A note on the Heston model

David Šiška
School of Mathematics, University of Edinburgh

11th January 2018

1 Introduction

The classical Black–Scholes model assumes that the volatility is constant for the life-time of the option. This assumption is contradicted by empirical evidence. The so-called volatility smile is observed: the volatility implied by options with different strikes and maturities is different. When plotted as function of time-to-maturity (or strike) one often observes a rough "smile-like" shape of the graph.

Two essentially different approaches to improving the Black–Scholes model in order to fit empirical observations have been developed in the last decades of 20th century. The first approach, so-called "local volatility", assumes that the volatility is a deterministic function of time and risky asset price. It must be chosen as to match observed market option prices. This non-trivial problem has been independently solved by Duprie [7] and Derman, Kani [6]. The second approach, so-called "stochastic volatility", assumes that the volatility is itself a stochastic process. Here we focus on the second approach.

2 Heston stochastic volatility model

2.1 The model

Let $T > 0$ be given. Let us say we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a 2-dimensional Wiener process $W = (W(t))_{t\in[0,T]}$. Assume that we have a risk-free asset with price at time t given by $B(t)$ such that $B(0) = 1$ and $dB(t) = rB(t)dt$. Moreover there is a risky asset with the dynamics given by the stochastic differential equation

$$dS(t) = \mu S(t)dt + \sqrt{V(t)}S(t)dW_1(t), \;\; S(0) = S \tag{1}$$

where

$$dV(t) = \kappa(\theta - V(t))dt + \sigma\sqrt{V(t)}\left[\rho dW_1(t) + \sqrt{1-\rho^2}dW_2(t)\right], \;\; V(0) = v. \tag{2}$$

It can be shown that (2) has a unique solution such that, if $2\kappa\theta > \sigma^2$ then $V(t) > 0$ (provided of course that the initial value is strictly positive). Thus the square-root in (1) makes sense in the domain of real numbers and S is a well defined stochastic process that is also strictly positive for all t (provided that the initial value is strictly positive).

2.2 Contingent claim valuation and option pricing

The model consists of two traded assets: the risk-free B and the risky S. Such model is incomplete but it is free of arbitrage. Indeed we can use Girsanov's Theorem with the change of drift process

$$\varphi_1(t) = \frac{\mu - r}{\sqrt{V(t)}}, \;\; \varphi_2(t) = \frac{\lambda}{\sigma\sqrt{1-\rho^2}}\sqrt{V(t)} + \frac{\rho(r-\mu)}{\sqrt{1-\rho^2}\sqrt{V(t)}},$$

for arbitrary $\lambda \in \mathbb{R}$, which we choose later. With $\varphi = (\varphi_1, \varphi_2)^T$ we can define a new measure $\mathbb{Q}$ via

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^T \varphi(s)^T dW(s) - \frac{1}{2}\int_0^T |\varphi(s)|^2 ds\right) =: L(T)\,.$$

See Remark 3.2 for details why $\mathbb{E}L(T) = 1$ and hence why Girsanov's theorem applies.[1] Thus

$$W_1^{\mathbb{Q}}(t) := W_1(t) + \int_0^t \varphi_1(s)ds, \;\; W_2^{\mathbb{Q}}(t) := W_2(t) - \int_0^t \varphi_2(s)ds.$$

is a Wiener process under $\mathbb{Q}$. Then under measure $\mathbb{Q}$ we have

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sqrt{V(t)}S(t)\left[dW_1^{\mathbb{Q}}(t) + \frac{r-\mu}{\sqrt{V(t)}}dt\right] \\ &= rS(t)dt + \sqrt{V(t)}S(t)dW^{\mathbb{Q}}(t) \end{aligned}$$

and

$$\begin{aligned} dV(t) &= \kappa(\theta - V(t))dt \\ &\quad + \sigma\sqrt{V(t)}\left[\rho\left(dW_1^{\mathbb{Q}}(t) - \frac{\mu - r}{\sqrt{V(t)}}dt\right) + \sqrt{1-\rho^2}\left(dW_2^{\mathbb{Q}}(t) - \varphi_2(t)dt\right)\right] \\ &= \left[\kappa\theta - \kappa V(t) - \lambda V(t)\right]dt + \sigma\sqrt{V(t)}\left[\rho dW_1^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}dW_2^{\mathbb{Q}}(t)\right] \\ &= \kappa^*\left(\theta^* - V(t)\right)dt + \sigma\sqrt{V(t)}\left[\rho dW_1^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}dW_2^{\mathbb{Q}}(t)\right], \end{aligned}$$

where

$$\kappa^* := \kappa + \lambda, \;\; \theta^* := \frac{\kappa\theta}{\kappa + \lambda}.$$

To summarise, under some risk neutral measure $\mathbb{Q}$ we have

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)dW_1^{\mathbb{Q}}(t), \;\; S(0) = S \tag{3}$$

$$dV(t) = \kappa^*\left(\theta^* - V(t)\right)dt + \sigma\sqrt{V(t)}\left[\rho dW_1^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}dW_2^{\mathbb{Q}}(t)\right], \;\; V(0) = v. \tag{4}$$

The constant λ is related to the *market price of risk* and we see that it is not unique. So the risk-neutral measure $\mathbb{Q}$ is not unique. Nevertheless we see that the model is free of arbitrage and the price of any contingent claim X (up to the choice of λ) is given by

$$p(t) := B(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{X}{B(T)}\middle|\mathcal{F}_t\right]. \tag{5}$$

[1] If one wishes to use Novikov's condition then (applying Proposition 1 from [5]) there is some $T^* < \infty$ (which can be calculated from κ, θ, σ) such that $\mathbb{E}L(T) = 1$ for $T \le T^*$. However such T^* is too small for reasonable choice of model parameters and so one needs a different argument.

We will initially be interested in call (or put) options and so we will have

$$X = \max(S(T) - K, 0), \text{ for call, } X = \max(K - S(T), 0), \text{ for put.}$$

Moreover from the definition of B we obtain

$$p(t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[X|\mathcal{F}_t]$$

and in particular for a call option

$$C(t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\max(S(T) - K, 0)|\mathcal{F}_t]. \tag{6}$$

This fact that $\mathbb{Q}$ is not unique (and hence the price is not unique) is an interesting theoretical issue but in practice it just means that one has one more parameter, namely λ to choose when using the model. In practice one can either

1. Use time series and statistics to choose κ, θ and σ and then choose λ such that $C(t)$ matches one specific traded option or perhaps such that it minimises some error when considering more than one option. We would say that κ, θ, ρ and σ are calibrated to the "real-world measure" $\mathbb{P}$. The parameter λ is then related to the "market price of risk".

2. One can also do the entire calibration in the "risk-neutral" setting by taking at least five traded options and choosing κ^*, θ^*, ρ, σ and v such that the option prices implied by our model match the observed prices.

So the non-uniqueness of $\mathbb{Q}$ is not a problem in practice. What one has to be able to do however is to evaluate (6) efficiently. There are effectively three choices:

1. A Monte-Carlo based algorithm,

2. An algorithm based on a representation of (6) as a solution to some partial differential equation (PDE) and then a finite difference or perhaps finite element method,

3. A method based on "Heston formula". This is the most efficient.

Here we want to use the last method i.e. Heston formula for calibration. Thus we have to find out what the formula is.

2.3 Heston formula derivation

Consider a simple claim $X = g(S(T))$ where $g : [0, \infty) \to [0, \infty)$ is a given function e.g. $g(S) = \max(S - K, 0)$. Feynmann-Kac theorem tells us that in fact, the option price

$$c(t, S, v) := e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\max(S(T) - K, 0)|S(t) = S, v(t) = v]$$

satisfies

$$\begin{aligned} &\frac{\partial c}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 c}{\partial S^2} + \sigma\rho v S\frac{\partial^2 c}{\partial S \partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 c}{\partial v^2} \\ &\quad + rS\frac{\partial c}{\partial S} + \kappa^*(\theta^* - v)\frac{\partial c}{\partial v} - rc = 0 \quad \text{on} \quad [0, T) \times (0, \infty) \times (0, \infty), \end{aligned} \tag{7}$$

with the terminal condition

$$c(T,S,v) = \max(S-K,0) \quad \forall (S,v) \in [0,\infty) \times [0,\infty). \tag{8}$$

To solve the PDE we follow the approach in Heston [10]. See also Gatheral [8, Ch. 2]. Recall that the price of a call in the Black–Scholes model is given by

$$SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

and in fact we can think of $N(d_1)$ and $N(d_2)$ as some probabilities.[2] With this in mind we make the ansatz that $c(t,S,v)$ has the form[3]

$$c(t,S,v) = SP_1(t,S,v) - Ke^{-r(T-t)}P_2(t,S,v). \tag{9}$$

Furthermore we change variables by introducing $x = \ln S$ which means that (7) becomes

$$\begin{aligned}\frac{\partial c}{\partial t} + \frac{1}{2}v\frac{\partial^2 c}{\partial x^2} + \sigma\rho v\frac{\partial^2 c}{\partial x \partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 c}{\partial v^2} + \left(r - \frac{1}{2}v\right)\frac{\partial c}{\partial x} \\ + \kappa^*(\theta^* - v)\frac{\partial c}{\partial v} - rc = 0 \quad \text{on} \quad [0,T) \times (0,\infty) \times (0,\infty),\end{aligned} \tag{10}$$

Now we substitute (9) into (10) to derive PDEs for P_1 and P_2. Let

$$u_1 = 1/2, \;\; u_2 = -1/2, \;\; a = \kappa^*\theta^*, \;\; b_1 = \kappa^* - \rho\sigma, \;\; b_2 = \kappa^*.$$

If for $j = 1,2$

$$\begin{aligned}\frac{\partial P_j}{\partial t} + \frac{1}{2}v\frac{\partial^2 P_j}{\partial x^2} + \sigma\rho v\frac{\partial^2 P_j}{\partial x \partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_j}{\partial v^2} + (r + u_j v)\frac{\partial P_j}{\partial x} \\ + (a - b_j v)\frac{\partial P_j}{\partial v} = 0 \quad \text{on} \quad [0,T) \times (0,\infty) \times (0,\infty),\end{aligned}$$

then (10) must hold. The terminal condition (8) will be satisfied provided that for $j = 1,2$

$$P_j(t,x,v) = \mathbb{1}_{\{x \geq \ln K\}}.$$

We can see, using the Feynman–Kac theorem, that, in fact,

$$\begin{aligned}P_j(t,x,v) &= \mathbb{E}\left[\mathbb{1}_{\{x_j(T) \geq \ln K\}} | x_j(t) = x, v_j(t) = v\right] \\ &= \mathbb{Q}\left[x_j(T) \geq \ln K | x_j(t) = x, v_j(t) = v\right],\end{aligned}$$

where the processes $x = (x(t))_{t \in [t,T]}$ and $(v(t))_{t \in [t,T]}$ that are given by

$$\begin{aligned}dx_j(t) &= (r + u_j v(t))dt + \sqrt{v(t)}dW_1^{\mathbb{Q}}(t) \\ dv_j(t) &= (a - b_j v_j(t))dt + \sigma\sqrt{v(t)}\left[\rho dW_1^{\mathbb{Q}}(t) + \sqrt{1-\rho}dW_2^{\mathbb{Q}}(t)\right].\end{aligned}$$

To calculate such probabilities we need to know the distribution of $x_j(T)$. We know that the distribution of any random variable is fully characterised by its characteristic function. Thus we try to obtain these characteristic functions.

[2] You can find the form of d_1 and d_2 in any book that covers Black–Scholes formula. We do not need them here.

[3] This is just a wild stab into the dark; but it is worth a try and if you're an applied mathematician or a physicist you would have seen such tricks before.

Let us omit the subscript j in the next paragraph for brevity. We know that to obtain the characteristic function of $x(T)$, we need to calculate

$$\varphi(t,x,v,\phi) = \mathbb{E}[e^{i\phi X(T)}|x(t)=x, v(t)=v]$$

which, due to Feynman–Kac theorem, amounts to solving

$$\begin{aligned} \frac{\partial\varphi}{\partial t} + \frac{1}{2}v\frac{\partial^2\varphi}{\partial x^2} + \sigma\rho v\frac{\partial^2\varphi}{\partial x\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2\varphi}{\partial v^2} + (r+u_j v)\frac{\partial\varphi}{\partial x} \\ + (a-b_j v)\frac{\partial\varphi}{\partial v} = 0 \quad \text{on} \quad [0,T)\times(0,\infty)\times(0,\infty), \end{aligned} \tag{11}$$

with the terminal condition $\varphi(T,x,v,\phi) = e^{i\phi x}$. Now we make a guess at the form of φ. Assume there are $C(T-t)$ and $D(T-t)$ such that

$$\varphi(t,x,v,\phi) = \exp\left[C(T-t) + D(T-t)v + i\varphi x\right].$$

Then the PDE above becomes

$$-\frac{\partial C}{\partial t} - v\frac{\partial D}{\partial t} - \frac{1}{2}v\phi^2 - \sigma\rho viD + \frac{1}{2}\sigma^2 vD^2 + (r+uv)i\phi + (a-bv)D = 0.$$

Collecting terms that include $v > 0$ we get two ordinary differential equations (ODEs):

$$\begin{aligned} -\frac{\partial D}{\partial t} - \frac{1}{2}\phi^2 - \sigma\rho i\phi D + \frac{1}{2}\sigma^2 D^2 + ui\phi - bD = 0 \\ -\frac{\partial C}{\partial t} + ri\phi + aD = 0 \end{aligned}$$

with the initial conditions $D(0) = C(0) = 0$. If these ODEs hold then so does (11).

Solving the ODEs will give us the two characteristic functions φ_j, $j = 1,2$ with

$$\varphi_j(t,x,v,\phi) = \exp\left[C_j(T-t,\phi) + D_j(T-t,\phi)v + i\varphi x\right] \tag{12}$$

where (with $\tau := T-t$)

$$\begin{aligned} C_j(\tau,\phi) &= r\phi i\tau + \frac{a}{\sigma^2}\left\{(b_j - \rho\sigma\phi i + d_j(\phi))\tau - 2\ln\left[\frac{1-g_j(\phi)e^{\tau d_j(\phi)}}{1-g_j(\phi)}\right]\right\}, \\ D_j(\tau,\phi) &= \frac{b_j - \rho\sigma\phi i + d_j(\phi)}{\sigma^2}\left[\frac{1-e^{\tau d_j(\phi)}}{1-g_j(\phi)e^{\tau d_j(\phi)}}\right], \end{aligned} \tag{13}$$

and

$$\begin{aligned} g_j(\phi) &= \frac{b_j - \rho\sigma\phi i + d_j(\phi)}{b_j - \rho\sigma\phi i - d_j(\phi)}, \\ d_j(\phi) &= \sqrt{(\rho\sigma\phi i - b_j)^2 - \sigma^2(2u_j\phi i - \phi^2)}. \end{aligned} \tag{14}$$

Finally we can use the characteristic function to obtain

$$\begin{aligned} P_j(t,x,v) &= \mathbb{Q}[x(T) \geq \ln K | x(t)=x, v(t)=v] \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \operatorname{Re}\left[\frac{e^{-i\phi\ln K}\varphi_j(t,x,v,\phi)}{i\phi}\right]d\phi. \end{aligned} \tag{15}$$

This expression can be integrated numerically as the integrand decays rapidly as ϕ increases. Recall now the form of the call option price (9). We thus have the call option price in the Heston model given by

$$c(t,S,v) = SP_1(t,\ln S,v) - Ke^{-r(T-t)}P_2(t,\ln S,v).$$

To calculate this we need a numerical integration algorithm for evaluating (15) and recalling (12), (13) and (14) in order to evaluate the integrand.

2.4 Heston formula avoiding the little Heston trap

There is no way to choose a logarithm of each nonzero complex number z so as to make a function $\ln(z)$ continuous. One can only define "branches" of the logarithm function (open subsets of $\mathbb{C}\setminus\{0\}$ on which $z\ln z$ is analytic). Evaluating (13) one would need to go from one branch to another in the standard computer implementation of complex logarithm function. This means we will be jumping over a "branch cut". This (with details in [1]) leads to the following preferable choice of C_j, D_j and g_j which avoids branch cut issues.

$$
\begin{aligned}
C_j(\tau,\phi) &= r\phi i\tau + \frac{a}{\sigma^2}\left\{(b_j - \rho\sigma\phi i - d_j(\phi))\tau - 2\ln\left[\frac{1-g_j(\phi)e^{-\tau d_j(\phi)}}{1-g_j(\phi)}\right]\right\},\\
D_j(\tau,\phi) &= \frac{b_j-\rho\sigma\phi i - d_j(\phi)}{\sigma^2}\left[\frac{1-e^{-\tau d_j(\phi)}}{1-g_j(\phi)e^{-\tau d_j(\phi)}}\right],
\end{aligned}
\tag{16}
$$

where

$$
\begin{aligned}
g_j(\phi) &= \frac{b_j-\rho\sigma\phi i - d_j(\phi)}{b_j-\rho\sigma\phi i + d_j(\phi)},\\
d_j(\phi) &= \sqrt{(\rho\sigma\phi i - b_j)^2 - \sigma^2(2u_j\phi i - \phi^2)}.
\end{aligned}
\tag{17}
$$

Notice that d_j is same as before. And as before we have

$$
\varphi_j(t,x,v,\phi) = \exp\left[C_j(T-t,\phi) + D_j(T-t,\phi)v + i\phi x\right], \tag{18}
$$

and

$$
P_j(t,x,v) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \operatorname{Re}\left[\frac{e^{-i\phi\ln K}\varphi_j(t,x,v,\phi)}{i\phi}\right]d\phi. \tag{19}
$$

as before and, as before, we have

$$
c(t,S,v) = SP_1(t,\ln S,v) - Ke^{-r(T-t)}P_2(t,\ln S,v). \tag{20}
$$

3 Derivative pricing using Monte Carlo

Once we have a calibrated model (calibrated to observed put / call option prices) the price of any contingent claim is given by (5) where X is the time T payoff of the claim. Since this expression is in general impossible to evaluate exactly we need numerical approximation. Here we focus on Monte Carlo. A good introductory book for Monte Carlo Methods is Glasserman [9].

To use a Monte Carlo methods we need to "discretise" the SDEs (3) and (4). To that end let $N \in \mathbb{N}$ be given (denoting the number of time steps). Let $\tau := T/N$ be the time step. Let $t_k := k\tau$. To simplify notation let

$$
Z_1(t) := W_1^{\mathbb{Q}}(t),\ \ Z_2(t) := \rho W_1^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}W_2^{\mathbb{Q}}(t).
$$

Let $\Delta Z_j(t_k) := Z_j(t_k) - Z_j(t_{k-1},\ j=1,2,\ k=1,\dots,N$. A "naive" explicit Euler discretisation of (3) and (4) is

$$
S_N(t_{k+1}) = S_N(t_k) + rS_N(t_k)\tau + \sqrt{V_N(t_k)}S_N(t_k)\Delta Z_1(t_{k+1}) \tag{21}
$$

and

$$V_N(t_{k+1}) = V_N(t_k) + \kappa^*(\theta^* - V_N(t_k))\tau + \sigma\sqrt{V_N(t_k)}\Delta Z_2(t_{k+1}).$$

We can immediately notice that if $\Delta Z_2(t_{k+1})$ is sufficiently large and negative then $V(t_{k+1})$ will also be negative and the square root will not be defined as a real number (in the next time step).

We know that if the Feller condition $2\kappa\theta > \sigma^2$ is satisfied then $V(t) > 0$. Thus we can "correct" the scheme by taking

$$V_N(t_{k+1}) = \left| V_N(t_k) + \kappa^*(\theta^* - V_N(t_k))\tau + \sigma\sqrt{V_N(t_k)}\Delta Z_2(t_{k+1}) \right|. \tag{22}$$

We effectively "reflect" any "overshoot" into negative territory back into positive one. One can use this scheme (21) together with (22) for pricing derivatives and will get satisfactory results in most situations. However the convergence $\mathbb{E}|V_N(T) - V(T)|^2 \to 0$ as $N \to \infty$ does not follow from standard results on numerics of SDEs and one should only use this scheme with extreme caution.

An alternative scheme has been proposed in Dereich, Neuenkirch and Szpruch [4] and it is better to use that for approximation of $V(t)$. The scheme is as follows. Let

$$\alpha := \frac{4\kappa^*\theta^* - \sigma^{*2}}{8}, \quad \beta := -\frac{\kappa^*}{2}, \quad \gamma := \frac{\sigma}{2}.$$

Instead of $V(t)$ we discretise $Y(t) := \sqrt{V(t)}$ as follows:

$$Y_N(t_{k+1}) = \frac{Y_N(t_k) + \gamma\Delta Z_2(t_{k+1})}{2(1-\beta\tau)} + \sqrt{\frac{(Y_N(t_k) + \gamma\Delta Z_2(t_{k+1}))^2}{4(1-\beta\tau)^2} + \frac{\alpha\tau}{1-\beta\tau}}. \tag{23}$$

We discretise $S(t)$ exactly as in (21) but use $V_N(t_k) = Y(t_k)^2$.

To get one approximated sample path $s_N^0, s_N^1, \dots, s_N^N$ which can be used for derivative pricing we have to do the following.

i) Set $s_N^0 := S$, the initial asset price, set $y_N^0 := \sqrt{v}$, i.e. square root of the initial variance.

ii) Generate $\Delta z_{1,N}^k$ and $\Delta z_{2,N}^k$, by twice getting N iid samples from $N(0,1)$, call them $x_{1,N}^k$ and $x_{2,N}^k$. Let

$$\begin{aligned}
\Delta z_{1,N}^k &:= \sqrt{\tau}x_{1,N}^k, \\
\Delta z_{2,N}^k &:= \sqrt{\tau}\left(\rho x_{1,N}^k + \sqrt{1-\rho^2}x_{2,N}^k\right).
\end{aligned}$$

iii) Generate s_N^k and y_N^k recursively for $k = 1, \dots, N$ by replacing $S_N(t_k)$ with s_N^k, $\sqrt{V_N(t_k)}$ with y_N^k and $\Delta Z_1(t_k)$ with $\Delta z_{1,N}^k$ in (21) and by replacing $Y_N(t_k)$ by y_N^k and $\Delta Z_2(t_k)$ with $\Delta z_{2,N}^k$ in (23).

You can now calculate they derivative payoff for this one approximated sample path. By considering many approximated sample paths and averaging you get an approximation for the payoff of the derivative.

3.1 Change of Measure

The following Lemma is a simpler version of the result in [3] and is proved following the argument used in [2].

Lemma 3.1. *Let* $\lambda : \mathbb{R}^d \to \mathbb{R}^d$ *be locally bounded and measurable and* $\mu : \mathbb{R}^d \to \mathbb{R}^d$ *and* $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d'}$ *be measurable. Assume that*

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)dW_t^{\mathbb{P}},\ \ X_0 \sim \xi \tag{24}$$

and

$$dY_t = [\mu(Y_t) - \sigma(Y_t)\lambda(Y_t)]\ dt + \sigma(Y_t)dW_t^{\mathbb{P}},\ \ Y_0 \sim \psi \tag{25}$$

for some initial distributions ξ *and* ψ *have a weak solution* $(\Omega, \mathcal{F}, \mathbb{P}, W^{\mathbb{P}}, X, Y)$ *on* $[0,T]$*. Assume further that the solution to* (25) *is unique in the sense of probability law. Let*

$$Z_t := \exp\left(-\int_0^t \lambda(X_s)\,dW_s^{\mathbb{P}} - \frac{1}{2}\int_0^t |\lambda(X_s)|^2\,ds\right).$$

Then $\mathbb{E}^{\mathbb{P}}[Z_T] = 1$*.*

Proof. For $n \in \mathbb{N}$ let $T_n := \inf\{t \geq 0 : |\lambda(X_t)|^2 \geq n\}$ and $S_n := \inf\{t \geq 0 : |\lambda(Y_t)|^2 \geq n\}$. These are stopping times with respect to the filtration generated by $W^{\mathbb{P}}$. Since λ is locally bounded and since the processes X and Y are solutions to (24) and (25) on $[0,T]$ (so the trajectories cannot explode before T) we have[4]

$$\mathbb{P}\left(\lim_{n\to\infty} T_n = \infty\right) = 1 \quad \text{and} \quad \mathbb{P}\left(\lim_{n\to\infty} S_n = \infty\right) = 1\,. \tag{26}$$

Now let $\lambda_t^n := \lambda(X_t)1_{t\leq T_n}$ and

$$Z_t^n := \exp\left(-\int_0^t \lambda_s^n\,dW_s^{\mathbb{P}} - \frac{1}{2}\int_0^t |\lambda_s^n|^2\,ds\right).$$

It is easy to check that Z^n are $\mathbb{P}$-martingales. Hence, due to Girsanov's theorem, the processes

$$W_t^n := W_t^{\mathbb{P}} + \int_0^t \lambda_s^n\,ds$$

are Wiener processes under the measures $\mathbb{Q}^n \sim \mathbb{P}$, given by the Radon–Nikodym derivatives $d\mathbb{Q}^n/d\mathbb{P} = Z_T^n$. Moreover

$$dX_t = [\mu(X_t) - \sigma(X_t)\lambda_t^n]\ dt + \sigma(X_t)\,dW_t^n$$

and so

$$dX_{t\wedge T_n} = [\mu(X_{t\wedge T_n}) - \sigma(X_{t\wedge T_n})\lambda(X_{t\wedge T_n})]\ dt + \sigma(X_{t\wedge T_n})\,dW_t^n.$$

The uniqueness in law of solution to (25) can be used to deduce that the stopped processes $X_{t\wedge T_n}$ and $Y_{t\wedge S_n}$ must have the same law. This implies the crucial identity

$$\mathbb{Q}^n(T_n \geq T) = \mathbb{P}(S_n \geq T). \tag{27}$$

The proof can then be concluded as follows. First $Z_T^n 1_{T_n\geq T} = Z_T 1_{T_n \geq T}$ and this increasing sequence converges $\mathbb{P}$-almost surely to Z_T due to the first identity in (26).

[4] We follow the convention that infimum of an empty set is ∞.

Then, by monotone convergence, definition of $\mathbb{Q}^n$, by (27), and by the second identity in (26) we get

$$\mathbb{E}^{\mathbb{P}}[Z_T] = \lim_{n\to\infty} \mathbb{E}^{\mathbb{P}}[Z_T 1_{T_n \geq T}] = \lim_{n\to\infty} \mathbb{Q}^n[T_n \geq T] = \lim_{n\to\infty} \mathbb{P}[S_n \geq T] = 1.$$

□

Remark 3.2. We now wish to apply this in the case of the CIR process

$$dX(t) = \kappa(\theta - X(t))dt + \sigma\sqrt{X(t)}\left[\rho dW_1^{\mathbb{P}}(t) + \sqrt{1-\rho^2}dW_2^{\mathbb{P}}(t)\right], \quad X(0) = x. \tag{28}$$

and a change of measure function λ given by

$$\lambda_1(x) = \frac{r-\mu}{\sqrt{x}}, \quad \lambda_2(x) = -\frac{\gamma}{\sigma\sqrt{1-\rho^2}}\sqrt{x} - \frac{\rho(r-\mu)}{\sqrt{1-\rho^2}\sqrt{x}},$$

with $r, \mu, \gamma, \sigma, \rho \in \mathbb{R}$ constants. So the other SDE must consider is

$$\begin{aligned} dY(t) = &\left[\kappa(\theta - Y(t)) - \sigma\sqrt{Y(t)}\rho\lambda_1(Y(t)) - \sigma\sqrt{Y(t)}\sqrt{1-\rho^2}\lambda_2(Y(t))\right] dt \\ &+ \sigma\sqrt{Y(t)}\left[\rho dW_1(t) + \sqrt{1-\rho^2}dW_2(t)\right], \quad Y(0) = x. \end{aligned}$$

This is

$$dY(t) = \kappa^*(\theta^* - Y(t))\, dt + \sigma\sqrt{Y(t)}\left[\rho dW_1^{\mathbb{P}}(t) + \sqrt{1-\rho^2}dW_2^{\mathbb{P}}(t)\right], \quad Y(0) = x \tag{29}$$

where

$$\kappa^* := \kappa + \lambda, \quad \theta^* := \frac{\kappa\theta}{\kappa+\lambda}.$$

Both (28) and (29) have unique strong solutions and so Lemma 3.1 can be applied with the locally bounded function λ.

References

[1] Albrecher, H. and Mayer, P. and Schoutens, W and Tistaert, J., The Little Heston Trap, *Wilmott Magazine*, 2007.

[2] Cheridito, P. and Filipović, D. and Kimmel, R. L., Market price of risk specifications for affine models: Theory and evidence, *Journal of Financial Economics*, 83, 123–170, 2007.

[3] Cheridito, P. and Filipović, D. and Yor, M., Equivalent and Absolutely Continuous Measure Changes for Jump-Diffusion Processes. *Ann. Appl. Probab.*, 15(3), 1713–1732, 2005.

[4] Dereich, S. and Neuenkirch, A. and Szpruch, L., An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process, *Proceedings of the Royal Society A*, Vol. 468, No. 2140, pp. 1105–1115, 2011.

[5] Duffie, D. and Pan, J. and Singleton, K., Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 68(6), 1343–1376, 2000.

[6] Derman, E. and Kani, I., Riding on a smile, *Risk Magazine*, 1994.

[7] Dupire, B., Pricing and hedging with smiles, *Mathematics of Derivative Securities*, Cambridge Uni. Press 1997.

[8] Gatheral, J., *The Volatility Surface: A Practitioner's Guide*, Wiley 2006.

[9] Glasserman, P., *Monte Carlo Methods in Financial Engineering*, Springer, 2003.

[10] Heston, S. L., A Closed-Form Solution for Options with Stochastic Volatility and Applications to Bond and Currency Options, *The Review of Financial Studies*, 6(2) pp. 327–343 1997.