INTEREST RATE DYNAMICS AND CONSISTENT FORWARD RATE CURVES*

TOMAS BJÖRK†

Department of Finance, Stockholm School of Economics

Box 6501, S-113 83 Stockholm, Sweden

BENT JESPER CHRISTENSEN‡

School of Economics and Management, University of Aarhus

Bldg. 350, University Park, DK-8000 Aarhus C, Denmark

November 18 1997

We derive general necessary and sufficient conditions for the mutual consistency of a given parametrized family of forward rate curves and the dynamics of a given interest rate model. Consistency in this context means that the interest rate model will produce forward rate curves belonging to the parameterized family. The interest rate model may be driven by a multidimensional Wiener process as well as by a marked point process. As an application, the Nelson-Siegel family of forward curves is shown to be inconsistent with the Ho-Lee interest rate model, and with the Hull-White extension of the Vasiček model, but it may be adjusted to achieve consistency with these models and with extensions that allow for jumps in interest rates. For a natural exponential-polynomial generalization of the Nelson-Siegel family, we give necessary and sufficient conditions for the existence of a

*We are grateful for useful comments from the audiences in Aarhus, Hindsgavl (Odense), Oberwohlfach, Stockholm, Trento and Zürich. In particular, we have benefitted from comments by D. Filipović.

†Support from ITM is gratefully acknowledged.

‡Support from the Danish Science and Social Science Research Councils is gratefully acknowledged.

consistent interest rate model with deterministic volatility.

Key Words: forward rate curves, interest rate models, invariant manifolds, marked point processes

1. INTRODUCTION

A standard procedure when dealing with concrete interest rate models (e.g. those found in Heath, Jarrow, and Morton (1992), Ho and Lee (1986), or Hull and White (1990)) on a high frequency (say, daily) basis can be described as follows:

1. At time $t = 0$, use market data to fit (calibrate) the model to the observed bond prices.
2. Use the calibrated model to compute prices of various interest rate derivatives.
3. The following day ($t = 1$), repeat the procedure in 1. above in order to recalibrate the model, etc..

To carry out the calibration in step 1. above, the analyst typically has to produce a forward rate curve $\{f^\star(0,T); T \geq 0\}$ from the observed data. However, since only a finite number of bonds actually trade in the market, the data consist of a discrete set of points, and a need to fit a curve to these points arises. This curve-fitting may be done in a variety of ways. One way is to use splines, but also a number of parametrized families of smooth forward curves have become popular in applications—the most well-known probably being the Nelson-Siegel (1987) family. Once the curve $\{f^\star(0,T); T \geq 0\}$ has been obtained, the parameters of the interest rate model may be calibrated to this.

Now, from a purely logical point of view, the recalibration procedure in step 3. above is of course slightly nonsensical: If the interest rate model at hand is an exact picture of reality, then there should be no need to recalibrate. The reason that everyone insists on recalibrating is of course that any model in fact only is an approximate picture of the financial market under consideration, and recalibration allows incorporating newly arrived information in the approximation. Even so, the calibration procedure itself ought to take into account that it will be repeated. It appears that the optimal way to do so would involve a combination of time series and cross-section data, as opposed to the purely cross-sectional curve-fitting, where the information contained in previous curves is discarded in each recalibration.

The cross-sectional fitting of a forward curve and the repeated recalibration is thus, in a sense, a pragmatic and somewhat non-theoretical endeavour. Nonetheless, there are some nontrivial theoretical problems to be dealt with in this context. The problem to be studied in the present paper concerns the **consistency** between, on the one hand, the dynamics of a given interest rate model, and, on the other hand, the forward curve family employed.

What, then, is meant by consistency in this context? Assume that a given interest rate model $\mathcal{M}$, e.g. the Hull-White model, in fact *is* an exact picture of the financial market. Now consider a particular family $\mathcal{G}$ of forward rate curves, e.g. the Nelson- Siegel family, and assume that the interest rate model is calibrated using this family. We then say that the pair $(\mathcal{M}, \mathcal{G})$ is **consistent** (or, that $\mathcal{M}$ and $\mathcal{G}$ are consistent) if all forward curves which may be produced by the interest rate model $\mathcal{M}$ are contained within the family $\mathcal{G}$, provided that the initial forward curve is in $\mathcal{G}$. Otherwise, the pair $(\mathcal{M}, \mathcal{G})$ is inconsistent.

Thus, if $\mathcal{M}$ and $\mathcal{G}$ are consistent, then the interest rate model actually produces forward curves which belong to the relevant family. In contrast, if $\mathcal{M}$ and $\mathcal{G}$ are inconsistent, then the interest rate model will produce forward curves outside the family used in the calibration step, and this will force the analyst to change the model parameters all the time—not because the model is an approximation to reality, but simply because the family does not go well with the model.

Put into more operational terms this can be rephrased as follows.

- Suppose that you are using a fixed interest rate model $\mathcal{M}$. If you want to do recalibration, then your family $\mathcal{G}$ of forward rate curves should be chosen is such a way as to be consistent with the model $\mathcal{M}$.

Note however that the argument also can be run backwards, yielding the following conclusion for empirical work.

- Suppose that a particular forward curve family $\mathcal{G}$ has been observed to provide a good fit, on a day-to-day basis, in a particular bond market. Then this gives you modelling information about the choice of an interest rate model in the sense that you should try to use/construct an interest rate model which is consistent with the family $\mathcal{G}$.

We now have a number of natural problems to study.

I Given an interest rate model $\mathcal{M}$ and a family of forward curves $\mathcal{G}$, what are necessary and sufficient conditions for consistency?

II Take as given a specific family $\mathcal{G}$ of forward curves (e.g. the Nelson-Siegel family). Does there exist any interest rate model $\mathcal{M}$ which is consistent with $\mathcal{G}$?

III Take as given a specific interest rate model $\mathcal{M}$ (e.g. the Hull-White model). Does there exist any finitely parametrized family of forward curves $\mathcal{G}$ which is consistent with $\mathcal{M}$?

In the present paper we will mainly adress problem (I) above. Problems (II) and (III) will be solved in special cases. The harder problem of determining all models $\mathcal{M}$ consistent with a given family $\mathcal{G}$ will be solved under the additional restriction of deterministic volatility. The most general version of the companion problem of determining all $\mathcal{G}$ consistent with a given $\mathcal{M}$ is the topic of ongoing research.

The more detailed structure and basic results of our paper are as follows.

- In Section 2 we set the stage and formalize our basic consistency problem.

- In Section 3 we study the finite-dimensional version of the problem in order to gain some geometric intuition.

- The main results of the paper are presented in Section 4. We give necessary and sufficient conditions for consistency in the infinite-dimensional case (i.e. we solve problem (I) above) for Wiener-driven interest rate models.

- In Section 5 we illustrate the ideas by applying the theory to the case of the Nelson-Siegel family and the interest rate models of Ho-Lee and Hull-White. Not surprisingly, it turns out that neither model is consistent with Nelson-Siegel, but we solve problem (II) for a simpler family. We do obtain positive results for the two interest rate models—we solve problem (III) for both.

- The general theory is extended to include a driving multivariate point process, besides the Wiener processes. This is done in Section 6, where we solve problem (I) in this generalized setting, and we illustrate the ideas by studying some examples.

- In Section 7 we present a new and fairly flexible family of forward rate curves. The Nelson-Siegel family is contained as a special case, and we give a complete characterization of the class of linear interest rate models consistent with the new family.

2. FORMAL PROBLEM STATEMENT

As in Heath, Jarrow, and Morton (1992) (henceforth HJM), we consider a default free bond market with $p(t,T)$ denoting the price at time $t \geq 0$ of a zero coupon bond maturing at $T \geq t$. We assume frictionless markets, i.e. there are no transaction costs, taxes, or short sale restrictions, and the bonds are perfectly divisible (Jarrow (1988)). The forward rates $f(t,T)$ are defined as

$$f(t,T) = -\frac{\partial \log p(t,T)}{\partial T}\ , \tag{1}$$

and the short rate is $s(t) = f(t,t)$. Suppose that we are given a concrete term structure model, defined on a filtered probability space $(\Omega, \mathcal{F}, Q, \{\mathcal{F}_t\}_{t\geq 0})$, and that the model is free of arbitrage in the sense that the probability measure Q is a martingale measure. The physical probability measure P will play no role below, so all calculations are carried out under Q.

More specifically, we assume that under the martingale measure Q, the dynamics of the forward rates are of the form

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t), \tag{2}$$

where W is an m-dimensional Q-Wiener process, and α and σ are the drift and volatility functions. Thus, for each maturity date T, the evolution across calendar time t (where $t \leq T$) of the forward rate $f(t,T)$ is governed by this stochastic differential equation.

We furthermore assume that we are given a parametrized family of forward rate curves

$$G : \mathcal{Z} \to C[0,\infty), \tag{3}$$

with $\mathcal{Z} \subseteq R^d$ the parameter space, i.e. for each parameter value $z \in \mathcal{Z}$ we have a smooth curve $G(z)$. By slight abuse of notation we will sometimes write the curve as

$$x \longmapsto G(x; z), \tag{4}$$

where the variable x is interpreted as the time **to** maturity, as opposed to the time **of** maturity T, i.e. $x = T - t$. The main problem is to determine under which conditions the interest rate model (2) is **consistent** with the parametrized family of forward curves (3), in the following sense:

- Assume that, at an arbitrarily chosen time $t = s$, we have fitted a forward curve G to market data. Technically, this means that we have specified an initial forward curve, i.e. for some $z_0 \in \mathcal{Z}$ we have

  $$f^*(s, s + x) = G(x; z_0), \quad \forall x \geq 0. \tag{5}$$

- Is it then the case that the the subsequent forward curves produced by the interest rate model (2) always stay within the given forward curve family, i.e. does there at every fixed time $t \geq s$ exist some $z \in \mathcal{Z}$ such that

  $$f(t, t + x) = G(x; z), \tag{6}$$

  $\forall x \geq 0$? Here, z may depend on t and on the elementary outcome $\omega \in \Omega$.

Remark 2.1 Since we want to consider fairly general interest rate models, including those who are not time homogeneous, we are forced to consider an arbitrary initial time $t = s$, rather than just $t = 0$.

Remark 2.2 Note that we take the volatility structure $\sigma(t, T)$ as given. This structure is then kept fixed, so the calibration procedure only concerns the choice of initial forward rate curve. Put into measure theoretic terms this means that, by fixing the volatility structure, we have fixed a class of equivalent martingale measures, and the calibration step will then pin down a particular member of this class. In many practical calibration schemes you also update volatility parameters, but this leads to questions outside the scope of the present paper.

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 2.1 (Forward curve manifold). *Given a smooth mapping* $G : \mathcal{Z} \to C[0,\infty)$, *we define the* **forward curve manifold** $\mathcal{G}$ *by*

$$\mathcal{G} = Im\ G\ , \tag{7}$$

i.e.

$$\mathcal{G} = \{G(\cdot; z) \in C[0,\infty) :\ z \in \mathcal{Z}\}\ . \tag{8}$$

As we will see below it will be convenient to use the Musiela (1993) (see also Brace-Musiela (1994)) parameterization of forward rates

$$r(t,x) = f(t, t+x), \tag{9}$$

where the symbol x again denotes time **to** maturity as opposed to T, which denotes time **of** maturity. We denote the induced dynamics for the r-process by

$$dr(t,x) = \beta(t,x)dt + \sigma_0(t,x)dW, \tag{10}$$

and it is easy to see that there is a one-to-one correspondence between the formulations (2) and (10), namely

$$\beta(t,x) = \frac{\partial}{\partial x} r(t,x) + \alpha(t, t+x), \tag{11}$$

$$\sigma_0(t,x) = \sigma(t, t+x). \tag{12}$$

The Heath-Jarrow-Morton drift condition (see HJM) can now be transferred to the Musiela parameterization case. The result from Musiela (1993) is as follows.

Proposition 2.1 (Interest rate model). *Under the martingale measure* Q*, the* r*-dynamics must be of the form*

$$dr(t,x) = \left\{ \frac{\partial}{\partial x} r(t,x) + \sigma_0(t,x) \int_0^x \sigma_0(t,u)' du \right\} dt + \sigma_0(t,x) dW(t)\ . \tag{13}$$

Thus the interest rate model $\mathcal{M}$ is completely characterized by the volatility function $\sigma_0(t,x)$. With this in hand, we will use the following definition.

DEFINITION 2.2 (Invariant manifold). *Let a forward curve manifold $\mathcal{G}$ and a forward rate process $r(t,x)$ be given. We say that $\mathcal{G}$ is* **invariant under the action of** r *if, for every fixed initial time s, the condition $r(s,\cdot) \in \mathcal{G}$ implies that $r(t,\cdot) \in \mathcal{G}$, $\forall t \geq s$, Q-a.s..*

We may now phrase our main problem in the following way.

- Suppose that we are given (i) a forward rate model $\mathcal{M}$, specifying a process r as in (13); and (ii) a forward curve manifold $\mathcal{G}$.
- Is $\mathcal{G}$ then invariant under the action of r?

Thus, the pair $(\mathcal{M}, \mathcal{G})$ is consistent if and only if the manifold $\mathcal{G}$ is invariant under the action of r, and the question we pursue is when this happens.

3. THE FINITE-DIMENSIONAL CASE

In order to gain some geometric intuition, we will start out by analyzing the finite-dimensional version of our problem. Since the entire section is purely motivational, the reasoning is slightly sketchy, and from a logical point of view the entire section can be skipped.

3.1. The Finite-Dimensional Deterministic Case

We start with the deterministic case. We take as given:

- A deterministic n-dimensional 'process' $Y : R_+ \to R^n$,

$$Y(t) = \begin{bmatrix} Y_1(t) \\ Y_2(t) \\ \vdots \\ Y_n(t) \end{bmatrix}, \tag{14}$$

with differential given by

$$\frac{dY}{dt} = \mu(t, Y(t)), \tag{15}$$

where $\mu : R_+ \times R^n \to R^n$ is some smooth vector field.

- A smooth mapping $G : \mathcal{Z} \to R^n$,

$$G(z) = \begin{bmatrix} G_1(z) \\ G_2(z) \\ \vdots \\ G_n(z) \end{bmatrix}. \tag{16}$$

The interpretation is that the forward rate process r corresponds to the process Y, each entry in $Y(t)$ containing a certain specified x-coordinate of the infinite-dimensional vector $r(t,x)$ (say, picking out a certain key rate at time t). Thus, this is the interest rate model $\mathcal{M}$.

In the same spirit as above, we now define the manifold $\mathcal{G}$ as

$$\mathcal{G} = \{G(z) : \; z \in \mathcal{Z}\}. \tag{17}$$

We assume that $Y(s) \in \mathcal{G}$, i.e. for some $z_0 \in \mathcal{Z}$, the initial term structure is given by $G(z_0)$, and we wish to determine when the following relation holds Q-a.s.

$$Y(t) \in \mathcal{G}, \quad \forall t \geq s. \tag{18}$$

The answer is geometrically obvious. We have the relation (18) if and only if the velocity vector $\frac{dY}{dt}$ belongs to the tangent space $T_{Y(t)}(\mathcal{G})$ for each $t \geq s$.

A generic point of $\mathcal{G}$ is written as $y = G(z)$, and the tangent space at this point is given as the span of the tangent vectors

$$\frac{\partial G(z)}{\partial z_i}, \; i = 1, \ldots, d. \tag{19}$$

Let $DG(z)$ denote the Frechet derivative (Jacobian) of G at z, i.e., the columns of the matrix representation of DG are the tangent vectors above, so the tangent space $T_y(\mathcal{G})$ to $\mathcal{G}$ at $y = G(z)$ coincides with the image $Im[DG(z)]$.

We thus have 18 if and only if

$$\mu(t, G(z)) \in Im[DG(z)] \tag{20}$$

for all $t \geq s$ and all z such that $G(z) = Y(t)$. Since (18) is to hold for all initial times s, and for all initial points $Y(s)$, the relation (20) must in fact hold for all t and all z, thus giving us the following result.

Proposition 3.1 (Simple invariance). *The manifold $\mathcal{G}$ is invariant under the action of Y if and only if*

$$\mu(t, G(z)) \in Im[DG(z)] \ , \tag{21}$$

for each $z \in \mathcal{Z}$ and each $t \geq 0$.

Thus, the pair $(\mathcal{M}, \mathcal{G})$ is consistent under the indicated condition.

3.2. The Finite-Dimensional Controlled Case

Now we go one step further and consider a controlled system Y of the form

$$\frac{dY}{dt} = \mu(t, Y(t)) + \sigma(t, Y(t))u(t) \ , \tag{22}$$

where the deterministic control u acts as input to the system, with $u(t) \in R^m$ and σ an $n \times m$-matrix, and we are allowed to choose the control freely among some (large) class of R^m-valued functions $u(\cdot)$. This system plays the role of $\mathcal{M}$ and is considered in conjunction with a concrete manifold $\mathcal{G}$. Suppose that $Y(s) \in \mathcal{G}$ for some arbitrarily fixed initial time s. The question is now when the Y-trajectory stays on the manifold $\mathcal{G}$ for all $t \geq s$, regardless the choice of control signal u. Problems of this kind ('reachability problems') are standard in control theory, and the answer is again intuitively obvious: $\mathcal{G}$ is invariant if and only if the velocity vector of Y is in the tangent space of $\mathcal{G}$, regardless the choice of u. We thus have:

Proposition 3.2 (Controlled invariance). *The manifold $\mathcal{G}$ is invariant under the action of Y if and only if*

$$\mu(t, G(z)) + \sigma(t, G(z))u \in Im[DG(z)] \ , \quad \forall u \in R^m \ , \tag{23}$$

for each $z \in \mathcal{Z}$ and each $t \geq 0$.

Thus, $\mathcal{M}$ is consistent with $\mathcal{G}$ if and only if at every point of $\mathcal{G}$ all directions spanned by the system parameters μ and σ are contained within the tangent space to $\mathcal{G}$ at that point. Otherwise, the term structure trajectory is bound to leave the specified family of forward curves.

It is easily seen that (23) is equivalent to the conditions

$$\mu(t, G(z)) \in Im[DG(z)] \ , \tag{24}$$

$$\sigma(t, G(z)) \in Im[DG(z)] \ , \tag{25}$$

for all $t \in R_+$ and all $z \in \mathcal{Z}$. Here the condition (25) is to be interpreted in the sense that each column of the matrix σ belongs to $Im[DG(z)]$. For notational simplicity we shall throughout write this requirement as in (25).

We note that another equivalent representation of the condition (23) is that there exist mappings $\gamma : R_+ \times \mathcal{Z} \to R^d$ and $\psi : R_+ \times \mathcal{Z} \to R^{d\times m}$ (the space of $d \times m$-matrices) such that $\mu(t, G(z)) = DG(z)\gamma(t, z)$ and $\sigma(t, G(z)) = DG(z)\psi(t, z)$, for every $(t, z) \in R_+ \times \mathcal{Z}$.

3.3. The Finite-Dimensional Stochastic Case

Now we consider the case where the n-dimensional process Y possesses a stochastic differential of the form

$$dY(t) = \mu(t, Y(t))dt + \sigma(t, Y(t))dW(t) \ , \tag{26}$$

with W an m-dimensional Q-Wiener process. The manifold $\mathcal{G}$ is defined as before, and we assume that $Y(s) \in \mathcal{G}$. The question is again when Y stays on $\mathcal{G}$ with Q-probability one.

To build some geometric intuition we rewrite equation (26) somewhat informally as

$$\frac{dY}{dt} = \mu(t, Y(t)) + \sigma(t, Y(t))\dot{W}(t), \tag{27}$$

thinking of $\dot{W}$ as 'white noise.' Technically, of course, $\dot{W}$ does not exist, but the idea gives rise to nice heuristics. The interpretation is that someone (the God of Chance, or perhaps your worst enemy) 'chooses' the 'input signal' $\dot{W}$. A natural guess, inspired by the results in the preceding subsection, is now that $\mathcal{G}$ is invariant if and only if

$$\mu(t, G(z)) + \sigma(t, G(z))u \in Im[DG(z)] \ , \quad \forall u \in R^m \ , \tag{28}$$

for every $(t, z) \in R_+ \times \mathcal{Z}$. Unfortunately, this 'result' is wrong, since it neglects the discrepancy between ordinary differential calculus and the Itô calculus for continuous time stochastic processes. Nonetheless, the geometric heuristics above are not all bad, provided that we bridge this gap. This is done by rewriting the analysis in terms of Stratonovich integrals (see, e.g., Karatzas and Shreve (1988)) instead of the usual Itô integrals.[1]

[1]Note that the stochastic differentials considered so far only have precise meaning in terms of their Itô integral representations.

DEFINITION 3.3 (Stratonovich integral). *For given semimartingales X and Y, the* **Stratonovich integral** *of X with respect to Y, $\int_0^t X(s) \circ dY(s)$, is defined as*

$$\int_0^t X(s) \circ dY(s) = \int_0^t X(s)dY(s) + \frac{1}{2}\langle X, Y \rangle_t \ . \tag{29}$$

The first term on the RHS is the Itô integral. In the present case, with only Wiener processes as driving noise, we can define the 'quadratic variation process' $\langle X, Y \rangle$ in (29) by

$$\langle X, Y \rangle_t = \int_0^t dX(s)dY(s), \tag{30}$$

with the usual 'multiplication rules' $dW \cdot dt = dt \cdot dt = 0, \ dW \cdot dW = dt$. We now recall the main result and *raison d'être* for the Stratonovich integral.

PROPOSITION 3.3 (Chain rule). *Assume that the function $F(t, y)$ is smooth. Then we have*

$$dF(t, Y(t)) = \frac{\partial F}{\partial t}(t, Y(t))dt + \frac{\partial F}{\partial y} \circ dY(t) \ . \tag{31}$$

Thus, in the Stratonovich calculus, the Itô formula takes the form of the standard chain rule of ordinary calculus. In view of the above arguments, it is now easy to guess our main result concerning invariance in the finite-dimensional case.

PROPOSITION 3.4 (Stochastic invariance). *Assume that the process Y possesses a Stratonovich differential given by*

$$dY(t) = \mu(t, Y(t))dt + \sigma(t, Y(t)) \circ dW(t). \tag{32}$$

Then the manifold $\mathcal{G}$ is invariant for Y if and only if there exist mappings $\gamma : R_+ \times \mathcal{Z} \to R^d$ and $\psi : R_+ \times \mathcal{Z} \to R^{d \times m}$ such that

$$\mu(t, G(z)) = DG(z)\gamma(t, z) \ , \tag{33}$$

$$\sigma(t, G(z)) = DG(z)\psi(t, z) \ , \tag{34}$$

for each $t \geq 0$ *and* $z \in \mathcal{G}$.

We omit the proof of the proposition since it is a special case of the result for the infinite-dimensional case treated below. Again, as in (24)-(25) of the previous subsection, (33)-(34) may be replaced by the conditions

$$\mu(t, G(z)) \in Im[DG(z)] \ , \tag{35}$$

$$\sigma(t, G(z)) \in Im[DG(z)] \ , \tag{36}$$

where as before (36) is interpreted columnwise for $\sigma(t, G(z))$. Thus, the consistency condition for the pair $(\mathcal{M}, \mathcal{G})$ indeed coincides with the conjecture (28), provided that the stochastic differential equation for Y is interpreted in the Stratonovich sense.

4. INVARIANT FORWARD CURVE MODELS

We now return to our original problem, i.e. the infinite-dimensional stochastic case. Assume that, by the Musiela parametrization of the HJM no arbitrage drift condition, the Itô dynamics for the forward rates are given by

$$dr(t,x) = \left\{ \frac{\partial}{\partial x} r(t,x) + \sigma_0(t,x) \int_0^x \sigma_0(t,u)' du \right\} dt + \sigma_0(t,x) dW(t) \tag{37}$$

under the martingale measure Q. This implies that the Stratonovich dynamics are given by

$$\begin{aligned} dr(t,x) &= \left\{ \frac{\partial}{\partial x} r(t,x) + \sigma_0(t,x) \int_0^x \sigma_0(t,u)' du \right\} dt \\ &\quad - \frac{1}{2} d\langle \sigma_0(\cdot,x), W \rangle (t) + \sigma_0(t,x) \circ dW(t) \end{aligned} \tag{38}$$

under Q. In the infinite-dimensional stochastic case, this is the model $\mathcal{M}$. In fact, in many cases the quadratic variation process may be written in intensity form as

$$-\frac{1}{2} d\langle \sigma_0(\cdot,x), W \rangle (t) = \varphi(t,x) dt. \tag{39}$$

In particular, in the (rather common) case of a deterministic volatility function $\sigma_0(t,x)$, the Stratonovich formulation coincides with the Itô formulation, viz.

$$dr(t,x) = \left\{ \frac{\partial}{\partial x} r(t,x) + \sigma_0(t,x) \int_0^x \sigma_0(t,u)' du \right\} dt + \sigma_0(t,x) \circ dW(t) \ . \tag{40}$$

We again take as given the forward curve manifold $\mathcal{G}$, and the idea is now to copy the finite-dimensional reasoning from the preceding section. Of course, for the proofs to go through, we need to impose some amount of regularity. Thus, in the infinite-dimensional stochastic case, the relevant invariance concept is the following.

Definition 4.1 (r-invariance). *Consider a given interest rate model $\mathcal{M}$, specifying a forward rate process $r(t,x)$, as well as a forward curve manifold $\mathcal{G}$. We say that $\mathcal{G}$ is* **r-invariant** *under the action of the forward rate process*

$r(t,x)$ if there exists a stochastic process Z with state space $\mathcal{Z}$ and possessing a Stratonovich differential of the form

$$dZ(t) = \gamma(t, Z(t))dt + \psi(t, Z(t)) \circ dW(t), \tag{41}$$

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

$$y(t,x) = G\left(x; Z(t)\right), \quad \forall t \geq s, \; x \geq 0, \tag{42}$$

solves the SDE (38) with initial condition $r(s,\cdot) = y(s,\cdot)$.

Remark 4.1 The interpretation is that the Z-process describes the evolution of the z-parameters as the forward rate curve moves on the manifold $\mathcal{G}$. It is obvious that r-invariance implies invariance, and our conjecture is that invariance plus some minimal smoothness of the mapping G implies r-invariance.

We may now state and prove our main invariance result. We assume that the forward rate Itô dynamics of $\mathcal{M}$ are given by (37), and that the quadratic variation process has the structure (39). We also assume that (41) has a solution that is unique in distribution, since in this case any solution y, say, to (38) in the form (42) is unique in distribution, too, and hence y may be used in place of r, say, for derivative pricing and so forth.

Theorem 4.1 (**Main Theorem**). *The forward curve manifold $\mathcal{G}$ is r-invariant for the forward rate process $r(t,x)$ in $\mathcal{M}$ if and only if*

$$G_x(\cdot; z) + \sigma_0(t,\cdot)\int_0^{\cdot} \sigma_0(t,u)'du + \varphi(t,\cdot) \;\in\; Im[G_z(\cdot;z)]\,, \tag{43}$$

$$\sigma_0(t,\cdot) \;\in\; Im[G_z(\cdot;z)]\,, \tag{44}$$

for all $(t,z) \in R_+ \times \mathcal{Z}$.

Here, G_z and G_x denote the Frechet derivatives of G with respect to z and x, respectively, which are assumed to exist. As before, the condition (44) is interpreted columnwise for σ_0. Condition (43) is called *the consistent drift condition*, and (44) is called *the consistent volatility condition.*

Proof. To prove necessity we assume r-invariance. Then we may take the differential in equation (42), producing

$$dy(t,x) = G_z(x; Z(t))\gamma(t, Z(t))dt + G_z(x; Z(t))\psi(t, Z(t)) \circ dW(t). \tag{45}$$

Comparing this with equation (38) and equating coefficients yields, (43)-(44) for every $t \geq s$, but with $z = Z(t)$. Since the initial time s as well as the initial point $r(s, \cdot)$, and thus $Z(s)$, can be chosen arbitrarily, we finally obtain (43)-(44) in full generality.

To prove sufficiency, assume (43)-(44). Then we may select $\gamma : R_+ \times \mathcal{Z} \to R^d$ and $\psi : R_+ \times \mathcal{Z} \to R^{d \times m}$ satisfying

$$
\begin{aligned}
&(46) \qquad & G_x(\cdot, z) + \sigma_0(t, \cdot) \int_0^{\cdot} \sigma_0(t, s)' ds + \varphi(t, \cdot) &= G_z(\cdot, z)\gamma(t, z) \ , \\
&(47) \qquad & \sigma_0(t, \cdot) &= G_z(\cdot, z)\psi(t, z) \ ,
\end{aligned}
$$

for all $(t, z) \in R_+ \times \mathcal{Z}$. Let $y(s, \cdot) \in \mathcal{G}$, i.e. $y(s, \cdot) = G(\cdot; z_0)$, for some $z_0 \in \mathcal{Z}$. Define Z as the solution to (41) with initial condition $Z(s) = z_0$, and define the infinite-dimensional process $y(x, t)$ by $y(t, x) = G(x; Z(t))$. Then

$$
\begin{aligned}
dy(t, x) &= G_z(x; Z(t))\gamma(t, Z(t))dt + G_z(x; Z(t))\psi(t, Z(t)) \circ dW(t) \\
&= \left\{ G_x(x; Z(t)) + \sigma_0(t, x) \int_0^x \sigma_0(t, s)' ds + \varphi(t, x) \right\} dt + \sigma_0(t, x) \circ dW(t) \\
&= \left\{ \frac{\partial}{\partial x} y(t, x) + \sigma_0(t, x) \int_0^x \sigma_0(t, s)' ds + \varphi(t, x) \right\} dt + \sigma_0(t, x) \circ dW(t).
\end{aligned}
$$

Thus, y solves the SDE (38). □

The moral is that, subject to regularity conditions, we have invariance if and only if conditions (43)-(44) hold. This leads us to the following definition.

Definition 4.2 (Consistency). *We say that the interest rate model $\mathcal{M}$ is* **consistent** *with the forward rate manifold $\mathcal{G}$ if the consistent drift and volatility conditions (43)-(44) hold.*

For all practical purposes, this implies r-consistency, as well, since the additional regularity condition on the coefficients γ and ψ usually is satisfied in the models of interest in finance.

Remark 4.2 It is easily seen that if the family G is invariant under shift in the x-variable, then we will automatically have the relation

$$G_x(\cdot, z) \in Im[G_z(\cdot; z)],$$

so in this case the relation (43) can be replaced by

$$\sigma_0(t, \cdot) \int_0^{\cdot} \sigma_0(t, u)' du + \varphi(t, \cdot) \in Im[G_z(\cdot; z)].$$

Thus, we have solved problem (I) from the Introduction for Wiener-driven interest rate models. In addition, we have along the way identified the relevant partial differential equations that hold the key to the solution of problems (II) and (III). We solve these problems in specific cases in the following section.

5. APPLICATIONS

In order to illustrate the ideas we now move from abstract theory to the investigation of a number of concrete forward curve families and interest rate models. The leading example is the popular forward curve family introduced by Nelson and Siegel (1987) (henceforth NS), considered in 5.1. We analyze the consistency of this family, and several variations of it, with the Ho-Lee interest rate model, in 5.2, and with the Hull-White model, in 5.3.

5.1. The Nelson-Siegel Family

The NS forward curve manifold $\mathcal{G}$ is parametrized by $z \in \mathcal{Z} = R^4$, the curve shape G being given by the well-known expression

$$G(x; z) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x} \; . \tag{48}$$

Frequently, it is the associated functional form for the yield curve (obtained by integration) that is used in the empirical curve-fitting procedure, but the original justification in NS (based on Laguerre functions) was explicitly in terms of the forward rate curve. Essentially, G specifies a parsimonious, yet flexible functional form. For $z_4 \neq 0$ the Frechet derivatives are easily obtained as

$$G_z(x; z) = \left[1,\; e^{-z_4 x},\; x e^{-z_4 x},\; -(z_2 + z_3 x) x e^{-z_4 x}\right] \; , \tag{49}$$

$$G_x(x; z) = (z_3 - z_2 z_4 - z_3 z_4 x) e^{-z_4 x} \; . \tag{50}$$

For the degenerate case $z_4 = 0$ (which we return to below) we have

$$G(x; z) = z_1 + z_2 + z_3 x \; , \tag{51}$$

so z_2 is redundant, i.e. we may set $z_2 = 0$. Thus, the two-dimensional degenerate family takes the form

$$G(x; z) = z_1 + z_3 x \; , \tag{52}$$

and we find

$$G_z(x;z) = [1, x] \ , \tag{53}$$

$$G_x(x;z) = z_3 \ . \tag{54}$$

Henceforth, we write $\mathcal{Z}_{NS} = \{(z_1, z_2, z_3, z_4) : z_4 \neq 0\}$ for the non-degenerate NS parameter space and $\mathcal{G}_{NS} = G(\mathcal{Z}_{NS})$ for the associated manifold, whereas $\mathcal{Z}_0 = \{z \in \mathcal{Z} : z_2 = z_4 = 0\}$ and $\mathcal{G}_0 = G(\mathcal{Z}_0)$ indicate the corresponding objects in the degenerate case.

5.2. The Ho-Lee Model

As a laboratory case, we will now study the interest rate model of Ho and Lee (1986) (henceforth HL) and see if it is consistent with the NS manifold. In its short rate formulation, the continuous-time limit of the HL model takes the form

$$ds(t) = \Phi(t)dt + \sigma dW(t), \tag{55}$$

where the volatility $\sigma > 0$ is constant, and the drift function Φ is calibrated to obtain a perfect fit between the observed and the theoretical forward rate curve at $t = 0$. The exact expression is

$$\Phi(t) = \sigma^2 \cdot t + \frac{\partial f^*}{\partial T}(0, t) \ , \tag{56}$$

where $\{f^*(0, T) : T \geq 0\}$ has been obtained from bond prices using some form of curve-fitting procedure—say, Nelson-Siegel. The HJM formulation of HL is

$$df(t, T) = \alpha(t, T)dt + \sigma dW(t) \ , \tag{57}$$

with the same σ as above. Since the volatility is common across maturities, the BM parametrization (12) is obtained by setting $\sigma_0(t, x) = \sigma$. Further, since this is deterministic, we have $\langle \sigma, W \rangle = 0$, i.e. $\varphi = 0$ in (43). Thus, by Theorem 4.1, consistency of the HL model $\mathcal{M}$ and any given manifold $\mathcal{G}$ requires that the consistent drift and volatility conditions be met, viz.

$$\begin{aligned} G_x(\cdot, z) + \sigma^2 x \ &\in \ Im[G_z(\cdot, z)], && (58) \\ \sigma \ &\in \ Im[G_z(\cdot, z)]. && (59) \end{aligned}$$

The interpretation of the consistent volatility condition requires no special convention in the present situation, since $m = 1$. It is easily seen that the relation (59) is satisfied, so it remains to check the relation (58). For the specific case of the Nelson-Siegel manifold $\mathcal{G}_{NS}$, we must investigate whether there for each fixed choice of

$z = (z_1, \ldots, z_4)$ exist constants A, B, C, D (which may depend on z) such that

$$\begin{aligned} & [z_3 - z_2 z_4 - z_3 z_4 x]\, e^{-z_4 x} + \sigma^2 x \\ = \; & A + Be^{-z_4 x} + Cxe^{-z_4 x} - D(z_2 + z_3 x) x e^{-z_4 x} \ , \end{aligned} \tag{60}$$

for all $x \geq 0$. Because of the term $\sigma^2 x$ we see that (60) cannot possibly hold unless $z_4 = 0$. Thus, the HL model is inconsistent with the non-degenerate NS manifold $\mathcal{G}_{NS}$. Some NS initial forward rate curves are such that under the action of the HL dynamics, the subsequent forward curves escape from the NS manifold.

It remains to examine whether HL is consistent with the degenerate NS manifold $\mathcal{G}_0$ (see 52). In this case, the relation (59) is satisfied. In order to check (60) we now fix z_1 and z_3, and we look for constants A and B such that

$$z_3 + \sigma^2 x = A + Bx \ , \quad \forall x \geq 0 \ . \tag{61}$$

This holds with $A = z_3$ and $B = \sigma^2$, so we have proved the following result.

Proposition 5.1 (Nelson-Siegel and Ho-Lee).

(a) *The full Nelson-Siegel family is inconsistent with the Ho-Lee interest rate model.*

(b) *The degenerate Nelson-Siegel family $G(x; z) = z_1 + z_3 x$ is in fact consistent with Ho-Lee.*

It follows that if the initial term structure is affine, then all subsequent term structures remain affine under Ho-Lee dynamics. Hence, we have solved problem (III) from the Introduction for HL, i.e. we have determined a consistent forward curve manifold for this model. Of course, we have at the same time solved problem (II), that of finding a consistent interest rate model, for the class of affine forward curves.

5.3. The Hull-White Model

As our next test case we analyze the following model, studied by Hull and White (1990) (henceforth HW):

$$ds(t) = \{\Phi(t) - as(t)\}\, dt + \sigma dW(t)\ , \tag{62}$$

where a, $\sigma > 0$. The HW model generalizes the HL model to allow for mean reversion (at rate a), and as in HL, the drift function Φ is calibrated from the observed initial forward rate curve $\{f^*(0, T) : T \geq 0\}$—where the latter in turn has been obtained from bond prices using, say, Nelson-Siegel. The HW model also generalizes the Vasičec (1977) model, where Φ is assumed constant. The HJM forward rate formulation is

$$df(t, T) = \alpha(t, T)dt + \sigma e^{-a(T-t)} dW(t). \tag{63}$$

We thus have

$$\sigma_0(t, x) = \sigma e^{-ax}, \tag{64}$$

and the conditions of Theorem 4.1 become

$$G_x(x, z) + \frac{\sigma^2}{a}\left[e^{-ax} - e^{-2ax}\right] \ \in\ Im[G_z(x, z)], \tag{65}$$

$$\sigma e^{-ax} \ \in\ Im[G_z(x, z)]. \tag{66}$$

To investigate whether $\mathcal{G}_{NS}$ is invariant under HW dynamics, we start with (66) and as usual fix a z-vector. We then look for constants A, B, C, and D, such that for all $x \geq 0$ we have

$$\sigma e^{-ax} = A + Be^{-z_4 x} + Cxe^{-z_4 x} - D(z_2 + z_3 x)xe^{-z_4 x}. \tag{67}$$

This is possible if and only if $z_4 = a$, so we immediately see that HW is inconsistent with the full Nelson-Siegel manifold.

One might of course ask whether the HW model for a fixed choice of a is consistent with the three-dimensional restricted NS manifold defined in terms of the curves

$$G(x; z) = z_1 + z_2 e^{-ax} + z_3 xe^{-ax}. \tag{68}$$

More precisely, writing $\mathcal{Z}_a = \{z \in \mathcal{Z}_{NS} : z_4 = a\}$, the candidate manifold is $\mathcal{G}_a = G(\mathcal{Z}_a)$. In this case, the Frechet derivatives are given by

$$G_z(x; z) = \left[1,\ e^{-ax},\ xe^{-ax}\right] , \tag{69}$$

$$G_x(x; z) = [z_3 - az_2 - az_3 x]\, e^{-ax} . \tag{70}$$

For this submanifold, the relation (66) is by definition satisfied. In order to check (65) we fix $z \in \mathcal{Z}_a$ and seek constants $A,\ B$, and C, such that

$$[z_3 - az_2 - az_3 x]\, e^{-ax} + \frac{\sigma^2}{a}\left[e^{-ax} - e^{-2ax}\right] = A + Be^{-ax} + Cxe^{-ax}, \tag{71}$$

$\forall x \geq 0$. Since there are two different exponents on the left hand side, this relation can never hold unless $a = 0$. In this case the HW model collapses into the HL model, which we have already investigated, so we have proved the following.

Proposition 5.2 (Nelson-Siegel and Hull-White). *The Hull-White model is inconsistent with the Nelson-Siegel family.*

We have thus obtained a negative result for the HW model. The NS manifold is 'too small' for HW, in the sense that if the initial forward rate curve is on the NS manifold, then the HW dynamics will force the term structure off the manifold within an arbitrarily short period of time.

A more positive approach is now to ask whether it is possible to extend the NS family (as opposed to restricting it, as in the cases $\mathcal{G}_a$ and $\mathcal{G}_0$) in such a way that it becomes consistent with the HW model. This is in fact possible, and the calculations above immediately tell us how. First, we must again require that $z_4 = a$. Secondly, to avoid that the term structure escapes, we must expand the manifold by introducing an exponential of the form e^{-2ax}. Thus, define an augmented NS manifold $\tilde{\mathcal{G}}_a$ by the curve shape

$$G(x; z) = z_1 + z_2 e^{-ax} + z_3 x e^{-ax} + z_4 e^{-2ax}, \tag{72}$$

where the constant a in (72) is the same as in (62). Working through the by now standard procedure we arrive at the following result.

Proposition 5.3 (Augmented Nelson-Siegel and Hull-White). *The augmented Nelson-Siegel family (72) is consistent with the Hull-White model.*

REMARK 5.1 Note that the extended Nelson-Siegel model above is not the samllest possible family which is consistent with Hull-White. The minimal consistent family is in fact given by

$$G(x; z) = z_1 e^{-ax} + z_2 e^{-2ax}$$

Thus, we have solved problem (III) from the Introduction for the HW model. That is, we have identified a forward curve manifold consistent with HW. In addition, the analysis is indicative of the procedure that may be used for other concrete interest rate models. Sometimes $\mathcal{G}$ must be restricted since otherwise the tangent space is too rich, tangent vectors sometimes pointing in directions not spanned by the drift and volatility functions that define the given model dynamics; and sometimes $\mathcal{G}$ must be extended in order for the tangent space to contain these functions, thus avoiding that the model dynamics drive the term structure off the manifold.

6. MODELS WITH JUMPS

In this section we generalize the theory to cover interest rate models which are driven jointly by Wiener processes (as in the previous sections) and a general marked point process. We first solve problem (I) from the Introduction, then for a specific case problems (II) and (III).

6.1. The general case

The model $\mathcal{M}$ now specifies that the forward rate dynamics under a martingale measure Q be given as

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t) + \int_E \delta(t,y,T)m(dt,dy)\ , \tag{73}$$

where $m(dt,dy)$ is a marked point process with measurable mark space $(E,\mathcal{E})$ and compensator $\nu(dt,dy)$. We assume that $\nu([0,t]\times E) < \infty$, Q-a.s., for all finite t, i.e. m is a multivariate point process in the terminology of Jacod and Shiryaev (1987), and δ is the jump size. For simplicity, we assume that the compensator may be cast in intensity form, viz., $\nu(dt,dy) = \lambda(t,dy)dt$. Furthermore, we assume that W and m are independent and that the filtration under consideration is the one generated by these two processes.

Under these assumptions, Björk, Kabanov, and Runggaldier (1997) (henceforth BKR) show the following result on the relationship between α, σ, δ and λ, generalizing the HJM drift condition, as well as the corresponding results in Musiela (1993) and Brace-Musiela (1994).

Proposition 6.1 (The model with jumps). *Given the forward rate dynamics (73) under Q, we must have the following relations:*

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)'ds - \int_E \delta(t,y,T)e^{\Delta(t,y,T)}\lambda(t,dy), \tag{74}$$

where

$$\Delta(t,y,T) = -\int_t^T \delta(t,y,s)ds. \tag{75}$$

Furthermore, the r-dynamics are given by

$$\begin{aligned} dr(t,x) &= \left\{\frac{\partial}{\partial x}r(t,x) + \sigma_0(t,x)\int_0^x \sigma_0(t,s)'ds - \int_E \delta_0(t,y,x)e^{\Delta_0(t,y,x)}\lambda(t,dy)\right\}dt \\ &+ \sigma_0(t,x)dW(t) + \int_E \delta_0(t,y,x)m(dt,dy)\ , \end{aligned}$$

where

$$\begin{aligned}\sigma_0(t,x) &= \sigma(t,t+x),\\ \delta_0(t,y,x) &= \delta(t,y,t+x),\\ \Delta_0(t,y,x) &= \Delta(t,y,t+x).\end{aligned}$$

Suppose now that we, in addition to the forward rate model $\mathcal{M}$ above, also are given a forward curve manifold $\mathcal{G}$. The question is when $\mathcal{G}$ is invariant under the action of r. The answer in this generalized framework is that firstly, the manifold $\mathcal{G}$ must be invariant under the continuous action, i.e. under the dt and dW dynamics of $\mathcal{M}$. This invariance may be accounted for by appeal to our earlier results. Secondly, $\mathcal{G}$ must be invariant under the jump-action. At a jump-time t, the forward rate $r(t,x)$ exhibits a jump of size $\delta_0(t,y,x)$, given that the jump carries the mark $y \in E$. Since $\mathcal{G}$ must be invariant under jumps, regardless the mark of the jump, we have the following.

THEOREM 6.1 (Main Theorem with jumps). *Suppose that the forward rate dynamics are given as in Proposition 6.1, and that the quadratic variation process has the structure (39). Then the conclusion of Theorem 4.1 carries over to the jump case, provided the conditions (43)-(44) are replaced with*

$$\begin{aligned}G_x(\cdot;z) + \sigma_0(t,\cdot)\int_0^{\cdot}\sigma_0(t,s)'ds \quad &+ \quad \varphi(t,\cdot)\\ -\int_E \delta_0(t,y,\cdot)e^{\Delta_0(t,y,\cdot)}\lambda(t,dy) \quad &\in \quad Im[G_z(\cdot;z)]\,, \qquad (76)\\ \sigma_0(t,\cdot) \quad &\in \quad Im[G_z(\cdot;z)]\,, \qquad (77)\\ G(\cdot;z) + \delta_0(t,y,\cdot) \quad &\in \quad \mathcal{G}, \quad \forall t \geq 0,\ \forall y \in E. \qquad (78)\end{aligned}$$

Thus, we have also solved problem (I) from the Introduction in the presence of jump processes.

6.2. The jump-extended Ho-Lee model

As an example, we will now study the simplest possible interest rate model with jumps. It is a straightforward generalization of the Ho-Lee model for the short rate $s(t)$, and the Q-dynamics are given as

$$ds(t) = \Phi(t)dt + \sigma dW(t) + \delta dN(t) \ , \tag{79}$$

where W is a Wiener process and N is a Poisson process with intensity λ under Q. It may be shown (see Björk (1995), BKR, and Shirakawa (1991)) that the forward rate formulation of this model is given by

$$df(t,T) = \alpha(t,T) + \sigma dW(t) + \delta dN(t) \ , \tag{80}$$

so $\sigma_0 = \sigma$ and $\delta_0 = \delta$. To determine whether this model $\mathcal{M}$ is consistent with the NS manifold $\mathcal{G}_{NS}$, our results show that it suffices to investigate the conditions (76)-(78), which in the present case (by slight abuse of notation) read as

$$G_x(x,z) + \sigma^2 x - \delta e^{-\delta x}\lambda \ \in \ Im[G_z(x;z)] \ , \tag{81}$$

$$\sigma \ \in \ Im[G_z(x;z)] \ , \tag{82}$$

$$G(\cdot,z) + \delta \ \in \ \mathcal{G} \ , \tag{83}$$

for all $(t,z) \in R_+ \times \mathcal{Z}$. Condition (82) is clearly satisfied (it has not changed by the introduction of jumps). Since $\mathcal{G}_{NS}$ is translation invariant (z_1 is exactly a translation parameter), we also see that condition (83) is met. In order to check (81) we fix z and seek constants $A,\ B,\ C$, and D such that

$$\begin{aligned} & [z_3 - z_2 z_4 - z_3 z_4 x]\, e^{-z_4 x} + \sigma^2 x - \delta e^{-\delta x} \\ = \ & A + Be^{-z_4 x} + Cxe^{-z_4 x} - D(z_2 + z_3 x)xe^{-z_4 x} \ . \end{aligned} \tag{84}$$

This requires that $z_4 = \delta$, but also (due to the term $\sigma^2 x$) that $z_4 = 0$. Thus, the jump-extended HL model is consistent with $\mathcal{G}_{NS}$ if and only if (i) $\delta = 0$, i.e. there are no jumps, so we are back in the original HL model, and (ii) attention is restricted to the degenerate NS manifold $\mathcal{G}_0$. No non-trivial jump-component is permitted.

Thus, the results so far are negative in the case of point processes. One might ask whether the NS family can be expanded to be consistent with the point process generalization of HL in the same manner that we were able to introduce the augmented NS manifold $\tilde{\mathcal{G}}_a$ in order to capture the HW model. This can in fact be done, and we have the following result.

Proposition 6.2 (Modified Nelson-Siegel and Ho-Lee with jumps). *The following modified Nelson-Siegel family $\tilde{\mathcal{G}}_\delta$ is consistent with the jump-extended Ho-Lee model:*

$$G(x; z) = z_1 + z_2 x + z_3 e^{-\delta x}$$

Thus, we have solved problem (III) for the jump-extended HL model, and by definition therefore also problem (II) for the modified NS family. The modified forward curve family $\tilde{\mathcal{G}}_\delta$ presupposes knowledge of the jump size δ, just like the relevant manifold $\tilde{\mathcal{G}}_a$ for the HW model requires knowledge of the rate of mean reversion a.

We have seen that we can handle interest rate models with jumps. In ongoing research, we consider the extension to non-deterministic volatility functions, such as the Cox-Ingersoll-Ross model.

7. A NEW FORWARD CURVE MANIFOLD

In this section we will introduce and study a particular forward curve family of exponential-polynomial type. This family seems to be fairly natural from an applied point of view, and it contains the Nelson-Siegel family as a special case.

7.1. The full $EP(K, n)$ family

We restrict ourselves to purely Wiener driven forward rate dynamics, where furthermore the volatility $\sigma_0(x)$ is assumed to be a **deterministic** function of x, only. Thus, the r-dynamics are of the form

$$dr(t, x) = \left\{ \frac{\partial}{\partial x} r(t, x) + \sigma_0(x) \int_0^x \sigma_0(u)' du \right\} dt + \sigma_0(x) dW(t), \tag{85}$$

and we see that r is an infinite-dimensional Gaussian process.

In order to specify the forward curve manifold, let us fix a positive integer K, and a vector $n = (n_1, ..., n_K)$ having non-negative integers as its components.

Definition 7.1 (Exponential-polynomial family). *The forward curve manifold $EP(K, n)$ is defined as the set of all curves of the form*

$$G(x) = \sum_{i=1}^{K} p_i(x) e^{-\alpha_i x}, \tag{86}$$

where $\alpha_i \in R$ for all i, and where p_i is any polynomial with $deg(p_i) \leq n_i$ for all i.

If we write the polynomial p_i as

$$p_i(x) = \sum_{j=0}^{n_i} z_{ij} x^j, \tag{87}$$

we see that a particular polynomial p_i is determined by its (n_i+1)-dimensional vector of coefficients $z_i = (z_{i0}, \ldots, z_{in_i})$. In terms of the previous sections, the family $EP(K, n)$ is thus defined as the mapping

$$G : R^{|n|+K} \times R^K \to C[0, \infty), \tag{88}$$

where $|n| = \sum n_i$, and $G(x; z, \alpha)$ is given by the RHS of (86), with $\alpha = (\alpha_1, \ldots, \alpha_K)$, and $z = (z_1, \ldots, z_K)$. We see that the dimension of the manifold is given by $|n| + 2K$.

The various partial derivatives of the mapping G are easily obtained as

$$\frac{\partial G}{\partial z_{ij}} = x^j e^{-\alpha_i x}, \tag{89}$$

$$\frac{\partial G}{\partial \alpha_i} = -x p_i(x) e^{-\alpha_i x}, \tag{90}$$

$$\frac{\partial G}{\partial x} = \sum_{i=1}^{K} (p_i'(x) - \alpha_i p_i(x)) e^{-\alpha_i x}. \tag{91}$$

From (89)-(90) we obtain the tangent space of $\mathcal{G}$.

LEMMA 7.1. *Fix (z, α) as above. Then we have*

$$g(\cdot) \in Im\,[G_{z,\alpha}(\cdot; z, \alpha)] \tag{92}$$

if and only if g is of the form

$$g(x) = \sum_{i=1}^{K} q_i(x) e^{-\alpha_i x} \tag{93}$$

where q_i is an arbitrary polynomial with $deg(q_i) \leq n_i + 1$.

We now turn to the question of whether there exists any forward rate model of the form (85) which is consistent with the $EP(K,n)$ family—this is problem (II) from the Introduction. Given our experiences from the Nelson-Siegel family it is not surprising that the answer is in the negative.

PROPOSITION 7.1 (Inconsistency). *No non-trivial forward rate model of the form (85) is consistent with the forward curve family $EP(K,n)$.*

Proof. Consider any volatility function σ_0. Combining the consistent volatility condition of Theorem 4.1 with Lemma 7.1 we see that, for every fixed α (and z), we must have the relation

$$\sigma_0(x) = \sum_{i=1}^{K} q_i(x) e^{-\alpha_i x}, \tag{94}$$

where the choice of the polynomials $q_1, ..., q_K$ may depend on α. Since the functions $e^{-\alpha_i x}$ are independent over the ring of polynomials, this can only be satisfied for all choices of α if $\sigma_0 \equiv 0$. □

It is obvious from the proof of Proposition 7.1 that to guarantee the existence of a consistent forward rate model, we must keep some or all exponents fixed in the $EP(K,n)$ family. In fact, we have already in Section 5 seen such restrictions at work. Thus, the NS family $\mathcal{G}_{NS}$ is the case where $K = 2$ and $n = (0,1)$, i.e. $EP(2,(0,1))$, but with a zero restriction on the leading exponent, and the degenerate NS family $\mathcal{G}_0$ is the case $K = n = 1$, i.e. $EP(1,1)$, and again with $\alpha_1 = 0$ imposed. Such families, where some or all of the exponents $\alpha_1, \ldots, \alpha_K$ are kept fixed, will be termed *restricted families*, whereas the models without restrictions will be referred to as the *full families*.

Since the second exponent is unrestricted in $\mathcal{G}_{NS}$, the method of proof in Proposition 7.1 produces the following.

COROLLARY 7.1. *No non-trivial forward rate model with deterministic volatility structure is consistent with Nelson-Siegel.*

Of course, this actually implies some of the inconsistency results of Section 5, but what we did in that section was in addition to achieve consistency in specific cases by modifying the relevant manifolds.

In view of the above, we now turn to the restricted families.

7.2. The $REP(K,n;\beta)$ family

We consider a restricted exponential-polynomial forward curve family where in fact all exponents are fixed, and in order to emphasize this we will denote these fixed exponents by the new symbol β.

DEFINITION 7.2 (Restricted family). *Consider a fixed choice of* (K, n) *and a vector* $\beta = (\beta_1, ..., \beta_K) \in R^K$. *The restricted forward curve manifold* $REP(K, n, \beta)$ *is defined as the set of all curves of the form*

$$G(x) = \sum_{i=1}^{K} p_i(x) e^{-\beta_i x} , \tag{95}$$

where p_i *is any polynomial with* $deg(p_i) \leq n_i$ *for all* i.

For the restricted manifold we thus have the functionality

$$G : R^{|n|+K} \to C[0, \infty), \tag{96}$$

and the dimension of the manifold is now $|n| + K$.

Although restricted, this class is still very general and retains the property that any continuous term structure may be approximated arbitrarily closely on any compact interval by suitable choice of (K, n, β) and the parameters z. If we can find a consistent interest rate model for every $REP(K, n, \beta)$, then this will allow unforeseen flexibility in term structure modelling.

We have already in Section 5 seen several special instances of the $REP(K, n, \beta)$ at work. Thus, the degenerate NS manifold $\mathcal{G}_0$ consistent with the Ho-Lee model is $REP(1, 1, 0)$, the restricted NS manifold $\mathcal{G}_a$ coincides with $REP(2, (0, 1), (0, a))$, the augmented version $\tilde{\mathcal{G}}_a$ consistent with the Hull-White model is reckognized as $REP(3, (0, 1, 0), (0, a, 2a))$, and the extended family $\tilde{\mathcal{G}}_\delta$ consistent with the jump-extended Ho-Lee model is $REP(2, (1, 0), (0, \delta))$.

Using (89) and (91), we immediately have the following result.

LEMMA 7.2. *Given* z, *we have*

$$g(\cdot) \in Im \left[G_z(\cdot; z) \right] \tag{97}$$

if and only if g *is of the form*

$$g(x) = \sum_{i=1}^{K} q_i(x) e^{-\beta_i x} , \tag{98}$$

where q_i is an arbitrary polynomial with $deg(q_i) \leq n_i$. Furthermore, for every z we have the relation

$$G_x(\cdot; z) \in Im\left[G_z(\cdot; z)\right] \ . \tag{99}$$

When investigating the consistency of a pair $(\mathcal{M}, \mathcal{G})$, with $\mathcal{G}$ of the form $REP(K, n, \beta)$, the latter result implies a simplification when checking the consistent drift condition.

We now turn to the question of the existence of a consistent choice of σ_0. It turns out that the behavior of the restricted manifold depends on whether all exponents are nonzero or not. Therefore, we introduce the notational convention that if one β_i equals zero then it is given the index zero. Thus, for every $i \neq 0$ we have $\beta_i \neq 0$, whereas by definition $\beta_0 = 0$. We will treat the cases with and without a purely polynomial part separately. Note that all instances of the NS manifold fall within the first category (the 'zero-beta' case).

7.2.1. The $REP(K, n, \beta)$ manifold with no polynomial part

Assume now that we have fixed the family $REP(K, n, \beta)$, where all β_i are different from zero. Our quest is for a volatility function σ_0 for a one-factor model ($m = 1$) consistent with the given family. Then, from Theorem 4.1 and Lemma 7.2, we know that (upon permutation of indices) σ_0 must be of the form

$$\sigma_0(x) = \sum_{i=1}^{L} \hat{p}_i(x) e^{-\beta_i x} \ , \tag{100}$$

where $L \leq K$, and $\hat{p}_1, ..., \hat{p}_L$ are fixed polynomials with $deg(\hat{p}_i) \leq n_i$ for all i. Any σ_0 of this form will obviously satisfy the second condition of Theorem 4.1, so by (91), we only have to test the condition

$$\sigma_0(\cdot) \int_0^{\cdot} \sigma_0(s)' ds \in Im\left[G_z(\cdot; z)\right] \ . \tag{101}$$

Integration by parts shows that the integral above takes the form

$$\int_0^x \sigma_0(s) ds = \sum_{i=1}^{L} p_i^{\star}(x) e^{-\beta_i x} \ , \tag{102}$$

where $deg(p_i^\star) = deg(\hat{p}_i)$. Thus, we obtain

$$\sigma_0(x) \int_0^x \sigma_0(s)' ds = \sum_{i,j=1}^{L} \hat{p}_i(x) p_j^\star(x) e^{-(\beta_i+\beta_j)x}, \tag{103}$$

with $deg(\hat{p}_i p_j^\star) = deg(\hat{p}_i) + deg(\hat{p}_j)$. The main result now follows at once.

PROPOSITION 7.2 (Consistency without polynomial). *Consider a fixed manifold $REP(K, n, \beta)$ with no purely polynomial part. The volatility function $\sigma_0(x)$ is consistent with this manifold if and only if the following conditions hold.*

1. *The volatility function must be of the form*

$$\sigma_0(x) = \sum_{i=1}^{L} \hat{p}_i(x) e^{-\beta_i x}, \tag{104}$$

   with $L \le K$ and $deg(\hat{p}_i) \le n_i$ for all $i = 1, \dots, L$.

2. *For all $i, j \in \{1, \dots, L\}$ there exists an index $k \in \{1, \dots, K\}$ such that*

$$\begin{aligned} \beta_i + \beta_j &= \beta_k \ , \\ deg(\hat{p}_i) + deg(\hat{p}_j) &\le n_k \ . \end{aligned}$$

We also have the following easily testable criterion.

COROLLARY 7.2. *Under the assumptions of Proposition 7.2, there exists a consistent volatility function σ_0 if and only if there exist two indices $i, k \le K$ such that*

$$2\beta_i = \beta_k \ .$$

If this condition is met, then any σ_0 of the form $\sigma_0(x) = \hat{p}_i(x) e^{-\beta_i x}$, with $deg(\hat{p}_i) \le n_i$ and $2 \cdot deg(\hat{p}_i) \le n_k$, is consistent.

7.2.3. The 'zero-beta' $REP(K, n, \beta)$ manifold

We now consider the case where the manifold $REP(K, n, \beta)$ possesses a nontrivial purely polynomial part, i.e. when G takes the form

$$G(x) = p_0(x) + \sum_{i=1}^{K} p_i(x) e^{-\beta_i x}. \tag{105}$$

Recall that the variations $\mathcal{G}_0$, $\mathcal{G}_a$, $\tilde{\mathcal{G}}_a$ and $\tilde{\mathcal{G}}_\delta$ of the NS manifold all fall in this category. As in the previous subsection we see that σ_0 satisfies the consistent volatility condition if and only if it takes the form

$$\sigma_0(x) = \hat{p}_0(x) + \sum_{i=1}^{L} \hat{p}_i(x) e^{-\beta_i x} \ , \tag{106}$$

where $\hat{p}_i$ is a polynomial with $deg(\hat{p}_i) \leq n_i$ for all i. Again, it is easy to see that we have the relation $G_x(\cdot; z) \in Im\, [G_z(\cdot; z)]$, so to test the consistent volatility condition, it remains to verify

$$\sigma_0(\cdot) \int_0^{\cdot} \sigma_0(s)' ds \in Im\, [G_z(\cdot; z)] \ . \tag{107}$$

As before, we get

$$\sigma_0(x) \int_0^x \sigma_0(s)' ds = \sum_{i,j=0}^{L} \hat{p}_i(x) p_j^\star(x) e^{-(\beta_i + \beta_j)x}, \tag{108}$$

where $deg(p_j^\star) = deg(\hat{p}_j)$ for all $j \neq 0$, whereas $deg(p_0^\star) = deg(\hat{p}_0) + 1$ (if $\hat{p}_0 \neq 0$). Using these relations we have the following characterization.

Proposition 7.3 ('Zero-beta' consistency). *Consider a fixed family $REP(K, n, \beta)$ with a non-trivial purely polynomial part. The volatility function $\sigma_0(x)$ is consistent with this family if and only if the following conditions hold.*

1 *The volatility function must be of the form*

$$\sigma_0(x) = \hat{p}_0(x) + \sum_{i=1}^{L} \hat{p}_i(x) e^{-\beta_i x},$$

where $deg(\hat{p}_i) \leq n_i$ for all $i = 1, \ldots, L$.

2 *For all $i, j \in \{1, \ldots, L\}$ there exists an index $k \in \{1, \ldots, K\}$ such that*

$$\begin{aligned} \beta_i + \beta_j &= \beta_k \ , \\ deg(\hat{p}_i) + deg(\hat{p}_j) &\leq n_k \ . \end{aligned}$$

If $\hat{p}_0 = 0$ these are all the conditions. If $\hat{p}_0 \neq 0$ the following conditions must be added.

3 *For every $i \in \{1, \ldots, L\}$ it holds that*

$$deg(\hat{p}_i) + deg(\hat{p}_0) + 1 \le n_i \ .$$

4 *For $i = 0$ we have*

$$2 \cdot deg(\hat{p}_0) + 1 \le n_0 \ .$$

In this case, we have the following simple test for the existence of a consistent volatility function σ_0.

COROLLARY 7.3. *Under the assumptions of Proposition 7.3, there exists a consistent volatility function $\sigma_0(x)$ if and only if either (i) there exist some $i, j \in \{1, ..., K\}$ such that*

$$2 \cdot \beta_i = \beta_j, \tag{109}$$

or (ii) we have

$$n_0 \ge 1. \tag{110}$$

We remark that if the first conditions of the corollary is satisfied, then, for any choice of the constant $c > 0$, a consistent σ_0 is given by $\sigma_0(x) = c \cdot e^{-\beta_i x}$, and if the second condition is satisfied, $\sigma_0(x) = c$ is consistent. The last case is of course the Ho-Lee model.

Let us illustrate the application of Proposition 7.3. The degenerate NS manifold $\mathcal{G}_0$ is $REP(0, 1, 0)$ under the convention $\beta_0 = 0$. By the first condition of the proposition, σ_0 must be a polynomial $\hat{p}_0$. Since $n_0 = 1$, the last condition implies that $deg(\hat{p}_0) = 0$. We have proved:

COROLLARY 7.4. *Ho-Lee is the only model with deterministic volatility consistent with affine forward rate curves.*

As a second illustration, consider the augmented NS manifold $\tilde{\mathcal{G}}_a$, i.e. $REP(2, (0, 1, 0), (0, a, 2a))$. By the first condition of Proposition 7.3, any consistent volatility function must be of the form

$$\sigma_0(x) = \hat{z}_0 + (\hat{z}_1 + \hat{z}_2 x)e^{-ax} + \hat{z}_3 e^{-2ax} \ .$$

From the next condition, $\hat{z}_3 = 0$, since otherwise with $\beta_i = 2a$ we cannot find k such that $2\beta_i = \beta_k$. By the same condition, $\hat{z}_2 = 0$, since otherwise

$2deg(\hat{p}_1) = 2 > n_1$ (note that $n_1 = 1$). Thus, we have $deg(\hat{p}_1) = 0$. All the conditions are now satisfied if we let $\hat{z}_0 = 0$. In this case $\sigma_0(x) = \hat{z}_1 e^{-ax}$. On the other hand, if $\hat{z}_0$ (and thus $\hat{p}_0$) does not vanish, then the last condition must be checked, and since $n_0 = 0$ it yields a contradiction. We have proved:

COROLLARY 7.5. *Hull-White is the only model with deterministic volatility consistent with the augmented Nelson-Siegel forward curve family.*

In Propositions 5.1.b and 5.3, we proved the existence of consistent interest rate models for the degenerate and augmented Nelson-Siegel manifolds. Using the tools of the present section, we have tightened these results to uniqueness within the deterministic volatility class.

7.3. The Heath-Jarrow-Morton Model

So far in this section, we have considered one-factor ($m = 1$) models. We now show how our findings for the HL and HW models in Section 5 may be easily combined within the restricted exponential-polynomial framework to an analysis of the 'encompassing' two-factor model that naturally nests both these simple one-factor models. Thus, it is natural to combine the forward rate models (57) and (63) by specifying

$$df(t,T) = \alpha(t,T)dt + \sigma_1 dW_1(t) + \sigma_2 e^{-a(T-t)} dW_2(t) \ , \tag{111}$$

where W_1 and W_2 are independent Wiener processes. This is the two-factor example studied by HJM, and for this reason we refer to (111) as the HJM model—to be distinguished, of course, from the HJM forward rate framework which runs through the entire theory.

The analysis in Section 5 suggests that a forward curve manifold consistent with the HJM model $\mathcal{M}$ must combine the characteristics of the degenerate and augmented NS manifolds $\mathcal{G}_0$ and $\tilde{\mathcal{G}}_a$, much in the same manner that HJM encompasses both HL and HW. Thus, define the linear-exponential manifold $\hat{\mathcal{G}}_a$ to be the set of curves

$$G(x;z) = z_1 + z_2 x + z_3 e^{-ax} + z_4 x e^{-ax} + z_5 e^{-2ax} \ . \tag{112}$$

The consistent volatility condition in this case requires the existence of $A,\ B,\ C,\ D$, and E such that

$$\left\{ \sigma_1,\ \sigma_2 e^{-ax} \right\} = A + Bx + Ce^{-ax} + Dxe^{-ax} + Ee^{-2ax} \ , \tag{113}$$

but again, $A, \ldots, E$ may be different for the two functions on the left hand side. The consistent drift condition requires that

$$(z_4 - z_3 a - z_4 a x)e^{-ax} - 2z_5 a e^{-2ax} + \sigma_1^2 x + \frac{\sigma_2^2}{a}(e^{-ax} - e^{-2ax}) \in Im[G_z(\cdot; z)] ,$$

i.e. the left hand side must be represented in the same form as the right hand side of (113), but possibly for different constants $A, \ldots, E$. Analyzing the conditions as earlier, and drawing on our previous findings, we get the following.

PROPOSITION 7.4 (HJM and the linear-exponential family). *The HJM model is inconsistent with the Nelson-Siegel family (and with both the degenerate and augmented variations of this). However, the HJM model (as well as both HL and HW) is consistent with the linear-exponential family $\hat{\mathcal{G}}_a$ in (112).*

Finally, note that the theory has several empirical applications. For example, suppose that NS curves have been observed to produce good fits on a day-to-day basis in a particular bond market. By Propositions 5.1, 5.2 and 7.4, this would be evidence against the HL, HW and HJM models. Such evidence is largely ignored in the current literature on the estimation of continuous-time models for the short rate, using time series data on yields with fixed (and short) term to maturity, and discarding the information content of the long segment. In other markets, the relation between time series and cross-sectional information is clearly reckognized—e.g. the implied-realized volatility relation for foreign exchange options (Jorion, 1995) and stock index options (Christensen and Prabhala, 1997). In ongoing work, we employ the paradigms of the present paper as the basis for empirical research on the term structure of interest rates.

REFERENCES

BJÖRK, T. (1995): "On the term structure of discontinuous interest rates." *Obozrenie Prikladnoi i Promyshlennoi Matematiki*, **2**, 627-657. (In Russian. To be translated in *Surveys in Industrial and Applied Mathematics*).

BJÖRK, T. & DI MASI, G. & KABANOV, Y. & RUNGGALDIER, W. (1997): "Towards a general theory of bond markets." *Finance and Stochastics,* **1**, 141-174.

BJÖRK, T. & KABANOV, Y. & RUNGGALDIER, W. (1997): "Bond market structure in the presence of marked point processes." *Mathematical Finance,* **7**, 211-239.

BRACE, A. & MUSIELA, M. (1994): "A multi factor Gauss Markov implementation of Heath Jarrow and Morton." *Mathematical Finance*, **4**, 563-576.

BRÉMAUD, P. (1981): *Point Processes and Queues: Martingale Dynamics.* Berlin: Springer-Verlag.

CHRISTENSEN, B.J. & PRABHALA, N.R. (1997): "The Relation Between Implied and Realized Volatility." *Journal of Financial Economics*, forthcoming.

HEATH, D. & JARROW, R. & MORTON, A. (1992): "Bond pricing and the term structure of interest rates." *Econometrica*, **60**, 77-106.

HO, T. & LEE, S. (1986): "Term structure movements and pricing interest rate contingent claims." *Journal of Finance*, **41**, 1011-1029.

HULL, J & WHITE, A. (1990): "Pricing interest-rate-derivative securities." *The Review of Financial Studies*, **3**, 573-592.

JACOD, J. & SHIRYAEV A.N. (1987): *Limit Theorems for Stochastic Processes.* Berlin: Springer-Verlag.

JARROW, R. (1988): *Finance Theory.* Englewood Cliffs, New Jersey: Prentice-Hall.

JORION, P. (1995): "Predicting Volatility in the Foreign Exchange Market." *Journal of Finance*, **50**, 507-528.

KARATZAS, I. & SHREVE, S.E. (1988): *Brownian Motion and Stochastic Calculus.* Berlin: Springer-Verlag.

MUSIELA, M. (1993): "Stochastic PDEs and term structure models." Preprint, University of New South Wales.

NELSON, C. & SIEGEL, A. (1987): "Parsimonious modeling of yield curves." *Journal of Business*, **60**, 473-489.

SHIRAKAWA, H. (1991): "Interest rate option pricing with Poisson–Gaussian forward rate curve processes." *Mathematical Finance*, **1**, 77-94.

VASIČEK, O. (1977): "An equilibrium characterization of the term structure." *Journal of Financial Economics*, **5**, 177-188.