

Applications of Information Geometry to Interest Rate Theory¹

Dorje C. Brody* and Lane P. Hughston[†]

**Blackett Laboratory, Imperial College, London SW7 2BZ, UK
and Centre for Mathematical Science, Cambridge University,
Wilberforce Road, Cambridge CB3 0WA, UK*

[†]*Department of Mathematics, King's College London,
The Strand, London WC2R 2LS, UK*

Abstract. There is an information geometry associated with the theory of interest rates. This arises from the fact that the system of smooth yield curves is isomorphic to the convex space of density functions on the positive real line. The arbitrage-free interest rate dynamics of Heath, Jarrow & Morton (1992) can be represented as a process on this space. Properties of this process are investigated and some special cases examined.

Recently there has been a growing interest in modelling the interest rate term structure as a dynamical system (see, e.g., Musiela 1993, Björk & Svensson 1999, Björk & Christensen 1999, Björk & Gombani 1999, Björk 2000, Brody & Hughston 2000, Filipović 2000). The idea is a simple one: we treat the yield curve as a mathematical object in its own right, a “point” ρ lying in a space \mathfrak{M} , which we identify as “the space of all possible smooth yield curves”. Then with the specification of an initial yield curve ρ_0 we wish to understand the resulting dynamics, which we model as a random trajectory ρ_t in \mathfrak{M} . The following questions then arise: How do we best model \mathfrak{M} ? What natural structures exist on \mathfrak{M} that might assist us in determining the dynamical laws governing ρ_t ? These questions are interesting because the standard HJM theory (Heath, Jarrow & Morton 1992) of arbitrage-free interest rate dynamics makes no use of specific structures on \mathfrak{M} . The thought is that by bringing the structure of \mathfrak{M} into play it will be possible both to clarify the status of existing interest models, and to devise new interest rate models that might form a more adequate representation of the properties of observed interest rates. In what follows we shall sketch some tentative examples of new models that can be developed by following this line of argument.

The key point we emphasise here is that there is a natural information geometry

¹) Appeared in *Disordered and Complex Systems* (p. 281-287), P. Sollich, A.C.C. Coolen, L.P. Hughston & R.F. Streater, eds., AIP.

associated with the space of yield curves. This comes about as follows. Let $t = 0$ denote the present, and P_{0T} a family of discount bond prices satisfying $P_{00} = 1$, where T is the maturity date ($0 \leq T < \infty$). We impose the economic condition that interest rates should always be positive by means of the following criterion:

Definition. *A term structure will be said to be admissible if the discount function P_{0T} is of class C^2 and satisfies $0 < P_{0T} \leq 1$, $\partial_T P_{0T} < 0$, and $\lim_{T \rightarrow \infty} P_{0T} = 0$.*

The remarkable fact about an admissible discount function P_{0T} is that it can be viewed as a complementary probability distribution: we think of the maturity date as an abstract random variable X , and for its distribution we write $\Pr[X < T] = 1 - P_{0T}$. Then the associated density function $\rho(T) = -\partial_T P_{0T}$ clearly satisfies $\rho(T) > 0$ for all T , and $\int_T^\infty \rho(u) du = P_{0T}$. Now suppose we define the positive real line $\mathbb{R}_+^1 = [0, \infty)$. We shall say a density function is *smooth* if it is of class C^2 on \mathbb{R}_+^1 . Then we have the following characterisation of the space \mathfrak{M} :

Proposition 1. *The system of admissible term structures is isomorphic to the convex space $\mathcal{D}(\mathbb{R}_+^1)$ of smooth density functions on the positive real line.*

The requirement that P_{0T} should be C^2 is to some extent arbitrary, and we may consider strengthening or weakening this condition slightly. It is reasonable, however, to insist that the forward short rate curve $f_{0T} = -\partial_T \ln P_{0T}$ should be continuous and nonvanishing for all $T < \infty$.

The space of probability distributions on a given sample space has natural geometric properties. In particular, given a pair of term structure densities $\rho_1(x)$ and $\rho_2(x)$ we can define a distance function ϕ_{12} on \mathfrak{M} by the formula:

$$\phi_{12} = \cos^{-1} \int_0^\infty \xi_1(x) \xi_2(x) dx, \quad (1)$$

where $\xi(x) = \sqrt{\rho(x)}$. We call this ‘angle’ the Bhattacharyya distance between the given yield curves. The geometrical interpretation of ϕ_{12} is as follows. The map $\rho(x) \rightarrow \xi(x)$ associates to each point of \mathfrak{M} a point in the positive orthant \mathcal{S}^+ of the unit sphere in the Hilbert space $L^2(\mathbb{R}_+^1)$, and ϕ is the resulting spherical angle on \mathcal{S}^+ . We note that $0 \leq \phi < \frac{1}{2}\pi$ and that orthogonality can never quite be achieved on account of the requirement that forward rates are always nonvanishing.

Example 1. The discount bond family $P_{0T} = \exp(-rT)$ for constant r determines a ‘flat’ term structure with a continuously compounded rate of interest r for each value of the maturity date T . The spherical distance between two such yield curves is $\phi_{12} = \cos^{-1} (2\sqrt{r_1 r_2} / (r_1 + r_2))$. \diamond

We note that in this case the distance function is determined by the ratio of the geometric and arithmetic means of the corresponding rates.

Example 2. The discount bond family $P_{0T} = (1 + \kappa^{-1}rT)^{-\kappa}$ for constant r and κ corresponds to a ‘flat’ term structure with a constant annualised rate of interest r assuming a compounding frequency κ over the life of each bond (κ need not be an integer). In particular, the spherical distance between two such yield curves for $\kappa = 1$ is $\phi_{12} = \cos^{-1} (\ln(r_1/r_2) \sqrt{r_1 r_2} / (r_1 - r_2))$. \diamond

The spherical distance computation is applicable when we wish to make a comparison in nonparametric situations. In the case of a parametric family of yield curves we consider the corresponding submanifold of \mathfrak{M} and the Fisher-Rao geometry induced on this submanifold (cf. Brody & Hughston 1998).

Now suppose we write P_{tT} for the random value at time t of a discount bond that matures at time T , where $T \in \mathbb{R}_+^1$ and $0 \leq t \leq T$. We assume for each value of T that P_{tT} is an Ito process defined on the interval $t \in [0, T]$, for which the dynamics are given by $dP_{tT} = m_{tT}dt + \Sigma_{tT}dW_t$. The process W_t , to which P_{tT} is assumed to be adapted, is a standard Brownian motion taking values in a separable Hilbert space \mathcal{H} (cf. Da Prato & Zabczyk 1992, Filipović 2000). The absolute drift m_{tT} thus defined, along with the absolute volatility process Σ_{tT} , which takes values in the dual Hilbert space \mathcal{H}^* , are assumed to satisfy regularity conditions sufficient to ensure that $\partial_T P_{tT}$ is also an Ito process. For interest rate positivity we require $0 < P_{tT} \leq 1$ and $\partial_T P_{tT} < 0$. Additionally we impose the boundary conditions $P_{TT} = 1$, $\lim_{T \rightarrow \infty} P_{tT} = 0$, and $\lim_{T \rightarrow \infty} \partial_T P_{tT} = 0$.

We define the short rate $r_t = \rho_{tt}$, and the forward short rate $f_{tT} = \rho_{tT}/P_{tT}$. Because P_{tT} is positive, f_{tT} is an Ito process iff ρ_{tT} is an Ito process. We note that f_{tT} is the forward rate fixed at time t for the short rate at time T . The formula $\int_t^\infty \rho_{tu}du = 1$ says that the value at time t of a continuous cash flow that pays the small amount $f_{tu}du$ at time u is unity. For no arbitrage we require the existence of an \mathcal{H}^* -valued process λ_t , independent of T , such that $m_{tT} = r_t P_{tT} + \lambda_t \Sigma_{tT}$. We do not assume the bond market is complete.

To proceed further, we introduce the Musiela parameterisation $x = T - t$ for time to maturity and write $B_{tx} = P_{t,t+x}$ for the bond price thus parameterised. Thus B_{tx} represents the price at time t of a bond for which the time left until maturity is x , and we have the dynamical relation

$$dB_{tx} = (r_t - f_{t,t+x})B_{tx}dt + \Sigma_{t,t+x}(dW_t + \lambda_t dt), \quad (2)$$

which has a simple intuitive interpretation. Now suppose we write $\rho_t(x) = -\partial_x B_{tx}$. Because $\rho_t(x) > 0$ and $\int_0^\infty \rho_t(x)dx = 1$ for all $t > 0$, it follows that $\rho_t(x)$ is a density valued process. Then writing $\omega_{tx} = -\partial_x \Sigma_{t,t+x}$ we obtain the following dynamics:

$$d\rho_t(x) = (r_t \rho_t(x) + \partial_x \rho_t(x))dt + \omega_{tx}(dW_t + \lambda_t dt). \quad (3)$$

The process ω_{tx} is subject to the constraints $\int_0^\infty \omega_{tx}dx = 0$ and $\lim_{x \rightarrow \infty} \omega_{tx} = 0$, which implies that it can be expressed in the form $\omega_{tx} = \rho_t(x)(\nu_t(x) - \bar{\nu}_t)$, where $\bar{\nu}_t = \int_0^\infty \rho_t(u)\nu_t(u)du$ and $\nu_t(x)$ is unconstrained. The process $\nu_t(x)$ plays a role similar to the HJM forward short rate volatility. However, if the volatility structure is specified arbitrarily in the HJM theory there is no guarantee that the interest rate system will be admissible, whereas that feature is built into the present dynamics. The resulting bond volatility structure Σ_{tx} is invariant under transformations of the form $\nu_t(x) \rightarrow \nu_t(x) + \alpha_t$. This freedom can be used to set $\lambda_t = -\bar{\nu}_t$ without loss of generality. Then, both λ_t and Σ_{tx} are determined by $\nu_t(x)$.

Proposition 2. *The general admissible term structure evolution based on the information set generated by a Brownian motion W_t on a Hilbert space \mathcal{H} is given by a measure valued process $\rho_t(x)$ in $\mathcal{D}(\mathbb{R}_+^1)$ satisfying*

$$\frac{d\rho_t(x)}{\rho_t(x)} = (\rho_t(0) + \partial_x \ln \rho_t(x)) dt + (\nu_t(x) - \bar{\nu}_t) (dW_t - \bar{\nu}_t dt), \quad (4)$$

where $\bar{\nu}_t = \int_0^\infty \rho_t(u) \nu_t(u) du$. The process $\nu_t(x)$ can be specified exogenously along with the initial term structure density $\rho_0(x)$.

It follows from the relation $\rho_t(0) = r_t$, that the process for the short rate satisfies

$$dr_t = (r_t^2 + \partial_x \rho_t(x)|_{x=0}) dt + r_t(\nu_t(0) - \bar{\nu}_t)(dW_t - \bar{\nu}_t dt). \quad (5)$$

The appearance of r_t^2 in the drift might seem counterintuitive. This is compensated by the second term appearing in the drift. For example, if we consider the CIR model (Cox et al. 1985), a calculation shows that the r_t^2 term cancels a similar term arising from $\partial_x \rho_t(x)|_{x=0}$, leaving the correct mean-reverting behaviour.

Proposition 3. *The solution of the dynamical equation for $\rho_t(x)$ in terms of the volatility structure $\nu_t(x)$ and the initial term structure density $\rho_0(x)$ is*

$$\rho_t(T-t) = \rho_0(T) \frac{\exp\left(\int_{s=0}^t V_{sT} dW_s - \frac{1}{2} \int_{s=0}^t V_{sT}^2 ds\right)}{\int_{u=t}^\infty \rho_0(u) \exp\left(\int_{s=0}^t V_{su} dW_s - \frac{1}{2} \int_{s=0}^t V_{su}^2 ds\right) du}, \quad (6)$$

where $V_{tu} = \nu_t(u-t)$. The corresponding formula for the bond price process is

$$P_{tT} = \frac{\int_{u=T}^\infty \rho_0(u) \exp\left(\int_{s=0}^t V_{su} dW_s - \frac{1}{2} \int_{s=0}^t V_{su}^2 ds\right) du}{\int_{u=t}^\infty \rho_0(u) \exp\left(\int_{s=0}^t V_{su} dW_s - \frac{1}{2} \int_{s=0}^t V_{su}^2 ds\right) du}, \quad (7)$$

and for the unit-initialised money market account $B_t = \exp\left(\int_0^t r_s ds\right)$ we have

$$B_t = \frac{\exp\left(\int_{s=0}^t \bar{\nu}_s dW_s - \frac{1}{2} \int_{s=0}^t \bar{\nu}_s^2 ds\right)}{\int_{u=t}^\infty \rho_0(u) \exp\left(\int_{s=0}^t V_{su} dW_s - \frac{1}{2} \int_{s=0}^t V_{su}^2 ds\right) du}. \quad (8)$$

The formulae for $\rho_t(x)$, P_{tT} and B_t indicated in Proposition 3 are surprisingly simple, given the nonlinearities of the dynamics of $\rho_t(x)$. These results can be used as a starting point for the derivation of new interest rate models, some of which we remark on briefly below.

Example 3. If $\nu_t(x)$ is chosen to be a deterministic function of t and x then we obtain a class of ‘quasi-lognormal’ models. In such a model the bond prices are given by ratios of superpositions of lognormally distributed random variables. \diamond

Example 4. Martingale representations for discount bonds. It follows from Proposition 3 that an alternative expression for $\rho_t(x)$ is given by

$$\rho_t(T-t) = \frac{\rho_0(T)M_{tT}}{\int_{u=t}^{\infty} \rho_0(u)M_{tu}du}, \quad (9)$$

where for each value of T the process M_{tT} is a martingale over the interval $0 \leq t \leq T$, such that $M_{tT} > 0$ and $M_{0T} = 1$. We see that M_{tT} is the exponential martingale associated with V_{tT} . This expression for $\rho_t(T-t)$ arises also in the Flesaker-Hughston model (Flesaker & Hughston 1996, 1997, 1998; Rutkowski 1997; Musiela & Rutkowski 1997; Brody 2000; Hunt & Kennedy 2000), for which we have

$$P_{tT} = \frac{\int_{u=T}^{\infty} \rho_0(u)M_{tu}du}{\int_{u=t}^{\infty} \rho_0(u)M_{tu}du} \quad (10)$$

as a representation for the bond price. \diamond

Example 5. Quasi-linear models. Suppose we write $\rho_0(u) = \int_0^{\infty} e^{-uR}\phi(R)dR$, where $\phi(R)$ is the inverse Laplace transform of the initial term structure density. Then for certain choices of the martingale family M_{tu} the integration $\int_T^{\infty} e^{-uR}M_{tu}du$ can be carried out explicitly. For instance, let M_t be an arbitrary martingale ($0 \leq t < \infty$) and Q_t the associated quadratic variation, so $(dM_t)^2 = dQ_t$, and set $M_{tT} = \exp\left((\alpha + \beta T)M_t - \frac{1}{2}(\alpha + \beta T)^2Q_t\right)$. This model is obtained by putting $\nu_t(T-t) = (\alpha + \beta T)\sigma_t$ in Proposition 3, where $dM_t = \sigma_t M_t dW_t$. Then the u -integration can be carried out explicitly in the expressions for $\rho_t(x)$ and P_{tT} , and the results expressed in closed form:

$$\begin{aligned} \int_{u=T}^{\infty} e^{-uR}M_{tu}du &= \frac{1}{|\beta|\sqrt{Q_t}} \exp\left(\frac{1}{2}\frac{(M_t - \beta^{-1}R)^2}{Q_t} + \alpha\beta^{-1}R\right) \\ &\times \mathcal{N}\left(\pm\frac{(M_t - \beta^{-1}R)}{\sqrt{Q_t}} \mp (\alpha + \beta T)\sqrt{Q_t}\right), \end{aligned} \quad (11)$$

where $\mathcal{N}(x) = (2\pi)^{-1/2} \int_{-\infty}^x \exp(-\frac{1}{2}\xi^2)d\xi$ is the normal distribution function, and the \pm sign is chosen according to the sign of β (cf. Brody 2000). These models are ‘quasi-linear’ in the sense that we have a closed form solution for a simple flat-rate initial term structure, and more general solutions can be constructed by taking linear superpositions of the resulting ‘simple’ martingale families. \diamond

Example 6. Canonical exponential models. If we let $\rho_t(x)$ be of the form

$$\rho_t(x) = \frac{\exp(g_t(x) - \theta_t h_t(x))}{\int_{u=0}^{\infty} \exp(g_t(u) - \theta_t h_t(u)) du}, \quad (12)$$

where θ_t is a one-dimensional diffusion and the processes $g_t(x)$ and $h_t(x)$ are deterministic, then at each time t the term structure belongs to an exponential family parameterised by the value of θ_t . The no-arbitrage condition implies that θ_t is a

time inhomogeneous square-root process, and that $g_t(x)$ and $h_t(x)$ satisfy Riccati-type equations. \diamond

Example 7. State-variable models. Here we consider the structure of admissible term structure dynamics driven by a stationary Markov process. In this case we write $\rho_t(T-t) = N_{tT} / \int_{u=t}^{\infty} N_{tu} du$, and we assume that the martingale family N_{tT} is of the form $N_{tT} = F(\psi_t, T-t)$, where the vector state variable ψ_t is a homogeneous diffusion process satisfying $d\psi_t = \alpha(\psi_t)dt + \beta(\psi_t)dW_t$. The condition that N_{tT} has no drift is $\partial_T F = \Delta F$, where $\Delta = \alpha\partial_\psi + \frac{1}{2}\beta^2\partial_\psi^2$ is the generator of the diffusion. The solution is formally $F(\psi_t, T-t) = e^{(T-t)\Delta}\Phi(\psi_t)$, where $\Phi(\psi_t)$ is a positive bounded function of class C^∞ . We assume that Δ is negative and possesses an inverse Δ^{-1} . The solution for $\rho_t(x)$ is

$$\rho_t(x) = \frac{e^{x\Delta}\Phi(\psi_t)}{\int_{u=0}^{\infty} e^{u\Delta}\Phi(\psi_t)du}, \quad (13)$$

or, equivalently, $\rho_t(x) = -e^{x\Delta}\Phi(\psi_t)/\Delta^{-1}\Phi(\psi_t)$. It follows, further, on account of the Markov property, that $N_{tT} = E_t[\Phi(\psi_T)]$. The stationarity property ensures that $\alpha(\psi_t)$, $\beta(\psi_t)$ and $\Phi(\psi_t)$ determine the term structure at any time t as a function of the state variable ψ_t . This applies in particular to the initial term structure. The short rate is given by $r_t = -\Delta Z(\psi_t)/Z(\psi_t)$, where $Z(\psi_t) = \Delta^{-1}\Phi(\psi_t)$, and for the bond prices we have $B_{tx} = e^{x\Delta}Z(\psi_t)/Z(\psi_t)$. The forward short rates are then given by $r_{tx} = -e^{x\Delta}\Delta Z_t/e^{x\Delta}Z_t$, where $r_{tx} = f_{t,t+x}$, and we have $\rho_t(x) = -e^{x\Delta}\Delta Z_t/Z_t$. The significance of the process $Z(\psi_t)$ is that it represents the state-price density (cf. Rogers 1997, Hunt & Kennedy 2000), and a calculation shows that $dZ_t/Z_t = -r_t dt - \lambda_t dW_t$, where $\lambda_t = (\partial_\psi \ln Z)\beta_t$ is the risk premium process. \diamond

Let us turn now to the problem of the relative movement of term structures, which can be given a simple characterisation by the use of information geometry. Given a density process $\rho(x)$ we form the associated square root process $\xi(x)$. Letting $\nu(x)$ and λ be exogenous and writing $\bar{\nu} = \int_0^\infty \rho(x)\nu(x)dx$, we have

$$d\xi = \left(\frac{1}{2}r\xi(x) + \partial_x\xi(x) - \frac{1}{8}(\nu(x) - \bar{\nu})^2\xi(x)\right) dt + \frac{1}{2}\xi(x)(\nu(x) - \bar{\nu})(dW + \lambda dt), \quad (14)$$

where for convenience here we suppress the time index. Given a pair of such processes $\xi_1(x)$ and $\xi_2(x)$ we form the process for the cosine of the spherical distance ϕ_{12} between the corresponding yield curves given in formula (1). An increase in $\cos \phi_{12}$ means a decrease in the distance. Writing $\sigma(x) = \nu(x) - \int_0^\infty \xi^2(x)\nu(x)dx$, for the dynamics of $\cos \phi_{12}$ we obtain

$$d \cos \phi_{12} = \left(\frac{1}{2}(r_1 + r_2) \cos \phi_{12} - \sqrt{r_1 r_2} - \frac{1}{8} \int_0^\infty (\sigma_1(x) - \sigma_2(x))^2 \xi_1(x) \xi_2(x) dx \right) dt + \frac{1}{2} \left(\int_0^\infty (\sigma_1(x) + \sigma_2(x)) \xi_1(x) \xi_2(x) dx \right) (dW + \lambda dt). \quad (15)$$

A special case arises when we consider the relative dynamics of two yield curves with differing initial conditions but governed by the same volatility and risk premium $\nu(x)$ and λ . Defining $\bar{\nu}_{12} = \int_0^\infty \xi_1(x)\xi_2(x)\nu(x)dx$, we have:

Proposition 4. *The Bhattacharyya distance process for two yield curves subject to the same underlying interest rate dynamics satisfies*

$$d \cos \phi_{12} = \left(\frac{1}{2}(r_1 + r_2) \cos \phi_{12} - \sqrt{r_1 r_2} - \frac{1}{8}(\bar{\nu}_1 - \bar{\nu}_2)^2 \cos \phi_{12} \right) dt + \left(\bar{\nu}_{12} - \frac{1}{2}(\bar{\nu}_1 + \bar{\nu}_2) \cos \phi_{12} \right) (dW + \lambda dt). \quad (16)$$

The distance process is clearly invariant under transformations of the form $\nu(x) \rightarrow \nu(x) + \alpha$. It is interesting to note that the ratio of the geometric and arithmetic means of the short rates plays a critical role in determining the behaviour of $\cos \phi$. In particular, in a risk neutral world the two curves will necessarily tend to diverge once the distance is sufficiently great to ensure that $\cos \phi \leq 2\sqrt{r_1 r_2}/(r_1 + r_2)$.

Acknowledgements. The authors are grateful to P. Balland, T. Björk, D. Filipović, B. Flesaker, P. Glasserman, R. A. Jarrow, Y. Jin, T. Knudsen, D. Madan, and R. F. Streater for helpful remarks. DCB acknowledges the support of The Royal Society. LPH acknowledges the hospitality of the Finance Department of the Graduate School of Business of the University of Texas at Austin.

REFERENCES

1. Björk, T. & Christensen, B. J., *Mathematical Finance* **9**, 323 (1999).
2. Björk, T. & Gombani, A., *Finance and Stochastics* **3**, 413 (1999).
3. Björk, T., A Geometric View of Interest Rate Theory. In *Handbook of Mathematical Finance*, Cambridge: Cambridge University Press (2000).
4. Brody, D. C., *Mathematical Theory of Finance*, Tokyo: Nihon-Hyoronsya (2000).
5. Brody, D. C. & Hughston, L. P., *Proc. Roy. Soc. London* **454**, 2445 (1998).
6. Brody, D. C. & Hughston, L. P., Interest Rates and Information Geometry (Preprint, Imperial College London and King's College London, 2000).
7. Cox, J. C., Ingersoll, J. & Ross, S., *Econometrica* **53**, 385 (1985).
8. Da Prato, G. & Zabczyk, J., *Stochastic Equations in Infinite Dimensions*, Cambridge: Cambridge University Press (1992).
9. Filipović, D., Consistency Problems for HJM Interest Rate Models, PhD thesis, Swiss Federal Institute of Technology (2000).
10. Flesaker, B. & Hughston, L. P., *Risk Magazine* **9**, 46 (1996).
11. Flesaker, B. & Hughston, L. P., *Net Exposure* **3**, 55 (1997).
12. Flesaker, B. & Hughston, L. P., Positive Interest: An Afterword, in *Hedging with Trees*, ed. Broadie, M. & Glasserman, P., London: Risk Publications (1998).
13. Hunt, P. J. & Kennedy, J. E., *Financial Derivatives in Theory and Practice*, Amsterdam: North-Holland (2000).

14. Heath, D., Jarrow, R. & Morton, A., *Econometrica* **60**, 77 (1992).
15. Musiela, M., Stochastic PDEs and Term Structure Models, Journées Internationales de Finance, La Baule (1993).
16. Musiela, M. & Rutkowski, M., *Martingale Methods in Financial Modelling*, Berlin: Springer-Verlag (1997).
17. Rogers, L. C. G., *Math. Finance* **7**, 157 (1997).
18. Rutkowski, M., *Applied Math. Finance* **4**, 151 (1997).