

**A REGULARIZED SOLUTION TO THE  
BIRKHOFF INTERPOLATION PROBLEM**

C. F. MARTIN and Y. ZHOU

REPORT No. 10, 2002/2003, spring

ISSN 1103-467X

ISRN IML-R- -10-02/03- -SE+spring



**INSTITUT MITTAG-LEFFLER**  
THE ROYAL SWEDISH ACADEMY OF SCIENCES

# A REGULARIZED SOLUTION TO THE BIRKHOFF INTERPOLATION PROBLEM

Y. ZHOU AND C.F. MARTIN

ABSTRACT. In this paper a relationship the Birkhoff Interpolation Problem and the problem of control theoretic splines is established. It is shown that the Birkhoff interpolation problem has a solution if and only if a certain cost function remains bounded under perturbation of a suitably chooser variable.

## 1. INTRODUCTION

In 1906 G. Birkhoff published a paper, [2], that introduced a problem that has remained unsolved for almost one hundred years. The problem that was later restated by Polya, [11], as the following. Given a differential equation

$$u^{(n)}(t) = 0$$

what combination of initial and terminal values suffice to construct a unique solution. This generalizes the concepts of interpolation that were then current. Lagrange interpolation (more properly Warring interpolation, [15, 6]), and the most general Hermite interpolation are special cases if only two time points are considered. See [3] for the most readable account of several different interpolation schemes. The most notable thing about the problem posed by Birkhoff is that there is not always a solution. For example, if  $n = 2$  and we are given  $u'(0) = 1$  and  $u'(1) = 2$  then there is no solution. The most general version of the problem, as poised by Birkhoff, is the following: Let  $0 = t_0 < t_1 < \dots < t_N = 1$ . What combination of  $n$  functional values and higher derivatives at the points  $t_i$  suffices to give a unique solution to the problem. It is this problem that has been studied in the literature for almost one hundred years. Primary references are [12, 8].

An other problem that has been studied and will be shown to be related to Birkhoff interpolation is the study of optimal digital to analog conversion using linear system theory. Two versions of this have been studied, interpolation and smoothing, but in this paper we restrict attention to interpolation. The simplest problem is the following. Given a set of data points

$$D = \{(t_i, \alpha_i) : i = 1, \dots, N\},$$

---

*Key words and phrases.* Birkhoff Interpolation, optimal control systems, control theoretic splines .

The work of the first author is supported in part by the Swedish Research Council (VR), the Magnossun's Foundation through Swedish Royal Academy of Sciences and the SEB Bergvall's Foundation. The work of the second author is supported in part by NSF Grants ECS 9720357 and ECS 9705312. The support of KTH in the spring of 2003 is gratefully acknowledged. The second 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 a linear control system

$$\dot{x} = Ax + bu, \quad y = cx, \quad x(0) = x_0$$

find a control  $u$  that drives the output  $y(t)$  through the data points at the required times and that uses minimal control energy. To state the problem precisely

$$\min_u \int_0^T u^2(t) dt$$

subject to the constraints

$$y(t_i) = \alpha_i.$$

This problem can be generalized to include Birkhoff type constraints. See, for example, the papers [16, 14] for additional references and results.

In Section 2 a brief recap of linear control theory is presented. The material that is presented is what is needed to read this paper. In Section 3 the theory of digital to analog (D-A) conversion using optimal linear control is presented. This material can be found in a series of papers by Martin and his coauthors but sufficient references and background occur in [14]. These functional approximations need have continuous derivatives of all possible orders. The spline is different than that used by Melkman, [10]. In Section 4 we begin the formulation of the main problem of this paper. We recast the D-A conversion in terms of polynomials and consider several special cases. This will demonstrate the method and will provide at least one new example. In Section 5 we obtain the main theorem of the paper; an equivalence between the Birkhoff interpolation problem and a sequence of problems from linear optimal control.

## 2. LINEAR CONTROL THEORY

In this section a few facts about linear controlled systems are established. We consider the following system. Let

$$(1) \quad \dot{x} = Ax + bu, \quad x(0) = x_0$$

$$(2) \quad y = cx$$

where  $A$  is  $n \times n$ ,  $b$  is  $n \times 1$  and  $c$  is  $1 \times n$ . The system is said to be controllable if and only if for each initial point  $x_0$  and each terminal point  $x_T$  there exists a control  $u(t)$  such that

$$x_T = x(T) = e^{AT} x_0 + \int_0^T e^{A(T-s)} bu(s) ds.$$

The following theorem gives two equivalent criteria for controllability.

**Theorem 2.1.** *The linear system given by equation 1 is controllable if and only if either of the following two conditions apply*

$$(3) \quad \text{rank}[b, Ab, A^2b, \dots, A^{n-1}b] = n,$$

$$(4) \quad \text{rank} \int_0^T e^{-As} b^T b e^{-As} ds = n.$$

The proof of the theorem is elementary and can be found in any book on control theory.

Another key concept is observability. The system 1 and 2 with  $u = 0$  is observable if and only if knowing  $y(t)$  on any open interval allows the reconstruction of  $x(t)$  on the interval  $[0, T]$ . The following theorem is classical.

**Theorem 2.2.** *The system 1 and 2 with  $u = 0$  is observable if and only if either of the following conditions hold*

$$(5) \quad \text{rank} \begin{bmatrix} c \\ cA \\ \vdots \\ cA^{n-1} \end{bmatrix} = n,$$

$$(6) \quad \text{the system } \dot{x} = A^T x + c^T u \text{ is controllable.}$$

In this paper we will be concerned with optimal control of an elementary sort. The classical linear quadratic control problem is to minimize a cost function of the form

$$(7) \quad J(u) = \int_0^T x^T(t) c^T c x(t) + u^2(t) dt.$$

The following theorem is classical.

**Theorem 2.3.** *There exists a unique control  $u^*(t)$  such that*

$$J(u^*) \leq J(u),$$

*with equality holding only if  $u = u^*$  a.e., if and only if the system 1 is controllable and the system 1 and 2 with  $u = 0$  is observable. Furthermore the optimal control is given by*

$$u(t) = -b^T P(t)x(t)$$

*where  $P(t)$  is the solution to the Riccati equation*

$$\dot{P} = -A^T P - PA + P b b^T P - c^T c, \quad P(T) = 0.$$

### 3. CONTROL THEORETIC SPLINES AND INTERPOLATION

In this section we consider the basic optimization routines that results in control theoretic splines and we consider basic interpolation results using linear control systems. For the material on splines the history and basic results are found in [16, 14].

Again let

$$\begin{aligned} \dot{x} &= Ax + bu, & x(0) &= x_0 \\ y &= cx \end{aligned}$$

where  $A$  is  $n \times n$ ,  $b$  is  $n \times 1$  and  $c$  is  $1 \times n$ . Since our interest is in smooth approximation and not in control theory we make the following assumptions on the

form of the system matrices.

$$(8) \quad A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ & & & & \ddots & \\ 0 & \cdots & & & 0 & 1 \\ \gamma_1 & \cdots & & & \cdots & \gamma_n \end{pmatrix},$$

$$(9) \quad b = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

$$(10) \quad c = (1, 0, \dots, 0).$$

Note that this implies

$$(11) \quad cb = cAb = \dots = cA^{n-2}b = 0.$$

Also assume that there is given a set of data points

$$(12) \quad D = \{(t_i, \alpha_i) : 0 < t_1 < \dots < t_N \leq T\}.$$

We further assume that the control

$$u \in L^2[0, T].$$

We also assume that  $x_0 \in R^n$  and that  $R^n$  is equipped with the norm  $\|x\|^2 = \lambda x^T x$  with  $\lambda > 0$ . We define a set of functions that are basic to the problem. Let

$$(13) \quad \ell_i(s) = \begin{cases} ce^{A(t_i-s)}b & t_i - s > 0 \\ 0 & \text{otherwise} \end{cases}$$

It is important to note that by equation 11  $\ell_i(t)$  is  $n - 2$  times continuously differentiable. Let

$$\mathcal{H} = R^n \times L_2[0, T]$$

be the Hilbert space with norm

$$\begin{aligned} \|(x, u)\|_{\mathcal{H}}^2 &= \|x\|^2 + \|u\|^2 \\ &= \lambda x^T x + \int_0^T u^2(t) dt. \end{aligned}$$

We now prove the following theorem.

**Theorem 3.1.** *Let*

$$J(x_0, u) = \|x_0\|^2 + \|u\|^2$$

and

$$y(t) = ce^{At}x_0 + \int_0^t ce^{A(t-s)}bu(s)ds.$$

*There exists a unique pair  $(x_0^*, u^*)$  that has the following two properties.*

- (1) *for all  $(x, u) \in \mathcal{H}$ ,  $J(x_0^*, u^*) \leq J(x, u)$*
- (2) *for  $i = 1, \dots, N$ ,  $\alpha_i = ce^{At_i}x_0^* + \int_0^{t_i} \ell_i(s)u^*(s)ds$ .*

Furthermore the optimal solution is

$$(x_0^*, u^*) = \sum_{i=1}^N \tau_i (\lambda^{-1} (ce^{At_i})^T, l_i(t)),$$

where

$$\tau_i = e_i^T (\lambda^{-1} G + H) \alpha,$$

and  $G$  and  $H$  are grammians associated with the constraints.

**Proof:**  $J(x_0, u) = \|(x_0, u)\|_{\mathcal{H}}^2$ . Now let

$$W = \{(x_0, u) : \text{For every } i, \alpha_i = y(t_i) = ce^{At_i} x_0 + \int_0^T \ell_i(s) u(s) ds\}$$

Thus the optimization problem is just the problem of finding a point of minimum norm on an affine subspace. We now prove an important lemma

**Lemma 3.2.** *The set  $S = \{(ce^{At_i}, \ell_i) : i = 1, \dots, N\}$  is linearly independent in  $\mathcal{H}$ .*

**Proof of lemma:** Suppose that there exists a set of nonzero numbers  $\rho_i$  such that for all  $t \in R$

$$\sum_{i=1}^N \rho_i (ce^{At_i}, \ell_i(t)) = 0.$$

Now by definition we have that for  $t_{N-1} < t < t_N$ ,  $\ell_j(t) = 0$  for all  $j \leq N-1$ . Thus we have that  $\rho_N (ce^{At_N}, \ell_N(t)) = 0$  and on the interval  $(t_{N-1}, t_N)$   $ce^{At} b = 0$  and so we must have that  $\rho_N = 0$ , a contradiction. Continuing in this manner we conclude that each  $\rho_i$  is 0.

Let  $G$  be the matrix with  $g_{ij} = ce^{At_i} e^{A^T t_j} c^T$  and let  $H$  be the matrix with

$$h_{ij} = \int_0^T \ell_i(s) \ell_j(s) ds.$$

Thus  $\lambda G + H$  is the grammian for the set  $S$  and since the elements of  $S$  are linearly independent we have

$$(\lambda G + H)^{-1}$$

exists for all  $\lambda \in [0, \infty)$ . We state, for future reference, this as a corollary to the lemma.

**Corollary 3.3.**  $(\lambda^{-1} G + H)^{-1}$  exists for all  $\lambda \in (0, \infty]$ .

Now by the projection theorem in Hilbert Space [9] the point of minimum norm is the unique point in

$$\text{span } S \cap W.$$

Thus we see that the optimal point  $(x_0^*, u^*(t))$  has the representation

$$(14) \quad (x_0^*, u^*(t)) = \sum_{i=1}^N \tau_i (\lambda^{-1} (ce^{At_i})^T, l_i(t)).$$

It remains to calculate the  $\tau_i$ . To do so we substitute this representation into equation 2 of the theorem and after a bit of calculation we see that we have

$$e_i (\lambda^{-1} G + H) \tau = \alpha_i$$

and then note that this is just the  $i$ th row of the equation

$$(\lambda^{-1} G + H) \tau = \alpha$$

and hence

$$\tau_i = e_i(\lambda G + H)^{-1}\alpha.$$

#### 4. THE POLYNOMIAL CASE

In this section we restrict attention to the case when  $A$  is nilpotent. Then  $e^{At}$  is a matrix of polynomials and hence the spline and interpolation problem both reduce to well understood cases of polynomial interpolation and polynomial splines. The assumptions that were made on the form of system matrices assures that these are classical forms.

The problem of non-Lagrange (or more precisely non-Waring) interpolation has not been considered from the viewpoint of systems theory. In this section we consider the relatively simple problem of osculatory splines and interpolation.

Let our data set be given as

$$D = \{(t_i, \alpha_i, \beta_i) : i = 1, \dots, N\}$$

We ask again to minimize the cost functional

$$J(u, x_0) = \lambda x^T x + \int_0^T u^2(t) dt$$

subject to the constraints

$$(15) \quad \alpha_i = ce^{At_i}x_0 + \int_0^{t_i} ce^{A(t_i-s)}bu(s)ds$$

and

$$(16) \quad \beta_i = cAe^{At_i}x_0 + \int_0^{t_i} cAe^{A(t_i-s)}bu(s)ds.$$

We begin by defining two sets of functions: the first is just new nomenclature for the  $\ell_i$ 's defined in the previous section and the second is the derivatives of those functions.

$$(17) \quad \ell_{(i,0)}(s) = \ell_i(s)$$

and

$$(18) \quad \ell_{(i,1)}(s) = \begin{cases} cAe^{At_i}b & t_i - s > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Note that for  $\ell_{(1,1)}(s)$  to be continuous that  $n > 2$ . To use the projection theorem we construct the affine subspace

$$(19) \quad W = \{(x_0, u(t)) : \text{for every } 0 < i \leq N, \text{ equations 15 and 16 hold}\}.$$

Controllability guarantees that the set is nonempty and hence there exist a linear subspace  $W_0$  parallel to  $W$ . We construct the orthogonal complement of  $W_0$ . A simple calculation shows that

$$(20) \quad W_0^\perp = \text{span}(\{(\lambda^{-1}(ce^{At_i})^T, \ell_{(i,0)}(s))\} \cup \{(\lambda^{-1}(cAe^{At_i})^T, \ell_{(i,1)}(s))\}).$$

Thus if the elements of the two spaces are linearly independent we are guaranteed a unique solution.

**Proposition 4.1.** *The set*

$S = \{(\lambda^{-1}(ce^{Ai})^T, \ell_{(i,0)}(s)) : i = 1, \dots, N\} \cup \{(\lambda^{-1}(cAe^{Ai})^T, \ell_{(i,1)}(s)) : i = 1, \dots, N\}$   
*is linearly independent provided that  $n > 2$ .*

**Proof:** Suppose that for all  $s \in [0, T]$

$$\sum_{i=1}^N \tau_{(i,0)}((ce^{Ai})^T, \ell_{(i,0)}(s)) + \tau_{(i,1)}((cAe^{Ai})^T, \ell_{(i,1)}(s)) = 0.$$

Then for  $t \in [t_{N-1}, t_N]$  we have that

$$\tau_{(N,0)}((ce^{At_N})^T, \ell_{(i,0)}(s)) + \tau_{(N,1)}((cAe^{At_N})^T, \ell_{(i,1)}(s)) = 0$$

and hence we have

$$\tau_{(N,0)}(ce^{At_N})^T + \tau_{(N,1)}(cAe^{At_N})^T = 0$$

Multiplying on the left by  $(e^{-At_N})^T$  we have that

$$\tau_{(N,0)}c + \tau_{(N,1)}cA = 0$$

but this implies that  $\tau_{(N,0)} = 0$  and then it follows that  $\tau_{(N,1)} = 0$  and the proposition follows.

By the projection theorem there exists a unique point in

$$W_0^\perp \cap W$$

and hence the point must be of the form

$$(21) \quad (x_0^*, u^*(t)) = \tau_{(i,0)}(\lambda^{-1}ce^{At_i}, \ell_{(i,0)}(t)) + \tau_{(i,1)}(\lambda^{-1}cAe^{At_i}, \ell_{(i,1)}(s))$$

where the  $\tau$  are chosen to satisfy the constraints.

We order the elements of  $S$  with the lexicographic order and construct the grammian. We immediately see that the grammian of  $S$  is of the form

$$\lambda^{-1}G + H$$

where  $G$  is the grammian of the ordered set

$$S_1 = \{\lambda^{-1}ce^{At_i}, \lambda^{-1}cAe^{At_i}, \lambda^{-1}ce^{At_2}, \dots, \lambda^{-1}cAe^{At_N}\}$$

and  $H$  is the grammian of the ordered set

$$S_2 = \{\ell_{(1,0)}, \ell_{(1,1)}, \ell_{(2,0)}, \dots, \ell_{(N,1)}\}$$

We have proved that  $\lambda^{-1}G + H$  is invertible but note that  $G$  is not necessarily invertible. In fact  $G$  is clearly not invertible unless  $n \leq 2N$ . As before it is seen after a bit of manipulation that the optimal

$$\tau = (\tau_{(1,0)}, \tau_{(1,1)}, \tau_{(2,0)}, \dots, \tau_{(n,1)})^T$$

is the solution of

$$(\lambda^{-1}G + H)\tau = \alpha$$

where  $\alpha$  is the vector of data given by

$$(\alpha_1, \beta_1, \alpha_2, \dots, \beta_N)^T.$$

Note that this statement is true for every  $\lambda \in (0, \infty)$ .

If we go back to the proof of the proposition we see that the same method of proof can be used to prove that  $H$  is invertible. We now make the assumption that

$n = 2N$  and note the following theorem. Also note that if  $\lambda = \infty$  then this just the the osculatory interpolation problem that is known to be solvable, [3].

**Theorem 4.2.** *If  $n = 2N$  and  $N > 1$  then  $\lambda G + H$  is invertible for all  $\lambda \in [0, \infty]$ .*

It is a fairly routine extension of this section to construct the general Hermite interpolation and spline problem in the polynomial case. It is important to examine the differentiability of the control and the resulting spline in this case. The basis element  $\ell_{i,1}$  is the culprit. Recall that  $l_{i1}(t) = cAe^{A(t_i-t)}b$  if  $t_i > t$  and 0 otherwise. Thus we can continuously differentiate  $l_{i1}$  only  $n - 3$  times. Calculating  $y(t)$  we see that  $y$  is only  $2n - 2$  times continuously differentiable.

The matrix  $H$  is invertible and basically the same proof holds as for  $G + \lambda H$ .

## 5. BIRKHOFF-HERMITE INTERPOLATION AND SPLINES

In this section we construct controls and initial values that solve a mixed version of the classical Birkhoff interpolation and Birkhoff spline problems.

**5.1. Statement of Problem.** We first assume that the system given by equations 1 and 2 is of dimension  $n$ , that the matrix  $A$  is nilpotent, and the system matrices are in the form given above. This insures that the optimal controls are polynomial.

We assume as given an incidence matrix  $J$  that is  $n \times n$  and whose entries consist of 0s and 1s. The number of 1s is assumed to be exactly  $n$ . Let

$$E = \{(i, j) : e_i^T J e_j = 1\}$$

and further assume that  $E$  is equipped with the lexicographic order.

We define the data set  $D^*$ .

$$(22) \quad D = \{(t_i, \alpha_{(i,j-1)}) : \text{for every } (i, j) \in E \ 0 \leq t_1 < t_2 < \dots < t_n \leq T\}$$

We now extend the definition of the basis functions from the previous section to the following: for every  $(i, j) \in E$

$$(23) \quad \ell_{(i,j)}(s) = \begin{cases} cA^{j-1}e^{A(t_i-s)}b & t_i - s > 0 \\ 0 & \text{otherwise} \end{cases}$$

Let the cost function be given as before as

$$J(x_0, u) = \lambda x_0^T x_0 + \int_0^t u^2(t) dt.$$

As before we assume that

$$(x_0, u) \in \mathcal{H}.$$

We define  $n$  constraints: for  $(i, j) \in E$

$$(24) \quad \alpha_{(i,j)} = cA^{j-1}e^{At_i}x_0 + \int_0^T l_{(i,j)}(s)u(s)ds.$$

We can state the problem as

**Problem 1.**

$$\min_{(x_0, u) \in \mathcal{H}} J(x_0, u)$$

subject to the  $n$  constraints of equation 24.

**5.2. The optimal solution.** We again will appeal to the Hilbert Projection Theorem. We will carefully define the affine subspace, the orthogonal complement of the parallel linear space and their intersection, thus finding the unique point that minimizes the cost.

First recall that

$$\mathcal{H} = R^n \times L_2[0, T]$$

with inner product

$$\langle (x, u), (z, v) \rangle = \lambda x^T z + \int_0^T u(t)v(t)dt.$$

Let

$$(25) \quad W = \{(x_0, u) \in R^n \times \mathcal{H} : \text{for all } (i, j) \in E \text{ equation 24 holds}\}$$

Note that  $W$  is not empty by controllability.

Let  $W_0$  denote the linear subspace parallel to  $W$ . Now an easy calculation shows that

$$(26) \quad W_0^\perp = \text{span} \{(\lambda^{-1}cA^{j-1}e^{At_i}, \ell_{(i,j)}) : (i, j) \in E\}$$

The key proposition is the following.

**Proposition 5.1.** *The set*

$$S_1 = \{(\lambda^{-1}cA^{j-1}e^{At_i}, \ell_{(i,j)}) : (i, j) \in E\}$$

*is linearly independent.*

**Proof:** Suppose on the contrary that there exist constants  $\tau_{(i,j)}$  such that

$$0 = \sum_{(i,j) \in E} \tau_{(i,j)} (\lambda^{-1}cA^{j-1}e^{At_i}, \ell_{(i,j)}(t)).$$

Let  $(k_1, j_{k_1})$  be the maximal element in  $E$  and let  $(k_2, j_{k_2})$  be the largest element less than all elements of the form  $(k_1, j)$  for  $j$ . Then for every  $t \in (t_{k_2}, t_{k_1})$  we have that  $\ell_{(k,j)}(t) = 0$  and thus that

$$0 = \sum_{(k_1, j) \in E} \tau_{(k_1, j)} (\lambda^{-1}cA^{j-1}e^{At_{k_1}}, \ell_{(k_1, j)}(t))$$

We now consider the sum

$$(27) \quad 0 = \sum_{(k_1, j) \in E} \tau_{(k_1, j)} \lambda^{-1}cA^{j-1}e^{At_{k_1}}$$

$$(28) \quad = \sum_{(k_1, j) \in E} \tau_{(k_1, j)} cA^{j-1}$$

However this last sum is just a linear combination of rows of the identity matrix and hence we have that  $\tau_{(k_1, j)} = 0$  whenever  $(k_1, j) \in E$ . It then follows that the set  $S_1$  is linearly independent. Thus the proof depends on the independence of  $\ell_{(k_i, j)}$  for initial  $i$ s and the independence of the  $cA^{j-1}e^{At_{k_i}}$  for fixed  $k_i$ .

Define the set  $S_2$

$$S_2 = \{\ell_{(i,j)} : (i, j) \in E\}$$

We prove the following proposition.

**Proposition 5.2.** *The set  $S_2$  is linearly independent.*

**Proof:** The idea of the proof is similar to the above. We first isolate the last block of functions and then prove that the functions in that block are linearly independent. Suppose that there exist numbers  $\tau_{ij}$  such that

$$\sum_{(i,j) \in E} \tau_{ij} \ell_{(i,j)}(t) = 0$$

for all  $t$ . Let  $(k_1, j_{k_1})$  be the maximal element in  $E$  and let  $(k_2, j_{k_2})$  be the largest element less than all elements of the form  $(k_1, j)$  for  $j$ . Then for every  $t \in (t_{k_2}, t_{k_1})$  this sum reduces to

$$\sum_{(k_2, j) \in E} \tau_{k_2 j} \ell_{(k_2, j)}(t) = 0$$

for all  $t \in (t_{k_2}, t_{k_1})$  or rewriting the sum we have

$$\sum_{(k_1, j) \in E} \tau_{k_1 j} c A^j e^{A(t_{k_1} - s)b} = 0.$$

Thus we have by controllability

$$\sum_{(k_1, j) \in E} \tau_{k_1 j} c A^j = 0.$$

this is just a set of distinct rows of the identity matrix and hence we have that this set is linearly independent. Thus we have that  $\tau_{k_1 j} = 0$  for all  $(k_1, j) \in E$ . The proof is finished by backwards induction.

Let

$$(29) \quad S_3 = \{\lambda^{-1} c A^{j-1} e^{At_i} : (i, j) \in E\}$$

and note that the following:  $S_3$  is linearly independent if and only if the corresponding Birkhoff interpolation problem is solvable.

Thus there exists a unique control of the form

$$(30) \quad u^*(t) = \sum_{(i,j) \in E} \tau_{(i,j)} \ell_{(i,j)}(t)$$

and an initial value

$$(31) \quad x_0^* = \sum_{(i,j) \in E} \tau_{(i,j)} e^{A^T t_i} (A^{j-1})^T c^T$$

that minimizes the cost function and satisfies the constraints.

To determine the the values of the  $\tau$ s we substitute the optimal control and initial value into the constraints. Let the sets  $S_i$  be ordered by the ordering induced from  $E$ . If we let  $G$  be the Grammian of the set  $S_3$ ,  $H$  the grammian of the set  $S_2$  then after some manipulation we have the following important fact. The grammian of the set  $S_1$  is

$$(32) \quad \lambda^{-1} G + H$$

and the optimal  $\tau$ s are the solution of

$$(33) \quad (\lambda^{-1} G + H) \tau = \alpha.$$

Thus we have that the interpolation problem is regularized by the optimal spline problem and we have a "new" necessary and sufficient condition for the existence

of a polynomial that solves the Birkhoff interpolation problem. Recall the cost function

$$J(x_0, u) = \lambda x_0^T x_0 + \int_0^T u^2(t) dt.$$

When  $\lambda = 0$  we have that all of the cost is on the control. Thus if there is a solution  $x_0$  to the Birkhoff interpolation problem the cost function is minimized by setting  $u = 0$ . Letting  $\lambda = \infty$  is equivalent to using the cost function

$$J(x_0, u) = x_0^T x_0.$$

There is no penalty for using large control and thus the if the spline interpolation problem is solvable for arbitrary initial data the the optimal cost is obtained by setting  $x_0 = 0$ .

Now assume that the Birkhoff problem is solvable with initial data  $x_0^*$  and consider the pair  $(x_0^*, 0)$  and cost function

$$J_\lambda(x_0, u) = \lambda x_0^T x_0 + \int_0^T u^2(t) dt$$

and let  $(x_\lambda^\#, u_\lambda^\#)$  be the optimal point. Then we have that for all  $\lambda \in (0, \infty)$  that

$$(34) \quad J_\lambda(x_\lambda^\#, u_\lambda^\#) \leq J(x_0^*, 0)$$

On the other hand suppose that there is no solution to the Birkhoff interpolation problem and consider the equivalent cost function

$$J_\beta(x_0, u) = x_0^T x_0 + \beta \int_0^T u^2(t) dt.$$

Let  $(x_\beta, u_\beta)$  be the optimal solution for  $\beta \in (0, \infty)$ . As  $\beta$  becomes large then either  $\int_0^T u^2(t) dt$  becomes unbounded or it is bounded and bounded away from 0 since there is no solution to the Birkhoff interpolation problem. Thus we have that

$$\lim_{\beta \rightarrow \infty} J_\beta(x_\beta, u_\beta) = \infty.$$

We have proven the following theorem.

**Theorem 5.3.** *Birkhoff interpolation problem with incidence matrix  $J$  has a solution if and only if there exists a positive constant  $M$  so that for ever  $\lambda \in (0, T)$  with optimal solution  $(x_\lambda^\#, u_\lambda^\#)$*

$$J_\lambda(x_\lambda^\#, u_\lambda^\#) \leq M.$$

#### REFERENCES

- [1] G. Ammar, W. Dayawansa and C. Martin, Exponential interpolation and numerical algorithms, Appl. Math. Comp. 41, 1991, 189-232
- [2] G. Birkhoff, General mean value theorems with applications to mechanical differentiation and quadrature. Tran.Amer.Math.Soc. vol. 7, 1906, 107-136
- [3] P. Davis, Interpolation and Approximation, Dover Pub. Co. 1975
- [4] R. de Prony, Essai expérimental et analytique, J. École Polytech. (Paris)1:24–76, 1795.
- [5] F.B. Hildebrand, Introduction to Numerical Analysis, Dover Publishig Co. 1983
- [6] J-L. Lagrange, Memoire sur la Méthod d'Interpolation, Nouveaux Mémoires de l'Academie Royale des Sciences ar Belle-Lettres de Berlin années 1792-1793 in Serret, J-A (ed) Oeuvres de LaGrange vol 5. Paris: Gauthiers-Villers, 1870
- [7] S. Karlin, J.M. Karon, On Hermite-Birkhoff interpolation, J. of Approx. Theory, 6, (1972), 90-115
- [8] G.G. Lorentz, K. Jetter, S.D. Riemenschneider, Birkhoff Interpolation, Encyclopedia of Mathematics and its Applications, 19, Addison Wesley Publishing Company, Reading, MA, 1983

- [9] D.Luenberger, Optimization by Vector Space Methods, John Wiley and Sons, Inc., 1968
- [10] A.A. Melkman, Hermite-Birkhoff interpolation by splines, J. Approximation Theory, 19,1977, 259-279
- [11] G. Polya, Bemerkungen zur Interpolation und der Naherungstheorie der Balkenbiegung, Z. Angew.Math.Mech. 11,1931,445-449.
- [12] A. Sharma, Some poised and nonpoised problems of interpolation, SIAM Review, 14, 1972,129-151.
- [13] J. Smith, C. Martin, Approximation, interpolation and sampling, in *Differential geometry: the interface between pure and applied mathematics*, 227-252,Cont. Math. Amer.Math. Soc. 1997.
- [14] S.Sun, M.Egerstedt, C.Martin Control theoretic smoothing splines, IEEE Tran. Auto. Cont.,45, 2000,2271-2279.
- [15] E. Waring, Problems concerning interpolation, Philosophical Transactions of the Royal Society of London, Vol 69, 1779, 59-67
- [16] Z. Zhang, J.Tomlinson, C. Martin, Splines and linear control theory, Acta. Appl.Math., 49, 1997,1-34

STOCKHOLM UNIVERSITET, STOCKHOLM SWEDEN, TEXAS TECH UNIVERSITY, LUBBOCK, TEXAS  
E-mail address: zhou@math.su.se, martin@math.ttu.edu