ISRN KTH/OPT SYST/FR 98/3 SE
Coden: TRITA/MAT-98-OS3
ISSN 0348-405X

A note on maximum-smoothness approximation of forward interest rate

by

Anders Forsgren

Optimization and Systems Theory

DEPARTMENT OF MATHEMATICS
ROYAL INSTITUTE OF TECHNOLOGY
SE-100 44 STOCKHOLM, SWEDEN

ISSN 0348-405X

A NOTE ON MAXIMUM-SMOOTHNESS APPROXIMATION OF FORWARD INTEREST RATE

Anders FORSGREN*

Technical Report TRITA-MAT-1998-OS3
Department of Mathematics
Royal Institute of Technology
March 1998

Abstract

Maximum-smoothness approximation of forward interest rate is considered. The smoothness is measured as the integral of the square of the second-derivative of the forward interest rate. Well-known results on natural splines are utilized in order to characterize the optimal solution.

1. Introduction

Assume that we have m zero-coupon bonds, such that bond i, $i = 1, \ldots, m$, gives unit value at time t_i, and its market price at time zero, v_i, is known. Here, and throughout this note, the current time is set to zero and denoted by t_0. Without loss of generality, we assume that $0 = t_0 < t_1 < \cdots < t_m$. It is of interest to extend these market prices to a smooth curve $v(t)$ which is consistent with the information that is known, i.e., $v(t_0) = 1$ and $v(t_i) = v_i$, $i = 1, \ldots, m$. Alternatively, a function $f(t)$, which models the forward interest rate, may be formed, where for $t \in [0, t_m]$, the relationship between $v(t)$ and the function f is given by

$$v(t) = e^{-\int_{t_0}^{t} f(s)ds}. \tag{1.1}$$

This note concerns finding a suitable smooth f. We focus on a specific model proposed by Adams and van Deventer [AvD94] and extended by Bjerksund and Stensland [BS96]. For a more general discussion, see, e.g., Adams and van Deventer [AvD94], Bjerksund and Stensland [BS96], Frishling and Yamamura [FY96], Tanggaard [Tan97], and the references given in these papers.

Using (1.1), the consistency with the market price v_i of bond i, for $i = 1, \ldots, m$, may be expressed as

$$v_i = e^{-\int_{t_0}^{t_i} f(s)ds}, \quad i = 1, \ldots, m,$$

or equivalently, as

$$-\ln v_i = \int_{t_0}^{t_i} f(s)ds, \quad i = 1, \ldots, m. \tag{1.2}$$

*Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden (andersf@math.kth.se). Research supported by the Swedish Natural Science Research Council (NFR).

Adams and van Deventer [AvD94] have suggested a choice of f that gives "maximum smoothness" in the sense that they include a final time of the planning horizon, t_{m+1}, $(t_{m+1} > t_m)$, and suggest choosing the f that has minimal L_2-norm of its second-derivative over the planning horizon, i.e., they minimize

$$\int_{t_0}^{t_{m+1}} (f''(s))^2 ds.$$

In addition, they require that $f \in C^1[t_0, t_{m+1}]$ and $f \in C^5[t_{i-1}^+, t_i^-]$, $i = 1, \ldots, m+1$. Here, and throughout the note, superscripts "$-$" and "$+$" are used to denote limits from left and right respectively. Taking into account the consistency requirement (1.2), this measure of smoothness leads to an optimization problem on the form

$$\begin{array}{ll} \underset{f \in C^1[t_0, t_{m+1}]}{\text{minimize}} & \displaystyle\int_{t_0}^{t_{m+1}} (f''(s))^2 ds \\ \text{subject to} & \displaystyle\int_{t_0}^{t_i} f(s) ds = -\ln v_i, \quad i = 1, \ldots, m, \\ & f \in C^5[t_{i-1}^+, t_i^-], \quad i = 1, \ldots, m+1. \end{array} \tag{1.3}$$

For this problem, Adams and van Deventer [AvD94, p. 55] claim that the optimal f may be written as a piecewise fourth-order polynomial, where the cubic and quadratic terms are absent on each interval. This is not correct. As stated in Section 2, the solution is given by so-called *natural splines*, see Schwartz [Sch89, pp. 126–128]. This means that the fourth-order polynomials have continuous derivatives up to order three over the whole planning horizon.

The purpose of this note is to give the properties of an optimal solution to problem (1.3), thereby correcting the abovementioned mistake in Adams and van Deventer [AvD94]. In addition, we characterize an optimal solution of the related problem that arises when the values of v_i are not uniquely determined. This latter problem has been proposed by Bjerksund and Stensland [BS96]. The solutions to both these problems are given by natural splines. The results concerning the former problem are given in Section 2. These results follow immediately from the analysis of Schwartz [Sch89, pp. 126–128]. In Section 3, the results concerning the latter problem are given. The proof of the extended results of Section 3 are given in Appendix A. (In addition, the proof of the results of Section 2 follows as a by-product.) In Appendix B, explicit nonlinear equations that characterize an optimal solution to the latter problem are given.

2. Characterization of the optimal solution

If the primitive function $F(t) = \int_{t_0}^{t} f(s)ds$ is introduced in (1.3), the equivalent result regarding minimizing the L_2-norm of the third-derivative of F over the planning horizon is given in Schwartz [Sch89, p. 128]. This result is reviewed in the following proposition, stated for $f(t)$, taking into account that $f(t) = F'(t)$. As mentioned above, superscripts "$-$" and "$+$" are used to denote limits from left and right respectively, such as $f(t_i^-) = \lim_{\epsilon \downarrow 0} f(t_i - \epsilon)$ and $f(t_i^+) = \lim_{\epsilon \downarrow 0} f(t_i + \epsilon)$. Superscripts at endpoints of the interval considered in the specific optimization problem, such as t_0, will be omitted for brevity.

Proposition 2.1. (See Schwartz [Sch89, p. 128]) *If f is a feasible solution to* (1.3)*, then f is a global minimizer if and only if*

1. $f''(t_0) = f''(t_m) = 0, \quad f''(t_i^-) = f''(t_i^+),\ i = 1, \ldots, m-1.$
2. $f^{(3)}(t_0) = f^{(3)}(t_m) = 0, \quad f^{(3)}(t_i^-) = f^{(3)}(t_i^+),\ i = 1, \ldots, m-1.$
3. $f^{(5)}(t) = 0,\ t_{i-1} < t < t_i,\ i = 1, \ldots, m.$
4. $f''(t) = 0,\ t_m < t < t_{m+1}.$

In Appendix A, the analysis of Schwartz [Sch89, pp. 126–128] is extended to the more general case of Section 3 when the values of v_i are not uniquely determined. Schwartz' proof of Proposition 2.1 is then obtained as a by-product. The consequence of Proposition 2.1 is that the optimal solution is a piecewise quadratic polynomial, which is three times continuously differentiable on $[t_0, t_{m+1}]$. This result is a correction to a previous result given by Adams and van Deventer [AvD94], where the continuity properties of the second and third derivatives are not correctly stated. Adams and van Deventer [AvD94] claim that the optimal f may be written as a piecewise fourth-order polynomial, where the cubic and quadratic terms vanish on each interval. In general, this is *not* a property which is given by Proposition 2.1, but rather the second and third derivatives are continuous over the interval limits.

Note that since there are no constraints on f on the interval $[t_m, t_{m+1}]$, the optimal f will be a linear function on that interval, determined by the function value and derivative at t_m. Hence, the optimization may be limited to the interval $[t_0, t_m]$, which is done for the remainder of this note.

By Proposition 2.1, the optimal f is a piecewise quadratic polynomial, and (1.3) may be posed as an equality-constrained convex quadratic programming problem, see the discussion in Section 4. However, following Schwartz [Sch89, pp. 126–128], the properties of Proposition 2.1 uniquely determine the coefficients of the polynomial. Hence, it is not necessary to explicitly formulate or solve the quadratic programming problem. To see this, for $i = 1, \ldots, m$, let

$$f_i(t) = a_i(t - t_{i-1})^4 + b_i(t - t_{i-1})^3 + c_i(t - t_{i-1})^2 + d_i(t - t_{i-1}) + e_i, \qquad (2.1)$$

for $t_{i-1} < t < t_i$. (As discussed above, we limit the discussion to the interval $[t_0, t_m]$.) This gives $5m$ unknown coefficients a_i, b_i, c_i, d_i and e_i, $i = 1, \ldots, m$. The continuity of f and f', imposed by $f \in C^1[t_0, t_{m+1}]$ give $2(m-1)$ conditions. The requirements on f'' and $f^{(3)}$, imposed by properties 1 and 2 of Proposition 2.1, give $2(m+1)$ conditions. Finally, v_i, $i = 1, \ldots, m$, are known, giving an additional m conditions. Consequently, there is a total of $5m$ conditions for the $5m$ unknowns.

3. Extensions

A slightly more general model has been proposed by Bjerksund and Stensland [BS96]. It is assumed that, for each bond j, $j = 1, \ldots, n$, today's price p^j and the payments

p_i^j at times t_i, $i \in I_j$, are known, where $I_j \subseteq \{1, \dots, m\}$. The consistency requirement on v_i, $i = 1, \dots, m$, then becomes

$$p^j = \sum_{i \in I_j} p_i^j v_i, \quad j = 1, \dots, n. \tag{3.1}$$

By defining $p_i^j = 0$ if $i \notin I_j$, an $m \times n$ matrix P with element (i, j) given by p_i^j may be created, and (3.1) may be written as $p^j = \sum_{i=1}^m p_i^j v_i$, $j = 1, \dots, n$, or equivalently as $p = P^T v$, if p is defined as the n-dimensional vector with jth component p^j and v here denotes the m-dimensional vector with ith component v_i. The associated optimization problem becomes

$$\begin{array}{ll} \underset{\substack{f \in C^1[t_0, t_m] \\ v \in I\!R^m}}{\text{minimize}} & \displaystyle\int_{t_0}^{t_m} (f''(s))^2 ds \\ \text{subject to} & \displaystyle\int_{t_0}^{t_i} f(s) ds = -\ln v_i, \quad i = 1, \dots, m, \\ & \displaystyle\sum_{i=1}^m p_i^j v_i = p^j, \quad j = 1, \dots, n, \\ & f \in C^5[t_{i-1}^+, t_i^-], \quad i = 1, \dots, m, \end{array} \tag{3.2}$$

see Bjerksund and Stensland [BS96]. Throughout the note, it is assumed that each bond affects the value of at least one v_i, $i = 1, \dots, m$, i.e., that the matrix P of above has full column rank. (If this is not the case, and there is a feasible solution, bonds associated with linearly dependent columns of P can simply be removed.)

If $n = m$, then P is square and nonsingular, and in this situation, v_i, $i = 1, \dots, m$, are uniquely determined. Hence, in this situation, (3.2) is equivalent to (1.3). This is the case, if for example $|I_j| = 1$ for $j = 1, \dots, n$. If $m > n$, then the values of v_i are not uniquely determined, and the problem becomes nonlinearly constrained and nonconvex. The properties of Proposition 2.1 become necessary optimality conditions for problem (3.2), with an additional condition on $f^{(4)}$, as stated in the following proposition.

Proposition 3.1. *If f and v_i, $i = 1, \dots, m$, form a local minimizer to* (3.2)*, then*

1. $f''(t_0) = f''(t_m) = 0, \quad f''(t_i^-) = f''(t_i^+)$, $i = 1, \dots, m-1$.
2. $f^{(3)}(t_0) = f^{(3)}(t_m) = 0, \quad f^{(3)}(t_i^-) = f^{(3)}(t_i^+)$, $i = 1, \dots, m-1$.
3. *There is a* $\lambda \in I\!R^n$ *such that*

$$f^{(4)}(t_i^-) - f^{(4)}(t_i^+) - v_i \sum_{j=1}^n p_i^j \lambda_j = 0, \quad i = 1, \dots, m-1,$$

$$f^{(4)}(t_m) - v_m \sum_{j=1}^n p_m^j \lambda_j = 0.$$

4. $f^{(5)}(t) = 0$, $t_{i-1} < t < t_i$, $i = 1, \dots, m$.

A proof based on the analysis of Schwartz [Sch89, pp. 126–128] is given in Appendix A.

Rather than requiring the prices to be satisfied exactly, as in (3.1), other approaches have been proposed, where the equations (3.1) are not required to be satisfied exactly, but the deviation is penalized by a least-squares term, see Tanggaard [Tan97].

4. Solution of the optimization problem

By introducing piecewise quadratic polynomials as defined in (2.1), Proposition 3.1 suggests that problem (3.2) may be posed as

$$\begin{array}{ll}\underset{a,b,c,d,e,v}{\text{minimize}} & \displaystyle\sum_{i=1}^{m}\int_{t_{i-1}}^{t_i}(12a_i(s-t_{i-1})^2+6b_i(s-t_{i-1})+2c_i)^2 ds\\ \text{subject to} & a_i(t_i-t_{i-1})^4+b_i(t_i-t_{i-1})^3+c_i(t_i-t_{i-1})^2\\ & \quad +d_i(t_i-t_{i-1})+e_i=e_{i+1}, \quad i=1,\ldots,m-1,\\ & 4a_i(t_i-t_{i-1})^3+3b_i(t_i-t_{i-1})^2+2c_i(t_i-t_{i-1})\\ & \quad +d_i=d_{i+1}, \quad i=1,\ldots,m-1,\\ & \displaystyle\int_{t_{i-1}}^{t_i}(a_i(s-t_{i-1})^4+b_i(s-t_{i-1})^3+c_i(s-t_{i-1})^2\\ & \quad +d_i(s-t_{i-1})+e_i)ds=-\ln v_i+\ln v_{i-1}, \quad i=1,\ldots,m,\\ & \displaystyle\sum_{i=1}^{m}p_i^j v_i=p^j, \quad j=1,\ldots,n,\\ & a\in I\!R^m, b\in I\!R^m, c\in I\!R^m, d\in I\!R^m, e\in I\!R^m, v\in I\!R^m, v_0=1.\end{array} \tag{4.1}$$

If the values of v_i are uniquely determined by the equations $\sum_{i=1}^{m} p_i^j v_i = p^j$, $j=1,\ldots,n$, then (4.1) is a convex equality-constrained quadratic programming problem in the variables $a\in I\!R^m$, $b\in I\!R^m$, $c\in I\!R^m$, $d\in I\!R^m$ and $e\in I\!R^m$. In this situation, Proposition 2.1 ensures that (4.1) is equivalent to (1.3). Otherwise, (4.1) is in general a nonconvex optimization problem. A Newton-type method based on the Karush-Kuhn-Tucker first-order optimality conditions of (4.1) may be employed in order to identify a local minimizer, see, e.g., Fletcher [Fle87, pp. 304–317] and Nash and Sofer [NS96, Chapter 15.5]. We do not focus on not how best to solve the optimization problem, but note that an alternative to solving the first-order optimality conditions is is to solve a system of nonlinear equations that arises from the alternative optimality conditions of Proposition 3.1. Such a system of nonlinear equations is presented in Appendix B. Both approaches lead to Newton iterations involving sparse matrices, with a structured sparsity pattern. Our limited computational experience on test problems suggests that these matrices in both cases may become ill-conditioned. An advantage of a method based on the Karush-Kuhn-Tucker first-order optimality conditions of (4.1) is that such a method also contains second-derivative information, and hence, second-order optimality conditions can be verified, see, e.g., Nash and Sofer [NS96, Chapter 14].

5. Concluding remarks

Natural splines have been used to characterize the solution of the maximum smoothness problems (1.3) and (3.2). We note that the analysis of Schwartz [Sch89, pp. 126–128] can be used to give similar results also if the smoothness measure $\int_{t_0}^{t_m} (f'(s))^2 ds$ is used, as proposed by Frishling and Yamamura [FY96]. In addition, from the proof of Proposition A.1, it can directly be deduced from (A.4) how the characterization of the optimal f is altered, if different boundary conditions are imposed on f. For example, if $f(t_0)$ is assumed to be fixed, this condition replaces the boundary condition $f^{(3)}(t_0) = 0$ in Propositions 2.1 and 3.1. Finally, Bjerksund and Stensland [BS96] propose the model in which equations (3.1) are replaced by

$$\underline{p}^j \leq \sum_{i \in I_j} p_i^j v_i \leq \bar{p}^j, \quad j = 1, \ldots, n, \tag{5.1}$$

where $\underline{p}^j$ denotes the bid price of bond j and $\bar{p}^j$ denotes its ask price. For the sake of brevity and clarity, we have not put problem (3.2) in this form, but we note that the discussion can be modified to cover this case. This would lead to a modified optimization problem of the form (4.1), the difference being that the inequalities (5.1) replace the equations (3.1).

Acknowledgements

I thank Magnus Nyström at Algorithmica Research AB for having initiated this research and for giving bibliographical assistance and valuable comments on practical aspects of the problem. Thanks are also due to Tomas Björk and Camilla Landén for many helpful discussions and suggestions clarifying the results presented.

References

[AvD94] K. J. Adams and D. R. van Deventer. Fitting yield curves and forward rate curves with maximum smoothness. *Journal of Fixed Income*, pages 52–62, 1994.

[BS96] P. Bjerksund and G. Stensland. Utledning av rentens terminstruktur ved "maksimum glatthet"-prinsippet. *Beta*, pages 2–6, 1996. ISSN 0801-3322 (In Norwegian).

[Fle87] R. Fletcher. *Practical Methods of Optimization*. John Wiley and Sons, Chichester, New York, Brisbane, Toronto and Singapore, second edition, 1987. ISBN 0-471-91547-5.

[FY96] V. Frishling and J. Yamamura. Fitting a smooth forward rate curve to coupon instruments. *Journal of Fixed Income*, pages 97–103, 1996.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1985. ISBN 0-521-38632-2.

[NS96] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, New York, 1996. ISBN 0-07-046065-5.

[Sch89] H. R. Schwartz. *Numerical Analysis: A Comprehensive Introduction*. John Wiley and Sons, Inc., New York, 1989. ISBN 0-471-92064-9.

[Tan97] C. Tanggaard. Nonparametric smoothing of yield curves. *Review of Quantitative Finance and Accounting*, 9, 251–267, 1997.

A. Calculus of variation

In this appendix, the analysis of Schwartz [Sch89, pp. 126–128] is reviewed and applied to problems (1.3) and (3.2). This is straightforward if the primitive function $F(t) = \int_{t_0}^{t} f(s)ds$ is introduced. The results for problem (1.3) follow immediately from the results of Schwartz, whereas the following proposition gives the required generalization of Schwartz' analysis to the case when the values of v_i are not unique.

Proposition A.1. *Consider the problem defined by*

$$\begin{array}{ll} \underset{\substack{F\in C^2[t_0,t_m] \\ v\in \mathbb{R}^m}}{\text{minimize}} & \int_{t_0}^{t_m} (F^{(3)}(s))^2 ds \\ \text{subject to} & F(t_i) = -\ln v_i, \quad i = 0,\dots,m, \\ & v_0 = 1, \\ & \sum_{i=1}^{m} p_i^j v_i = p^j, \quad j = 1,\dots,n, \\ & F \in C^6[t_{i-1}^+, t_i^-], \quad i = 1,\dots,m. \end{array} \tag{A.1}$$

If F and v_i, $i = 1,\dots,m$, form a local minimizer to (A.1)*, then*

1. $F^{(3)}(t_0) = F^{(3)}(t_m) = 0, \quad F^{(3)}(t_i^-) = F^{(3)}(t_i^+),\ i = 1,\dots,m-1.$
2. $F^{(4)}(t_0) = F^{(4)}(t_m) = 0, \quad F^{(4)}(t_i^-) = F^{(4)}(t_i^+),\ i = 1,\dots,m-1.$
3. *There is a $\lambda \in \mathbb{R}^n$ such that*

$$F^{(5)}(t_i^-) - F^{(5)}(t_i^+) - v_i \sum_{j=1}^{n} p_i^j \lambda_j = 0, \quad i = 1,\dots,m-1,$$

$$F^{(5)}(t_m) - v_m \sum_{j=1}^{n} p_m^j \lambda_j = 0.$$

4. $F^{(6)}(t) = 0,\ t_{i-1} < t < t_i,\ i = 1,\dots,m.$

Moreover, if the equations $\sum_{i=1}^{m} p_i^j v_i = p^j$, $j = 1,\dots,n$, have a unique solution, then the above conditions are sufficient for F and v_i, $i = 1,\dots,m$, to form a global minimizer to (A.1)*.*

Proof. Let F and v_i, $i = 1,\dots,m$, form a locally optimal solution to (A.1). Let $G \in C^2[t_0,t_m]$ be such that $G \in C^6[t_{i-1}^+, t_i^-]$, $i = 1,\dots,m$, $G(t_0) = 0$ and $\sum_{i=1}^{m} p_i^j v_i G(t_i) = 0$, $j = 1,\dots,n$. For τ near zero, define

$$F_\tau(t) = F(t) - \ln(1 + \tau G(t)). \tag{A.2}$$

Then $F_\tau(t_0) = 0$ and $F_\tau(t_i) = -\ln(v_i + \tau v_i G(t_i))$, $i = 1,\dots,m$. Moreover, $\sum_{i=1}^{m} p_i^j (v_i + \tau v_i G(t_i)) = p^j$, $j = 1,\dots,n$. Hence, for τ near zero, F_τ and $v_i + \tau v_i G(t_i)$,

$i = 1, \ldots, m$, form a feasible solution to (A.1). A Taylor-series expansion of (A.2) gives

$$\int_{t_0}^{t_m} (F_\tau^{(3)}(s))^2 ds = \int_{t_0}^{t_m} (F^{(3)}(s))^2 ds - 2\tau \int_{t_0}^{t_m} F^{(3)}(s) G^{(3)}(s) ds + o(\tau),$$

where $o(\tau)$ is a remainder term such that $\lim_{\tau \to 0}(o(\tau)/\tau) = 0$. Consequently, it follows that necessary conditions for F and v_i, $i = 1, \ldots, m$, to form a local minimizer to (A.1) is that

$$\int_{t_0}^{t_m} F^{(3)}(s) G^{(3)}(s) ds = 0. \tag{A.3}$$

Integration of (A.3) by parts, taking into account that $G \in C^2[t_0, t_m]$ and $F \in C^6[t_{i-1}^+, t_i^-]$, $i = 1, \ldots, m$, gives

$$\begin{aligned} 0 = \int_{t_0}^{t_m} F^{(3)}(s) G^{(3)}(s) ds &= \sum_{i=1}^{m} \int_{t_{i-1}}^{t_i} F^{(3)}(s) G^{(3)}(s) ds \\ &= F^{(3)}(t_m^-) G''(t_m) - F^{(3)}(t_0^+) G''(t_0) \\ &+ \sum_{i=1}^{m-1} (F^{(3)}(t_i^-) - F^{(3)}(t_i^+)) G''(t_i) \\ &- F^{(4)}(t_m^-) G'(t_m) + F^{(4)}(t_0^+) G'(t_0) \\ &- \sum_{i=1}^{m-1} (F^{(4)}(t_i^-) - F^{(4)}(t_i^+)) G'(t_i) \\ &+ F^{(5)}(t_m^-) G(t_m) - F^{(5)}(t_0^+) G(t_0) \\ &+ \sum_{i=1}^{m-1} (F^{(5)}(t_i^-) - F^{(5)}(t_i^+)) G(t_i) \\ &- \sum_{i=1}^{m} \int_{t_{i-1}}^{t_i} F^{(6)}(s) G(s) ds. \end{aligned} \tag{A.4}$$

Since there are only constraints on $G(t_i)$, $i = 0, \ldots, m$, conditions 1, 2 and 4 follow. From the values of $G(t_i)$, $i = 0, \ldots, m$, it follows that

$$F^{(5)}(t_m^-) G(t_m) + \sum_{i=1}^{m-1} (F^{(5)}(t_i^-) - F^{(5)}(t_i^+)) G(t_i) = 0,$$

where $G(t_i)$ are arbitrary numbers such that $\sum_{i=1}^m p_i^j v_i G(t_i) = 0$, $j = 1, \ldots, n$. Since the range space of a matrix is the orthogonal complement to the null space of its transpose, this is equivalent to condition 3, see, e.g., Horn and Johnson [HJ85, pp. 16-17].

Finally, if v_i, $i = 1, \ldots, m$, are uniquely determined by the equations $\sum_{i=1}^m p_i^j v_i = p^j$, $j = 1, \ldots, n$, let F be a feasible solution such that the above conditions hold, and let $\tilde{F}$ be any other feasible solution. Then, with $G(t) = \tilde{F}(t) - F(t)$, it follows that $G(t_i) = 0$, $i = 0, \ldots, m$. Hence, for this choice of G, the expansion of (A.4) in conjunction with conditions 1, 2 and 4 imply that (A.3) holds, and it follows that

$$\int_{t_0}^{t_m} (\tilde{F}^{(3)}(s))^2 ds = \int_{t_0}^{t_m} (F^{(3)}(s))^2 ds + \int_{t_0}^{t_m} (G^{(3)}(s))^2 ds \geq \int_{t_0}^{t_m} (F^{(3)}(s))^2 ds.$$

Consequently, F is a global minimizer. ∎

This proposition is the key to the analysis, in that the proofs of Propositions 2.1 and 3.1 follow directly. The proof of Proposition 2.1 is a review of the proof given by Schwartz [Sch89, pp. 126–128]. It is given here for the sake of completeness, since it appears as a by-product of the slightly more general result of Proposition A.1.

Proof of Proposition 2.1. The integral in the objective function may be split into two parts, as

$$\int_{t_0}^{t_{m+1}} (f''(s))^2 ds = \int_{t_0}^{t_m} (f''(s))^2 ds + \int_{t_m}^{t_{m+1}} (f''(s))^2 ds.$$

There is no constraint other than continuity of f and f' at t_m on the second term. Hence, f is linear on this part, determined by $f(t_m)$ and $f'(t_m)$. For the first term, let $F(t) = \int_{t_0}^{t} f(s)ds$. The result then follows from Proposition A.1. Condition 3 of Proposition A.1 is superfluous, since the values of v_i are unique. ∎

Proof of Proposition 3.1. Let $F(t) = \int_{t_0}^{t} f(s)ds$. The result then follows from Proposition A.1. ∎

B. Alternative nonlinear equations

Similar to the equations derived by Schwartz [Sch89, pp. 128–130], an alternative system of nonlinear equations, based on Proposition 2.1, can be derived, that characterizes the first-order optimality conditions of (3.2). Let $F(t) = \int_{t_0}^{t} f(s)ds$. For $i = 1, \ldots, m$, let $f_i(t)$ be defined by (2.1) on the interval $[t_{i-1}^+, t_i^-]$ and let $h_i = (t_i - t_{i-1})$. We want the condition $F(t_i) = -\ln v_i$ to hold for $i = 0, \ldots, m$. To simplify the notation, $F(t_i)$ is denoted by y_i initially. A substitution $y_i = -\ln v_i$ will be made later. Then, we have for $i = 1, \ldots, m$,

$$\begin{aligned}
F_i(t_{i-1}) &= y_{i-1}, & F_i(t_i) &= \frac{a_i h_i^5}{5} + \frac{b_i h_i^4}{4} + \frac{c_i h_i^3}{3} + \frac{d_i h_i^2}{2} + e_i h_i + y_{i-1},\\
F_i'(t_{i-1}) &= e_i, & F_i'(t_i) &= a_i h_i^4 + b_i h_i^3 + c_i h_i^2 + d_i h_i + e_i,\\
F_i''(t_{i-1}) &= d_i, & F_i''(t_i) &= 4a_i h_i^3 + 3b_i h_i^2 + 2c_i h_i + d_i,\\
F_i^{(3)}(t_{i-1}) &= 2c_i, & F_i^{(3)}(t_i) &= 12a_i h_i^2 + 6b_i h_i + 2c_i,\\
F_i^{(4)}(t_{i-1}) &= 6b_i, & F_i^{(4)}(t_i) &= 24a_i h_i + 6b_i.
\end{aligned}$$

In addition to $y_i = F(t_i)$, we may introduce $y_i'' = F''(t_i)$ and $y_i^{(4)} = F^{(4)}(t_i)$ for $i = 0, \ldots, m$. If the function values as well as second- and fourth-order derivatives are required to be continuous, we may express a_i, b_i, c_i, d_i, and e_i in these quantities for $i = 1, \ldots, m$ as

$$\begin{aligned}
a_i &= \frac{1}{24h_i}(y_i^{(4)} - y_{i-1}^{(4)}),\\
b_i &= \frac{1}{6}y_{i-1}^{(4)},\\
c_i &= \frac{1}{2h_i}(y_i'' - y_{i-1}'') - \frac{h_i}{12}(y_i^{(4)} + 2y_{i-1}^{(4)}),\\
d_i &= y_{i-1}'',\\
e_i &= \frac{1}{h_i}(y_i - y_{i-1}) - \frac{h_i}{6}(y_i'' + 2y_{i-1}'') + \frac{h_i^3}{360}(7y_i^{(4)} + 8y_{i-1}^{(4)}).
\end{aligned}$$

The continuity of the first- and third-order derivatives must be handled too. From above,

$$\begin{aligned}
F_i'(t_{i-1}) &= \frac{1}{h_i}(y_i - y_{i-1}) - \frac{h_i}{6}(y_i'' + 2y_{i-1}'') + \frac{h_i^3}{360}(7y_i^{(4)} + 8y_{i-1}^{(4)})\\
F_i'(t_i) &= \frac{1}{h_i}(y_i - y_{i-1}) + \frac{h_i}{6}(2y_i'' + y_{i-1}'') - \frac{h_i^3}{360}(8y_i^{(4)} + 7y_{i-1}^{(4)}),\\
F_i^{(3)}(t_{i-1}) &= \frac{1}{h_i}(y_i'' - y_{i-1}'') - \frac{h_i}{6}(y_i^{(4)} + 2y_{i-1}^{(4)}),\\
F_i^{(3)}(t_i) &= \frac{1}{h_i}(y_i'' - y_{i-1}'') + \frac{h_i}{6}(2y_i^{(4)} + y_{i-1}^{(4)}),
\end{aligned}$$

for $i = 1, \ldots, m$. Requiring continuity gives

$$\begin{aligned}
F'(t_i) &= \frac{1}{h_{i+1}}(y_{i+1} - y_i) - \frac{h_{i+1}}{6}(y_{i+1}'' + 2y_i'') + \frac{h_{i+1}^3}{360}(7y_{i+1}^{(4)} + 8y_i^{(4)})\\
&= \frac{1}{h_i}(y_i - y_{i-1}) + \frac{h_i}{6}(2y_i'' + y_{i-1}'') - \frac{h_i^3}{360}(8y_i^{(4)} + 7y_{i-1}^{(4)}),\\
F^{(3)}(t_i) &= \frac{1}{h_{i+1}}(y_{i+1}'' - y_i'') - \frac{h_{i+1}}{6}(y_{i+1}^{(4)} + 2y_i^{(4)})\\
&= \frac{1}{h_i}(y_i'' - y_{i-1}'') + \frac{h_i}{6}(2y_i^{(4)} + y_{i-1}^{(4)}),
\end{aligned}$$

for $i = 1, \ldots, m-1$. This gives $2(m-1)$ equations as

$$\begin{aligned}
&-7h_i^3 y_{i-1}^{(4)} - 8(h_i^3 + h_{i+1}^3)y_i^{(4)} - 7h_{i+1}^3 y_{i+1}^{(4)}\\
&+ 60h_i y_{i-1}'' + 120(h_i + h_{i+1})y_i'' + 60h_{i+1}y_{i+1}''\\
&- \frac{360}{h_i}y_{i-1} + 360\left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right)y_i - \frac{360}{h_{i+1}}y_{i+1} = 0, \quad i = 1, \ldots, m-1, \quad \text{(B.1a)}\\
&h_i y_{i-1}^{(4)} + 2(h_i + h_{i+1})y_i^{(4)} + h_{i+1}y_{i+1}^{(4)}\\
&- \frac{6}{h_i}y_{i-1}'' + 6\left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right)y_i'' - \frac{6}{h_{i+1}}y_{i+1}'' = 0, \quad i = 1, \ldots, m-1. \quad \text{(B.1b)}
\end{aligned}$$

The boundary conditions on $F^{(3)}$ and $F^{(4)}$ give

$$F^{(3)}(t_0) = \frac{1}{h_1}(y_1'' - y_0'') - \frac{h_1}{6}(y_1^{(4)} + 2y_0^{(4)}) = 0, \tag{B.2a}$$

$$F^{(3)}(t_m) = \frac{1}{h_m}(y_m'' - y_{m-1}'') + \frac{h_m}{6}(2y_m^{(4)} + y_{m-1}^{(4)}) = 0, \tag{B.2b}$$

$$y_0^{(4)} = 0, \tag{B.2c}$$

$$y_m^{(4)} = 0, \tag{B.2d}$$

i.e., four additional constraints.

If we replace y_i by $-\ln v_i$ for $i = 0, \ldots, m$ in (B.1), and add together (B.1), (B.2), the m constraints imposed by condition 3 of Proposition 3.1, plus the additional constraints $\sum_{i=1}^{m} p_i^j v_i = p^j$, $j = 1, \ldots, n$, and $v_0 = 1$, imposed by feasibility in (3.2), a total of $3(m+1)+n$ equations results for the $3(m+1)+n$ unknowns, v_i, y_i'' and $y_i^{(4)}$, $i = 0, \ldots, m$, and λ_j, $j = 1, \ldots, n$. Such a system of equations may be written on the form

$$\begin{aligned}
v_0 &= 1,\\
2h_1 y_0^{(4)} + h_1 y_1^{(4)} + \frac{6}{h_1} y_0'' - \frac{6}{h_1} y_1'' &= 0,\\
y_0^{(4)} &= 0,\\
-7h_i^3 y_{i-1}^{(4)} - 8(h_i^3 + h_{i+1}^3) y_i^{(4)} - 7h_{i+1}^3 y_{i+1}^{(4)} &\\
+ 60 h_i y_{i-1}'' + 120(h_i + h_{i+1}) y_i'' + 60 h_{i+1} y_{i+1}'' &\\
+ \frac{360}{h_i} \ln v_{i-1} - 360\left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right) \ln v_i + \frac{360}{h_{i+1}} \ln v_{i+1} &= 0, \quad i = 1, \ldots, m-1,\\
h_i y_{i-1}^{(4)} + 2(h_i + h_{i+1}) y_i^{(4)} + h_{i+1} y_{i+1}^{(4)} &\\
- \frac{6}{h_i} y_{i-1}'' + 6\left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right) y_i'' - \frac{6}{h_{i+1}} y_{i+1}'' &= 0, \quad i = 1, \ldots, m-1,\\
-\frac{1}{h_i} y_{i-1}^{(4)} + \left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right) y_i^{(4)} - \frac{1}{h_{i+1}} y_{i+1}^{(4)} - v_i \sum_{j=1}^{n} p_i^j \lambda_j &= 0, \quad i = 1, \ldots, m-1,\\
h_m y_{m-1}^{(4)} + 2h_m y_m^{(4)} - \frac{6}{h_m} y_{m-1}'' + \frac{6}{h_m} y_m'' &= 0,\\
y_m^{(4)} &= 0,\\
-\frac{1}{h_m} y_{m-1}^{(4)} + \frac{1}{h_m} y_m^{(4)} - v_m \sum_{j=1}^{n} p_m^j \lambda_j &= 0,\\
\sum_{i=1}^{m} p_i^j v_i &= p^j, \quad j = 1, \ldots, n.
\end{aligned}$$

The associated Newton equations are sparse with a structured sparsity pattern. However, our limited computational experiments suggest that they may be ill-conditioned.

Addendum to *A note on maximum-smoothness approximation of forward interest rate*

After completion of the report, it has been brought to our attention by Donald R. van Deventer that the incorrect result of Adams and van Deventer [AvD94], regarding the shape of the fourth-order polynomials that give an optimal solution to (1.3), has been corrected by van Deventer and Imai [vDI97, Chapter 2].

Unfortunately, the name Schwarz is misspelled throughout the report.

Reference to correction

[vDI97] D. R. van Deventer and K. Imai. *Financial Risk Analytics: A Term Structure Model Approach for Banking, Insurance and Investment Management.* Irwin Professional, Chicago, IL, 1997. ISBN 0-7863-0964-4.