A Complete Parameterization of All Positive Rational Extensions of a Covariance Sequence

Christopher I. Byrnes, *Fellow, IEEE*, Anders Lindquist, *Fellow, IEEE*, Sergei V. Gusev, and Alexei S. Matveev

Abstract— **In this paper we formalize the observation that filtering and interpolation induce complementary, or "dual," decompositions of the space of positive real rational functions of degree less than or equal to n. From this basic result about the geometry of the space of positive real functions, we are able to deduce two complementary sets of conclusions about positive rational extensions of a given partial covariance sequence. On the one hand, by viewing a certain fast filtering algorithm as a nonlinear dynamical system defined on this space, we are able to develop estimates on the asymptotic behavior of the Schur parameters of positive rational extensions. On the other hand we are also able to provide a characterization of all positive rational extensions of a given partial covariance sequence. Indeed, motivated by its application to signal processing, speech processing, and stochastic realization theory, this characterization is in terms of a complete parameterization using familiar objects from systems theory and proves a conjecture made by Georgiou. Our basic result, however, also enables us to analyze the robustness of this parameterization with respect to variations in the problem data. The methodology employed is a combination of complex analysis, geometry, linear systems, and nonlinear dynamics.**

I. Introduction

Given a partial covariance sequence of length n, the problem of finding all positive rational extensions of degree at most n is a fundamental open problem with important applications in signal processing and speech processing [36], [30], [40], [22], [18], [35], [47], [39] and in stochastic realization theory and system identification [4], [56], [44], [45]. Such extension problems, of course, have a long history. Indeed, if one suppresses the rationality, the degree, or the positivity requirement, then the problem becomes considerably easier and solutions are known.

On the one hand, the problem of finding all positive real functions interpolating the given covariance sequence is known as the Carathéodory extension problem, dating from [19] and [20]. Carathéodory gave conditions for the existence of such an extension, conditions which were later reformulated by Toeplitz [55] in terms of positive definiteness of what is now known as the Toeplitz matrix of the covariance sequence. The problem of parameterizing all such solutions was solved by Schur [52]. In particular, given a covariance sequence, one obtains a parameterization of all interpolating positive real functions in terms of all extensions of a corresponding partial sequence of Schur parameters satisfying inequalities known as the "Schur conditions." Since the set of rational positive extensions of given degree is a finite dimensional manifold embedded in the set of infinite sequences satisfying the Schur conditions, however, finite sets of inequalities will not be sufficient to characterize rational positive extensions, underscoring the difficulty of using the Schur parameterization to directly characterize rationality. While a linear fractional parameterization of all positive real rational interpolants can be found, e.g., in [6], the degree of the interpolants represented in this parameterization may only be estimated except in certain special cases.

On the other hand, this problem can also be approached in the purely algebraic context of the partial realization problem [37], [38], [51], [32], [9], [2], with the additional constraint given by positivity, a point of view pioneered by Kalman [36]. Indeed, in [40], Kimura refines the linear fractional parameterization of positive real rational interpolants to incorporate the degree constraints at the expense of maintaining positivity. Solutions to the partial realization problem with the additional constraint of stability have also been obtained [3], [50]. In contrast to the Schur parameterization, however, parameterizations of partial realizations guarantee rationality (or even stability) of the appropriate degree, but do leave open the problem of characterizing positivity.

In this paper, we present several contributions to the rational covariance extension problem, including some basic asymptotics on the Schur parameters of rational positive extensions and a complete parameterization of all positive rational extensions of a given partial covariance sequence. These results follow as immediate corollaries of a more fundamental theorem concerning the geometry of the space of positive real rational functions of degree at most n. Building on earlier work in the literature, this result involves a blend of analytic and algebraic methods, explaining the extensive use of geometric concepts in the formulation of our main result.

The space $\mathcal{P}_n$ of positive real rational functions of degree at most n may be identified by viewing coefficients of the rational function as parameters, with an open subset of $2n$-dimensional Euclidean space. Our main result begins with the observation that filtering and interpolation define two "dual" or "complementary" decompositions of this space. Very briefly, the recent global analysis of certain fast filtering algorithms [41]–[43] as nonlinear dynamical systems [14]–[16] defined

Manuscript received January 5, 1994; revised April 15, 1995. Recommended by Associate Editor, P. van Dooren. This research was supported in part by grants from AFOSR, NSF, TFR, the Göran Gustafsson Foundation, the Royal Swedish Academy of Sciences, and Southwestern Bell.

C. I. Byrnes is with the Department of Systems Science and Mathematics, Washington University, St. Louis, MO 63130 USA.

A. Lindquist is with the Division of Optimization and Systems Theory, Royal Institute of Technology, 100 44 Stockholm, Sweden.

S. V. Gusev and A. S. Matveev are with the Department of Mathematics and Mechanics, St. Petersburg University, St. Petersburg 198904, Russia.

IEEE Log Number 9414725.

0018-9286/95$04.00 © 1995 IEEE

on $\mathcal{P}_n$, partition $\mathcal{P}_n$ into leaves of a foliation, where the leaves consist of the stable manifolds of the filtering algorithms. On the other hand, each choice of "window," consisting of the first n correlation coefficients or equivalently Schur parameters, also defines a leaf of a second foliation of $\mathcal{P}_n$. Our main result is that these two decompositions are complementary in the sense that these foliations are everywhere transverse.

From this basic result on the geometry of positive real rational functions, we derive several corollaries. First of all, while Schur parameters do characterize all meromorphic positive real covariance extensions, the basic question of which extensions are rational is open. Partial results in this direction are provided in [30] in terms of asymptotic properties of the Schur parameter sequence. For example, it is noted that for rational modeling filters, the Schur sequence is square summable and asymptotically rational. As it turns out, these properties are a consequence of more general asymptotic properties which we derive from the foliation of $\mathcal{P}_n$ into stable manifolds for the fast filtering algorithm.

As a second corollary, we give a complete bianalytic parameterization of all positive, rational extensions of a given degree. Our derivation proves a conjecture due to Georgiou [30], yielding a complete parameterization of rational positive real extensions in terms of the choice of zeros of the associated spectral density.

Obtaining such a complete parameterization ultimately boils down to the existence and uniqueness of solutions to a system of nonlinear equations, with inequality constraints reflecting the positivity requirements. In such settings, several questions arise for both analytical and numerical reasons: How many solutions exist, and are there *a priori* bounds on the norm of solutions given bounds on the norm of the data? In this connection, degree theory is a very powerful methodology derived earlier in this century, motivated by the study of solvability of algebraic and transcendental equations [49], [53]. Using an innovative application of topological degree theory, Georgiou was able to prove existence of a positive rational extension for any desired choice of spectral density zero structure. To provide a bona fide parameterization of all positive rational extensions, however, one would need to know that this correspondence is also unique.

Although a very useful tool for the study of existence, degree theory cannot be used, in general, to enumerate solutions to equations. Indeed, the definition of degree for differentiable functions involves sums of the signs of Jacobian determinants of the relevant function and these typically can assume either positive or negative values. Our main result on the transversality of the two basic foliations of $\mathcal{P}_n$ implies that these Jacobians can never vanish and hence can only be positive, reflecting the positivity of the associated covariance sequence. Thus, a simple argument using differentiable degree theory allows us to conclude from our main result that the correspondence studied by Georgiou is actually a complete bianalytic parameterization.

In fact, our proof of the Georgiou conjecture actually shows more, namely that the problem of parameterizing rational covariance extensions by means of covariance data and modeling filter zeros is well posed. Recall that a problem is well posed if solutions exist, are unique, and depend continuously on the data of the problem, so that small (or *a priori* bounded) perturbations in the problem data give rise to small (or *a priori* bounded) perturbations in the solution. The issue of small perturbations is typically addressed by showing that the appropriate Jacobian is everywhere nonsingular. *A priori* boundedness of solutions, phrased in a coordinate-free formulation, follow from the fact that the appropriate maps are proper, i.e., the solution set to a compact set of problem data is also compact. For example, spectral factorization is well posed, as we will illustrate for polynomials in Section III. For the rational covariance extension problem, we do more, obtaining continuity by proving analyticity. This increased regularity of solutions is important for reasons of analysis, approximation, and computation.

The analytic dependence of solutions on the data of the rational covariance extension problem also suggests several interesting questions concerning the analysis and computation of stochastic realizations of a rational covariance extension. This analysis requires some new ideas and is too involved to be included as a corollary. It turns out, however, that the modeling filters and the data in their state-space realizations can be determined from the covariance data and modeling filter zeros by solving a nonstandard Riccati equation [11]–[13]. This formulation also sheds some light on the important open problem of computing the minimal degree of partial stochastic realizations.

The body of the paper is organized as follows. In Section II, we state our main result about the geometry of positive real functions. In this section we also review some relevant features of the dynamics of a fast filtering algorithm and present some geometric constructions which are preliminary to the proof of our main result. The corollaries outline the consequences of the main result for rational covariance extensions, in terms of Schur parameters, modeling filters, and spectral densities.

In Section III, for the sake of illustration, we consider a simple special case of the parameterization conjecture, namely, the spectral factorization problem for polynomials. As well known as this result is, this derivation serves as a hint to the general methodology and also allows us to review basic degree theory and its use in determining well posedness in a familiar setting.

In his important paper [52], Schur also established a result asserting that the correspondence between partial covariance sequences and Schur sequences is birational and entire, for each finite window. A starting point for our proof is an extension, discussed in Section IV, of Schur's birational change of coordinates to include other data in the problem. We complete the proof of the transversality result for $\mathcal{P}_n$, by geometrically characterizing the tangent spaces of each of the leaves in terms of polynomials with particular properties. From an application of complex analysis and the positivity of the covariance sequence, it follows that the foliations defined by filtering and by interpolation are everywhere transverse.

In Section V, we conclude the paper by deriving some consequences of the main theorem and its proof, thereby giving proofs of the remaining assertions stated in Section III. By using this interpretation of filtering as a nonlinear dynamical system, we deduce a partial result concerning which sequences of Schur parameters correspond to rational positive extensions. Moreover, the proof given for spectral factorization in Section II can be easily extended to prove well posedness for the case of general correlation coefficients.

To set notation and for the sake of completeness, we give a proof of both existence and uniqueness, verifying that the technical hypotheses (e.g., properness) underlying degree theory are satisfied. That degree theory applies in this case reposes upon the fact that *a priori* bounds on the parameters in a modeling filter for a partial sequence are implied by *a priori* bounds on the zeroes of the candidate spectral density. This follows from a filtering interpretation, since such *a priori* bounds are inversely proportional to the Kalman filter steady state estimation error, which is bounded away from zero in a continuous manner.

Existence follows from a calculation of the degree for the maximum entropy filter, about which a great deal is known. Positivity of the appropriate Jacobian, and hence uniqueness of the parameterization, follows immediately from our main transversality result. Analyticity of the inverse, and hence well posedness of the problem, follows from the inverse function theorem.

II. Statement of the Main Result and its Corollaries

One of the goals of this paper is to provide a complete parameterization of all positive rational extensions of a given partial covariance sequence. Given a partial sequence

$$(1, c_1, \cdots, c_n); \quad c_i := E\{y(t+i)y(t)\} \tag{2.1}$$

of real correlation coefficients for a stationary stochastic process $\{y(t)\}_{t\in \mathbf{Z}}$ (normalized so that $c_0 = 1$), a strictly positive real[1] rational function v satisfying

$$v(z) = 1/2 + \sum_{i=1}^{\infty} \hat{c}_i z^{-i}; \quad \hat{c}_i = c_i \quad \text{for} \quad i = 1, 2, \cdots, n \tag{2.2}$$

and having degree at most n is said to be a positive rational extension of sequence (2.1). The problem of finding all such positive rational extensions is a fundamental open problem with important applications in signal processing and in speech processing (see, e.g., [36], [30], [40], [22], [18], [47], [39]). Ideally, one would like a complete parameterization of such extensions given in systems theoretic terms.

Of course the function v has a representation

$$v(z) = \frac{1}{2}\frac{b(z)}{a(z)} \tag{2.3}$$

where $a(z), b(z)$ are monic degree n polynomials having all roots inside the unit disc, i.e., where a and b are Schur polynomials. Moreover, $v(z)$ must satisfy

$$\mathrm{Re}\{v(z)\} > 0 \quad \text{for} \quad |z| = 1 \tag{2.4}$$

[1] A function $v(z)$ is strictly positive real if it is analytic for $|z| \geq 1$ and satisfies $v(z)+v(1/z) > 0$ on the unit circle; it is positive real if it is analytic for $|z| > 1$ and satisfies $v(z) + v(1/z) \geq 0$ at each point of the unit circle where v has no pole.

and therefore

$$\mathrm{Re}\{v(z)\} > 0 \quad \text{for} \quad |z| \geq 1.$$

One can also formulate this problem in terms of modeling filters $w(z)$ for the partial covariance sequence $(1, c_1, c_2, \cdots, c_n)$, i.e., minimum-phase stable rational functions w of degree at most n which satisfy

$$v(z) + v(1/z) = w(z)w(1/z) \tag{2.5}$$

for some positive real solution to the problem described above. Each $w(z)$ is a stable minimum-phase spectral factor for a spectral density

$$\Phi(z) = 1 + \sum_{i=1}^{\infty} \hat{c}_i(z^i + z^{-i});$$
$$\hat{c}_i = c_i \quad \text{for} \quad i = 1, 2, \cdots, n \tag{2.6}$$

corresponding to a positive rational extension of the given partial covariance sequence (2.1).

Such extension problems have a long history. In fact, if one drops either the rationality, the bound on the degree, or the positivity requirement, then the problem becomes considerably easier and solutions are known. On the one hand, the problem of finding all positive real functions $v(z)$, analytic outside the unit disc, which satisfy (2.2) is known as the Carathéodory extension problem, dating from [19] and [20]. The problem of parameterizing all such solutions was solved by Schur [52] and has exerted an important influence on classical function theory, interpolation theory, and operator theory. On the other hand, this problem can also be approached in the purely algebraic context of the partial realization problem, with the additional constraint given by positivity. Indeed, following Kalman (see, e.g., [36]) one can view $v(z)$ as the transfer function of the sequence (2.1), thought of as a sequence of Markov parameters.

Concerning the Carathéodory extension problem, it is well known [33], [52] that to any sequence $(1, c_1, c_2, \cdots, c_m)$ one can bijectively assign a sequence $(\gamma_0, \gamma_1, \cdots, \gamma_{m-1})$ of Schur parameters defined in terms of the Szegö polynomials $\{\varphi_0(z), \varphi_1(z), \varphi_2(z), \cdots\}$, a family of monic polynomials

$$\varphi_t(z) = z^t + \varphi_{t1}z^{t-1} + \cdots + \varphi_{tt}$$

which are orthogonal on the unit circle [1], [33]. The Schur parameters are then given by

$$\gamma_t = \frac{1}{r_t}\sum_{k=0}^{t} \varphi_{t,t-k}c_{k+1} \tag{2.7}$$

where $(r_0, r_1, r_2, \cdots)$ and the coefficients $\{\varphi_{ti}\}$ can be determined recursively [1] by

$$r_{t+1} = (1-\gamma_t^2)r_t; \qquad r_0 = 1 \tag{2.8}$$

and the Szegö–Levinson equations

$$\begin{aligned} \varphi_{t+1}(z) &= z\varphi_t(z) - \gamma_t\varphi_t^*(z); & \varphi_0(z) &= 1 \\ \varphi_{t+1}^*(z) &= \varphi_t^*(z) - \gamma_t z\varphi_t(z); & \varphi_0^*(z) &= 1 \end{aligned} \tag{2.9}$$

with $\varphi_t^*(z)$ being the reversed polynomials

$$\varphi_t^*(z) = \varphi_{tt}z^t + \varphi_{t,t-1}z^{t-1} + \cdots + 1.$$

Moreover [52], the Schur parameters satisfy the condition

$$|\gamma_i| < 1, \qquad i = 0, 1, \cdots, m-1 \tag{2.10}$$

if and only if the Toeplitz matrix

$$T_m = \begin{bmatrix} 1 & c_1 & \cdots & c_m \\ c_1 & 1 & \cdots & c_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_m & c_{m-1} & \cdots & 1 \end{bmatrix} \tag{2.11}$$

is positive definite. Furthermore, there is a bijection [52] between the class of positive real functions $v(z)$ and the class of sequences $(\gamma_0, \gamma_1, \gamma_2, \cdots)$ satisfying

$$|\gamma_i| < 1 \quad \text{for} \quad i = 0, 1, 2, \cdots. \tag{2.12}$$

In particular, given sequence (2.1), one obtains a parameterization of all positive real functions $v(z)$ satisfying (2.2), in terms of all extensions $(\gamma_n, \gamma_{n+1}, \cdots)$ of the corresponding partial sequence $(\gamma_0, \gamma_1, \cdots \gamma_{n-1})$ of Schur parameters satisfying (2.12). Using Schur's method, it is also possible [6, Chapter 22] to give a parameterization of all rational, positive real, interpolating functions, but incorporating bounds on the degree (as arise in the formulation of the rational covariance extension problem) has only been achieved at the expense of the positivity constraint (see e.g., [40]). This parameterization, which we shall refer to as the Kimura–Georgiou parameterization, was originally derived from different points of view.

Briefly, to the sequence $(\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$ one can associate the Szegö polynomials of the first and second kind, two bases for the vector space of polynomials of degree less than or equal to n. The Szegö polynomials of the first kind are the orthogonal polynomials, $\{\varphi_0(z), \varphi_1(z), \cdots \varphi_n(z)\}$, defined above, and those of the second kind, $\{\psi_0(z), \psi_1(z), \cdots, \psi_n(z)\}$, are merely the first kind polynomials corresponding to $(-\gamma_0, -\gamma_1, \cdots, -\gamma_{n-1})$, the Schur sequence obtained by switching signs. More explicitly, the Szegö polynomials of the second kind are generated by the recursion

$$\begin{aligned} \psi_{t+1}(z) &= z\psi_t(z) + \gamma_t\psi_t^*(z); \quad \psi_0(z) = 1 \\ \psi_{t+1}^*(z) &= \psi_t^*(z) + \gamma_t z\psi_t(z); \quad \psi_0^*(z) = 1. \end{aligned} \tag{2.13}$$

If the sequence $(\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$ satisfies (2.12), then Kimura [40] and Georgiou [30] independently showed that any positive real $v(z)$, of degree at most n, satisfying (2.2) has a representation

$$v(z) = \frac{1}{2}\frac{\psi_n(z) + \alpha_1\psi_{n-1}(z) + \cdots + \alpha_n\psi_0(z)}{\varphi_n(z) + \alpha_1\varphi_{n-1}(z) + \cdots + \alpha_n\varphi_0(z)} \tag{2.14}$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n$ are real numbers. In fact, it was shown in [30] and later also in [15] that the representation in (2.14) of rational functions with fixed Schur parameters holds for all α, regardless of positivity.

We now give a geometric interpretation of the Kimura–Georgiou parameterization. It will be convenient to regard polynomials $a(z)$ as points in various Euclidean spaces, using the sequence of coefficients, denoted by a as parameters. Thus, if $a(z)$ has degree n, then a is a vector in $\mathbf{R}^{n+1}$. If, in addition, $a(z)$ is monic of degree n, then we shall emphasize this point and suppress the monic leading coefficient, so that $a(z)$ is represented as a vector a in $\mathbf{R}^n$. Consider the open subset $\mathcal{P}_n \subset \mathbf{R}^{2n}$ of pairs (a, b) of monic polynomials such that

$$v(z) = \frac{1}{2}\frac{b(z)}{a(z)}$$

is strictly positive real. We first note (see Section IV) that the Kimura–Georgiou parameterization induces a birational diffeomorphic change of coordinates, from (a,b)-coordinates to (α, γ)-coordinates, on $\mathcal{P}_n$. We next introduce the subset $\mathcal{P}_n(\gamma)$ of $\mathcal{P}_n$ obtained by fixing a partial Schur sequence $(\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$ with $|\gamma_i| < 1, i = 1, 2, \cdots, n-1$. Thus, $\mathcal{P}_n(\gamma)$ is parameterized by choices of α such that

$$\frac{1}{2}\frac{b(z)}{a(z)} = \frac{1}{2}\frac{\psi_n(z) + \alpha_1\psi_{n-1}(z) + \cdots + \alpha_n\psi_0(z)}{\varphi_n(z) + \alpha_1\varphi_{n-1}(z) + \cdots + \alpha_n\varphi_0(z)}$$

is positive real. Geometrically, the decomposition

$$\mathcal{P}_n = \bigcup_\gamma \mathcal{P}_n(\gamma)$$

is an important example of what is known as a foliation of the open manifold $\mathcal{P}_n$. Intuitively, a foliation is a decomposition of a manifold into disjoint connected submanifolds, called leaves, with the additional property that in the neighborhood of any point the leaves vary in a sufficiently smooth way (see Section IV for a precise definition and for proofs). Foliations have served as the nonlinear enhancement of linear subspaces in the development of nonlinear control over the past two decades. For the example at hand, in [10] it was shown that $\mathcal{P}_n(\gamma)$ is diffeomorphic to Euclidean space and is therefore connected. With this in mind, we note (see [17]) that the Kimura–Georgiou parameterization shows that this decomposition is sufficiently regular to define a foliation Γ of $\mathcal{P}_n$ into the leaves $\mathcal{P}_n(\gamma)$.

Finally, there is a complementary geometric construction, arising from an interpretation of the fast filtering algorithm [41], [42] as a nonlinear dynamical system [16]. This nonlinear dynamical system is a reformulation of a fast algorithm for Kalman filtering [41], [42] and is also related to the Schur algorithm (see, for example, [21]). More precisely, if $(\alpha, \gamma) \in \mathcal{P}_n$ and the maps $A, G: \mathbf{R}^n \to \mathbf{R}^{n\times n}$ are defined as the matrix in (2.15) (as shown at the bottom of the next page) and

$$G(\alpha) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_n & -\alpha_{n-1} & -\alpha_{n-2} & \cdots & -\alpha_1 \end{bmatrix} \tag{2.16}$$

then the dynamical system

$$\alpha(t+1) = A(\gamma(t))\alpha(t), \qquad \alpha(0) = \alpha \tag{2.17a}$$
$$\gamma(t+1) = G(\alpha(t+1))\gamma(t), \qquad \gamma(0) = \gamma \tag{2.17b}$$

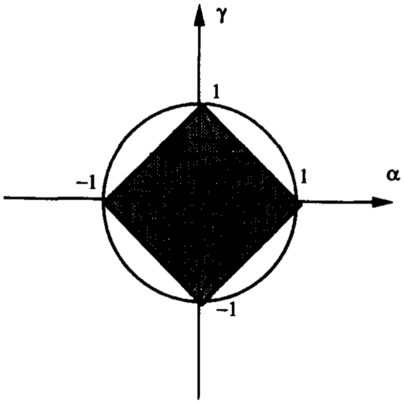

Fig. 1.

initiated at (α, γ) evolves on an invariant manifold X_{α_∞} and converges to $(\alpha_\infty, 0)$, where $\alpha_\infty \in \mathcal{P}_n(0) = \mathcal{S}_n$, the space of monic Schur polynomials of degree n [16].

This dynamical system is essentially the Kalman filter rewritten in a universal form so that the system and covariance data appear in the initial conditions. It also has the feature that the Kalman gain may be computed recursively as the dynamical system evolves. More precisely, consider a linear stochastic system

$$\begin{aligned} x_{t+1} &= Fx_t + Bu_t \\ y_t &= h'x_t + d'u_t \end{aligned} \tag{2.18}$$

driven by (normalized) white noise $\{u_t\}$, where x_t is an n-dimensional state process, y_t is a scalar output process, and, for simplicity, (h, F) is in observer-canonical form. Then the linear least-squares estimate $\hat{x}_t$ given the observed data $y_0, y_1, \cdots, y_{t-1}$ is generated by the Kalman filter

$$\hat{x}_{t+1} = \hat{x}_t + k_t(y_t - h'\hat{x}_t); \qquad \hat{x}_0 = 0. \tag{2.19}$$

Now, suppose $\Phi(z)$ is the spectral density of the process $\{y_t\}$ and $v(z)$ is its positive real part, i.e., the unique positive real function such that $v(z) + v(z^{-1}) = \Phi(z)$. Using as initial conditions the corresponding (α, γ) in the Kimura–Georgiou parameterization (2.14), we may propagate $(\alpha(t), \gamma(t))$ or, in (a, b)-coordinates $(a(t), b(t))$, using the dynamical system (2.17). Then, as explained in more detail in [16], the Kalman gain is given by

$$k_t = \tfrac{1}{2}[b(t) - a(t)] - a(0). \tag{2.20}$$

As it turns out [16], the invariant manifold

$$\mathcal{W}^s(\alpha_\infty, 0) = X_{\alpha_\infty} \cap \mathcal{P}_n$$

is the global stable manifold in $\mathcal{P}_n$ at $(\alpha_\infty, 0)$, and it plays an important role in the geometry of the space $\mathcal{P}_n$. More explicitly, denote by a_t and b_t the monic polynomials corresponding to the updated vectors $a(t)$ and $b(t)$. Then X_{α_∞} is described by a system

$$F_i(\alpha, \gamma) = 0 \qquad i = 1, 2, \cdots, n$$

of nonlinear equations. These are derived in [16] by eliminating r_t in the $n+1$ algebraic relations

$$\begin{aligned} r_t[a_t(z)b_t(z^{-1}) &+ a_t(z^{-1})b_t(z)] \\ &= 2r_\infty \alpha_\infty(z)\alpha_\infty(z^{-1}) \end{aligned} \tag{2.21}$$

or, equivalently

$$\frac{1}{2}\frac{b_t(z)}{a_t(z)} + \frac{1}{2}\frac{b_t(z^{-1})}{a_t(z^{-1})} = r_\infty \frac{\alpha_\infty(z)}{a_t(z)}\frac{\alpha_\infty(z^{-1})}{a_t(z^{-1})}$$

where r_t is defined by (2.8).

It is shown in [16] that the infinite Schur sequence $(\gamma_0, \gamma_1, \gamma_2, \cdots)$ corresponding to $(\alpha, \gamma) \in \mathcal{P}_n$ is generated by (2.17) via

$$\gamma_k(t) = \gamma_{t+k} \tag{2.22}$$

and hence the sequence $(\alpha(t), \gamma(t))_{t \in \mathbf{Z}}$ is completely contained in $\mathcal{P}_n$. Since $|\gamma_i| < 1$ for $i = 0, 1, 2, \cdots, r_0 = 1 \geq r_1 \geq r_2 \geq \cdots$, and it can be shown [16] that $r_t \to r_\infty > 0$ as $t \to \infty$.

The global stable manifold $\mathcal{W}^s(\alpha_\infty, 0)$ at $(\frac{1}{3}, 0)$ is depicted as a subset of $\mathcal{P}_1$ in Fig. 1. Also depicted is $\mathcal{P}_1(\frac{1}{2})$ and a closed curve, exterior to $\mathcal{P}_1$ whose importance we will discuss later.

In general the decomposition of $\mathcal{P}_n$ as a union of the global stable manifolds $\mathcal{W}^s(\alpha_\infty, 0)$ defines a second foliation Ω of $\mathcal{P}_n$. As Fig. 1 suggests, the leaves of the foliations Γ and Ω are transverse, i.e., at a point of intersection of the leaves of these two foliations the corresponding tangent spaces are complementary subspaces. Indeed, using the characterization of the tangent spaces to the stable manifold $\mathcal{W}^s(\alpha_\infty, 0)$ developed in [16], in Section IV we prove that the intersection of $\mathcal{W}^s(\alpha_\infty, 0)$ with $\mathcal{P}_n(\gamma)$ is in fact always transverse. In such a case, one says that two foliations are complementary.

Main Theorem: The positive real region $\mathcal{P}_n$ is connected and invariant under the filtering algorithm (2.17), which is globally convergent on $\mathcal{P}_n$. In fact, $\mathcal{P}_n$ is foliated by the stable manifolds $\mathcal{W}^s(\alpha_\infty, 0)$ of the equilibrium set $\mathcal{P}_n(0)$. The set $\mathcal{P}_n$ is also foliated into leaves given by the submanifolds $\mathcal{P}_n(\gamma)$. Moreover, these foliations, Γ and Ω, are complementary.

$$A(\gamma) = \begin{bmatrix} \dfrac{1}{1-\gamma_{n-1}^2} & \dfrac{\gamma_{n-1}\gamma_{n-2}}{(1-\gamma_{n-1}^2)(1-\gamma_{n-2}^2)} & \cdots & \dfrac{\gamma_{n-1}\gamma_0}{(1-\gamma_{n-1}^2)\cdots(1-\gamma_0^2)} \\ 0 & \dfrac{1}{1-\gamma_{n-2}^2} & \cdots & \dfrac{\gamma_{n-2}\gamma_0}{(1-\gamma_{n-2}^2)\cdots(1-\gamma_0^2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{1}{1-\gamma_0^2} \end{bmatrix} \tag{2.15}$$

There are several important corollaries of our main theorem. The first is related to asymptotic properties of the Schur parameter sequence. That the Schur sequence of a rational modeling filter tends to zero has been noted by several authors, some of whom derive asymptotic estimates also implying absolute or conditional summability (see e.g., [8] and [31]). For example, it is noted in [30] that for rational modeling filters the Schur sequence is square summable and asymptotically rational. In fact, the Schur sequence is actually in ℓ_p for any p satisfying $p \geq 1$. As it turns out, these properties are a consequence of stable manifold theory for the dynamical system (2.17), and they can be strengthened in the form of lower and upper bounds on the decay rates of the Schur sequence. We shall use vector norms defined by a positive definite matrix P, i.e., in terms of quantities

$$\|x\|_P^2 = x'Px.$$

Corollary 2.1: Given a positive sequence $(1, c_1, \cdots, c_n)$, and consequently a sequence of Schur parameters $(\gamma_0, \gamma_1, \cdots \gamma_{n-1})$ satisfying (2.12), consider the positive real function $v(z)$ corresponding to a sequence of Schur parameters $\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \cdots$ satisfying

$$\hat{\gamma}_i = \gamma_i, \qquad i = 0, 1, \cdots, n-1$$

and

$$|\hat{\gamma}_i| < 1, \qquad i = 0, 1, 2, \cdots.$$

Then a necessary condition for v to be rational is

$$|\hat{\gamma}_i| = O(\lambda^i)$$

for some $\lambda \in [0, 1)$. In fact, λ is the maximum of the moduli of the zeroes of the corresponding polynomial $\alpha_\infty(z)$. Moreover, if $v(z)$ is rational of degree at most n, then, for some $m, 1 \leq m \leq n$, and some sufficiently large T, there exist $\lambda_1, \lambda_2 \in (0, 1)$ and a positive definite $m \times m$ matrix P so that the vector sequence $\gamma(t) = (\hat{\gamma}_t, \hat{\gamma}_{t+1}, \cdots, \hat{\gamma}_{t+m-1})', t = 0, 1, 2, 3, \cdots,$ satisfies

$$\lambda_1 \|\gamma(t)\|_P \leq \|\gamma(t+1)\|_P \leq \lambda_2 \|\gamma(t)\|_P \tag{2.23}$$

for all $t \geq T$, where $\|x\|_P^2 := x'Px$.

Since the set of rational positive extensions of degree n is a finite dimensional manifold embedded in the "infinite cube" of extensions satisfying (2.12), it cannot be characterized by a finite set of inequalities in $(\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$, which would define an open (and hence infinite dimensional) subspace (in the product topology). Moreover, inequalities (2.23) show that the set of rational positive extensions of degree n is a finite dimensional manifold embedded in ℓ_2, as observed by Georgiou [30] using different methods. Since the inequalities (2.23) define a subset of ℓ_2 with a nonempty interior, for the same reason they will not be sufficient to characterize rational positive extensions. These observations illustrate the difficulty of using inequalities in the Schur parameters to directly characterize rationality, as suggested by Kalman [36].

In contrast to the Schur parameterization, the Kimura–Georgiou parameterization (2.14) guarantees that $v(z)$ will be rational of degree n, but leaves open the rather challenging problem of characterizing positivity in terms of the α parameters. Such a characterization would be especially interesting if it expressed the design freedom available in the choice of suitable α-parameters in familiar systems theoretic terms. For example, the desirability of obtaining a parameterization of partial stochastic realizations in terms of poles or zeros of a candidate spectral density has been noted by several authors [57], [7] and has important applications to signal and to speech processing. (See, e.g., [34].) In this direction, Burg developed an algorithm for computing the partial Schur parameter sequence from observed data and then proposed considering the modeling filter obtained from the simple extension

$$(\gamma_0, \gamma_1, \cdots \gamma_{n-1}, 0, 0, 0, \cdots).$$

In harmony with Proposition 2.1, this extension does in fact yield a rational, strictly positive real $v_0(z)$ having degree n, corresponding to the choice

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

in the Kimura–Georgiou parameterization (2.14). The associated modeling filter $w_0(z)$ is known as the maximum entropy filter, since it can be obtained by maximum entropy methods (see, e.g., [34], [26], [5]). This filter is also sometimes referred to as the autoregressive (AR) model since, as we shall see below, the maximal entropy filter $w_0(z)$ corresponds to a spectral density

$$\Phi_0(z) = w_0(z) w_0(1/z)$$

with no zeros. Due to its simplicity, this solution is widely used but for many reasons (arising, for example, in speech processing [22], in spectral analysis [48], and in recursive prediction and identification [47], [7]), it is desirable to allow for solutions to (2.2) corresponding to modeling filters with nontrivial zeros.

Indeed, several algorithms which yield more general modeling filters are now available. Such filters satisfy

$$w(z) w(z^{-1}) = \frac{d(z, z^{-1})}{a(z) a(z^{-1})} \tag{2.24}$$

where

$$d(z, z^{-1}) = d_0 + d_1(z + z^{-1}) + \cdots + d_n(z^n + z^{-n}) \tag{2.25}$$

is a pseudopolynomial which is positive on the unit circle. Earlier work by Dewilde and Dym [24], [25] gave a systematic construction, starting with a nonrational $v(z)$ satisfying (2.12), of a rational modeling filter with prescribed zeros, but for which the interpolation condition (2.2) will not be satisfied.

On the other hand, Georgiou's work on the covariance extension problem focused on the freedom in choosing zeros of the spectral density while retaining the interpolation constraint (2.2). Using degree theory, in 1983 Georgiou [29] proved that for any sequence (2.1) and any choice of pseudopolynomial d satisfying

$$d(z, z^{-1}) > 0 \quad \text{for} \quad |z| = 1 \tag{2.26}$$

and having degree less than or equal to n, there is a rational positive real function $v(z)$ satisfying condition (2.2). If such a v were unique, as conjectured by Georgiou [30], then this result would lead to a complete parameterization, in the desired systems theoretic terms, of all rational positive real interpolants $v(z)$ of the sequence (2.1).

In geometric terms, we denote by $\mathcal{D}_n$ the set of pseudopolynomials having degree at most n and satisfying (2.26). In this language, Georgiou's conjecture is equivalent to the assertion that the mapping of $(n+1)$-manifolds

$$f: \mathbf{R}_+ \times \mathcal{P}_n(\gamma) \to \mathcal{D}_n$$

defined via

$$f(a_0, a, b) = a_0^2[a(z)b(z^{-1}) + a(z^{-1})b(z)]$$

is one-one and onto. In fact, our main theorem implies that this factorization problem is well posed, i.e., f is one-one, onto, and has an analytic (and hence continuous) global inverse.

Corollary 2.2: Let $(1, c_1, \cdots, c_n)$ be a positive sequence, i.e., a sequence satisfying the condition $T_n > 0$. Then, the mapping f is a proper analytic bijection with an analytic inverse.

Our proof of Corollary 2.2 will repose on the fact that $\mathrm{Jac}(f)$ never vanishes on $\mathbf{R}_+ \times \mathcal{P}_n(\gamma)$. Of course, f is defined on all of $\{a_0\} \times \mathcal{P}_n$ and the kernel of its Jacobian at a point (a_0, α, γ) consists of the tangent space to the invariant manifold $\mathcal{W}^s(\alpha_\infty, 0)$. Therefore to say that $\mathrm{Jac}(f)$ does not vanish on $\{a_0\} \times \mathcal{P}_n(\gamma)$ at the point (a_0, α, γ) is to say that no vector tangent to $\mathcal{P}_n(\gamma)$ at (α, γ) is also tangent to $\mathcal{W}^s(\alpha_\infty, 0)$, which is part of our main result. Fig. 1 illustrates that this is true for the case $n = 1$. Indeed, the closed curve in Fig. 1 is in fact the circle in $\mathbf{R}^2$ exterior to $\mathcal{P}_1$ on which the Jacobian vanishes. Of course, in higher dimensions the hypersurface on which the Jacobian vanishes becomes much more complex. Nonetheless, our main result proves that this hypersurface never intersects $\mathcal{P}_n$.

Corollary 2.3: Let $(1, c_1, \cdots, c_n)$ be a positive sequence, i.e., a sequence satisfying the condition $T_n > 0$. Then, to any pseudopolynomial $d(z, z^{-1})$ of degree less than or equal to n, which is positive on the unit circle, there corresponds one and only one strictly positive real rational function (2.3) of degree at most n which satisfies (2.2) and

$$v(z) + v(z^{-1}) = a_0^{-2} \frac{d(z, z^{-1})}{a(z)a(z^{-1})} \tag{2.27}$$

where $a(z)$ is a monic Schur polynomial and a_0 is a positive real number. Conversely, to any strictly positive real rational function (2.3) satisfying (2.2), there exists a pseudopolynomial $d(z, z^{-1})$, uniquely defined up to multiplication by a positive number, with the properties described above. Moreover, the solution depends analytically on the covariance data and the choice of pseudopolynomial.

This result allows for an interesting interpretation and refinement concerning the corresponding parameterization of modeling filters. Corresponding to each positive pseudopolynomial $d(z, z^{-1})$, as defined in this theorem, there corresponds a unique spectral density

$$\Phi(z) = a_0^{-2} \frac{d(z, z^{-1})}{a(z)a(z^{-1})}$$

such that

$$v(z) + v(z^{-1}) = \Phi(z).$$

Although the spectral factorization problem

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

does not have a unique solution, there is exactly one spectral factor w which is a modeling filter in the sense defined earlier. Such a w is stable and minimum phase in the sense that all zeros lie in the open unit disc and the numerator polynomial has degree n. Consequently, $\rho := w(\infty) \neq 0$. For example, in this representation the numerator polynomial of the maximum entropy modeling filter $w_0(z)$ will be ρz^n, but, as far as the question of shaping the process $\{y(t)\}_{t \in \mathbf{Z}}$ from white noise is concerned, the power z^n may be deleted. With this in mind, we are ready to state the following corollary.

Corollary 2.4: Let $(1, c_1, \cdots, c_n)$ be a given positive partial covariance sequence. Then given any Schur polynomial

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \cdots + \sigma_n$$

there exists a unique monic Schur polynomial $a(z)$ of degree n and a unique $\rho \in (0, 1]$ such that

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

is a minimum phase spectral factor of a spectral density $\Phi(z)$ satisfying

$$\Phi(z) = 1 + \sum_{i=1}^{\infty} \hat{c}_i (z^i + z^{-i});$$
$$\hat{c}_i = c_i \quad \text{for} \quad i = 1, 2, \cdots, n.$$

In particular, the solutions of the rational positive extension problem are in one-one correspondence with self-conjugate sets of n points (counted with multiplicity) lying in the open unit disc, i.e., with all possible zero structures of modeling filters. Moreover, the modeling filter depends analytically on the covariance data and the choice of zeros of the spectral density.

Well posedness can be expressed in terms of the commutative diagram

$$\begin{array}{ccc} \mathbf{R}_+ \times \mathcal{P}_n(\gamma) & \xrightarrow{f} & \mathcal{D}_n \\ g \searrow & & \nearrow h \\ & \mathbf{R}_+ \times \mathcal{S}_n & \end{array}$$

where

$$h(a_0, a) = a_0^2 a(z) a(z^{-1})$$

arises in spectral factorization (see Section III) and where

$$g = h^{-1} \circ f.$$

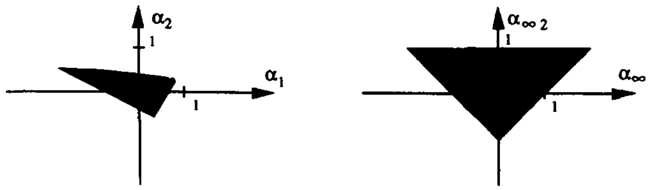

Fig. 2.

The function g can also be expressed as a function

$$g(a_0, \alpha, \gamma) = (a_0, \alpha_\infty)$$

where α_∞ can be defined in terms of the dynamical system in (α, γ)-coordinates [16] and where $r_\infty = \rho^2$ in Corollary 2.4. In particular

$$0 < r_\infty \leq 1 \quad \text{and so} \quad 0 < \rho \leq 1 \tag{2.28}$$

as claimed in Corollary 2.4.

As an illustration of Corollary 2.4, Fig. 2 depicts the connected open submanifolds $\mathcal{P}_2(\gamma)$ and $\mathcal{S}_2$, for $\gamma = (\frac{1}{2}, \frac{1}{3})$.

These sets form the domain and codomain of the map g, restricted to the surface defined by $a_0 = 1$, for this case. Corollary 2.4 asserts that any α such that $(\alpha, \gamma) \in \mathcal{P}_2(\gamma)$ determines [for example, via the convergence of the dynamical system (2.17)] a (limit) $\alpha_\infty(z)$ which is a Schur polynomial. Conversely, for any point α_∞ in $\mathcal{S}_2$, there is one and only one $(a, b) \in \mathcal{P}_2(\gamma)$ and hence one, and only one, α such that $(\alpha, \gamma) \in \mathcal{P}_2(\gamma)$ that defines a modeling filter $w(z)$ having the zeros of $\alpha_\infty(z)$.

III. Spectral Factorization and Degree Theory

In this section, we describe the general proof of Corollary 2.3 in the context of the much simpler problem when $\gamma := (\gamma_0, \gamma_1, \cdots, \gamma_{n-1}) = 0$. When $\gamma = 0$, the Szegö polynomials coincide with the standard monomials $\{z^i\}$, and the rational covariance extension problem reduces to a geometric proof of a very familiar problem, spectral factorization for polynomials. This analysis gives us an opportunity to introduce, in a very familiar context, the basic concepts from degree theory, a very powerful methodology derived earlier in this century motivated by the study of solvability of algebraic and transcendental equations (see especially [49]).

If $c_1 = c_2 = \cdots = c_n = 0$, then all covariance extensions generated by a rational function of degree at most n also vanish, so the basic parameterization problem is to find all a, b such that

$$\frac{1}{2}\frac{b(z)}{a(z)} = \frac{1}{2}$$

which is then given by $b(z) = a(z)$, for any choice of Schur polynomial $a(z)$ having degree n. Corollary 2.3, however, is still interesting in this case, the question being whether the equation

$$a_0^2 a(z) a(z^{-1}) = d(z, z^{-1}) \tag{3.1}$$

is solvable, in a continuous or regular fashion for a monic Schur polynomial a and a real constant $a_0 > 0$. We refer to the space of all Schur polynomials

$$a(z) = z^n + a_1 z^{n-1} + \cdots + a_n \tag{3.2}$$

as the Schur region $\mathcal{S}_n$. In this context, the spectral factorization problem is whether the function

$$f\colon \mathbf{R}_+ \times \mathcal{S}_n \to \mathcal{D}_n \tag{3.3}$$

defined via

$$f(a_0, a) = a_0^2 a(z) a(z^{-1})$$

is one-one and onto. Since $\mathbf{R}_+ \times \mathcal{S}_n \subset \mathbf{R} \times \mathbf{R}^n$ and $\mathcal{D}_n \subset \mathbf{R}^{n+1}$ are open connected subsets, this problem can be approached using differential analysis. For example, a tangent vector (v_0, v) to $\mathbf{R}_+ \times \mathcal{S}_n$ at a point $(a_0, a) \in \mathbf{R}_+ \times \mathcal{S}_n$ corresponds to a pair (v_0, v) where $v_0 \in \mathbf{R}$ and v is a polynomial

$$v(z) = v_1 z^{n-1} + \cdots + v_n$$

of degree less than or equal to $n - 1$. In particular, for $(a_0, a) \in \mathbf{R}_+ \times \mathcal{S}_n$ and a sufficiently small $\epsilon > 0$, $(a_0 + \epsilon v_0, a + \epsilon v) \in \mathbf{R}_+ \times \mathcal{S}_n$. The Jacobian matrix of f at a point (a_0, a), denoted by $\mathrm{Jac}_{(a_0,a)}(f)$, assigns to a tangent vector (v_0, v) the directional derivative of f in the direction (v_0, v), i.e.,

$$\mathrm{Jac}_{(a_0,a)}(f)(v_0, v) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}\{f(a_0 + \epsilon v_0, a + \epsilon v) - f(a_0, a)\}.$$

Explicitly

$$\begin{aligned}\mathrm{Jac}_{(a_0,a)}(f)(v_0, v) &= a_0^2[a(z)v(z^{-1}) + a(z^{-1})v(z)] \\ &\quad + 2a_0 v_0 a(z) a(z^{-1}).\end{aligned}$$

We define the operator $S(a)\colon V_n \to W_n$ from the vector space V_n of polynomials having degree less than or equal to n into the vector space W_n of symmetric pseudopolynomials of degree less than or equal to n via

$$S(a)v = a(z)v(z^{-1}) + a(z^{-1})v(z). \tag{3.4}$$

In this notation, we have

$$\mathrm{Jac}_{(a_0,a)}(f)(v_0, v) = S(a)(a_0^2 v + a_0 v_0 a). \tag{3.5}$$

We recall that, for any $a \in \mathcal{S}_n$ satisfying (3.1), the unit circle formulation of Orlando's formula [28] shows that

$$\det\ S(a) \neq 0 \quad \text{for} \quad a \in \mathcal{S}_n.$$

In fact, it is easily seen [23] that

$$\det\ S(a) = \prod_i \prod_j (1 - \rho_i \rho_j) > 0$$

where ρ_i are the roots of $a(z)$, which all lie in $\{z\colon |z| < 1\}$.

In the light of (3.5), a straightforward calculation yields

$$\det\ \mathrm{Jac}_{(a_0,a)}(f) = a_0^{2n+1} \det\ S(a)$$

so that

$$\det\ \mathrm{Jac}_{(a_0,a)}(f) > 0 \quad \text{for all} \quad (a_0, a) \in \mathbf{R}_+ \times \mathcal{S}_n. \tag{3.6}$$

Our interest in this calculation lies in its consequences for the solvability of the (3.1). If (3.1) is solvable, then several questions arise for both analytical and numerical reasons: How many solutions exist, and are there *a priori* bounds on $||a||$ given bounds on $||d||$?

First note that, since $\mathrm{Jac}_{(a_0,a)}(f)$ is everywhere nonsingular, the inverse function theorem implies that solutions to (3.1), for a given d, form a set of isolated points, i.e., a set with no cluster points. That the number of solutions must be finite then follows from the existence of *a priori* bounds. Phrased in a coordinate-free formulation, we shall need to know that f is proper, i.e., if K is a compact set in $\mathcal{D}_n$ then $f^{-1}(K)$ is also compact. Since $\{d\}$ is compact and since the solution set has been shown to consist of isolated points, a consequence of properness is that the number of solutions of (3.1) will be finite.

Before using degree theory to enumerate the solution set, we first verify that f is proper. Using (3.1), we see that f has a continuous extension

$$\overline{f}\colon \overline{\mathbf{R}_+ \times \mathcal{S}_n} \to \overline{\mathcal{D}}_n.$$

Moreover, it is straightforward to check that

$$\overline{f}(\partial(\mathbf{R}_+ \times \mathcal{S}_n)) \subset \partial\mathcal{D}_n. \tag{3.7}$$

Now, if $K \subset \mathcal{D}_n$ is compact, $\overline{f}^{-1}(K)$ is closed in $\overline{\mathcal{S}}_n$ by continuity of $\overline{f}$. Next, note that $f^{-1}(K) = \overline{f}^{-1}(K)$ since

$$\overline{f}^{-1}(K) \cap \partial(\mathbf{R}_+ \times \mathcal{S}_n) = \emptyset$$

by virtue of (3.7). In particular, $f^{-1}(K)$ is closed, and it remains only to check that $f^{-1}(K)$ is bounded. This follows from two observations. For a point $(a_0, a) \in f^{-1}(K), a$ is a monic Schur polynomial and therefore has bounded coefficients. As for a_0, the constant term in $d \in K$ is

$$d_0 = a_0^2(1 + a_1^2 + \cdots + a_n^2)$$

which achieves a maximum on K providing an upper bound on a_0^2.

We now review some basic facts from degree theory. Suppose more generally that $U, V \subset \mathbf{R}^{n+1}$ are open connected subsets and that

$$F\colon U \to V$$

is an infinitely differentiable (C^∞), proper function on U. We are interested in solutions to the equation

$$y = F(x). \tag{3.8}$$

For $x \in U$, we denote the Jacobian matrix of F at x by $\mathrm{Jac}_x(F)$. A point $y \in V$ is called a regular value for F if either

i) $F^{-1}(y)$ is empty or
ii) For each $x \in F^{-1}(y), \mathrm{Jac}_x(F)$ is nonsingular.

Regular values not only exist but, according to Sard's Theorem [49], are dense. Since for a regular value y of type ii), $F^{-1}(y)$ is finite, we may then compute the finite sum

$$\deg_y(F) = \sum_{F(x)=y} \text{sign det } \mathrm{Jac}_x(F). \tag{3.9}$$

If y is a regular value of type i), we set $\deg_y(F) = 0$.

For example, if $U = V = \mathbf{R}$ and $F(x) = x^2$, then any nonzero y is a regular value and $\deg_y(F) = 0$. If $F(x) = x^3$, then any nonzero y is again a regular value and $\deg(F) = 1$. More generally, if $F(x)$ is any odd order polynomial

$$F(x) = a_0 x^{2n+1} + a_1 x^{2n} + \cdots + a_{2n+1}, \qquad a_0 > 0$$

then F is proper and $\deg_y(F) = 1$ for every regular value y. In contrast, however, the computation of the degree of such polynomials, regarded as complex polynomials so that $U = V = \mathbf{C}$, is remarkably simpler. Indeed, the Cauchy–Riemann equations imply that $\det \mathrm{Jac}_x(F) \geq 0$ so that, for any regular value $y, \deg_y(F)$ is equal to the algebraic degree of the polynomial.

For our purposes, the main conclusions of degree theory [49] assert:

i) The degree, $\deg_y(F)$, of F with respect to y is independent of the choice of regular value y.
ii) Therefore, we may define the degree of F as

$$\deg(F) = \deg_y(F)$$

for any regular y.

iii) If $\deg(F) \neq 0$, then F maps U onto V.

The proof of iii) is simple: Regular values y are dense, and for each such $y, F^{-1}(y)$ is nonempty. Therefore $F(U) \subset V$ is dense, but $F(U)$ is closed in V since F is proper, and so $F(U) = V$.

Returning to the spectral factorization problem, we note that all positive pseudopolynomials are regular values for f. For either (3.1) is not solvable, so that $f^{-1}(d) = \emptyset$, or (3.6) holds for all solutions of (3.1).

One first concludes that

$$\deg(f) > 0$$

so that spectral factorization is always possible, but a more careful summation in (3.9), in the light of (3.6), shows that

$$\deg(f) = \#\{(a_0, a)\colon f(a_0, a) = d\} \tag{3.10}$$

for any d. The choice $d(z, z^{-1}) \equiv 1$ then leads to the unique solution

$$a_0 = 1, \quad a(z) = z^n$$

so that we conclude

$$\deg(f) = 1.$$

From this it follows [see (3.10)] that $f\colon \mathbf{R}_+ \times \mathcal{S}_n \to \mathcal{D}_n$ is one-one and onto, i.e., for each positive pseudopolynomial there is a unique spectral factor.

Finally, we observe that f^{-1} exists and, from the implicit function theorem, that f^{-1} is analytic (since f is), and hence continuous.

As we shall see in Section V, this proof applies, *mutatis mutandis*, to the case of general correlation coefficients $\{c_1, c_2, \cdots, c_n\}$.

IV. Proof of the Main Theorem

We have already seen in Section II that the fast filtering algorithm leaves $\mathcal{P}_n$ invariant. We now turn to the geometric assertions in our main theorem.

Proposition 4.1: The open submanifold $\mathcal{P}_n$ is connected.

Proof: Since $\mathcal{P}_n(\gamma)$ has been shown to be connected, to show that $\mathcal{P}_n$ is connected it suffices to show that for $\gamma^{(1)}, \gamma^{(2)}$ with $\mathcal{P}_n(\gamma^{(1)})$ and $\mathcal{P}_n(\gamma^{(2)})$ nonempty, there exists a path from some point in $\mathcal{P}_n(\gamma^{(1)})$ to some point in $\mathcal{P}_n(\gamma^{(2)})$. For this purpose we choose the points $(0, \gamma^{(1)})$ and $(0, \gamma^{(2)})$, corresponding to the maximum entropy solutions. Since the Schur conditions (2.10) define a convex set, the path

$$\{0, \lambda\gamma^{(1)}) + (0, (1-\lambda)\gamma^{(2)}\}$$

lies in $\mathcal{P}_n$. □

Recall that a foliation $\mathcal{F}$ of dimension m on a smooth manifold M of dimension n is a partition of M into a family of disjoint, connected m-dimensional submanifolds L_β, called the leaves of the foliation, such that i) $M = \cup_\beta L_\beta$, and ii) each point $x \in M$ has a Euclidean neighborhood U and coordinates $(x_1, \cdots, x_n)$ for which the equations

$$x_1 = 0, \quad x_2 = 0, \quad \cdots, \quad x_{n-m} = 0$$

define the connected components of the nonempty intersections $U \cap L_\beta$.

Foliations arise naturally in several ways:

1) A linear subspace and its parallel translations are the leaves of a (linear) foliation on $\mathbf{R}^n$.
2) If U is an open subset of $\mathbf{R}^n$ and $(x_1, \cdots, x_n)$ are Euclidean coordinates, the connected components of the submanifolds defined by
$$x_1 = 0, \quad x_2 = 0, \quad \cdots, \quad x_{n-m} = 0$$
define a foliation of U of dimension m.
3) The connected components of the level sets of a smooth function $h: M \to N$, between an m-manifold M and an n-manifold N with an everywhere surjective Jacobian are the leaves of an $m-n$ dimensional foliation on M.

An important related concept is that of the distribution defined by the foliation. The distribution Δ is the collection of m-dimensional subspaces $\Delta(x)$ of the tangent spaces $T_x(M)$ defined by

$$\Delta(x) = \{v \in T_x(M) | v \text{ is tangent to } L_x\}$$

where L_x is the leaf of $\mathcal{F}$ passing through x. Finally, we say that two foliations, $\mathcal{F}_1$ and $\mathcal{F}_2$, of M are complementary as foliations provided their distributions define complementary subspaces at each x in M, i.e., for each x the subspaces $\Delta_1(x)$ and $\Delta_2(x)$ are complementary.

In our calculations, we can represent tangent vectors in familiar algebraic terms. Given a point $(a, b) \in \mathcal{P}_n$ it follows that tangent vectors to $\mathcal{P}_n$ at (a, b) can be represented as a pair of polynomials, (u, v), where

$$u(z) = u_1 z^{n-1} + \cdots + u_n$$
$$v(z) = v_1 z^{n-1} + \cdots + v_n.$$

Proposition 4.2: Γ defines a smooth foliation of $\mathcal{P}_n$ into the leaves $\mathcal{P}_n(\gamma)$. The distribution of Γ is the collection of subspaces

$$\Delta_\Gamma(a, b) = T_{(a,b)}\mathcal{P}_n(\gamma) = \{(u, v) | \; av - bu = r, \deg r \le n-1\}. \quad (4.1)$$

Proof: In fact, Γ is an illustration of the second manner, itemized above, in which foliations arise. More precisely, in his important paper [52], Schur established a bijective relation between the covariance sequence $(1, c_1, c_2, \cdots, c_n)$ and the Schur parameters $(\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$ defined by (2.7)–(2.9). Less known is his result asserting that this correspondence is birational and entire, for each n, provided (2.10) holds. The Kimura–Georgiou parameterization is an extension of Schur's birational change of coordinates to the other data in the problem, while relaxing the constraint that (2.10) holds. This additional data is essentially contained in the Kimura–Georgiou parameterization (2.14), which we shall now treat in detail following our treatment of this in [16].

Lemma 4.3: Let $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_{n-1})$ be an arbitrary vector in $\mathbf{R}^n$ such that $\gamma_k^2 \ne 1$ for $k = 0, 1, \cdots, n-2$, let $\{\varphi_k(z), \psi_k(z); k = 0, 1, \cdots, n-1\}$ be the corresponding polynomials generated by (2.9) and (2.13), and set $c_1 := \gamma_0$ and

$$c_{k+1} := r_k\gamma_k - \sum_{j=0}^{k-1} \varphi_{k,k-j} c_{j+1} \quad (4.2)$$

for $k = 1, 2, \cdots, n-1$, where $r_1, r_2, \cdots, r_n$ are defined by (2.8). Let $a(z)$ and $b(z)$ be arbitrary monic polynomials of degree n such that

$$\frac{b(z)}{2a(z)} = \frac{1}{2} + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_n z^{-n} + \cdots. \quad (4.3)$$

Then there is a unique $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbf{R}^n$ such that

$$a(z) = \varphi_n(z) + \alpha_1\varphi_{n-1}(z) + \cdots + \alpha_n \quad (4.4a)$$
$$b(z) = \psi_n(z) + \alpha_1\psi_{n-1}(z) + \cdots + \alpha_n. \quad (4.4b)$$

We note that this parameterization constitutes in fact a bona fide change of coordinates. In the language of classical algebraic geometry, the map defined by (4.3) is a birational isomorphism [54], i.e., a rational map with a rational inverse. More explicitly, consider the set

$$U_\gamma = \{(\alpha, \gamma) \in \mathbf{R}^{2n} | \gamma_i^2 \ne 1, i = 0, 1, \cdots, n-2\}.$$

Also, by virtue of (4.3), the generalized "correlation" coefficients $c_1, c_2, \cdots, c_n$ are functions of (a, b) so that we may define the open, dense set

$$V_c = \{(a, b) \in \mathbf{R}^{2n} | \det T_i \ne 0, i = 1, 2, \cdots, n-1\}.$$

In [16] it is shown that the polynomial map $\mathcal{T}: U_\gamma \to V_c$, defined by (4.4), sending $(\alpha, \gamma) \in \mathbf{R}^{2n}$ to $(a, b) \in \mathbf{R}^{2n}$, is a bijection with a rational inverse $\mathcal{T}^{-1}$. Hence $\mathcal{T}$ is indeed a birational isomorphism.

In these coordinates on $\mathcal{P}_n$, the connected submanifolds $\mathcal{P}_n(\gamma)$ are defined by coordinate constraints on γ and hence define the leaves of an n-dimensional foliation, which we refer to as Γ. The following result characterizes which tangent vectors (u, v) to $\mathcal{P}_n$ at (a, b) lie in $T_{(a,b)}\mathcal{P}(\gamma)$.

Lemma 4.4: For any $(a,b) \in \mathcal{P}_n(\gamma)$

$$T_{(a,b)}\mathcal{P}_n(\gamma) = \{(u,v)|av - bu = r, \deg\, r \leq n-1\}. \quad (4.5)$$

Proof: Let $T_{(a,b)}\mathcal{P}_n$ be the tangent space of $\mathcal{P}_n$ at (a,b). Denote by V_k the vector space of polynomials having degree at most k and by W the subspace of $T_{(a,b)}\mathcal{P}_n$ defined via

$$W = \{(u,v)|\ \deg\,(av - bu) \leq n-1\}.$$

We first note that

$$W = M_{(a,b)}^{-1}(V_{n-1})$$

where the linear map $M_{(a,b)}: T_{(a,b)}\mathcal{P}_n \to V_{2n-1}$ is defined via

$$M_{(a,b)}(u,v) = av - bu.$$

Since ker $M_{(a,b)} \subset W$, by complementarity of rank and nullity for the linear operator $M_{(a,b)}|_W$ we have

$$\dim\, W = \dim\, M_{(a,b)}(W) + \dim\, \ker\, M_{(a,b)}. \quad (4.6)$$

Let δ be the degree of the greatest common divisor θ of a and b. Then $a = \tilde{a}\theta$ and $b = \tilde{b}\theta$ where of course $\tilde{a}$ and $\tilde{b}$ are coprime. Now

$$av - bu = \theta(\tilde{a}v - \tilde{b}u)$$

so, if $(u,v) \in W$

$$\deg\,(\tilde{a}v - \tilde{b}u) \leq n - 1 - \delta$$

and consequently

$$\dim\, M_{(a,b)}(W) = n - \delta. \quad (4.7)$$

Also, if $(u,v) \in \ker\, M_{(a,b)}$, we have $av - bu = 0$, i.e.,

$$\frac{v(z)}{u(z)} = \frac{b(z)}{a(z)} = \frac{\tilde{b}(z)}{\tilde{a}(z)}$$

so u and v must have a common divisor ϑ of degree $\delta - 1$. Since $\vartheta \in V_{\delta-1}$ is arbitrary

$$\dim\, \ker\, M_{(a,b)} = \delta. \quad (4.8)$$

From (4.6)–(4.8) we may conclude that

$$\dim\, W = n$$

and it therefore suffices to prove

$$W \subset T_{(a,b)}\mathcal{P}_n(\gamma) \quad (4.9)$$

because $\mathcal{P}_n(\gamma)$ is an open subset of $\mathbf{R}^n$. Since the sequence $\{\gamma_0, \gamma_1, \cdots, \gamma_{n-1}\}$ determines the sequence $\{c_1, c_2, \cdots, c_n\}$ and conversely, to see that (4.9) holds it suffices to compute the first n Laurent coefficients in the Newton quotient

$$\frac{1}{\epsilon}\left[\frac{1}{2}\frac{b+\epsilon v}{a+\epsilon u} - \frac{1}{2}\frac{b}{a}\right]$$

and to prove that, if

$$\frac{1}{2}\frac{b+\epsilon v}{a+\epsilon u} = \frac{1}{2} + c_1(\epsilon)z^{-1} + c_2(\epsilon)z^{-2} + \cdots$$

then

$$\lim_{\epsilon\to 0}\frac{c_i(\epsilon) - c_i}{\epsilon} = 0 \quad \text{for} \quad i = 1, 2, \cdots, n.$$

This, however, follows immediately from the expansion

$$\frac{b+\epsilon v}{a+\epsilon u} - \frac{b}{a} = \frac{\epsilon(av - bu)}{a^2} - \frac{\epsilon^2 u(av - bu)}{a^2(a+\epsilon u)}$$

and the condition $(u,v) \in W$. □

This concludes the proof of Proposition 4.2. □

The next result will be important both in proving that Ω defines a foliation and that Γ and Ω are complementary as foliations. It also forms the basis for a degree theoretic proof of Georgiou's conjecture.

Lemma 4.5—Transversality Lemma: There are no nonzero polynomials p and q of degree at most n such that

$$S(a)q + S(b)p = 0 \quad (4.10)$$

and

$$aq - bp = r \quad (4.11)$$

where r is a polynomial of degree less than n.

Proof: Suppose that p and q are polynomials of at most degree n satisfying (4.10) and (4.11). We want to prove that $p = q = 0$. To this end, first note that, in view of (4.11), the function

$$g(z) := \frac{q(z)}{b(z)} - \frac{p(z)}{a(z)} = \frac{r(z)}{a(z)b(z)}$$

has relative degree at least $n+1$ and is analytic outside a disc contained in the open unit disc so that it has the Laurent expansion

$$g(z) = g_0 z^{-n-1} + g_1 z^{-n-2} + \cdots \quad (4.12)$$

there. Likewise $g(z^{-1})$ is analytic in an open disc containing the closed unit disc, and in this region it has the Taylor expansion

$$g(z^{-1}) = g_0 z^{n+1} + g_1 z^{n+2} + \cdots.$$

Now a simple calculation shows that

$$g(z) - g(z^{-1}) = \frac{h(z)}{b(z)a(z^{-1})} - \frac{h(z^{-1})}{a(z)b(z^{-1})}$$

where

$$h(z) := a(z^{-1})q(z) + b(z)p(z^{-1})$$

so that

$$g(z) - g(z^{-1}) = -h(z^{-1})\frac{d(z,z^{-1})}{a(z)a(z^{-1})b(z)b(z^{-1})}$$

and therefore

$$\int_{|z|=1} |h(z)|^2 \frac{d(z,z^{-1})}{|a(z)|^2|b(z)|^2}\frac{dz}{z} = \int_{|z|=1} h(z)[g(z^{-1}) - g(z)]\frac{dz}{z}. \quad (4.13)$$

h, however, is a pseudopolynomial of degree less than or equal to n, i.e.,

$$h(z) = h_0 + h_1(z + z^{-1}) + \cdots + h_n(z^n + z^{-n})$$

and therefore $h(z)g(z^{-1})z^{-1}$ is holomorphic, having no poles in the open disc containing the closed unit disc. Similarly, the Laurent expansion of $h(z)g(z)z^{-1}$ in the region where (4.12) holds has only negative powers of z of order larger than one. Consequently (4.13) is zero, which implies that $h(e^{i\theta}) \equiv 0$, because $d(z, z^{-1}), |a(z)|^2$ and $|b(z)|^2$ are all positive on the unit circle. Therefore, by the identity theorem, $h \equiv 0$ in the whole complex plane so that

$$g(z) = g(z^{-1}).$$

But $g(z)$ has only negative powers of z and $g(z^{-1})$ only nonnegative powers of z in an annulus containing the unit circle and hence $g \equiv 0$. Since, therefore, $r \equiv 0$, we have

$$q(z) = \frac{b(z)}{a(z)}p(z)$$

which substituted into (4.10) yields

$$\left[\frac{b(z)}{a(z)}\frac{b(z^{-1})}{a(z^{-1})}\right][a(z)p(z^{-1}) + a(z^{-1})p(z)] = 0.$$

Since $(a, b) \in \mathcal{P}_n(\gamma)$, the first factor is positive on the unit circle and so

$$a(e^{i\theta})p(e^{-i\theta}) + a(e^{-i\theta})p(e^{i\theta}) = 0$$

for all θ, and therefore, by the identity theorem

$$S(a)p = 0.$$

Since a is a Schur polynomial, and hence has no reciprocal roots, the unit circle version of Orlando's formula [28] (also see [23] and [16, Lemma 5.5]) implies that p, and hence q, is identically zero. $\square$

Remark 4.6: Since this is so important in the geometry of $\mathcal{P}_n$, we found it instructive to give an alternative, algebraic proof of the key argument in the Transversality Lemma using positivity of a Toeplitz form. The proof that $h(z) \equiv 0$ depends in a crucial way on the fact that the function

$$k(z) = \frac{d(z, z^{-1})}{a(z)a(z^{-1})b(z)b(z^{-1})}$$

is positive on the unit circle and hence is a spectral density. If

$$k(z) = k_0 + k_1(z + z^{-1}) + k_2(z^2 + z^{-2}) + \cdots$$

is its Laurent expansion in an open annulus containing the unit circle, then the positivity of k on the unit circle implies that the Toeplitz matrix

$$K = \begin{bmatrix} k_0 & k_1 & \cdots & k_{2n} \\ k_1 & k_0 & \cdots & k_{2n-1} \\ \vdots & \vdots & \ddots & \vdots \\ k_{2n} & k_{2n-1} & \cdots & k_0 \end{bmatrix}$$

is positive definite. Since the Laurent expansion of

$$g(z) - g(z^{-1}) = -h(z^{-1})k(z)$$

in a neighborhood of the unit circle lacks powers of orders $0, \pm 1, \pm 2, \cdots, \pm n$ in z, we also have

$$Kh = 0$$

where h is the column vector of the coefficients $(h_n, h_{n-1}, \cdots, h_{-n})$ of the pseudopolynomial h. Finally, since the degree of h is n, we must have $h \equiv 0$, as claimed.

Proposition 4.7: Ω defines a smooth foliation of $\mathcal{P}_n$. The distribution of Ω is the collection of subspaces

$$\begin{aligned}\Delta_\Omega(a, b) &= T_{(a,b)}(\mathcal{W}^s(\alpha_\infty, 0)) \\ &= \{(u, v) | S(a)q + S(b)p = 0\} \end{aligned} \tag{4.14}$$

where $p = d_0 u - v_0 a, q = d_0 v - v_0 b$. Here $d_0 = 1 + \langle a, b \rangle$ and

$$v_0 = \tfrac{1}{2}(\langle a, v \rangle + \langle b, u \rangle) = \tfrac{1}{2}(\langle a, q \rangle + \langle b, p \rangle)$$

where $\langle a, b \rangle$ is the inner product $\Sigma_{i=1}^n a_i b_i$ of a and b regarded as vectors in $\mathbf{R}^n$.

Proof: As we shall demonstrate, Ω is an illustration of the third manner, discussed above, in which foliations arise, viz. as the level sets of a mapping h: $M \rightarrow N$, between an m-manifold M and an n-manifold N with an everywhere surjective Jacobian.

In [16, Theorem 5.9] it is shown that the stable manifolds $\mathcal{W}^s(\alpha_\infty, 0)$ are in fact closed n-dimensional submanifolds of $\mathcal{P}_n$ which decompose $\mathcal{P}_n$ into their disjoint union.

In this light, recall (Section II) that

$$\mathcal{W}^s(\alpha_\infty, 0) = X_{\alpha_\infty} \cap \mathcal{P}_n$$

and that X_{α_∞} is defined by (A.1). Inspection shows that $d_0 > 0$ on $\mathcal{D}_n$, so that $X_{\alpha_\infty} \cap \mathcal{P}_n$ may be defined [16, p. 759] in (α, γ)-coordinates by the invariant quantities

$$\frac{d_i(\alpha, \gamma)}{d_0(\alpha, \gamma)} = \kappa_i, \qquad i = 1, 2, \cdots, n$$

where $d_i(\alpha, \gamma) := d_i^{(n)}(1, \alpha, \gamma)$ as defined in Lemma A.1, and where the κ_i are determined by evaluating the left-hand side at any point (α, γ) lying on $\mathcal{W}^s(\alpha_\infty, 0)$.

We shall consider the smooth n-dimensional manifold N which is the subset of $\mathcal{D}_n$ defined by the constraint $d_0 = 1$. We then define h: $\mathcal{P}_n \rightarrow N$ by

$$h_i(a, b) = \frac{d_i(a, b)}{d_0(a, b)} \quad \text{for} \quad i = 0, 1, \cdots, n.$$

In particular, for $d \in N$ we see that $h^{-1}(d) = \mathcal{W}^s(\alpha_\infty, 0)$ where $d = 2r_\infty \alpha_\infty(z)\alpha_\infty(z^{-1})$. Moreover, each submanifold $\mathcal{W}^s(\alpha_\infty, 0)$ is realized as $h^{-1}(d)$ for some $d \in N$. If (u, v) is tangent to the point (a, b) in $\mathcal{P}_n$ then it is straightforward to compute the directional derivative of h, and hence its Jacobian, via a Newton quotient to obtain

$$\mathrm{Jac}_{(a,b)}(h) = S(a)q + S(b)p$$

where $p = d_0 u - v_0 a$, $q = d_0 v - v_0 b$, $d_0 = 1 + \langle a, b \rangle$ and $v_0 = \frac{1}{2}$ $(\langle a, v \rangle + \langle b, u \rangle)$. Also, it is straightforward to show that $v_0 = \frac{1}{2}$ $(\langle a, q \rangle + \langle b, p \rangle)$.

We claim that $\mathrm{Jac}_{(a,b)}(h)$ maps $T_{(a,b)}(\mathcal{P}_n(\gamma))$ onto $T_{h(a,b)}(N)$. Since each tangent space has dimension n this is equivalent to the assertion that no nonzero tangent vector (u,v) in $T_{(a,b)}(\mathcal{P}_n(\gamma))$ is annihilated by $\text{Jac}_{(a,b)}(h)$. Suppose (u,v) is tangent to $\mathcal{P}_n(\gamma)$ and note that

$$r = aq - bp = av - bu \tag{4.15}$$

also has degree less than or equal to $n-1$. Therefore, if (u,v) is annihilated by $\text{Jac}_{(a,b)}(h)$, we must have $p = q = 0$, and hence $u = v = 0$, by the transversality lemma.

In particular the Jacobian of h is everywhere surjective and therefore Ω defines a foliation. It also follows that the tangent space of the level set $h^{-1}(d)$ is given by ker $\text{Jac}_{(a,b)}(h)$, proving our second claim as well. □

Taken together, these characterizations of the foliations Γ and Ω, and their respective distributions, and the tranversality lemma allow us to conclude the heart of our main theorem.

Proposition 4.8: Γ and Ω are complementary foliations of $\mathcal{P}_n$. In particular any intersection of $\mathcal{W}^s(\alpha_\infty, 0)$ with $\mathcal{P}_n(\gamma)$ is transverse. That is, no nonzero vector tangent to $\mathcal{P}_n(\gamma)$ at (α, γ) is also tangent to $\mathcal{W}^s(\alpha_\infty, 0)$.

Proof: Let $(u,v) \in \Delta_\Gamma(a,b) \cap \Delta_\Omega(a,b)$. Then (4.10) holds with p and q defined as in Proposition 4.2, and, by Proposition 4.7, (4.15) has degree less or equal to $n-1$. Consequently, by the Transversality Lemma, $p = q = 0$ and hence $u = v = 0$.

V. Proofs of the Corollaries

In this section, we derive some consequences of the main theorem and its proof, thereby giving proofs of the remaining assertions in Section II. As noted in the discussion of filtering as a dynamical system in Section II, for any strictly positive real $v(z)$ the sequence γ_t converges to zero. We shall now show that this convergence is actually geometric as claimed in Proposition 2.1.

Proof of Corollary 2.1: We shall analyze the asymptotic behavior of the scalar sequence γ_t interpreted as the vector sequence $\gamma(t)$, which is propagated along with $\alpha(t)$ by the nonlinear dynamical system (2.17). It is known [16] that, if $(\alpha(0), \gamma(0)) \in \mathcal{W}^s(\alpha_\infty, 0)$ then $(\alpha(t), \gamma(t)) \in \mathcal{W}^s(\alpha_\infty, 0)$ for $t \geq 0$ and $(\alpha(t), \gamma(t)) \to (\alpha_\infty, 0)$ as $t \to \infty$. The first necessary condition, geometric decay, follows from the fact [16] that $\mathcal{W}^s(\alpha_\infty, 0)$ is the global stable manifold for the nonlinear system (2.17) and that, on this invariant manifold the linear approximation to the dynamical system is defined by a linear operator with characteristic polynomial $\alpha_\infty(z)$. More explicitly, the Transversality Lemma (Lemma 4.5) asserts, in particular, that the tangent space to $\mathcal{W}^s(\alpha_\infty, 0)$ at the point $(\alpha_\infty, 0)$ intersects the tangent space to $\gamma = 0$ only in the zero vector. By the implicit function theorem, for γ in a neighborhood of zero there exists an analytic function F for which $\mathcal{W}^s(\alpha_\infty, 0)$ is locally the graph of $\alpha = F(\gamma)$. Therefore, to analyze the asymptotic properties of $\gamma(t)$ we may consider the autonomous nonlinear dynamical system

$$\gamma(t+1) = G(F(\gamma(t)))\gamma(t)$$

which can be expanded as

$$\gamma(t+1) = G(\alpha_\infty)\gamma(t) + H(\gamma(t))$$

where

$$\|H(\gamma(t))\| = O(\|\gamma(t)\|^2).$$

Since $\alpha_\infty(z)$ is a Schur polynomial, $G(\alpha_\infty)$ is a stability matrix from which our first claim follows.

The second assertion also follows from this fact with a little more work. There exist a positive definite matrix P satisfying the Lyapunov equation

$$P = G(\alpha_\infty)' P G(\alpha_\infty) + I.$$

Now, let us first consider the case that $\alpha_n \neq 0$. Then $\alpha_{n\infty} \neq 0$ also so that $G(\alpha_\infty)$ is nonsingular. In fact, if $\alpha_{n\infty} = 0$, then $d_n = 0$, and hence, by Proposition A.1, $\alpha_n = 0$. Therefore, $G(\alpha_\infty)' P G(\alpha_\infty) > 0$, and hence all eigenvalues of P are greater than one. Then, setting $m = n$, it is straightforward to compute that

$$\|\gamma(t+1)\|_P^2 = \|\gamma(t)\|_P^2 - \|\gamma(t)\|^2 + \epsilon(\gamma)\|\gamma(t)\|^2$$

where ϵ is $O(\|\gamma(t)\|)$ for $t \geq T$ for some $T \geq 0$. Since

$$-\frac{1}{\lambda_{\min}}\|x\|_P^2 \leq -\|x\|^2 \leq -\frac{1}{\lambda_{\max}}\|x\|_P^2$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimal and maximal eigenvalue of P, respectively, we obtain the desired inequalities for suitable choices of λ_1 and λ_2 satisfying

$$0 < \lambda_1 < \sqrt{1 - \frac{1}{\lambda_{\min}}} < 1$$

and

$$0 < \sqrt{1 - \frac{1}{\lambda_{\max}}} < \lambda_2 < 1.$$

Secondly, if $\alpha_n = \cdots = \alpha_{m+1} = 0$ but $\alpha_m \neq 0$, it is easily seen that the dynamical system reduces in a few iterations to one for which α and γ have dimensions m. Then applying the above argument to this reduced system, the required result is obtained. Finally, if $\alpha = 0$, we have $\gamma_k = 0$ for $k \geq n$ (maximum entropy solution) so statement (2.23) holds trivially for $t \geq n$. □

As described in Section II, Corollary 2.2 is equivalent to the assertion that the function

$$f: \mathbf{R}_+ \times \mathcal{P}_n(\gamma) \to \mathcal{D}_n \tag{5.1}$$

is a bijective analytic diffeomorphism, where

$$f(a_0, a, b) = a_0^2 S(a) b = a_0^2[a(z)b(z^{-1}) + a(z^{-1})b(z)]. \tag{5.2}$$

Since f is proper (see Appendix), f has a well-defined degree, deg f. To compute the Jacobian effectively, we need to obtain an intrinsic description of the tangent vectors to $\mathbf{R}_+ \times \mathcal{P}_n(\gamma)$ at a point (a_0, a, b). Denoting the tangent space to $\mathbf{R}_+ \times \mathcal{P}_n(\gamma)$ at (a_0, a, b) by $T_{(a_0,a,b)}\mathbf{R}_+ \times \mathcal{P}_n(\gamma)$ and the tangent space to $\mathcal{P}_n(\gamma)$ at (a,b) by $T_{(a,b)}\mathcal{P}_n(\gamma)$, there is a natural direct sum decomposition

$$T_{(a_0,a,b)}\mathbf{R}_+ \times \mathcal{P}_n(\gamma) \simeq T_{a_0}\mathbf{R}_+ \oplus T_{(a,b)}\mathcal{P}_n(\gamma).$$

Hence, for a tangent vector $(u_0, u, v) \in T_{(a_0,a,b)}\mathbf{R}_+ \times \mathcal{P}_n(\gamma)$, the Jacobian of f at (a_0, a, b) becomes

$$\begin{aligned}\mathrm{Jac}_{(a_0,a,b)}(f)(u_0, u, v) &= S(a)(a_0^2 v + a_0 u_0 b)\\ &\quad + S(b)(a_0^2 u + a_0 u_0 a). \quad (5.3)\end{aligned}$$

For simplicity of notation, we define polynomials, having degree less or equal to n, via

$$p = a_0^2 u + a_0 u_0 a \quad (5.4a)$$

$$q = a_0^2 v + a_0 u_0 b. \quad (5.4b)$$

We observe that p and q also satisfy

$$aq - bp = r, \quad \deg\, r \le n - 1$$

because (u, v) is tangent to $\mathcal{P}_n(\gamma)$. Finally, note that to say $p = q = 0$ is to say that $u_0 = 0$ and that $u = v = 0$. In this language, the following result is a direct consequences of the Transversality Lemma (Lemma 4.5).

Lemma 5.1: For each $(a, b) \in \mathcal{P}_n(\gamma), \mathrm{Jac}_{(a_0,a,b)}(f)$ is nonsingular.

Thus far, we have shown that the Jacobian of f is always nonsingular on the connected open manifold $\mathbf{R}_+ \times \mathcal{P}_n(\gamma)$ and therefore the sign of its determinant cannot change. In the next lemma, we compute this sign showing it to be positive and, in fact, evaluate the degree in this case, obtaining

$$1 = \deg_d(f) = \#\{(a_0, a, b) | f(a_0, a, b) = d\}$$

for all $d \in \mathcal{D}_n$. Hence, f is both one-one and onto; see iii) in Section III.

Lemma 5.2: For all $d \in \mathcal{D}_n$

$$\deg_d f = 1.$$

Proof: Since $\mathbf{R}_+ \times \mathcal{P}_n$ is connected (Proposition 4.1) and

$$\det\, \mathrm{Jac}_{(a_0,a,b)}\,(f) \neq 0 \quad \text{for all } (a_0, a, b) \in \mathbf{R}_+ \times \mathcal{P}_n \quad (5.5)$$

it follows that $\mathrm{Jac}_{(a_0,a,b)}\,(f)$ is sign definite on $\mathbf{R}_+ \times \mathcal{P}_n$. To prove that

$$\deg_d\,(f) = \#\{(a_0, a, b) | f(a_0, a, b) = d\} \quad (5.6)$$

for all $d \in \mathcal{D}_n$, it therefore suffices to evaluate $\mathrm{Jac}_{(a_0,a,b)}\,(f)$ at any point in $\mathbf{R}_+ \times \mathcal{P}_n$. Choosing a point with $\gamma = 0$, we know from Section III that

$$\det\, \mathrm{Jac}_{(a_0,a,a)}\,(f) > 0 \quad \text{for all} \quad (a_0, a) \in \mathbf{R}_+ \times \mathcal{S}_n$$

and hence (5.6) follows from the definition (3.9) of degree.

The lemma will follow if we can show that there is one and only one point $(a_0, a, b) \in \mathbf{R} \times \mathcal{P}(\gamma)$ corresponding to $d(z, z^{-1}) \equiv 1$. This is easy if $\gamma = 0$, in which case the problem reduces to the spectral factorization problem of Section III yielding $a(z) = b(z) = z^n$ and $a_0 = 1$, i.e., $\alpha_\infty = 0$. For an arbitrary γ consider the mapping g defined in Section II. From (2.17) we see that $\alpha = 0$ yields $\alpha_\infty = 0$ so there is at least one α with $(\alpha, \gamma) \in \mathcal{P}_n(\gamma)$ corresponding to $\alpha_\infty = 0$, namely the maximum entropy solution. If we can show that there is no nonzero α with $(\alpha, \gamma) \in \mathcal{P}_n(\gamma)$ for which the limit α_∞ equals zero, we will have shown that there is a unique (a, b) corresponding to $d \equiv 1$. In this case, a_0 is also uniquely determined via

$$a(z)b(z^{-1}) + a(z^{-1})b(z) = 1.$$

This can be derived from Lemma A.1 but also from an analysis of the nonlinear dynamical system introduced in Section II. To this end, consider an arbitrary α with $(\alpha, \gamma) \in \mathcal{P}_n(\gamma)$ such that $g(a_0, \alpha) = (a_0, 0)$. The corresponding infinite Schur sequence $\{\gamma_0, \gamma_1, \gamma_2, \cdots\}$ determines α_∞ via (2.17) as

$$\alpha_\infty = \lim_{t \to \infty} \prod_{\tau=0}^{t} A(\gamma(\tau))\alpha$$

from which we have $\alpha_{\infty n} = (1/r_\infty)\alpha_n$. Therefore, if $\alpha_{\infty n} = 0$, then $\alpha_n = 0$ and so

$$\alpha_{\infty, n-1} = \frac{r_1}{r_\infty}\alpha_{n-1}.$$

Proceeding in this manner we show that $\alpha_\infty = 0$ implies that $\alpha = 0$, concluding the proof of the lemma. $\square$

This concludes the proof that

$$f\colon \mathbf{R}_+ \times \mathcal{P}_n(\gamma) \to \mathcal{D}_n$$

is one-one and onto with an everywhere invertible Jacobian. By the implicit function theorem, the inverse function

$$f^{-1}\colon \mathcal{D}_n \to \mathbf{R}_+ \times \mathcal{P}_n(\gamma)$$

is analytic, since f is analytic, thus proving Corollary 2.3.

Proof of Corollary 2.4: From the commutative diagram introduced in Section II, it now follows that the function g, defined as

$$g = h^{-1} \circ f$$

is a diffeomorphism and hence a bijection, since h is a diffeomorphism (see Section III) and f is a diffeomorphism. This establishes the main claim in Corollary 2.4. Finally, the assertion concerning ρ follows from (2.28). $\square$

Remark 5.3: It is clear that, in general, degree theory cannot be used to enumerate solutions to (3.8) since det $\mathrm{Jac}_x\,(F)$ can assume either positive or negative values. One well-known exceptionally tractable case is degree theory for complex polynomials, for which the degree equals the algebraic degree of the polynomial. Indeed, in sharp contrast to the situation for real polynomials, the Cauchy–Riemann equations imply that the Jacobian determinant of a complex analytic function can only assume positive values. Quite remarkably, a similar situation interrelating algebra and analysis prevails here: Positivity of the covariance sequence in fact implies a similar positivity condition on the Jacobian determinant. This nontrivial fact underlies our proof of uniqueness and follows directly from the transversality lemma, which itself was shown using complex analytic methods.

Remark 5.4: One of the referees pointed out that there is a simple, more direct complex analysis argument which proves uniqueness. This argument is quite similar to the proof of the transversality lemma, but does not require degree theory to show that f is one-to-one.

Remark 5.5: The actual degree-theoretic proof of Georgiou's conjecture can also be presented more succinctly, as a direct consequence of the transversality lemma. Indeed, it reposes only on verifying that the Jacobian never vanishes on $\mathcal{P}_n$ and then computing the degree at the maximum entropy filter (see also [30]). On the other hand, our proof of Corollaries 2.2–2.4 shows much more. Aside from also providing a self-contained proof of existence (surjectivity of f), we are able to address well posedness of the rational covariance extension problem (locally) by proving that the Jacobian is everywhere invertible and (globally) by verifying properness of the map f. Finally, we observe that a more involved differential analysis is implicit in an analysis of such robustness issues, since even locally well posedness will imply invertibility of the Jacobian (see e.g., [46]).

APPENDIX

In this appendix we prove that f is a proper map (see also [30]). To this end, it will be useful to have an explicit description of f in (a_0, α, γ)-coordinates. For this reason we cite a lemma from [16], trivially modified for our purposes.

Lemma A.1: Suppose

$$d(z, z^{-1}) = d_0 + d_1(z + z^{-1}) + \cdots + d_n(z^n + z^{-n})$$

is a pseudopolynomial which satisfies

$$f(a_0, a, b) = a_0^2[a(z)b(z^{-1}) + a(z^{-1})b(z)] = 2d(z, z^{-1}). \tag{A.1}$$

If (α, γ) corresponds to (a, b), then

$$d_0 = a_0^2(\alpha_n^2 + r_1\alpha_{n-1}^2 + \cdots + r_n) \tag{A.2}$$

where $r_1, r_2, \cdots, r_n$ are defined by (2.8), and $d_i := d_i^{(n)}(a_0, \alpha, \gamma)$ for $i = 1, 2, \cdots, n$, where $d_i^{(n)}$ is determined recursively by

$$d_1^{(1)}(a_0, \alpha_1, \gamma_0) = a_0^2\alpha_1$$

$$\begin{aligned} d_i^{(k)}&(a_0, \alpha_1, \cdots, \alpha_k, \gamma_0, \cdots, \gamma_{k-1}) \\ &= (1 - \gamma_0^2)d_i^{(k-1)}(a_0, \alpha_1, \cdots, \alpha_{k-1}, \gamma_1, \cdots, \gamma_{k-1}) \\ &\quad + a_0^2\alpha_k \sum_{j=1}^{k} \alpha_{k-j}\pi_{j,j-i}, \quad \text{for} \quad i = 1, 2, \ldots, k-1 \end{aligned}$$

$$d_k^{(k)}(a_0, \alpha_1, \cdots, \alpha_k, \gamma_0, \cdots, \gamma_{k-1}) = a_0^2\alpha_k$$

where $\alpha_0 = 1$ and $\{\pi_{jl}\}$ are the coefficients of the polynomials

$$\pi_j(z) = z^j + \pi_{j1}z^{j-1} + \cdots + \pi_{jj}$$

generated by the polynomial recursion

$$\begin{aligned} \pi_{t+1}(z) &= (1 + z)\pi_t(z) + (\gamma_t\gamma_{t-1} - 1)z\pi_{t-1}(z); \\ \pi_0 &= 1, \quad \pi_1(z) = z \end{aligned}$$

and $\pi_{ji} = 0$ for $i > j$.

We can now proceed with our proof.

Lemma A.2: The smooth map (5.1) is proper.

Proof: We first note that

$$\partial(\mathbf{R}_+ \times \mathcal{P}_n(\gamma)) = \left(\{0\} \times \overline{\mathcal{P}_n(\gamma)}\right) \cup (\mathbf{R}_+ \times \partial\mathcal{P}_n(\gamma)).$$

The boundary $\partial\mathcal{P}_n(\gamma)$ of $\overline{\mathcal{P}_n(\gamma)}$ consists of pairs (a, b) for which (2.3) is positive real but not strictly positive real. Similarly, $\partial\mathcal{D}_n$ consists of those pseudopolynomials which are nonnegative on the unit circle and have at least one zero there (including the zero pseudopolynomial). From these facts it follows that

$$f(\partial(\mathbf{R}_+ \times \mathcal{P}_n(\gamma))) \subset \partial\mathcal{D}_n. \tag{A.3}$$

Now suppose $K \subset \mathcal{D}_n$ is a compact subset. We wish to show that $f^{-1}(K)$ is closed and bounded in $\mathbf{R} \times \mathbf{R}^n$. Denote by $\overline{f}$ the continuous extension of f to $\overline{\mathbf{R}_+ \times \mathcal{P}_n(\gamma)}$ defined via

$$\overline{f}(a_0, a, b) = a_0^2 S(a)b = a_0^2[a(z)b(z^{-1}) + a(z^{-1})b(z)].$$

The set $\overline{f}^{-1}(K)$ is closed, but, in the light of (A.3), $f^{-1}(K) = \overline{f}^{-1}(K)$. Finally, boundedness of $f^{-1}(K)$ follows from two observations. For a point $(a_0, a, b) \in f^{-1}(K)$, a and b are monic Schur polynomials and therefore have bounded coefficients. As for a_0, the constant term of the pseudopolynomials in K achieve a maximum which provides a bound on a_0^2 by virtue of (A.2). □

ACKNOWLEDGMENT

The authors would like to thank Profs. I. Gohberg and V. A. Yakubovich for their interest in this research program and for stimulating discussions and technical advice. They would also like to thank the editor and referees for several very good suggestions and for their time and effort in helping to improve this paper.

REFERENCES

[1] N. I. Akhiezer, *The Classical Moment Problem.* Hafner, 1965.
[2] A. C. Antoulas, J. A. Ball, J. Kang, and J. C. Willems, "On the solution of minimal rational interpolation problem," *Linear Algebra and Its Applications,* vol. 50, pp. 277–319, 1983.
[3] A. C. Antoulas and B. D. O. Anderson, "On the problem of stable rational interpolation," *Linear Algebra and Its Applications,* vol. 122, pp. 301–229, 1989.
[4] M. Aoki, *State Space Modeling of Time Series.* New York: Springer-Verlag, 1987.
[5] D. Z. Arov and M. G. Krein, "Problem of search of minimum of entropy in indeterminate extension problems," *Functional Anal. Appl.,* vol. 15, pp. 123–126, 1981.
[6] J. A. Ball, I. Gohberg, and L. Rodman, *Interpolation of Rational Matrix Functions.* Boston, MA: Birkhäuser Verlag, 1990.

[7] A. Benveniste and C. Chaure, "AR and ARMA identification algorithms of Levinson type: An identification approach," *IEEE Trans. Automat. Contr.* vol. AC-26, pp. 1243–1261, 1981.
[8] A. Bultheel, "Convergence of Schur parameters and transmission zeros of a meromorphic spectrum," *Modelling, Identification and Robust Control,* C. I. Byrnes and A. Lindquist, Eds. Amsterdam: North-Holland, 1986, pp. 281–296.
[9] C. I. Byrnes and A. Lindquist, "The stability and instability of partial realizations," *Syst. Contr. Lett.*, vol. 2, pp. 99–105, 1982.
[10] ———, "On the geometry of the Kimura–Georgiou parameterization of modeling filter," *Int. J. Contr.*, vol. 50, pp. 2301–2312, 1989.
[11] ———, "Toward a solution of the minimal partial stochastic realization problem," *Comptes Rendus Acad. Sci. Paris,* t. 319, Série I, 1994, pp. 1231–1236.
[12] ———, "Some recent advances on the rational covariance extension problem," *Proc. IEEE European Workshop on Computer-Intensive Methods in Control and Signal Processing,* Sep. 1994.
[13] ———, "On the partial stochastic realization problem," to appear.
[14] C. I. Byrnes, A. Lindquist, and T. McGregor, "Predictability and unpredictability in Kalman filtering," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 563–579, 1991.
[15] C. I. Byrnes, A. Lindquist, and Y. Zhou, "Stable, unstable and center manifolds for fast filtering algorithms," *Modeling, Estimation and Control of Systems with Uncertainty,* G. B. Di Masi, A. Gombani, and A. Kurzhanski, Eds. Boston, MA: Birkhauser, 1991.
[16] ———, "On the nonlinear dynamics of fast filtering algorithms," *SIAM J. Contr. Optim.*, vol. 32, pp. 744–789, 1994.
[17] W. A. Boothby, *A Introduction to Differentiable Manifolds and Riemannian Geometry.* New York: Academic, 1975.
[18] J. A. Cadzow, "Spectral estimation: An overdetermined rational model equation approach," *Proc. IEEE,* vol. 70, pp. 907–939, 1982.
[19] C. Carathéodory, "Über den variabilitätsbereich der koeffizienten von potenzreihen, die gegebene werte nicht annehmen," *Math. Ann.*, vol. 64, pp. 95–115, 1907.
[20] ———, "Über den variabilitätsbereich der fourierschen konstanten von positiven harmonischen functionen, *Rend. di Palermo,* vol. 32, pp. 193–217, 1911.
[21] C. Chang and T. T. Georgiou, "The Schur algorithm: Connections with the difference Riccati equation, the Chandrasekhar algorithm, and square-root filtering," *Signal Processing, Scattering and Operator Theory, and Numerical Methods,* M. A. Kaashoek, J. H. van Schuppen, and A. C. M. Ran, Eds. Boston, MA: Birkhäuser, 1990, pp. 207–213.
[22] P. Delsarte, Y. Genin, Y. Kamp, and P. van Dooren, "Speech modeling and the trigonometric moment problem," *Philips J. Res.*, vol. 37, pp. 277–292, 1982.
[23] C. J. Demeure and C. T. Mullis, "The Euclid algorithm and the fast computation of cross-covariance and autocovariance sequences," *IEEE Trans. Acoustics, Speech, Signal Proces.*, vol. 37, pp. 545–552, 1989.
[24] P. Dewilde and H. Dym, "Schur recursions, error formulas, and convergence of rational estimators for stationary stochastic sequences," *IEEE Trans. Inform. Theory,* vol. IT-17, pp. 446–461, 1981.
[25] ———, "Lossless chain scattering matrices and optimum linear prediction: The vector case," *Circuit Theory Appl.*, vol. 9, pp. 135–175, 1981.
[26] H. Dym and I. Gohberg, "A maximum entropy principle for contractive interpolants," *J. Functional Analysis,* vol. 65, pp. 83–125, 1986.
[27] P. Faurre, M. Clerget, and F. Germain, *Opérateurs Rationnels Positifs,* Dunod, 1979.
[28] F. R. Gantmacher, *The Theory of Matrices.* New York: Chelsea, 1959.
[29] T. T. Georgiou, "Partial realization of covariance sequences," *CMST,* Univ. Florida, Gainesville, 1983.
[30] ———, "Realization of power spectra from partial covariance sequences," *IEEE Trans. Acoustics, Speech, Signal Proces.*, vol. ASSP-35, pp. 438–449, 1987.
[31] Y. L. Geronimus, *Orthogonal Polynomials.* New York: Consultants Bureau, 1961.
[32] W. B. Gragg and A. Lindquist, "On the partial realization problem," *Linear Algebra and Its Applications,* vol. 50, pp. 277–319, 1983.
[33] U. Grenander and G. Szegö, *Toeplitz Forms and their Applications.* California: Univ. California Press, 1958.
[34] S. Haykin, *Toeplitz Forms and Their Applications.* New York: Springer-Verlag, 1979.
[35] T. Kailath, "Signal processing applications for some moment problems," in *Moments in Mathematics,* H. Landau, Ed. Providence, RI: American Mathematical Society, 1987, vol. 37, pp. 71–109.
[36] R. E. Kalman, "Realization of covariance sequences," *Proc. Toeplitz Memorial Conference (1981),* Tel Aviv, Israel, 1981.
[37] ———, "On minimal partial realizations of a linear input/output map," in *Aspects of Network and System Theory,* R. E. Kalman and N. de Claris, Eds. New York: Holt, Reinhart and Winston, 1971, pp. 385–408.
[38] ———, "On partial realizations, transfer functions and canonical forms," *Acta Polytech. Scand.*, vol. MA-31, pp. 9–39, 1979.
[39] S. M. Kay and S. L. Marple, Jr., "Spectrum analysis–A modern perspective," *Proc. IEEE,* vol. 69, pp. 1380–1419, 1981.
[40] H. Kimura, "Positive partial realization of covariance sequences," *Modeling, Identification and Robust Control,* C. I. Byrnes and A. Lindquist, Eds. Amsterdam: North-Holland, 1987, pp. 499–513.
[41] A. Lindquist, "A new algorithm for optimal filtering of discrete-time stationary processes," *SIAM J. Contr.*, vol. 12, pp. 736–746, 1974.
[42] ———, "Some reduced-order non-Riccati equations for linear least-squares estimation: The stationary, single-output case," *Int. J. Contr.*, vol. 24, pp. 821–842, 1976.
[43] ———, "On Fredholm integral equations, Toeplitz equations and Kalman-Bucy filtering," *Applied Math. Optim.*, vol. 1, pp. 355–373, 1975.
[44] A. Lindquist and G. Picci, "On subspace method identification," *Proc. MTNS93,* 1993.
[45] ———, "Canonical correlation analysis, approximate covariance extension, and identification of stationary time series," *Automatica,* to appear.
[46] P. Linz, *Theoretical Numerical Analysis.* New York: Wiley, 1979.
[47] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE,* vol. 63, pp. 561–580, 1975.
[48] S. L. Marple, Jr., *Digital Spectral Analysis and Applications.* New York: Prentice-Hall, 1987.
[49] J. W. Milnor, *Topology from Differentiable Viewpoint.* Virginia: Univ. of Virginia Press.
[50] A. Rantzer, "Equivalence between stability of partial realizations and feedback stabilization—Applications to reduced order stabilization," *Linear Algebra and Its Applications,* vol. 124, pp. 641–653, 1989.
[51] J. Rissanen, "Recursive identification of linear systems," *SIAM J. Contr.*, vol. 9, pp. 420–430, 1971
[52] I. Schur, "On power series which are bounded in the interior of the unit circle I and II," *J. fur die reine und angewandte Mathematik,* vol. 148, pp. 122–145, 1918.
[53] J. T. Schwartz, *Nonlinear Functional Analysis.* New York: Courant Institute of Mathematical Sciences, 1965.
[54] I. R. Shafarevitch, *Basic Algebraic Geometry.* Heidelberg: Springer-Verlag, 1974.
[55] O. Toeplitz, "Über die fouriersche entwicklung positiver funktionen," *Rendiconti del Circolo Matematico di Palermo,* vol. 32, pp. 191–192, 1911.
[56] P. van Overschee and B. De Moor, "Subspace algorithms for stochastic identification problem," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 382–387, 1982.
[57] A. S. Willsky, *Digital Signal Processing and Control and Estimation Theory.* Cambridge, MA: MIT Press, 1979.

Christopher I. Byrnes (M'81–SM'86–F'89) received the B.S. degree from Manhattan College, Bronx, NY, in 1971 and the M.S. and Ph.D. degrees from the University of Massachusetts, Amherst, in 1973 and 1975, respectively.

He came to Washington University, St. Louis, MO, in 1989 as Chairman of and Professor in the Department of Systems Science and Mathematics. Since 1991, he has been Dean of the School of Engineering and Applied Science. Prior to that, he held positions at Arizona State University, Harvard, and the University of Utah and was also an Adjunct Professor of mathematical system theory at the Royal Institute of Technology (KTH), Stockholm, Sweden. He has also held visiting positions at Bremen, Groningen, Harvard, IIASA, Soviet Union, Austria, Germany, Kansas, KTH, Osaka, Paris-Dauphine, Rome La Sapienza, Stanford, and the University of Tokyo. His research interests include adaptive control, algebraic system theory, distributed parameter systems, linear multivariable control, nonlinear control, and the applications of nonlinear dynamics in control and estimation. He is the editor of 11 research volumes and the author of 150 technical articles.

Dr. Byrnes is a member of AAAS, AMS, and SIAM and was named a Case Centennial Scholar by Case Western Research University in 1980, a Fellow of the Japan Society for the Promotion of Science in 1986, and the Graduate School Distinguished Research Professor at Arizona State University in 1988.

Anders Lindquist (M'77–SM'86–F'89) was born in Lund, Sweden, in 1942. He received the Ph.D. degree from the Royal Institute of Technology, Stockholm, Sweden. In 1972 he was appointed a Docent of Optimization and Systems Theory there.

From 1972 to 1974, he held visiting positions at the University of Florida, Brown University, and State University of New York at Albany. In 1974 he became an Associate Professor and in 1980 a Professor at the University of Kentucky, where he remained until 1983. He is presently a Professor at the Royal Institute of Technology, where, in 1982, he was appointed to the Chair of Optimization and Systems Theory. He is also an Affiliate Professor at Washington University, St. Louis, MO. He has also held visiting positions at University of Padova, Italy, University of Arizona, USSR Academy of Sciences, Moscow, East China Normal University, Shanghai, and Technion, Haifa, Israel. From 1975–1976 he was a SIAM Visiting Lecturer. He is the author of many papers in the area of systems and control, especially stochastic control, filtering, stochastic systems theory, realization theory, and applications of nonlinear dynamics in estimation and control. He is also the editor of four research volumes.

Dr. Lindquist is a honorary member of the Hungarian Operations Research Society. He is a Communicating Editor of the *Journal of Mathematical Systems, Estimation, and Control* and, until 1993, was an Associate Editor of *Systems and Control Letters*. He also serves on the editorial boards of *Adaptive Control and Signal Processing* and of the book series *Systems and Control: Foundations and Applications* (Boston, MA: Birkhäuser) and *Progress in Systems and Control* (Boston, MA: Birkhäuser).

Sergei V. Gusev was born in Murmansk, USSR, in 1952. He received the Diploma degree in mathematics in 1974 and the Ph.D. degree in mathematical cybernetics in 1981 from the Leningrad (now St. Petersburg) State University.

Since 1974 he has been with the Department of Mathematics and Mechanics of St. Petersburg State University, where he is now a Research Associate Professor. His research interests include robust control, adaptive control, system identification, mechanical systems control, and robotics. He has published over 60 papers in this area.

Alexei S. Matveev received the Ph.D. degree from Leningrad State University in 1980.

He was a Research Fellow at Leningrad (now St. Petersburg) State University. Since 1987 he has been with the Department of Mathematics and Mechanics, where he is currently an Associate Professor. His current research interests include nonlinear control theory, theory of optimization, and operator theoretic methods.