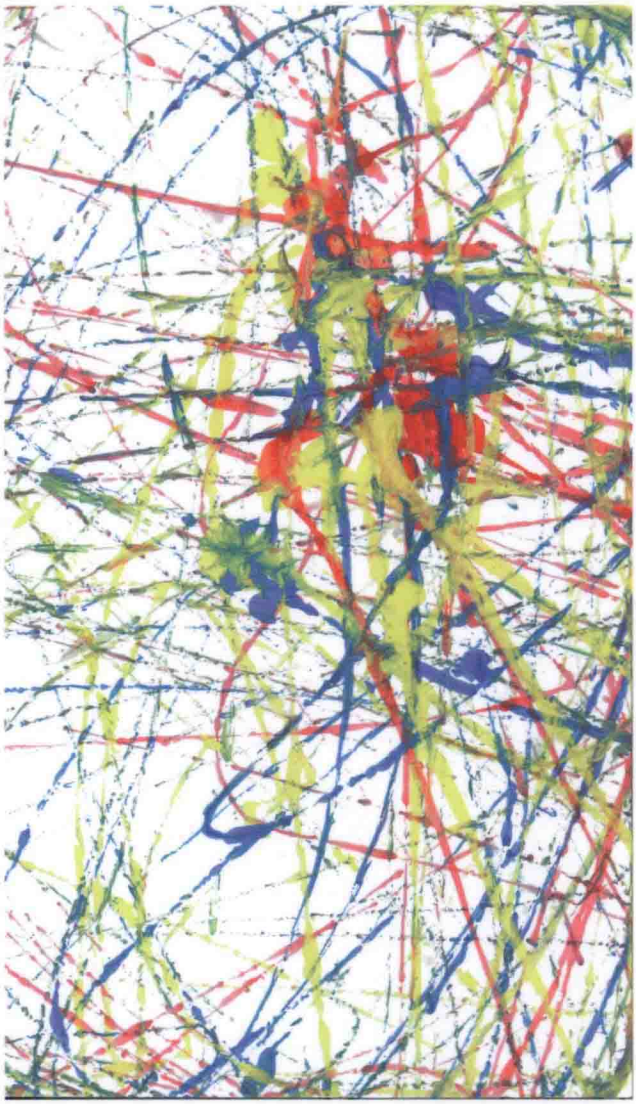
Princeton Series in APPLIED MATHEMATICS

Control Theoretic Splines

Optimal Control, Statistics, and Path Planning



Magnus Egerstedt
Clyde Martin

Control Theoretic Splines

Optimal Control, Statistics, and Path Planning

Magnus Egerstedt and Clyde Martin

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Published by Princeton University Press,

41 William Street, Princeton, New Jersey 08540

In the United Kingdom: Princeton University Press,

6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

Library of Congress Cataloging-in-Publication Data

Egerstedt, Magnus.
Control theoretic splines : optimal control, statistics, and path planning / Magnus Egerstedt and Clyde Martin.
p. cm. – (Princeton series in applied mathematics)
Includes bibliographical references and index.
ISBN 978-0-691-13296-9 (hardcover : alk. paper) 1. Interpolation. 2. Smoothing (Numerical analysis) 3. Smoothing (Statistics) 4. Curve fitting. 5. Splines. 6. Spline theory. I. Martin, Clyde. II. Title.
QA297.6.E44 2010
511'.42–dc22

2009034177

The publisher would like to acknowledge the authors of this volume
for providing the camera-ready copy from which this book was printed.

British Library Cataloging-in-Publication Data is available

Printed on acid-free paper. ∞

press.princeton.edu

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To our wives Danielle and Joyce
and
to our friend and mentor Roger Brockett.

PREFACE

In 1998, Clyde Martin visited the Royal Institute of Technology in Stockholm and taught a course on advanced topics in systems theory. Among the students were Magnus Egerstedt, and what started as a homework assignment quickly led to the discovery that it was possible to generalize the smoothing splines concepts, as defined by Grace Wahba in the area of statistics, using standard control theoretic ideas. The key enabling (yet rather obvious) observation was that a rich class of smoothing curves can be traced by the output of a linear control system, driven by an appropriately selected input. However, that the corresponding class of curves captured almost all of the traditional splines, as well as leading to useful new areas of investigation, including monotone splines, splines with continuous data, and splines on manifolds, was more of a surprise. During the last ten years, a rather large body of work has been developed, connecting splining concepts to those found in the systems theory literature. This book is the outcome of that study.

Everyone who reads this book will realize that the basic material owes a great deal to the "red book" of David Luenberger. His concept of optimization using vector space methods is one of those ideas that has had a major influence in engineering, economics, mathematics, and every other area that is concerned with basic optimization. Both authors have taught from and have been taught from his basic book. We have strived to bring to this book some of the readability properties that David has mastered, as well as to connect with the tools and techniques developed by him.

Although the material in this book covers a lot of ground, a word of caution is in order. We have made no attempt to survey the huge field of splines, even of smoothing splines. The purely statistical approach to smoothing splines differs from our approach in the end application, but there is a huge overlap in basic concepts. Anyone who is interested in the statistical approach to smoothing splines is urged to read the seminal monograph of Grace Wahba and the excellent monograph of Randy Eubanks on this topic.

The material in this book relies heavily on a very fruitful collaboration with Professor Yishao Zhou, and it is fair to say that the book would not have been the same without her. In particular the chapter on smoothing splines as integral filters owes much to her and to Professor W. P. (Daya) Dayawansa. Daya has contributed to the content of the book, but, more important, he has been a pioneer in the interface between mathematics and engineering. Both authors are indepted to him for his influence on their scientific philosophy.

The authors have also been heavily involved with Professor Hiroyuki Kano in the development of applications of B-splines. We decided not to

include that material in this book because of our focus on curves generated by linear control systems with a particular optimization format. However, there certainly is some overlap in applications areas that the interested reader is urged to explore through the extensive publications of Professor Kano.

When developing smoothing splines, statistics cannot be ignored. The second author was patiently reminded to use statistics correctly by Shan Sun. Professor Sun was a coauthor on the first two major papers in the development of control theoretic smoothing splines. When statistics are used correctly, she deserves the credit, while the authors assume the responsibility for all statistical bobbles.

The very first paper in this series was a collaborative effort with Professor Zhimin Zhang. That paper mimicked the classical spline construction, and the approach used was quite different from the approach used in this book. Nonetheless, it served as a starting point for our study of the connection between splines and linear systems theory.

There have been many graduate students involved with the development and application of control theoretic splines at Texas Tech University, Royal Institute of Technology in Stockholm, Georgia Institute of Technology, and Stockholm University. Many of their names can be found on papers in the bibliography. We are so very grateful for the work that they have done on this long term project. We thank the past, present, and future students for their diligence.

Modern research cannot be done without financial support. We have been fortunate to have been supported by many agencies: NSF, NASA, NSA, AFOSR, ARO, EPA, NIH, and DARPA. We gratefully acknowledge the support these agencies have provided over the years.

On a personal note, we want to thank our families and friends for supporting our work. In particular, Danielle Hanson has been a constant source of joy, energy, and inspiration to Magnus. Not only has she kept Magnus's mind (somewhat) straight in terms of providing a big picture, but she has also been involved in many technical discussions at the dinner table. Thank you!

Joyce Martin has stood beside Clyde for 45 years and has never flinched. She has understood when mathematics was first on his mind, and she has patiently stood by as he traveled even when there were four small children at home. She has always been ready to be the occasional pro bono editor of his papers and books. Not only does she deserve the credit for this book but for all of the work that Clyde has done!

Atlanta and Lubbock – February 2009

Contents

Chapter One

INTRODUCTION

Splines are ubiquitous in science and engineering. Sometimes they play a leading role as generators of paths or curves, but often they are hidden inside, for example, software packages for solving dynamic equations, in graphics, and in numerous other applications.

The standard, classic spline is an interpolating curve. In contrast to this, smoothing splines are only required to pass "close" to the data points. Such smoothing splines are well know by name in statistics, but not so well known outside of this area. The goal of this book is to show that smoothing splines arise as a natural part of control theory, and that, by using control theoretic concepts, we can construct and interpret smoothing splines in an efficient, algorithmic manner.

Throughout the book, this connection between control theory and smoothing splines will be made explicit, and we will find numerous applications for smoothing splines in path planning for mobile robots, in numerical analysis, graphics, and other basic applications. This introductory chapter presents a brief background to interpolating and smoothing splines, as well as sets up their connection to linear systems theory.

1.1 FROM INTERPOLATION TO SMOOTHING

The basic problem that the classical spline was constructed to solve was as follows: Given a finite set of data points, find a smooth curve that interpolates through these points. Of course, there are infinitely many such curves, and the real task is to devise an algorithm that selects a unique (hopefully exhibiting certain desirable properties) curve. In fact, classical splines solve this problem by requiring that the curve be piecewise polynomial, that is, that it be polynomial between the data points, and that the pieces be connected as smoothly as possible. Often additional conditions must be applied as well at the endpoints to ensure uniqueness.

This idea of producing interpolating polynomials, stitched together at the data points, works wonderfully if the data are exact, or nearly so. Unfortunately, data often have significant error associated with them, and classical splines tend to accent these errors. *Smoothing splines* were developed to

remedy this very problem, that is, to handle cases when there is error associated with the data points. Naturally enough, these smoothing splines were developed in statistics, where noise is a fact of life, and where error is assumed in almost all data. As such, the restriction of exact interpolation was dropped, while the restriction remained that the curves should be piecewise polynomial and as smooth as possible.

Statistics aside, this notion of producing smoothing rather than interpolating curves is rather natural as well in engineering in general, and control theory in particular. In fact, various notions of controllability have always played fundamental roles in engineering through the canonical problem of moving an object at a known position with known dynamics to a new position. For example, in air traffic control, ground control typically will dictate to the pilot of an airplane where it should be at a fixed set of times, and what its corresponding directions should be, for example, the command could be to be at 10,000 feet in 2 minutes with a given heading. The pilot will in fact receive a string of such commands as the plane approaches an airport. Typically, some deviations from the exact locations are allowed, and the size of the deviation depends on many factors. For example, passenger comfort requires that accelerations are minimized, and that transitions are smooth. As a consequence, exact interpolation is not desirable in this case. In fact, the pilot is constructing a type of smoothing spline.

Based on this rather informal observation, it seems natural to give a more explicit description of the general controllability problem in the context of smoothing splines. It was from this rather straightforward idea that the concept under investigation in this book arose, that is, the concept of control theoretic splines.

1.2 BACKGROUND

The problem of approximation is almost as old as modern mathematics. In fact, polynomial interpolation dates back to the mid 1700s, with the work of Edward Waring (Lagrange interpolation). The ideas of polynomial approximation were central during the 1800s, with the development of various families of orthogonal polynomials, and what later became known as the related Hilbert space theories. The polynomial interpolation problems were of such importance that a significant part of modern mathematics can trace its history back to these developments in one form or another. But, if polynomial interpolation is such a well-studied and powerful tool, then why were polynomial splines invented?[1]

[1] By splines, we here mean piecewise polynomial curves that are stitched together at given nodal points in order to ensure certain regularity properties.

1.2.1 Polynomial Interpolating Splines

Traditional (pre-spline) polynomial interpolation has at least two very serious drawbacks, which limit its use in many applications. The first is that a polynomial of degree $n+1$ may have as many as n local extrema. This causes the curve to be very complex. If, for example, we have $n+2$ data points that are connected by curves that are approximately linear, then the interpolating polynomial will have degree $n+1$, and hence will not at all be approximately (piecewise) linear. As such, while we may have a locally good fit, we cannot have a good fit over an arbitrarily large interval.

The second major drawback is an algorithmic problem. To find a polynomial that interpolates a given set of data is equivalent to inverting a van der Monde matrix. The condition number of a van der Monde matrix can grow as 2^{2^N}, with N being the size of the matrix, making the inversion rather intractable in that numerically, the problem of polynomial interpolation may become highly unstable (see e.g., [42]). So, as beautiful as the theory of polynomial interpolation is, it is not particularly useful for large problems.

To remedy this, during the early 1940s, splines as we know them were invented by Isaac Schoenberg at the U.S. Army Ballistic Research Laboratory in Aberdeen, Maryland (the Aberdeen Proving Ground). The splines' early uses are somewhat shrouded in mystery, as this was highly classified research, and it was not until after the Second World War that Schoenberg publicly described his invention.

Schoenberg formulated the spline problem in the following manner. Let $D = \{(t_i, \alpha_i) : i = 1, \ldots, n\}$ be a set of time-stamped data points (with t_i the time stamp and α_i the data point), and let F be the set of twice continuously differentiable functions that interpolate the data, that is, $F = \{f \in C^2[0,T] \mid f(t_i) = \alpha_i\}$. Now, the spline problem is given in terms of the following optimization problem:

$$\min_{f \in F} \max_{t \in [0,T]} |f''(t)|,$$

where f'' denotes second derivative. What this problem entails is to find the interpolating function f that has the smallest maximal second derivative on the interval in question.

While this formulation is very elegant, it is (at least at first glance) not an easy problem to solve. In fact, this formulation constitutes an optimization problem over a notoriously difficult Banach space–the space of continuous functions on a compact set. Luckily, the solution is the classical *cubic spline*.

The observation that the cubic spline (piecewise polynomial curves of degree three) solved Schoenberg's problem led to the development of a host of good numerical algorithms for the construction of the optimal solution,

based explicitly on the cubic nature of the solution polynomials. As a consequence, the focus shifted to an in-depth study of such piecewise (cubic) polynomials, while the original optimization problem was largely ignored for nearly three decades. A comprehensive overview of classical splines can be found in Carl de Boor's *A Practical Guide to Splines* [24].

1.2.2 Polynomial Smoothing Splines

It was not until the early 1970s that Grace Wahba (who appropriately enough happens to be the I. J. Schoenberg Professor of Statistics at the University of Wisconsin-Madison) began to study the use of splines with noisy data, that the underlying optimization problem was revisited. In fact, one of Wahba's most important contributions to the subject was to replace the Banach space problem with the much simpler Hilbert space problem

$$\min_{f \in L_2[0,T]} \int_0^T f''(t)^2 dt + \lambda \sum_{i=1}^{n} (f(t_i) - \alpha_i)^2. \tag{1.1}$$

Here L_2 denotes the Hilbert space of square integrable functions, and $\lambda > 0$ is a weight that determines the tradeoff between the smoothness of the solution and the closeness between curve and data points. An example of interpolating and smoothing curves, as formulated by Schoenberg and Wahba, is given in Figures 1.1 and 1.2.

1.3 THE INTRODUCTION OF CONTROL THEORY

It should be noted already at this point that the formulation in (1.1) requires a certain leap of faith, since most L_2 functions are not differentiable, that is, the second derivative, f'', may not be well defined. However, this small inconvenience can be easily remedied by the use of a little control theory.

Let

$$f''(t) = u(t),$$

and let

$$y(t) = \int_0^t (t - s)u(s)ds. \tag{1.2}$$

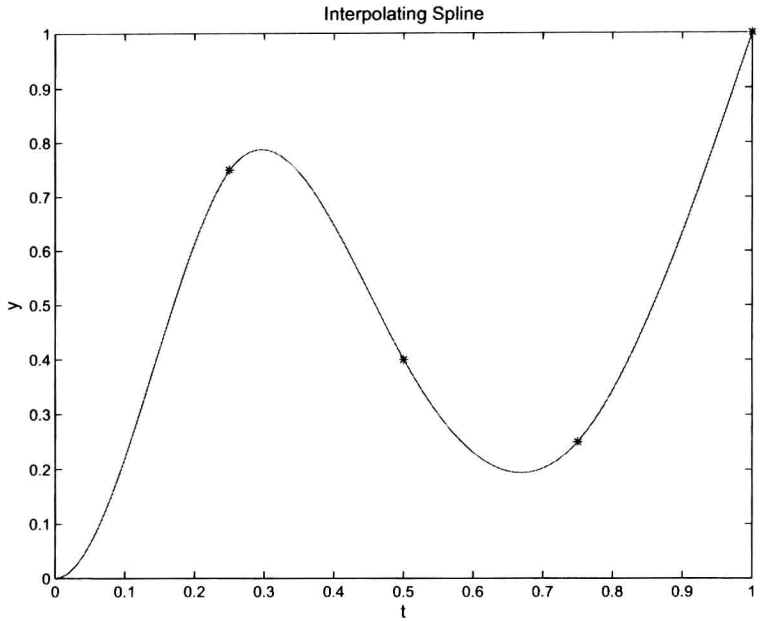


Figure 1.1 Interpolating cubic splines.

Then Wahba's optimization problem can be reformulated as

$$\min_{u \in L_2[0,T]} \int_0^T u(t)^2 dt + \lambda \sum_{i=1}^{n} (y(t_i) - \alpha_i)^2.$$

Based on this formulation, we are only one small step away from the full-fledged control theoretic formulation that will be pursued in this book. In fact, if we simply assume a control system of the form

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

so that

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

we are ready to apply a century of results from linear control theory to the problem of smoothing splines. Note, for example, that the choice of

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

corresponds to the situation in (1.2).

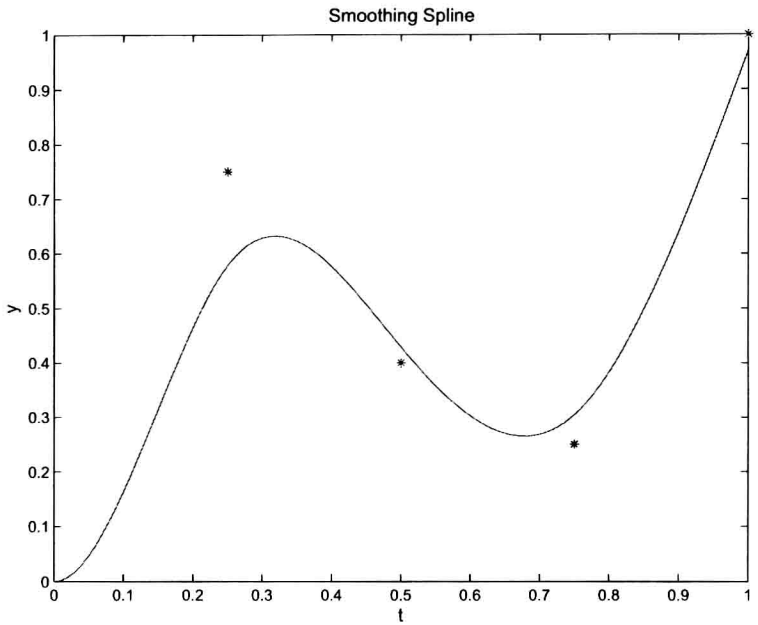


Figure 1.2 Smoothing cubic splines.

1.3.1 When Do Solutions Exist?

It is an easy matter to add additional constraints to the control theoretic formulation of the smoothing spline problem. For example, we can introduce

$$C = \{u \in L_2[0,T] \mid \ell_i(u) = 0,\ i = 1,\ldots,M\},$$

where each ℓ_i is an affine linear functional on $L_2[0,T]$.

We can then ask for the solution to the constrained problem

$$\min_{u \in C} \int_0^T u(t)^2 dt + \lambda \sum_{i=1}^{n} (y(t_i) - \alpha_i)^2.$$

For example, we might let

$$\ell_1(u) = \int_0^1 (t-s)u(s)ds - 1$$

and

$$\ell_2(u) = \int_0^1 (t-s)u(s)ds + 1.$$

Here there is obviously no solution since the constraints are contradictory. A general condition for the solution to this problem to exist is of course that C is nonempty. In fact, as we see later in the book, C defines an affine subspace in $L_2[0,T]$, and the optimization problem is simply asking for the point of minimum norm in that affine space. *And, as long as the affine space is closed there is guaranteed to be a unique solution, as a direct consequence of Hilbert's famous projection theorem.* (See Section 2.3.) As a consequence, we do not need the full machinery of convex optimization, as developed by Rockafellar [82].

As a final comment, it should be noted that if the constraints are nonlinear, the problem is much more difficult. In fact, even if the constraints define a "nice" subspace of L_2, the problem of constructing the optimal control in this case can be (and usually is) very difficult. We will examine a few such problems in this book.

1.4 APPLICATIONS

One of the major goals of this book is to provide tools for applications. To that end, we consider two main categories of applications to which the control theoretic spline is particularly well suited: *path planning* and *statistics*. In fact, even though the major impetus for this work came from path planning–originally the air traffic control problem–it has evolved into a much more general problem involving many autonomous vehicles or even biological entities.

1.4.1 Path Planning

We consider this problem in several chapters in the book. The basic idea is that we are given a set of way points and times, and we ask that the system be at, or near, those points at specified times. We are not very interested in the nature of the error at the way points unless it is too large. If, for example, we are trying to design a path for an autonomous vehicle, we may have to impose restrictions on the curvature of the path at particular points, and this may require iterations over different choices of smoothing parameters (λ) to deliver a suitable path.

1.4.2 Statistics

The major role of splines in statistics is to smooth noisy data. To this end, it is important that the residues be well behaved. This observation has led to a science studying the selection of the smoothing parameter λ in (1.1) to achieve residues that have suitable statistical properties. Hopefully, λ can

be chosen so that the residues are identically normally distributed. This is seldom a goal in engineering. In this book, we use the parameter λ to control bandwidth and do not study the residues as such. However, we will tie into a number of statistically motivated applications, including the production of probability densities using smoothing splines.

1.5 TOPICAL OUTLINE OF THE BOOK

This book is organized into ten chapters (plus the introduction). In Chapter 2, the basic material from control theory is presented as well as the setup for the solution of the optimal control problem. Fundamental concepts from the theory of Hilbert spaces are summarized. Notation is established and, as an example, we revisit the classical controllability problem in the context of Hilbert's projection theorem.

In Chapter 3, we describe eight problems that are fundamental in the area of interpolation and smoothing, and that will serve as motivation for the subsequent chapters. Rather than providing complete solutions to these problems, this chapter should be thought of more as a road-map for this book (and beyond). The eight problems are (1) interpolating splines, (2) interpolating splines with constraints, (3) smoothing splines, (4) smoothing splines with constraints, (5) monotone smoothing splines, (6) dynamic time warping, (7) model following, and (8) trajectory planning. Problems (1) and (3) are basic to the material in the book. The other six problems are important as applications and constitute refinements of the two basic problems.

In Chapter 4, we consider the general problem of smoothing splines from the viewpoint of Hilbert spaces. In some sense, this chapter is the major contribution of control theory techniques to the spline problems, and constitutes the core of the book. We show that the general smoothing splines result from an application of Hilbert's projection theorem, and that we are able to add any finite number of linear constraints to the formulation and still have an effective algorithmic solution.

In Chapter 5, we show that control theoretic splines have the properties that we expect from splines–suitable approximation properties. We show, for example, that if we are given a smooth curve, then, as the number of data points approaches a dense set, the sequence of splines converges in an appropriate manner to this underlying curve. We also show that if noise is added, we still maintain convergence.

In Chapter 6, we consider an extension of the smoothing spline problem with finite/discrete data (a finite/discrete collection of data points) to the problem of smoothing splines with continuous data. This problem is in some real sense a filtering problem. The data can be considered to be the

output of some machine, and we are trying to find a smooth approximation of these data. The smoothing spline formulation lends itself well to the problem. In this chapter we also consider the problem of recursive splines as a natural tool for tackling the continuous data problem.

Chapter 7 deals with the question of how to produce splines with certain regularity properties. In particular, we discuss how to produce splines that are monotone in the sense of having nonnegative first or second derivatives. The main theorem in this chapter is that for nilpotent systems, the optimal curve is still piecewise polynomial despite the monotonicity constraints, while the problem is completely solved using dynamic programming for the case of monotone cubic splines. The monotone smoothing problem is of importance in a number of applications ranging from economics to biology. In this chapter we also discuss the related problem of constructing probability density functions from data, which is an example of a much larger problem involving continuous constraints.

In Chapter 8, we further consider the application of smoothing splines to statistics by showing that the smoothing spline can be considered as an approximation to an explicit linear filter. The resulting construction will based on linear-quadratic optimization and its associated theories of Hamiltonians and Riccati transforms.

In Chapter 9, we consider a variation of the smoothing spline problem–transfer between affine varieties. An example where this problem arises is considered in detail (path planning for multi-robot systems), and the problem of transfer between affine varieties is solved in its full generality, although introduced and motivated by this particular example problem. Interestingly, this transfer problem can be considered as a control problem on the manifold of affine subspaces.

In Chapter 10, we consider some applications to path planning. In particular, we study the problem of planning paths for multiple airplanes close to an airport and the problem of reconstructing the paths executed by sea turtles, based on telemetric data. As a consequence of this, we are forced to construct splines on spheres instead of on "flat" Euclidean spaces.

Finally, in Chapter 11, we show that there are classes of problems that do not fall into the Hilbert space setting but are still important and can be solved. The particular application under investigation in this context is the classic problem of selecting appropriate nodal (or data) points. In other words, where do you put the sensors to obtain the information that you need for control? In the context of polynomial interpolation, this problem was of interest a hundred years ago, and it still remains an important and primary problem in certain engineering fields.