INTERIOR POINT SOLUTIONS OF VARIATIONAL PROBLEMS AND GLOBAL INVERSE FUNCTION THEOREMS*

CHRISTOPHER I. BYRNES† AND ANDERS LINDQUIST‡

Abstract. Variational problems and the solvability of certain nonlinear equations have a long and rich history beginning with calculus and extending through the calculus of variations. In this paper, we are interested in "well-connected" pairs of such problems which are not necessarily related by critical point considerations. We also study constrained problems of the kind which arise in mathematical programming as well as constraints of a geometric nature where a solution is sought on a leaf of a foliation. In these cases we are interested in interior minimizing points for the variational problem and in the well-posedness (in the sense of Hadamard) of solvability of the related systems of equations. We first prove a general result which implies the existence of interior points and which also leads to the development of certain generalization of the Hadamard-type global inverse function theorem, along the theme that uniqueness quite often implies existence. This result is illustrated by proving the non-existence of shock waves for certain initial data for the vector Burgers' equation. The global inverse function theorem is also illustrated by a derivation of the existence of positive definite solutions of matrix Riccati equations without first analyzing the nonlinear matrix Riccati differential equation. The main results on the existence of solutions to geometrically constrained well-connected pairs are then presented and illustrated by a universal solution to the generalized moment problem, with a nonclassical complexity constraint, given by a a geometrically derived minimization problem for a strictly convex nonlinear functional. A more general result is then illustrated by a geometric analysis of the existence of interior points for linear programming problems. In a final section, our solution to the generalized moment problem is applied to two interpolation problems, arising in signal processing and systems theory. These moment problems are of the Carathéodory, and of the Nevanlinna-Pick, type, respectively. The nonclassical complexity constraints reflect constraints on the physical synthesis of the corresponding filters or circuits.

Key words. Variational problems, well-posedness of solvability of nonlinear equations, global inverse function theorems, constrained optimization, foliations, interior point methods, Burgers' equation, generalized moment problems, interpolation problems.

AMS subject classifications. Primary: 49J99, 49Q99, 57R30. Secondary: 30E05, 49K40, 49N45.

---

* This research was supported in part by grants from AFOSR, TFR, the Göran Gustafsson Foundation, and Southwestern Bell.

† Dept. of Systems Science and Mathematics, Washington University, St. Louis, MO 63130, USA.

‡ Department of Mathematics, Royal Institute of Technology, 100 44 Stockholm, Sweden.

1. Introduction

Consider two n-manifolds M and N and a continuous map

$$f : M \to N. \tag{1.1}$$

We assume that N is connected. Following Hadamard, one is typically interested in the existence of solutions to the equations

$$f(x) = y, \tag{1.2}$$

the uniqueness of solutions, and the continuous dependence of solutions on the y. More formally, the problem of finding solutions to (1.2) is said to be *well-posed* provided the map (1.1) satisfies:

(i) f is surjective
(ii) f is injective
(iii) f has a continuous inverse.

In geometric terms, the solvability problem is well-posed provided f is a homeomorphism. One consequence of well-posedness is therefore that the map f is *proper*, i.e. that, for any compact subset K of N, $f^{-1}(K)$ is compact. In other words, given *a priori* bounds on the right hand side of (1.2), there exists bounds on the solutions to (1.2).

We shall also consider variational problems of the following form. For each $y \in N$, let $\mathcal{Q}_y$ be a closed connected subset of a topological space $\mathcal{Z}_y$. In many applications $\mathcal{Q}_y$ and $\mathcal{Z}_y$ will not vary with y, in which case we suppress the subscripts. For each $y \in N$, suppose there is a function

$$J_y : \mathcal{Q}_y \to \mathbb{R} \cup \{\infty\}. \tag{1.3}$$

Here, one is often interested in computing

$$\inf_{q \in \mathcal{Q}_y} J_y(q), \tag{1.4}$$

and in finding all q_0 for which

$$J_y(q_0) = \inf_{q \in \mathcal{Q}_y} J_y(q). \tag{1.5}$$

Denoting by $\mathrm{int}(\mathcal{Q}_y)$ the set of interior points of $\mathcal{Q}_y$ in its relative topology, we are particularly interested in finding minimizers in $\mathrm{int}(\mathcal{Q}_y)$. We refer to such a point as an *interior minimizing point for J_y*. While for concave problems we are interested in interior maximizing points, there are important problems when we would want to know about the uniqueness of saddle points. For these reasons, and others illustrated throughout this paper, our basic definition will include general critical points.

Definition 1.1. Consider a pair of problems of the form formulated above. This pair is said to *well-connected* if, for all $y \in N$, there exists an injection from the solution set of (1.2) into the set of interior critical points of J_y.

Of course, one instance of well-connectedness arises from an analysis of the stationarity of interior points for variational problems. Conversely, there is a finite dimensional analogue of the inverse problem, or Dirichlet Principle, of the calculus of variations. More explicitly, suppose

$$f : \mathbb{R}^n \to (\mathbb{R}^n)^*$$

is a smooth map. This map defines a 1-form on $\mathbb{R}^n$ which can be expressed as

$$\omega = \sum_{i=1}^{n} f_i dx_i$$

in terms of a global coordinate system $(x_1, x_2, \cdots, x_n)$. If $d\omega = 0$, then there exists a smooth functional J such that these two problems are well-connected. Indeed, since ω is closed on $\mathbb{R}^n$, it is exact; i.e.,

$$dJ = \omega$$

for some smooth $J : \mathbb{R}^n \to \mathbb{R}$. Since

$$D^2 J(x) = \mathrm{Jac}(f)(x),$$

$\mathrm{Jac}(f)$ is symmetric, and we now assume that $\mathrm{Jac}(f)(x) > 0$ for all $\in \mathbb{R}^n$ so that

$$J_y(x) = J(x) - y(x)$$

is strictly convex for each $y \in (\mathbb{R}^n)^*$. Alternatively, by the second derivative test, J_y has only minima as critical points, and to say that J_y achieves its minimum at x is to say that $f(x) = y$. Now suppose f is proper. By Hadamard's global inverse function theorem [30, 31, 32], $f : \mathbb{R}^n \to (\mathbb{R}^n)^*$ is a diffeomorphism so that the problem, $f(x) = y$, is well-posed, and, in fact, J_y always achieves its minimum.

As an elementary example consider the equation

$$F(z) = c, \tag{1.6}$$

where F a holomorphic function of one variable and $c = c_1 + ic_2$. Since the 1-form $(F(z) - c)dz$ is exact, taking the imaginary part of a primitive yields a functional $J_c : \mathbb{C} \to \mathbb{R}$ such that

$$dJ_c = (v - c_2)dx + (u - c_1)dy.$$

where $F(x, y) = u(x, y) + iv(x, y)$. In particular, the pair (F, J_c) is well-connected. Moreover, by the Cauchy-Riemann equations, the Hessian $D^2 J_c$ is an indefinite symmetric matrix satisfying

$$\det D^2 J_c \le 0$$

with equality at z_0 if and only if $F'(z_0) = 0$. In particular, if $F'(z) \neq 0$, (1.6) is well-connected to a variational problem for which the critical points are nondegenerate saddle-points. Interestingly, J_c has a unique critical point for all c if and only if F is univalent.

Remarkably, a similar situation arises in the analysis of the generalized moment problem of classical analysis. For simplicity, we restrict to the case of real quantities, deferring the complex case to Section 3. Consider a sequence of real numbers $c_0, c_1, \cdots, c_n$ and a sequence of continuous, linearly independent real-valued functions $\alpha_0, \alpha_1, \cdots, \alpha_n$ defined on the real interval $[a, b]$. The moment problem is then to find all monotone, nondecreasing functions μ of bounded variation such that

$$\int_a^b \alpha_k(t) d\mu(t) = c_k, \quad k = 0, 1, \cdots, n, \tag{1.7}$$

where the sequence $c_0, c_1, \cdots, c_n$ is positive in the following sense. Let $\mathfrak{P}$ be the subspace of $C[a, b]$ spanned by the functions

$$\alpha_0, \alpha_1, \cdots, \alpha_n$$

and let $\mathfrak{P}_+$ be the subset of $p \in \mathfrak{P}$ that are positive on $[a, b]$. It is typically assumed that $\mathfrak{P}_+$ is nonempty, in which case $\mathfrak{P}_+$ is an open convex subspace of $\mathfrak{P}$. One says that the sequence $c_0, c_1, \cdots, c_n$ is *positive* if and only if

$$\langle c, q\rangle := \sum_{k=0}^{n} q_k c_k > 0 \tag{1.8}$$

for all $q := (q_0, q_1, \cdots, q_n) \in \mathbb{R}^{n+1}$ such that

$$\sum_{k=0}^{n} q_k \alpha_k \in \mathfrak{P}_+. \tag{1.9}$$

Denote by $\mathfrak{C}_+$ the space of positive sequences. The fact that $\mathfrak{P}_+$ is nonempty implies that $\mathfrak{C}_+$ is nonempty. Therefore, $\mathfrak{C}_+$ is also a convex open subset of the space of real sequences of length $n + 1$.

Motivated by several applications in speech synthesis, robust control and signal processing (see Section 4), we introduce the complexity constraint

$$\frac{d\mu}{dt} = \Phi(t) = \frac{P(t)}{Q(t)}, \quad P, Q \in \mathfrak{P}_+. \tag{1.10}$$

Fix P and define the function

$$F : \mathfrak{P}_+ \to \mathfrak{C}_+$$

componentwise via

$$F_k(Q) = \int_a^b \alpha_k(t) d\mu(t)$$

Parameterizing Q via $Q = \sum_{k=0}^{n} q_k \alpha_k$, we construct the 1-form

$$\omega_c = \sum_{k=0}^{n} \left[c_k - F_k(\mu)\right] dq_k,$$

on $\mathfrak{P}_+$ and observe that ω_c is closed. Therefore, by the Poincaré Lemma, there exists a smooth function J_c such that $dJ_c = \omega_c$, leading to the construction of a well-connected pair for the generalized moment problem with complexity constraints. More explicitly,

$$J_c = \int \omega_c,$$

with the integral being independent of the path between two endpoints. Therefore, since

$$\begin{aligned}\omega_c &= \sum_{k=0}^{n} \left[c_k - \int_a^b \alpha_k \frac{P}{Q} dt\right] dq_k \\ &= \sum_{k=0}^{n} c_k dq_k - \int_a^b \frac{P}{Q} dQ dt,\end{aligned}$$

computing the path integral

$$\int_{Q_0}^{Q_1} \omega_c = \left[\langle c, q\rangle - \int_a^b P \log Q dt\right]_{Q_0}^{Q_1},$$

we obtain, modulo a constant of integration,

$$J_c(Q) = \langle c, q\rangle \ - \int_a^b P \log Q dt.$$

In general, since J_c is strictly convex, any interior critical point is a nondegenerate minimum, yielding uniqueness of a distribution of constrained complexity solving the moment problem. In Section 3, we shall address existence.

As a final example, these conditions are also always satisfied for the pair of well-connected problems arising from the Ritz approximations of a strictly convex variational problem and, as pointed out by Hilbert, its fully elliptic Euler Lagrange equation. For example, following [52], let G be a bounded region in $\mathbb{R}^N$ with a piecewise smooth boundary ∂G, and let $\bar{G} = G \cup \partial G$ be the closure of G. Consider the variational problem

$$\inf_{u\in C^1(\bar{G})} J_h(u), \quad u = 0 \text{ on } \partial G$$

where

$$J_h(u) = \int_G \left(\frac{1}{2}\sum_{k=1}^{N} (\frac{\partial u}{\partial x_k})^2 + g(u) - uh \right) dx.$$

The corresponding Euler-Lagrange equations are given by

$$\begin{aligned} \Delta u - g'(u) &= -h \qquad && \text{on } G \\ u &= 0 && \text{on } \partial G, \end{aligned}$$

and they are equivalent to the variational problem, for suitable g and sufficiently smooth u ($u \in C^2(\bar{G})$). If $V \subset C^2(\bar{G})$ is finite-dimensional subspace, Ritz's method gives two problems, namely the optimization problem

$$\inf_{v\in V} \varphi(v) - \langle b, v\rangle,$$

where φ is a smooth function, and the problem to solve the equation

$$\varphi'(v) = b$$

for $b \in V$. In this case, φ is strictly convex, φ' is strictly monotone and proper (coercive), and

$$D^2\varphi > 0.$$

In this paper, however, we are also interested in well-connected pairs where f need not generate an exact 1-form and need not have an everywhere nonsingular Jacobian. And, we are particularly interested in constrained problems of two types: the constraints arising in mathematical programming and the geometric constraints that solutions lie on a particular leaf of a foliation.

In an ascending order of complexity, in Section 2 we begin with the case where the leaf is the entire manifold M and prove a "duality" theorem for well-connected pairs which asserts that problem (1.2) is well-posed and the variational problem (1.4) has an interior critical point provided f is proper and J_y has at most one critical point. Of course, the case when $\mathcal{Q}_y$ is convex and J_y is strictly convex is particularly easy to apply. The "duality" theorem is then illustrated by a well-posed pair of problems concerning shock waves in an interval $[t_0, T]$ for solutions of the (vector) inviscid Burgers' equation. We begin by defining a variational problem for which, when viewed as an optimal control problem, the Riccati partial differential equation [6] for the gradient of the value function turns out to be the Burgers' equation in backwards time. We then construct, for each $t \in [t_0, T]$, a map f_t for which well-posedness is equivalent to the absence of shock waves at time t. Several criteria for properness of this map are given in terms of qualitative behavior of the initial data. After rendering the variational problem finite dimensional , it is shown that these two problems are well-connected. In particular, convexity of the terminal constraint condition is equivalent to monotonicity of the intial condition, which is related (equivalent in the scalar case) to the absence of shock waves. Moreover, in the nonconvex case, the formation of multiple critical points is related to the formation of shock waves. Finally, an example is given where a parameter variation in the initial condition causes an onset of shock waves that is reflected, as a finite-dimensional "shadow," in the critical point equation undergoing a pitchfork bifurcation.

The proof of our basic result in Section 2 belongs to a long tradition of global inverse function theorems, relying on a lemma that uniqueness of solutions to certain finite dimensional nonlinear problems implies existence. The lemma is illustrated by showing existence of positive definite solutions to the algebraic matrix Riccati equation, for which it is relatively simple to establish uniqueness.

In Section 3, we formulate a theorem for well-posedness and for existence of interior minima for problems constrained to an arbitrary leaf of a foliation. To illustrate how such problems arise, we consider here the generalized moment problem over the complex number field. In the real case considered above, fixing either P or Q defines leaves of two complementary foliations. While in this paper we will encounter less trivial foliations, the remarkable fact is that, restricted to these leaves, there is a natural variational problem which is well-connected to the moment problem.

For constrained well-connected pairs, the existence of a pair of foliations arises very naturally. While the mappings obtained by restricting f to the leaves of a constraint foliation do not always generate closed (or exact) 1-forms, we are able to obtain enhancements of the basic duality theorem in terms of geometric properties of the corresponding foliations. In Section 3, we illustrate this for the nontrivial problem of the analysis of of interior point methods for finite-dimensional linear programming.

In section 4, the results obtained in Section 3 for the generalized moment problem are illustrated by interpolation problems arising in spectral estimation and signal processing, and in applications to robust controller design. As it turns out, the first interpolation problem is the classical Carathéodory extension problem with a nonclassical complexity constraint defined by the rationality of, and a degree constraint on, the interpolant. The second interpolation problem is the classical Nevanlinna-Pick interpolation, again with the same nonclassical complexity constraints. In each case the complexity constraints reflect the physical synthesizabilty of a filter, or circuit, that solves the relevant interpolation problem.

2. The unconstrained problem

Not surprisingly, a great deal can be said about well-connected pairs of problems, especially when either the number of solutions to (1.2) or the number of minimizers for J_y is independent of y. One such example is given in the following result.

Theorem 2.1. *Consider a well-connected pair of problems. If f is proper, and, for each $y \in N$, J_y has at most one critical point, then*

(i) *problem* (1.2) *is well posed;*
(ii) *for each* $y \in N$, J_y *has a unique critical point which is an interior point.*

Example 2.2. (*Variational methods for the absence of shock waves in Burgers' equation.*) Consider the (vector) inviscid Burgers' equation

$$\frac{\partial \Pi}{\partial t} = \frac{\partial \Pi}{\partial x}\Pi, \quad \Pi(x, 0) = \Pi(x), \tag{2.1}$$

where $x \in \mathbb{R}^n$ and Π is a smooth map $\Pi : \mathbb{R}^n \to \mathbb{R}^n$.

We are interested in long-time existence of solutions on arbitrary intervals $[t_0, T]$, with the obstructions to existence being finite escape time or the existence of a shock wave at some time t. We shall first define a variational problem for which, when viewed as an optimal control problem, the Riccati partial differential equation [6] for the gradient of the value function turns out to be the vector inviscid Burgers' equation in backwards time.

For $Q : \mathbb{R}^n \to \mathbb{R}$ a smooth map, the canonical equations

$$\begin{aligned} \dot{x} &= p, & x(t_0) &= x_0 \in \mathbb{R}^n, \\ \dot{p} &= 0, & p(T) &= -\nabla Q(x(T)) \in \mathbb{R}^n \end{aligned} \tag{2.2}$$

for the variational problem

$$V(x, t_0) = \inf \left\{ \int_{t_0}^{T} \frac{1}{2}\|\dot{x}\|^2 dt + Q(x(T)) \right\} \tag{2.3}$$

generate characteristic curves

$$p = \Theta(x(t)) = \Theta(x, t),$$

which satisfy a related equation

$$0 = \dot{p} = \frac{\partial \Theta}{\partial t} + \frac{\partial \Theta}{\partial x}\Theta$$

with terminal condition

$$\Theta(x, T) = \nabla Q(x).$$

In particular, a solution of Burgers' equation with

$$\Pi(x, 0) = \nabla Q(x)$$

yields a solution (backwards in time)

$$p(T - t) = \Theta(x, T - t) = \Pi(x, t)$$

of the canonical equations (2.2). Denote the "time t" map of the canonical equations by

$$\Phi_t : (x_0, p_0) \mapsto \Phi_t(x, p) = (x(t; x_0), p(t; p_0)).$$

In order to define the other ingredient in a well-connected pair, we also consider the two projection maps

$$\mathrm{proj}_1(x, p) = x \quad \text{and} \quad \mathrm{proj}_2(x, p) = p,$$

defined on $\mathbb{R}^{2n}$. If Q is C^{k+1}, then

$$M_t = \Phi_t(\mathrm{graph}(\Pi(x, 0)))$$

is a smooth (C^k) n-manifold for all $t \geq 0$. Now define the smooth map $f_t : \mathbb{R}^n \to \mathbb{R}^n$ via

$$f_t(x_0) = \text{proj}_1\{\Phi_t(x_0, \nabla Q(x_0))\}.$$

To say that the problem of solving

$$f_t(x) = y$$

is well-posed is to say that M_t is the graph of a C^k function. If this is the case for all $t \in [t_0, T]$, then (see [7, 8])

$$M_t = \text{graph}(\Pi(x,t)) = \text{graph}(V'(x,t))$$

so that the "value" function V of the variational problem is C^{k+1} and is a classical solution of the Hamilton-Jacobi Theorem. In general, if V is C^2, then it is C^{k+1}, and, in fact, a shock wave for Burgers' equation occurs precisely when V fails to be twice differentiable, and in this case the (Lagrangian) submanifold M_t is a "generalized solution" of Burgers' equation.

We now show that these problems are well-connected. Integrating the canonical equations (or the Euler-Lagrange equations) to obtain

$$x(t; x_0) = pt + x_0, \quad \text{for } p \text{ a constant vector in } \mathbb{R}^n,$$

one can see that f_t is proper provided either ∇Q is bounded or $\langle \nabla Q(x), x\rangle \geq 0$ for all $x \in \mathbb{R}^n$. Now, the critical points of

$$W(x_0, p) = \frac{1}{2}\int_{t_0}^{T} \|\dot{x}\|^2 dt + Q(p(T-t_0) + x_0)$$

are characterized by the terminal constraint equations

$$p = -\nabla Q(p(T-t_0) + x_0).$$

Therefore, every solution x_0 of

$$f_t(x) = y_0$$

determines, in a $1-1$ fashion, a critical point

$$p_0 = \text{proj}_2(\Phi_t(x_0, \nabla Q(x_0)))$$

of W, but not necessarily a minimum. Indeed, if x_i satisfy $f_t(x_i) = y_0$ for $i = 1, 2, \cdots, N$, and p_i are the corresponding "costates", then $(y_0, p_i) \in M_t$ generate extremal trajectories for the variational problems, which may correspond to local minima, local maxima, or inflections, as we shall illustrate by example.

We next investigate when the variational problem has a minimum as its only critical point. Analytically, the Hessian of $W(x_0, p)$ at a critical point is given by

$$D^2W(x_0, p) = I + (T-t_0)D^2Q(p(T-t_0) + x_0).$$

In particular, if $D^2Q(x) > 0$ for $x \in \mathbb{R}^n$, then $W(x_0, p)$ has only nondegenerate minima as critical points.

One can also see this directly from the variational problem. Indeed, suppose that Q is strictly convex. Then, the variational problem (2.3) is a strictly convex optimization problem defined on an affine subset of a Hilbert space and therefore has a unique minimum point. Moreover, the problem of solving

$$f_t(x) = y$$

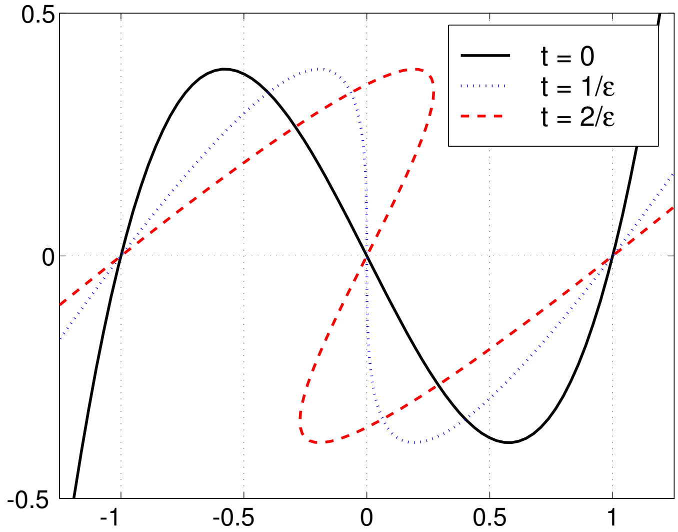

Figure 1.

is well-posed so that there do not exist shock waves. Conversely, suppose $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is an initial condition for (2.1), and consider the submanifold

$$\mathrm{graph}(\varphi) = \{(x, p) \mid p = \varphi(x)\} \subset \mathbb{R}^{2n},$$

where we now consider $\mathbb{R}^{2n}$ as a symplectic manifold with the standard symplectic form $dx \wedge dp$. To say that graph(φ) is a Lagrangian submanifold is of course to say that there exists a Q such that

$$\nabla Q(x) = \varphi(x),$$

which also occurs if, and only if, the 1-form

$$\omega = \sum_{i=1}^{n} \varphi_i dx_i$$

is exact, $\omega = d\varphi$. We call φ Lagrangian in this case. Suppose this is the case and that φ is strictly monotone; i.e.,

$$\langle \varphi(x) - \varphi(y), x - y \rangle > 0$$

for $x, y \in \mathbb{R}^n$ with $x \neq y$. In this case, Q is strictly convex. Therefore, there exists no shock waves as we solve (2.1) with $\Pi(x, 0) = \varphi(x)$ whenever φ is a strictly monotone, Lagrangian map.

If $n = 1$, every function is Lagrangian, so this argument yields a variational proof of the classical fact that shock waves do not occur for monotone increasing initial data.

Finally, consider the explicit example with

$$Q(x) = \frac{1}{4}(x^4 - 2\epsilon x^2).$$

When $\epsilon = 0$, there are no shocks, but for $\epsilon > 0$, a shock occurs at $T = 1/\epsilon$. For $T > 1/\epsilon$ the pitchfork bifurcation for $Q'(x) = 0$ produces a shock wave yielding three extremals, two minima and a maximum, for sufficiently small x, as can be seen in Figure 1.

The proof of the duality theorem follows from Brouwer's Theorem on Invariance of Domain [46, page 3].

Proof of Theorem 2.1. Since J_y has at most one minimizing point, and the pair of problems is well-connected, f is injective. In this case, uniqueness implies existence of solutions, as stated in the ensuing lemma. Matters being so, since the pair of problems is well-connected, for each $y \in N$, J_y has a minimizing interior point, which is therefore the unique minimizer for J_y. □

Lemma 2.3. *Suppose M and N are n-dimensional, topological manifolds and that N is connected. Consider a continuous map $f : M \to N$. Then, f is a homeomorphism if and only if f is injective and proper. In this case, M is connected.*

Proof. If f is a homeomorphism, then f is injective, and since f^{-1} is continuous f is proper. Conversely, if f is injective, f is an open mapping by Brouwer's Theorem on Invariance of Domain. In particular, $f(M)$ is a nonempty open submanifold of N. Since f is proper, f is a closed mapping so that $f(M)$ is a closed subset of N. Since N is connected, f is surjective. Finally, since f is a closed mapping, f^{-1} is continuous. □

That uniqueness implies existence for a class of nonlinear problems has several precedents. Generalizing the linear case, it is known [5, 43] that injective polynomial maps from $\mathbb{R}^n$ to $\mathbb{R}^n$ are surjective. One can also extend Lemma 2.3 to the case where f is locally injective and N is simply connected. This would also follow from the Banach-Mazur Theorem [4, p. 221] after one observes, by Brouwer's Theorem, that locally injective maps of n-manifolds are also local homeomorphisms. In the smooth category for the same class of smooth manifolds, Hadamard's Theorem asserts that smooth, proper maps with nonvanishing Jacobian are diffeomorphisms.

These global inverse function theorems give criteria for problems to be well-posed. The following example illustrates the use of uniqueness implying existence for the algebraic Riccati equation, replacing a standard argument requiring an analysis of the nonlinear matrix Riccati differential equation.

Example 2.4. (*Uniqueness implies existence for algebraic Riccati equations.*) As is well-known, the algebraic Riccati equation

$$PA + A^{\mathsf{T}}P - PBB^{\mathsf{T}}P = -C^{\mathsf{T}}C \tag{2.4}$$

plays a crucial role in infinite-horizon linear-quadratic optimal control. Recall that, in that problem, one considers the optimal control problem to infimize

$$\int_0^\infty (\|y(t)\|^2 + \|u(t)\|^2)dt, \tag{2.5}$$

subject to the constraints

$$\dot{x} = Ax + Bu, \quad x(0) = x_0 \tag{2.6}$$

$$y = Cx \tag{2.7}$$

Here x is square integrable as a function on $(0, \infty)$ taking values in $\mathbb{R}^d$ and u is square integrable as a function on $(0, \infty)$ taking values in $\mathbb{R}^k$. Thus, A, B and C are real matrices of dimensions $d \times d$, $k \times k$ and $m \times d$, respectively. Without loss of generality, we may take $m \leq d$. We assume that the system (2.6) is *controllable*; i.e., for any time

$t_1 > 0$ and any initial state $x_0 \in \mathbb{R}^d$, there exists a control u defined on the interval $[0, t_1]$, which drives the trajectory of (2.6) from state x_0 and time $t = 0$ to state 0 at time $t = t_1$. Controllability ensures that the integral (2.5) has a finite infimum. We also assume that the system is *observable*; i.e., if $y(t) \equiv 0$ on some interval $[0, t_1]$, then $x_0 = 0$.

In the solution of such optimal control problems it is important to show the existence of a unique positive definite solution to (2.4). Here we shall use Lemma 2.3 to do this. To this end, we first use a standard argument to show uniqueness. In fact, let P be any positive definite solution of (2.4), and consider the control law

$$u(t) = -B^{\mathsf{T}} P x(t), \tag{2.8}$$

which yields the solution $x(t) = e^{\Gamma t} x_0$, where $\Gamma := A - BB^{\mathsf{T}} P$. Then P is also a positive definite solution to

$$\Gamma^{\mathsf{T}} P + P\Gamma = -PBB^{\mathsf{T}} P - C^{\mathsf{T}} C, \tag{2.9}$$

and consequently, since the right member of the Lyapunov equation (2.9) is negative semidefinite, Γ must have all its eigenvalues in the closed left half-plane $\operatorname{Re} z \le 0$. However, observability implies that none of these eigenvalues lies on the imaginary axis, for if there were such eigenvalues we could choose the initial value x_0 in the corresponding eigenspace, thus producing a solution $x(t) = e^{\Gamma t} x_0$ which is periodic; i.e., $x(t_0) = x(t_1)$ for some $t_0 \ne t_1$. Computing the rate of change of the quadratic "candidate" Lyapunov function $x^{\mathsf{T}} P x$ along this trajectory, we obtain

$$\frac{d}{dt} x^{\mathsf{T}} P x = x^{\mathsf{T}} (\Gamma^{\mathsf{T}} P + P\Gamma) x = -x^{\mathsf{T}} (PBB^{\mathsf{T}} P + C^{\mathsf{T}} C) x,$$

so integrating from t_0 to t_1 we obtain

$$\int_{t_0}^{t_1} (\|B^{\mathsf{T}} P x\|^2 + \|Cx\|^2) dt = 0.$$

However, then $B^{\mathsf{T}} P x \equiv 0$ and $Cx(t) \equiv 0$ on $[t_0, t_1]$, contradicting observability. Consequently, Γ has all its eigenvalues in the open left half-plane, and hence the control law (2.8) is stabilizing so that, in particular, both x and u are square integrable.

Now, if u is any square-integrable control for which x is also square-integrable, then first computing the rate of change of $x^{\mathsf{T}} P x$ along this trajectory, we have

$$\frac{d}{dt} x^{\mathsf{T}} P x = x^{\mathsf{T}} P A x + u^{\mathsf{T}} P B u + x^{\mathsf{T}} A^{\mathsf{T}} P x + u^{\mathsf{T}} B^{\mathsf{T}} P x.$$

Next, integrating from 0 to ∞ yields

$$-x_0{}^{\mathsf{T}} P x_0 = \int_0^\infty (x^{\mathsf{T}} P A x + u^{\mathsf{T}} P B u + x^{\mathsf{T}} A^{\mathsf{T}} P x + u^{\mathsf{T}} B^{\mathsf{T}} P x) dt.$$

In particular,

$$\int_0^\infty (\|y(t)\|^2 + \|u(t)\|^2) dt = x_0{}^{\mathsf{T}} P x_0 + \int_0^\infty \|u + B^{\mathsf{T}} P x\|^2 dt. \tag{2.10}$$

From this one sees that, if P were to exist, the control law (2.8) would be optimal for all initial data x_0, resulting in a minimum cost $J(\hat{x}, \hat{u}) = x_0{}^{\mathsf{T}} P x_0$. Supposing that $\tilde{P}$

is another positive definite solution of (2.4), we obtain, as above, $J(\hat{x},\hat{u}) = {x_0}^{\mathsf{T}}\tilde{P}x_0$, i.e.,

$$ {x_0}^{\mathsf{T}}(\tilde{P}-P)x_0 = 0. $$

However, this is valid for all $x_0 \in \mathbb{R}^n$, and hence $\tilde{P} = P$. Therefore, the algebraic Riccati equation (2.4) has at most one positive definite solution.

To show that uniqueness of the positive definite solution implies existence of such a solution, we apply Lemma 2.3. To this end, we first assume that $Q := C^{\mathsf{T}}C$ is positive definite. Consider the two n-manifolds, where $n := d(d+1)/2$,

$$ M = \{P \mid P > 0 \text{ and } -PA - A^{\mathsf{T}}P + PBB^{\mathsf{T}}P > 0\} $$

and

$$ N = \{Q \mid Q > 0\}. $$

Note that N is connected. We define the continuous map $f : M \to N$ to be

$$ f(P) = -PA - A^{\mathsf{T}}P + PBB^{\mathsf{T}}P. $$

We have just established that f is injective.

To see that f is proper, we first note that

$$ f(\partial M) \subset \partial N. $$

Then, all that remains to be proven to establish properness is show that, if $Q_k \to Q \in N$ as $k \to \infty$ and $P_k \in M$ with $f(P_k) = Q_k$, the $\|P_k\| \le c$ for some constant c. To this end, let $\tilde{u}$ be a control that drives the controllable system (2.6) from state x_0 at time $t = 0$ to state 0 at time $t = t_1$ and that is identically zero for $t \ge t_1$, and let $\tilde{x}$ be the corresponding state trajectory. Then, since P_k is a solution to the algebraic Riccati equation $f(P) = Q_k$, it follows from (2.10) that

$$ {x_0}^{\mathsf{T}}P_kx_0 \le \int_0^{t_1} \left(\tilde{x}^{\mathsf{T}}Q_k\tilde{x} + \|\tilde{u}\|^2\right) dt. $$

However, since $Q_k \to Q$, there is a matrix $\tilde{Q}$ such that $Q_k \le \tilde{Q}$. Consequently, for each $x_0 \in \mathbb{R}^n$, there is a bound $c(x_0)$ such that ${x_0}^{\mathsf{T}}P_kx_0 \le c(x_0)$ for all k. Bounding ${x_0}^{\mathsf{T}}P_kx_0$ in this way for each x_0 in a basis in $\mathbb{R}^d$ provides an upper bound for P_k and thus establishes the required bound on $\|P_k\|$. Consequently, f is proper, so, by Lemma 2.3, f is a homeomorphism. We also observe that solvability of the equation $f(P) = Q$ and the linear-quadratic optimization problem are well-connected.

In particular, the equation $f(P) = Q$ has a solution that depends continuously on Q. To prove that there exists a solution to the algebraic Riccati equation (2.4), we need to extend this result to the boundary. Let $\tilde{C}$ be the $d \times d$ matrix obtained from C by amending zero rows as needed, and define the square matrix $C_\epsilon := \tilde{C} + \epsilon I$, where $\epsilon > 0$. If we exchange C for C_ϵ, we do not change the dynamical system (2.6), and, since C_ϵ is full rank, we still have observability. Also, $Q_\epsilon := {C_\epsilon}^{\mathsf{T}}C_\epsilon \in N$. Next, let P_ϵ be the unique solution to $f(P) = Q_\epsilon$. Again applying the control $\tilde{u}$, driving the trajectory of (2.6) to zero at time $t = t_1$, (2.10) implies that there is a bound $\mu(x_0)$ for all $x_0 \in \mathbb{R}^n$ such that

$$ {x_0}^{\mathsf{T}}P_\epsilon x_0 \le \int_0^{t_1} (\|C\tilde{x}(t)\|^2 + \|\tilde{u}(t)\|^2)dt + \epsilon\int_0^{t_1} \tilde{x}^{\mathsf{T}}(\tilde{C}^{\mathsf{T}} + \tilde{C} + \epsilon I)\tilde{x}dt \le \mu(x_0) $$

for all $\epsilon \in [0, \delta]$. Precisely as above, this provides an upper bound for P_ϵ on $[0, \delta]$. Hence, as $\epsilon \to 0$, some subsequence of P_ϵ tends to a limit $P_0 \geq 0$, which clearly must satisfy (2.4). Since $\Gamma_\epsilon := A - BB^{\mathsf{T}} P_\epsilon$ has all its eigenvalues in the open right half-plane, the limit $\Gamma_0 := A - BB^{\mathsf{T}} P_0$ has all its eigenvalues in the closed left half-plane. However just as above, we can show that observability ensures that none of these eigenvalues can lie on the imaginary axis. Consequently (2.10) holds for $P = P_0$ and $u = -B^{\mathsf{T}} P_0 x$ so that

$$x_0{}^{\mathsf{T}} P_0 x_0 = \int_0^\infty (\|y(t)\|^2 + \|u(t)\|^2) dt.$$

Therefore, if there were an $x_0 \neq 0$ such that $x_0 P_0 x_0 = 0$, this would contradict observability. Hence, P_0 must be positive definite.

Consequently, there is always a unique positive definite solution P of the Riccati equation (2.4), which depends continuously on C.

As these examples suggest, there is a broad class of problems for which the number of interior points for J_y is at most one. We conclude this section with specializations of the duality theorem to the case where the variational problems are convex.

Corollary 2.5. *Consider a well-connected pair of problems for which $\mathcal{Q}_y$ is a convex subset of a topological vector space and for which J_y is strictly convex for all $y \in N$. Then the problem* (1.2) *is well-posed if and only if f is proper. In this case, J_y has a (unique) minimizing point which, for each $y \in N$, is an interior point.*

Corollary 2.6. *Let M be a connected, open subset of $\mathbb{R}^n$, and let $\varphi : M \to \mathbb{R}$ be a C^1 function, which is proper and bounded from below. Then, if the derivative $\varphi' : M \to (\mathbb{R}^n)^*$ is locally injective and proper, then φ has a unique minimum on M, and φ' is a homeomorphism between M and $(\mathbb{R}^n)^*$.*

Remark 2.7. This formulation and conclusion remain valid, mutatis mutandis, whenever M is a parallelizable n-manifold.

Corollary 2.8. *Let M be an open convex subset of $\mathbb{R}^n$, and let $\varphi : M \to \mathbb{R}$ be a C^1 convex function. Consider the problem to minimize*

$$\psi(x) = \varphi(x) - \langle b, x \rangle.$$

If the derivative $\varphi' : M \to (\mathbb{R}^n)^*$ is locally injective and proper, then, for each $b \in (\mathbb{R}^n)^*$, ψ achieves a unique minimum on M and φ' is a homeomorphism between M and $(\mathbb{R}^n)^*$.

3. A geometric approach to constrained problems

Let L, M and N be manifolds of dimensions ℓ, $(n + \ell)$ and n, respectively, with M connected and N an open subset of $(\mathbb{R}^n)^*$, and consider two functions $g : M \to L$ and $f : M \to N$. For $a \in L$ and $b \in N$, we seek solutions to the equation

$$f^a(x) = b, \tag{3.1}$$

where $f^a := f|_{M_a}$ is the restriction of f to $M_a = g^{-1}(a)$. In this setting, we are interested in the well-posedness of solving the equation (3.1) subject to the constraint $g(x) = a$. There is of course also a dual problem which involves $M_b = f^{-1}(b)$ and the map $g^b := g|_{M_b}$ obtained by restricting g to M_b.

In this way, one may think of the constrained well-posedness problem as defining two (possibly singular) foliations on M. More precisely, we assume that g is a submersion and that each M_a is connected. Then, $\{M_a \mid a \in L\}$ are the leaves of a foliation $\mathcal{F}_1$ of M. We now consider the question of when f^a is well-connected to a variational problem, constrained to the leaf M_a. To this end, let d_1 denote the tangential derivative of $\mathcal{F}_1$; i.e., the exterior derivative relative to each leaf M_a. If there were a variational functional J_a defined on each M_a, then well-connectedness would arise if the level sets of $d_1 J_a$ defined a (possibly singular) foliations on M whose leaves were the level sets of f. There are several ways this can occur, leading to a pair of theorems, each of which we illustrate by example.

First, assume that M_a is diffeomorphic to $\mathbb{R}^n$ and that $(x_1, x_2, \cdots, x_n)$ are global coordinates for M_a. Any tangent vector $v \in T_x(M_a)$ to M_a can be represented, using these coordinates, as an n-vector $v \in \mathbb{R}^n$. In this sense, f^a defines a 1-form ω_0. More generally, for each $x \in M$ and each $b \in N$, $f^a - b$ is a linear functional on $T_x M_a$; i.e., an element of the cotangent space $T_x^* M_a$. As a functional which depends on x, $f^a - b$ is a 1-form on M_a. In the chosen coordinates, this 1-form ω_b is given by

$$\omega_b = \sum_{i=0}^{n} (f_i^a - b_i) dx_i,$$

where f_i^a and b_i represent the i-th component of f^a and b, respectively. We note that

$$\operatorname{graph}(\omega_b) = \operatorname{graph}(f^a - b) \subset T^* M_a \simeq M_a \times (\mathbb{R}^n)^*$$

where $T^* M_a$ denotes the cotangent bundle $\bigcup_{x \in M_a} T_x^* M_a$ of M_a The same construction can, of course, be carried out if M_a were just parallelizable.

Theorem 3.1. *Let f^a be proper for all $a \in L$, and suppose that M and each M_a are connected and N is Euclidean. Then, the following sets of statements are equivalent:*

(i) *For all $a \in L$, M_a is diffeomorphic to $\mathbb{R}^n$.*

(ii) *There exists a functional $J_b : M_a \to \mathbb{R}$ whose tangential exterior derivative $d_1 J_b$ is given, as a map $d_1 J_b : M_a \to (\mathbb{R}^n)^*$, by*

$$d_1 J_b = f^a(x) - b.$$

(iii) *The functional J_b has only nondegenerate critical points.*

and

(i)′ *The connected components of the inverse images $\{f^{-1}(b) \mid b \in N\}$ form the leaves of a foliation $\mathcal{F}_2$ of M, which is transverse to $\mathcal{F}_1$.*

(ii)′ *For any $a \in L$, ω_0 is d_1-closed, i.e., $d_1 \omega_0 = 0$.*

(iii)′ *For some $x \in M_a$, $\det \operatorname{Jac} f^a(x) \neq 0$.*

Moreover, if either set of conditions holds, f^a is well-posed and J_b has a unique critical point.

Proof. Suppose that M_a is diffeomorphc to $\mathbb{R}^n$. Choosing global coordinates we compute

$$d_1 J_b = \sum_{i=1}^{n} (f_i^a - b_i) dx_i.$$

At each solution x of equation (3.1),

$$\det D^2 J_b(x) = \det \operatorname{Jac}(f^a)(x)$$

is nonzero. Since f^a is proper, f^a is a diffeomorphism by Hadamard's Theorem. In particular, $f : M \to N$ is a submersion, and hence $\mathfrak{F}_2$ is a foliation.

Choose $x \in M$, and suppose that any v in the tangent space $T_x(M)$ of M at x satisfies

$$v \in T_x(M_a) \quad \text{and} \quad v \in T_x(f^{-1}(f^a(x))),$$

that is, suppose v is tangent to a leaf of both foliations. In this case, $\mathrm{Jac}(f^a)(x)v = 0$ so that

$$D^2 J_b(x)v = 0, \quad \text{where } b = f^a(x),$$

contradicting nondegeneracy of the critical points of J_b on M_a. In particular, the foliations $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are transverse.

Since $\omega_b = d_1 J_b$, taking $b = 0$ we obtain

$$d_1\omega_0 = d_1^2 J_0 = 0$$

so that ω_0 is d_1-closed.

Finally, since f^a is a diffeomorphism solving $f^a(x) = b$ is well-posed and since f^a is injective, for each b there is a unique critical point for J_b, which is nondegenerate by hypothesis.

We shall now prove the converse. To say that $\mathfrak{F}_2$ is a foliation transverse to $\mathfrak{F}_1$ is to say that if $v \in T_x(M_a)$ satisfies

$$\mathrm{Jac}(f^a)(x)v = 0$$

then v must be zero. Therefore, since M_a is connected and f^a is proper, $f^a : M_a \to N$ is a diffeomorphism by Hadamard's Theorem, and hence M_a is diffeomorphic to $\mathbb{R}^n$. We also conclude that J_b can only have one critical point, which is nondegenerate.

Choosing global coordinates $(x_1, \cdots, x_n)$ on M_a, suppose that

$$\omega_0 = \sum_{i=1}^{n} f_i^a dx_i$$

is d_1-closed. In this case,

$$\omega_b = \sum_{i=0}^{n} (f_i^a - b_i) dx_i$$

is also d_1-closed. Then, since M_a is Euclidean, the Poincaré Lemma implies that there exists a functional $J_b : M_a \to \mathbb{R}$ such that $dJ_b = \omega_b$, or $d_1 J_b = f - b$. Indeed, since path integrals of ω_b are dependent only on the end points of the path, choosing any point $x_0 \in M_a$,

$$J_b(x) = \int_{x_0}^{x} \omega_b = \int_{x_0}^{x} \sum_{i=1}^{n} \left(f_i^a(x) - b_i\right) dx_i$$

is well-defined and smooth, satisfying $d_1 J_b = f^a - b$.

Finally, since f^a is a diffeomorphism and $d_1 J_b = \omega_b$, $\mathrm{Jac}\, f^a(x)$ is a nonsingular, symmetric matrix-valued function on M satisfying

$$D^2 J_b(x) = \mathrm{Jac}(f^a)(x)$$

□

Corollary 3.2. *With the same notation and hypotheses as in Theorem 3.1, suppose that either of the sets of three conditions hold. If g^b is proper, then each leaf of $\mathcal{F}_1$ intersects each leaf of $\mathcal{F}_2$ in exactly one point, and the map*

$$(f,g):\, M \to N \times L$$

is a diffeomorphism.

Corollary 3.3. *With the same notation and hypotheses as in Theorem 3.1 but with hypothesis* (iii) *specialized to*

(iiia) *The functional J_b has nondegenerate minima at its critical points,*

and (iii)$'$ *specialized to*

(iiia)$'$ *For some $x \in M_a$,* $\operatorname{Jac} f^a(x) > 0$.

the conclusions hold, but with J_b having a unique minimum. In particular, this holds if each M_a is convex and each J_b is strictly convex.

Example 3.4. *The generalized moment problem with complexity constraint.* There is a vast literature on the generalized moment problem (see, e.g., [1, 2, 28, 38]), in part because so many problems and theorems in pure and applied mathematics, physics and engineering can be formulated as moment problems. Recall, that we are given a sequence of complex numbers $c_0, c_1, \cdots, c_n$ and a sequence of continuous, linearly independent complex-valued functions $\alpha_0, \alpha_1, \cdots, \alpha_n$ defined on the real interval $[a, b]$. In order for the moment equations (1.7) to hold it is necessary that c_k be real whenever α_k is real, with a similar statement holding for the case that α_k is purely imaginary. Indeed, a purely imaginary moment condition can always be reduced to a real one, and henceforth we shall assume that this is the case. In fact, we assume that $\alpha_0, \dots, \alpha_{r-1}$ are real functions and $\alpha_r, \dots, \alpha_n$ are complex-valued functions whose real and imaginary parts, taking together with $\alpha_0, \dots, \alpha_{r-1}$ are linearly independent over $\mathbb{R}$. In analogy with the real case discussed in the introduction, we introduce the real vector space $\mathfrak{P}$ which is the sum of the real span of $\alpha_0, \dots, \alpha_{r-1}$ and the complex span of $\alpha_r, \dots, \alpha_n$. In particular, the real dimension of $\mathfrak{P}$ is $2n - r + 2$.

The moment problem is then to find all monotone, nondecreasing functions μ of bounded variation such that (1.7) is satisfied whenever the sequence $c_0, c_1, \cdots, c_n$ is *positive* in the sense that

$$\langle c, q\rangle := \operatorname{Re} \sum_{k=0}^{n} q_k c_k > 0 \tag{3.2}$$

for all $q := (q_0, q_1, \cdots, q_n) \in \mathbb{R}^r \times \mathbb{C}^{n-r+1}$ such that

$$\sum_{k=0}^{n} q_k \alpha_k \in \overline{\mathfrak{P}}_+ \smallsetminus \{0\}, \tag{3.3}$$

where $\mathfrak{P}_+$ consists of all functions in $\mathfrak{P}$ that have positive real part, and $\overline{\mathfrak{P}}_+$ is its closure. Denote by $\mathfrak{C}_+$ the space of positive sequences.

We shall assume that $\mathfrak{P}_+$ is nonempty, so that it is also open and convex, and therefore diffeomorphic to $\mathbb{R}^{2n-r+2}$. Then $\mathfrak{C}_+$ is also nonempty, and, since it is an open convex subset of $\mathbb{R}^{2n-r+2}$, it is diffeomorphic to a Euclidean space of the same dimension as $\mathfrak{P}_+$. Moreover, we introduce the complexity constraint

$$\frac{d\mu}{dt} = \Phi(t) = \frac{P(t)}{Q(t)}, \tag{3.4}$$

where however we now ask that $P = \mathrm{Re}\{\tilde{P}\}$ and $Q = \mathrm{Re}\{\tilde{Q}\}$ where $\tilde{P}, \tilde{Q} \in \mathfrak{P}_+$. Since the real and imaginary parts of $\alpha_r, \dots, \alpha_n$, taken together with $\alpha_0, \dots, \alpha_{r-1}$, are linearly independent over $\mathbb{R}$, any $\tilde{P} \in \mathfrak{P}_+$ is uniquely determined by its real part $P := \mathrm{Re}\{\tilde{P}\}$. Each choice of P defines the leaf $\{P\} \times \mathfrak{P}_+$ of a foliation $\mathfrak{F}_1$ of $M := \mathfrak{P}_+ \times \mathfrak{P}_+$ having Euclidean leaves. On such a leaf we consider the function

$$F^{\tilde{P}} : \mathfrak{P}_+ \to \mathfrak{C}_+$$

componentwise via

$$F_k^{\tilde{P}}(\tilde{Q}) = \int_a^b \alpha_k(t) d\mu(t).$$

Parameterizing $\tilde{Q}$ via $\tilde{Q} = \sum_{k=0}^n q_k \alpha_k$, we construct the 1-form

$$\omega_c = \mathrm{Re}\left\{ \sum_{k=0}^n \left[c_k - F_k(\tilde{Q}) \right] dq_k \right\},$$

on $\mathfrak{P}_+$. Explicitly, we have

$$\begin{aligned} \omega_c &= \mathrm{Re}\left\{ \sum_{k=0}^n c_k dq_k - \int_a^b \sum_{k=0}^n \alpha_k \frac{P}{Q} dq_k dt \right\} \\ &= \mathrm{Re} \sum_{k=0}^n c_k dq_k - \int_a^b \frac{P}{Q} dQ dt \end{aligned}$$

so taking the exterior derivative (on $\mathfrak{P}_+$) we obtain

$$d\omega_c = \int_a^b \frac{P}{Q^2} dQ \wedge dQ dt = 0,$$

establishing that the 1-form ω_c is closed.

Therefore, by the Poincaré Lemma, there exist a smooth function J_c such that

$$J_c = \int \omega_c = \int \left(\mathrm{Re} \sum_{k=0}^n c_k dq_k - \int_a^b \frac{P}{Q} dQ dt \right),$$

with the integral being independent of the path between two endpoints. Computing the path integral along the lines in Section 1, one finds that

$$J_c(Q) = \langle c, q \rangle - \int_a^b P \log Q \, dt, \tag{3.5}$$

which is strictly convex and bounded from below for positive sequences $c_0, c_1, \cdots, c_n$. As in the real case, J_c has an interior critical point precisely at the solution of the (complex) generalized moment problem. To see this on the second factor of $\mathbb{R}^r \times \mathbb{C}^{n-r+1}$, we decompose the exterior differential as the sum $d = \partial + \bar{\partial}$, where $\bar{\partial}$ is the Cauchy-Riemann differential. Since J_c is real, to say that $dJ_c = 0$ is to say that $\partial J_c = 0$ or, equivalently, that $\bar{\partial} J_c = 0$. Finally, by inspection we see that $\partial J_c = 0$ is the set of defining equations of the generalized moment problem.

In order to apply Theorem 3.1, we need now only show that $F^P : \mathfrak{P}_+ \to \mathfrak{C}_+$ is proper.

Lemma 3.5. *Suppose $\mathfrak{P}$ is a vector space consisting of C^2-smooth functions. Then $F^P : \mathfrak{P}_+ \to \mathfrak{C}_+$ is proper.*

Proof. First note that

$$\operatorname{Re} \sum_{k=0}^{n} f_k F_k(\tilde{Q}) = \int_a^b \operatorname{Re}\left\{ \sum_{k=0}^{n} f_k \alpha_k \right\} \frac{P}{Q} dt > 0 \tag{3.6}$$

whenever $\sum_{k=0}^{n} f_k \alpha_k \in \mathfrak{P}_+$. Given any compact set K in $\mathfrak{C}_+$, $(F^P)^{-1}(K)$ is bounded. In fact, if $\|Q\|$ tends to infinity, then $F^P(Q)$ tends to zero, which belongs to the boundary of $\mathfrak{C}_+$ but not to $\mathfrak{C}_+$ itself. Thus the preimage of a convergent sequence in K has a cluster point in the closure of $\mathfrak{P}_+$. Such a cluster point, let us call it $\tilde{Q}$, cannot lie on the boundary of $\mathfrak{P}_+$, for then Q is a smooth, nonnegative function having a zero on the interval so that (3.6), and hence the integral defining F^P, is divergent. Hence, $\tilde{Q} \in \mathfrak{P}_+$, establishing that F^P is proper. □

Corollary 3.6. *Let $\mathfrak{P}$ be spanned by C^2 functions $\alpha_0, \alpha_1, \dots, \alpha_n$, and let $c \in \mathfrak{C}_+$. Suppose also that the nonzero real and imaginary parts of $\alpha_0, \alpha_1, \dots, \alpha_n$ form a linearly indepentent set over $\mathbb{R}$. Then, for any choice of $\tilde{P} \in \mathfrak{P}_+$ there is one and only one $\tilde{Q} \in \mathfrak{P}_+$ solving the generalized moment problem*

$$\int_a^b \alpha_k(t) \frac{P(t)}{Q(t)} dt = c_k, \quad k = 0, 1, \dots, n,$$

where $P = \operatorname{Re}\{\tilde{P}\}$ and $Q = \operatorname{Re}\{\tilde{Q}\}$. Moreover, the generalized moment problem determines, for each choice of $c = (c_0, c_1, \dots, c_n)$, a leaf of a foliation $\mathcal{F}_2$ of $\mathfrak{P}_+ \times \mathfrak{P}_+$. The foliation $\mathcal{F}_2$ is transverse to the foliation $\mathcal{F}_1$, whose leaves are defined by fixing a choice of P. Each leaf of $\mathcal{F}_1$ meets each leaf of $\mathcal{F}_2$ in exactly one point. In particular, the generalized moment problem is well-posed with $\tilde{Q}$ depending smoothly on c and $\tilde{P}$. Indeed, c and P determine Q as the unique minimum in $\mathfrak{P}_+$ of the strictly convex functional

$$J_c P(q) = \langle c, q \rangle - \int_a^b P \log Q dt.$$

Remark 3.7. We have shown that the solutions of the generalized moment problem are completely parameterized by P. It is therefore natural to ask whether there is also a complete parameterization in terms of Q. In fact, all the steps, save one, can be reproduced for this setting, and one can even derive a functional J_c whose minimum would be the unique solution if a solution would exist. However, what fails is the proof that $F^{\tilde{Q}}$ is proper. Consequently, for an arbitrary fixed Q, there may be no solution.

In Theorem 3.1, we assumed that J_b had a very classical form. Of course, one can have a pair of well-connected problems for more general situations, e.g., when the solution set of $f_a(x) = b$ is a branch of $J_b(x) = 0$. This indeed occurs in example that illustrates the next theorem. In this theorem, well-posedness and uniqueness of minima again have interesting implications about the geometry of the corresponding foliations.

Theorem 3.8. *Suppose the manifolds M, N, L and each leaf M_a are connected and that, for each $a \in L$ and $b \in N$, there exists a functional J_b defined on $\mathcal{Q}_b$ such that the problems of solving $f^a(x) = b$ and minimizing J_b are well-connected. Suppose g is a submersion so that each M_a is a leaf of a foliation $\mathcal{F}_1$ of M. Suppose further that f^a is proper and that each J_b has at most one minimum. Then*

(i) $f^a : M_a \to N$ *is a homeomorphism.*
(ii) J_b *achieves its minimum at an interior point.*

Furthermore, if f^a is a submersion and $g^b : M_b \to L$ is proper, then

(iii) g^b *is a homeomorphism.*
(iv) $f^{-1}(b)$ *is the leaf of a foliation $\mathcal{F}_2$ of M.*
(v) *Each leaf of $\mathcal{F}_1$ intersects each leaf of $\mathcal{F}_2$ in exactly one point.*
(vi) $(f, g) : M \to N \times L$ *is a homeomorphism.*

Proof. Conditions (i) and (ii) follow from Theorem 2.1. We now show that g^b is an injection. To say that $g^b(x_1) = g^b(x_2)$ is to say that $x_1, x_2 \in M_a$, when $g^b(x_1) = a$. Since $x_1, x_2 \in f^{-1}(b)$, we have $f^a(x_1) = f^a(x_2)$ so that $x_1 = x_2$. The map g^b is a proper injection, and therefore, by Lemma 2.3, g^b is a homeomorphism and $f^{-1}(b)$ is connected, proving (iii) and (iv). Given any $a \in L$ and $b \in N$, there exists one and only one $x \in M_a$ such that $f^a(x) = b$, which proves (v). Finally, (f, g) is a continuous bijection between M and $N \times L$. By Brouwer's Theorem, (f, g) is an open map and therefore has a continuous inverse. □

Example 3.9. (*Interior-point methods for linear programming problems.*) Recall that linear programming problems are often given in the standard form

$$\min c^{\mathsf{T}} x \quad \text{subject to } Ax = b,\ x \geq 0, \tag{3.7}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$ are given and $x \in \mathbb{R}^n$ is the variable. For simplicity, we assume that the matrix A is fixed and has full rank. The corresponding dual problem is

$$\max b^{\mathsf{T}} u \quad \text{subject to } A^{\mathsf{T}} u \leq c,$$

where $u \in \mathbb{R}^m$ is the variable. This can also be written

$$\max b^{\mathsf{T}} u \quad \text{subject to } A^{\mathsf{T}} u + s = c,\ s \geq 0, \tag{3.8}$$

by introducing slack variables $s_1, s_2, \ldots, s_n$. From now on, we shall refer to (3.7) as the *primal* problem and to (3.8) as the *dual* problem.

It is well-known that the primal problem has an optimal solution if and only if the dual one does, and this happens if and only if the the *primal-dual feasibility set*

$$\mathcal{F} = \{(x, u, s) \mid Ax = b,\ A^{\mathsf{T}} u + s = c,\ x \geq 0,\ s \geq 0\}$$

is nonempty. Then $(\hat{x}, \hat{u}, \hat{s}) \in \mathcal{F}$ is optimal for the dual and the primal problems if and only if the *complementary-slackness condition*

$$\hat{x}_k \hat{s}_k = 0 \quad k = 1, 2, \ldots, n \tag{3.9}$$

is fulfilled. Let $M = \{(x, u, s) \in \mathbb{R}^{2n+m} \mid x > 0,\ s > 0\}$.

Introducing the notation $a = \begin{bmatrix} b \\ c \end{bmatrix}$, we also consider the strictly feasible set

$$M_a = \{(x, u, s) \in \mathcal{F} \mid x > 0,\ s > 0\}.$$

Then $M_a = g^{-1}(a)$ where $g:\ M \to L := \mathbb{R}^{n+m}$ is defined via

$$\begin{bmatrix} Ax \\ A^{\mathsf{T}}u + s \end{bmatrix}.$$

On M_a the variable u is uniquely determiner by s, and therefore we parameterize a point in M_a by the coordinate pair (x, s). A point in M_a is called an *interior point.* The tangent space to M_a at an interior point (x, s) is given by

$$T_{(x,s)}(M_a) = \{(h, k) \in \mathbb{R}^n \times \mathbb{R}^n \mid h \in \ker A,\ k \in (\ker A)^{\perp}\}.$$

In fact, since $\mathrm{Im}A^{\mathsf{T}} = (\ker A)^{\perp}$, $x_1 - x_2 \in \ker A$ and $s_1 - s_2 \in (\ker A)^{\perp}$ for all (x_1, s_1) and (x_2, s_2) in M_a. In particular,

$$(x_1 - x_2)^{\mathsf{T}}(s_1 - s_2) = 0. \tag{3.10}$$

Since $\dim T_{(x,s)}(M_a) = n$, M_a is an n-manifold.

The basic idea in *interior-point methods* is to construct a parameterized set of interior points satisfying

$$x_k s_k = \tau_k, \quad k = 1, 2, \dots, n, \tag{3.11}$$

where the parameters $\tau_1, \tau_2, \dots, \tau_n$ are positive real numbers. This set is called the *central path*, and the idea is to construct a sequence of points in this set converging to an optimal solution as $\tau := (\tau_1, \tau_2, \dots, \tau_n)^{\mathsf{T}} \to 0$. For details, see, e.g.,[50]), where, however, τ is chosen so that $\tau_1 = \tau_2 = \cdots = \tau_n$. Here, we prefer the more general parameterization.

Now, define $\mathbb{R}^n_+$ to be the n-manifold of vectors $\tau := (\tau_1, \tau_2, \dots, \tau_n)^{\mathsf{T}}$ such that $\tau_k > 0, k = 1, 2, \dots, n$. Next, consider the smooth function $f:\ M_a \to N := \mathbb{R}^n_+$ given by

$$f^a(x, s) = \begin{bmatrix} x_1 s_1 \\ x_2 s_2 \\ \vdots \\ x_n s_n \end{bmatrix}.$$

If M_a is nonempty, f^a is proper. To see this, let $\bar{f}^a : M_a \to \mathbb{R}^n$ be the continuous extension of f^a to the boundary ∂M_a. Then

$$\bar{f}^a(\partial M_a) \subset \partial\mathbb{R}^n_+,$$

and hence $f^{-1}(K) \subset M_a$ for any compact set $K \in \mathbb{R}^n_+$. Now, $f^{-1}(K)$ is closed by continuity, so it just remains to prove that $f^{-1}(K)$ is bounded. Let $(\tilde{x}, \tilde{s}) \in M_a$, and let $(x, s) \in f^{-1}(K)$ be arbitrary. In view of (3.10), $(x - \tilde{x})^{\mathsf{T}}(s - \tilde{s}) = 0$, and hence

$$x^{\mathsf{T}}s + \tilde{x}^{\mathsf{T}}\tilde{s} = \tilde{x}^{\mathsf{T}}s + \tilde{s}^{\mathsf{T}}x \geq \epsilon e^{\mathsf{T}}(x + s),$$

where $e := (1, 1, \dots, 1)^{\mathsf{T}}$ and where

$$\epsilon = \min\{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n, \tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_n\} > 0.$$

Consequently,

$$0 < e^{\mathsf{T}}[x + s] \leq \frac{1}{\epsilon} e^{\mathsf{T}}[f(x, s) + f(\tilde{x}, \tilde{s})],$$

which is bounded since K is. Hence, $f^{-1}(K)$ is bounded, as claimed.

We shall now construct a variational problem which is well-connected to f. To this end, for each $\tau \in \mathbb{R}^n_+$, let $J_\tau : M_a \to \mathbb{R} \cup \{\infty\}$ be the function

$$J_\tau(x, s) = x^\mathsf{T} s - \sum_{k=1}^{n} \tau_k \log(x_k s_k), \tag{3.12}$$

and consider the problem of solving the problem

$$\inf_{M_a} J_\tau(x, s). \tag{3.13}$$

The function J_τ is strictly convex for each $\tau \in \mathbb{R}^n_+$. In fact, since the functions $-\log x_k s_k = -\log x_k - \log s_k$ are strictly convex, convexity of J_τ follows from the fact that the function $g(x, s) = x^\mathsf{T} s$ is linear on M_a. To see this, note that

$$[x_2-\lambda(x_1-x_2)]^\mathsf{T}[s_2-\lambda(s_1-s_2)] = {x_2}^\mathsf{T} s_2+(x_1-x_2)^\mathsf{T}(s_1-s_2)-\lambda {x_2}^\mathsf{T}(s_1-s_2)-\lambda(x_1-x_2)^\mathsf{T} s_2$$

for $(x_i, s_i) \in M_a$, $i = 1, 2$, which is the same as

$$g(\lambda x_1 + (1-\lambda)x_2, \lambda s_1 + (1-\lambda)s_2) = \lambda g(x_1, s_1) + (1-\lambda)g(x_2, s_2).$$

In fact, in view of (3.10), ${x_2}^\mathsf{T}(s_1 - s_2) = {x_1}^\mathsf{T}(s_1 - s_2)$.

Next, note that

$$\begin{aligned} \frac{\partial J_\tau}{\partial x_k} &= s_k - \frac{\tau_k}{x_k}, \qquad & k = 1, 2, \dots, n, \\ \frac{\partial J_\tau}{\partial s_k} &= x_k - \frac{\tau_k}{s_k}, \qquad & k = 1, 2, \dots, n, \end{aligned}$$

and hence the directional derivative of J_τ in $(x, s) \in M_a$ in the direction (h, k) is given by

$$d_1 J_\tau(x, s; h, k) = \sum_{k=1}^{n} (x_k s_k - \tau_k) \left[\frac{h_k}{x_k} + \frac{k_k}{s_k}\right].$$

Consequently, $(\hat{x}, \hat{s}) \in M_a$ is a critical point for the optimization problem (3.13) whenever $(\hat{x}, \hat{s})$ is a solution of

$$f^a(x, s) = \tau \tag{3.14}$$

so that (3.14) is a branch of $d_1 J_\tau = 0$. In particular, the family of optimization problems (3.13) and the family of equations (3.14) form a well-connected pair. Since J_τ is strictly convex, it has at most one critical point, which is a minimum.

By Theorem 2.1, J_τ has an interior minimizing point, and the map f is a homeomorphism between M_a and $\mathbb{R}^n_+$. In particular, as $\|\tau\| \to 0$, (x_τ, s_τ) tends to a limit $(\hat{x}, \hat{s}) \in \overline{M_a}$, which satisfies the complementary-slackness condition (3.9) and hence is optimal for the primal and dual linear programming problems (3.7) and (3.8).

Remark 3.10. From Theorem 3.8 we know that f^a is a homeomorphism so that the problem of solving (3.14) is well-posed. We also know that the unique solution of (3.14) is the unique minimizing interior point for the strictly convex functional (3.12). In fact, one can actually prove smooth dependence of solutions on the data in this particular case. $\mathbb{R}^n$ with $(\mathbb{R}^n)^*$ using the standard inner product, consider the subsets of $M_a \times N$ defined by the following two conditions:

$$d_1 J(x, s) = 0 \tag{3.15}$$

$$\tau - f^a(x, s) = 0 \tag{3.16}$$

The second set, being the graph of a smooth function is a smooth submanifold of dimension n. Since $\frac{\partial}{\partial\tau}d_1 J_\tau(x,s)$ is full rank, the first set is also a submanifold of dimension n, which is everywhere locally the graph of a smooth function of (x,s). One such branch is given by (3.16), and therefore graph(f^a) is a connected component of the submanifold (3.15). Now, consider the projection $p_2(x,s,\tau)=\tau$ restricted to this connected component of (3.15). The Jacobian of this map, computed with respect to (x,s) at the point (x,s,τ) is $D^2 J_\tau(x,s)$, which is positive definite by strict convexity of J_τ. Therefore, the Jacobian Jac(f^a) is everywhere nonsingular on M_a. Since f^a is proper, f^a is a diffeomorphism by Hadamard's Theorem.

Theorem 3.11. *For the interior-point method for linear programming, the problem of solving $f^a(x^i,s^i)=\tau_i$ on the (generalized) central path $\tau=(\tau_1,\tau_2,\dots,\tau_n)^\mathsf{T}\to 0$ as $i\to\infty$ is well-posed. In particular, the solution sequence (x^i,s^i) exists, is unique, and lies in a compact set. Moreover, the solution (x,s) to* (3.14) *is a smooth function of τ. Indeed, (x,s) is the unique minimizing interior point for the strictly convex functional*

$$J_\tau(x,s)=x^\mathsf{T}s-\sum_{k=1}^n \tau_k \log(x_k s_k).$$

4. Some interpolation problems occurring in signal processing and systems theory

Interpolation problems for meromorphic functions are an important class of moment problems, with a history going back to Carathéodory, Schur, Toeplitz, Nevanlinna and Pick. One class of meromorphic functions studied in this context are those functions of a complex variable which are analytic in the open unit disc and map points there into the open left half-plane. In the mathematics literature these are typically referred as *Carathéodory functions*, while in the engineering literature they are often called *positive real functions.* Indeed, given that the impedance of a circuit with finitely many active components is a positive real rational function, it is not surprising that important problems involving interpolation by positive real functions at points in the finite complex plane emerged in circuit theory. They also abound in robust stabilization and control and in signal processing; see [13, 14, 19, 33, 34, 36, 51].

As an example, the mathematical basis for the design of the Linear Predictive Coding (LPC) filters used for speech synthesis by most existing cellular telephones involves the interpolation problem of finding a positive-real rational function (filter) which matches a given window of Laurent (or covariance) coefficients. More explicitly, in signal processing it is common to model a signal $\{y(t)\mid t\in\mathbb{Z}\}$ as a convolution

$$y(t)=\sum_{k=-\infty}^{t} w_{t-k}u_k$$

of some excitation signal $\{u(t)\mid t\in\mathbb{Z}\}$. In the language of systems and control, this amounts to passing the excitation signal u through a linear filter with the transfer function

$$w(z)=\sum_{k=0}^{\infty} w_k z^{-k}$$

which is assumed to be rational, thus obtaining the signal y as the output. More specifically, we take $w(z)$ to be rational with $w_0 \neq 0$ and all zeros and poles in the open unit disc. In other words, the function w is outer with respect to the complement of the unit disc. Such a filter is often called a *shaping filter* when it is used to "shape" a white noise input into a stationary stochastic process.

Indeed, consider a signal y for which the excitation signal u is white noise, i.e., $\mathrm{E}\{u(t)u(s)\} = \delta_{ts}$, where δ_{ts} is one if $t = s$ and zero otherwise. Then y is a stationary stochastic process with a rational spectral density

$$\Phi(e^{i\theta}) = |w(e^{i\theta})|^2$$

that is positive for all θ. The Fourier coefficients

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\Phi(e^{i\theta})d\theta, \quad k = 0, 1, 2, \cdots$$

are then the covariance lags $c_k = \mathrm{E}\{y(t+k)y(t)\}$, and a relatively short sequence of these, $c_0, c_1, \cdots, c_n$, can be determined, via ergodic limits, from a record of observed data from the output process y.

As pointed out by Delsarte et al [20], to find a rational Φ matching this window of covariance lags is to find a rational solution of the classical trigonometric moment problem, since whenever

$$\Phi(e^{i\theta}) = \tilde{c}_0 + \sum_{k=1}^{\infty} \tilde{c}_k \cos(k\theta), \quad \tilde{c}_k = c_k, k = 0, 1, \dots, n$$

then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(k\theta)\Phi(e^{i\theta})d\theta = c_k, \quad k = 0, 1, \cdots, n.$$

If we write

$$\Phi(e^{i\theta}) = 2\mathrm{Re}\{F(e^{i\theta})\}, \tag{4.1}$$

where

$$F(z) = \frac{1}{2}c_0 + c_1 z + c_2 z^2 + \cdots, \tag{4.2}$$

then this problem is also an interpolation problem for the value and the first n derivatives at $z = 0$ of the positive real function F, with the complexity constraint that it be rational of degree at most n.

According to Corollary 3.6, given any positive trigonometric polynomial P of degree at most n and any positive sequence $(c_0, c_1, \dots, c_n)$, there is a unique positive trigonometric polynomial Q of degree at most n, given as the minimum of the strictly convex functional

$$J_c(q) = \langle c, q\rangle - \int_{-\pi}^{\pi} P \log Q d\theta,$$

such that

$$\Phi = \frac{P}{Q}$$

solves the trigonometric moment problem.

Either by applying the Riesz-Fejér Theorem to P and Q and substituting $z = e^{i\theta}$ or by substituting $z = e^{i\theta}$ and applying spectral factorization, from Φ we obtain

$$\Phi(z) = w(z)w(z^{-1}),$$

where

$$w(z) = \frac{\sigma(z)}{a(z)} \tag{4.3}$$

with $\sigma(z)$ and $a(z)$ Schur polynomials of degree n. (A *Schur polynomial* is a polynomial with all its roots in the open unit disc.) In this way, the choice of spectral zeros (i.e., the zeros of $\sigma(z)$) parameterizes all shaping filters having a given window of covariance lags as there is one and only one $a(z)$ for each choice of zeros. In this context, the condition that the sequence $(c_0, c_1, \cdots, c_n)$ be positive is a positivity condition on the corresponding Toeplitz matrix:

$$T_n = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_n \\ c_1 & c_0 & c_1 & \cdots & c_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_n & c_0 & c_{n-1} & \cdots & c_0 \end{bmatrix} > 0.$$

In fact, for $Q(z) = a(z)a(z^{-1})$,

$$\langle c, q \rangle = a^{\mathsf{T}} T_n a,$$

where $a := (a_0, a_1, \cdots, a_n)^{\mathsf{T}}$ is the vector of coefficients of $a(z)$.

Classically, it is well-known that the choice $P = 1$, corresponding to $\sigma(z) = z^n$, always yields such a rational filter, known as the maximum entropy solution or the LPC filter. Indeed, this filter is characterized as the maximum of the entropy functional

$$E(\Phi) = \int_{-\pi}^{\pi} \log \Phi d\theta$$

subject to the covariance constraints. Since the parameters of the LPC filter can be computed very easily from the covariance data, LPC filters are in widespread commercial use. However, it is well-documented in the literature on speech processing (see [3, p. 1726], [39, pp. 271–272], [44, pp. 76–78,105]) that nontrivial choices of spectral zeros are necessary for high quality speech synthesis, an observation which led to many attempts to incorporate a choice of zeros into shaping filter design.

In 1983, using degree theory applied to spaces of rational positive real functions, Georgiou [24] showed that for every choice of $\sigma(z)$ in the open unit disc there exists an $a(z)$ having all roots in the open unit disc such that the positive real function (4.2) corresponding to the shaping filter (4.3). He also conjectured uniqueness which, of course, in this setting would itself imply existence. Interestingly, the space of positive real functions on which Georgiou applied a degree theoretic argument was a leaf of the foliation of the space of rational positive real functions, of degree at most n, defined by fixing the choice of $\sigma(z)$. As it turns out, this is also the foliation whose leaves are the stable manifolds of a variety of equilibria for a fast form of Kalman filtering, viewed as a dynamical system on this space of rational positive real functions [17]. In [9] it was shown that the interpolation conditions define a second foliation, transverse to the filtering foliation. The geometry of these foliations then allowed for a refinement of degree theoretic arguments which proved uniqueness, proving Georgiou's conjecture.

The computation of solutions by solving a strictly convex optimization problem was then developed in [11, 16], starting with a generalized maximum entropy integral as a primal problem and deriving J_c as the functional defining the dual problem. The direct derivation of J_c in terms of tangentially closed one-forms presented here is new. Further applications of these methods to problems of speech processing and signal processing are given in [15].

Generalizing the trigonometric moment problem described above to the situation that F interpolates at arbitrary points in the open unit disc, rather than just at $z = 0$, leads to Nevanlinna-Pick interpolation with a complexity constraint, which amounts to imposing a degree constraint on a rational interpolant. During the last two decades it has been discovered that this analytic interpolation problem is closely related to several robust control problems, for example, the gain-margin maximization problem [47, 48, 36], robust stabilization problem [37], sensitivity shaping in feedback control, simultaneous stabilization [27], robust regulation problem [18] and general H_∞ control problem [23]. In fact, by now such Nevanlinna-Pick interpolation problems appear even in textbooks on a first course in control [21].

Given $n + 1$ distinct points $z_0, z_1, \dots, z_n$ in the open unit disc, our interpolation problem consist in determining the rational Carathéodory functions F of degree at most n satisfying the interpolation condition

$$F(z_k) = w_k \quad \text{for } k = 0, 1, \dots, n, \tag{4.4}$$

where $w_0, w_1, \dots, w_n$ are prescribed values in the open right half of the complex plane. Without loss of generality, we assume for convenience that $z_0 = 0$ and w_0 is real. If the points $z_0, z_1, \dots, z_n$ are not distinct, the interpolation conditions are modified in the following way. If $z_k = z_{k+1} = \cdots = z_{k+m-1}$, the corresponding interpolation conditions are replaced by

$$\begin{aligned} F(z_k) &= w_k \\ F'(z_k) &= w_{k+1} \\ &\vdots \\ \frac{1}{(m-1)!}F^{(m-1)}(z_k) &= w_{k+m-1} \end{aligned}$$

This formulation of the Nevanlinna-Pick interpolation problem differs from the classical one in that a degree constraint on the interpolant F has been introduced, a restriction motivated by applications.

Next, we reformulate this interpolation problem as a generalized moment problem. From the Herglotz Theorem we see that the rational Carathéodory functions of degree at most n can be represented as

$$F(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z}\Phi(e^{i\theta})d\theta \tag{4.5}$$

for some rational spectral density Φ of degree at most $2n$, which is given by

$$\Phi(e^{i\theta}) = \mathrm{Re}\{F(e^{i\theta})\}. \tag{4.6}$$

Differentiating (4.5), we obtain

$$\frac{1}{\nu!}F^{(\nu)}(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{2e^{i\theta}}{(e^{i\theta}-z)^{\nu+1}}\Phi(e^{i\theta})d\theta, \quad \nu = 1,2,\dots.$$

Consequently, the interpolation conditions can be written

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\alpha_k(\theta)\Phi(e^{i\theta})d\theta = w_k, \quad k = 0,1,\dots,n,$$

where

$$\alpha_k(\theta) = \frac{e^{i\theta}+z_k}{e^{i\theta}-z_k}$$

for single interpolation points and

$$\begin{aligned}
\alpha_k(\theta) &= \frac{e^{i\theta}+z_k}{e^{i\theta}-z_k}\\
\alpha_{k+1}(\theta) &= \frac{2e^{i\theta}}{(e^{i\theta}-z_k)^2}\\
&\vdots\\
\alpha_{k+m-1}(\theta) &= \frac{2e^{i\theta}}{(e^{i\theta}-z_k)^m}
\end{aligned}$$

whenever $z_k = z_{k+1} = \cdots = z_{k+m-1}$. In particular, $\alpha_0 = 1$.

In order for this moment problem to have a solution, the sequence $w_0, w_1, \dots, w_n$ must be positive with respect to $\alpha_0, \alpha_1, \dots, \alpha_n$. For distinct interpolation points $z_0, z_1, \dots, z_n$, this is equivalent to the Pick matrix

$$P_n = \left[\frac{w_k + \bar{w}_\ell}{1 - z_k\bar{z}_\ell}\right]_{k,\ell=0}^{n} \tag{4.7}$$

being positive definite. To see this, first note that there is a factorization

$$\operatorname{Re}\sum_{k=0}^{n} q_k\alpha_k(\theta) = \left|\sum_{k=0}^{n} a_k\frac{\overline{\alpha_k(\theta)}+1}{2}\right|^2,$$

from which a straight-forward calculation shows that

$$\langle w, q\rangle = a^* P_n a, \tag{4.8}$$

where $a = (a_0, a_1, \dots, a_n)^*$.

Then, by Corollary 3.6, to each $P = \operatorname{Re}\{\tilde{P}\}$ with $\tilde{P} \in \mathfrak{P}_+$, there is a unique $Q = \operatorname{Re}\{\tilde{Q}\}$ with $\tilde{Q} \in \mathfrak{P}_+$ so that

$$\Phi(z) = \frac{P(z)}{Q(z)} \tag{4.9}$$

solves the corresponding moment problem, and this Q is given as the minimum of the strictly convex functional

$$J_w(Q) = \langle w, q\rangle - \int_{-\pi}^{\pi} P(e^{i\theta})\log Q(e^{i\theta})d\theta. \tag{4.10}$$

In the language of Hardy spaces of functions analytic in the unit disc $\mathbb{D}$, the space $\mathfrak{P} = \operatorname{span}\{\alpha_0, \alpha_1, \dots, \alpha_n\}$ can be identified with the coinvariant subspace

$$H(B) := H(\mathbb{D}) \ominus BH(\mathbb{D}),$$

where B is the Blaschke product

$$B(z) = z\prod_{k=1}^{n} \frac{z - z_k}{1 - \bar{z}_k z}.$$

In fact, the analytic functions $g_0, g_1, \dots, g_n$ defined via

$$g_k(e^{i\theta}) = \overline{\alpha(\theta)}, \quad k = 0, 1, \dots, n$$

span $H(B)$, and P and Q are invariant under conjugation. Moreover, $H(B)$ consists precisely of those functions that can be written as

$$g(z) = \frac{\pi(z)}{\tau(z)}, \quad \tau(z) = \prod_{k=1}^{n}(1 - \bar{z}_k z),$$

for some polynomial $\pi(z)$ of degree at most n, and hence any Carathéodory function F of degree at most n has a representation

$$F(z) = \frac{b(z)}{a(z)}, \tag{4.11}$$

where a and b are outer functions in $H(B)$. Since therefore

$$F + F^* = \frac{a^*b + ab^*}{aa^*},$$

it follows from (4.6) and (4.9) that

$$Q(z) = a(z)a^*(z) \tag{4.12}$$

$$P(z) = a^*(z)b(z) + a(z)b^*(z). \tag{4.13}$$

Hence, there is unique interpolant (4.11) corresponding to P which can be determined in the following way. Given the unique minimizer Q of (4.10), determine a as its outer spectral factor. Then b is the unique solution to the linear equation (4.13), since the kernel of the linear map $S(a)$ defined by $S(a)v = a^*v + av^*$ is zero. In fact, if $\zeta_1, \zeta_1, \dots, \zeta_n$ are the zeros of the outer function a, $a^*(\zeta_k) \neq 0$ for $k = 1, 2, \dots, n$, and therefore $S(a)v(\zeta_k) = a^*(\zeta_k)v(\zeta_k) = 0$ for $k = 1, 2, \dots, n$ implies that $v = 0$.

Determining the Q that minimizes (4.10) is particularly simple when $P = 1$, as this problem can then be reduced to solving a system of linear equations. The corresponding interpolant F is called the *central solution*. In fact, replacing $Q(e^{i\theta})$ in (4.10) by $|a(e^{i\theta})|^2$, where $a(z)$ is outer, and observing that, for such an $a(z)$, $\log|a(e^{i\theta})|$ is harmonic in the unit disc, we see that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log Q(e^{i\theta})d\theta = 2\log|a(0)|,$$

and consequently (4.10) can be replaced by

$$\hat{J}_w(a) = a^*P_n a - 2\log|a(0)|,$$

where the positive definite quadric form $a^*P_n a$ is formed as in (4.8), P_n being the Pick matrix (4.7) when the interpolation points are distinct or a generalized Pick matrix otherwise. Modifying the basis in $H(B)$ so that $a(0) = a_0$, it is not hard to see that the unique minimizer can be obtained by solving a system of linear equations in a_k/a_0, $k = 1, 2, \ldots, n$. This approach is taken in [41], where the case of distinct interpolation points is worked out in detail, and where also an efficient homotopy continuation method for solving the convex optimization problem (4.10) for an arbitrary P is presented. (This builds on some previous work [22] on the the Caratheodory extension problem.)

Consequently, the set of all rational solutions of degree at most n to the Nevanlinna-Pick interpolation problem are completely parameterized by P, and each such solution can be determined by solving the convex optimization problem to minimize J_w. Traditionally, one has only been able to determine the central solution corresponding to $P = 1$, and then in a much less direct way than described above. Our present method and parameterization in terms of P, first presented in the Nevanlinna-Pick setting in [13], provides us with an extra set of "tuning parameters", which, in the applications mentioned above, may be used to satisfy additional design specifications without having to increase the degree of the interpolant.

The existence of solutions to the Nevanlinna-Pick interpolation problem with degree constraint was established in [25] using topological degree theory. Uniqueness was proven in [26], using methods from [10], and in [13], using the variational problem to minimize the functional J_w defined above and Hardy space theory with both approaches making extensive use of Hardy spaces. In [12] we proved that the problem is well-posed by establishing that the foliation $\mathcal{F}_1$ defined by the fast filtering algorithm and the foliation $\mathcal{F}_2$ defined by Nevanlinna-Pick interpolation are transverse, each leaf in $\mathcal{F}_1$ intersecting each leaf in $\mathcal{F}_2$ in one and only one point. Applications of these result to high-resolution spectral estimation can be found in [14]. These results are also used in the design of robust controllers in [41, 42].

References

[1] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Hafner Publishing, New York, 1965.

[2] N.I. Ahiezer and M. Krein, *Some Questions in the Theory of Moments*, American Mathematical Society, Providence, Rhode Island, 1962.

[3] C. G. Bell, H. Fujisaki, J. M. Heinz, K. N. Stevens, and A. S. House, *Reduction of speech spectra by analysis-by-synthesis techniques*, J. Acoust. Soc. Amer., 33 (1961), pp. 1725–1736.

[4] M. S. Berger, *Nonlinearity and Functional Analysis*, Academic press, New York, 1977.

[5] A. Bialynicki-Birula and M. Rosenlicht, *Injective morphisms of real algebraic varieties*, Proc. Am. Math. Soc. **13** (1962), 200–203.

[6] C. I. Byrnes and A. Jhemi, *Shock waves for Riccati partial differential equations arising in nonlinear optimal control*, in Systems, Models and Feedback: Theory and Applications, A. Isidori and T. J. Tarn, editors, Birkhäuser, Boston, 1992, pp. 211–227.

[7] C. I. Byrnes and H. Frankowska, *Unicite des Controles Optimales et Chocs pour les Equationes Hamilton-Jacobi-Bellman et Riccati*, Comptes Rendus Acad. Sci. Paris, t.312 Ser I (1992) 427-431.

[8] C. I. Byrnes, The solution of nonlinear Lagrange and Bolza problems via Riccati partial differential equations, J. Math Systems, Estimation and Control **8** (1998), No.4, 1-54. The retrieval code is 87727.

[9] C. I. Byrnes, A. Lindquist, S.V. Gusev and A. S. Matveev, *A complete parametrization of all positive rational extensions of a covariance sequence*, IEEE Trans. Automatic Control **AC-40** (1995), 1841-1857.

[10] C. I. Byrnes, H. J. Landau and A. Lindquist, *On the well-posedness of the rational covariance extension problem*, in *Current and Future Directions in Applied Mathematics*, eds. M. Alber, B. Hu, J. Rosenthal, Birkhäuser, pp. 83-106, 1997.

[11] C. I. Byrnes, S.V. Gusev, and A. Lindquist, *A convex optimization approach to the rational covariance extension problem,* SIAM J. Control and Opt. **37:** 211–229, 1999.

[12] C. I. Byrnes and A. Lindquist, *On the duality between filtering and Nevanlinna-Pick interpolation*, SIAM J. Control and Optimization **39** (2000), 757–775.

[13] C. I. Byrnes, T. T. Georgiou and A. Lindquist, *A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint*, IEEE Trans. Automatic Control **AC-46** (Sept. 2001), to be published.

[14] C.I. Byrnes, T.T. Georgiou, and A. Lindquist, "A new approach to spectral estimation: A tunable high-resolution spectral estimator," IEEE Trans. on Signal Processing *SP-49* (Nov. 2000), 3189–3205.

[15] C. I. Byrnes, P. Enqvist, and A. Lindquist, *Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach*, SIAM J. Control and Optimization, to be published.

[16] C. I. Byrnes, S.V. Gusev, and A. Lindquist, *From finite covariance windows to modeling filters: A convex optimization approach*, SIAM Review **43** (Dec. 2001), 645–675.

[17] C. I. Byrnes, A. Lindquist, and Y. Zhou, *On the nonlinear dynamics of fast filtering algorithms*, SIAM J. Control and Optimization, 32 (1994), pp. 744–789.

[18] *M. K. K. Cevik and J. M. Schumacher*, *The robust regulation problem with robust stability*, Technical Report MAS-R9827, CWI, Amsterdam, 1999.

[19] Ph. Delsarte, Y. Genin and Y. Kamp, *On the role of the Nevanlinna-Pick problem in circuits and system theory*, Circuit Theory and Applications **9** (1981), 177–187.

[20] Ph. Delsarte, Y. Genin, Y. Kamp and P. van Dooren, *Speech modelling and the trigonometric moment problem*, Philips J. Res. **37** (1982), 277–292.

[21] J. C. Doyle, B. A. Frances and A. R. Tannenbaum, *Feedback Control Theory*, Macmillan Publ. Co., New York, 1992.

[22] P. Enqvist, *Spectral estimation by Geometric, Topological and Optimization Methods*, PhD thesis, Division of Optimization and Systems Theory, Department of Mathematics, Royal Instute of Technology, 2001.

[23] B. A. Francis, *A Course in H_∞ Control Theory*, Springer-Verlag, 1987.

[24] T.T. Georgiou, *Realization of power spectra from partial covariance sequences*, IEEE Transactions on Acoustics, Speech, and Signal Processing, **ASSP-35:** 438-449, 1987.

[25] T. T. Georgiou, *A topological approach to Nevanlinna-Pick interpolation*, SIAM J. Math. Analysis **18** (1987), 1248–1260.

[26] T. T. Georgiou, *The interpolation problem with a degree constraint*, IEEE Transactions on Automatic Control **44(3)**: 631-635, 1999.

[27] B. K. Ghosh, *Transcendental and interpolation methods in simultaneous stabilization and simultaneous partial pole placement problems*, SIAM J. Control and Optim., 24 (1986), pp. 1091–1109.

[28] U. Grenander and G. Szegö, *Toeplitz forms and their applications*, Univ. California Press, 1958.

[29] V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, Englewood Cliffs, New Jersey, 1974.

[30] J. Hadamard, *Sur les transformations planes*, C. R. Acad. Sci. Paris **142** (1906), 74.

[31] J. Hadamard, *Sur les transformations ponctuelles*, Bull. Soc.Math. France **34** (1906), 71–84; Oeuvres, 349–363.

[32] J. Hadamard, *Sur les correspondances ponctuelles*, Oeuvres, Editions du Centre Nationale de la Researche Scientifique, Paris (1968), 383–384.

[33] J. W. Helton, *Non-Euclidean analysis and electronics*, Bull. Amer. Math. Soc. **7** (1982), 1–64.

[34] J. W. Helton, *The distance from a function to H^∞ in the Poincaré metric: Electrical power transfer*, J. of Functional Analysis, **38** (1980), 273-314.

[35] R. E. Kalman, *Realization of covariance sequences*, Proc. Toeplitz Memorial Conference (1981), Tel Aviv, Israel, 1981.

[36] P. P. Khargonekar and A. Tannenbaum, *Non-Euclidean metrics and robust stabilization of systems with parameter uncertainty*, IEEE Trans. Automatic Control **AC–30** (1985), 1005–1013.

[37] H. Kimura, *Robust stabilizability for a class of transfer functions*, IEEE Trans. Automat. Control, AC–29 (1984), pp. 788–793.

[38] M.G. Krein and A.A. Nudelman, *The markov Moment Problem and Extremal Problems*, American Mathematical Society, Providence, Rhode Island, 1977.

[39] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*, Springer-Verlag, Berlin, 1976.

[40] J. W. Milnor, *Topology from Differentiable Viewpoint*, Revised Edition, Princeton University Press, Princeton, New Jersey, 1997.

[41] R. Nagamune, forthcoming PhD disertation, Division of Optimization and Systems Theory, Department of Mathematics, Royal Institute of Technology, Stockholm Sweden.

[42] R. Nagamune and A. Lindquist, *Sensitivity shaping in feedback control and analytic interpolation theory*, Optimal Control and Partial Differential Equations,J.L. Medaldi et al(editors), IOS Press, Amsterdam, 2001, pp. 404–413.

[43] D. J. Newman, *One-one polynomial maps*, Proc. Am. Math. Soc. **11** (1960), 867–870.

[44] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice–Hall, Englewood Cliffs, NJ, 1978.

[45] I. Schur, *On power series which are bounded in the interior of the unit circle I and II*, Journal fur die reine und angewandte Mathematik **148** (1918), 122–145.

[46] M. Spivak, *A comprehensive Introduction to Differential Geometry*, Volume 1, Publish or Perish, Berkeley, 1979.

[47] A. Tannenbaum, *Feedback stabilization of linear dynamical plants with uncertainty in the gain factor*, Int. J. Control, 32 (1980), pp. 1–16.

[48] A. Tannenbaum, *Modified Nevanlinna-Pick interpolation of linear plants with uncertainty in the gain factor*, Int. J. Control, 36 (1982), pp. 331–336.

[49] J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, Amer. Math.Soc. Colloquium Publications, **20**, Providence, R. I., 1956.

[50] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.

[51] D. C. Youla and M. Saito, *Interpolation with positive-real functions*, J. Franklin Institute **284** (1967), 77–108.

[52] E. Zeidler, *Nonlinear Functional Analysis and its Applications*, Vol III, Springer-Verlag, New York, 1985.