Chapter 4

Geometric Interest Rate Theory

4.1 The Problem

Any acceptable model which prices interest rate derivatives must fit the observed term structure. This idea pioneered by Ho and Lee [22], has been explored in the past by many other researchers like Black and Karasinski [12] and Hull and White [23].

The contemporary models are more complex because they consider the evolution of the whole forward curve as an infinite system of stochastic differential equations (Heath, Jarrow and Morton [21]). In particular, they use a continuous forward rate curve as initial input. In reality, one only observes a discrete set composed either by bond prices or swap rates. So, in practice, the usual approach is to interpolate the forward curve by using splines or other parametrized families of functions.

A very plausible question arises at this point: Choose a specific parametric family, $\mathcal{G}$, of functions that represent the forward curve, and also an arbitrage free interest rate model $\mathcal{M}$. Assume that we use an initial curve that lay within as input for model $\mathcal{M}$. Will this interest rate model evolve through forward curves that lay within the family? Motivated by this question, Björk and Christensen [10] define the so–called consistent pairs ($\mathcal{M}$, $\mathcal{G}$) as ones whose answer to the above question is positive. In particular, they studied the problem of consistency between the family of curves proposed by Nelson and Siegel [29] and any HJM interest rate model with deterministic volatility, obtaining that there is no such interest model consistent with it.

We remark that the Nelson and Siegel interpolating scheme is an important example of a parametric family of forward curves, because it is widely adopted by central banks (see for instance BIS [3]). Its forward curve shape, $G_{NS}(z, \cdot)$ is given by the expression

$$G_{NS}(z, x) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x},$$

where x denotes time to maturity and z the parameter vector

$$z = [z_1\ z_2\ \ldots]^T.$$

Despite all the positive empirical features and general acceptance by the financial community, Filipoviè [19] has shown that there is no Itô process that is consistent with the Nelson-Siegel family. In a recent study De Rossi [16] applies consistency results to propose a consistent exponential dynamic model, and estimates it using data on LIBOR and UK swap rates. On the other hand, Buraschi and Corielli [14] add results to theoretical framework indicating that the use of inconsistent parametric families to obtain smooth interest rate curves, violates the standard self financing arguments of replicating strategies, with direct consequences in risk management procedures.

In order to illustrate this situation, we describe a very common fixed-income market procedure. In the real world, practitioners usually re-estimate yield curve and HJM model parameters on a daily basis. This procedure consists of two steps:

- They fit the initial yield curve from discrete market data (bond prices, swap rates, short-term zero rates), and
- They obtain an estimate of the parameters of the HJM model, minimizing the pricing error of some actively traded (plain vanilla) interest rate derivatives (commonly swap options or caps).

In contrast with the parsimonious assumption that model parameters are constant, an unstable HJM model parameter estimation it is often observed. Perhaps, this fact is not relevant for mark to market, but it could have practical consequences on the hedging portfolios associated with these financial instruments. Recall that such dynamic strategies depend on the model assumptions. Thus, re-calibration is conceivable because the practitioners are aware of *model risk*. A particular HJM model is not a perfect description of reality, and they are forced to re-estimate day to day model parameters in order to include new information that arrives from the market. On the other hand, unstable estimates may be caused by reasons that are more theoretical, because the above mentioned set-up does not take into account that HJM model parameters are linked, in general, to the initial yield curve fit parameters. If a practitioner uses an interpolation scheme which is not consistent with the model, then the parameters will be artificially forced to change. Thus, it seems to be that there are a plethora of motivations for the study of the empirical evidence and the practical implications that are predicted by a consistent HJM build model.

4.2 Setup

We consider, as earlier in Chap. 2, a given forward rate model under a risk neutral martingale measure $\mathbb{Q}$. We will adopt the Musiela parameterization [28] and use the notation

$$f(t,x) := F(t,t+x).$$

The reasons why this parameterization is better suited to address the main problems that are the subject of the present work have been accurately explained, for instance, by Filipoviċ [18]. As we will see, this parametrization makes changes in the specification of the stochastic process originally proposed by HJM for the instantaneous forward rate process, $F(t,t+x)$, since now x is not considered any more a fixed variable which parametrizes the dynamics of these rates.

Proposition 6 (The Musiela HJM formulation.) *Under the martingale measure $\mathbb{Q}$ the f-dynamics are given by*

$$\begin{cases} df(t,x) &= \left(\frac{\partial f(t,x)}{\partial x} + \widetilde{\sigma}(t,x) \int_0^x \widetilde{\sigma}(t,u)^T \, du \right) dt + \widetilde{\sigma}(t,x) dW(t) \\ f(0,x) &= f^o(0,x). \end{cases} \tag{4.1}$$

where $\widetilde{\sigma}(t,x) := \sigma(t,t+x)$.

Proof. See Appendix B. □

From now, with a clear abuse of notation we remove the symbol $\widetilde{}$ from σ because in the whole chapter we will consider the HJM model under the Musiela parametrization. Thus the interest rate model $\mathcal{M}$ will be characterized by the particular volatility function $\sigma(t,x)$ used in the following f-dynamics:

$$df(t,x) = \left(\partial_x f(t,x) + \sigma(t,x)\int_0^x \sigma(t,u)^T\,du\right)dt + \sigma(t,x)dW(t). \tag{4.2}$$

4.3 The Formalized Problem

4.3.1 The Forward Curve Manifold

Assume that we have a parametrized family of forward rate curves

$$G : \mathcal{Z} \longrightarrow \mathcal{H}, \tag{4.3}$$

with $\mathcal{Z} \subseteq \mathbb{R}^d$ the parameter space. For each parameter value $z \in \mathcal{Z}$ we have a smooth curve $G(z)$. The value of this curve at the point $x \in \mathbb{R}_+$ will be written as $G(z,x)$, so we see that G can also be viewed as the mapping

$$G : \mathcal{Z} \times \mathbb{R}_+ \longrightarrow \mathbb{R}. \tag{4.4}$$

The main problem is to determine under which conditions the f-dynamics given by (4.2) is *consistent* with the parametrized family of forward rate curves (4.3) as follows:

- Assume that, at an arbitrarily chosen time $t = s$, we have fitted a forward curve G to market data, i.e. for some $z^o \in \mathcal{Z}$ we have

  $$f^o(s, s+x) = G(z^o, x), \quad \forall x \geq 0\ ,$$

- the *future* forward curves produced by the interest rate model (4.2) always stay within the given forward curve family? In other words, does there exist at every fixed time $t \geq s$ some $z \in \mathcal{Z}$ such that

  $$f(t, t+x) = G(z, x), \quad \forall x \geq 0\ ?$$

First, to see more clearly what is going on in differential geometric terms, we define the *forward curve manifold* $\mathcal{G}$, as the set of all forward curves produced by the parametrized family.

Definition 18 *The* **forward curve manifold** $\mathcal{G} \subseteq \mathcal{H}$ *is defined as*

$$\mathcal{G} = Im(G).$$

We now move on to give precise mathematical definition of the consistency property discussed above.

Definition 19 (Invariant manifold.) *Take as given the f-dynamics (4.2). Consider also the forward curve manifold $\mathcal{G}$. We say that $\mathcal{G}$ is invariant under the action of f if, for each point $(s, f) \in R_+ \times \mathcal{G}$, the condition $f_s \in \mathcal{G}$ implies that $f_t \in \mathcal{G}$ on a time interval $t - s > 0$.*

The purpose of the following section will be to characterize invariance in terms of local characteristics of both $\mathcal{G}$ and $\mathcal{M}$.

4.3.2 The Space

As the space of forward rate curves we will use a weighted Sobolev space where a generic point will be denoted by f.

Definition 20 *Consider a fixed real number $\gamma > 0$. The space $\mathcal{H}_\gamma$ is defined as the space of all differentiable (in the distributional sense) functions*

$$f : \mathbb{R}_+ \longrightarrow \mathbb{R}$$

satisfying the norm condition $\|f\|_\gamma < \infty$. Here the norm is defined as

$$\|f\|_\gamma^2 = \int_0^\infty f^2(x) e^{-\gamma x}\, dx + \int_0^\infty \left(\frac{df}{dx}(x)\right)^2 e^{-\gamma x}\, dx$$

Intuitively, as a specific Sobolev space, $\mathcal{H}_\gamma$ is a vector space of functions equipped with a norm that is a combination of L^2-norms of the function itself as well as its first derivative. Recall that x is the time to maturity, as defined in Chapter 1.

In fact, if we introduce the inner product

$$(f, g) = \int_0^\infty f(x) g(x) e^{-\gamma x}\, dx + \int_0^\infty \left(\frac{df}{dx}(x)\right)\left(\frac{dg}{dx}(x)\right) e^{-\gamma x}\, dx,$$

the space $\mathcal{H}_\gamma$ becomes a Hilbert space as proved by Björk and Landen [6].

4.3.3 The Interest Rate Model

Finally, let us consider as given a volatility function σ of the form

$$\sigma : \mathcal{H}_\gamma \times \mathbb{R}_+ \to \mathbb{R}^q.$$

$\sigma(f, x)$ is thus a functional of the infinite dimensional f-variable, and a function of the real variable x. Denoting the forward curve at time t by f_t we then have the following forward rate equation.

$$df_t(x) = \left\{ \frac{\partial}{\partial x} f_t(x) + \sigma(f_t, x) \int_0^x \sigma(f_t, u)^T \, du \right\} dt + \sigma(f_t, x) \, dW_t. \tag{4.5}$$

4.4 The Invariance Conditions

As we see before, the pair $(\mathcal{M}, \mathcal{G})$ is consistent if and only if the forward curve manifold $\mathcal{G}$ is invariant under the action f, and the question we pursue from now is when it happens. In order to guess the precise answer we have to rewrite the analysis in terms of Stratonovich integrals instead of Itô integrals.

Definition 21 *For given semimartingales X and Y driven by a multidimensional Wiener process, the* **Stratonovich integral** *of X w.r.t Y,*

$$\int_0^t X_s \circ dY(s) \, ,$$

is defined as

$$\int_0^t X_s \circ dY_s := \int_0^t X_s \cdot dY_s + \frac{1}{2} dX_t \cdot dY_t$$

Remark 3 *For computing the "quadratic variation process" $dX_t \cdot dY_t$ the usual "multiplication rules" $dW \cdot dt = dt \cdot dt$, $dW \cdot dW = dt$ must be applied.*

Proposition 7 (Chain rule.) *Assume that the function $F(t, y)$ is smooth. Then we have*

$$dF(t, Y_t) = \frac{\partial F}{\partial t}(t, Y_t) dt + \frac{\partial F}{\partial y}(t, Y_t) \circ dY_t.$$

Note that under the Stratonovich formulation of the stochastic integral, the Itô formula takes the form of the standard chain rule of ordinary calculus. Now, returning to the above f-dynamics (4.5), we can write it in terms of Stratonovich calculus as the following

$$df_t(x) = \left\{ \frac{\partial}{\partial x} f_t(x) + \sigma(f_t, x) \int_0^x \sigma(f_t, u)^T \, du \right\} dt - \frac{1}{2} d\sigma(f_t, x) \cdot dW_t + \sigma(f_t, x) \circ dW_t \; , \tag{4.6}$$

and, as it can be seen, it appears a *quadratic variation term* commonly known as the *Stratonovich correction*. Note that σ is not a function but a functional, however in practical terms we can work with the ordinary Itô formula which is still correct as in the finite dimensional case. Then,

$$\begin{aligned} d\sigma \cdot dW(t) &= \big(\{ \dots \} dt + \sigma'_f(f_t)\sigma(f_t)^T dW(t) \big) \cdot dW(t) \\ &= \sigma'_f(f_t)\sigma(f_t)^T dt. \end{aligned} \tag{4.7}$$

where σ'_f denotes the Frechet derivative of σ w.r.t the f-variable. This derivative extends the concept of Jacobian matrix to the infinite dimensional case. Its formal definition, which is somewhat technical, is left out. See [15].

Finally, we may write the Stratonovich formulation of the Musiela equation (4.6) as

$$df_t = \mu(f_t) + \sigma(f_t) \circ dW_t \tag{4.8}$$

where

$$\mu(f_t, x) = \partial_x f_t(x) + \sigma(f_t, x) \int_0^x \sigma(f_t, u)^T \, du - \frac{1}{2} \left[\sigma'_f(f_t)\sigma_f(f_t)^T \right](x). \tag{4.9}$$

Let us consider as given the forward curve manifold $\mathcal{G}$: the relevant concept is the following.

Definition 22 *Consider a given interest rate model $\mathcal{M}$, specifying a forward rate process $f_t(x)$, as well as a forward curve manifold $\mathcal{G}$. We say that $\mathcal{G}$ is f-invariant under the action of the forward rate process $f_t(x)$ if there exists a stochastic process $Z(t)$ with state space $\mathcal{Z}$ and possessing a differential of the form*

$$dZ(t) = \gamma(t, Z(t))dt + \psi(t, Z(t)) \circ dW_t, \tag{4.10}$$

such that, for every fixed choice of initial time s, whenever $y_s(\cdot) \in \mathcal{G}$, the stochastic process defined by

$$y_t(x) = G(Z(t), x), \; \forall t \geq s, \; x \geq 0, \tag{4.11}$$

solves the SDE (4.6) with initial condition $f_s(\cdot) = y_s(\cdot)$.

In fact, the stochastic Z-process is describing how evolves the vector parameter z as the forward rate curve moves on the manifold $\mathcal{G}$.

We assume that the forward rate Itô dynamics of $\mathcal{M}$ are given by (4.5), and that the quadratic variation process may be written in intensity form:

$$-\frac{1}{2}\left[\sigma'_f(f_t)\sigma(f_t)^T\right](x)dt = \phi(t,x)dt$$

Now we can state and prove the main invariance result.

Theorem 4 (Consistency Conditions.) *The forward curve manifold $\mathcal{G}$ is f invariant for the forward rate process $f(t,x)$ in $\mathcal{M}$ iff*

$$G_x(z,\cdot) + \sigma(t,\cdot)\int_0^{\cdot} \sigma(t,u)^T\,du + \phi(t,\cdot) \in Im\left[G_z(z,\cdot)\right], \tag{4.12}$$

$$\sigma(t,\cdot) \in Im\left[G_z(z,\cdot)\right]. \tag{4.13}$$

$\forall (t,z) \in \mathbb{R}_+ \times \mathcal{Z}$. Here, G_z and G_x denote the Jacobian of G w.r.t. to z and x, provided some minimal smoothness of the mapping G.

Proof. See Appendix B. □

Condition (B.8) is called *the consistent drift condition* (henceforth CDC) and condition (B.9) is called *the consistent volatility condition* (henceforth CVC). It is said that we have invariance if and only if the latter conditions hold which brings us the following definition.

Definition 23 (Consistency.) *We say that the interest rate model $\mathcal{M}$ is* **consistent** *with the forward rate manifold $\mathcal{G}$ if the CDC and CVC conditions (B.8)-(B.9) prevail.*

4.4.1 Simple Invariance

In order to obtain some geometric intuition, we will analyze a bidimensional deterministic version of our problem as a motivational example. Let us consider:

- first, a deterministic vector function $Q : \mathbb{R}_+ \to \mathbb{R}^2$

$$Q(t) = \begin{bmatrix} Q_1(t) \\ Q_2(t) \end{bmatrix}, \tag{4.14}$$

with differential given by

$$\dot{Q} = \mu(t, Q(t)), \tag{4.15}$$

where $\mu : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}^2$ is some smooth vector field.

- Next, a smooth mapping $G : \mathcal{Z} \to \mathbb{R}^2$,

$$G(z) = \begin{bmatrix} G_1(z) \\ G_2(z) \end{bmatrix}. \tag{4.16}$$

It could be interpreted that the process Q corresponds to two specific coordinates of the infinite-dimensional object $f(t, x)$ in a purely deterministic world. Say, for instance, picking out the 3-month and 10-year key rates at any time t. Thus, this is our toy model $\mathcal{M}$.

Now define the manifold $\mathcal{G}$ as

$$\mathcal{G} = \{G(z) : z \in \mathcal{Z}\}, \tag{4.17}$$

and assume that $Q(s) \in \mathcal{G}$ for some initial value $z^o \in \mathcal{Z}$, that is

$$q^o = Q(s) = G(Z(s) = z^o), \tag{4.18}$$

where Z is also another deterministic d-dimensional process $Z : \mathbb{R}_+ \to \mathbb{R}^d$. When the relation for the future times $t \geq s$,

$$Q(t) \in \mathcal{G}, \quad \forall\, t \geq s, \tag{4.19}$$

prevails? The answer is geometrically obvious. We have the relation (4.19) iff the velocity vector

$$\frac{dQ}{dt},$$

belongs to the tangent space $T_{Q(t)}(\mathcal{G})$ for each $t \geq s$. Note, that a generic point of $\mathcal{G}$ is written as $q = G(z)$, and the tangent space at this point is given as the span of the tangent vectors

$$\frac{\partial G(z)}{\partial z_i}, \quad i = 1, \ldots, d. \tag{4.20}$$

Let us return to the first order deterministic system written on normal form

$$\dot{Q} = \mu(t, Q(t)).$$

Physically, we may interpret this system as a condition set for the velocity of a particle at position q at time t. Indeed, geometrically, we may think in this context that $\mu : U \to \mathbb{R}^2$, where U is an open set $U \subset \mathbb{R}_+ \times \mathbb{R}^2$, is a velocity field defined over the plane $\mathbb{R}^2$. By definition, the function $Q : I \to \mathbb{R}^2$ is a solution of the aforementioned linear dynamical system whenever the relation

$$\dot{Q} = \mu(t, Q(t)), \quad \forall t \in I \subset \mathbb{R}_+$$

holds. That is, $Q(t)$ is a solution for the differential system iff the trajectory of the particle is tangent to the vector field μ in all its points. Thus, we have related the problem of finding the solution for the differential system with the geometric problem of finding tangent trajectories for the velocity field μ. In fact, the theoretical problem for the infinite-dimensional stochastic case is the same, and the use of the Stratonovich form for the differential is the convenience trick in order to bridge the gap between stochastic differential calculus and ordinary calculus for real variable functions. Consider for instance the concrete bidimensional system in $\mathbb{R}^2$:

$$\begin{cases} \dot{Q}(t) &= [-Q_2(t)\ Q_1(t)]^T \\ Q(0) &= [1\ 0]^T \end{cases} \tag{4.21}$$

For this system the unit circle manifold

$$\mathcal{S}^1 = \{G(z) = [\cos z\ \ \sin z]^T : z \in \mathbb{R}\}$$

is invariant. The deterministic system (4.21) has the exact solution

$$Q(t) = [\cos t\ \ \sin t]^T\ ,$$

and if we start the system on $\mathcal{S}^1$ it will stay forever on $\mathcal{S}^1$. In fact, the latter may be easily seen by introducing the trivial one-dimensional deterministic realization $Z(t) = t$, because making such a choice

$$Q(t) = G(Z(t)), \quad \forall t \geq s.$$

Now, it is easy to check the corresponding invariance condition first introduced

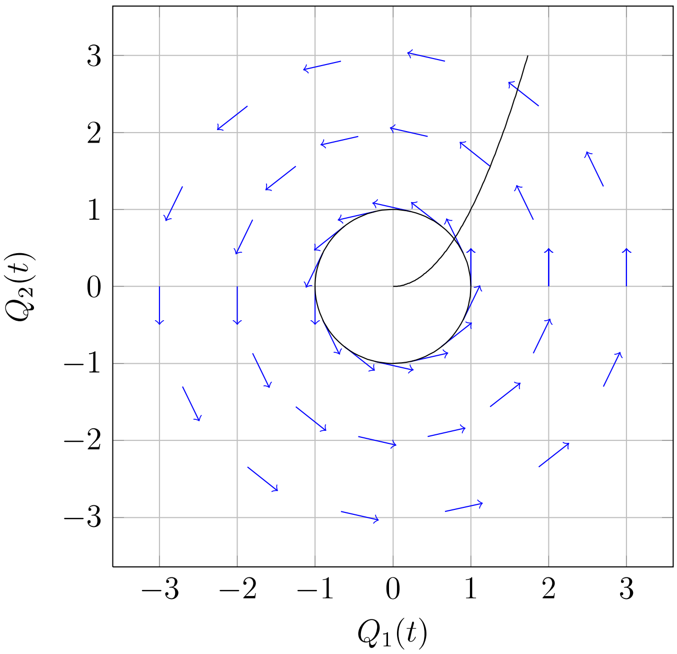

Figure 4.1: The vector field from the the system (4.21) with $\mathcal{S}^1$ and a parabollic noninvariant manifold.

in (4.20). Let $G_z(z)$ denote the Frechet derivative (which in turns is the Jacobian in this deterministic and finite-dimensional case) of G at z. The columns of the matrix representation of G_z are the tangent vectors above (4.20), so the tangent space $T_q(\mathcal{S}^1)$ to $\mathcal{S}^1$ at $q = G(z)$ coincides with the image $Im[G_z(z)]$:

$$G_z(z) = [-\sin(z)\ \ \cos(z)]^T.$$

Recall that the vector field μ at $q = G(z)$ is in turns:

$$\mu(G(z)) = [-\sin(z)\ \ \cos(z)]^T.$$

We thus have trivially

$$\mu(G(z)) \in Im[G_z(z)],$$

and, in fact, this is the consistency condition we have to check out for pairs ($\mathcal{M}$,$\mathcal{G}$) when $\mathcal{M}$ is an autonomous and deterministic differential system of finite dimension.