TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 358, Number 3, Pages 965–987
S 0002-9947(04)03616-5
Article electronically published on December 9, 2004

GENERALIZED INTERPOLATION IN H^∞ WITH A COMPLEXITY CONSTRAINT

CHRISTOPHER I. BYRNES, TRYPHON T. GEORGIOU, ANDERS LINDQUIST,
AND ALEXANDER MEGRETSKI

Abstract. In a seminal paper, Sarason generalized some classical interpolation problems for H^∞ functions on the unit disc to problems concerning lifting onto H^2 of an operator T that is defined on $\mathcal{K} = H^2 \ominus \phi H^2$ (ϕ is an inner function) and commutes with the (compressed) shift S. In particular, he showed that interpolants (i.e., $f \in H^\infty$ such that $f(S) = T$) having norm equal to $\|T\|$ exist, and that in certain cases such an f is unique and can be expressed as a fraction $f = b/a$ with $a, b \in \mathcal{K}$. In this paper, we study interpolants that are such fractions of $\mathcal{K}$ functions and are bounded in norm by 1 (assuming that $\|T\| < 1$, in which case they always exist). We parameterize the collection of all such pairs $(a, b) \in \mathcal{K} \times \mathcal{K}$ and show that each interpolant of this type can be determined as the unique minimum of a convex functional. Our motivation stems from the relevance of classical interpolation to circuit theory, systems theory, and signal processing, where ϕ is typically a finite Blaschke product, and where the quotient representation is a physically meaningful complexity constraint.

1. Introduction

In 1967, Sarason published a seminal paper [29] which contained, among other results, a theorem generalizing classical interpolation problems in the class $H(\mathbb{D})$ of analytic functions on the open unit disc $\mathbb{D}$ to a problem concerning operators commuting with a shift operator on a certain Hilbert space. This work marked the beginning of a series of important developments in operator theory [31, 1, 3, 27].

In more detail, denote by $\mathbb{T}$ the unit circle, by $L^p(\mathbb{T})$ or, simply L^p, the usual Lebesgue (p-)integrable functions on $\mathbb{T}$, and by H^p the usual Hardy space of functions analytic on $\mathbb{D}$. If U denotes the shift operator in L^2 defined by $U : f(z) \to zf(z)$, then by Beurling's Theorem all U-invariant subspaces in H^2 have the form ϕH^2 for ϕ an inner function. For the remainder of this paper, we fix ϕ to be a nonconstant inner function. Denote by $\mathcal{K}$ the coinvariant subspace

$$\mathcal{K} := H^2 \ominus \phi H^2,$$

invariant under U^*, and by S the compressed shift

$$S = \mathrm{P}^{\mathcal{K}} U \mid_{\mathcal{K}}.$$

Received by the editors October 27, 2003 and, in revised form, January 21, 2004.
2000 *Mathematics Subject Classification.* Primary 47A57, 30E05; Secondary 46N10, 47N10, 93B15.
This research was supported in part by Institut Mittag-Leffler and by grants from AFOSR, NSF, VR, the Göran Gustafsson Foundation, and Southwestern Bell.

©2004 American Mathematical Society

For $f \in H^\infty$, $f(S)$ denotes the compression onto $\mathcal{K}$ of the multiplication by f in L^2 and, following [29], when an operator T on $\mathcal{K}$ can be written as $f(S)$ for a suitable $f \in H^\infty$, then we say that f interpolates T. The operators $f(S)$ are precisely those that commute with S. Sarason's Theorem deals with the converse. It states that, *if T is a bounded operator on $\mathcal{K}$ commuting with S, then there is a function $f \in H^\infty$ such that $T = f(S)$ with $\|f\|_\infty = \|T\|$*. He also shows that, if T has a maximal vector, then there is a unique such interpolant, which takes the form

$$f = \frac{b}{a}, \quad a, b \in \mathcal{K}, \tag{1.1}$$

and, in this case, $f/\|T\|$ is inner [29, Proposition 5.1].

Sarason also raised the question of describing, for $\|T\| \leq 1$, the class of all interpolants $f \in H^\infty$ with $\|f\|_\infty \leq 1$ [29, p. 190]. This question is now classical and has been answered in various forms of generality in the literature. Here we will focus on the subclass of interpolants satisfying (1.1). We refer to the representation (1.1) of an interpolant as a quotient of two functions in $\mathcal{K}$ as a complexity constraint. As so often happens, there is a two-way street between results in pure mathematics and problems in engineering and the sciences. In particular, the complexity constraint in Sarason's framework arises naturally in engineering, where interpolants give rise to a controller or a filter which meets certain performance criteria and design specifications. In some important engineering applications, the problem specifications dictate that ϕ be a finite Blasckhe product, in which case the complexity constraint requires that f be a rational function with degree bounded by the degree of ϕ. In these applications, f is viewed as the transfer function of a circuit to be designed, and the degree bounds the number of dynamical components required to realize the circuit, while the design specifications are encapsulated in the pair T, ϕ and a bound for $\|f\|_\infty$. In our earlier work [17, 18, 19, 4, 5, 6, 7, 8, 9, 20] which dealt with the case of finite Blaschke products, we have discovered that all rational interpolants of bounded degree can be conveniently parameterized by the roots of $1 - f^*f$ inside $\mathbb{D}$, where $f^*(z) = \overline{f(\frac{1}{\bar{z}})}$.

In this paper, we consider this question for arbitrary inner functions and for the case of commutants that are strict contractions, i.e., $\|T\| < 1$. We give a complete parameterization of those interpolants f, with $\|f\|_\infty \leq 1$, which are quotients of two functions in $\mathcal{K}$. Indeed, suppose $f = b/a$, with $a, b \in \mathcal{K}$, interpolates T and satisfies $\|f\|_\infty \leq 1$. It follows that $b = Ta$, but, in contrast to the case treated by Sarason, $f/\|T\|$ is not an inner function in general. The function

$$\Psi(z) := |a(z)|^2 - |b(z)|^2$$

is nonnegative and not identically zero on $\mathbb{T}$. We observe that $|a|^2$ and $|b|^2$, and hence Ψ, belong to the (closed) subspace $M \subset L^1_{\mathbb{R}}(\mathbb{T})$ of real-valued functions spanned by $\{\mathrm{Re}(g\bar{h}) \mid g, h \in \mathcal{K}\}$ and hence, in particular, to the subset $\mathcal{Q}$ of nonnegative functions in M. It can be shown that $\Psi \in \mathcal{Q}$ if and only if $\Psi = |\sigma|^2$ for some outer $\sigma \in \mathcal{K}$ (Proposition 9). In fact, we may choose a unique such σ in $\mathcal{K}_0$, the subset of all outer functions in $\mathcal{K}$ which are positive at the origin.

The function σ determined from Ψ by such an interpolant f is on one hand a measure of how far f is from being an inner function and on the other hand determines f uniquely. Moreover, it gives a complete parameterization of all such interpolants by $\mathcal{K}_0$. This is made precise in the statement of our first main result.

Theorem 1. *Let T be an operator in $\mathcal{K}$ that commutes with S and has norm $\|T\| < 1$, and let σ be an arbitrary function in $\mathcal{K}_0$. Then there exists a unique pair of elements $(a, b) \in \mathcal{K}_0 \times \mathcal{K}$ such that*

(i) $f = b/a \in H^\infty$ *with* $\|f\|_\infty \leq 1$,
(ii) $f(S) = T$, *and*
(iii) $|a|^2 - |b|^2 = |\sigma|^2$ *a.e. on* $\mathbb{T}$.

Conversely, any pair $(a, b) \in \mathcal{K}_0 \times \mathcal{K}$ satisfying (i) *and* (ii) *determines, via* (iii), *a unique $\sigma \in \mathcal{K}_0$.*

Sarason's Theorem [29] provides us with the existence of a function $w \in H^\infty$ such that $w(S) = T$ and $\|w\|_\infty < 1$, in terms of which the interpolation condition (ii) can be written

$$f = w + \phi v, \quad v \in H^\infty. \tag{1.2}$$

When such an interpolant satisfies condition (i), it is said to belong to the Schur class $\mathcal{S}$. A theorem analogous to Theorem 1 can be formulated for the Carathèodory class $\mathcal{C}$ of functions in $H(\mathbb{D})$ with a nonnegative real part (or, more precisely, in the Smirnov class N^+). In fact, it is well known that the (involutory) linear fractional transformation

$$\mathcal{S} \to \mathcal{C}: \quad f \mapsto \varphi = \frac{1-f}{1+f} \tag{1.3}$$

is a bijective correspondence between $\mathcal{C}$ and $\mathcal{S}$, and, if $f = b/a$ as in Theorem 1, then

$$\begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix}, \tag{1.4}$$

provides a corresponding representation $\varphi = \beta/\alpha$, for which $\mathrm{Re}\{\bar{\alpha}\beta\} = \Psi$.

Theorem 2. *Let $\mathcal{T}$ be a bounded operator on $\mathcal{K}$ that commutes with S and satisfies $\mathrm{Re}\,\mathcal{T} > 0$, i.e., $\frac{1}{2}\langle x, (\mathcal{T} + \mathcal{T}^*)x\rangle \geq \varepsilon\|x\|^2$ for some $\varepsilon > 0$, and let σ be an arbitrary function in $\mathcal{K}_0$. Then, there exists a unique pair of outer functions $(\alpha, \beta) \in \mathcal{K}_0 \times \mathcal{K}$ such that*

(i) $\varphi = \beta/\alpha \in \mathcal{C}$,
(ii) $\varphi = c + \phi v$ *for some* $v \in H(\mathbb{D})$ *and any* $c \in H^\infty$ *such that* $c(S) = \mathcal{T}$, *and*
(iii) $\mathrm{Re}\{\bar{\alpha}\beta\} = |\sigma|^2$ *a.e. on* $\mathbb{T}$.

Conversely, any $(\alpha, \beta) \in \mathcal{K}_0 \times \mathcal{K}$ satisfying (i) *and* (ii) *determines via* (iii), *a unique $\sigma \in \mathcal{K}_0$.*

The interpolation data in the two theorems relate via

$$\mathcal{T} = (I - T)(I + T)^{-1} \tag{1.5}$$

which sets up a bijective correspondence between contractive and positive operators.

Theorems 1 and 2 extend the earlier work [18, 19, 4, 5, 20] to the case of a general inner function ϕ and characterize all interpolants f that can be expressed as a fraction of two functions in $\mathcal{K}$. The ideas of proof used in [18, 19, 4, 5, 8, 20] are topological in nature and are not directly extendable to infinite dimensions without the very strong assumption that T is compact. A constructive method of proof in the subsequent work [6, 7, 9, 11, 21] uses a nonlinear convex optimization approach inspired by entropy-theoretic methods. Indeed, in the next section we shall give an intrinsic derivation of one of these convex optimization schemes, by reinterpreting the generalized interpolation problem with the complexity constraint (1.1) in the context of differential forms on a convex subset of M.

The proofs of Theorem 1 and Theorem 2 will be given in Section 3. In Section 2 we develop a circle of ideas that lead to our method of proof. In Section 4 we show that the problem of generalized interpolation with the complexity constraint (1.1) is well-posed, in the sense of Hadamard, with respect to natural choices of topologies. Finally, in Section 5 we discuss connections to Carathéodory extension and Nevanlinna-Pick interpolation and provide an example, with ϕ being a singular inner function, motivated by systems theory. In this example, condition (iii) in Theorem 1 has a natural physical interpretation as energy dissipated in a passive system.

We would like to thank Harold Shapiro for many valuable discussions on various aspects of this work and Institut Mittag-Leffler for providing the ideal meeting place for finishing this paper.

2. Generalized interpolation, differential forms and optimization

We begin our analysis by restricting our attention to interpolants $f \in \mathcal{S}$ that, in the context of Theorem 1, are strictly contractive, i.e., $\|f\|_\infty < 1$. These are in bijective correspondence, via (1.3), with interpolants φ in the framework of Theorem 2 that belong to $\mathcal{C}_+ \cap H^\infty$, where $\mathcal{C}_+$ is the subclass of $\mathcal{C}$ of functions with real part bounded away from zero. The interpolant in Sarason's Theorem is of this type.

Starting out in the framework of Theorem 2, we note that, for $\varphi \in \mathcal{C}_+ \cap H^\infty$, condition (ii) of Theorem 2 can now be written as $\varphi(S) = \mathcal{T}$, and, in view of (i),

$$\Phi := \mathrm{Re}\{\varphi\} = \frac{|\sigma|^2}{|\alpha|^2} = \frac{\Psi}{Q} \in L^\infty_{\mathbb{R}}(\mathbb{T}), \tag{2.1}$$

where $Q := |\alpha|^2$ and $L^\infty_{\mathbb{R}}(\mathbb{T})$ is the space of real functions in $L^\infty(\mathbb{T})$.

Denote by $\mathcal{Q}_+$ the subset of $\mathcal{Q}$ consisting of those functions in M for which the essential infimum is positive, and suppose that $\Psi = |\sigma|^2 \in \mathcal{Q}_+$. From (2.1) it follows that $Q \in \mathcal{Q}_+$. To determine an interpolant φ from σ, it suffices to find $\alpha \in \mathcal{K}_0$, in terms of which $\beta \in \mathcal{K}$ is given by $\beta = \mathcal{T}\alpha$. The problem is thus reduced to finding Q, from which we can obtain α as its outer factor. Thus our next goal is to express all the constraints in terms of Q. We note that $\mathcal{Q}_+$ is a convex subset of M, a fact which will make calculus and optimization very applicable.

We shall temporarily relax the interpolation condition $\varphi(S) = \mathcal{T}$ by replacing it by $\Phi(S) = \mathrm{Re}\,\mathcal{T}$, where $\Phi(S) : \mathcal{K} \to \mathcal{K}$ is the bounded operator sending α to $\mathrm{P}^{\mathcal{K}}\Phi\alpha$, or, equivalently, by

$$\Phi(S) = C(S), \tag{2.2}$$

where C is given by

$$C := \mathrm{Re}\{c\} = \frac{1-|w|^2}{|1+w|^2} \in L^\infty_{\mathbb{R}}(\mathbb{T}) \tag{2.3}$$

with w is as in (1.2). In fact, $\varphi(S) = \mathcal{T} = c(S)$ implies that

(2.4)
$$\Phi(S)\alpha = \frac{1}{2}\mathrm{P}^{\mathcal{K}}(\varphi + \bar{\varphi})\alpha = \frac{1}{2}\left(\varphi(S) + \varphi(S)^*\right)\alpha = \frac{1}{2}\left(c(S) + c(S)^*\right)\alpha = C(S)\alpha$$

for all $\alpha \in \mathcal{K}$. In particular,

$$\langle \alpha, \Phi(S)\alpha \rangle = \langle \alpha, \mathrm{Re}(\mathcal{T})\alpha \rangle,$$

which is the Pick form.

We observe that, for any $V \in L^\infty_{\mathbb{R}}$, to say that $V(S) = 0$ is to say that

$$\int_{\mathbb{T}} g\bar{h}V\, dm = \langle h, Vg \rangle = \langle h, \mathrm{P}^{\mathcal{K}} Vg \rangle = 0 \quad \text{for all } g, h \in \mathcal{K},$$

where dm denotes the Lebesgue measure on $\mathbb{T}$. This is in turn equivalent to

$$\int_{\mathbb{T}} PV\, dm = 0 \quad \text{for all } P \in M \subset L^1_{\mathbb{R}}(\mathbb{T}).$$

Hence the space of annihilators $M^\perp$ consists of all $V \in L^\infty_{\mathbb{R}}$ such that $V(S) = 0$. In particular, the interpolation condition (2.2) holds if and only if $\Phi = C$ in $M^* := L^\infty_{\mathbb{R}}/M^\perp$.

With this in mind, we provisionally assume that Ψ is an L^∞ function and (following [10]; see also [9]) consider the 1-form

$$\omega = \int_{\mathbb{T}} \left(C - \frac{\Psi}{Q} \right) dQ\, dm \tag{2.5}$$

which is defined at points in the convex set $\mathcal{Q}_+ \subset M$. The tangent space to M at a point Q is canonically isomorphic to M, and the value $\omega(Q)(v)$ at the tangent vector $v \in M$ is given by

$$\omega(Q)(v) = \int_{\mathbb{T}} \left(C - \frac{\Psi}{Q} \right) v\, dm.$$

In particular, to say that ω vanishes at Q is to say that Ψ/Q interpolates C. We claim that this occurs at the unique minimum of a potential function. Indeed, fixing a base point Q_0 in $\mathcal{Q}_+$, we consider the integral

$$\mathbb{J}_\Psi(Q) = \int_{Q_0}^{Q} \omega$$

along the line from Q_0 to Q as a function of its upper limit. Suppose Q_n tends to Q in $\mathcal{Q}_+$ and consider the triangle bounded by traversing the line from Q_0 to Q, the line from Q to Q_n and the line from Q_n to Q_0. On this triangle,

$$d\omega = \int_{\mathbb{T}} \frac{\Psi}{Q^2} dQ \wedge dQ\, dm = 0,$$

so that the integral along the path is zero, by Green's Theorem. In particular, $\mathbb{J}_\Psi(Q_n)$ tends to $\mathbb{J}_\Psi(Q)$.

More generally, a similar argument shows that the integral is independent of the path. In particular, modulo a constant of integration we have

$$\mathbb{J}_\Psi(Q) = \int_{\mathbb{T}} CQdm - \int_{\mathbb{T}} \Psi \log Q dm. \tag{2.6}$$

Clearly, $\mathbb{J}_\Psi : \mathcal{Q} \to \bar{\mathbb{R}}$ is a strictly convex functional on its effective domain. The first term is a Pick form, which is required to be positive. In fact, since $Q \in \mathcal{Q}$, there is an $\alpha \in \mathcal{K}$ such that $Q = |\alpha|^2$, and hence

$$\int_{\mathbb{T}} CQdm = \langle \alpha, \mathrm{Re}(\mathcal{T})\alpha \rangle. \tag{2.7}$$

This derivation of $\mathbb{J}_\Psi$ is meant to underscore the fact that nonlinear convex optimization methods are intrinsic to the problem of generalized interpolation with the complexity constraint (1.1). Indeed, the assumption that Ψ be an L^∞ function is not needed for the rigorous analysis of this minimization problem, as stated in the following theorem.

Theorem 3. *Suppose that* $\|T\| < 1$ *and that* $f \in H^\infty$ *satisfies the conditions*

(i) $f = b/a \in H^\infty$ *with* $\|f\|_\infty < 1$ *and* $(a,b) \in \mathcal{K}_0 \times \mathcal{K}$,
(ii) $f(S) = T$, *and*
(iii) $\Psi := |a|^2 - |b|^2$ *satisfies*

$$\int_{\mathbb{T}} \Psi \log^+ \Psi \, dm < \infty.$$

Then the functional $\mathbb{J}_\Psi$ *has a unique minimum, and*

$$a = \sqrt{2}(I+T)^{-1}\alpha, \quad b = \sqrt{2}T(I+T)^{-1}\alpha \tag{2.8}$$

where $\alpha \in \mathcal{K}_0$ *is uniquely defined by*

$$|\alpha|^2 = \arg\min_{Q \in \mathcal{Q}} \mathbb{J}_\Psi. \tag{2.9}$$

Moreover, setting $\beta = (a-b)/\sqrt{2}$ *and* $\varphi = \beta/\alpha$, *we have* $\varphi \in \mathcal{C}$ *and*

$$\varphi(S) = (I-T)(I+T)^{-1}.$$

Proof. First we note that $\mathbb{J}_\Psi$ is a strictly convex functional in its effective domain. Hence, if a minimizer exists, then it is unique. Next, set $\alpha = (a+b)/\sqrt{2}$, $\beta = (a-b)/\sqrt{2}$, and $\hat{Q} = |\alpha|^2$. In view of condition (i), $\varphi = \beta/\alpha$ satisfies

$$\Phi := \frac{\Psi}{\hat{Q}} = \mathrm{Re}\{\varphi\} \in L^\infty(\mathbb{T}), \tag{2.10}$$

and hence $\Phi(S) = C(S)$ by the calculation (2.4), or, equivalently, $\Phi = C$ in M^*. Moreover, because of (iii) and (2.10), $\mathbb{J}_\Psi(\hat{Q}) < \infty$. Taking now a Newton quotient in the direction v, we have

$$d\mathbb{J}_\Psi(\hat{Q})(v) = \int_{\mathbb{T}} \left(C - \frac{\Psi}{\hat{Q}} \right) v \, dm = 0 \tag{2.11}$$

for all $v \in T_{\hat{Q}}\mathcal{Q} = M$, i.e., $\hat{Q}$ is a stationary point and hence the unique minimizer. Finally, observing that $b = Ta$, the rest follows from $\varphi = (1-f)(1+f)^{-1}$, (1.4) and (1.5). □

The convex optimization problem of Theorem 3 is the dual of the concave maximization problem to find a Φ in

$$\mathcal{F} := \{\Phi \in L^\infty_{\mathbb{R}}(\mathbb{T}) \mid \operatorname{ess\,inf} \Phi(e^{i\theta}) > 0\},$$

that maximizes the functional

$$\Phi \mapsto \int_{\mathbb{T}} \Psi \log \Phi \, dm \tag{2.12}$$

subject to $\Phi(S) = C(S)$. Indeed, setting up the Lagrangian for this constrained optimization problem, Q in (2.6) appears as a Lagrange multipier.

In fact, since the constraint $\Phi(S) = C(S)$ is equivalent to $\Phi = C$ on M^*, we may form the Lagrangian

$$L(\Phi, Q) = \int_{\mathbb{T}} \Psi \log \Phi \, dm + \int_{\mathbb{T}} Q(C - \Phi) dm,$$

where the Lagrange multiplier Q is a real-valued function in the predual M. We want to minimize the dual function $Q \mapsto \sup_{\phi \in \mathcal{F}} L(\Phi, Q)$, which can only take finite values for $Q \in \mathcal{Q}$. Therefore, without lack of generality, we may restrict Q to $\mathcal{Q}$.

The derivative

$$dL(\Phi, Q)(v) = \int_{\mathbb{T}} \left(\frac{\Psi}{\Phi} - Q \right) v \, dm$$

vanishes in all directions $v \in T_\Phi \mathcal{F}$ if $\Phi = \Psi/Q$. Inserting this into the Lagrangian, we obtain

$$L\left(\frac{\Psi}{Q}, Q \right) = \mathbb{J}_\Psi(Q) + \kappa,$$

where $\mathbb{J}_\Psi$ is defined by (2.6) and

$$\kappa := \int_{\mathbb{T}} \Psi(\log \Psi - 1) dm. \tag{2.13}$$

Assuming condition (iii) of Theorem 3, κ is guaranteed to be well defined. In fact, $0 \geq \int_{\mathbb{T}} \Psi \log^- \Psi \, dm > -\infty$, since $\Psi = |\sigma|^2$ with $\sigma \in H^2$. Now, it can be seen that the primal problem to maximize (2.12) over $\mathcal{F}$ subject to $\Phi(S) = C(S)$ can be reformulated as maximizing the concave functional

$$\mathbb{I}_\Psi(\varphi) = \int_{\mathbb{T}} \Psi \log \left(\operatorname{Re}\{\varphi\} \right) dm, \tag{2.14}$$

subject to the original interpolation condition $\varphi(S) = c(S)$.

Theorem 4. *Let $\mathcal{T}$ be a bounded operator on $\mathcal{K}$ that commutes with S and satisfies $\operatorname{Re} \mathcal{T} > 0$. Suppose $\varphi \in \mathcal{C}_+ \cap H^\infty$ satisfies $\varphi(S) = \mathcal{T}$ and $\varphi = \beta/\alpha$ with $(\alpha, \beta) \in \mathcal{K}_0 \times \mathcal{K}_0$, and set $\Psi = \operatorname{Re}\{\bar{\alpha}\beta\}$. Then $\mathbb{I}_\Psi : \mathcal{C}_+ \cap H^\infty \to \mathbb{R} \cup \{-\infty\}$ has a unique maximizer in the class of interpolants $\varphi(S) = \mathcal{T}$, and this maximizer is precisely φ.*

Proof. For notational convenience, let $\hat{\varphi} \in H^\infty$ be the function φ in the theorem, i.e., $\hat{\varphi}(S) = \mathcal{T}$ and $\hat{\varphi} = \beta/\alpha$, and set $\hat{Q} = |\alpha|^2$. Then

$$\hat{\Phi} := \operatorname{Re}\{\hat{\varphi}\} = \frac{\Psi}{\hat{Q}}$$

belongs to $L^\infty_{\mathbb{R}}(\mathbb{T})$ and satisfies $\hat{\Phi}(S) = C(S)$. Moreover, since $\hat{\varphi} \in \mathcal{C}_+ \cap H^\infty$, we have $\operatorname{Re}\{\hat{\varphi}\}$ bounded away from zero and infinity, and therefore $\mathbb{I}_\Psi(\hat{\varphi})$ is finite. Since $\Phi \mapsto L(\Phi, \hat{Q})$ is strictly concave, and since

$$dL(\hat{\Phi}, \hat{Q})(v) = \int_{\mathbb{T}} \left(\frac{\Psi}{\hat{\Phi}} - \hat{Q} \right) v \, dm = 0, \quad \text{for all } v \in T_{\hat{\Phi}} \mathcal{F},$$

it follows that

$$L(\Phi, \hat{Q}) \leq L(\hat{\Phi}, \hat{Q}), \quad \text{for all } \Phi \in \mathcal{F}, \tag{2.15}$$

with equality if and only if $\Phi = \hat{\Phi}$. However, since $\hat{\Phi} = C$ in M^*, $L(\hat{\Phi}, \hat{Q}) = \mathbb{I}_\Psi(\hat{\Phi})$. Moreover, for each $\varphi \in H^\infty \cap \mathcal{C}$ satisfying $\varphi(S) = \mathcal{T}$, it holds that $\Phi := \operatorname{Re}\{\varphi\}$ satisfies $\Phi = C$ in M^*. Hence

$$\mathbb{I}_\Psi(\varphi) \leq \mathbb{I}_\Psi(\hat{\varphi}), \quad \text{for all } \varphi \in H^\infty \cap \mathcal{C} \text{ such that } \varphi(S) = \mathcal{T},$$

with equality if and only if $\mathrm{Re}\{\varphi\} = \mathrm{Re}\{\hat{\varphi}\}$. In view of the Riesz-Herglotz representation

$$\varphi(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{i\theta}-z}{e^{i\theta}+z}\mathrm{Re}\{\varphi(e^{i\theta})\}d\theta + i\gamma, \quad \gamma \in \mathbb{R},$$

$\mathrm{Re}\{\varphi\} = \mathrm{Re}\{\hat{\varphi}\}$ if and only if φ and $\hat{\varphi}$ differ by an imaginary constant. However, this constant must be zero. Indeed, if $\varphi(S) = \hat{\varphi}(S)$, then $\bar{\phi}(\varphi - \hat{\varphi}) \in H^\infty$, which implies that $\varphi = \hat{\varphi}$, as ϕ is nonconstant. Consequently, $\hat{\varphi}$ is the unique solution to the optimization problem of Theorem 4, as claimed. □

Translating Theorem 4 to the Schur setting, we see that a function $f \in \mathcal{S}$ satisfying conditions (i) and (ii) of Theorem 3, is the unique maximizer of

$$f \mapsto \mathbb{I}_\Psi\left(\frac{1-f}{1+f}\right) = \mathbb{K}_\Psi(f) - \rho(f) \tag{2.16}$$

in the class of functions $f \in \mathcal{S}$ satisfying $f(S) = T$, where $\Psi = |a|^2 - |b|^2$,

$$\mathbb{K}_\Psi(f) = \int_{\mathbb{T}} \Psi \log\left(1 - |f|^2\right) dm, \tag{2.17}$$

and

$$\rho(f) = \int_{\mathbb{T}} \Psi \log(|1+f|^2) dm.$$

Now the derivative of ρ at any point f and in any direction ϕv in which the interpolation conditions are preserved is

$$d\rho(f)(v) = 2\mathrm{Re}\int_{\mathbb{T}} \Psi \frac{\phi v}{1+f} dm = 0,$$

since $\Psi \in \bar{\phi} H_0^1$ (Lemma 10) and $v(1+f)^{-1} \in H^\infty$. Moreover, it is easy to check that $\mathbb{K}_\Psi$ is strictly concave, and it therefore follows that $\mathbb{K}_\Psi$ has a maximum at the same point as (2.16). Consequently, we have shown that a Schur function satisfying (i) and (ii) in Theorem 3 is the unique maximizer of $\mathbb{K}_\Psi$ as well. In fact, we shall establish next that, for *any* $\sigma \in \mathcal{K}_0$, the corresponding interpolant f satisfying the conditions of Theorem 1 can be obtained as the unique maximizer of $\mathbb{K}_\Psi$ for $\Psi = |\sigma|^2$.

Theorem 5. *Let T be an operator in $\mathcal{K}$ that commutes with S and has norm $\|T\| < 1$, and let σ be an arbitrary function in $\mathcal{K}_0$. Then, setting $\Psi = |\sigma|^2$, the functional $\mathbb{K}_\Psi$ has a unique maximizer in the class of functions satisfying $f(S) = T$, and this maximizer is precisely the unique $f \in \mathcal{S}$ satisfying conditions* (i), (ii) *and* (iii) *in Theorem* 1.

3. Proofs of Theorems 1, 2 and 5

Optimization is the main ingredient in the proof of Theorem 1. Therefore, we begin by establishing that the functional (2.17) has a unique maximizer in the class of functions satisfying $f(S) = T$ or, equivalently, satisfying $f = w + \phi v$ for some $v \in H^\infty$, where $w \in H^\infty$ is an arbitrary function satisfying $w(S) = T$ and $\|w\| < 1$. Sarason [29] ensures the existence of such a function and that

$$X = \{v \in H^\infty \mid \|w + \phi v\| \leq 1\} \tag{3.1}$$

is nonempty. Elements in X are in bijective correspondence via (1.2) with contractive interpolants of T, i.e., with elements of the set

$$Y = \{f \in H^\infty \mid \|f\| \le 1 \text{ and } f(S) = T\}. \tag{3.2}$$

Setting $F(v) := \mathbb{K}_\Psi(w + \phi v)$ and $\Psi = |\sigma|^2$, we obtain a strictly concave functional $F:\ X \to [-\infty, 0)$ given by

$$F(v) = \int_{\mathbb{T}} |\sigma|^2 \log(1 - |w + \phi v|^2) dm \tag{3.3}$$

with the value $-\infty$ if the Lebesgue measure of the set $\{z \in \mathbb{T} \mid |w + \phi v| = 1\}$ is positive.

Main Lemma 6. *The functional* (3.3) *has a unique maximum on* X*.*

More specifically, we will show that this functional has a unique maximum which yields the unique pair $(a, b) \in \mathcal{K}_0 \times \mathcal{K}$ corresponding to a given σ as claimed in Theorem 1.

3.1. **Proof of the Main Lemma.** We begin with the existence part, for which we need the following lemma.

Lemma 7. *There exists a set* $\mathcal{L}$ *of continuous affine functionals on* H^∞ *of the form*

$$\lambda(v) = \lambda_0 + \int_{\mathbb{T}} \operatorname{Re}(hv) dm \tag{3.4}$$

with $\lambda_0 \in \mathbb{R}$*,* $h \in L^1$*, such that*

$$F(v) = \inf_{\lambda \in \mathcal{L}} \lambda(v) \quad \textit{for every } v \in X. \tag{3.5}$$

Proof. For all $a, x > 0$, $\log(x) \le \log(a) + \frac{x-a}{a}$. Hence for all $|s| \le 1$, $|z| \le 1$, and $\varepsilon > 0$ we have

$$\begin{aligned} \log(1 + \varepsilon - |s|^2) &\le \log(1 + \varepsilon - |z|^2) + \frac{|z|^2 - |s|^2}{1 + \varepsilon - |z|^2} \\ &\le \log(1 + \varepsilon - |z|^2) + 2\frac{|z|^2 - \operatorname{Re}(\bar{z}s)}{1 + \varepsilon - |z|^2} \end{aligned}$$

with equality whenever $z = s$. Consequently,

$$\begin{aligned} F(v) &= \inf_{\varepsilon > 0} \int_{\mathbb{T}} |\sigma|^2 \log(1 + \varepsilon - |w + \phi v|^2) dm \\ &\le \inf_{u \in X} \inf_{\varepsilon > 0} \int_{\mathbb{T}} |\sigma|^2 \Bigg(\log(1 + \varepsilon - |w + \phi u|^2) \\ &\qquad + 2\frac{|w + \phi u|^2 - \operatorname{Re}(\overline{w + \phi u})(w + \phi v)}{1 + \varepsilon - |w + \phi u|^2} \Bigg) dm \end{aligned}$$

where the outer infimum is achieved at $u = v$. Therefore, if $\mathcal{L}$ is the class (3.4) of affine functionals on H^∞ defined via

$$\lambda_0 = \int_{\mathbb{T}} |\sigma|^2 \left(\log(1 + \varepsilon - |w + \phi u|^2) + 2\frac{|w + \phi u|^2 - \operatorname{Re}(\overline{w + \phi u})w}{1 + \varepsilon - |w + \phi u|^2} \right) dm \in \mathbb{R}$$

and

$$h = -2|\sigma|^2 \frac{(\overline{w + \phi u})\phi}{1 + \varepsilon - |w + \phi u|^2} \in L^1$$

for all $\varepsilon > 0$ and $u \in X$, then the representation (3.5) holds. $\square$

We now show that there exists a $\hat{v} \in X$ such that

$$F(\hat{v}) \geq F(v) \text{ for all } v \in X.$$

To this end, we first note that X is sequentially weak* compact. To see this, take a sequence $\left(v_k\right)_{k=1}^{\infty}$ in X, and consider the corresponding sequence $\left(f_k\right)_{k=1}^{\infty}$ defined by $f_k = w + \phi v_k$. Since the unit ball in H^∞ is sequentially weak* compact, there exists a subsequence $\left(f_j\right)_{j=1}^{\infty}$ converging to a weak* limit $\hat{f} \in H^\infty$ with norm $\|\hat{f}\| \leq 1$. Then, for any $h \in L^1$, $\int h(\hat{f} - f_j)dm \to 0$ as $j \to \infty$, and, in particular, this is true for $h = \bar{\phi}u$ and any $u \in H_0^1$. But in this case, the integral $\int \bar{\phi}u(\hat{f} - f_j)dm = \int \bar{\phi}u(\hat{f} - w)dm$ is independent of j and hence identically zero. From this fact it follows readily that $\hat{f}$ is of the form $w + \phi\hat{v}$ for some $\hat{v} \in H^\infty$ which then is the weak* limit of the sequence $\left(v_j\right)_{k=1}^{\infty}$ and belongs to X.

Now, let

$$\rho := \sup_{v \in X} F(v) \geq F(0) > -\infty,$$

and note that $\rho \leq 0$. Then, let $\left(v_k\right)_{k=1}^{\infty}$ be a (maximizing) sequence in X such that $F(v_k) \to \rho$ as $k \to \infty$, and let $\left(v_j\right)_{j=1}^{\infty}$ be a subsequence having a weak* limit $\hat{v} \in X$. Since the predual of H^∞ is L^1/H_0^1,

$$\int_{\mathbb{T}} h v_j \, dm \to \int_{\mathbb{T}} h v \, dm$$

for all $h \in L_1$, and hence

$$\lambda(\hat{v}) = \lim_{j \to \infty} \lambda(v_j)$$

for all $\lambda \in \mathfrak{L}$. From Lemma 7 we also have that

$$\lim_{j \to \infty} \lambda(v_j) \geq \liminf_{j \to \infty} F(v_j) = \rho$$

for any $\lambda \in \mathfrak{L}$. Therefore, in view of (3.5), we conclude that $F(\hat{v}) \geq \sup_{v \in X} F(v)$, as claimed. This establishes the existence part of the theorem.

Next, we turn to the question of uniqueness of the maximum of $F(v)$.

Lemma 8. *The functional* (3.3) *is strictly concave on its effective domain.*

Proof. Since, for each $z \in \mathbb{T}$, we have that $v(z) \mapsto 1 - |w(z) + \phi(z)v(z)|^2$ is strictly concave for $|w(z) + \phi(z)v(z)| \leq 1$, and since the log function is strictly concave and monotonically increasing on $(0, \infty)$, it follows that $v(z) \mapsto \log(1 - |w(z) + \phi(z)v(z)|^2)$ is strictly convex, and hence so is F on its effective domain, i.e., where $F(v) > -\infty$. $\square$

Consequently, the maximizer, whose existence we have established, is also unique. This concludes the proof of the main lemma.

3.2. **Factorization in $\mathcal{K}$ of nonnegative functions in M.**

Proposition 9. *Let $\mathcal{Q}$ be the set of nonnegative functions in M. Then, $Q \in \mathcal{Q}$ if and only if there is an outer function $a \in \mathcal{K}$ such $Q = |a|^2$.*

Proof. Clearly, if $Q = |a|^2$ with $a \in \mathcal{K}$, then $Q \in \mathcal{Q}$. For the converse, we first note that for $g, h \in \mathcal{K}$, we have $g\bar{h} \in \bar{\phi}H_0^1$ since $\bar{\mathcal{K}} \subset \bar{\phi}H_0^2$. Hence $\phi M \subset H_0^1$. Now, since $Q \in \mathcal{Q} \subset \bar{\phi}H_0^1$,

$$\int_{\mathbb{T}} \log Q dm > -\infty \tag{3.6}$$

(see, e.g., [14, p. 17]), and hence $Q = |a|^2$ for some $a \in H^2$; see, e.g., [23, p. 53]. Then, since $\phi Q \in H_0^1$, the rest of the proposition follows from the following lemma. □

Lemma 10. *If $g \in \mathcal{K}$, then $\phi|g|^2 \in H_0^1$. Conversely, if $g \in H^2$ and $\phi|g|^2 \in H_0^1$, then the outer factor of g is in $\mathcal{K}$.*

Proof. Since $\mathcal{K} = (\phi\bar{H}_0^2) \cap H^2$, we have

$$g \in \mathcal{K} \Leftrightarrow \phi\bar{g} \in H_0^2 \text{ and } g \in H^2. \tag{3.7}$$

Then, it immediately follows that $\phi|g|^2 \in H_0^1$. For the converse statement, the fact that $\phi|g|^2 \in H_0^1$ implies that $\phi|g_o|^2 \in H_0^1$, where g_o is the outer factor of g. Since therefore $\phi\bar{g}_o g_o \in H_0^1$ with g_o outer in H^2, it follows that $\phi\bar{g_o}$ belongs to the Smirnov class N^+. However, $\phi\bar{g_o}$ is also in $L^2(\mathbb{T})$, and hence $\phi\bar{g_o} \in H_0^2$ (see, e.g., [14, Theorem 2.11]). Consequently, it follows from (3.7) that g_o is in $\mathcal{K}$. □

Corollary 11. $M \subset \bar{\phi}H_0^1 \cap \phi\bar{H}_0^1$.

3.3. **Proofs of Theorems 1 and 5.** We show that the maximizer of the functional F gives rise to an interpolant of the form claimed in the statement of Theorem 1. This follows from the stationarity conditions. Uniqueness follows from strict concavity. Recalling that $F(v) = \mathbb{K}_\Psi(w + \phi v)$, this also proves Theorem 5.

Lemma 12. *Let $f = w + \phi v$, where $v = \operatorname{argmax}(F)$. Then*

$$\frac{|\sigma|^2}{1 - |f|^2} \in L^1.$$

Proof. If $v = 0$, then $f = w$ and thus $\|f\| < 1$. Now suppose that $v \neq 0$, and consider, for $t \in (0, 1)$, the differentiable function $\psi \ : \ t \mapsto F((1-t)v)$ with derivative

$$\dot{\psi}(t) = \int_{\mathbb{T}} \varphi_t dm \quad \text{ where } \varphi_t := |\sigma|^2 \frac{\operatorname{Re}\{[\bar{f} + t(\bar{w} - \bar{f})][f - w]\}}{1 - |f + t(w - f)|^2}.$$

Since F has a maximum at v, $\dot{\psi}(t) \leq 0$ for t small. Choose a γ such that $\|w\| \leq \gamma < 1$, and let I_1 and I_2 be the subintervals of $\mathbb{T}$ where $|f| < \frac{1}{2}(1+\gamma)$ and $|f| \geq \frac{1}{2}(1+\gamma)$, respectively. On I_1, the denominator of φ_t is bounded away from zero, and hence

$$\left| \int_{I_1} \varphi_t \, dm \right| < \infty,$$

while on I_2 $\varphi_t > 0$ for sufficiently small t. Therefore, to say that $\frac{|\sigma|^2}{1-|f|^2} \notin L^1$ is to say that

$$\int_{I_2} \varphi_t \, dm \to +\infty,$$

contradicting the nonpositivity of $\dot{\psi}(t)$ for t small. □

Since σ is in H^2, we have $\int_{\mathbb{T}} \log |\sigma|^2 dm > -\infty$; see, e.g., [23, p. 53]. Therefore, since $1 - |f|^2$ is bounded,

$$\int_{\mathbb{T}} \log\left(\frac{|\sigma|^2}{1-|f|^2}\right) dm = \int_{\mathbb{T}} \log(|\sigma|^2) dm - \int_{\mathbb{T}} \log(1-|f|^2)\, dm > -\infty.$$

Hence, there is a unique outer function $a \in H^2$ with $a(0) > 0$ such that

$$|a|^2 = \frac{|\sigma|^2}{1-|f|^2}, \tag{3.8}$$

and, if we define $b := fa \in H^2$, a and b satisfy

$$f = \frac{b}{a}, \tag{3.9}$$

$$|\sigma|^2 = |a|^2 - |b|^2. \tag{3.10}$$

To show that $a, b \in \mathcal{K}$ we need the following lemma.

Lemma 13. *Let $f = w + \phi v$, where $v = \operatorname{argmax}(F)$. Then*

$$\frac{|\sigma|^2}{1-|f|^2} \bar{f}\phi \in H_0^1. \tag{3.11}$$

Proof. From Lemma 12 and the fact that $\int_{\mathbb{T}} \log |\sigma|^2 dm > -\infty$ we see that

$$\int_{\mathbb{T}} \log(1-|f|^2) dm > -\infty,$$

which implies that $1 - |f|^2 = |g|^2$ for some outer $g \in H^2$. For any $h \in H^\infty$ with norm $\|h\|_\infty \le 1$, we have $v + t\phi g^2 h \in X$ for $t \in (-\varepsilon, \varepsilon)$ and ε sufficiently small. In fact, since $\operatorname{Re}\{t\bar{f}\phi g^2 h\} \le |t| \cdot |g|^2 |h|$, we have

$$1 - |f + t\phi g^2 h|^2 \ge |g|^2 (1 - 2|t| \cdot |h| - t^2 |g|^2 |h|^2),$$

which is nonnegative for $|t|$ sufficiently small. Since $v = \operatorname{argmax}(F)$, the derivative of F must be zero at v in the directions $\pm\phi g^2 h$ for arbitrary $h \in H^\infty$ of norm $\|h\|_\infty \le 1$, i.e.,

$$\int_{\mathbb{T}} |\sigma|^2 \frac{\operatorname{Re}\{\bar{f}\phi g^2 h\}}{1-|f|^2} dm = 0$$

for all $h \in H^\infty$ in the unit ball. Therefore,

$$|\sigma|^2 \frac{\bar{f}\phi g^2}{1-|f|^2} \in H_0^1,$$

so, since g is outer and $|\sigma|^2 \frac{\bar{f}\phi}{1-|f|^2} \in L_1$ (Lemma 12), (3.11) follows. □

In view of (3.8), we may write (3.11) as

$$|a|^2 \bar{f}\phi \in H_0^1. \tag{3.12}$$

Since $f \in H^\infty$, (3.12) implies that $|a|^2 |f|^2 \phi \in H_0^1$, and hence, $\phi |b|^2 \in H_0^1$. Therefore, since $|a|^2 = |\sigma|^2 + |b|^2$, where a is outer and $\sigma \in \mathcal{K}$, it follows from Lemma 10 that $a \in \mathcal{K}$. To prove that $b \in \mathcal{K}$, note that (3.12) implies that $\bar{b}a\phi \in H_0^1$ and, since a is outer, $\bar{b}\phi \in N^+$. Therefore, since $\bar{b}\phi$ is also in L^2, we have $\bar{b}\phi \in H_0^2$ [14, p. 28], which, by (3.7), implies that $b \in \mathcal{K}$.

This establishes the existence of an interpolant f of the form stated in Theorems 1. It remains to prove uniqueness. To this end, suppose that there are two such interpolants, i.e., that there exist $v_k \in X$, $k = 1, 2$, such that

$$f_k = \frac{b_k}{a_k} = w + \phi v_k, \quad b_k, a_k \in \mathcal{K}, \quad |a_k|^2 - |b_k|^2 = |\sigma|^2.$$

If so, then

$$\frac{|\sigma|^2}{1 - |f_k|^2} = |a_k|^2 \in L^1, \quad k = 1, 2. \tag{3.13}$$

Since X is convex, $v_1 + t(v_2 - v_1) \in X$ for $t \in [0, 1]$. Then, given (3.3), the function $\psi : [0, 1] \to \mathbb{R}$, defined by

$$\psi(t) = F(v_1 + t(v_2 - v_1)),$$

is differentiable at $t = 0$ and has the derivative

$$\dot{\psi}(0^+) = \int_{\mathbb{T}} |\sigma|^2 \frac{2\mathrm{Re}\{\bar{f}_1(f_1 - f_2)\}}{1 - |f_1|^2} dm$$

there. In fact, taking $q_1 = 1 - |f_1 + t(f_2 - f_1)|^2$ and $q_2 = 1 - |f_1|^2$ in

$$\log q_1 - \log q_2 \le \frac{q_1 - q_2}{q_2},$$

we obtain

$$\frac{1}{t}\left[\log\left(1 - |f_1 + t(f_2 - f_1)|^2\right) - \log\left(1 - |f_1|^2\right)\right] \\ - \frac{2\mathrm{Re}\{\bar{f}_1(f_1 - f_2)\}}{1 - |f_1|^2} \le -t\frac{|f_1 - f_2|^2}{1 - |f_1|^2} \le 0,$$

and hence we are allowed to differentiate inside the integral.

Now, in view of (3.13), we have

$$\dot{\psi}(0^+) = 2\mathrm{Re}\int_{\mathbb{T}} a_1\bar{b}_1\phi(v_1 - v_2)dm,$$

which equals zero since $\bar{b}_1\phi \in H_0^2$ and therefore $a_1\bar{b}_1\phi(v_1 - v_2) \in H_0^1$. The same argument can then be used to show that $\dot{\psi}(1^-) = 0$. This contradicts strict concavity of F unless $v_1 = v_2$. Therefore, $f_1 = f_2$, and hence $|a_1|^2 = |a_2|^2$ by (3.13). Consequently, since $a_1, a_2 \in \mathcal{K}_0$, we must have $a_1 = a_2$ and $b_1 = f_1a_1 = f_2a_2 = b_2$.

For the converse statement of Theorem 1, observe that (i), (ii) and (1.2) imply that $b = wa + \phi va$ and hence that $b = \mathrm{P}^{\mathcal{K}}wa = Ta$. Therefore, since $\|T\| < 1$, $\Psi := |a|^2 - |b|^2 \ge 0$ and hence, by Proposition 9, there is a unique $\sigma \in \mathcal{K}_0$ such that $\Psi = |\sigma|^2$.

3.4. **Proof of Theorem 2.** In view of the bijective correspondence (1.3) between $\mathcal{C}$ and $\mathcal{S}$, Theorem 2 follows quite directly from Theorem 1. To see this, note that, since $\mathcal{T}$ is bounded and $\mathrm{Re}\,\mathcal{T} > 0$, $T = (I - \mathcal{T})(I + \mathcal{T})^{-1}$ is a bounded linear operator on $\mathcal{K}$ and $\|T\| < 1$. Moreover, T commutes with S if and only if $\mathcal{T}$ does. Now, according to Theorem 1, there exist $a, b \in \mathcal{K}$ with $|a|^2 - |b|^2 = |\sigma|^2$ such that $f = b/a \in H^\infty$ satisfies $\|f\|_\infty \le 1$ and $f(S) = T$. Then, define α, β via (1.4). Since a is outer and $|a| \ge |b|$ a.e. on $\mathbb{T}$, by an analogue of Rouché's Theorem developed in [1, Lemma 3.1, p. 47], α and β are both outer. It is now easy to check that φ satisfies conditions (i), (ii) and (iii) of Theorem 2.

4. Well-posedness in the sense of Hadamard

Given any inner function ϕ, Theorem 1 yields a complete parameterization of solutions to the generalized interpolation problem with the complexity constraint (1.1) in terms of two sets of data, T and $\Psi := |\sigma|^2$. Fixing a choice of Ψ, for every choice of the operator T, there is a unique generalized interpolant f satisfying the complexity constraint. On the other hand, fixing T, for every choice of Ψ there is a unique generalized interpolant as well. In each case, the resulting bijections show that a generalized interpolant exists and is uniquely determined by the choice of problem data. Hadamard was among the first to emphasize that, in addition to existence and uniqueness of solutions to a (linear or) nonlinear problem, it is highly desirable to also prove continuous dependence of the solution with respect to variations in the problem data. We will now make these statements precise with appropriate choices of topologies.

Theorem 14. *Let T be an operator in $\mathcal{K}$ that commutes with S and has norm $\|T\| < 1$, let (Ψ_n) be a sequence in $\mathcal{Q}$ converging strongly to $\Psi \in \mathcal{Q}$, and let $(f_n), f \in H^\infty$ be the corresponding generalized interpolants prescribed by Theorem 1. Then $f_n \to f$ weak*.*

Proof. By the Main Lemma 6, the functional F defined by (3.3) with $\Psi = |\sigma|^2$ has a unique maximizer, which we denote $\hat{v}$. Analogously, for $n = 1, 2, 3, \dots$, define $\hat{v}_n = \operatorname{argmax} F_n$ where

$$F_n(v) = \int_{\mathbb{T}} \Psi_n \log(1 - |w + \phi v|^2) dm.$$

Lemma 15. *$\big(F_n(\hat{v}_n)\big)$ is a bounded sequence in $\mathbb{R}$.*

Proof. Let w be a Sarason interpolant, i.e., an interpolant of minimum norm. Then $\|w\|_\infty = \|T\| < 1$. By optimality of $\hat{v}_n$,

$$F_n(0) \le F_n(\hat{v}_n) \le 0.$$

However, $F_n(0)$ is bounded from below by $\log(1 - \|T\|^2)\|\Psi_n\|_1$, and $\|\Psi_n\|_1 \to \|\Psi\|_1$, and hence there is a uniform lower bound. $\square$

Therefore, since $(\hat{v}_n) \in X$, there is a subsequence $(\hat{v}_j)$ converging weakly* to some v_∞ in X and a Cauchy subsequence $\big(F_k(\hat{v}_k)\big)$ of $\big(F_j(\hat{v}_j)\big)$ converging to some $F^* \le 0$.

Lemma 16. *For all $v \in X$, we have $F(v) \le F^*$.*

Proof. For any $v \in X_+ = \{v \mid \|w + \phi v\|_\infty < 1\}$,

$$F_k(v) = \int_{\mathbb{T}} \Psi_k \log(1 - |w + \phi v|^2) dm \to \int_{\mathbb{T}} \Psi \log(1 - |w + \phi v|^2) dm = F(v),$$

since $\log(1 - |w + \phi v|^2) \in H^\infty$. By optimality, $F_k(v) \le F_k(\hat{v}_k)$, and consequently $F(v) \le F^*$ for all $v \in X_+$. Now, for any $\lambda \in [0, 1)$,we have $\lambda\hat{v} \in X_+$, and hence $F(\lambda\hat{v}) \le F^*$. However, F is strictly concave and $\hat{v} = \operatorname{argmax} F$, and therefore $\lambda \mapsto F(\lambda\hat{v})$ is monotonely increasing. Consequently, $F(\hat{v}) \le F^*$, and *a fortiori* $F(v) \le F^*$ for all $v \in X_+$. $\square$

Lemma 17. *Let v_∞ be the weak* limit of $(\hat{v}_k)$. Then, $F(v_\infty) = F^*$.*

Proof. For $\varepsilon > 0$, by Lemma 7, there is $\lambda \in \mathcal{L}$, such that

$$F(v_\infty) \geq \lambda(v_\infty) - \varepsilon.$$

Let

$$\lambda_k(v) = \lambda_{0k} + \mathrm{Re}\langle h_k, v\rangle$$

be defined from λ by exchanging Ψ for Ψ_k. Then $h_k \to h$ strongly as $k \to \infty$, and therefore, since $\hat{v}_k \to v_\infty$ weak*,

$$\langle h_k, v_k\rangle \to \langle h, v_\infty\rangle.$$

(See, e.g., [33, p. 148].) Since, in addition, $\lambda_k(v) \to \lambda_{0k}$, we have $\lambda_k(\hat{v}_k) \to \lambda(v_\infty)$ so that

$$F(v_\infty) \geq \lambda_k(\hat{v}_k) - 2\varepsilon$$

for k sufficiently large. However, by Lemma 7, $\lambda_k(\hat{v}_k) \geq F_k(\hat{v}_k)$, which tends to F^* as $k \to \infty$. Consequently, $F(v_\infty) \geq F^*$. However, by Lemma 16, $F(v_\infty) \leq F^*$, and hence $F(v_\infty) = F^*$ as claimed. □

Then, by Lemma 16 and Lemma 17,

$$F(\hat{v}) \leq F^* = F(v_\infty).$$

However, $\hat{v} = \operatorname{argmax} F$, and hence we must have $F(\hat{v}) = F(v_\infty)$. Therefore, since F is strictly concave, $\hat{v} = v_\infty$. Consequently, $v_k \to \hat{v}$, and thus $f_k = w + \phi v_k \to f$, weak*.

Therefore, the full sequence (f_n) has the property that every convergent subsequence (f_k) converges weak* to *the same* f. It follows that $f_n \to f$ weak*. In more detail, if this were not the case, there would be an (infinite) subsequence (v_ν) of (v_n) that must be excluded in order for the remaining sequence to converge to v_∞. Then, since (v_ν) is bounded, there is a subsequence (v_μ) converging weak* to some limit $\tilde{v}_\infty$ that does not equal v_∞. However, going again through the proof above, we see that $\tilde{v}_\infty = \hat{v}$ and hence that $\tilde{v}_\infty = v_\infty$, contrary to our assumption. Consequently, (f_n) tends to f weak*, as claimed.

This concludes the proof of Theorem 14. □

Theorem 18. *Let T be an operator in $\mathcal{K}$ that commutes with S and has norm $\|T\| < 1$, and let (T_n) be a sequence of operators in $\mathcal{K}$ with norm $\|T_n\| < 1$ also commuting with S such that $T_n \to T$ in operator norm. Moreover, given any $\Psi \in \mathcal{Q}$, let $(f_n), f \in H^\infty$ be the corresponding generalized interpolants prescribed by Theorem* 1*. Then $f_n \to f$ weak*.*

Proof. Let w be a Sarason interpolant of norm $\|w\|_\infty = \|T\| < 1$. In the same way, let u_n be a Sarason interpolant of $T_n - T$ of minimum norm. It follows that $\|u_n\|_\infty = \|T_n - T\| \to 0$. Setting $w_n = w + u_n$, we have that $w_n \to w$ strongly. Then, along the lines of the proof of Theorem 14, we define $\hat{v}_n = \operatorname{argmax} F_n$, where now

$$F_n(v) = \int_{\mathbb{T}} \Psi \log(1 - |w_n + \phi v|^2) dm.$$

Lemma 19. *For some sufficiently large N, $\big(F_n(\hat{v}_n)\big)_{n=N}^{\infty}$ is a bounded sequence in $\mathbb{R}$.*

Proof. Again by optimality of $\hat{v}_n$, we have $F_n(0) \leq F_n(\hat{v}_n) \leq 0$. Moreover,

$$F_n(0) - F(0) = \int_{\mathbb{T}} \Psi \log\left(\frac{1-|w_n|^2}{1-|w|^2}\right) dm.$$

Now, $1-|w|^2 > 0$ on $\mathbb{T}$. Since $w_n \to w$ strongly, $1-|w_n|^2 > 0$ on $\mathbb{T}$ for $n \geq N$ with N sufficiently large, and $F_n(0) \to F(0)$ as $n \to \infty$. Hence $F(0) - \varepsilon \leq F_n(\hat{v}_n) \leq 0$ for $\varepsilon > 0$ and n sufficiently large. □

Consequently, as above, there is a subsequence $(\hat{v}_j)$ converging weakly* to some v_∞ in X and a Cauchy subsequence $\big(F_k(\hat{v}_k)\big)$ of $\big(F_j(\hat{v}_j)\big)$ converging to some $F^* \leq 0$.

Lemma 20. *For all $v \in X$, we have $F(v) \leq F^*$.*

Proof. In the same manner as above, we observe that

$$F_k(0) - F(0) = \int_{\mathbb{T}} \Psi \log\left(\frac{1-|w_k+\phi v|^2}{1-|w+\phi v|^2}\right) dm \to 0$$

for $v \in X_+$ as $k \to \infty$. Then the rest of the proof is similar to that for Lemma 16. □

Lemma 21. *Let v_∞ be the weak* limit of $(\hat{v}_k)$. Then, $F(v_\infty) = F^*$.*

Proof. The proof is the same as for Lemma 17 after observing that

$$h_k - h = -2\Psi\left(\frac{(\overline{w_k+\phi u})\phi}{1+\varepsilon-|w_k+\phi u|^2} - \frac{(\overline{w+\phi u})\phi}{1+\varepsilon-|w+\phi u|^2}\right)$$

tends to zero strongly for all $\varepsilon > 0$ and $u \in X$. □

Finally, as in the proof of Theorem 14, we now see that $v_k \to \hat{v}$ weak*, and in fact $v_n \to \hat{v}$ weak*. Then the same holds for $f_n = w_n + \phi v_n \to f$. This concludes the proof of Theorem 18. □

5. Applications to interpolation

We now present three different examples where the theory is relevant. The first two are drawn from the trigonometric moment problem and from classical analytic interpolation, respectively. They correspond to cases where the inner function ϕ is a finite Blaschke product, and hence $\mathcal{K}$ is finite-dimensional. The last example is drawn from systems theory and corresponds to a case where ϕ is a singular inner function.

5.1. The Carathéodory extension problem.

Given $n+1$ complex numbers

$$c_0, c_1, \ldots, c_n,$$

consider the class of functions in the Carathéodory class $\mathcal{C}$ whose power series begins with the term $c_0 + c_1 z + \cdots + c_n z^n$. If c is one such function, all the others take the form

$$\varphi = c + \phi v, \quad v \in H(\mathbb{D}),$$

where ϕ is the inner function

$$\phi(z) = z^{n+1}.$$

The problem to characterize this class of interpolants has been widely studied by, among others, Carathéodory and Fejer [12], Toeplitz [32] and Schur [30], the latter of whom completely parameterized all interpolants in terms of what are now known as the *Schur parameters*.

In this context, the interpolants of the form β/α with $\alpha, \beta \in \mathcal{K} := H^2 \ominus \phi H^2$, studied in this paper, are the rational interpolants of degree at most n. In fact, $\mathcal{K}$ is the $n+1$-dimensional space spanned by the monomials $1, z, z^2, \dots, z^n$, and hence the functions in $\mathcal{K}$ are the polynomials of degree less or equal to n. In this basis, the shift has the matrix representation

$$S = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

and consequently an operator T on $\mathcal{K}$ commutes with S if and only if its matrix has the form

$$T = \begin{bmatrix} c_0 & 0 & 0 & \cdots & 0 \\ c_1 & c_0 & 0 & \cdots & 0 \\ c_2 & c_1 & c_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix}. \tag{5.1}$$

The interpolation problem has a solution if and only if the Toeplitz matrix

$$\operatorname{Re} T = \frac{1}{2} \begin{bmatrix} c_0 + \bar{c}_0 & \bar{c}_1 & \bar{c}_2 & \cdots & \bar{c}_n \\ c_1 & c_0 + \bar{c}_0 & \bar{c}_1 & \cdots & \bar{c}_{n-1} \\ c_2 & c_1 & c_0 + \bar{c}_0 & \cdots & \bar{c}_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & c_{n-2} & \cdots & c_0 + \bar{c}_0 \end{bmatrix} \tag{5.2}$$

is positive semidefinite. If this Toeplitz matrix is singular, there is a unique interpolant, namely the one discussed by Sarason in [29]. This solution, which has all its poles and zeros in the unit circle $\mathbb{T}$ (see, e.g., [22, p. 148] for an explicit expression) has been used extensively in signal processing and statistics in the context of spectral analysis of time series. In fact, even if $\operatorname{Re} T > 0$, the interpolation problem can be reduced to one where the Toeplitz matrix $\operatorname{Re} T$ is singular by subtracting a multiple of the identity. An interpolant for the original problem can be reconstructed by adding back that constant. Such a solution is known in signal processing as the *Pisarenko* solution.

In general, if $\operatorname{Re} T > 0$, there are infinitely many rational interpolants of degree at most n, and it is precisely these we characterize in this paper. We reformulate Theorems 2 and 3 in the present setting. To this end, let $\mathcal{P}(n)$ be the class of polynomial with degree at most n and with all its roots in the complement of the open unit disc $\mathbb{D}$, and let $\mathcal{P}_+(n)$ be the subclass of $\sigma \in \mathcal{P}(n)$ such that $\sigma(0) > 0$. Moreover, let $\mathcal{Q}_+$ be the class of trigonometric polynomials

$$\mathcal{Q}_+ = \{Q = \operatorname{Re}\{q\} > 0 \text{ on } \mathbb{T} \mid q \text{ is a polynomial of degree at most } n\}.$$

Theorem 22 ([18, 4, 20, 6]). *Let $c_0, c_1, \dots, c_n$ be complex numbers such that the Toeplitz matrix* (5.2) *is positive definite, and let T be given by* (5.1)*. Then, to each $\sigma \in \mathcal{P}_+(n)$ there corresponds a unique pair $(\alpha, \beta) \in \mathcal{P}_+(n) \times \in \mathcal{P}(n)$ such that*

(i) $\varphi = \beta/\alpha \in \mathcal{C}$,
(ii) $\varphi(z) = c_0 + c_1 z + \cdots + c_n z^n + \cdots$, *and*

(iii) $\bar{\alpha}\beta + \alpha\bar{\beta} = 2|\sigma|^2$ *on* $\mathbb{T}$.

Conversely, any φ *satisfying* (i) *and* (ii) *determines via* (iii)*, a unique* $\sigma \in \mathcal{P}_+(n)$*. Moreover, if* σ *has no zeros on the unit circle, the functional* $\mathbb{J}_\sigma : \mathcal{Q}_+ \to \mathbb{R}$*, given by*

$$\mathbb{J}_\sigma(Q) = \mathrm{Re}\langle c, q\rangle - \int_{\mathbb{T}} |\sigma|^2 \log Q dm,$$

has a unique minimum in $\mathcal{Q}_+$*,* α *is the unique solution in* $\mathcal{P}_+(n)$ *of*

$$|\alpha|^2 = \hat{Q} := \mathrm{argmin}(\mathbb{J}_\sigma),$$

and $\beta = \mathfrak{T}\alpha$*.*

The existence part of this theorem was proven in [18], where uniqueness was conjectured. Uniqueness was proved in [4], except when σ has a zero on $\mathbb{T}$, which was settled in [20]. These proofs were nonconstructive, using degree theory and the theory of foliations. The (constructive) optimization method was introduced in [6], which was selected by the SIAM editorial board as a SIGEST paper and republished in enhanced form as [9].

Rational Carathéodory interpolants of bounded degree are of particular interest in several engineering applications, since the degree relates to the complexity of the apparatus to be constructed. For lack of computational procedure, one has traditionally been confined to two solutions, namely the Pisarenko solution and the so-called *maximum entropy solution* φ_0, the interpolant that maximizes the entropy rate

$$\int_{\mathbb{T}} \log\left(\mathrm{Re}\{\varphi\}\right) dm.$$

The maximum entropy solution, which is the interpolant corresponding to $\sigma = 1$, can be determined by solving simple linear (normal) equations. In fact, $\varphi_0 = \beta/\alpha$, where α and β are the n:th Szegö polynomials (orthogonal on the unit circle) of the first and second kind, respectively; see, e.g., [22]. Theorem 4, interpreted in the present setting, describes the interpolant of degree at most n corresponding to an arbitrary σ as the maximizer of the generalized entropy rate

$$\int_{\mathbb{T}} |\sigma|^2 \log\left(\mathrm{Re}\{\varphi\}\right) dm.$$

It is interesting to notice that this optimization problem is equivalent to minimizing

$$\int_{\mathbb{T}} \Psi \log\left(\frac{\Psi}{\Phi}\right) dm,$$

over Φ, where $\Psi = |\sigma|^2$ and $\Phi = \mathrm{Re}\{\varphi\}$, which functional is precisely the Kullback-Leibler divergence between Φ and Ψ, when these are suitably normalized and interpreted as probability distributions. Such functionals have been extensively studied in probability and statistics [24].

5.2. **The Nevanlinna-Pick interpolation problem.** The following problem was first studied by Nevanlinna [26] and Pick [28]. Given n distinct points $z_0, z_1, \dots, z_n$ in the unit disc $\mathbb{D}$ and $n+1$ complex numbers $w_0, w_1, \dots, w_n$, find a function f in $\mathcal{S}$ (or in $\mathcal{C}$) such that

$$f(z_k) = w_k, \quad k = 0, 1, \dots, n. \tag{5.3}$$

For simplicity of notation, we also assume that $z_0 = 0$. (This can be done without loss of generality since the the bianalytic map $z \mapsto (z - z_0)/(1 - \bar{z}_0 z)$ maps the disc into itself and sends z_0 to zero.)

To transform this problem to the Sarason setting, it is standard to take the inner function ϕ to be the finite Blaschke product

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

Then, as is well known, the coinvariant subspace $\mathcal{K} := H^2 \ominus \phi H^2$ is the n-dimensional subspace spanned by $g_0, g_1, \dots, g_n$, where

$$g_0 = 1, \quad g_k(z) = \frac{1}{1 - \bar{z}_k z}, \quad k = 0, 1, \dots, n,$$

are the evaluation kernels satisfying $\langle g, g_k \rangle = g(z_k)$ for all $g \in H^2$. Since

$$\langle g, f(S^*) g_k \rangle = \langle fg, g_k \rangle = f(z_k) g(z_k) = \langle g, \overline{f(z_k)} g_k \rangle$$

for all $g \in \mathcal{K}$, we have

$$f(S)^* g_k = \overline{f(z_k)} g_k, \quad k = 0, 1, \dots, n, \tag{5.4}$$

i.e., $g_0, g_1, \dots, g_n$ are the eigenvectors of $f(S)^*$ with eigenvalues $\overline{f(z_0)}, \overline{f(z_1)}, \dots, \overline{f(z_n)}$. In particular,

$$S^* g_k = \bar{z}_k g_k \quad \text{and} \quad T^* g_k = \bar{w}_k g_k, \quad k = 0, 1, \dots, n, \tag{5.5}$$

if T is an operator on $\mathcal{K}$ that commutes with S.

The condition $\|T\| < 1$ in Theorem 1 can be written $1 - TT^* > 0$. Therefore, if $g = \alpha_0 g_0 + a_1 g_1 + \dots + a_n g_n$, then, in view of (5.5), the condition $\|T\| < 1$ is equivalent to

$$\langle g, (1 - TT^*) g \rangle = \langle g, g \rangle - \langle T^* g, T^* g \rangle = a^* P_n a > 0,$$

where P_n is the Pick matrix

$$P_n = \left[\frac{1 - \bar{w}_j w_k}{1 - \bar{z}_j z_k} \right]_{j,k=0}^{n}. \tag{5.6}$$

Likewise, the condition $\operatorname{Re} T > 0$ in Theorem 2 can be written

$$\langle T^* g, g \rangle + \langle g, T^* g \rangle = a^* \Pi_n a > 0,$$

where Π_n is the Pick matrix

$$\Pi_n = \left[\frac{\bar{w}_j + w_k}{1 - \bar{z}_j z_k} \right]_{j,k=0}^{n}. \tag{5.7}$$

Again, if the Pick matrix, given by (5.6) in the case when $f \in \mathcal{S}$ and by (5.7) in the case when $f \in \mathcal{C}$, is singular, there is only one interpolant. If the Pick matrix is positive definite, there are infinitely many solutions, and we are interested in the ones that can be written as a quotient of two functions in $\mathcal{K}$. Since any $g \in \mathcal{K}$, takes the form π/τ, where

$$\tau(z) = \prod_{k=1}^{n} (1 - \bar{z}_k z)$$

and π is an arbitrary polynomial of degree at most n, the class of interpolants under consideration consists precisely of all rational functions of degree at most n, the same complexity constraint as in the Carathéodory extension problem. In the present setting, Theorems 2 and 3 can then be formulated in the following way, where we now take $\mathcal{Q}_+$ to be

$$\mathcal{Q}_+ = \left\{ q = (q_0, q_1, \dots, q_n) \in \mathbb{R} \times \mathbb{C}^n \mid Q := \mathrm{Re}\left(\sum_{k=0}^{n} q_k g_k \right) > 0 \text{ on } \mathbb{T} \right\}.$$

Theorem 23 ([19, 20, 7]). *Let $w_1, w_2, \dots, w_n$ be complex numbers such that the Pick matrix* (5.7) *is positive definite, and let T be given by* (5.5). *Then, to each $\rho \in \mathcal{P}_+(n)$ there is a unique pair $(\pi, \chi) \in \mathcal{P}_+(n) \times \in \mathcal{P}(n)$ such that*

(i) $f = \chi/\pi \in \mathcal{C}$,

(ii) $f(z_k) = w_k, \quad k = 1, 2, \dots, n$, *and*

(iii) $\bar{\pi}\chi + \pi\bar{\chi} = 2|\rho|^2$ *on* $\mathbb{T}$.

Conversely, any f satisfying (i) *and* (ii) *determines via* (iii), *a unique $\rho \in \mathcal{P}_+(n)$. Moreover, if ρ has no zeros on the unit circle, the functional $\mathbb{J}_\rho : \mathcal{Q}_+ \to \mathbb{R}$, given by*

$$\mathbb{J}_\rho(q) = \mathrm{Re} \sum_{k=0}^{n} w_k q_k - \int_{\mathbb{T}} \left| \frac{\rho}{\tau} \right|^2 \left(\log \mathrm{Re} \sum_{k=0}^{n} q_k g_k \right) dm,$$

has a unique minimum in $\mathcal{Q}_+$, π is the unique solution in $\mathcal{P}_+(n)$ of

$$|\pi|^2 = |\tau|^2 \mathrm{argmin}(\mathbb{J}_\rho),$$

and $\chi = \tau T(\pi/\tau)$.

The existence part of this theorem was proved in [19], and a uniqueness proof was given in [20]. A constructive proof, introducing the optimization problem of the theorem, was presented in [7]. The corresponding theorem for $f \in \mathcal{S}$ is obtained by replacing Pick matrix (5.7) by (5.6) and conditions (i) by

(i) $f = (\pi - \chi)/(\pi + \chi) \in \mathcal{S}$.

The unique interpolant corresponding to ρ can also be obtained as the unique maximizer of a generalized entropy rate over the space of functions satisfying the interpolation condition (5.3). By Theorem 4, this entropy rate is

$$\mathbb{I}_\rho(f) = \int_{\mathbb{T}} \left| \frac{\rho}{\tau} \right|^2 \log\left(\mathrm{Re}\{f\} \right) dm$$

for $f \in \mathcal{C}$, and, by Theorem 5, it is

$$\mathbb{K}_\rho(f) = \int_{\mathbb{T}} \left| \frac{\rho}{\tau} \right|^2 \log\left(1 - |f|^2 \right) dm$$

for $f \in \mathcal{S}$. Nevanlinna-Pick interpolation problems with degree constraint abound in applications, ranging from signal processing to robust control engineering (see, e.g., [7, 9] for references), and often one has had to content oneself with the central (or maximum entropy) solution obtained by taking $\rho = \tau$, which can be computed using linear algebra. For example, in robust H^∞ control, the problem to determine a control system whose "sensitivity function" f maximizes

$$\int_{\mathbb{T}} \log\left(1 - |f|^2 \right) dm$$

is known as the maximum entropy controller [25].

Needless to say, more general Nevanlinna-Pick interpolation problems, allowing for multiple interpolation points, can also be handled by the theory presented in this paper by choosing the appropriate inner function ϕ.

5.3. **A problem in systems theory.** Let $L^p(i\mathbb{R})$ be the L^p space over the imaginary axis and $H^p(\mathbb{C}_+)$ the corresponding Hardy space of functions analytic in the right half of the complex plane. Given $w \in H^\infty(\mathbb{C}_+)$, consider the map Σ_w defined by the commutative diagram

$$\begin{array}{ccc} L^2[0,\infty) & \xrightarrow{\Sigma_w} & L^2[0,\infty) \\ \mathcal{F}\big\downarrow & & \big\downarrow\mathcal{F} \\ H^2(\mathbb{C}_+) & \xrightarrow{M_w} & H^2(\mathbb{C}_+) \end{array}$$

where M_w denotes the multiplication with w and

$$\mathcal{F} : \hat{a} \mapsto a(s) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{st} \hat{a}(t) dt.$$

Consider the singular inner function $\phi(s) = e^{-s}$ and the subspace

$$\mathcal{K} = H^2(\mathbb{C}_+) \ominus \phi H^2(\mathbb{C}_+),$$

and suppose that the compressed operator

$$T = \mathrm{P}^{\mathcal{K}} M_w|_{\mathcal{K}}$$

has norm less than one. Since $\mathcal{F}^{-1} M_\phi \mathcal{F}$ sends $\hat{a}(t)$ to $\hat{a}(t-1)$, the subspace $\mathcal{K}$ is mapped via $\mathcal{F}^{-1}$ onto $L^2[0,1]$ which is identified with the subspace of $L^2[0,\infty)$ of functions which are zero on the interval $(1,\infty)$.

In systems theory the map Σ_w constitutes a model for a linear stable dynamical system that transforms an input signal $\hat{a}$ to an output signal $\hat{b}$. The domain and range of Σ_w are referred to as the input and output spaces, and the square of the L_2-norm represents the energy of the signal. Thus, if $\|w\|_\infty < 1$, then Σ_w is said to be passive since the energy of the input always exceeds the energy of the output; see, e.g., [2]. The operator T, on the other hand, represents observed data about Σ_w collected on the time interval $[0,1]$. The problem we address is therefore to parametrize the family of all input-output pairs $(\hat{a}, \hat{b})$ that are consistent with the data T, correspond to a passive underlying system, and have support on the time interval $[0,1]$. Each such pair gives rise to a possible model Σ_f, where $f = b/a$ with $a = \mathcal{F}\hat{a}$ and $b = \mathcal{F}\hat{b}$. Such a model is consistent with the data in the sense that $\mathrm{P}^{\mathcal{K}} M_f|_{\mathcal{K}} = T$. This can be seen as an inverse problem, where the underlying model Σ_w is unknown and only information about its behavior on the interval $[0,1]$ is available in the form of T.

Transforming an interpolation problem in $H^\infty(\mathbb{C}_+)$ into one in $H^\infty(\mathbb{D})$ is standard and relies on the linear fractional transformation

$$\mathbb{C}_+ \to \mathbb{D} : s \mapsto z = \frac{s-1}{s+1}.$$

Accordingly, a function $\tilde{f} \in H^\infty(\mathbb{D})$ is transformed to a function $f \in H^\infty(\mathbb{C}_+)$ via

$$f(s) = \tilde{f}\left(\frac{s-1}{s+1}\right),$$

while a function $\tilde{a} \in H^2(\mathbb{D})$ is transformed to a function $a \in H^2(\mathbb{C}_+)$ via

$$a(s) = \frac{1}{s+1} \tilde{a}\left(\frac{s-1}{s+1}\right).$$

(Cf. [23, p. 130].) Under this correspondence the inner function $\phi(s) = e^{-s}$ is transformed to the singular inner function $\tilde{\phi}(z) = \exp\left(\frac{z+1}{z-1}\right)$, while the ordinary shift M_z in $H^2(\mathbb{D})$ corresponds to multiplication with $(s-1)/(s+1)$. It is also interesting to note that this operator corresponds in $L^2[0,\infty)$ to

$$\mathcal{F}^{-1} M_{\left(\frac{s-1}{s+1}\right)} \mathcal{F} = (I - V)(I + V)^{-1},$$

where $V : L^2[0,\infty) \to L^2[0,\infty)$ is the Volterra operator $V\hat{a}(t) = \int_0^t \hat{a}(\tau)d\tau$. (Cf. [29].)

Thus applying Theorem 1 to the present setting via the above correspondence we have the following result. *For any* σ *in*

$$\mathcal{K}_0 = \{\sigma \in \mathcal{K} \mid \sigma \text{ outer in } H^2(\mathbb{C}_+) \text{ and } \sigma(1) > 0\},$$

there exists a unique pair $(a, b) \in \mathcal{K}_0 \times \mathcal{K}$ *such that*

(i) $f = b/a \in H^\infty(\mathbb{C}_+)$ *with* $\|f\|_\infty \le 1$,
(ii) $P^{\mathcal{K}} M_f|_{\mathcal{K}} = T$, *and*
(iii) $|a|^2 - |b|^2 = |\sigma|^2$ *a.e. on* $i\mathbb{R}$.

The correspondence between the pair $(a, b) \in \mathcal{K}_0 \times \mathcal{K}$ *and* $\sigma \in \mathcal{K}_0$ *is bijective.*

Consequently, for each $\hat{\sigma} \in \mathcal{F}^{-1}\mathcal{K}_0$, there is a unique model Σ_f that explains the data T, and a corresponding input-output pair $(\hat{a}, \hat{b}) \in L^2[0,1] \times L^2[0,1]$. It is interesting to remark that

$$\|a\|^2 - \|b\|^2 = \|\sigma\|^2$$

represents the total energy dissipation in the model Σ_f when $\hat{a}$ is the input. In fact, $|a|^2 - |b|^2 = |\sigma|^2$ describes the distribution of dissipation across frequencies. Thus σ may be selected to reflect additional prior information about the underlying system.

References

[1] V. M. Adamjan, D. Z. Arov and M. G. Krein, Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem, *Math. USSR Sbornik* **15** (1971), 31–73. MR45:7505

[2] D. Z. Arov and M. A. Nudelman. Passive linear stationary dynamical scattering systems with continuous time, *Integral Equations Operator Theory* **24** (1996), 1–45. MR96k:47016

[3] J. A. Ball and J. W. Helton, A Beurling-Lax theorem for the Lie group $U(m,n)$ which contains most classical interpolation theory, *J. Operator Theory* **9** (1983), 107–142. MR84m:47046

[4] C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Matveev, A complete parameterization of all positive rational extensions of a covariance sequence, *IEEE Trans. Automat. Control* **40** (1995), 1841–1857. MR96i:93015

[5] C. I. Byrnes, H. J. Landau and A. Lindquist, *On the well-posedness of the rational covariance extension problem*, in *Current and Future Directions in Applied Mathematics*, eds. M. Alber, B. Hu, J. Rosenthal, Birkhäuser, pp. 83-106, 1997. MR98c:93026

[6] C. I. Byrnes, S. V. Gusev, and A. Lindquist, A convex optimization approach to the rational covariance extension problem, *SIAM J. Contr. and Optimiz.* **37** (1998) 211–229. MR99f:93135

[7] C. I. Byrnes, T. T. Georgiou, and A. Lindquist, A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint, *IEEE Trans. Automat. Control* **46** (2001), 822–839. MR2002c:93038

[8] C. I. Byrnes and A. Lindquist, On the duality between filtering and Nevanlinna-Pick interpolation, *SIAM J. Contr. and Optimiz.* **39** (2000), 757–775. MR2001h:93090

[9] C. I. Byrnes, S. V. Gusev, and A. Lindquist, From finite covariance windows to modeling filters: A convex optimization approach, *SIAM Review* **43** (2001), 645–675. MR2002k:93080

[10] C. I. Byrnes and A. Lindquist, Interior point solutions of variational problems and global inverse function theorems, Report TRITA/MAT-01-OS13, 2001, Royal Institute of Technology, Stockholm, Sweden, 2001.

[11] C. I. Byrnes and A. Lindquist, A convex optimization approach to generalized moment problems, *Control and Modeling of Complex Systems: Cybernetics in the 21st Century: Festschrift in Honor of Hidenori Kimura on the Occasion of his 60th Birthday*, K. Hashimoto, Y. Oishi and Y. Yamamoto, Editors, Birkhäuser, 2003, 3–21. MR2004a:93022

[12] C. Carathéodory and L. Fejer, Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizenten und über den Picard-Landau'schen Satz, *Rend. Circ. Mat. Palermo* **32** (1911), 218–239.

[13] R. G. Douglas, H. S. Shapiro, and A. L. Shields, Cyclic vectors and invariant subspaces for the backward shift operator, *Ann. Inst. Fourier, Grenoble* **20** (1970), 37–76. MR42:5088

[14] P. L. Duren, *Theory of H^p Spaces*, Academic Press, 1970. MR42:3552

[15] H. Dym and I. Gohberg, A maximum entropy principle for contractive interpolants, *J. Functional Analysis* **65** (1986), 83–125. MR87h:47065

[16] J. B. Garnett, *Bounded Analytic Functions*, Academic Press, 1981. MR83g:30037

[17] T. T. Georgiou, *Partial Realization of Covariance Sequences*, Ph.D. thesis, CMST, University of Florida, Gainesville 1983.

[18] T. T. Georgiou, Realization of power spectra from partial covariance sequences, *IEEE Trans. Acoustics, Speech and Signal processing* **35** (1987), 438–449.

[19] T. T. Georgiou, A Topological Approach to Nevanlinna-Pick Interpolation, *SIAM J. Math. and Anal.* **18** (1987), 1248–1260. MR88j:30076

[20] T. T. Georgiou, The Interpolation Problem with a Degree Constraint, *IEEE Trans. Automat. Control* **44** (1999), 631–635.

[21] T. T. Georgiou and A. Lindquist, Kullback-Leibler approximation of spectral density functions, *IEEE Trans. Information Theory* **49** (2003), 2910–2917.

[22] U. Grenander and G. Szegö, *Toeplitz forms and their applications*, Univ. California Press, 1958. MR20:1349

[23] K. Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall, 1962. MR24:A2844

[24] S. Kullback, *Information Theory and Statistics*, 2nd edition, New York: Dover Books, 1968 (1st ed. New York: John Wiley, 1959). MR21:2325

[25] D. Mustafa and K. Glover, *Minimum Entropy H_∞ Control*, Springer-Verlag, Berlin Heidelberg, 1990. MR92g:93002

[26] R. Nevanlinna, Über beschränkte Funktionen die in gegebenen Punkten vorgeschriebene Werte annehmen, *Ann. Acad, Sci. Fenn.* Ser A 13(1), 1919.

[27] N. K. Nikol'skiĭ, *Treatise on the Shift Operator*, Springer-Verlag: Berlin, 1986. MR87i:47042

[28] G. Pick, Über die Beschränkungen analytischer Funktionen, welche durch vorgegebene Funktionswerte bewirkt werden, *Math. Ann.* **77** (1916), 7–23.

[29] D. Sarason, Generalized Interpolation in H^∞, *Trans. Amer. Math. Society* **127** (1967), 179-203. MR34:8193

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

[31] B. Sz.-Nagy and C. Foias, *Harmonic Analysis of Operators on Hibert Space,* North-Holland, Amsterdam, 1970. MR43:947

[32] O. Toeplitz, Über die Fouriersche Entwicklung positiver Funktionen, *Rendiconti del Circolo Matematico di Palermo* **32** (1911), 191–192.

[33] E. Zeidler, *Nonlinear Functional Analysis and its Applications*, Vol. III, Springer-Verlag, New York, 1985. MR90b:49005

Department of Electrical and Systems Engineering, Washington University, St. Louis, Missouri 63130

Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota 55455

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

Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307