SIAM J. CONTROL OPTIM.
Vol. 37, No. 1, pp. 211–229

A CONVEX OPTIMIZATION APPROACH TO THE RATIONAL COVARIANCE EXTENSION PROBLEM*

CHRISTOPHER I. BYRNES†, SERGEI V. GUSEV‡, AND ANDERS LINDQUIST§

Abstract. In this paper we present a convex optimization problem for solving the rational covariance extension problem. Given a partial covariance sequence and the desired zeros of the modeling filter, the poles are uniquely determined from the unique minimum of the corresponding optimization problem. In this way we obtain an algorithm for solving the covariance extension problem, as well as a constructive proof of Georgiou's seminal existence result and his conjecture, a stronger version of which we have resolved in [Byrnes et al., *IEEE Trans. Automat. Control*, AC-40 (1995), pp. 1841–1857].

Key words. rational covariance extension, partial stochastic realization, trigonometric moment problem, spectral estimation, speech processing, stochastic modeling

AMS subject classifications. 30E05, 60G35, 62M15, 93A30, 93E12

PII. S0363012997321553

1. Introduction.

In [7] a solution to the problem of parameterizing all rational extensions of a given window of covariance data has been given. This problem has a long history, with antecedents going back to potential theory in the work of Carathéodory, Toeplitz, and Schur [9, 10, 31, 30], and continuing in the work of Kalman, Georgiou, Kimura, and others [18, 14, 21]. It has been of more recent interest due to its significant interface with problems in signal processing and speech processing [11, 8, 25, 20] and in stochastic realization theory and system identification [2, 32, 22]. Indeed, the recent solution to this problem, which extended a result by Georgiou and confirmed one of his conjectures [13, 14], has shed some light on the stochastic (partial) realization problem through the development of an associated Riccati-type equation, whose unique positive semidefinite solution has as its rank the minimum dimension of a stochastic linear realization of the given rational covariance extension [6]. In both its form as a complete parameterization of rational extensions to a given covariance sequence and as an indefinite Riccati-type equation, one of the principal problems which remains open is that of developing effective computational methods for the approximate solution of this problem. In this paper, motivated by the effectiveness of interior point methods for solving nonlinear convex optimization problems, we recast the fundamental problem as such an optimization problem.

In section 2 we describe the principal results for the rational covariance extension problem and set notation we shall need throughout. The only solution to this problem for which there have been simple computational procedures is the so-called *maximum entropy* solution, which is the particular solution that maximizes the entropy gain.

*Received by the editors May 15, 1997; accepted for publication December 3, 1997; published electronically October 7, 1998. 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.

http://www.siam.org/journals/sicon/37-1/32155.html

†Department of Systems Science and Mathematics, Washington University, St. Louis, MO 63130 (chrisbyrnes@seas.wustl.edu).

‡Department of Mathematics and Mechanics, St. Petersburg University, St. Petersburg 198904, Russia (sergei@gusev.niimm.spb.su).

§Division of Optimization and Systems Theory, Royal Institute of Technology, 100 44 Stockholm, Sweden (alq@math.kth.se).

In section 3 we demonstrate that the infinite-dimensional optimization problem for determining this solution has a simple finite-dimensional dual. This motivates the introduction in section 4 of a nonlinear, strictly convex functional defined on a closed convex set naturally related to the covariance extension problem. We first show that any solution of the rational covariance extension problem lies in the interior of this convex set and that, conversely, an interior minimum of this convex functional will correspond to the unique solution of the covariance extension problem. Our interest in this convex optimization problem is, therefore, twofold: as a starting point for the computation of an explicit solution and as a means of providing an alternative proof of the rational covariance extension theorem.

Concerning the existence of a minimum, we show that this functional is proper and bounded below, i.e., that the sublevel sets of this functional are compact. From this, it follows that there exists a minimum. Since uniqueness follows from strict convexity of the functional, the central issue which needs to be addressed in order to solve the rational covariance extension problem is whether, in fact, this minimum is an interior point. Indeed, our formulation of the convex functional, which contains a barrier-like term, was inspired by interior point methods. However, in contrast to interior point methods, the barrier function we have introduced does not become infinite on the boundary of our closed convex set. Nonetheless, we are able to show that the gradient, rather than the value, of the convex functional becomes infinite on the boundary. The existence of an interior point which minimizes the functional then follows from this observation.

In section 5, we apply these convex minimization techniques to the rational covariance extension problem, noting that, as hinted above, we obtain a new proof of Georgiou's conjecture. Moreover, this proof, unlike our previous proof [7] and the existence proof of Georgiou [14], is constructive. Consequently, we have also obtained an algorithmic procedure for solving the rational covariance extension problem. In section 6 we report some computational results and present some simulations.

2. The rational covariance extension problem.

It is well known that the spectral density $\Phi(z)$ of a purely nondeterministic stationary random process $\{y(t)\}$ is given by the Fourier expansion

$$\Phi(e^{i\theta}) = \sum_{-\infty}^{\infty} c_k e^{ik\theta} \tag{2.1}$$

on the unit circle, where the covariance lags

$$c_k = \mathrm{E}\{y_{t+k} y_t\}, \quad k = 0, 1, 2, \dots \tag{2.2}$$

play the role of the Fourier coefficients

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Phi(e^{i\theta}) d\theta. \tag{2.3}$$

In spectral estimation [8], identification [2, 22, 32], speech processing [11, 25, 24, 29], and several other applications in signal processing and systems and control, we are faced with the inverse problem of finding a spectral density which is *coercive*, i.e., positive on the unit circle, given only

$$c = (c_0, c_1, \dots, c_n), \tag{2.4}$$

which is a *partial covariance sequence* positive in the sense that

$$T_n = \begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_0 & \cdots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & \cdots & c_0 \end{bmatrix} > 0, \tag{2.5}$$

i.e., the Toeplitz matrix T_n is positive definite.

In fact, the covariance lags (2.2) are usually estimated from an approximation

$$\frac{1}{N-k+1} \sum_{t=0}^{N-k} y_{t+k} y_t$$

of the ergodic limit

$$c_k = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} y_{t+k} y_t,$$

since only a finite string

$$y_0, y_1, y_2, y_3, \ldots, y_N$$

of observations of the process $\{y(t)\}$ is available, and therefore we can only estimate a finite partial covariance (2.4), where $n << N$.

The corresponding inverse problem is a version of the *trigonometric moment problem* [1, 16]: Given a sequence (2.4) of real numbers satisfying the positivity condition (2.5), find a coercive spectral density $\Phi(z)$ such that (2.3) is satisfied for $k = 0, 1, 2, \ldots, n$. Of course there are infinitely many such solutions, and we shall shortly specify some additional properties which we would like the solution to have.

The trigonometric moment problem, as stated above, is equivalent to the *Carathéodory extension problem* to determine an extension

$$c_{n+1}, c_{n+2}, c_{n+3}, \ldots, \tag{2.6}$$

with the property that the function

$$v(z) = \frac{1}{2} c_0 + c_1 z^{-1} + c_2 z^{-2} + \cdots \tag{2.7}$$

is *strictly positive real*, i.e., is analytic on and outside the unit circle (so that the Laurent expansion (2.7) holds for all $|z| \geq 1$) and satisfies

$$v(z) + v(z^{-1}) > 0 \quad \text{on the unit circle.} \tag{2.8}$$

In fact, given such a $v(z)$,

$$\Phi(z) = v(z) + v(z^{-1}) \tag{2.9}$$

is a solution to the trigonometric moment problem. Conversely, any coercive spectral density $\Phi(z)$ uniquely defines a strictly positive real function $v(z)$ via (2.9).

These problems are classical and go back to Carathéodory [9, 10], Toeplitz [31], and Schur [30]. In fact, Schur parameterized all solutions in terms of what are now known as the *Schur parameters*, or, more commonly in the circuits and systems literature, as *reflection coefficients*, and which are easily determined from the covariance lags via the Levinson algorithm [27]. More precisely, modulo the choice of c_0, there is a one-to-one correspondence between infinite covariance sequences $c_0, c_1, c_2, \dots$ and Schur parameters $\gamma_0, \gamma_1, \dots$ such that

$$|\gamma_t| < 1 \quad \text{for } t = 0, 1, 2, \dots, \tag{2.10}$$

under which partial sequences (2.4) correspond to partial sequences $\gamma_0, \gamma_1, \dots, \gamma_{n-1}$ of Schur parameters. Therefore, covariance extension (2.6) amounts precisely to finding a continuation

$$\gamma_n, \gamma_{n+1}, \gamma_{n+2}, \dots \tag{2.11}$$

of Schur parameters satisfying (2.10). Each such solution is only guaranteed to yield a $v(z)$ which is meromorphic.

In circuits and systems theory, however, we are generally only interested in solutions which yield a rational $v(z)$ of at most degree n, or, equivalently, a rational spectral density $\Phi(z)$ of at most degree $2n$. Then the unique rational, stable, minimum-phase function $w(z)$ having the same degree as $v(z)$ and satisfying

$$w(z)w(z^{-1}) = \Phi(z) \tag{2.12}$$

is the transfer function of a *modeling filter*, which shapes white noise into a random process with the first $n+1$ covariance lags given by (2.4); see, e.g., [7, 6] for more details.

Setting all free Schur parameters (2.11) equal to zero, which clearly satisfies the condition (2.10), yields a rational solution

$$\Phi(z) = \frac{1}{a(z)a(z^{-1})}, \tag{2.13}$$

where $a(z)$ is a polynomial given by

$$a(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n \quad (a_0 > 0), \tag{2.14}$$

which is easily computed via the Levinson algorithm [27]. This so-called *maximum entropy solution* is an all-pole or AR solution, and the corresponding modeling filter

$$w(z) = \frac{z^n}{a(z)} \tag{2.15}$$

has all its zeros at the origin.

However, in many applications a wider variety in the choice of zeros is required in the spectral density $\Phi(z)$. To illustrate this point, consider in Figure 2.1 a spectral density in the form of a periodogram determined from a speech signal sampled over 20 milliseconds (in which time interval it represents a stationary process) together with a maximum entropy solution corresponding to $n = 6$. As can be seen, the latter yields a rather flat spectrum which is unable to approximate the valleys or the "notches" in the speech spectrum, and therefore in speech synthesis, the maximum entropy solution results in artificial speech which sounds quite flat. This is a manifestation of the fact that all the zeros of the maximum entropy filter (2.15) are located at the origin and thus do not give rise to a frequency where the power spectrum vanishes. However,

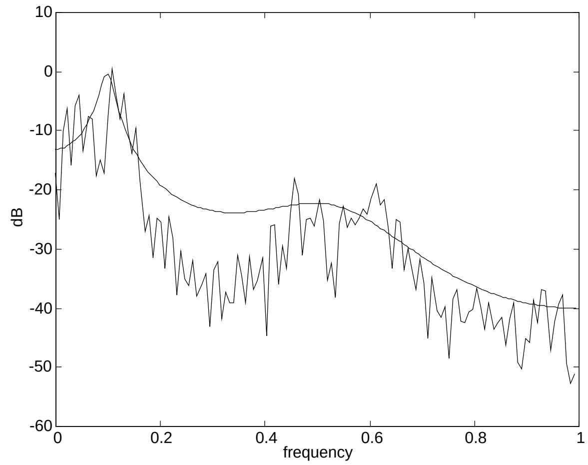

FIG. 2.1. *Spectral envelope of a maximum entropy solution.*

were we able to place some zeros of the modeling filter reasonably close to the unit circle, these would produce notches in the spectrum at approximately the frequency of the arguments of those zeros.

For this reason, it is widely appreciated in the signal and speech processing community that regeneration of human speech requires the design of filters having nontrivial zeros [3, p. 1726], [24, pp. 271–272], [29, pp. 76–78]. Indeed, while all-pole filters can reproduce many human speech sounds, the acoustic theory teaches that nasals and fricatives require both zeros and poles [24, pp. 271–272], [29, p. 105].

Therefore, we are interested in modeling filters

$$w(z) = \frac{\sigma(z)}{a(z)}, \tag{2.16}$$

for which (2.14) and

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \cdots + \sigma_n \tag{2.17}$$

are *Schur polynomials*, i.e., polynomials with all roots in the open unit disc. In this context, the maximum entropy solution corresponds to the choice $\sigma(z) = z^n$.

An important mathematical question, therefore, is to what extent it is possible to assign desired zeros and still satisfy the interpolation condition that the partial covariance sequence (2.4) is as prescribed. In [13] (see also [14]), Georgiou proved that for any prescribed zero polynomial $\sigma(z)$ there exists a modeling filter $w(z)$ and conjectured that this correspondence would yield a complete parameterization of all rational solutions of at most degree n, i.e., that the correspondence between v and a choice of positive sequence (2.4) and a choice of Schur polynomial (2.14) would be a bijection. This is a nontrivial and highly nonlinear problem, since generally there is no method to see which choices of free Schur parameters will yield rational solutions. In [7] we resolved this long-standing conjecture by proving the following theorem as a corollary of a more general theorem on complementary foliations of the space of all rational positive real functions of degree at most n.

THEOREM 2.1 (see [7]). *Given any partial covariance sequence* (2.4) *and Schur polynomial* (2.17), *there exists a unique Schur polynomial* (2.14) *such that* (2.16) *is* *a minimum-phase spectral factor of a spectral density* $\Phi(z)$ *satisfying*

$$\Phi(z) = c_0 + \sum_{k=1}^{\infty} \hat{c}_k (z^k + z^{-k}),$$

where

$$\hat{c}_k = c_k \quad \textit{for} \quad i = 1, 2, \ldots, n.$$

In particular, the solutions of the rational positive extension problem are in one-to-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, this correspondence is bianalytic.*

Consequently, we not only proved Georgiou's conjecture that the family of all rational covariance extensions of (2.4) of degree at most n is completely parameterized in terms of the zeros of the corresponding modeling filters $w(z)$, but also that the modeling filter $w(z)$ depends analytically on the covariance data and the choice of zeros, a strong form of well-posedness increasing the likelihood of finding a numerical algorithm.

In fact, both Georgiou's existence proof and our proof of Theorem 2.1 are nonconstructive. However, in this paper we present for the first time an algorithm which, given the partial covariance sequence (2.4) and the desired zero polynomial (2.17), computes the unique pole polynomial (2.14). This is done via the convex optimization problem to minimize the value of the function $\varphi : \mathbb{R}^{n+1} \to \mathbb{R}$, defined by

$$\varphi(q_0, q_1, \ldots, q_n) = c_0 q_0 + c_1 q_1 + \cdots + c_n q_n - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log Q(e^{i\theta}) |\sigma(e^{i\theta})|^2 d\theta \tag{2.18}$$

over all $q_0, q_1, \ldots, q_n$ such that

$$Q(e^{i\theta}) = q_0 + q_1 \cos\theta + q_2 \cos 2\theta + \cdots + q_n \cos n\theta > 0 \quad \text{for all } \theta. \tag{2.19}$$

In sections 4 and 5 we show this problem has a unique minimum. In this way we shall also provide a new and constructive proof of the weaker form of Theorem 2.1 conjectured by Georgiou.

Using this convex optimization problem, a sixth-degree modeling filter with zeros at the appropriate frequencies can be constructed for the speech segment represented by the periodogram of Figure 2.1. In fact, Figure 2.2 illustrates the same periodogram together with the spectral density of such a filter. As can be seen, this filter yields a much better description of the notches than does the maximum entropy filter.

Before turning to the main topic of this paper, the convex optimization problem for solving the rational covariance extension problem for arbitrarily assigned zeros, we shall provide a motivation for this approach in terms of the maximum entropy solution.

3. The maximum entropy solution. As a preliminary we shall first consider the maximum entropy solution discussed in section 2. The reason for this is that, as indicated by its name, this particular solution corresponds to an optimization problem. Hence, this section will be devoted to clarifying the relation between this particular optimization problem and the class of problems solving the general problem. Thus

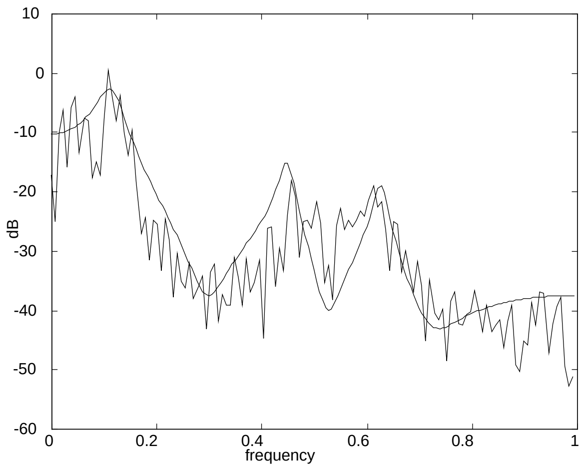

FIG. 2.2. *Spectral envelope obtained with appropriate choice of zeros.*

our interest is not in the maximum entropy solution per se, but in showing that it can be determined from a constrained convex minimization problem in $\mathbb{R}^{n+1}$, which naturally is generalized to a problem with arbitrary prescribed zeros.

Let us briefly recall the problem at hand. Given the partial covariance sequence

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

determine a coercive, rational spectral density

$$\Phi(z) = \hat{c}_0 + \sum_{k=1}^{\infty} \hat{c}_k (z^k + z^{-k}) \tag{3.1}$$

of degree at most $2n$ such that

$$\hat{c}_k = c_k \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{3.2}$$

Of course there are many solutions to this problem, and it is well known that the maximum entropy solution is the one which maximizes the entropy gain

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \Phi(e^{i\theta}) d\theta \tag{3.3}$$

(see, e.g., [19]), and we shall now consider this constrained optimization problem.

We begin by setting up the appropriate spaces. Recall from classical realization theory that a rational function

$$v(z) = \frac{1}{2}\hat{c}_0 + \hat{c}_1 z^{-1} + \hat{c}_2 z^{-2} + \cdots$$

of degree n has a representation

$$\hat{c}_k = h' F^{k-1} g \quad k = 1, 2, 3, \ldots$$

for some choice of $(F, g, h) \in \mathbb{R}^{n\times n} \times \mathbb{R}^n \times \mathbb{R}^n$. Therefore, if in addition $v(z)$ is strictly positive real, implying that all eigenvalues of F are less than one in modulus, $\hat{c}_k$ tends exponentially to zero as $k \to \infty$. Hence, in particular,

$$\hat{c} := (\hat{c}_0, \hat{c}_1, \hat{c}_2, \ldots)$$

must belong to ℓ_1. Moreover, the requirement that (3.1) be a coercive spectral density adds another constraint, namely that $\hat{c}$ belongs to the set

$$\mathcal{F} := \left\{ \hat{c} \in \ell_1 \mid \hat{c}_0 + \sum_{k=1}^{\infty} \hat{c}_k (e^{ik\theta} + e^{-ik\theta}) > 0 \right\}. \tag{3.4}$$

Now, let

$$\psi(\hat{c}) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left[\hat{c}_0 + \sum_{k=1}^{\infty} \hat{c}_k (e^{ik\theta} + e^{-ik\theta}) \right] d\theta \tag{3.5}$$

be a functional $\mathcal{F} \to \mathbb{R}$, and consider the infinite-dimensional convex constrained optimization problem to minimize $\psi(\hat{c})$ over $\mathcal{F}$ given the finite number of constraints (3.2). Thus we have relaxed the optimization problem to allow also for nonrational spectral densities.

Since the optimization problem is convex, the Lagrange function

$$L(\hat{c}, \lambda) = \psi(\hat{c}) + \sum_{k=0}^{n} \lambda_k (\hat{c}_k - c_k) \tag{3.6}$$

has a saddle point [26, p. 458] provided the stationary point lies in the interior of $\mathcal{F}$, and, in this case, the optimal Lagrange vector $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_n) \in \mathbb{R}^{n+1}$ can be determined by solving the *dual problem* to maximize

$$\rho(\lambda) = \min_{\hat{c} \in \mathcal{F}} L(\hat{c}, \lambda). \tag{3.7}$$

To this end, first note that

$$\frac{\partial L}{\partial \hat{c}_k} = -\frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{ik\theta} + e^{-ik\theta}) \Phi^{-1}(e^{i\theta}) d\theta + \lambda_k \quad \text{for } k = 0, 1, 2, \ldots, n, \tag{3.8}$$

and that

$$\frac{\partial L}{\partial \hat{c}_k} = -\frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{ik\theta} + e^{-ik\theta}) \Phi^{-1}(e^{i\theta}) d\theta \quad \text{for } k = n+1, n+2, \ldots. \tag{3.9}$$

Then, setting the gradient equal to zero, we obtain from (3.9) that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{ik\theta} + e^{-ik\theta}) \Phi^{-1}(e^{i\theta}) d\theta = 0 \quad \text{for } |k| > n,$$

from which it follows that Φ^{-1} must be a pseudopolynomial

$$Q(z) = q_0 + \frac{1}{2} q_1 (z + z^{-1}) + \cdots + \frac{1}{2} q_n (z^n + z^{-n}) \tag{3.10}$$

of degree at most n, i.e.,

$$\Phi^{-1}(z) = Q(z), \tag{3.11}$$

yielding a spectral density Φ which is rational and of at most degree $2n$, and thus belongs to the original (nonrelaxed) class of spectral densities. Likewise we obtain from (3.8) that

$$\lambda_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{ik\theta} + e^{-ik\theta}) \Phi^{-1}(e^{i\theta}) d\theta \tag{3.12}$$

for $k = 0, 1, 2, \dots, n$, which together with (3.11) yields

$$\lambda_k = q_k \quad \text{for } k = 0, 1, 2, \dots, n. \tag{3.13}$$

However, the minimizing $\hat{c}$ is given by

$$\hat{c}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{2}(e^{ik\theta} + e^{-ik\theta})Q(e^{i\theta})^{-1}d\theta \tag{3.14}$$

and consequently

$$\sum_{k=0}^{n} q_k \hat{c}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} Q(e^{i\theta})Q(e^{i\theta})^{-1}d\theta = 1. \tag{3.15}$$

To determine the optimal (saddle point) Lagrange multipliers we turn to the dual problem. In view of (3.11), (3.13), and (3.15), the dual function is

$$\rho(q) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log Q(e^{i\theta})d\theta + 1 - c'q,$$

where $c \in \mathbb{R}^{n+1}$ is the vector with components $c_0, c_1, \dots, c_n$. Consequently, the dual problem is equivalent to minimizing

$$\varphi(q) = c'q - \frac{1}{2\pi}\int_{-\pi}^{\pi} \log Q(e^{i\theta})d\theta \tag{3.16}$$

over all $q \in \mathbb{R}^{n+1}$ such that the pseudopolynomial (3.10) is nonnegative on the unit circle, i.e.,

$$Q(e^{i\theta}) > 0 \quad \text{for all } \theta, \tag{3.17}$$

and, if the dual problem has an optimal solution satisfying (3.17), the optimal Q solves the primal problem when inserted into (3.11).

The dual problem to minimize (3.16) given (3.17) is a finite-dimensional convex optimization problem, which is simpler than the original (primal) problem. Clearly it is a special case of the optimization problem (2.18)–(2.19), obtained by setting $|\sigma(e^{i\theta})|^2 = 1$ as required for the maximum entropy solution. Figure 3.1 depicts a typical cost function φ in the case $n = 1$. As can be seen, it is convex and attains its optimum in an interior point so that the spectral density Φ has all its poles in the open unit disc as required. That this is the case in general will be proven in section 5.

We stress again that the purpose of this section is not primarily to derive an algorithm for the maximum entropy solution, for which we already have the simple Levinson algorithm, but to motivate an algorithm for the case with prescribed zeros in the spectral density. This is the topic of the next two sections.

4. The general convex optimization problem. Given a partial covariance sequence $c = (c_0, c_1, \dots, c_n)'$ and a Schur polynomial $\sigma(z)$, we know from section 2 that there exists a Schur polynomial

$$a(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n \quad (a_0 > 0)$$

such that

$$\Phi(z) = \frac{\sigma(z)\sigma(z^{-1})}{a(z)a(z^{-1})} = c_0 + \sum_{k=1}^{\infty} \hat{c}_k(z^k + z^{-k}), \tag{4.1}$$

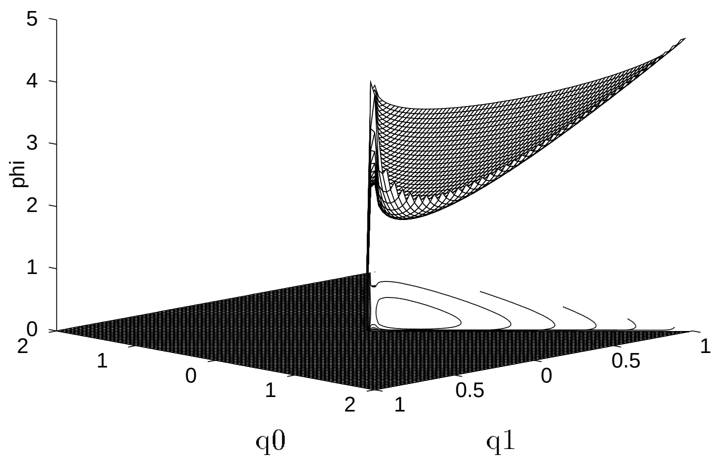

FIG. 3.1. *A typical cost function $\varphi(q)$ in the case $n = 1$.*

where

$$\hat{c}_k = c_k \quad \text{for } k = 1, 2, \ldots, n. \tag{4.2}$$

The question now is: How do we find $a(z)$? In this section, we shall construct a nonlinear, strictly convex functional on a closed convex domain. In the next section, we shall show that this functional always has a unique minimum and that if such a minimum occurs as an interior point, it gives rise to $a(z)$.

As seen from (2.3), the interpolation condition (4.2) may be written

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\frac{|\sigma(e^{i\theta})|^2}{Q(e^{i\theta})}d\theta \quad \text{for } k = 0, 1, \ldots, n, \tag{4.3}$$

where

$$Q(z) = a(z)a(z^{-1}), \tag{4.4}$$

so the problem is reduced to determining the variables

$$q = \begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_n \end{bmatrix} \in \mathbb{R}^{n+1} \tag{4.5}$$

in the pseudopolynomial

$$Q(z) = q_0 + \frac{1}{2}q_1(z + z^{-1}) + \frac{1}{2}q_2(z^2 + z^{-2}) + \cdots + \frac{1}{2}q_n(z^n + z^{-n}) \tag{4.6}$$

so that the conditions (4.3) and

$$Q(e^{i\theta}) > 0 \quad \text{for all } \theta \in [-\pi, \pi] \tag{4.7}$$

are satisfied.

Now, consider the convex functional $\varphi(q) : \mathbb{R}^{n+1} \to \mathbb{R}$ defined by

$$\varphi(q) = c'q - \frac{1}{2\pi}\int_{-\pi}^{\pi} \log Q(e^{i\theta})|\sigma(e^{i\theta})|^2 d\theta. \tag{4.8}$$

Our motivation in defining $\varphi(q)$ comes in part from the desire to introduce a barrier-like term, as is done in interior point methods, and in part from our analysis of the maximum entropy method in the previous section. As it turns out, by a theorem of Szegö the logarithmic integrand is in fact integrable for nonzero Q having zeros on the boundary of the unit circle, so that $\varphi(q)$ does not become infinite on the boundary of the convex set. On the other hand, $\varphi(q)$ is a natural generalization of the functional (3.16) in section 3, since it specializes to (3.16) when $|\sigma(e^{i\theta})|^2 \equiv 1$ as for the maximum entropy solution. As we shall see, minimizing (4.8) yields precisely via (4.4) the unique $a(z)$ which corresponds to $\sigma(z)$.

It is clear that if $q \in \mathcal{D}_n^+$, where

$$\mathcal{D}_n^+ = \{q \in \mathbb{R}^{n+1} \mid Q(z) > 0 \quad \text{for } |z| = 1\}, \tag{4.9}$$

then $\varphi(q)$ is finite. Moreover, $\varphi(q)$ is also finite when $Q(z)$ has finitely many zeros on the unit circle, as can be seen from the following lemma.

LEMMA 4.1. *The functional $\varphi(q)$ is finite and continuous at any $q \in \overline{\mathcal{D}_n^+}$ except at zero. The functional is infinite, but continuous, at $q = 0$. Moreover, φ is a C^∞ function on $\mathcal{D}_n^+$.*

Proof. We want to prove that $\varphi(q)$ is finite when $q \neq 0$. Then the rest follows by inspection. Clearly, $\varphi(q)$ cannot take the value $-\infty$; hence, it remains to prove that $\varphi(q) < \infty$. Since $q \neq 0$,

$$\mu := \max_{\theta} Q(e^{i\theta}) > 0.$$

Then setting $P(z) := \mu^{-1}Q(z)$,

$$\log P(e^{i\theta}) \leq 0 \tag{4.10}$$

and

$$\varphi(q) = c'q - \frac{1}{2\pi}\log\mu\int_{-\pi}^{\pi}|\sigma(e^{i\theta})|^2 d\theta - \frac{1}{2\pi}\int_{-\pi}^{\pi}\log P(e^{i\theta})|\sigma(e^{i\theta})|^2 d\theta,$$

and hence, the question of whether $\varphi(q) < \infty$ is reduced to determining whether

$$-\int_{-\pi}^{\pi}\log P(e^{i\theta})|\sigma(e^{i\theta})|^2 d\theta < \infty.$$

However, since $|\sigma(e^{i\theta})|^2 \leq M$ for some bound M, this follows from

$$\int_{-\pi}^{\pi}\log P(e^{i\theta}) d\theta > -\infty, \tag{4.11}$$

which is the well-known Szegö condition: (4.11) is a necessary and sufficient condition for $P(e^{i\theta})$ to have a stable spectral factor [17]. However, since $P(z)$ is a symmetric pseudopolynomial which is nonnegative on the unit circle, there is a polynomial $\pi(z)$ such that $\pi(z)\pi(z^{-1}) = P(z)$. But then $w(z) = \frac{\pi(z)}{z^n}$ is a stable spectral factor, and hence (4.11) holds. ∎

LEMMA 4.2. *The functional $\varphi(q)$ is strictly convex and defined on a closed, convex domain.*

Proof. We first note that $q = 0$ is an extreme point, but it can never be a minimum of φ since $\varphi(0)$ is infinite. In particular, in order to check the strict inequality

$$\varphi(\lambda q^{(1)} + (1-\lambda)q^{(2)}) < \lambda\varphi(q^{(1)}) + (1-\lambda)\varphi(q^{(2)}), \tag{4.12}$$

where one of the arguments is zero, we need only consider the case that either $q^{(1)}$ or $q^{(2)}$ is zero, in which case the strict inequality holds. We can now assume that none of the arguments is zero, in which case the strict inequality in (4.12) follows from the strict concavity of the logarithm. Finally, it is clear that $\overline{\mathcal{D}_n^+}$ is a closed convex subset. □

LEMMA 4.3. *Let $q \in \overline{\mathcal{D}_n^+}$, and suppose $q \neq 0$. Then $c'q > 0$.*

Proof. Consider an arbitrary covariance extension of c such as, for example, the maximum entropy extension, and let $\Phi(z)$ be the corresponding spectral density (2.9). Then c is given by (2.3), which may also be written

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

Therefore, in view of (4.6),

$$c'q = \frac{1}{2\pi}\int_{-\pi}^{\pi} Q(e^{ik\theta})\Phi(e^{i\theta})d\theta, \tag{4.13}$$

which is positive whenever $Q(z) \geq 0$ on the unit circle and $q \neq 0$. □

PROPOSITION 4.4. *For all $r \in \mathbb{R}$, $\varphi^{-1}(-\infty, r]$ is compact. Thus φ is proper (i.e., $\varphi^{-1}(K)$ is compact whenever K is compact) and bounded from below.*

Proof. Suppose $q^{(k)}$ is a sequence in $M_r := \varphi^{-1}(-\infty, r]$. It suffices to show that $q^{(k)}$ has a convergent subsequence. Each $Q^{(k)}$ may be factored as

$$Q^{(k)}(z) = \lambda_k \bar{a}^{(k)}(z)\bar{a}^{(k)}(z^{-1}) = \lambda_k \bar{Q}^{(k)}(z),$$

where λ_k is positive and $\bar{a}^{(k)}(z)$ is a monic polynomial, all of whose roots lie in the closed unit disc. The corresponding sequence of the (unordered) set of n roots of each $\bar{a}^{(k)}(z)$ has a convergent subsequence, since all (unordered) sets of roots lie in the closed unit disc. Denote by $\bar{a}(z)$ the monic polynomial of degree n which vanishes at this limit set of roots. By reordering the sequence if necessary, we may assume the sequence $a^{(k)}(z)$ tends to $\bar{a}(z)$. Therefore, the sequence $q^{(k)}$ has a convergent subsequence if and only if the sequence λ_k does, which will be the case provided the sequence λ_k is bounded from above and from below away from zero. Before proving this, we note that the sequences $c'\bar{q}^{(k)}$, where $\bar{q}^{(k)}$ is the vector corresponding to the pseudopolynomial $\bar{Q}^{(k)}$, and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log \bar{Q}^{(k)}(e^{i\theta})|\sigma(e^{i\theta})|^2 d\theta \tag{4.14}$$

are both bounded from above and from below, respectively, away from zero and $-\infty$. The upper bounds come from the fact that $\{\bar{a}^{(k)}(z)\}$ are Schur polynomials and hence have their coefficients in the bounded Schur region. As for the lower bound of $c'\bar{q}^{(k)}$, note that $c'\bar{q}^{(k)} > 0$ for all k (Lemma 4.3) and $c'\bar{q}^{(k)} \to \alpha > 0$. In fact, $\bar{Q}^{(k)}(e^{i\theta}) \to |\bar{a}(e^{i\theta})|^2$, where $\bar{a}(z)$ has all its zeros in the closed unit disc, and hence it follows from (4.13) that $\alpha > 0$. Then, since $\varphi(q) < \infty$ for all $q \in \overline{\mathcal{D}_n^+}$ except $q = 0$ (Lemma 4.1), (4.14) is bounded away from $-\infty$. Next, observe that

$$\varphi(q^{(k)}) = \lambda_k c' \bar{q}^{(k)} - \frac{1}{2\pi} \log \lambda_k \int_{-\pi}^{\pi} |\sigma(e^{i\theta})|^2 d\theta - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \bar{Q}^{(k)}(e^{i\theta}) |\sigma(e^{i\theta})|^2 d\theta.$$

From this we can see that if a subsequence of λ_k were to tend to zero, then $\varphi(q^{(k)})$ would exceed r. Likewise, if a subsequence of λ_k were to tend to infinity, φ would exceed r, since linear growth dominates logarithmic growth. ∎

5. Interior critical points and solutions of the rational covariance extension problem. In the previous section, we showed that φ has compact sublevel sets in $\overline{\mathcal{D}_n^+}$, so that φ achieves a minimum. Moreover, since φ is strictly convex and $\overline{\mathcal{D}_n^+}$ is convex, such a minimum is unique. We record these observations in the following statement.

PROPOSITION 5.1. *For each partial covariance sequence c and each Schur polynomial $\sigma(z)$, the functional φ has a unique minimum on $\overline{\mathcal{D}_n^+}$.*

In this paper we consider a question which is of independent interest: whether φ achieves its minimum at an interior point. The next result describes an interesting systems-theoretic consequence of the existence of such interior minima.

THEOREM 5.2. *Fix a partial covariance sequence c and a Schur polynomial $\sigma(z)$. If $\hat{q} \in \mathcal{D}_n^+$ is a minimum for φ, then*

$$\hat{Q}(z) = a(z)a(z^{-1}), \tag{5.1}$$

where $a(z)$ is the solution of the rational covariance extension problem.

Proof. Suppose that $\hat{q} \in \mathcal{D}_n^+$ is a minimum for φ. Then

$$\frac{\partial \varphi}{\partial q_k}(\hat{q}) = 0 \quad \text{for } k = 0, 1, 2, \ldots, n. \tag{5.2}$$

Differentiating inside the integral, which is allowed due to uniform convergence, (5.2) yields

$$c_k - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2}(e^{ik\theta} + e^{-ik\theta}) \frac{|\sigma(e^{i\theta})|^2}{\hat{Q}(e^{i\theta})} d\theta = 0 \quad \text{for } k = 0, 1, \ldots, n,$$

where $\hat{Q}(z)$ is the pseudopolynomial (4.6) corresponding to $\hat{q}$, or, equivalently,

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \frac{|\sigma(e^{i\theta})|^2}{\hat{Q}(e^{i\theta})} d\theta \quad \text{for } k = 0, 1, \ldots, n, \tag{5.3}$$

which is precisely the interpolation condition (4.3)–(4.4), provided (5.1) holds. ∎

As a corollary of this theorem, we have that the gradient of φ at *any* $\tilde{q} \in \mathcal{D}_n^+$ is given by

$$\frac{\partial \varphi}{\partial q_k}(\tilde{q}) = c_k - \tilde{c}_k, \tag{5.4}$$

where

$$\tilde{c}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \frac{|\sigma(e^{i\theta})|^2}{\tilde{Q}(e^{i\theta})} d\theta, \quad k = 0, 1, 2, \ldots, n \tag{5.5}$$

is the partial covariance sequence corresponding to a process with spectral density

$$\tilde{\Phi}(e^{i\theta}) = \frac{|\sigma(e^{i\theta})|^2}{\tilde{Q}(e^{i\theta})},$$

where $\tilde{Q}(z)$ is the pseudopolynomial corresponding to $\tilde{q}$. The gradient is thus the difference between the true and calculated partial covariance sequences.

We now state the converse result, underscoring our interest in this particular convex optimization problem.

THEOREM 5.3. *For each partial covariance sequence c and each Schur polynomial $\sigma(z)$, suppose that $a(z)$ gives a solution to the rational covariance extension problem. If*

$$\hat{Q}(z) = a(z)a(z^{-1}), \tag{5.6}$$

then the corresponding $(n+1)$-vector $\hat{q}$ lies in $\mathcal{D}_n^+$ and is a unique minimum for φ.

Proof. Let $a(z)$ be the solution of the rational covariance extension problem corresponding to c and $\sigma(z)$, and let $\hat{Q}(z)$ be given by (5.6). Then c satisfies the interpolation condition (5.3), which is equivalent to (5.2), as seen from the proof of Theorem 5.2. However, since $a(z)$ is a Schur polynomial, $\hat{Q}(z) > 0$ on the unit circle, and thus $\hat{q} \in \mathcal{D}_n^+$. Since φ is strictly convex on $\mathcal{D}_n^+$, (5.3) implies that $\hat{q}$ is a unique minimum for φ. □

Since the existence of a solution to the rational covariance extension problem has been established in [14] (see also [7]), we do in fact know the existence of interior minima for this convex optimization problem. On the other hand, we know from Proposition 5.1 that φ has a minimum for some $\hat{q} \in \overline{\mathcal{D}_n^+}$, so to show that φ has a minimum in the interior $\mathcal{D}_n^+$ it remains to prove the following lemma.

LEMMA 5.4. *The functional φ never attains a minimum on the boundary $\partial\mathcal{D}_n^+$.*

Proof. Denoting by $D_p\varphi(q)$ the directional derivative of φ at q in the direction p, it is easy to see that

$$D_p\varphi(q) := \lim_{\epsilon \to 0} \frac{\varphi(q+\epsilon p) - \varphi(q)}{\epsilon} \tag{5.7}$$

$$= c'p - \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{P(e^{i\theta})}{Q(e^{i\theta})}|\sigma(e^{i\theta})|^2 d\theta, \tag{5.8}$$

where $P(z)$ is the pseudopolynomial

$$P(z) = p_0 + \frac{1}{2}p_1(z+z^{-1}) + \frac{1}{2}p_2(z^2+z^{-2}) + \cdots + \frac{1}{2}p_n(z^n+z^{-n})$$

corresponding to the vector $p \in \mathbb{R}^{n+1}$. In fact,

$$\frac{\log(Q+\epsilon P) - \log Q}{\epsilon} = \frac{P}{Q}\log\left[\left(1+\epsilon\frac{P}{Q}\right)^{\frac{1}{\epsilon}\frac{Q}{P}}\right] \to \frac{P}{Q}$$

as $\epsilon \to +0$, and hence (5.7) follows by dominated convergence.

Now, let $q \in \mathcal{D}_n^+$ and $\bar{q} \in \partial\mathcal{D}_n^+$ be arbitrary. Then the corresponding pseudopolynomials Q and $\bar{Q}$ have the properties

$$Q(e^{i\theta}) > 0 \quad \text{for all } \theta \in [-\pi, \pi]$$

and

$$\bar{Q}(e^{i\theta}) \geq 0 \quad \text{for all } \theta \text{ and } \bar{Q}(e^{i\theta_0}) = 0 \text{ for some } \theta_0.$$

Since $q_\lambda := \bar{q} + \lambda(q - \bar{q}) \in \mathcal{D}_n^+$ for $\lambda \in (0, 1]$, we also have for $\lambda \in (0, 1]$ that

$$Q_\lambda(e^{i\theta}) := \bar{Q}(e^{i\theta}) + \lambda[Q(e^{i\theta}) - \bar{Q}(e^{i\theta})] > 0 \quad \text{for all } \theta \in [-\pi, \pi],$$

and we may form the directional derivative

$$D_{\bar{q}-q}\varphi(q_\lambda) = c'(\bar{q} - q) + \frac{1}{2\pi}\int_{-\pi}^{\pi} h_\lambda(\theta)d\theta, \tag{5.9}$$

where

$$h_\lambda(\theta) = \frac{Q(e^{i\theta}) - \bar{Q}(e^{i\theta})}{Q_\lambda(e^{i\theta})}|\sigma(e^{i\theta})|^2.$$

Now,

$$\frac{d}{d\lambda}h_\lambda(\theta) = \frac{[Q(e^{i\theta}) - \bar{Q}(e^{i\theta})]^2}{Q_\lambda(e^{i\theta})^2}|\sigma(e^{i\theta})|^2 \geq 0,$$

and hence $h_\lambda(\theta)$ is a monotonically nondecreasing function of λ for all $\theta \in [-\pi, \pi]$. Consequently, h_λ tends pointwise to h_0 as $\lambda \to 0$. Therefore,

$$\int_{-\pi}^{\pi} h_\lambda(\theta)d\theta \to +\infty \quad \text{as } \lambda \to 0. \tag{5.10}$$

In fact, if

$$\int_{-\pi}^{\pi} h_\lambda(\theta)d\theta \to \alpha < \infty \quad \text{as } \lambda \to 0, \tag{5.11}$$

then $\{h_\lambda\}$ is a Cauchy sequence in $L^1(-\pi, \pi)$ and hence has a limit in $L^1(-\pi, \pi)$ which must equal h_0 almost everywhere. However, h_0, having poles in $[-\pi, \pi]$, is not summable and hence, as claimed, (5.11) cannot hold.

Consequently, by virtue of (5.9),

$$D_{q-\bar{q}}\varphi(q_\lambda) \to +\infty \quad \text{as } \lambda \to 0$$

for all $q \in \mathcal{D}_n^+$ and $\bar{q} \in \partial\mathcal{D}_n^+$, and hence, in view of Lemma 26.2 in [28], φ is essentially smooth. Then it follows from Theorem 26.3 in [28] that the subdifferential of φ is empty on the boundary of $\mathcal{D}_n^+$, and therefore φ cannot have a minimum there. ∎

Thus we have proven the following result.

THEOREM 5.5. *For each partial covariance sequence c and each Schur polynomial $\sigma(z)$, there exists an $(n+1)$-vector $\hat{q}$ in $\mathcal{D}_n^+$ which is a minimum for φ.*

Consequently, by virtue of Theorem 5.2, there does exist a solution to the rational covariance extension problem for each partial covariance sequence and zero polynomial $\sigma(z)$, and, in view of Theorem 5.3, this solution is unique.

These theorems have the following corollary.

COROLLARY 5.6 (Georgiou's conjecture). *For each partial covariance sequence c and each Schur polynomial $\sigma(z)$, there is a unique Schur polynomial $a(z)$ such that* (4.1) *and* (4.2) *hold.*

Hence, we have given an independent proof of the weaker version of Theorem 2.1 conjectured by Georgiou, but not of the stronger version of [7] which states that the problem is well posed in the sense that the one-to-one correspondence between $\sigma(z)$ and $a(z)$ is a diffeomorphism.

6. Some numerical examples. Given an arbitrary partial covariance sequence $c_0, c_1, \ldots, c_n$ and an arbitrary zero polynomial $\sigma(z)$, the constructive proof of Georgiou's conjecture provides algorithmic procedures for computing the corresponding unique modeling filter, which are based on the convex optimization problem to minimize the functional (2.18) over all $q_0, q_1, \ldots, q_n$ such that (2.19) holds.

In general such procedures will be based on the gradient of the cost functional φ, which, as we saw in section 5, is given by

$$\frac{\partial \varphi}{\partial q_k}(q_0, q_1, \ldots, q_n) = c_k - \bar{c}_k, \tag{6.1}$$

where

$$\bar{c}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\frac{|\sigma(e^{i\theta})|^2}{Q(e^{i\theta})}d\theta \quad \text{for } k = 0, 1, 2, \ldots, n \tag{6.2}$$

are the covariances corresponding to a process with spectral density

$$\frac{|\sigma(e^{i\theta})|^2}{Q(e^{i\theta})} = \bar{c}_0 + 2\sum_{k=1}^{\infty} \bar{c}_k \cos(k\theta). \tag{6.3}$$

The gradient is thus the difference between the given partial covariance sequence $c_0, c_1, \ldots, c_n$ and the partial covariance sequence corresponding to the choice of variables $q_0, q_1, \ldots, q_n$ at which the gradient is calculated. The minimum is attained when this difference is zero.

The following simulations have been done by Per Enqvist, using Newton's method (see, e.g., [23, 26]), which of course also requires computing the Hessian (second-derivative matrix) in each iteration. A straightforward calculation shows that the Hessian is the sum of a Toeplitz and a Hankel matrix. More precisely,

$$H_{ij}(q_0, q_1, \ldots, q_n) = \frac{1}{2}(d_{i+j} + d_{i-j}), \quad i, j = 0, 1, 2, \ldots, n, \tag{6.4}$$

where

$$d_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\frac{|\sigma(e^{i\theta})|^2}{Q(e^{i\theta})^2}d\theta \quad \text{for } k = 0, 1, 2, \ldots, 2n \tag{6.5}$$

and $d_{-k} = d_k$. Moreover, $d_0, d_1, d_2, \ldots, d_{2n}$ are the $2n+1$ first Fourier coefficients of the spectral representation

$$\frac{|\sigma(e^{i\theta})|^2}{Q(e^{i\theta})^2} = d_0 + 2\sum_{k=1}^{\infty} d_k \cos(k\theta). \tag{6.6}$$

The gradient and the Hessian can be determined from (6.1) and (6.4), respectively, by applying the inverse Levinson algorithm (see, e.g., [27]) to the appropriate polynomial spectral factors of $Q(z)$ and $Q(z)^2$, respectively, and then solving the resulting linear equations for $\bar{c}_0, \bar{c}_1, \ldots, \bar{c}_n$ and $d_0, d_1, d_2, \ldots, d_{2n}$; see [12] for details.

To illustrate the procedure, let us again consider the sixth-order spectral envelopes of Figures 2.1 and 2.2 together with the corresponding zeros and poles. Hence, Figure 6.1 illustrates the periodogram for a section of speech data together with the corresponding sixth-order maximum entropy spectrum, which, since it lacks finite zeros,

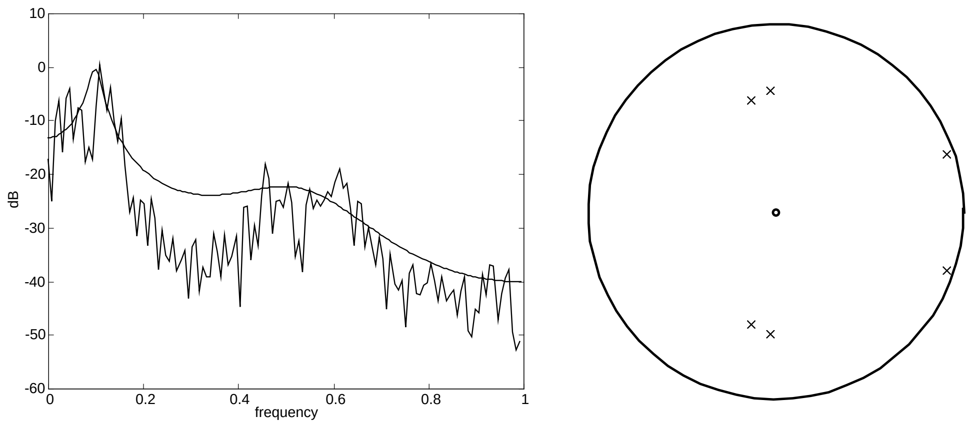

Fig. 6.1.

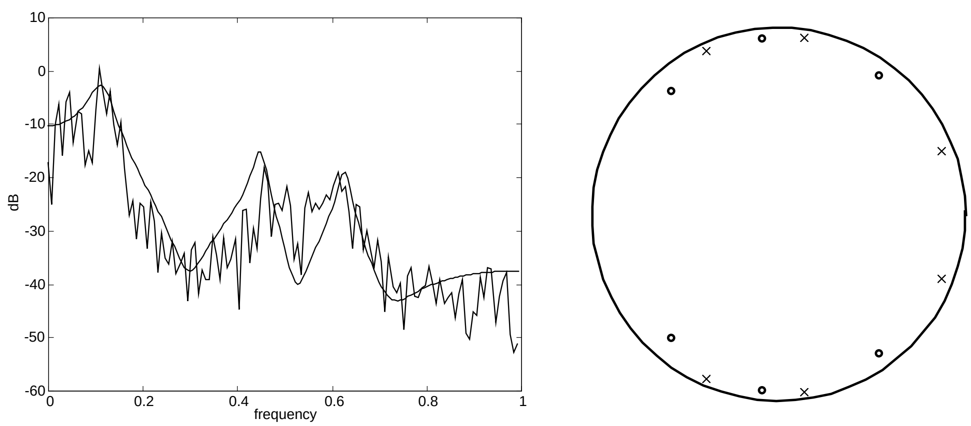

Fig. 6.2.

becomes rather "flat." The location of the corresponding poles (marked by $\times$) in the unit circle is shown next to it. The zeros (marked by $\circ$) of course all lie at the origin.

Now, selecting the zeros appropriately as indicated to the right in Figure 6.2, we obtain the poles as marked, and the corresponding sixth-order modeling filter produces the spectral envelope to the left in Figure 6.2. We see that the second solution has a spectral density that is less flat and provides a better approximation, reflecting the fact that the filter is designed to have transmission zeros near the minima of the periodogram.

7. Conclusions. In [13, 14] Georgiou proved that to each choice of partial covariance sequence and numerator polynomial of the modeling filter there exists a rational covariance extension yielding a pole polynomial for the modeling filter, and he conjectured that this extension is unique so that it provides a complete parameterization of all rational covariance extensions. In [7] we proved this long-standing conjecture in the more general context of a duality between filtering and interpolation and showed that the problem is well posed in a very strong sense. In [6] we connected this solution to a certain Riccati-type matrix equation that sheds further light on the structure of this problem.

However, our proof in [7], as well as the existence proof of Georgiou [14], is nonconstructive. In this paper we presented a constructive proof of Georgiou's conjecture, which, although it is weaker than our result in [7], provides us for the first time with an algorithm for solving the problem of determining the unique pole polynomial corresponding to the given partial covariance sequence and the desired zeros.

This is done by means of a constrained convex optimization problem, which can be solved without explicitly computing the values of the cost function and which has the interesting property that the cost function is finite on the boundary but the gradient is not. In this context, Georgiou's conjecture is equivalent to establishing that there is a unique minimum in the *interior* of the feasible region. Specialized to the maximum entropy solution, this optimization problem was seen to be a dual to the well-known problem of maximizing the entropy gain.

Acknowledgments. We would like to thank Per Enqvist and Massoud Amin for their help with simulations.

REFERENCES

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

[2] M. Aoki, *State Space Modeling of Time Series*, Springer-Verlag, Berlin, 1987.

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

[4] C. I. Byrnes and A. Lindquist, *On the geometry of the Kimura-Georgiou parameterization of modelling filter*, Internat. J. Control, 50 (1989), pp. 2301–2312.

[5] C. I. Byrnes and A. Lindquist, *Toward a solution of the minimal partial stochastic realization problem*, C. R. Acad. Sci. Paris Sér. I Math., 319 (1994), pp. 1231–1236.

[6] C. I. Byrnes and A. Lindquist, *On the partial stochastic realization problem*, IEEE Trans. Automat. Control, AC-42 (1997), pp. 1049–1070.

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

[8] J. A. Cadzow, *Spectral estimation: An overdetermined rational model equation approach*, Proc. IEEE, 70 (1982), pp. 907–939.

[9] C. Carathéodory, *Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen*, Math. Ann., 64 (1907), pp. 95–115.

[10] C. Carathéodory, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Functionen*, Rend. Circ. Mat. Palermo (2), 32 (1911), pp. 193–217.

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

[12] P. Enqvist, forthcoming Ph.D. dissertation, Division of Optimization and Systems Theory, Royal Institute of Technology, Stockholm, Sweden.

[13] T. T. Georgiou, *Partial Realization of Covariance Sequences*, Ph.D. thesis, Center for Mathematical Systems Theory, University of Florida, Gainesville, FL, 1983.

[14] T. T. Georgiou, *Realization of power spectra from partial covariance sequences*, IEEE Trans. Acoust. Speech Signal Process., ASSP-35 (1987), pp. 438–449.

[15] Y. L. Geronimus, *Orthogonal Polynomials*, Consultants Bureau, New York, 1961.

[16] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, CA, 1958.

[17] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*, Almqvist & Wiksell, Stockholm, 1956.

[18] R. E. Kalman, *Realization of covariance sequences*, in Proc. Toeplitz Centennial, Tel Aviv, Israel, 1981, Oper. Theory Adv. Appl. 4, Birkhäuser, Basel, 1982, pp. 331–342.

[19] S. A. Kassam and H. V. Poor, *Robust techniques for signal processing*, Proc. IEEE, 73 (1985), pp. 433–481.

[20] S. M. Kay and S. L. Marple, Jr., *Spectrum analysis—A modern perspective*, Proc. IEEE, 69 (1981), pp. 1380–1419.
[21] H. Kimura, *Positive partial realization of covariance sequences*, in Modelling, Identification and Robust Control, C. I. Byrnes and A. Lindquist, eds., North–Holland, Amsterdam, 1987, pp. 499–513.
[22] A. Lindquist and G. Picci, *Canonical correlation analysis, approximate covariance extension, and identification of stationary time series*, Automatica J. IFAC, 32 (1996), pp. 709–733.
[23] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed., Addison–Wesley, Reading, MA, 1984.
[24] J. D. Markel and A. H. Gray, *Linear Prediction of Speech*, Springer-Verlag, Berlin, 1976.
[25] J. Makhoul, *Linear prediction: A tutorial review*, Proc. IEEE, 63 (1975), pp. 561–580.
[26] M. Minoux, Jr., *Mathematical Programming: Theory and Algorithms*, John Wiley, New York, 1986.
[27] B. Porat, *Digital Processing of Random Signals*, Prentice–Hall, Englewood Cliffs, NJ, 1994.
[28] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[29] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice–Hall, Englewood Cliffs, NJ, 1978.
[30] I. Schur, *On power series which are bounded in the interior of the unit circle* I *and* II, J. Reine Angew. Math., 148 (1918), pp. 122–145.
[31] O. Toeplitz, *Über die Fouriersche Entwicklung positiver Funktionen*, Rend. Circ. Mat. Palermo (2), 32 (1911), pp. 191–192.
[32] P. van Overschee and B. De Moor, *Subspace algorithms for stochastic identification problem*, IEEE Trans. Automat. Control, AC-27 (1982), pp. 382–387.