

CONTROL THEORETIC SMOOTHING SPLINES ARE APPROXIMATE LINEAR FILTERS

Y. ZHOU*, W. DAYAWANSA†, AND C. MARTIN†

Abstract. The problem of constructing and approximating control theoretic smoothing splines is considered in this paper. It is shown that the optimal approximating function can be given as the solution of a forced Hamiltonian system, that can be explicitly solved using the Riccati transform, and an explicit linear filter can be constructed. We show that the bandwidth of the filter can be naturally controlled and thus for control theoretic smoothing splines the far past and the far future are unimportant. Hence smoothing splines are “local” in nature rather than “global”. We conclude that while spline approximations are not causal the far future is not important.

Key words: control theoretic smoothing splines, feedback, kernel approximation, forced Hamiltonian system

1. Introduction. In this paper we consider the problem of constructing and approximating control theoretic smoothing splines. Control theoretic splines have proven to be useful in trajectory planning for aircraft, [1], and in various trajectory planning problems in robotics, [2]. The general theory of control theoretic splines and numerous variations on the theme have appeared in [3, 4, 5]. Control theoretic splines are related to the theory of L-splines developed originally by R. Varga and M. Shultz, [6], and in the context of control theory by [7]. The theory of smoothing splines is based on the seminal work of G. Wahba. The monograph, [8], is used implicitly throughout this paper.

It is a well understood in folklore of statistics that polynomial smoothing splines act as smoothing filters on noisy data and that they are in some sense band limited, i.e., small changes in one data point has primary effect on the spline in a neighborhood of that point. This was studied explicitly by B. Silverman in [9] for the cubic spline. In this paper we construct an approximate linear filter for control theoretic splines. The construction is based on linear-quadratic optimal control and related filtering and tracking results. We use a quadratic cost function that contains the function representing the data to be approximated and the cost function is minimized subject to the constraints of the control system that is being used to generate the approximating curve. In Section 2 we state the central problem of the paper: How does the limiting value of the control theoretic spline compare to the spline itself? We also give a brief derivation of the control theoretic spline. In Section 3 we construct the optimal filter as an operator and show existence and uniqueness. We also prove some basic

*Department of Mathematics, Stockholm University, SE-10693, Stockholm, E-mail: yishao@math.su.se

†Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409, E-mail: daya@math.ttu.edu, martin@math.ttu.edu

properties of the filter. In Section 4 we show that the filter is given as a solution to a forced Hamiltonian where the forcing is the rough data. We show by means of the Riccati transformation that constructing the filter reduces to solving and storing the solution to a standard Riccati equation. In Section 5 we have the main technical result of the paper. We show that the construction can be reduced to 2 numerically feasible problems. The first is to construct the matrix exponential of the symplectic matrix and the second is to solve the algebraic Riccati equation for the positive definite solution. The explicit form of the operator is given. In Section 6 we apply the results of the previous sections to construct the filter corresponding to the cubic spline. The cubic spline is the standard for approximation. We show that the bandwidth of the filter is controlled by the weighting factor “ λ ”. Thus the spline and the filter ignore far future and far past values of the data. We compare our results to the results of Silverman and see that our filter is somewhat better than his approximation.

The main contributions of this paper are to 1) show that the control theoretic smoothing splines are a discrete approximation of an integral linear filter, 2) obtain an explicit, well motivated linear filter, [9], 3) show that the parameter λ (to be defined in Section 2) controls bandwidth and hence can be used to gain long term information or can be used to control the degree of approximation of the smoothing spline to the discrete numerical data, 4) show that splines are in fact “local” approximations rather than “global” as is emphasized in the numerical literature, and 5) show that while splines are not causal they depend primarily on the next few data points, not on the entire future.

2. Problem Formulation. In this section we state the basic assumptions of the paper and formulate the problems to be solved. For the sake of completeness we briefly derive the control used to generate the smoothing spline. A detailed derivation is found in [4].

We consider the control system

$$(2.1) \quad \dot{x} = Ax + bu$$

$$(2.2) \quad y = c'x$$

where $x \in \mathbb{R}^n$, $u, y \in \mathbb{R}$ and A, b and c are constant matrices of compatible dimension and $x(0) = x_0$. We further assume that the system is controllable and observable. Because we are primarily interested in approximation rather than control we make the assumption that

$$(2.3) \quad c'b = c'Ab = \dots = c'A^{n-2}b = 0.$$

This assumption simplifies some calculations and is natural for spline approximation. The goal of spline approximation is to obtain an approximation of a set of discrete data points that is real analytic between the data points and is as smooth as possible

at the data points. The assumption of equation (2.3) gives us the maximal smoothness at the data points. This is explicitly discussed in [4].

We assume as given a data set of the form

$$(2.4) \quad D_N = \{(t_{i,N}, \alpha_{i,N}) : 0 < t_{1,N} \leq t_{2,N} \leq \dots \leq t_{N,N} < T\}$$

and we make no assumptions about how the data was generated. That is, we do not assume that the data is generated by a system of the form $dx = Fxdt + bdW$ where dW is a probabilistic measure. We assume that T is fixed and finite for without this assumption the integrals do not exist. We assume that there exists a function $g_N \in C[0, T]$ so that $\alpha_{i,N} = g_N(t_{i,N}) + \epsilon_{i,N}$ where $\epsilon_{i,N}$ is a symmetrically distributed random variable. Let $S_N(t)$ be a smooth, piecewise polynomial function such that

$$S_N(t_{i,N}) = \alpha_{i,N},$$

i.e., S_N is an interpolating polynomial spline function. We assume that the data is such that there exists a function $f \in L_2[0, T]$ so that

$$(2.5) \quad \lim_{N \rightarrow \infty} \|f - S_N\|_2 = 0$$

and we further assume that there exists a function $g \in C[0, T]$ such that the sequence $\{g_N\}$ converges uniformly to g . These assumptions are evidently weaker than assuming that the data is generated by the output of a dynamical system driven by noise.

We assume as given a cost function of the form

$$(2.6) \quad J_N(u) = \lambda \int_0^T u^2(t) dt + \sum_{i=1}^N w_{i,N} (y(t_{i,N}) - \alpha_{i,N})^2.$$

We allow $u \in L_2[0, T]$. The choice of λ in this formula is very important. In the statistical literature the choice of λ is governed by the need for the differences between the spline and the data to be iid normal or that the approximation to the data is close. We will see that the choice of λ determines the bandwidth of the filter and hence whether the filter picks up the long term behavior of the data or if it closely approximates the data points. We assume that $w_{i,N} > 0$ for all indices i, N and a critical assumption is that for every $h \in C[0, T]$

$$(2.7) \quad \lim_{N \rightarrow \infty} \left| \int_0^T h(t) dt - \sum_{i=1}^N w_{i,N} h(t_{i,N}) \right| = 0,$$

that is, the sampling times and weights form a convergent quadrature algorithm. This assumption is satisfied as long as the weights are chosen to be the weights associated with a convergent quadrature scheme. See for example, [10].

We define a second cost function in terms of the function f of equation (2.5).

$$(2.8) \quad J(u) = \int_0^T \lambda u^2(t) + (y(t) - f(t))^2 dt.$$

We now formulate two optimal control problems. The first will produce the control that drives the output of the linear system to the control theoretic smoothing spline and the second will produce an approximation to the spline function and will be the object of interest for this paper.

PROBLEM 1. *The first problem is to minimize $J_N(u)$ in order to construct the classical smoothing splines,*

$$\min_{u \in L_2[0, T]} J_N(u)$$

subject to the constraints of the system (2.1) and (2.2).

PROBLEM 2. *The second problem is to minimize $J(u)$ to obtain the limiting value of the classical smoothing splines.*

$$\min_{u \in L_2[0, T]} J(u)$$

subject to the constraints of the system (2.1) and (2.2).

We define the function $g_t(s)$ as

$$(2.9) \quad g_t(s) = \begin{cases} c' e^{A(t-s)} b & t > s \\ 0 & t \leq s. \end{cases}$$

Following Luenberger [11], we calculating the derivative of $J_N(u)$ in the direction of w and setting the derivative equal to zero we see that a necessary condition for the optimal control is given by

$$\lambda u(s) + \sum_{i=1}^N w_{i,N} (y(t_{i,N}) - \alpha_{i,N}) g_{t_{i,N}}(s) = 0.$$

Now $y(t_{i,N})$ is a function of u . However J_N is convex and hence has a unique minimum. Thus we see that the optimal control must be of the form

$$u(s) = \sum_{i=1}^N \tau_i g_{t_{i,N}}(s)$$

and substituting this into $J_N(u)$ and doing some routine manipulation we have the finite dimensional problem of minimizing

$$J_N(\tau) = \lambda \tau' G \tau + \tau' G W_N G \tau - 2 \tau' G W_N \alpha_N + \text{constant}$$

where G is the Grammian formed from the linearly independent functions $g_{t_{i,N}}(s)$, τ is the vector of unknown coefficients τ_i , W_N is the diagonal matrix with the weights

$w_{i,N}$ on the diagonal and α_N is the vector of data. Calculating the derivative of $J_N(\tau)$ with respect to τ we have the necessary and sufficient condition for the optimal τ to be the solution of

$$\lambda G\tau + GW_N G\tau - GW_N \alpha_N = 0$$

and since G and W_N are non-singular we have that τ is the solution of

$$(2.10) \quad (\lambda W_N^{-1} + G)\tau = \alpha_N.$$

Let $u_N(t)$ be the unique solution to Problem (1) and let u be the unique solution to problem (2). It was shown in [12] that the sequence $\{u_N(t)\}$ converges to u in a point wise manner. Thus the solution of Problem (2) is an approximation to the control theoretic spline of Problem (1). A major goal in this paper is to find a solution to Problem (2) so that

$$(2.11) \quad y(t) = \int_0^T k(t,s)f(s)ds.$$

To simplify the exposition we will assign $w_i = 1$ in the rest of the paper.

3. Reduction to an Operator Equation. In this section we examine some of the properties of the filter from the somewhat formal view of operator theory. Basic properties of the filter are obtained in this manner as well as existence and uniqueness.

For typographical convenience define an operator

$$(3.1) \quad L_t(u) = \int_0^T g_t(s)u(s)ds.$$

It is clear that $y(t) = L_t(u) + c'e^{At}x_0$ and that by replacing the function f by $f(t) - c'e^{At}x_0$ the problem remains unchanged. In this section we will assume without loss of generality that $x_0 = 0$. We begin by calculating the Gateaux derivative of $DJ(u)(w)$. We have after a simple calculation

$$(3.2) \quad DJ(u)(w) = 2 \int_0^T \lambda u(t)w(t) + (L_t(u) - f(t))L_t(w)dt$$

and calculating the second derivative with respect to u and evaluating at w we have

$$(3.3) \quad D^2(u)(w) = 2 \int_0^T \lambda w^2(t) + L_t(w)^2 dt.$$

From this we see that the second derivative is nonnegative and is 0 if and only if $w(t) = 0$. Thus the functional is convex and hence has a unique minimum.

We now return to equation (3.2) and set it equal to zero to obtain a necessary and sufficient condition for optimality. After some manipulation

$$\begin{aligned} 0 &= \int_0^T \lambda u(t)w(t) + (L_t(u) - f(t))L_t(w)dt \\ &= \int_0^T \left[\lambda u(s) + \int_s^T (L_t(u) - f(t))g_t(s)dt \right] w(s)ds. \end{aligned}$$

Now this expression is 0 for all w and hence we have that the optimal u satisfies the integral equation

$$(3.4) \quad \lambda u(s) + \int_s^T (L_t(u) - f(t))g_t(s)dt = 0.$$

Multiplying this expression by $g_t(s)$ and integrating we have

$$\lambda y(t) + \int_0^t g_t(s) \int_s^T (y(r) - f(r))g_r(s)drds = 0$$

and after a little reorganization we have

$$(3.5) \quad \lambda y(t) + \int_0^t \int_s^T g_t(s)g_r(s)y(r)drds = \int_0^t \int_s^T g_t(s)g_r(s)f(r)drds.$$

We now define the operator K as

$$(3.6) \quad K(g) = \int_0^t \int_s^T g_t(s)g_r(s)g(r)drds$$

for $g \in L^2[0, T]$. Now for every $g \in L^2[0, T]$ $K(g)$ is smooth and hence in $L^2[0, T]$. Rewriting equation (3.5) we have

$$(3.7) \quad (\lambda I + K)(y) = K(f).$$

LEMMA 3.1. *The operator K is self adjoint.*

Proof. We prove the lemma by direct calculation. After substitution and interchanging the order of integration we have

$$\begin{aligned} \langle w, Ku \rangle &= \int_0^T w(t)K(u)(t)dt \\ &= \int_0^T \left[\int_s^T w(t)g_t(s)dt \right] \left[\int_s^T g_r(s)u(r)dr \right] ds \\ &= \langle Kw, u \rangle. \end{aligned}$$

We now decompose K as the sum of two operators by changing the order of integration. An elementary calculation shows that

$$(3.8) \quad K(u) = \int_0^t \left[\int_0^r g_t(s)g_r(s)ds \right] u(r)dr + \int_t^T \left[\int_0^t g_t(s)g_r(s)ds \right] u(r)dr.$$

Now define operators F and B (forward and backward) as

$$(3.9) \quad F(u) = \int_0^t \left[\int_0^r g_t(s)g_r(s)ds \right] u(r)dr$$

and

$$(3.10) \quad B(u) = \int_t^T \left[\int_0^t g_t(s)g_r(s)ds \right] u(r)dr.$$

Note that F and B are bounded and hence K is bounded. Also note that from the proof of the lemma the operator K is positive. Thus the spectrum of K is bounded below by 0 and hence the spectrum of $I + K$ is bounded away from 0 and the operator $I + K$ is thus injective.

LEMMA 3.2. *For $\lambda > 0$, the operator $\lambda I + K$ is one to one and onto.*

Proof. It only remains to prove that $I + K$ is onto. Suppose otherwise. Then there exists a function $x \in L^2[0, T]$ such that for all $y \in L^2[0, T]$ $\langle x, (I + K)(y) \rangle = 0$. We use the fact that the operator is self adjoint to conclude that $(I + K)x = 0$. This is equivalent to the fact that x is unique solution to the optimal control problem with cost function

$$J(u) = \int_0^T u^2(t) + y^2(t) dt$$

subject to the constraint of the system defined by equations (2.1) and (2.2). However, it is easy to see that the optimal control is identically 0 and hence that the corresponding $y(t)$ is identically zero. Thus we conclude that $x = 0$ and hence $\lambda I + K$ is onto.

We can thus solve equation (3.7) to obtain

$$(3.11) \quad y = (\lambda I + K)^{-1} K(f).$$

In the next section we will explicitly construct a representation of the operator $(\lambda I + K)^{-1} K$ in terms of an associated Riccati equation.

4. The optimal control problem. In this section we return to the optimal control problem to obtain a different representation of the operator K . With the representation we will obtain the detailed properties of the operator.

We return to equation (3.4). This representation of the optimal control can be rewritten as

$$u(t) = - \int_s^T \frac{1}{\lambda} c' e^{A(t-s)} b(y(t) - f(t)) dt$$

and in this form we see that it is in dynamic feedback form and is to be fed back through the system adjoint to the original system (2.1) and (2.2). Our first goal is to explicitly write out the relationship between the system and its adjoint.

We begin by letting

$$(4.1) \quad \ell(s) = \int_s^T e^{A'(t-s)} c(y(t) - f(t)) dt$$

where we have replaced $c' e^{A(t-s)} b$ by its transpose. We calculate the derivative of ℓ to obtain

$$\begin{aligned} \dot{\ell}(s) &= -A' \ell(s) - c(y(s) - f(s)) \\ &= -A' \ell(s) - c' x + c f(s) \end{aligned}$$

where we have used the fact from equation (2.2) that $y = c'x$. We now see that

$$(4.2) \quad u(s) = -\frac{1}{\lambda}b'\ell(s).$$

Now from equations (2.1) and (2.2) we have

$$\begin{aligned} \dot{x}(s) &= Ax(s) + bu(s) \\ &= Ax(s) - \frac{1}{\lambda}bb'\ell(s). \end{aligned}$$

From the definition of ℓ we have

$$\ell(T) = 0$$

and

$$x(0) = x_0.$$

Writing this in the more conventional form of a forced Hamiltonian system we have

$$(4.3) \quad \frac{d}{dt} \begin{pmatrix} x \\ \ell \end{pmatrix} = \begin{pmatrix} A & -\frac{1}{\lambda}bb' \\ -cc' & -A' \end{pmatrix} \begin{pmatrix} x \\ \ell \end{pmatrix} + \begin{pmatrix} 0 \\ c \end{pmatrix} f$$

with boundary conditions

$$(4.4) \quad \ell(T) = 0 \text{ and } x(0) = x_0.$$

Thus from the solution of this problem we can explicitly construct the approximate spline $y(t)$. Note that the symplectic matrix is the matrix associated with the Hamiltonian for the optimal control problem

$$(4.5) \quad \min_u \int_0^T y(t)^2 + u(t)^2 dt$$

subject to the constraint

$$\dot{x}(t) = Ax(t) + bu(t), \quad y(t) = cx(t).$$

To solve the two point boundary value problem we introduce the Riccati transform.

$$(4.6) \quad \begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} I & 0 \\ -P(t) & I \end{pmatrix} \begin{pmatrix} x \\ \ell \end{pmatrix}.$$

Applying this change of basis to the two point boundary value problem we have after a considerable amount of matrix multiplication

$$\frac{d}{dt} \begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} A - \frac{1}{\lambda}bb'P(t) & -\frac{1}{\lambda}bb' \\ R(t) & -(A - \frac{1}{\lambda}bb'P(t))' \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} + \begin{pmatrix} 0 \\ c \end{pmatrix} f(t)$$

where

$$R(t) = -\dot{P} - PA - cc' + P\frac{1}{\lambda}bb'P - A^T P.$$

We set $R(t) = 0$ and assign it the terminal value of $P(T) = 0$. Under the conditions we have imposed of observability and controllability of the original system this Riccati equation has a unique solution on the interval $[0, T]$. We thus have the following system of equations to solve

$$(4.7) \quad \dot{P} = -PA - cc' + P\frac{1}{\lambda}bb'P - A^T P, \quad P(T) = 0$$

$$(4.8) \quad \dot{w} = -\left(A - \frac{1}{\lambda}bb'P(t)\right)'w + cf, \quad w(T) = 0$$

$$(4.9) \quad \dot{x} = \left(A - \frac{1}{\lambda}bb'P(t)\right)x - \frac{1}{\lambda}bb'w, \quad x(0) = x_0.$$

We begin by solving and storing the solution of the Riccati equation and substituting this into equation (4.8). We now have a linear time varying terminal value problem to solve. Let $\Phi(t, \tau)$ be the solution of

$$\frac{d}{dt}\Phi(t, \tau) = \left(A - \frac{1}{\lambda}bb'P(t)\right)\Phi(t, \tau)$$

with initial data given by

$$\Phi(\tau, \tau) = I$$

and let $\Psi(t, \tau)$ be the solution of

$$\frac{d}{dt}\Psi(t, \tau) = -\left(A - \frac{1}{\lambda}bb'P(t)\right)'\Psi(t, \tau)$$

with initial data given by

$$\Psi(\tau, \tau) = I.$$

So the solution of equation (4.8) is given by

$$(4.10) \quad w(t) = -\int_t^T \Psi(t, \tau)cf(\tau)d\tau$$

and the solution of equation (4.9) is given by

$$(4.11) \quad x(t) = \Phi(t, 0)x_0 - \int_0^t \Phi(t, s)\frac{1}{\lambda}bb'w(s)ds.$$

Concatenating the two solutions we have

$$(4.12) \quad y(t) = c'\Phi(t, 0)x_0 + c' \int_0^t \Phi(t, s)\frac{1}{\lambda}bb' \int_s^T \Psi(s, r)cf(r)drds.$$

Changing the order of integration we have

$$\begin{aligned}
 (4.13) \quad y(t) &= c'\Phi(t, 0)x_0 + \int_0^t \int_0^r c'\Phi(t, s) \frac{1}{\lambda} bb'\Psi(r, s) c ds f(r) dr \\
 &+ \int_t^T \int_0^t c'\Phi(t, s) \frac{1}{\lambda} bb'\Psi(r, s) c ds f(r) dr.
 \end{aligned}$$

Thus we have

$$(4.14) \quad y(t) = c'\Phi(t, 0)x_0 + \int_0^T k(t, \sigma) f(\sigma) d\sigma$$

where

$$(4.15) \quad k(t, \sigma) = \begin{cases} \frac{1}{\lambda} c' \int_0^\sigma \Phi(t, \tau) bb'\Psi(\tau, \sigma) c d\tau & 0 \leq \sigma \leq t \\ \frac{1}{\lambda} c' \int_0^t \Phi(t, \tau) bb'\Psi(\tau, \sigma) c d\tau & t \leq \sigma \leq T. \end{cases}$$

5. Simplification of the formulae. We have shown that we can give an explicit expressions for the state transition matrix in terms of the system parameters and the solution of the Riccati equation. To solve and store the entire solution of the Riccati equation is very expensive. In this section we will see that it suffice to obtain and store the initial value for the Riccati equation. This simplification is the critical technical step of the paper.

We consider the system (4.3) with $x(0) = x_0$ and $\ell(T) = 0$. By the variation of parameters formula, we obtain

$$(5.1) \quad \begin{pmatrix} x(t) \\ \ell(t) \end{pmatrix} = e^{(t-T)H} \begin{pmatrix} x(T) \\ 0 \end{pmatrix} + \int_T^t e^{(t-s)H} \begin{pmatrix} 0 \\ cf(s) \end{pmatrix} ds,$$

where

$$H = \begin{pmatrix} A & -\frac{1}{\lambda} bb' \\ -cc' & -A' \end{pmatrix}.$$

This yields the relation between boundary values $x(T)$ and $\ell(0)$,

$$(5.2) \quad \begin{pmatrix} x_0 \\ \ell(0) \end{pmatrix} = e^{-TH} \begin{pmatrix} x(T) \\ 0 \end{pmatrix} + \int_T^0 e^{-sH} \begin{pmatrix} 0 \\ cf(s) \end{pmatrix} ds.$$

Now we partition the matrix e^{tH} as follows

$$e^{tH} = \begin{pmatrix} X_1(t) & X_2(t) \\ Y_1(t) & Y_2(t) \end{pmatrix}$$

where $X_i, Y_i, i = 1, 2$ are $n \times n$ matrices. Due to the semigroup properties we have the following identities which will be used later.

$$\begin{aligned}
 (5.3) \quad X_1(t-s) &= X_1(t)x_1(-s) + X_2(t)Y_1(-s) \\
 0 &= X_1(t)X_2(-t) + X_2(t)Y_2(-t).
 \end{aligned}$$

The unique positive definite solution to the Riccati equation is given by

$$(5.4) \quad P(t) = Y_1(t-T)X_1(t-T)^{-1}$$

using the standard Hamiltonian argument.

Then, solving the first block of equations in (5.2) yields

$$(5.5) \quad x(T) = X_1(-T)^{-1} \left(x_0 + \int_0^T X_2(-s)cf(s)ds \right)$$

which, together with (5.1), leads to the following expression

$$(5.6) \quad x(t) = X_1(t-T)X_1(-T)^{-1} \left(x_0 + \int_0^T X_2(-s)cf(s)ds \right) + \int_t^T X_2(t-s)cf(s)ds.$$

By the identities from (5.3) and (5.4) we have

$$\begin{aligned} x(t) &= (X_1(t) + X_2(t)P(0))x_0 + \int_0^T (X_1(t + X_2(t)P(0))X_2(-s)cf(s)ds \\ &\quad - \int_t^T (X_1(s)X_2(-s) + X_2(t)Y_2(-s)cf(s)ds \\ &= (X_1(t) + X_2(t)P(0))x_0 + \int_0^t (X_1(t + X_2(t)P(0))X_2(-s)cf(s)ds \\ &\quad + \int_t^T ((X_1(t) + X_2(t)P(0))X_2(-s) - (X_1(s)X_2(-s) + X_2(t)Y_2(-s)))cf(s)ds \\ &= (X_1(t) + X_2(t)P(0))x_0 + \int_0^t (X_1(t + X_2(t)P(0))X_2(-s)cf(s)ds \\ &\quad + \int_t^T (X_2(t)(P(0)X_2(-s) - Y_2(-s))cf(s)ds. \end{aligned}$$

Finally, we obtain the kernel

$$k(t, \sigma) = \begin{cases} c'(X_1(t) + X_2(t)P(0))X_2(-\sigma)c, & 0 \leq \sigma \leq t, \\ c'X_2(t)(P(0)X_2(-\sigma) - Y_2(-\sigma))c, & t \leq \sigma \leq T. \end{cases}$$

It is the same as the following explicit formula in terms of the system parameters and the Riccati solution

$$(5.7) \quad k(t, \sigma) = \begin{cases} c' \begin{pmatrix} I & 0 \\ P(0) & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I & 0 \\ P(0) & 0 \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix} c, & 0 \leq \sigma \leq t, \\ c' \begin{pmatrix} I & 0 \\ P(0) & -I \end{pmatrix} e^{tH} \begin{pmatrix} 0 & 0 \\ P(0) & -I \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix} c, & t \leq \sigma \leq T. \end{cases}$$

Furthermore, we have, by a simple observation, that

$$\begin{aligned} \int_0^\sigma \Phi(t, \tau) \frac{1}{\lambda} bb' \Phi(\sigma, \tau)' d\tau &= \begin{pmatrix} I & 0 \\ P(0) & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I & 0 \\ P(0) & 0 \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix}, \text{ for } 0 \leq \sigma \leq t \\ \int_0^t \Phi(t, \tau) \frac{1}{\lambda} bb' \Phi(\sigma, \tau)' d\tau &= \begin{pmatrix} I & 0 \\ P(0) & -I \end{pmatrix} e^{tH} \begin{pmatrix} 0 & 0 \\ P(0) & -I \end{pmatrix} e^{-\sigma H} \begin{pmatrix} 0 \\ I \end{pmatrix}, \text{ for } t \leq \sigma \leq T \end{aligned}$$

and

$$\Phi(t, 0) = \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I \\ P(0) \end{pmatrix}.$$

Therefore, the transition matrix $\Phi(t, s)$ is

$$\Phi(t, s) = \begin{pmatrix} I & 0 \end{pmatrix} e^{tH} \begin{pmatrix} I \\ P(0) \end{pmatrix} \left(\begin{pmatrix} I & 0 \end{pmatrix} e^{sH} \begin{pmatrix} I \\ P(0) \end{pmatrix} \right)^{-1}.$$

We now see that in order to explicitly construct the optimal filter we need only find the initial data for the Riccati equation associated with the optimal control problem of (4.5), and the matrix exponential of the symplectic matrix H .

6. The cubic spline. In this section we consider the most important of the splines, the cubic spline, and construct the explicit linear filter. In numerical analysis the cubic spline is the spline that is most commonly used. Recall that it is piecewise cubic polynomial and is twice continuously differentiable everywhere. In light of the previous section we begin by deriving the associated Riccati equation.

Let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad c = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Let

$$J(u) = \int_0^T y(t)^2 + \lambda^{-1} u(t)^2 dt$$

where λ is a positive constant. We now solve the following problem:

$$\min_{u(t)} J(u)$$

subject to the constraints that

$$\dot{x} = Ax + bu, \quad y = cx.$$

The Hamiltonian matrix associated with this problem is

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\lambda} \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Our immediate task is to calculate e^{Ht} . To do so we need the explicit powers of H .

$$H^2 = \begin{pmatrix} 0 & 0 & 0 & -\frac{1}{\lambda} \\ 0 & 0 & \frac{1}{\lambda} & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad H^3 = \begin{pmatrix} 0 & 0 & \frac{1}{\lambda} & 0 \\ -\frac{1}{\lambda} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\lambda} \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$H^4 = \begin{pmatrix} -\frac{1}{\lambda} & 0 & 0 & 0 \\ 0 & -\frac{1}{\lambda} & 0 & 0 \\ 0 & 0 & -\frac{1}{\lambda} & 0 \\ 0 & 0 & 0 & -\frac{1}{\lambda} \end{pmatrix}.$$

From this calculation we see that the eigenvalues of H are the fourth roots of $-\frac{1}{\lambda}$,

$$\left(\frac{1}{\lambda}\right)^{\frac{1}{4}}(\pm 2^{\frac{-1}{2}} \pm i 2^{\frac{-1}{2}}).$$

We now give a different form of $\exp Ht$.

$$\begin{aligned} e^{Ht} &= \sum_{n=0}^{\infty} \frac{H^n t^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{H^{4n} t^{4n}}{(4n)!} + \sum_{n=0}^{\infty} \frac{H^{4n+1} t^{4n+1}}{(4n+1)!} + \sum_{n=0}^{\infty} \frac{H^{4n+2} t^{4n+2}}{(4n)!} + \sum_{n=0}^{\infty} \frac{H^{4n+3} t^{4n+3}}{(4n+3)!} \\ &= I \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n}}{(4n)!} + H \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n+1}}{(4n+1)!} \\ &\quad + H^2 \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n+2}}{(4n+2)!} + H^3 \sum_{n=0}^{\infty} \frac{(-\frac{1}{\lambda})^n t^{4n+3}}{(4n+3)!} \\ &= f_0(t)I + f_1(t)H + f_2(t)H^2 + f_3(t)H^3. \end{aligned}$$

We note the following

$$He^{Ht} = f_0 H + f_1 H^2 + f_2 H^3 - \frac{1}{\lambda} f_3 I = f_0' I + f_1' H + f_2' H^2 + f_3' H^3$$

and hence by linear independence of the powers of H we have the following differential relations.

$$f_0' = -\frac{1}{\lambda} f_3, \quad f_0^{(2)} = -\frac{1}{\lambda} f_2, \quad f_0^{(3)} = -\frac{1}{\lambda} f_1$$

so that it suffices to find a closed form for f_0 .

LEMMA 6.1. *The function $f_0(t)$ is given by $f_0(t) = \cosh\left(\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\right) \cos\left(\frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t\right)$.*

Proof. Recall that

$$\cosh t = \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!}$$

and

$$\cos t = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!}$$

and so

$$\cosh t + \cos t = 2 \sum_{n=0}^{\infty} \frac{t^{4n}}{(4n)!}.$$

Using the fact that $(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})^4 = -1$ we have that

$$\cosh(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})t + \cos(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})t = 2 \sum_{n=0}^{\infty} \frac{(-1)^n (\frac{1}{\lambda})^n t^{4n}}{(4n)!}.$$

Now replacing cosh and cos in the above expression by their exponential representations we have after a little manipulation

$$\cosh(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})t + \cos(\frac{1}{\lambda})^{1/4}(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})t = 2 \cosh \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t \cos \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t.$$

Thus we have proved the lemma.

Using the differential relations we have after some tedious calculations

$$\begin{aligned} f_0(t) &= \cosh \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t \cos \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t, \\ f_1(t) &= \frac{\lambda^{1/4}}{\sqrt{2}}(\cos \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t \sinh \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t + \sin \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t \cosh \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t), \\ f_2(t) &= \lambda^{1/2} \sin \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t \sinh \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t, \\ f_3(t) &= \frac{-\lambda^{3/4}}{\sqrt{2}}(\sinh \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t \cos \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t - \cosh \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t \sin \frac{(\frac{1}{\lambda})^{1/4}}{\sqrt{2}}t). \end{aligned}$$

Thus we have a closed form representation of $\exp Ht$.

$$e^{Ht} = \begin{pmatrix} f_0 & f_1 & \frac{1}{\lambda}f_3 & -\frac{1}{\lambda}f_2 \\ -\frac{1}{\lambda}f_3 & f_0 & \frac{1}{\lambda}f_2 & -\frac{1}{\lambda}f_1 \\ -f_1 & -f_2 & f_0 & \frac{1}{\lambda}f_3 \\ f_2 & f_3 & -f_1 & f_0 \end{pmatrix}$$

Now let

$$F(t) = \begin{pmatrix} -f_1 & -f_2 \\ f_2 & f_3 \end{pmatrix} \begin{pmatrix} f_0 & f_1 \\ -\frac{1}{\lambda}f_3 & f_0 \end{pmatrix}^{-1}.$$

To show that

$$F(t)$$

is a solution to the Riccati equation take the derivative of $F(t)$ using the fact

$$\frac{d}{dt} \exp Ht = H \exp Ht$$

and then note that $F(0) = 0$ The particular solution we want is then given by

$$P(t) = F(t - T).$$

We now calculate $F(t)$ explicitly. We make the convention that $S = \sinh(\cdot)$, $s = \sin(\cdot)$, $C = \cosh(\cdot)$, $c = \cos(\cdot)$ for typographical convenience.

First

$$F(t) = \frac{1}{f_0^2 + \frac{1}{\lambda}f_1f_3} \begin{pmatrix} -f_1f_0 - \frac{1}{\lambda}f_2f_3 & f_1^2 - f_2f_0 \\ f_2f_0 + \frac{1}{\lambda}f_3^2 & -f_1f_2 + f_3f_0 \end{pmatrix}$$

after some calculations we have

$$\begin{aligned} f_1^2 - f_2f_0 &= \frac{\lambda^{1/2}}{\sqrt{2}}(c^2S^2 + s^2C^2) \\ f_1^2 - f_2f_0 &= f_2f_0 + \frac{1}{\lambda}f_3^3 \\ -f_1f_0 - \frac{1}{\lambda}f_2f_3 &= \frac{\lambda^{1/4}}{\sqrt{2}}(-SC - sc) \\ f_3f_0 - f_1f_2 &= \frac{\lambda^{3/4}}{\sqrt{2}}(-SC + cs) \\ f_0^2 + \frac{1}{\lambda}f_1f_3 &= \frac{1}{2}(C^2 + c^2). \end{aligned}$$

We have now found an explicit form for $F(t)$

$$F(t) = \begin{pmatrix} -\sqrt{2}\lambda^{1/4}\frac{SC-cs}{C^2+c^2} & \lambda^{1/2}\frac{S^2+s^2}{C^2+c^2} \\ \lambda^{1/2}\frac{S^2+s^2}{C^2+c^2} & \sqrt{2}\lambda^{3/4}\frac{cs-SC}{C^2+c^2} \end{pmatrix}.$$

Thus the explicit solution of the Riccati equation is

$$P(t) = F(t - T).$$

It is interesting to note that this derivation is closely related to the representation theory of the cyclic group of order 8. It is possible to construct the Riccati equation for general polynomial splines of degree $2n - 1$ and the construction is closely related to the representation theory of the cyclic group of order $4n$. At this point we do not know how to exploit this fact.

We now construct the explicit filter for the cubic spline and exploit the linear quadratic optimization theory to obtain a simplified form of the operator. We first exploit the fact that $P(0)$ is approximated by the positive definite solution of the algebraic Riccati equation. The steady state Riccati solution for the cubic splines is

$$P = \begin{pmatrix} \sqrt{2}\lambda^{1/4} & \sqrt{\lambda} \\ \sqrt{\lambda} & \sqrt{2}\lambda^{3/4} \end{pmatrix}.$$

It is positive definite. The transition matrix is

$$\Phi(t, s) = e^{(t-s)(A - \frac{1}{\lambda}bb'P)}$$

and consequently the kernel is,

$$(6.1) \quad \hat{k}(t, \sigma) = \begin{cases} \frac{\exp(-\frac{\sigma+t}{\sqrt{2}\lambda^{1/4}})}{2\sqrt{2}\lambda^{1/4}} \left[\left(-2 + \exp(\frac{\sqrt{2}\sigma}{\lambda^{1/4}}) \right) \cos \frac{\sigma-t}{\sqrt{2}\lambda^{1/4}} + \cos \frac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \right. \\ \left. - \exp(\frac{\sqrt{2}\sigma}{\lambda^{1/4}}) \sin \frac{\sigma-t}{\sqrt{2}\lambda^{1/4}} - \sin \frac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \right], & 0 \leq \sigma \leq t, \\ \frac{\exp(-\frac{\sigma+t}{\sqrt{2}\lambda^{1/4}})}{2\sqrt{2}\lambda^{1/4}} \left[\left(-2 + \exp(\frac{\sqrt{2}t}{\lambda^{1/4}}) \right) \cos \frac{\sigma-t}{\sqrt{2}\lambda^{1/4}} + \cos \frac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \right. \\ \left. + \exp(\frac{\sqrt{2}t}{\lambda^{1/4}}) \sin \frac{\sigma-t}{\sqrt{2}\lambda^{1/4}} - \sin \frac{\sigma+t}{\sqrt{2}\lambda^{1/4}} \right], & t \leq \sigma \leq T. \end{cases}$$

It is known that the feedback matrix $A - \frac{1}{\lambda}bb'P$ is Hurwitz, i.e., the eigenvalues of the matrix lie in the left-half complex plane and they are equal to the left-half plane eigenvalues of the matrix H . The following graphs illustrate the differences between our approximations and the kernel defined by Silverman in [9]. Silverman’s approximation is given by the kernel

$$(6.2) \quad \kappa(t, \sigma) = \frac{1}{2} \exp\left\{-\frac{|t-\sigma|}{\sqrt{2}}\right\} \sin\left(\frac{|t-\sigma|}{\sqrt{2}} + \frac{\pi}{4}\right).$$

It is worthwhile to compare Silverman’s approximation to our approximation in equation (6.1). Note that two graphs overlap in the plotted range in Figures 1 and Figure 3, where the differences occur near the outer limits but even there they are small.

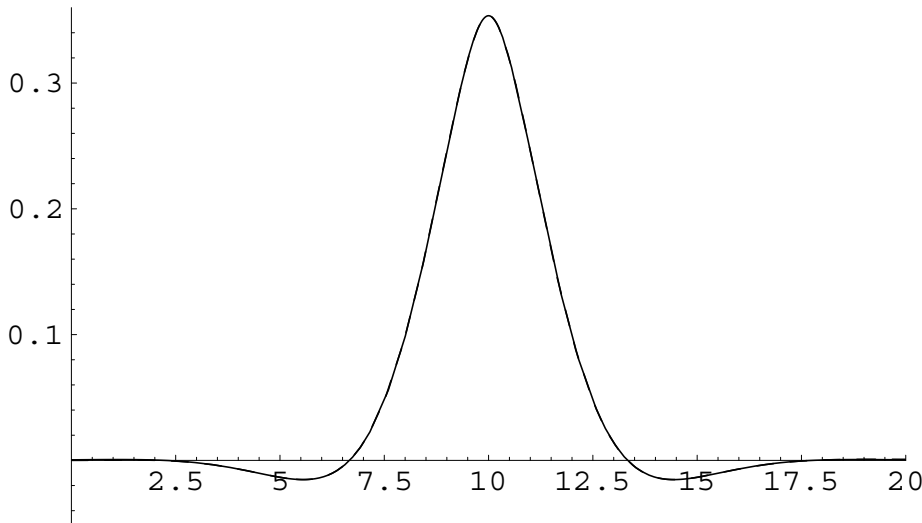


FIG. 1. The kernel $\kappa(t, \sigma)$ (solid line) and the kernel defined by $k(t, \sigma)$ dashed line

Our approximation is clearly better than Silverman’s. However the improvement is at the expense of complexity of the formula.

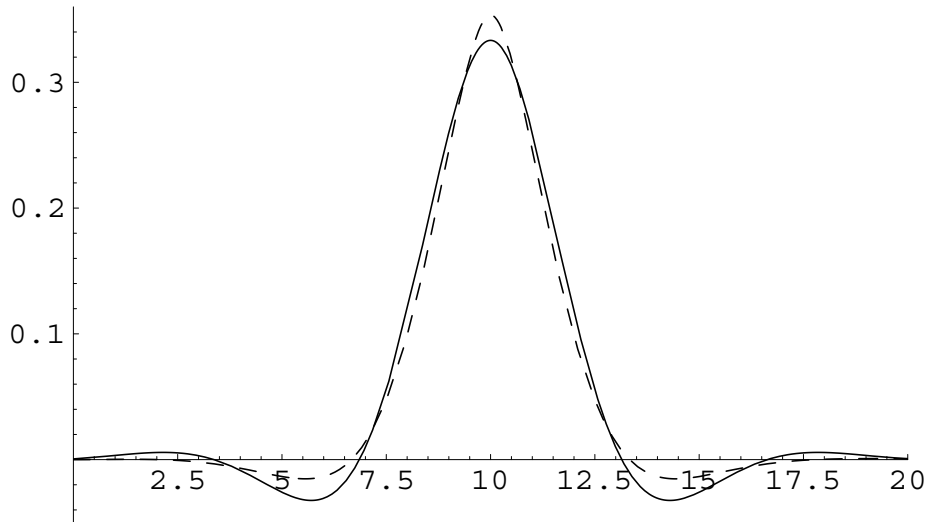


FIG. 2. The approximated kernel in [9] (solid line) and the kernel $k(t, \sigma)$ (dashed line)

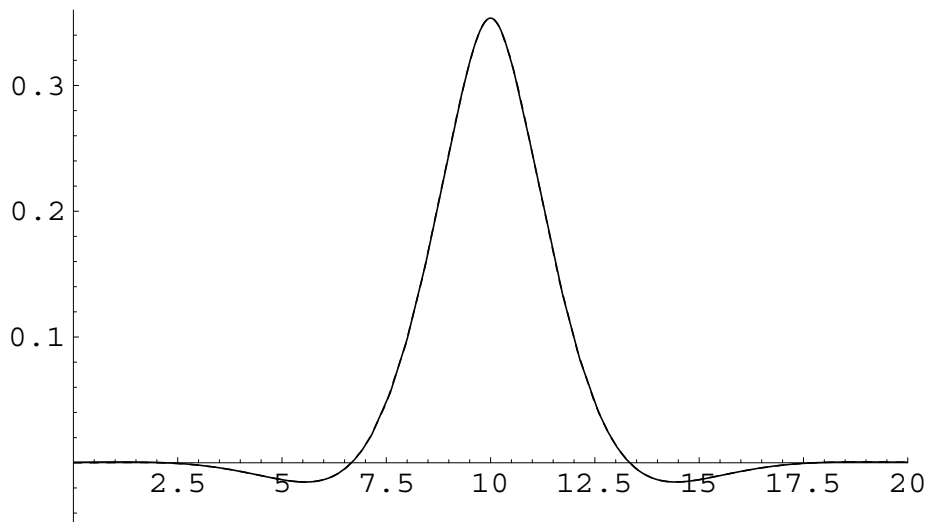


FIG. 3. The kernel $\kappa(t, \sigma)$ (solid line) and $\hat{\kappa}(t, \sigma)$ (dashed line)

Now we turn to a short discussion of the bandwidth of the kernel. The term bandwidth is frequently used in statistic literatures, but is seldom explicitly defined. The definition we use here is that the bandwidth is the interval where one obtains the most information. Concretely, we define the number β determined by the solution of

the following equation

$$(6.3) \quad \frac{\int_{\max(t-\beta,0)}^{\min(t+\beta,T)} k(t,\sigma)d\sigma}{\int_0^T k(t,\sigma)d\sigma} = 0.9.$$

Then 2β is the bandwidth of the kernel $k(t,\sigma)$.

This is a nonlinear equation which can be numerically solved by e.g., the Newton-Raphson method. We illustrate this by some examples. Since the approximated kernel does as well as the exact one. We compute the bandwidth for (6.1). After integration, we get the denominator, respectively the numerator, of the left hand side in (6.3) as follows

$$\begin{aligned} & \frac{1}{2}e^{-\frac{t+T}{\sqrt{2}\lambda^{1/4}}} \left(-2e^T \sqrt{2}\lambda^{1/4} \cos \frac{t}{\sqrt{2}\lambda^{1/4}} - \left(-1 + e^{\frac{T}{\sqrt{2}\lambda^{1/4}}} \right) \cos \frac{t-T}{\sqrt{2}\lambda^{1/4}} \right. \\ & \quad \left. + 2 \left(e^{\frac{t+T}{\sqrt{2}\lambda^{1/4}}} + \left(-e^{\frac{T}{\sqrt{2}\lambda^{1/4}}} + \cos \frac{T}{\sqrt{2}\lambda^{1/4}} \right) \sin \frac{t}{\sqrt{2}\lambda^{1/4}} \right) \right); \\ & \frac{1}{2}e^{-\frac{3t}{\sqrt{2}\lambda^{1/4}}} \left(2e^{\frac{3t}{\sqrt{2}\lambda^{1/4}}} - e^{\frac{-\beta+t}{\sqrt{2}\lambda^{1/4}}} \left(-1 + e^{\frac{\sqrt{2}\beta}{\lambda^{1/4}}} + 2e^{\frac{\sqrt{2}t}{\lambda^{1/4}}} \right) \cos \frac{\beta}{\sqrt{2}\lambda^{1/4}} \right. \\ & \quad \left. - \left(e^{\frac{-\beta+t}{\sqrt{2}\lambda^{1/4}}} + e^{\frac{\beta+t}{\sqrt{2}\lambda^{1/4}}} \right) \sin \frac{\beta}{\sqrt{2}\lambda^{1/4}} + e^{\frac{\beta+t}{\sqrt{2}\lambda^{1/4}}} \sin \frac{\beta-2t}{\sqrt{2}\lambda^{1/4}} + e^{\frac{-\beta+t}{\sqrt{2}\lambda^{1/4}}} \sin \frac{\beta+2t}{\sqrt{2}\lambda^{1/4}} \right). \end{aligned}$$

Solving the equation in some special cases, we obtain $\beta_{T,\lambda}$

$$\beta_{10,0.01} = 0.547334, \beta_{10,0.1} = 0.968005, \beta_{10,10} = 3.22425, \beta_{10,100} = 4.03161.$$

Remark: We have done an explicit construction for the cubic spline. The same construction can be done for arbitrary polynomial splines and for torsion splines. The limiting difficulty is the explicit calculation of the exponential of the Hamiltonian matrix. However for most data sets the case of a 4×4 Hamiltonian suffices. It is usually possible to explicitly construct the exponential in these cases.

7. Conclusion. In this paper we have shown that smoothing splines with discrete data lead to a problem with continuous data and we show that the solution to the continuous problem leads naturally to an integral filter. This filter is given as a function of the solution to a fixed optimal control problem. The kernel of the filter is obtained exactly. It is then shown that the kernel can be significantly simplified and that the kernel only depends on the matrix exponential of a well behaved Hamiltonian matrix and the initial value of a related Riccati equation. This allows another simplifying step in that we replace the $P(0)$ with the positive definite solution of the algebraic Riccati equation. We then apply this to an important special case—the cubic smoothing spline. We obtain the specific filter for this spline and compare our

results to Silverman's. Our filter is slightly better than his and the approximation we obtain by using the positive definite solution is very good indeed. We show that the bandwidth of the filter is related to the parameter λ and hence we conclude that smoothing splines can have a very narrow band width which is good for picking up local behavior of the data set and by choosing λ we can widen the band width and thus pick up long term behavior of the data. We also see that for narrow band width the smoothing spline while not casual only looks a few steps ahead. This coincides with the folklore of the statistical literature that a change of data at one point only has local effect. Perhaps one of the most important contributions of this paper is to show that by setting the smoothing splines in a control theoretic setting additional intuition about the behavior of splines is obtained.

It would be informative to construct smoothing splines and the corresponding filter in a nonlinear setting or on manifolds. There is a major obstacle in that it is important to have an explicit representation of the solution of the dynamics. This is only available in very special cases, none of which seem very important. On manifolds the major problem is a usable distance function. On a Riemannian manifold the metric could be used but significant complications arise. Restriction to a fixed chart can be used effectively but then the construction is really just local. However by properly choosing the weights on the data points good smoothing splines can be achieved. On simple manifolds geodesic distance is good but again significant technical complications occur.

Acknowledgements. The work of the first author is supported in part by the Swedish Research Council (VR), Magnusson's Foundation through Swedish Royal Academy of Sciences and the SEB Bergvall's Foundation. The work of the third author was supported in part by NSF Grants ECS 9720357 and ECS 9705312. The support of KTH in the spring of 2000 and 2003 is gratefully acknowledged. This work was supported in part by a fellowship from the Japanese Society for the Promotion of Science while the third author was in residence at Tokyo Denki University and RIKEN. This work was supported in part by a grant from the Texas Coordinating Board for Higher Education through the Advanced Research Projects Program. The third author wishes to thank the Mittag-Leffler Institute for support during the spring of 2003 and the Wenner-Gren Foundation for support during the summer of 2003, and the support of the Gustavsson Foundation.

REFERENCES

- [1] P. CROUCH AND J. JACKSON, *Dynamic interpolation and application to flight control*, Journal of Guidance, Control and Dynamics, 14(1991), pp. 814–822.
- [2] M. EGERSTEDT AND C. F. MARTIN, *Optimal trajectory planning and smoothing splines*, Automatica, 37(2001), pp. 1057–1064.

- [3] C. F. MARTIN, S. SUN, AND M. EGERSTEDT, *Optimal control, statistics and path planning*, Math. Comput. Modelling, 33:1-3(2001), pp. 237–253, Computation and Control, VI (Bozeman, MT, 1998).
- [4] S. SUN, M. EGERSTEDT, AND C. F. MARTIN, *Control theoretic smoothing splines*, IEEE Trans. Automat. Control, 45:12(2000), pp. 2271–2279.
- [5] Y. ZHOU, M. EGERSTEDT, AND C. MARTIN, *Optimal approximation of functions*, Com. Inf. Syst., 1(2001), pp. 101–112.
- [6] M. H. SCHULTZ AND R. S. VARGA, *l-splines*, Numer. Math., 10(1967), pp. 345–369.
- [7] F. SILVA-LEITE AND R. RODRIGUES, *l-splines - a manifestation of optimal control*, IMA Journal of Mathematical Control and Information, 19(2002), pp. 313–324.
- [8] G. WAHBA, *Spline models for observational data*, Number 59 in CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, PA, 1990.
- [9] B.W. SILVERMAN, *Spline smoothing: the equivalent variable kernel method*, Ann. Statist., 12(1984), pp. 898–916.
- [10] P. J. DAVIS AND P. RABINOWITZ, *Methods of numerical intergration*, Academic Press, 1984.
- [11] D. G. LUENBERGER, *Optimization by vector space methods*, John Wiley & Sons Inc., New York, 1969.
- [12] C. MARTIN AND M. EGERSTEDT, *Statistical estimates for generalized splines*, Control, Optimization and the Calculus of Variations, 9(2003), pp. 553–562.