

We start by following the hint and defining a process:

$$S_t = f(X_t, t) = X_t^2 \tag{1.19}$$

Using Itô's Lemma (1.4), we find the dynamics:

$$f_t = 0$$
$$f_x = 2X_t$$
$$f_{xx} = 2$$
$$\mu_t = \kappa(\theta - X_t)$$
$$\sigma_t = \sigma\sqrt{X_t}$$

$$
\begin{aligned}
dS_t = dX_t^2 =&(0 + 2X_t\kappa(\theta - X_t + \frac{1}{2}(\sigma\sqrt{X_t})^2 2)dt + \sigma\sqrt{X_t}2X_t dW_t \\
=&(2X_t\kappa(\theta - X_t) + \sigma^2 X_t)dt + 2\sigma X_t^{3/2} dW_t \\
=&X_t((2\kappa(\theta - X_t) + \sigma^2)dt + 2\sigma\sqrt{X_t}dW_t)
\end{aligned}
\tag{1.20}
$$

We now determine have the ODE for the second moment:

$$
\begin{aligned}
S_t =&S_0 + \int_0^t 2\kappa(\theta X_s - X_s^2) + \sigma^2 X_s ds + \int_0^t 2\sigma X_s^{3/2} dW_s) \Leftrightarrow \\
h(t) =&\mathrm{E}[X_t^2] = \mathrm{E}[S_t^2] = h_0 + \int_0^t 2\kappa(\theta\mathrm{E}[X_s] - \mathrm{E}[X_s^2]) + \sigma^2\mathrm{E}[X_s]ds + 0 = h_0 + \int_0^t 2\kappa(\theta m_s - h_s) + \sigma^2 m_s ds \Leftrightarrow \\
\frac{d}{dt}h_t =&2\kappa(\theta m_t - h_t) + \sigma^2 m_t \Leftrightarrow \\
h_t =&e^{-2\kappa t}(e^{\kappa t}(\theta + \frac{1}{2\kappa}\sigma^2)(\theta(e^{\kappa t} - 2) + 2X_0) + c), h_0 = X_0^2 \Leftrightarrow \\
h_t =&e^{-2\kappa t}(e^{\kappa t}\left(\theta + \frac{1}{2\kappa}\sigma^2)(\theta(e^{\kappa t} - 2) + 2X_0) + X_0^2 - (2X_0 - \theta)\left(\theta + \frac{1}{2\kappa}\sigma\right)\right) \\
=&e^{-2\kappa t}(X_0^2 + (\theta + \frac{1}{2\kappa}\sigma^2)(e^{\kappa t} - 1)(2X_0 + \theta(e^{\kappa t} - 1)))
\end{aligned}
\tag{1.21}
$$

We have hereby derived the ODE for the second moment and have solved it. We will also attempt with the original method below, creating new Itô process' that are simpler to work with and then substituting. The result is naturally the same and there is no real point to deriving this twice except to avoid the pain of deleting pretty formulas. The next part starts at equation (1.27).

We start by simplify the drift of the process. One idea for doing this is through the same multiplication as before, but also squared ($e^{2\kappa t}$).

$$
D_t = e^{2\kappa t}S_t \tag{1.22}
$$

Solving for the dynamics:

$$
f_t = 2\kappa e^{2\kappa t}S_t = 2\kappa D_t
$$

$$
f_s = e^{2\kappa t}
$$

$$
f_{ss} = 0
$$

$$
\mu_t = X_t((2\kappa(\theta - X_t) + \sigma^2))
$$

$$
\sigma_t = 2\sigma X_t\sqrt{X_t}
$$

$$
\begin{aligned}
dD_t = d(e^{2\kappa t}S_t) =&(2\kappa D_t + e^{2\kappa t}X_t(2\kappa(\theta - X_t) + \sigma^2) + \frac{1}{2}(2\sigma X_t\sqrt{X_t})^2 0)dt + 2\sigma X_t\sqrt{X_t}e^{2\kappa t}dW_t \\
=&e^{2\kappa t}X_t(2\kappa\theta + \sigma^2)dt + e^{2\kappa t}2\sigma X_t^{3/2}dW_t
\end{aligned}
\tag{1.23}
$$

We still have one X_t term in our drift, but as we have a formula for the expectation of X_t (m_t), then that is not a problem. We write our new process on integral form:

$$
\begin{aligned}
D_t =&D_0 + \int_0^t e^{2\kappa s}X_s(2\kappa\theta + \sigma^2)ds + \int_0^t e^{2\kappa s}2\sigma X_s^{3/2}dW_s \\
=&D_0 + (2\kappa\theta + \sigma^2)\int_0^t e^{2\kappa s}X_s ds + 2\sigma\int_0^t e^{2\kappa s}X_s^{3/2}dW_s
\end{aligned}
\tag{1.24}
$$

We can now compute the expectation of this process:

$$\begin{aligned}
\mathrm{E}[D_t] =& \mathrm{E}\left[D_0 + (2\kappa\theta+\sigma^2)\int_0^t e^{2\kappa s}X_s ds + 2\sigma\int_0^t e^{2\kappa s}X_s^{3/2}dW_s\right] \\
=& \mathrm{E}[D_0] + (2\kappa\theta+\sigma^2)\mathrm{E}\left[\int_0^t e^{2\kappa s}X_s ds\right] + 2\sigma\mathrm{E}\left[\int_0^t e^{2\kappa s}X_s^{3/2}dW_s\right] \\
=& D_0 + (2\kappa\theta+\sigma^2)\int_0^t e^{2\kappa s}\mathrm{E}\left[X_s\right]ds + 0 \\
=& D_0 + (2\kappa\theta+\sigma^2)\int_0^t e^{2\kappa s}(X_0e^{-\kappa s}+\theta(1-e^{-\kappa s}))ds \\
=& D_0 + (2\kappa\theta+\sigma^2)\int_0^t e^{\kappa s}(X_0+\theta(e^{\kappa s}-1))ds \\
=& e^{2\kappa 0}X_0^2 + (2\kappa\theta+\sigma^2)\frac{(e^{\kappa t}-1)(2X_0+\theta(e^{\kappa t}-1))}{2\kappa} \\
=& X_0^2 + (\theta+\frac{1}{2\kappa}\sigma^2)(e^{\kappa t}-1)(2X_0+\theta(e^{\kappa t}-1))
\end{aligned} \tag{1.25}$$

We now have a term for $\mathrm{E}[D_t] = \mathrm{E}[e^{2\kappa t}X_t^2]$, so deriving the term for $h_t = \mathrm{E}[X_t^2]$ is simple.

$$\begin{aligned}
h_t = \mathrm{E}[X_t^2] =& \mathrm{E}[e^{-2\kappa t}D_t] = e^{-2\kappa t}\mathrm{E}[D_t] \\
=& e^{-2\kappa t}(X_0^2 + (\theta+\frac{1}{2\kappa}\sigma^2)(e^{\kappa t}-1)(2X_0+\theta(e^{\kappa t}-1)))
\end{aligned} \tag{1.26}$$

This is as expected equivalent to the solution we found in equation (1.21).

For the final part of the question, we can simply use one of the basic formulas for variance and our 2 previous results:

$$\begin{aligned}
\mathrm{Var}[X_t] =& \mathrm{E}[X_t^2] - \mathrm{E}[X_t]^2 = h_t - m_t^2 \\
&= e^{-2\kappa t}(X_0^2 + (\theta+\frac{1}{2\kappa}\sigma^2)(e^{\kappa t}-1)(2X_0+\theta(e^{\kappa t}-1))) - (X_0e^{-\kappa t}+\theta(1-e^{-\kappa t}))^2 \\
&= e^{-2\kappa t}(X_0^2 + (\theta+\frac{1}{2\kappa}\sigma^2)(e^{\kappa t}-1)(2X_0+\theta(e^{\kappa t}-1))) - X_0^2e^{-2\kappa t} - \theta^2(1-e^{-\kappa t})^2 - 2X_0e^{-\kappa t}\theta(1-e^{-\kappa t}) \\
&= e^{-2\kappa t}(\theta+\frac{1}{2\kappa}\sigma^2)(2X_0(e^{\kappa t}-1)+\theta(e^{\kappa t}-1)^2) - \theta^2(1-e^{-\kappa t})^2 - 2X_0\theta(e^{-\kappa t}-e^{-2\kappa t}) \\
&= (\theta+\frac{1}{2\kappa}\sigma^2)(2X_0(e^{-\kappa t}-e^{-2\kappa t}) + \theta e^{-2\kappa t}(e^{2\kappa t}+1-2e^{\kappa t})) - \theta^2(1+e^{-2\kappa t}-2e^{-\kappa t}) - 2X_0\theta(e^{-\kappa t}-e^{-2\kappa t}) \\
&= \theta 2X_0(e^{-\kappa t}-e^{-2\kappa t}) + \theta^2(1+e^{-2\kappa t}-2e^{-\kappa t}) + \frac{1}{2\kappa}\sigma^2 2X_0(e^{-\kappa t}-e^{-2\kappa t}) + \frac{1}{2\kappa}\sigma^2\theta(1+e^{-2\kappa t}-2e^{-\kappa t}) \\
&-\theta^2(1+e^{-2\kappa t}-2e^{-\kappa t}) - 2X_0\theta(e^{-\kappa t}-e^{-2\kappa t}) \\
&= \frac{1}{2\kappa}\sigma^2 2X_0(e^{-\kappa t}-e^{-2\kappa t}) + \frac{1}{2\kappa}\sigma^2\theta(1+e^{-2\kappa t}-2e^{-\kappa t}) \\
&= X_0\frac{\sigma^2}{\kappa}(e^{-\kappa t}-e^{-2\kappa t}) + \theta\frac{\sigma^2}{2\kappa}(1-e^{-\kappa t})^2
\end{aligned} \tag{1.27}$$

Finally yielding exactly the expression for the variance of the process that we were attempting to prove. It is therefore clear that since:

$$\mathrm{Var}[X_t] = \mathrm{E}[X_t^2] - \mathrm{E}[X_t]^2 = h_t - m_t^2$$

Then:

$$\mathrm{Var}[X_t] = X_0\frac{\sigma^2}{\kappa}(e^{-\kappa t}-e^{-2\kappa t}) + \theta\frac{\sigma^2}{2\kappa}(1-e^{-\kappa t})^2$$

We have now determined the mean and variance of the CIR model's distribution. Sadly, we are not able to directly determine the distribution type. As the model has a stochastic volatility term $\sigma\sqrt{X_t}$ then Björk's Lemma 4.18 that we used above does not apply, and the future values of the process are not normally distributed. We also see that the variance is very different from the CIR level, as the volatility of the process is now directly related to the level of the process. The long term variance is however more similar:

$$\lim_{t\to\infty} \mathrm{Var}[X_t] = \theta\frac{\sigma^2}{2\kappa} \tag{1.28}$$

This result is equal to the long term variance of the Vasicek model, multiplied with the square root of our mean-reversion level for the model. Considering that our CIR model is equivalent to the Vasicek model, but with a multiplier of the square root of the process level on the volatility, then this is a very logical result. One must however not underestimate the differences of the two models, for example Vasicek allows for negative values (rates) while CIR does not. The paths and probabilities assigned to different values are therefore extremely different, even though the two first moments are similar.

2 The Bachelier Model

This part will be focused on the Bachelier Model, a model where asset prices have the dynamic:

$$dS_t = \mu_t dt + \sigma dW_t \tag{2.1}$$

So the deviations in asset prices have a constant volatility σ and do not depend on the asset prices themselves, like in the typical Black-Scholes which models asset prices as Geometric Brownian Motions (GMBs).

2.1 Call Option

We are told that the interest rate r is 0. We must show that the arbitrage-free price of an European call-option is:

$$\pi_t^{\text{call, Bach}} = (S_t - K)\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) \tag{2.2}$$

In the Bachelier model. $\Phi(x)$ and $\phi(x)$ denote the standard normal cumulative density distribution and the point density function. Next we must show that the delta of the option is:

$$\Delta_t^{\text{call,Bach}} = \Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) \tag{2.3}$$

We may use without proof that if $X \sim N\left(\mu, \sigma^2\right)$ then:

$$\mathrm{E}\left[X\mathbf{1}_{l\leq X\leq h}\right] = \mu\left(\Phi\left(\frac{h-\mu}{\sigma}\right) - \Phi\left(\frac{l-\mu}{\sigma}\right)\right) + \sigma\left(\phi\left(\frac{l-\mu}{\sigma}\right) - \phi\left(\frac{h-\mu}{\sigma}\right)\right) \tag{2.4}$$

To price the option, we use the risk neutral valuation formula with a constant interest rate r (Björk's proposition 9.3):

$$\pi_t = e^{-r(T-t)}\mathrm{E}_t^Q[F(S_T)] = e^{-0(T-t)}\mathrm{E}_t^Q[F(S_T)] = \mathrm{E}_t^Q[F(S_T)] \tag{2.5}$$

Where $F(S_T)$ is the terminal payoff and S_t has the following dynamics under the Bankbook martingale measure Q:

$$\begin{aligned} dS_t =& (\mu_t - \eta_t\sigma_t)dt + \sigma_t dW_t^Q = (\mu_t - \frac{\mu_t - rS_t}{\sigma}\sigma)dt + \sigma dW_t^Q = rS_t dt + \sigma dW_t^Q = 0S_t dt + \sigma dW_t^Q \\ =& \sigma dW_t^Q \end{aligned} \tag{2.6}$$

The reason why η_t is defined as such, is that it yields a process that is a martingale under Q, when discounted with the bankbook that has drift $dB_t = rB_t dt$. This gives the closed form solution for the conditioned asset price:

$$S_{t+s}|\mathcal{F}_t = S_t + \int_t^{t+s} \sigma dW_u^Q = S_t + \sigma\sqrt{s}Z \tag{2.7}$$

Where Z is a standard normally distributed variable, we can therefore determine the conditioned distribution of S_{t+s} under Q:

$$S_{t+s}|\mathcal{F}_t \sim N\left(S_t, \sigma^2 s\right) \tag{2.8}$$

We can plug this into the risk neutral valuation formula together with the call-option terminal condition $F(S_T) = (S_T - K)^+$:

$$\pi_t = \mathrm{E}_t^Q[F(S_T)] = \mathrm{E}_t^Q[(S_T - K)^+] = \mathrm{E}_t^Q[(S_T - K)\mathbf{1}_{K<S_T<\infty}] = \mathrm{E}_t^Q[S_T\mathbf{1}_{K<S_T<\infty}] - K\mathrm{E}_t^Q[\mathbf{1}_{K<S_T<\infty}]$$

For the first term we can use (2.4) and for the second term we can use that the expectation of an indicator function is equivalent to a probability function. In addition we use that $S_T < \infty$ almost surely (with probability 1) and that (from attributes of the normal distribution) it is clear that:

$$\Phi(\infty) = 1 \tag{2.9}$$

$$\phi(\infty) = 0 \tag{2.10}$$

$$\Phi(X) = 1 - \Phi(-X) \tag{2.11}$$

$$\phi(X) = \phi(-X) \tag{2.12}$$

$$\Phi(X) = P[Z < X] \tag{2.13}$$

We also need some knowledge about the derivatives later:

$$\frac{d}{dX}\Phi(X) = \phi(X) \tag{2.14}$$

$$\frac{d}{dX}\phi(X) = -X\phi(X) \tag{2.15}$$

Which yields:

$$\begin{aligned}
\pi_t =& S_t \left(\Phi\left(\frac{\infty - S_t}{\sigma\sqrt{T-t}}\right) - \Phi\left(\frac{K - S_t}{\sigma\sqrt{T-t}}\right)\right) + \sigma\sqrt{T-t}\left(\phi\left(\frac{K - S_t}{\sigma\sqrt{T-t}}\right) - \phi\left(\frac{\infty - S_t}{\sigma\sqrt{T-t}}\right)\right) - K\mathrm{P}_t^Q[K < S_T < \infty] \\
=& S_t\left(1 - \Phi\left(\frac{K - S_t}{\sigma\sqrt{T-t}}\right)\right) + \sigma\sqrt{T-t}\left(\phi\left(\frac{K - S_t}{\sigma\sqrt{T-t}}\right) - 0\right) - K\mathrm{P}_t^Q[K < S_t + \sigma\sqrt{T-t}Z] \\
=& S_t\Phi\left(\frac{-(K - S_t)}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{K - S_t}{\sigma\sqrt{T-t}}\right) - K\mathrm{P}_t^Q[Z < -\frac{K - S_t}{\sigma\sqrt{T-t}}] \\
=& S_t\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{K - S_t}{\sigma\sqrt{T-t}}\right) - K\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) \\
=& (S_t - K)\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right)
\end{aligned}$$

We have hereby reached the arbitrage-free pricing formula that we wished to prove. According to the risk neutral valuation theorem, a portfolio of the underlying asset and the risk free asset that perfectly replicates the payoff of the option has the above price. There would theoretically be arbitrage if the price of the option diverged from this, as one would be able to buy/sell the hedge portfolio and sell/buy the option to gain cash with no risk. This is obviously only true following the assumptions of the model, such as no spread, continuous hedging and the simple Bachelier asset price process.

We now have to determine the delta of the option, which is the position that one would have for hedging the option. We do this by simply differentiating the option with regards to the asset price S_t:

$$\begin{aligned}
\Delta_t = \frac{d}{dS_t}\pi_t =& \frac{d}{dS_t}\left((S_t - K)\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right)\right) \\
=& \frac{d}{dS_t}S_t\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) - K\frac{d}{dS_t}\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\frac{d}{dS_t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) \\
=& \Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + S_t\frac{d}{dS_t}\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) - K\frac{d}{dS_t}\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\frac{d}{dS_t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) \\
=& \Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + (S_t - K)\frac{d}{dS_t}\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\frac{d}{dS_t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right)
\end{aligned}$$

Where we have used the product rule. We must now use the chain rule to evaluate the other terms, using (2.14) and (2.15).

$$\frac{d}{dS_t}\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) = \phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right)\frac{d}{dS_t}\frac{S_t - K}{\sigma\sqrt{T-t}} = \frac{1}{\sigma\sqrt{T-t}}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right)$$

$$\frac{d}{dS_t}\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right) = -\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)\frac{d}{dS_t}\frac{S_t-K}{\sigma\sqrt{T-t}} = -\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)\frac{1}{\sigma\sqrt{T-t}}$$
$$= -\frac{S_t-K}{(\sigma\sqrt{T-t})^2}\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)$$

Inserting these differentials in our formula for Δ_t yields the final equation:

$$\begin{aligned}\Delta_t =&\Phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right) + (S_t-K)\frac{1}{\sigma\sqrt{T-t}}\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\left(-\frac{S_t-K}{(\sigma\sqrt{T-t})^2}\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)\right)\\ =&\Phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right) + \frac{S_t-K}{\sigma\sqrt{T-t}}\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right) - \frac{S_t-K}{\sigma\sqrt{T-t}}\phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)\\ =&\Phi\left(\frac{S_t-K}{\sigma\sqrt{T-t}}\right)\end{aligned}$$

We see that the additional terms cancel out, leaving the exact term that we were asked to show.

Both the price formula and the delta formula are reminiscent of the Black-Scholes equivalent formulas, especially the delta. There is however less complexity due to the 0 interest rate and the simple non-geometric volatility of the process.

2.2 Implied Volatility

We assume that $S_0 = 100, T = 0.25, \sigma = 15$ and still that $r = 0$. We have been asked to inspect and comment on the implied volatility of the Bachelier model across different strikes K. Afterwards we must compare with $S_0 = 50$.

We start by implementing Bachelier pricing in python, using the pricing formula we derived. This results in the following script:

```python
class bachelierCall(claim):

    def payoff(self,**v):
        return np.maximum(v['S']-v['K'],0)

    def price(self,**v):
        with stats.work_precision(self.precision), np.errstate(divide='ignore', invalid='ignore'):
              #for T=0
            d = (v['S'] - v['K']) / (v['sigma'] * stats.sqrt(v['T']))
            price = (v['S']-v['K']) * stats.norm_cdf(d) \
                  + (v['sigma'] * stats.sqrt(v['T'])) * stats.norm_pdf(d)
            return  np.where(v['T']==0,self.payoff(**v),price)
```

We can now use this method to price the option on the interval $K \in [0.1; 150]$:

```python
Bachlier = bachelierCall(precision = 1000)

df = pd.DataFrame(index = np.arange(1,1501)/10)
df.index.name = 'Strike'
df['opt price'] = Bachlier.price(S=100,K=df.index.to_numpy(),sigma=15,T=0.25)
```

We have plotted the resulting option prices in figure 1. We see that options far In The Money (ITM) has a high sensitivity to changes in the strike prices, at up to -1. This decreases as the "moneyness" of the option decreases, until the sensitivity approaches 0 for far out-of-the-money (OTM) options. This is a very reasonable result and it is very similar to that of Black-Scholes options.

Our next step now is to determine the implied volatility of these prices. That means that we have to determine the volatilities that yield the same option prices in the Black-Scholes model, holding all other parameters equal. To accomplish this we have to create a Black-Scholes pricing formula and a script that numerically converges on the implied volatility by repeatedly calculating option prices for different guesses.

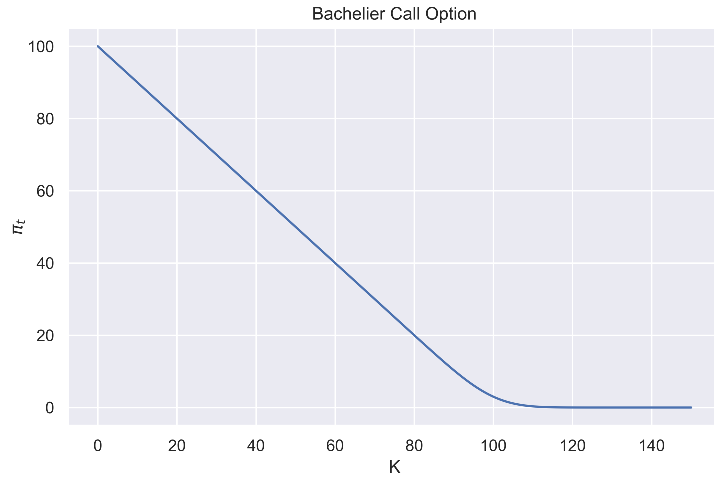

Figure 1: Call Option prices using the Bachlier model and $S = 100, T = 0.25, \sigma = 15$ and $r = 0$

This is however easier said that done, as we will be dealing with both far ITM and far OTM options, which can have very low sensitivity to changes in volatility. To ensure both fast convergence and overall stability, we have created an algorithm that combines both bisection methods and Newton-Raphson. The algorithm is as follows:

Algorithm Implied Volatility Calculation

0. Initialise by guessing that $\sigma^0 = \sqrt{\frac{2\pi}{T-t}} \frac{\pi^{real} - (S_t - K)/2}{S_t - (S_t - K)/2}$ and setting $\sigma_{lb} = 0, \sigma_{ub} = \min(\sigma^0 \cdot 1000, 10000)$
1. Compute option price $\pi^{BS,i}$ using Black-Scholes formula and σ^{i-1}, if close enough to real price then stop.
2. If $\pi^{BS,i} > \pi^{real}$ then set $\sigma_{up} = \sigma^{i-1}$ else set $\sigma_{down} = \sigma^{i-1}$.
3. Use the price gap divided with the sensitivity for a Newton-Raphson volatility guess: $\sigma^{NR} = \sigma^{i-1} + \frac{\pi^{real} - \pi^{BS,i}}{\nu^{BS}}$
4. If the guess is valid such that $\sigma^{NR} \in]1.1 \cdot \sigma_{lb}, 0.9 \cdot \sigma_{ub}[$ (requires well behaved ν), then use it $\sigma^i = \sigma^{NR}$, else use bisection $\sigma^i = (\sigma_{lb} + \sigma_{ub})/2$
5. Return to step 1

The validity check is required far OTM/ITM options, where the solver could diverge due to tiny ν's and machine errors. We will not derive the Black-Scholes formula π^{BS} or the volatility sensitivity $\nu^{BS} = \frac{d\pi^{BS}}{d\sigma}$, but simply use their definitions as they are well known.

$$\pi^{BS} = S_t \Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{2.16}$$

$$\nu^{BS} = S_t \sqrt{T-t}\,\phi(d_1) \tag{2.17}$$

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \tag{2.18}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} \tag{2.19}$$

The initialisation guess comes from *Computing the Black-Scholes implied volatility: Generalization of a simple formula* by *MAJ Bharadia, N Christofides* (1995). It is simple, but gives a fine starting point. We also use a special data type that with 1000 decimal precision. We implement the Black-Scholes functionality in the following code:

```python
class call(claim):

    def payoff(self,**v):
        return np.maximum(v['S']-v['K'],0)

    def price(self,**v):
        with stats.work_precision(self.precision), np.errstate(divide='ignore', invalid='ignore'):
            d1 = (stats.log(v['S'] / v['K']) + (v['r'] + 0.5 * v['sigma'] ** 2) * v['T']) / (v['sigma'] * stats.sqrt(v['T']))
            d2 = d1 - v['sigma'] * stats.sqrt(v['T'])

            price = v['S'] * stats.norm_cdf(d1) - v['K'] * stats.exp(-v['r'] * v['T']) * stats.norm_cdf(d2)
            return  np.where(v['T']==0,self.payoff(**v),price)

    def vega(self,**v):
        with stats.work_precision(self.precision),np.errstate(divide='ignore', invalid='ignore'):
            d1 = (stats.log(v['S'] / v['K']) + (v['r'] + 0.5 * v['sigma'] ** 2) * v['T']) / (v['sigma'] * stats.sqrt(v['T']))
            return np.where(v['T']==0,0,v['S'] * stats.norm_pdf(d1) * stats.sqrt(v['T']))

    def invertVol(self,**v):
         convergence = 1E-20
         sigma = np.sqrt(2*np.pi/v['T'])*(v['price']-(v['S']-v['K'])/2)/(v['S']-(v['S']-v['K'])/2)
         maxIterations = 1E6

         lb = 0
         ub = np.where(sigma * 1000>10000,sigma * 1000,10000)

        for i in range(0,int(maxIterations)):
            pi = self.price(sigma=sigma,S=v['S'],K=v['K'],r=v['r'],T=v['T']) #step 1
            diff = v['price'] - pi

            if np.max(np.abs(diff)) < convergence:
                return sigma

            sigma += diff / self.vega(sigma=sigma,S=v['S'],K=v['K'],r=v['r'],T=v['T']) #step 2

            with np.errstate(divide='ignore', invalid='ignore',over = 'ignore'): #step 3
                newton = sigma + diff/self.vega(sigma=sigma,S=v['S'],K=v['K'],r=v['r'],T=v['T'])

            sigma = np.where((1.1*lb < newton) & (newton < 0.9*ub),newton,(ub+lb)/2) #step 4

        raise Exception("Convergence not reached")
```

Note that this snippet is slightly simplified compared to the code in the GitHub, but it is fully functional. We plot the result in figure 2.

The result is still highly unstable for low K despite our attempts at stabilizing the solution. The sensitivity of π^{BS} to changes in σ is simply too low. The low time till maturity also plays a role in bringing down the sensitivity to the volatility, as the asset has less time to fluctuate. To illustrate the problem, we have plotted the price deviations of 3 different naive guesses for σ in figure 3. We see that the deviations are largest at the At The Money (ATM) level as expected, as this is where ν is largest. The sensitivity of an option price to the volatility comes from the optionality, and the optionality is much greater for options that are closer to the strike. This is where our implied volatility calculation is the most precise. We also see that the three very different guesses all almost perfectly price the option for low (and high) strike prices. There are differences in the prices, but they are tiny. For $K = 0.1$ then $\pi^{BS,\sigma=0.55} - \pi^{BS,\sigma=0.01} = 3\text{e}{-77}$, way below our convergence criteria of $1\text{e}{-15}$ and below levels where we could reliably calculate differences and reach convergence.

Even with high precision, the algorithm finds multiple satisfactory solutions for σ^{imp} that all have $\Phi(d_{1/2}) = 1$ and this gives the graph its high volatility. For such cases the algorithm will keep trying just as good σ values until it converges on a lower or upper bound. Meanwhile for K where the algorithm finds better solutions than $\Phi(d_{1/2}) = 1$ (such as $1 - (1e - 1000)$) then we can determine an unambiguous σ that minimises the price difference.

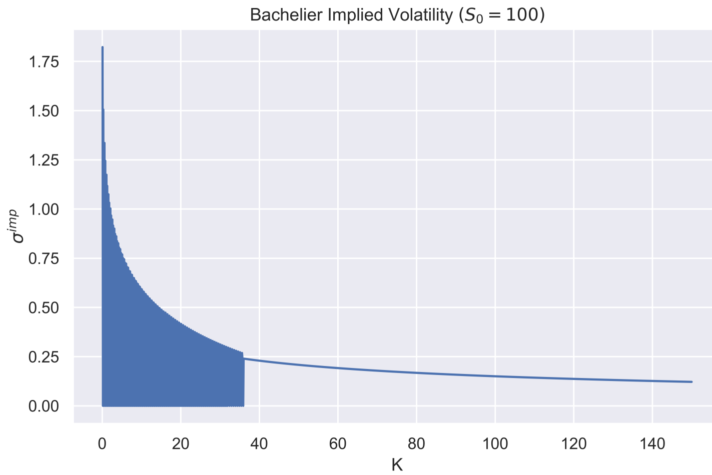

Figure 2

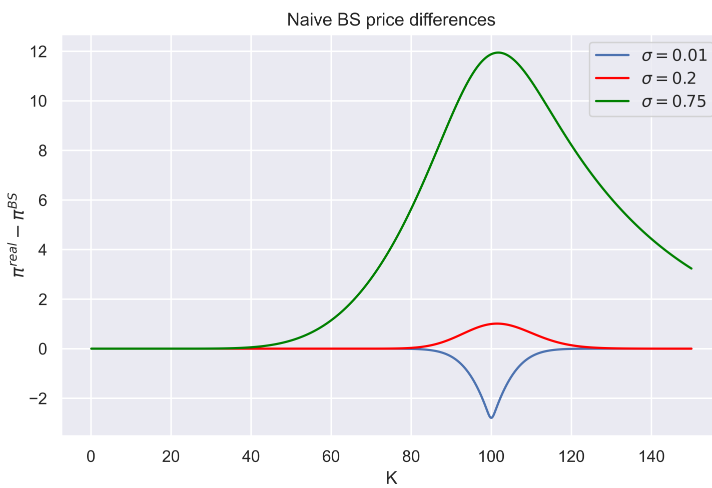

Figure 3

With that knowledge, we will in our next plot filter away all K where the optimal solution has $\Phi(d_1) = \Phi(d_2) = 1$, leaving only the certain solutions. Knowing that this is the issue, opens for the possibility of restructuring the algorithm in a more efficient way. We know that:

$$\Phi(x) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right) = \frac{1}{2}\left(1 + 1 - \text{erfc}\left(\frac{x}{\sqrt{2}}\right)\right) = 1 - \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right)$$

Where $\text{erf}(x)$ is the error function and $\text{erfc}(x)$ is the complementary error function. $\text{erfc}(x)$ is actually extremely effecient to calculate for absurdly large x, as the tiny result can be stored efficiently in a $xe - y$ format. The same is however not true for the error function as it is extremely inefficient to store 0.999.... This gives the idea, what if we can rewrite our optimisation problem as a function of $\text{erfc}(x)$? We have the problem:

$$\pi^{real} = \pi^{BS}(\sigma^{imp})$$

Inserting the Black-Scholes and rewriting we can get:

$$\pi^{real} = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

$$\pi^{real} = S_t\left(1 - \frac{1}{2}\text{erfc}\left(\frac{d_1}{\sqrt{2}}\right)\right) - Ke^{-r(T-t)}\left(1 - \frac{1}{2}\text{erfc}\left(\frac{d_2}{\sqrt{2}}\right)\right)$$

$$\pi^{real} = S_t - S_t\frac{1}{2}\text{erfc}\left(\frac{d_1}{\sqrt{2}}\right) - Ke^{-r(T-t)} + Ke^{-r(T-t)}\frac{1}{2}\text{erfc}\left(\frac{d_2}{\sqrt{2}}\right)$$

$$2\left(\pi^{real} - S_t + Ke^{-r(T-t)}\right) = -S_t\text{erfc}\left(\frac{d_1}{\sqrt{2}}\right) + Ke^{-r(T-t)}\text{erfc}\left(\frac{d_2}{\sqrt{2}}\right) \tag{2.20}$$

While the right hand side is much easier to implement for high precision here, it requires an extremely high precision π^{real}. This ruins the value of the restructuring, but if we insert our Bachelier formula, then possibly we can keep simplifying:

$$2\left((S_t - K)\Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) - S_t + Ke^{-0(T-t)}\right) = -S_t\text{erfc}\left(\frac{d_1}{\sqrt{2}}\right) + Ke^{-0(T-t)}\text{erfc}\left(\frac{d_2}{\sqrt{2}}\right)$$

$$2\left((S_t - K)\left(1 - \frac{1}{2}\text{erfc}\left(\frac{\frac{S_t-K}{\sigma\sqrt{T-t}}}{\sqrt{2}}\right)\right) + \sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) - S_t + K\right) = -S_t\text{erfc}\left(\frac{d_1}{\sqrt{2}}\right) + K\text{erfc}\left(\frac{d_2}{\sqrt{2}}\right)$$

$$2\left(-(S_t - K)\frac{1}{2}\text{erfc}\left(\frac{S_t - K}{\sqrt{2}\sigma\sqrt{T-t}}\right) + \sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right)\right) = -S_t\text{erfc}\left(\frac{d_1}{\sqrt{2}}\right) + K\text{erfc}\left(\frac{d_2}{\sqrt{2}}\right)$$

$$-(S_t - K)\text{erfc}\left(\frac{S_t - K}{\sqrt{2}\sigma\sqrt{T-t}}\right) + 2\sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) = -S_t\text{erfc}\left(\frac{d_1}{\sqrt{2}}\right) + K\text{erfc}\left(\frac{d_2}{\sqrt{2}}\right) \tag{2.21}$$

This new method works extremely well for small K, but it has the opposite effect for large K as the complementary error functions will be large. We can however handle this by using the definition erfc(x) $= 2 -$ erfc($-$x):

$$-(S_t - K)\left(2 - \text{erfc}\left(-\frac{S_t - K}{\sqrt{2}\sigma\sqrt{T-t}}\right)\right) + 2\sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) = -S_t\left(2 - \text{erfc}\left(-\frac{d_1}{\sqrt{2}}\right)\right) + K\left(2 - \text{erfc}\left(-\frac{d_2}{\sqrt{2}}\right)\right)$$

$$(S_t - K)\text{erfc}\left(-\frac{S_t - K}{\sqrt{2}\sigma\sqrt{T-t}}\right) + 2\sigma\sqrt{T-t}\phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right) = S_t\text{erfc}\left(-\frac{d_1}{\sqrt{2}}\right) - K\text{erfc}\left(-\frac{d_2}{\sqrt{2}}\right) \tag{2.22}$$

We now update our algorithm to solve for (2.21) if $K < S$ and (2.22) else. Remember that σ is the original volatility used for pricing the Bachelier options and that σ^{imp} is hidden in $d_{1/2}$. We first update our goal definition:

```
d = (v['S']-v['K'])/(v['sigma']*stats.sqrt(v['T']))
goalLow = -(v['S']-v['K'])*stats.erfc(d/stats.sqrt(2))+2*v['sigma']*stats.sqrt(v['T'])*stats.
    norm_pdf(d)
goalHigh = (v['S']-v['K'])*stats.erfc(-d/stats.sqrt(2))+2*v['sigma']*stats.sqrt(v['T'])*stats.
    norm_pdf(d)
goal = np.where(v['K']<v['S'],goalLow,goalHigh)
```

And then the calculated difference in our loop:

```
d1 = (stats.log(v['S'] / v['K']) + (v['r'] + 0.5 * sigma ** 2) * v['T']) / (sigma * stats.sqrt(v['
    T']))
d2 = d1 - sigma * stats.sqrt(v['T'])
hitLow = - v['S']*stats.erfc(d1/stats.sqrt(2)) + v['K']*np.exp(-v['T']*v['r'])*stats.erfc(d2/stats
    .sqrt(2))
hitHigh = v['S']*stats.erfc(-d1/stats.sqrt(2)) - v['K']*np.exp(-v['T']*v['r'])*stats.erfc(-d2/
    stats.sqrt(2))
diff = 0.5*(goal - np.where(v['K']<v['S'],hitLow,hitHigh))#scaled to remove 2x factor
```

The *diff* variable still denotes the price difference between the two models, but it now has much higher accuracy for far ITM or OTM options. We use this new algorithm to re-calculate the implied volatilities, but also with $S_0 = 50$ this time. We plot the results together in figure 4.

We see that there is a clear curvature shape on both of the options' volatilities, which is caused by the fact that the Black-Scholes volatility is dependent on the asset price through GBM. Large price movements upwards

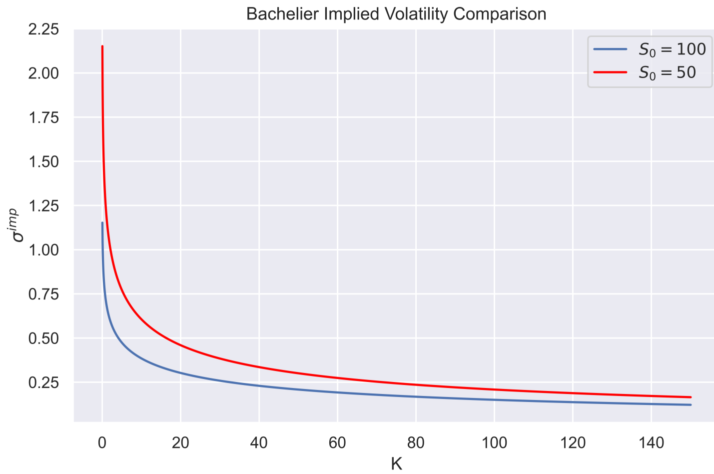

Figure 4

are more likely in the Black-Scholes than in the Bachelier model, leading to a very low required volatility to match the price for higher strikes. Meanwhile, steep declines in asset prices are more unlikely, as the volatility of the asset declines with the asset price. This also means that the optionality aspect is much less likely to become relevant for far ITM options. This is very different from the Bachelier model, where the asset prices even can go negative. In addition, the direct loading of the volatility in the Black-Scholes model would also fall with a decrease in the asset price, making rebounds far less likely than in the Bachelier model. For $S_t = 0.1$ then a $\sigma^{BS} = 150$ would be required to match the absolute volatility of the Bachelier model with $\sigma = 15$. This is however not the implied volatility we observe, as the upward skew of the lognormal model and the fact that σ^{BS} is not constant brings it down.As a result of these things, the Black-Scholes model requires higher volatility to match the Bachelier model for ITM options.

This slope is also reminiscent of the one that appears in the Heston model, when one models a correlation between the asset price and the volatility. In the Bachelier model there is constant volatility, while the Black-Scholes model has a deterministic and constant, but geometric volatility. Therefore, the Black-Scholes model has a relationship between the absolute volatility of the asset price and the price itself. This is obviously not the case in the Bachelier model.

Comparing the two lines, we see some interesting differences. Both curves are places such that they have approximately the same absolute volatility as the Bachelier model. That means that $\sigma^{imp,ATM} = 15/S_0$. We also see that the tangets at the ITM spot is approximately twice as steep for the $S_0 = 50$ model. Additionally, the general shape of the of the two models are identical. This gives us the exact relationship between models with different S_0. Changing the starting price of the underlying equates to shifting the curve such that $\sigma^{imp,ATM} \approx 15/S_0$ and then scaling the y-axis / the slope such that $\dfrac{d\sigma^{imp,ATM}}{dK} \approx -S_0^{-1}$, exactly the inverse loading factor in GBM. It must however be remembered that the slope is decreasing in K and that the approximation only holds at the ATM point.

For $K \to \infty$ we see the two curves appear to converge. The "slight" difference in starting points becomes less important compared to the extremely tiny probability of becoming ITM, the curves should however not fully converge before the option prices converge on 0, ignoring precision limits/machine errors.

2.3 Constant Interest Rate

We have now been asked to determine the arbitrage-free pricing formula for the Bachelier call option with a non-zero constant interest rate. Again, we start with the risk neutral valuation formula as in equation (2.5), but now without inserting $r = 0$:

$$\pi_t = e^{-r(T-t)}\mathrm{E}_t^Q[F(S_T)]$$

We also take our asset price process with r from equation (2.6), which was defined such that $S_t/B_t = e^{-rt}S_t$ is a Q martingale:

$$dS_t = rS_t dt + \sigma dW_t^Q$$

We now have a geometric drift, but still absolute volatility. To find the closed form solution, we can start by creating a discounted process:

$$D_t = f(S_t, t) = e^{-rt}S_t$$

We can now use Itô's Lemma (1.4) to determine the dynamics of this process.

$$f_t = -re^{-rt}S_t$$

$$f_s = e^{-rt}$$

$$f_{ss} = 0$$

$$\mu_t = rS_t$$

$$\sigma_t = \sigma$$

$$dD_t = df(X_t, t) = \left(-re^{-rt}S_t + e^{-rt}rS_t + \frac{1}{2}\sigma^2 0\right)dt + \sigma e^{-rt}dW_t = \sigma e^{-rt}dW_t^Q \tag{2.23}$$

As expected the discounted process has no drift, and we can write it on integral form:

$$D_{t+s}|\mathcal{F}_t = D_t + \int_t^{t+s} \sigma e^{-rt}dW_u^Q \tag{2.24}$$

We can now write up our asset process as a function of the discounted process:

$$S_{t+s}|\mathcal{F}_t = e^{r(t+s)}D_{t+s}|\mathcal{F}_t = e^{r(t+s)}e^{-rt}S_t + e^{r(t+s)}\int_t^{t+s} \sigma e^{-ru}dW_u^Q = e^{rs}S_t + \int_t^{t+s} \sigma e^{r(t+s-u)}dW_u^Q \tag{2.25}$$

While this equation is slightly complicated, we can use Björk's theorem 4.18 to determine the distribution of the asset price:

$$\mathrm{E}_t[S_{t+s}] = \mathrm{E}_t[e^{rs}S_t + \sigma e^{r(t+s)}\int_t^{t+s} e^{-ru}dW_u^Q] = e^{rs}S_t + 0 = e^{rs}S_t$$

$$\mathrm{Var}_t[S_{t+u}] = \int_t^{t+s} \sigma_u^2 du = \int_t^{t+s} (\sigma e^{r(t+s-u)})^2 du = \int_t^{t+s} \sigma^2 e^{2r(t+s-u)}du = \sigma^2 \frac{e^{2rs}-1}{2r}$$

$$S_{t+s}|\mathcal{F}_t \sim N\left(e^{rs}S_t, \sigma^2\frac{e^{2rs}-1}{2r}\right)$$

$$S_T|\mathcal{F}_t \sim N\left(e^{r(T-t)}S_t, \sigma^2\frac{e^{2r(T-t)}-1}{2r}\right) \tag{2.26}$$

We are however probably more interested in the discounted asset price, considering the risk neutral valuation formula. Using normal distribution rules for multiplying a coefficient on a distribution, we can determine the distribution of the discounted asset:

$$e^{-r(T-t)}S_T|\mathcal{F}_t \sim N\left(e^{-r(T-t)}e^{r(T-t)}S_t, e^{-2r(T-t)}\sigma^2\frac{e^{2r(T-t)}-1}{2r}\right)$$

$$D_T = e^{-r(T-t)}S_T|\mathcal{F}_t \sim N\left(S_t, \sigma^2\frac{1-e^{-2r(T-t)}}{2r}\right) \tag{2.27}$$

Plugging this into the risk neutral valuation formula, together with the call formula, we have:

$$
\begin{aligned}
\pi_t = e^{-r(T-t)}\mathrm{E}_t^Q[F(S_T)] =& \mathrm{E}_t^Q[e^{-r(T-t)}(S_T-K)^+] = \mathrm{E}_t^Q[(e^{-r(T-t)}S_T - e^{-r(T-t)}K)^+] \\
=& \mathrm{E}_t^Q[(D_T - e^{-r(T-t)}K)\mathbf{1}_{X_T > e^{-r(T-t)}K}] \\
=& \mathrm{E}_t^Q[D_T \mathbf{1}_{D_T > e^{-r(T-t)}K}] - e^{-r(T-t)}K\mathrm{E}_t^Q[\mathbf{1}_{D_T > e^{-r(T-t)}K}] \\
=& \mathrm{E}_t^Q[D_T \mathbf{1}_{e^{-r(T-t)}K < D_T < \infty}] - e^{-r(T-t)}K\mathrm{P}_t^Q[D_T > e^{-r(T-t)}K] \\
=& \mathrm{E}_t^Q[D_T \mathbf{1}_{e^{-r(T-t)}K < D_T < \infty}] - e^{-r(T-t)}K\mathrm{P}_t^Q[S_T > K]
\end{aligned}
$$

Considering that S_T is a normally distributed variable, then we can rewrite it as a function of a standard normal variable:

$$
S_T = \mathrm{E}_t[S_T] + \sqrt{\mathrm{Var}_t[S_T]}Z = e^{r(T-t)}S_t + \sqrt{\sigma^2 \frac{e^{2r(T-t)}-1}{2r}}Z = e^{r(T-t)}S_t + \sigma\sqrt{\frac{e^{2r(T-t)}-1}{2r}}Z
$$

Where Z is a standard normal random variable. We can now use this, equation (2.4) and the equations (2.9) to (2.15) to determine the pricing formula.

$$
\begin{aligned}
\pi_t =& S_t\left(\Phi\left(\frac{\infty - S_t}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right) - \Phi\left(\frac{e^{-r(T-t)}K - S_t}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right)\right) \\
&+\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}\left(\phi\left(\frac{e^{-r(T-t)}K - S_t}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right) - \phi\left(\frac{\infty - S_t}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right)\right) \\
&-e^{-r(T-t)}K\mathrm{P}_t^Q\left[e^{r(T-t)}S_t + \sigma\sqrt{\frac{e^{2r(T-t)}-1}{2r}}Z > K\right] \\
=& S_t\left(1 - \Phi\left(\frac{e^{-r(T-t)}K - S_t}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right)\right) + \sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}\left(\phi\left(-\frac{e^{-r(T-t)}K - S_t}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right) - 0\right) \\
&-e^{-r(T-t)}K\mathrm{P}_t^Q\left[Z > \frac{K - e^{r(T-t)}S_t}{\sigma\sqrt{\frac{e^{2r(T-t)}-1}{2r}}}\right] \\
=& S_t\Phi\left(\frac{S_t - e^{-r(T-t)}K}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right) - e^{-r(T-t)}K\Phi\left(\frac{e^{r(T-t)}S_t - K}{\sigma\sqrt{\frac{e^{2r(T-t)}-1}{2r}}}\right) \\
&+\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}\phi\left(\frac{S_t - e^{-r(T-t)}K}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right)
\end{aligned}
$$

To simplify this equation, we can create a variable:

$$
d = \frac{S_t - e^{-r(T-t)}K}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}} = \frac{e^{r(T-t)}S_t - e^{r(T-t)}e^{-r(T-t)}K}{e^{r(T-t)}\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}} = \frac{e^{r(T-t)}S_t - K}{\sigma\sqrt{\frac{e^{er(T-t)}-1}{2r}}} \tag{2.28}
$$

We also see that:

$$\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}} = \left(\frac{S_t - e^{-r(T-t)}K}{\sigma\sqrt{\frac{1-e^{-2r(T-t)}}{2r}}}\right)^{-1} \cdot \left(S_t - e^{-r(T-t)}K\right) = d^{-1} \cdot \left(S_t - e^{-r(T-t)}K\right)$$

We are now prepared to insert in our pricing equation:

$$\begin{aligned} \pi_t =& S_t\Phi\left(d\right) - e^{-r(T-t)}K\Phi\left(d\right) + d^{-1} \cdot \left(S_t - e^{-r(T-t)}K\right)\phi\left(d\right) \\ =& \left(S_t - e^{-r(T-t)}K\right)\left(\Phi(d) + d^{-1}\phi(d)\right) \end{aligned} \tag{2.29}$$

We have hereby determined the arbitrage free pricing equation in the Bachelier with a constant interests rate.

3 Quanto Puts

This section is focused on the pricing and hedging of Quanto puts. More specifically, it is puts on Japanese stocks that pay the payoff in USD. I.e. a contract that allows an investor to bet on the fall of the stock price of an individual Japenese compay, without having (direct) exposure to changes in the exchange rate. We have the following process' under the US-Martingale Measure:

$$dB_t^{US} = r_{US} B_t^{US} dt \tag{3.1}$$

$$dX_t = X_t \left(r_{US} - r_J\right) dt + X_t \sigma_X^T dW(t) \tag{3.2}$$

$$dB_t^J = r_J B_t^J dt \tag{3.3}$$

$$dS_t^J = S_t^J \left(r_J - \sigma_X^T \sigma_J\right) dt + S_t^J \sigma_J^T dW_t \tag{3.4}$$

In addition, the quanto put is defined as having the payoff:

$$F(T, S_T^J) = Y_0 (K - S_T^J)^+ \tag{3.5}$$

Finally, for numerical calculations we assume that $r_{US} = 3\%, r_J = 0\%, \sigma_X^T = \begin{pmatrix} 0.1 & 0.02 \end{pmatrix}$ and $\sigma_J^T = \begin{pmatrix} 0 & 0.25 \end{pmatrix}$.

3.1 Arbitrage-Free Pricing

We have to show that the arbitrage free time t price of the quanto put is:

$$F^{QP}(t, S_t^J) = Y_0 e^{-r_{US}(T-t)} \left(K\Phi(-d_2) - e^{(r_J - \sigma_X^T \sigma_J)(T-t)} S_t^J \Phi(-d_1)\right) \tag{3.6}$$

Where:

$$d_1 = \frac{\ln(S_t^J/K) + (r_J - \sigma_X^T \sigma_J + ||\sigma_J||^2/2)(T-t)}{\sqrt{T-t}||\sigma_J||} \tag{3.7}$$

$$d_2 = d_1 - \frac{||\sigma_J||^2 (T-t)}{\sqrt{T-t}||\sigma_J||} = d_1 - ||\sigma_J||\sqrt{T-t} \tag{3.8}$$

Afterwards, we have to show that the delta of the option is:

$$g(t, S_T) = \frac{\partial F^{QP}(t, S_t^J)}{\partial S_T^J} = Y_0 e^{(r_J - \sigma_X^T \sigma_{US} - r_{US})(T-t)} (\Phi(d_1) - 1) \tag{3.9}$$

We see that the following calculations are very similar to Black-Scholes, but that one must make sure to handle the more complicated drift and the different d_1 and d_2. In addition, inserting our well known Black-Formulas, making some small adjustments and then praying that they still hold feels unsatisfying. Therefore, we will calculate without any Black-Scholes tricks, but with the same method as if we were deriving the Black-Scholes price and delta from "scratch". To accomplish this, we use the Risk-Neutral valuation formula, that ensures we have a martingale under the US risk-free rate:

$$\begin{aligned} F^{QP}(t, S_t) =& e^{-r_{US}(T-t)} \mathrm{E}_t \left[F(T, S_T^J)\right] = e^{-r_{US}(T-t)} \mathrm{E}_t \left[Y_0 (K - S_T^J)^+\right] = Y_0 e^{-r_{US}(T-t)} \mathrm{E}_t \left[(K - S_T^J)\mathbf{1}_{K > S_T^J}\right] \\ =& Y_0 e^{-r_{US}(T-t)} \left(K\mathrm{E}_t \left[\mathbf{1}_{K > S_T^J}\right] - \mathrm{E}_t \left[S_T^J \mathbf{1}_{K > S_T^J}\right]\right) = Y_0 e^{-r_{US}(T-t)} \left(K\mathrm{P}_t \left[K > S_T^J\right] - \mathrm{E}_t \left[S_T^J \mathbf{1}_{K > S_T^J}\right]\right) \end{aligned}$$

We now see something very reminiscent of the Black-Scholes, but we must remember the added complexity of S_t^J from being a foreign stock price. A good next step would be to write our GBM for S_t^J as a closed form solution. From Björk's proposition 5.2, then if we have that a GBM defined as:

$$dX_t = aX_t dt + \sigma X_t dW_t \tag{3.10}$$

Then the solution is:

$$X_T = x_t e^{\left(a - \frac{1}{2}\sigma^2\right)(T-t) + \sigma W_{T-t}} \tag{3.11}$$

We have a two-dimensional Brownian motion, but our asset process is one dimensional so it is fine to use. It is clear that:

$$\sigma^2 = (\sigma_J^T)^2 = \sigma_J^T \sigma_J = ||\sigma_J||^2 \tag{3.12}$$

Inserting our case yields:

$$\begin{aligned} S_T^J =& S_t^J e^{\left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t) + \sigma_J^T W_{T-t}} = S_t^J e^{\left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t) + \sigma_J^T \sqrt{T-t} Z_2} \\ =& S_t^J e^{\left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t) + ||\sigma_J|| \sqrt{T-t} Z} \end{aligned} \tag{3.13}$$

Where Z_2 is a vector of 2 independent standard normal variables. In the last step we use the following idea:

$$\sigma_J^T Z_2 = \sigma_{J,1} Z_{2,1} + \sigma_{J,2} Z_{2,2} \sim N(0, \sigma_{J,1}^2 + \sigma_{J,2}^2) = N(0, ||\sigma_J||^2) \tag{3.14}$$

Using the definition of inner vector products and addition rules for independent normal distributed random variables. We can now insert this in our pricing formula, and continue as normal in the Black-Scholes deriviation. To simplify the next calculations we will first simplify the inequality seperately:

$$\begin{aligned} K > S_t^J e^{\left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t) + ||\sigma_J|| \sqrt{T-t} Z} &\Leftrightarrow \\ \ln(K/S_t^J) > \left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t) + ||\sigma_J|| \sqrt{T-t} Z &\Leftrightarrow \\ \frac{\ln(K/S_t^J) - \left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t)}{||\sigma_J|| \sqrt{T-t}} > Z &\Leftrightarrow \\ -\frac{\ln(S_t^J/K) + \left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t)}{||\sigma_J|| \sqrt{T-t}} > Z &\Leftrightarrow \\ -d_2 > Z& \end{aligned} \tag{3.15}$$

$$\begin{aligned} F^{QP}(t, S_t) =& Y_0 e^{-r_{US}(T-t)} \left(K \mathrm{P}_t\left[-d_2 > Z\right] - \mathrm{E}_t\left[S_t^J e^{\left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t) + ||\sigma_J|| \sqrt{T-t} Z} \mathbf{1}_{-d_2 > Z} \right] \right) \\ =& Y_0 e^{-r_{US}(T-t)} \left(K \Phi(-d_2) - e^{\left(r_J - \sigma_X^T \sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t)} S_t^J \mathrm{E}_t\left[e^{||\sigma_J|| \sqrt{T-t} Z} \mathbf{1}_{-d_2 > Z} \right] \right) \end{aligned}$$

We now need to do a substitution of the integrant to solve our expectation, to avoid clutter we do this separately.

$$\begin{aligned} &\mathrm{E}_t\left[e^{||\sigma_J|| \sqrt{T-t} Z} \mathbf{1}_{-d_2 > Z} \right] \\ &= \int_{-\infty}^{\infty} e^{||\sigma_J|| \sqrt{T-t} Z} \mathbf{1}_{-d_2 > Z} \phi(Z) dZ \\ &= \int_{-\infty}^{\infty} e^{||\sigma_J|| \sqrt{T-t} Z} \mathbf{1}_{-d_2 > Z} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} dZ \\ &= \int_{-\infty}^{\infty} \mathbf{1}_{-d_2 > Z} \frac{1}{\sqrt{2\pi}} e^{||\sigma_J|| \sqrt{T-t} Z - Z^2/2} dZ \end{aligned}$$

We want something standard normally distributed, and we also want something that makes our indicator function nice and simple. Thinking of the relationship between d_1 and d_2, we can try:

$$\begin{aligned} X =& Z - ||\sigma_J||\sqrt{T-t} \Leftrightarrow \\ Z =& X + ||\sigma_J||\sqrt{T-t} \end{aligned} \tag{3.16}$$

Inserting this in our integral:

$$\begin{aligned} &\mathrm{E}_t\left[e^{||\sigma_J||\sqrt{T-t}Z}\mathbf{1}_{-d_2>Z}\right] \\ &\quad= \int_{-\infty}^{\infty} \mathbf{1}_{-d_2>X+||\sigma_J||\sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{||\sigma_J||\sqrt{T-t}(X+||\sigma_J||\sqrt{T-t})-(X+||\sigma_J||\sqrt{T-t})^2/2} dX \\ &\quad= \int_{-\infty}^{\infty} \mathbf{1}_{-d_2-||\sigma_J||\sqrt{T-t}>X} \frac{1}{\sqrt{2\pi}} e^{X||\sigma_J||\sqrt{T-t}+||\sigma_J||^2(T-t)-X^2/2-||\sigma_J||^2(T-t)/2-2X||\sigma_J||\sqrt{T-t}/2} dX \\ &\quad= \int_{-\infty}^{\infty} \mathbf{1}_{-d_1>X} \frac{1}{\sqrt{2\pi}} e^{1/2||\sigma_J||^2(T-t)-X^2/2} dX \\ &\quad= e^{1/2||\sigma_J||^2(T-t)} \int_{-\infty}^{-d_1} \frac{1}{\sqrt{2\pi}} e^{-X^2/2} dX \\ &\quad= e^{1/2||\sigma_J||^2(T-t)} P_t[Z < -d_1] = e^{1/2||\sigma_J||^2(T-t)}\Phi(-d_1) \end{aligned} \tag{3.17}$$

We can now insert this, yielding the final arbitrage-free pricing formula:

$$\begin{aligned} F^{QP}(t,S_t) =& Y_0 e^{-r_{US}(T-t)} \left(K\Phi(-d_2) - e^{\left(r_J - \sigma_X^T\sigma_J - \frac{1}{2}||\sigma_J||^2\right)(T-t)} S_t^J e^{1/2||\sigma_J||^2(T-t)}\Phi(-d_1)\right) \\ =& Y_0 e^{-r_{US}(T-t)} \left(K\Phi(-d_2) - e^{\left(r_J-\sigma_X^T\sigma_J\right)(T-t)} S_t^J \Phi(-d_1)\right) \end{aligned} \tag{3.18}$$

We have hereby shown the derivation of the price of the quanto put option.

Next we will show the delta by taking the partial derivative of the pricing formula with regards to the asset price.

$$\begin{aligned} \frac{\partial F^{QP}(t,S_t^J)}{\partial S_t^J} =& \frac{\partial}{\partial S_t^J}\left(Y_0 e^{-r_{US}(T-t)}\left(K\Phi(-d_2) - e^{\left(r_J-\sigma_X^T\sigma_J\right)(T-t)} S_t^J\Phi(-d_1)\right)\right) \\ =& Y_0 e^{-r_{US}(T-t)}\left(K\frac{\partial}{\partial S_t^J}\Phi(-d_2) - e^{\left(r_J-\sigma_X^T\sigma_J\right)(T-t)}\Phi(-d_1) - e^{\left(r_J-\sigma_X^T\sigma_J\right)(T-t)} S_t^J \frac{\partial}{\partial S_t^J}\Phi(-d_1)\right) \end{aligned}$$

Where we have used the product rule. We now need to calculate the derivatives of the CDF's using the chain rule.

$$\frac{\partial}{\partial S_t^J}\Phi(-d_1) = \phi(-d_1)\frac{\partial}{\partial S_t^J}(-d_1) = \phi(-d_1)\frac{1}{S_t^J\sqrt{T-t}||\sigma_J||} \tag{3.19}$$

$$\begin{aligned} \frac{\partial}{\partial S_t^J}\Phi(-d_2) =& \phi(-d_2)\frac{\partial}{\partial S_t^J}(-d_2) = \phi(-d_2)\frac{1}{S_t^J\sqrt{T-t}||\sigma_J||} \\ =& \frac{1}{\sqrt{2\pi}} e^{-d_2^2/2}\frac{1}{S_t^J\sqrt{T-t}||\sigma_J||} = \frac{1}{\sqrt{2\pi}} e^{-(d_1-||\sigma_J||\sqrt{T-t})^2}\frac{1}{S_t^J\sqrt{T-t}||\sigma_J||} \\ =& e^{-1/2||\sigma_J||^2(T-t)+d_1||\sigma_J||\sqrt{T-t}}\phi(-d_1)\frac{1}{S_t^J\sqrt{T-t}||\sigma_J||} \\ =& e^{-1/2||\sigma_J||^2(T-t)+\ln(S_t^J/K)+(r_J-\sigma_X^T\sigma_J+||\sigma_J||^2/2)(T-t)}\phi(-d_1)\frac{1}{S_t^J\sqrt{T-t}||\sigma_J||} \\ =& S_t^J/K e^{(r_J-\sigma_X^T\sigma_J)(T-t)}\phi(-d_1)\frac{1}{S_t^J\sqrt{T-t}||\sigma_J||} \end{aligned} \tag{3.20}$$

Inserting these we get:

$$
\begin{aligned}
\frac{\partial F^{QP}(t,S_t^J)}{\partial S_t^J} =& Y_0 e^{-r_{US}(T-t)} (K S_t^J / K e^{(r_J - \sigma_X^T \sigma_J)(T-t)} \phi(-d_1) \frac{1}{S_t^J \sqrt{T-t} ||\sigma_J||} \\
& - e^{\left(r_J - \sigma_X^T \sigma_J\right)(T-t)} \Phi(-d_1) - e^{\left(r_J - \sigma_X^T \sigma_J\right)(T-t)} S_t^J \phi(-d_1) \frac{1}{S_t^J \sqrt{T-t} ||\sigma_J||}) \\
=& Y_0 e^{-r_{US}(T-t)} \left(- e^{\left(r_J - \sigma_X^T \sigma_J\right)(T-t)} \Phi(-d_1) \right) \\
=& Y_0 e^{(r_J - \sigma_X^T \sigma_J - r_{US})(T-t)} \left(-(1 - \Phi(d_1)) \right) \\
=& Y_0 e^{(r_J - \sigma_X^T \sigma_J - r_{US})(T-t)} \left(\Phi(d_1) - 1 \right)
\end{aligned}
$$

As in the classic Black-Scholes model, two terms even out, which leaving a simple expression based on $\Phi(d_1)$. We now however have an altered d_1, effects from the drift of the exchange rate and the locked rate Y_0. We have hereby proved the delta of the option.

3.2 A Failed Hedge

We are to consider a discrete hedge experiment, where we attempt to hedge the option through a portfolio that rebalances n times at equidistant time-points. We have our values from above and then that the fixed rate $Y_0 = X_0 = 1/300$ and ATM strike $K = S_0^J = 30,000$. The time horison is $T = 2$. The strategy is to at rebalance point i hold $\Delta_{t_i}^{QP} = \frac{g(t_i, S_J^{t_i})}{X_{t_i}}$ units of stock, balancing with the domestic bank account. This is a simple strategy, that can be expressed as financing domestically to keep a foreign delta hedge of the stock. This is also how one normally would replicate standard option payoffs. We have to illustrate experimentally that this strategy does **not** replicate the pay-off of the option.

We start by creating a class that can price the quanto option and that can calculate the delta. For this we use the equations from 3.1.

```python
class quantoPut(claim):

    def payoff(self,**v):
        return v['Y']*np.maximum(v['K']-v['S'],0)

    def price(self,**v):
        with stats.work_precision(self.precision), np.errstate(divide='ignore', invalid='ignore'):
            #for T=0
            norm = np.dot(v['sigmaF'].T,v['sigmaF'])
            comp = np.dot(v['sigmaX'].T,v['sigmaF'])
            d1 = (stats.log(v['S']/ v['K'])+(v['rF']-comp+1/2*norm)*v['T'])/(stats.sqrt(norm*v['T'
                ]))
            d2 = d1 - stats.sqrt(norm*v['T'])
            price = v['Y']*stats.exp(-v['rD']*v['T'])*(v['K']*stats.norm_cdf(-d2)-stats.exp((v['rF
                ']-comp)*v['T'])*v['S']*stats.norm_cdf(-d1))
            return  np.where(v['T']==0,self.payoff(**v),price)

    def delta(self,**v):
        with stats.work_precision(self.precision), np.errstate(divide='ignore', invalid='ignore'):
            #for T=0
            norm = np.dot(v['sigmaF'].T,v['sigmaF'])
            comp = np.dot(v['sigmaX'].T,v['sigmaF'])
            d1 = (stats.log(v['S']/ v['K'])+(v['rF']-comp+1/2*norm)*v['T'])/(stats.sqrt(norm*v['T'
                ]))
            return np.where(v['T']==0,0,v['Y']*stats.exp((v['rF']-comp-v['rD'])*v['T'])*(stats.
                norm_cdf(d1)-1))
```

We now have the tools needed to implement our hedge, we show 100,000 simulations of attempted hedging with 1,000 rebalances in figure 6.

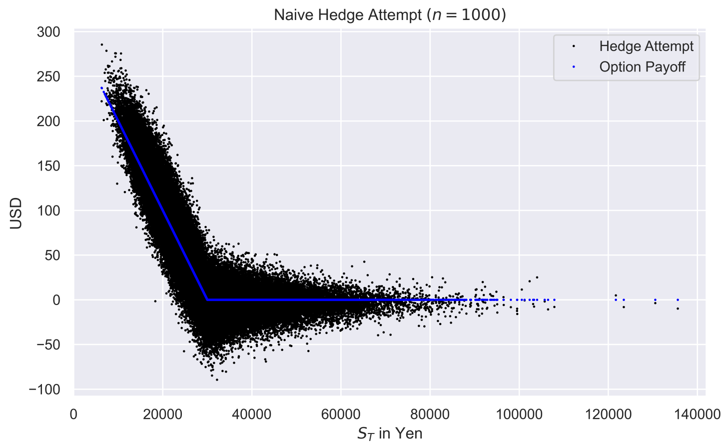

Figure 5

We observe that the hedge has some large deviations in performance. Especially for low S_T^J , the hedge appears to overshoot (likely related to the correlation between X_T and S_T^J). It is also slightly visible for high S_T^J, but not as clear due to the low values. Before we can conclude that the hedge fails, we should attempt to investigate the behaviour of the limit, as the errors simply could be the effect of trying to replicate a continuos hedge in discrete time. For this we use the measure known as the discounted hedge error:

$$e^{-r_{US}T}\varepsilon = e^{-r_{US}T}\left(F(T, S_T^J) - V_T^{Hedge}\right) \tag{3.21}$$

We can then compare the deviations of these errors for different n, hopefully it should converge to 0 as $n \to \infty$. The discounting is not important here, but it makes different options more comparable and is good practice. We have plotted the experiment in figure 6.

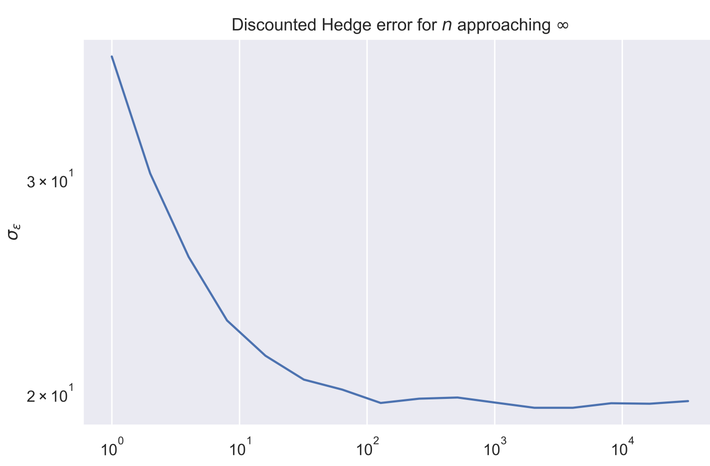

Figure 6: Standard deviation of discounted hedge errors for different amounts of hedge errors. 5,000 simulations were calculated for each n and used to compute σ_ε. Both σ_ε and n are on log-scales.

We see that the deviation of the discounted hedge errors decrease as n increases, especially for small n. It is also clear that the effect is diminishing as n increases, and that the errors stabilise at around 20, which is large. For a perfect hedge, we would expect that the errors would keep falling and converge on 0. We can therefore conclude that the errors are not simply the discrete time effect, but rather it is dominated by something else much larger. This is the effect of not hedging FX changes (X_T), which we will examine closer in the next two questions.

3.3 A Good Hedge

Simply using the foreign asset to attempt to reach the payoff of the option is poised to fail, as the quanto option pays in USD. Therefore the hedging portfolio will have full exposure to any changes in the exchange rate during the options life-time. To avoid this risk, we can try to include an FX hedge in our strategy.

We now include an additional position in our strategy, depositing $-\Delta_{t_i}^{QP} S_{t_i}^{J}$ Yen into the Japanese bank, to even out the original strategies direct exposure to changes in the exchange rate (delta exposure w.r.t. X_t) . We still use the domestic bank-account to keep the strategy self-financing. We must show that this strategy does replicate the pay-off of the quanto put, thereby being a suitable hedge.

We start by using our quanto implementation from before, and update our script to include the FX hedge. We make sure to keep the currencies correct and to keep the portfolio self-financing through B^D. Next we try recreating our first experiment with $n = 1000$ in figure 7.

Figure 7

We see that the new portfolio performs much better, demonstrating the need for the FX hedge. We still observe some errors, so our next step would be arguing that all the remaining errors are caused by the implementation of a continuous hedge in discrete time. We do this by again calculating the discounted hedge errors for different values of n. Now we should hopefully see a convergence to 0, such that all the remaining hedging errors stem from the discretization. We plot this in figure 8, together with a ln-ln regression.

We see that the plot shows a linear relationship between the ln transformed factors, which means that the rate of which the errors diminish as n increases is constant. The extremely well fitting regression also underscores this, and from its slope we can extract the exact relationship. We see that $\beta \approx -0.5$, which means that the errors have convergence order 1/2 and that the errors decrease with speed $O\left(1/\sqrt{n}\right)$. Interestingly, this is the same as that of standard Black-Scholes Call/Put options, which means that the FX element does not add any significant complexity, as long as we make sure to hedge it. We will not go into detail about why $O\left(1/\sqrt{n}\right)$ exactly is sufficient, but simply use the argument that if delta hedging in standard Black-Scholes does replicate the payoff in the limit, then so must this strategy with the same convergence order do.

3.4 Mathematical Proof

We must now prove mathematically that the strategy from 3.3 works, while the one from 3.2 does not. We start by disproving the first strategy. A good way to do this is by examining the process of a strategy that is long the hedge, but is short the option. If the hedge is perfect, then this process must by definition have no diffusion terms.

Figure 8: Standard deviation of the discounted hedge errors for the new hedge portfolio, calculated with 5,000 simulations for each n. Both scales are log transformed and a regression has been fit.

We first write up the USD value process for our a general portfolio:

$$\begin{aligned} dV^{\text{PF}}(t) =& V^{PF}_{SX} d(S^J_t X_t) + V^{PF}_{B^J X} d(B^J_t X_t) + V^{PF}_{B^D} dB^{US}_t \\ =& V^{PF}_{SX} \left(S^J_t dX_t + X_t dS^J_t + dS^J_t dX_t\right) + V^{PF}_{B^J X} \left(B^J_t dX_t + X_t dB^J_t + dB^J_t dX_t\right) + V^{PF}_{B^D} dB^{US}_t \\ =& V^{PF}_{SX} \left(S^J_t dX_t + X_t dS^J_t + S^J_t X^J_t \sigma^T_X \sigma_J dt\right) + V^{PF}_{B^J X} \left(B^J_t dX_t + r^J X_t B^J_t dt + 0\right) + r^{US} V^{PF}_{B^D} B^{US}_t dt \end{aligned} \tag{3.22}$$

Where $V^{PF}_{SX}, V^{PF}_{B^J X}$ and $V^{PF}_{B^D}$ is the exposure to each product in USD. We used Itô's product rule to reach the second equation and some basic stochastic calculation rules:

$$dtdt = 0 \tag{3.23}$$

$$dtdW_t = 0 \tag{3.24}$$

$$dW_t dW_t = dt \tag{3.25}$$

For the naive hedge we have:

$$\begin{aligned} dV^{\text{Naive}}(t) =& \frac{F^{QP}_S(t)}{X_t} \left(S^J_t dX_t + X_t dS^J_t + S^J_t X^J_t \sigma^T_X \sigma_J dt\right) + 0 \left(B^J_t dX_t + r^J X_t B^J_t dt\right) + r^{US} \left(-\frac{F^{QP}_S(t)}{X_t}\right) B^{US}_t dt \\ =& \frac{F^{QP}_S(t)}{X_t} \left(S^J_t dX_t + X_t dS^J_t + S^J_t X^J_t \sigma^T_X \sigma_J dt - r^{US} B^{US}_t dt\right) \end{aligned} \tag{3.26}$$

For the option, we can use Itô's lemma to get the dynamics.

$$\begin{aligned} dF^{QP}(t) =& F^{QP}_t dt + F^{QP}_S dS^J_t + \frac{1}{2} F^{QP}_{SS} (dS^J_t)^2 \\ =& F^{QP}_t dt + F^{QP}_S dS^J_t + \frac{1}{2} F^{QP}_{SS} (S^J_t)^2 ||\sigma_J|| dt \\ =& r^{US} F^{QP} dt + F^{QP}_S dS^J_t + \frac{1}{2} F^{QP}_{SS} (S^J_t)^2 ||\sigma_J|| dt \end{aligned} \tag{3.27}$$

Note, that here we have used the alternative Itô's formula from Björk's proposition 4.12. We have also used that the US discounted abitrage-free price must be a martingale under the US martingale measure. This does not really matter, as we do not care about the drift right now. We can now investigate the dynamics of the portfolio that is long the naive hedge and short the option.

$$
\begin{aligned}
dV^{Naive}(t) - dF^{QP} =& \frac{F_S^{QP}(t)}{X_t}\left(S_t^J dX_t + X_t dS_t^J + S_t^J X_t^J \sigma_X^T \sigma_J dt - r^{US} B_t^{US} dt\right) \\
&- \left(r^{US} F^{QP} dt + F_S^{QP} dS_t^J + \frac{1}{2} F_{SS}^{QP} (S_t^J)^2 ||\sigma_J|| dt\right) \\
=& \frac{F_S^{QP}(t)}{X_t}\left(S_t^J dX_t + S_t^J X_t^J \sigma_X^T \sigma_J dt - r^{US} B_t^{US} dt\right) - \left(r^{US} F^{QP} dt + \frac{1}{2} F_{SS}^{QP} (S_t^J)^2 ||\sigma_J|| dt\right)
\end{aligned}
\tag{3.28}
$$

Ignoring all the drift terms, we see that the naive portfolio has perfectly hedged the asset exposure of the option. It has however inadvertently gained an additional exposure to changes in the interest rate dX_t. This is an exposure that the quanto option does not have because of it's structure with a fixed exchange rate, and it therefore introduces new FX exposure to our portfolio. As the portfolio has risks (a dispersion term in dX_t) then we have proved mathematically that the strategy does not work for hedging the option.

We can now attempt with our new and improved strategy, it has dynamics:

$$
\begin{aligned}
dV^{\text{Good}}(t) =& \frac{F_S^{QP}(t)}{X_t}\left(S_t^J dX_t + X_t dS_t^J + S_t^J X_t^J \sigma_X^T \sigma_J dt\right) - \frac{F_S^{QP}(t)}{X_t} S_t^J \left(B_t^J dX_t + r^J X_t B_t^J dt\right) \\
&+ r^{US}\left(-\frac{F_S^{QP}(t)}{X_t} + \frac{F_S^{QP}(t)}{X_t} S_t^J\right) B_t^{US} dt \\
=& \frac{F_S^{QP}(t)}{X_t}\left(S_t^J dX_t + X_t dS_t^J + S_t^J X_t^J \sigma_X^T \sigma_J dt - S_t^J B_t^J dX_t - S_t^J r^J X_t B_t^J dt - r^{US}(S_t^J - 1) B_t^{US} dt\right) \\
=& \frac{F_S^{QP}(t)}{X_t}\left(X_t dS_t^J + S_t^J X_t^J (\sigma_X^T \sigma_J - r^J) dt - r^{US}(S_t^J - 1) B_t^{US} dt\right)
\end{aligned}
\tag{3.29}
$$

Where we in the last step used that $r^J = 0$ so $B_t^J = B_0^J = 1 \forall t$. It is evident that the FX risk of the position in the Japanese bank perfectly matches that of the Japanese asset, resulting in a portfolio that is delta neutral to changes in the exchange rate. We have thereby managed to hedge the additional FX exposure of the Japanese asset. We can now compute the dynamics of the portfolio that is long this hedge and short the option:

$$
\begin{aligned}
dV^{Good}(t) - dF^{QP} =& \frac{F_S^{QP}(t)}{X_t}\left(X_t dS_t^J + S_t^J X_t^J (\sigma_X^T \sigma_J - r^J) dt - r^{US}(S_t^J - 1) B_t^{US} dt\right) \\
&- \left(r^{US} F^{QP} dt + F_S^{QP} dS_t^J + \frac{1}{2} F_{SS}^{QP} (S_t^J)^2 ||\sigma_J|| dt\right) \\
=& \frac{F_S^{QP}(t)}{X_t}\left(S_t^J X_t^J (\sigma_X^T \sigma_J - r^J) dt - r^{US}(S_t^J - 1) B_t^{US} dt\right) \\
&- \left(r^{US} F^{QP} dt + \frac{1}{2} F_{SS}^{QP} (S_t^J)^2 ||\sigma_J|| dt\right)
\end{aligned}
\tag{3.30}
$$

We have now written out the dynamics of the portfolio and it is clear that there are no stochastic terms and only a drift. We can conclude that the second hedge perfectly replicates the option and that we can prove it mathematically. Our conclusion is that we can show mathematically that while the options price path does not directly depend on the exchange rate, then the asset that is required for replicating it does. This means that an effective hedge most also hedge this effect.

If we wanted to continue, then we would also see that the drifts would cancel out, leaving a portfolio with no drift (as no capital is demanded for selling the option and using the cash to replicate it), anything else would lead to arbitrage.