SIAM J. CONTROL AND OPTIMIZATION
Vol. 14, No. 4, July 1976

A SPECTRAL FACTORIZATION APPROACH TO THE DISTRIBUTED STABLE REGULAR PROBLEM; THE ALGEBRAIC RICCATI EQUATION*

J. WILLIAM HELTON†

Abstract. This paper is a study of the discrete-time infinite-dimensional "stable regulator problem" having a cost function which is not necessarily positive. We take a spectral factorization approach to the problem. Also there are results on the algebraic Riccatic equation which are equivalent to results about fixed points for a broad class of symplectic maps.

Introduction. This paper is a study of the infinite-dimensional "stable regulator problems" having a cost functional which is not necessarily positive. The control problem will have a solution or approximate solution in feedback form provided that one "completes" a certain square in a way familiar in control theory. In this paper, we use a spectral factorization method to obtain necessary and sufficient conditions for this to be possible (§ 2). Section 3 describes the stability of the feedback system resulting from the optimal control problem. Section 4 treats the infinite-dimensional algebraic Riccati equation associated with the control problem. This can also be described as a study of the fixed-point problem for certain infinite-dimensional "symplectic" maps (see Appendix plus § 4).

This paper follows in the footsteps of a paper by Willems [25] in which he gives necessary and sufficient conditions for solving a broad class of finite-dimensional continuous time algebraic Riccati equations. In addition to giving a discrete-time and an infinite-dimensional version of these results, our article gives proofs which in finite dimensions are rather simple. As this paper was being written, an elegant spectral factorization approach to Willems results was given by Molinari [15][16] and then applied to the stable regulator in [17]. His proof involves some basically finite-dimensional methods such as determinants and dimension counting while the key step in the proof here is subspace inclusion. The article [14] is a good reference for infinite-dimensional discrete-time systems having "positive cost operators". Our article gives a new approach to the time-invariant regulator results in that paper and extends them in several directions.

The results in this article apply to most least squares problems associated with the discretization of systems governed by a heavily damped variable coefficient wave equation (including the heat equation). A thorough list of applications of the finite-dimensional theory appears in [25]. Since we do not require our "cost operators" to be positive, the systems studied are capable of storing energy, that is, "cost". The basic principal which emerges in [25] and which is true to a large extent in infinite dimensions is that one can use the standard feedback approach to a control problem provided the zero state stores no energy. Roughly speaking,

---

* Received by the editors May 3, 1974, and in revised form September 25, 1975.

† Department of Mathematics, University of California, San Diego, La Jolla, California 92037.
This work was supported in part by the National Science Foundation.

conventional approaches suffice even when the system can store energy, but not when it can spontaneously produce energy.

1. Definitions and setting. We shall consider a system

$$x_{i+1} = Ax_i + Bu_i, \tag{1.1}$$

where the x_j are vectors in a Hilbert space $\mathcal{H}$ and the u_j are vectors in a Hilbert space $\mathcal{U}$. The cost of running the system from initial state x_0 for N time units is

$$J_N(x_0, u) = \sum_{i=0}^{N} [(x_i, Qx_i) + (u_i, Ru_i)]. \tag{1.2}$$

Here $Q : \mathcal{H} \to \mathcal{H}$ and $R : \mathcal{U} \to \mathcal{U}$ are bounded self-adjoint operators. The basic infinite time interval problem is: given a state x_0, determine exact or approximate a control sequence u which minimizes $J_\infty(x_0, \ )$. The admissible class of control sequences $u = \langle u_i \rangle_{i=1}^{\infty}$ in this paper will be those in $l^2(0, \infty, \mathcal{U})$, the set of all sequences from $\mathcal{U}$ whose norms are square summable. Also one usually requires that $x_n \to 0$ in some sense as $n \to \infty$. We shall always assume that A is stable, i.e., $\|A^n\| < M$ for all n, and that B is bounded.

Frequently in what follows it will be convenient to look at our system as one having an output. The natural choice for the output operator is $|Q|^{1/2}$. Thus the problem is equivalent to minimizing the cost

$$\sum (y_i, [\operatorname{sgn} Q] y_i) + (u_i, Ru_i) \tag{1.3}$$

of running the system

$$x_{i+1} = Ax_i + Bu_i, \qquad y_i = |Q|^{1/2} x_i. \tag{1.4}$$

Here sgn Q is the operator $P_+ - P_-$ on $\mathcal{H}$, where $P_+(P_-)$ is the projection onto the positive (negative) spectral subspace of Q. The frequency response function for the system $[A, B, |Q|^{1/2}]$ is

$$W(z) = z|Q|^{1/2}(I - zA)^{-1}B. \tag{1.5}$$

Since A is stable, the spectrum of A is contained inside the disk, and so $W(z)$ is well-defined and analytic inside the disk. It will be assumed that all systems we study have a uniformly bounded frequency response function. Let $l^2(0, \infty, \mathcal{U})$ denote all sequences $\langle u_i \rangle_{i=0}^{\infty}$ from $\mathcal{U}$ with square summable norm.

If $Q \geqq 0$, then it can be shown (see equations (2.2) and (2.3)) that the cost $J_\infty(0, u)$ of running the system initially at state 0 with l^2 input u is finite if and only if the frequency response function W is uniformly bounded on the disk. In dealing with the signed Q problem, we shall *always* assume that the cost $J_\infty(0, u)$ with $|Q|$ replacing Q is finite. This is also equivalent to the statement $W(z)$ is uniformly bounded on the unit disk, and we shall say that any (1.1) and (1.2) with this property have *absolutely finite cost.* This assumption will obviously be satisfied when A is very stable, for example, if (1.1) arises from discretizing a variable coefficient heat equation or heavily damped wave equation. This assumption can certainly be relaxed but the author suspects that the basic structure and proofs will change little while the technical complication will greatly increase. Thus it seems unwise to do so without first making a systematic list of compelling examples.

Henceforth, assume that (1.1) and (1.2) have absolutely finite cost. It is well known (see, [18, Chap. V]) that radial limits of such functions exist almost everywhere onto the unit circle, so we may consider W as being a function $W(e^{i\theta})$ defined for almost all θ.

The mathematical notation will be as follows. The unit circle will be denoted by $\mathbf{T}$, the set of all complex numbers by $\mathbb{C}$. If $\mathcal{H}$ is a separable complex Hilbert space, then $L^2(\mathcal{H})$ denotes the Hilbert space of norm square integrable Lebesgue-measurable $\mathcal{H}$-valued functions. We let $H^2(\mathcal{H})$ [resp., $\bar{H}^2(\mathcal{H})$] denote the closed subspace of functions in $L^2(\mathcal{H})$ with zero nonpositive (positive) Fourier coefficients. The operator P_{H^2} [resp., $P_{\bar{H}^2}$] is the orthogonal projection of L^2 onto this subspace. If $\mathcal{H}_1$ and $\mathcal{H}_2$ are separable complex Hilbert spaces, then $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ denotes the Banach space of bounded linear transformations from $\mathcal{H}_1$ to $\mathcal{H}_2$. We abbreviate $\mathcal{L}(\mathcal{H}, \mathcal{H})$ as $\mathcal{L}(\mathcal{H})$. Moreover $L^\infty(\mathcal{H}_1, \mathcal{H}_2)$ denotes the Banach space of essentially bounded weakly-measurable $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$-valued functions on $\mathbf{T}$, while $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ (resp., $\bar{H}^\infty(\mathcal{H}_1, \mathcal{H}_2)$) denotes the subspace of functions with negative (resp., nonnegative) Fourier coefficients equal to zero. When the context prevents ambiguities, we will only write H^∞ and H^2.

Functions in the Hardy spaces $H^2(\mathcal{H})$ [resp., $\bar{H}^2(\mathcal{H})$ and $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$] can be identified with boundary values of functions analytic inside [resp., outside] the unit disk; see [18, Chap. V]. If φ is a function in $L^\infty(\mathcal{H}_1, \mathcal{H}_2)$, then $\mathcal{M}_\varphi$ is the operator from $L^2(\mathcal{H}_1)$ to $L^2(\mathcal{H}_2)$ defined by $(\mathcal{M}_\varphi f)(e^{i\theta}) = \varphi(e^{i\theta})f(e^{i\theta})$ for f in $L^2(\mathcal{H}_1)$. A function φ in $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ is called *outer* provided that $\mathcal{M}_\varphi$ restricted to $H^2(\mathcal{H}_1)$ has dense range in $H^2(\mathcal{H}_2)$. A function φ in $\bar{H}^\infty(\mathcal{H}_1, \mathcal{H}_2)$ is called *conjugate* outer provided that it has the analogous properties on $\bar{H}^2(\mathcal{H}_1)$ and $\bar{H}^2(\mathcal{H}_2)$. Such a function is called *invertible* outer if its pointwise inverse is uniformly bounded. The invertible outer functions are precisely those in $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ whose inverse is in $H^\infty(\mathcal{H}_2, \mathcal{H}_1)$.

2. Optimality. In an infinite-dimensional problem, it is reasonable to expect the frequent occurrence of unbounded operators. This is so here. In fact, more unwieldy objects are necessary. For example, the cost of optimally driving a state x_0 to zero may not be finite on all states x_0 of the system, while it will frequently be finite for all states which actually occur in the running of the system. Thus it is only reasonable to expect the optimal cost functional to be a densely defined (quadratic) functional on the state space, and indeed that is what will be obtained. A good reference on such objects is [20, Chap. VIII]. The controllability map of a system $[A, B]$ is the densely defined map $\mathcal{C} : l^2(0, \infty, \mathcal{U}) \to \mathcal{H}$ given by

$$\mathcal{C}(u_0, u_1, \cdots) = \sum_{k=0}^{\infty} A^k B u_k.$$

We set $\mathcal{R} = \text{range } \mathcal{C}$ and $\tilde{\mathcal{R}} = \mathcal{C}$ (all sequences with only finite number of nonzero terms). These are domains which will be commonly used.

A trajectory of the system initially at 0 is a sequence of states $\{x_n\}_{n=0}^\infty$ which results from feeding some input sequences $\{u_i\}$ into the system. Unless otherwise specified, *trajectory* will refer to something arising from a $l^2(0, \infty, \mathcal{U})$ input. The finite cost assumption implies that $J_\infty(0, u)$ is well-defined and finite for all u in $l^2(\mathcal{U})$. In this section, we shall deal only with J_N for which $J_\infty(0, u) > 0$, all u in

$l^2(\mathscr{U})$. In this case, we may think of the quadratic functional $J_\infty(0, u)$ as giving a second degenerate norm on $l^2(\mathscr{U})$. Let σ be the completion of the orthogonal complement of $\mathscr{N}$ the nullspace of J_∞ in the J_∞-norm. This describes a space of input strings having finite cost. The elements of σ are not necessarily sequences; they are equivalence classes of sequences. However, under many circumstances they may be identified directly as sequences. This will be the case if $\mathscr{U}$ is finite-dimensional or, more generally, if $E(e^{i\theta})$ has closed range for almost all θ.

We shall approach the problem in the usual way (see [1], [2]), namely, by completing the square to put the cost functional J_N in a reduced form from which the optimal control law is apparent. This section is devoted to determining when this procedure is possible.

DEFINITION. The cost functional J_N for system (1.1) and (1.2) is in *reduced form* when there exists an auxiliary Hilbert space $\mathscr{H}_1$ with inner product $\langle\cdot,\cdot\rangle$, a continuous operator $G:\mathscr{U}\to\mathscr{H}_1$, an (possibly unbounded) operator $F:\mathscr{R}\to\mathscr{H}_1$, and a symmetric bilinear form $K(\,,)$ whose domain contains $\tilde{\mathscr{R}}$ such that for all N,

$$J_N(x_0, u)=\sum_{i=0}^{N}\langle Gu_i+Fx_i, Gu_i+Fx_i\rangle+K(x_0, x_0)-K(x_{N+1}, x_{N+1}) \tag{2.1}$$

for $u=\langle u_i\rangle$, an input to system (1.1), and x_i the corresponding states of the system.

If $K(x_N, x_N)\to 0$ along trajectories, then $J_\infty(0, u)$ is always $\geqq 0$, and the finite cost assumption implies that the map μ defined on $l^2(\mathscr{U})$ sequences by $\mu\{u_i\}=\{Gu_i+Fx_i\}$ satisfies $c\|u\|^2\geqq J_\infty(0, u)=\sum_{i=0}^{\infty}\|(\mu u)_i\|^2_{\mathscr{H}_1}$; thus $\mu: l^2(\mathscr{U})\to l^2(\mathscr{H}_1)$. A reduced form is called *outer* if μ has range dense in $l^2(\mathscr{H}_1)$. One could view the map μ as taking a dense subspace of σ isometrically to a dense subspace of $l^2(\mathscr{H}_1)$, and so we may extend μ to a unitary map $\tilde{\mu}:\sigma\xrightarrow{\text{onto}} l^2(\mathscr{H}_1)$.

Proceed formally for a moment. Once the reduced form is obtained, then to minimize J_∞ over inputs u with trajectories x_N on which $K(x_N, x_N)\to 0$, one would solve $Gu_i=-Fx_i$ to obtain a control sequence $\{u_i\}$. If $K(x_N, x_N)\to 0$ on the resulting trajectory, then clearly $\{u_i\}$ is the optimal control and the optimal cost is $K(x_0, x_0)$. To make this argument rigorous, let x_0 be the initial state to be controlled and let v be an input sequence of finite length n whose associated state sequence has $x_{n+1}=x_0$. The approach just described consists of extending v to an element $w=(v_0,\cdots,v_n, u_0, u_1\cdots)$ abbreviated (v, u) in σ with the property that each entry of the sequence $\tilde{\mu}u$ beyond the nth is zero. Such a u can be obtained by solving $\tilde{\mu}u=-(y_{n+1}, y_{n+2},\cdots)$ where $(y_0, y_1,\cdots)=\mu v$. This is because $\tilde{\mu}(v, u)=\mu(v, 0)+\tilde{\mu}(0, u)=(y_0, y_1,\cdots)-(0,\cdots,0, y_{n+1},\cdots)$. The element u of σ is an optimal control (also the unique one in σ) provided that $K(x_N, x_N)\to 0$ on the trajectory arising from u. It turns out that about the best we can expect is $K(x_N, x_N)\to 0$ along trajectories coming from $l^2(\mathscr{U})$ input strings. Under this circumstance, given $\varepsilon>0$, any $u(\varepsilon)$ in $l^2(\mathscr{U})$ with $J_\infty(0, u-u(\varepsilon))<\varepsilon$ is a control sequence which runs the system at within ε of $K(x_0, x_0)$, the optimal cost. Thus we have a reasonable sense in which to think of u as the optimal control. The approximate control is an approximate solution to $Gu_i=-Fx_i$ in the rather strong sense that $\sum_{i=0}^{\infty}\|Gu_i(\varepsilon)+Fx_i(\varepsilon)\|^2<\varepsilon$. Conversely, any $l^2(\mathscr{U})$ control u' within ε of being optimal satisfies $\varepsilon>J_\infty(0, u')-K(x_0, x_0)\geqq\sum_{i=0}^{\infty}\|Gu_i'+Fx_i\|^2$. Thus *the control problem is solved provided that J_N can be put in outer reduced form with $K(x_N, x_N)\to 0$ along any trajectory coming from an $l^2(\mathscr{U})$ input.* The goal of this section is to show when this can be done. After that is finished, we give some conditions under which the actual controlling sequence $u(\varepsilon)$ can be expressed concretely in terms of u.

The results of this section are given in terms of the power spectrum operator:

$$E(e^{i\theta}) = R + W(e^{i\theta})^*[\operatorname{sgn} Q]W(e^{i\theta}) \tag{2.2}$$

which is defined for almost all $e^{i\theta}$ on $\mathbf{T}$. In particular, we shall be concerned with spectral factorizations of it. Recall that an $\mathscr{L}(\mathscr{U})$ function $P(e^{i\theta})$ has a spectral factorization if it can be written in the form

$$P(e^{i\theta}) = M(e^{i\theta})^*M(e^{i\theta}),$$

where M is in $H^\infty(\mathscr{U}, \mathscr{H}_1)$ for some auxiliary Hilbert space $\mathscr{H}_1$. Given a nonnegative operator or matrix-valued function, a spectral factorization may or may not exist. This is a classical question, and the answer is that a factorization usually exists. In recent engineering literature, the results of Gohberg–Krein [10] are usually cited; however, necessary and sufficient conditions are available (see [18, Chap. V, § 4], [23]). The more applicable sufficient conditions for nonnegative P are

(I) $P(e^{i\theta}) \geqq \delta(e^{i\theta})I$ with δ a log integrable function [18. Chapt. V, § 7].
(II) For P matrix-valued log det $P(e^{i\theta})$ is integrable [11, Thm. 18].
(III) P has a (pseudo) meromorphic continuation to $\mathbb{C}$ [23, Thm. 3.1]. This includes the case where P is a rational function.

Thus there are quite a few ways to check if a function has a spectral factorization and so the hypotheses of the theorems appearing in this section are hopefully easy to apply.

Next we shall observe that placing J_N in reduced form is related to spectral factorization of E. Let $u(e^{i\theta})$ denote the Fourier transform of $\{u_n\}$ in l^2; it is a function in $H^2_{\mathbf{T}}(\mathscr{U})$. Fourier transforming (1.1) and the definition (1.2) of the cost function J_∞ gives

$$J_\infty(0, u) = \int_{-\pi}^{\pi} (u(e^{i\theta}), E(e^{i\theta})u(e^{i\theta}))\, d\theta. \tag{2.3}$$

Thus E is closely related to $J_\infty(0, u)$. Suppose that J_N is in reduced form with $K(x_N, x_N) \to 0$ on each trajectory. Then (2.1) can be Fourier transformed to give

$$J_\infty(0, u) = \int_{-\pi}^{\pi} (u(e^{i\theta}), M(e^{i\theta})^*M(e^{i\theta})u(e^{i\theta}))\, d\theta,$$

where M is the uniformly bounded function

$$M(z) = G + zF(I - zA)^{-1}B. \tag{2.4}$$

Comparing this with expression (2.3) for $J_\infty(0, u)$, we find that the Toeplitz operator generated by $E - M^*M$ is identically zero; thus (see [19]) we get that

$$E = M^*M.$$

That is, E has a spectral factorization. The main theorem of this section says that not only this but its converse is true.

THEOREM 2.1. *Suppose that the reachable states $\tilde{\mathcal{R}}$ for the system $[A, B]$ are dense in its state space and that the cost functional J_N is absolutely finite. Then the power density function $E(e^{i\theta})$ has a spectral factorization if and only if the cost J_N can be put in outer reduced form with an optimal cost functional K satisfying $K(x_N, x_N) \to 0$ on trajectories $\{x_N\}$ of the system arising from $l^2(\mathcal{U})$ inputs.*

Proof. One side of the theorem has already been proved. The converse requires the rest of this section. By hypothesis, E has a spectral factorization. Spectral factorizations are not unique and not all of these factorizations have the required form (2.4). However, since E has a factorization, it has [18, Chap. V, § 4] an outer factorization $M \in H^\infty(\mathcal{U}, \mathcal{H}_1)$ which is unique up to a constant multiple and which we will now prove has the form (2.4). The proof relies on realizability theory, in particular, on that developed in [8] in the one-dimensional case, in greater generality in [12], and surveyed in [13].

We begin with a quick sketch of realizability theory. The system $[A, B]$ is called *approximately* (*exactly*) controllable if the range of $\mathcal{C}$ is dense in $\mathcal{H}$ (is all of $\mathcal{H}$). It is *continuously controllable* if $\mathcal{C}$ is a continuous map. Exact controllability is equivalent to the standard pseudoinverse $\mathcal{C}^{-1}$ of $\mathcal{C}$ being a continuous operator. This follows immediately from the open mapping theorem. Similar considerations with adjoint systems give the obvious notions of *approximate* (*exact*) *observability*. A slight modification of Theorem 3C.1 of [12] is the

REALIZABILITY THEOREM. *Any $\mathcal{L}(\mathcal{U}, \mathcal{H}_1)$-valued function $F(z)$ analytic and bounded on the unit disk is the frequency response function of some exactly observable and approximately controllable system $[A, B, C, D]$.*

The operators A, B, C, D in the theorem are given explicitly: A is the restriction of $P_{H^2(\mathcal{H}_1)}\mathcal{M}_{e^{-i\theta}}$ to the subspace $X = \text{cl}\, P_{H^2(\mathcal{H}_1)}\mathcal{M}_F\bar{H}^2(\mathcal{U})$, $B : \mathcal{U} \to X$ is given by $Bu_0 = P_{H^2(\mathcal{H}_1)}\mathcal{M}_{F(e^{i\theta})}u_0e^{-i\theta}$, C is the projection of X onto the subspace of constant functions in $H^2(\mathcal{H}_1)$ and D is $F(0)$. The space X is the state space for the system. This particular realization of F is called the restricted shift realization by Fuhrmann. A fact critical to our control problem can be read off from this construction.

LEMMA 2.2. *If two functions $T_1(z)$ and $T_2(z)$ with $\mathcal{L}(\mathcal{S}_1, \mathcal{S}_2)$ and $\mathcal{L}(\mathcal{S}_3, \mathcal{S}_2)$ values, respectively, satisfy the hypotheses of the Realizability Theorem, if in the above representation, $T_2(z) = D + zC(I - zA)^{-1}B$, and if the state space X_1 for T_1 is contained in the state space X_2 for T_2, then $T_1(z)$ can be written in the form $T_1(z) = D_1 + zC(I - zA)^{-1}B_1$.*

It is now easy to show that M has the realization (2.4). Since M is in $H^\infty(\mathcal{U}, \mathcal{H}_1)$, the function $\tilde{M}$ defined by $\tilde{M}(e^{i\theta}) = M(e^{-i\theta})^*$ is in $H^\infty(\mathcal{H}_1, \mathcal{U})$. We now compare $\tilde{M}$ to the function $\tilde{W}$ defined by (1.5) using Lemma 2.3. Since M is outer,

$$X_{\tilde{M}} = \text{cl}\,(P_{H^2(\mathcal{U})}\mathcal{M}_{\tilde{M}}\bar{H}^2(\mathcal{H}_1)) = \text{cl}\,(P_{H^2(\mathcal{U})}\mathcal{M}_{\tilde{E}}\bar{H}^2(\mathcal{U}))$$

which, in turn, by the definition of E, is contained in

$$X_{\tilde{W}} = \text{cl}\,(P_{H^2(\mathcal{U})}\mathcal{M}_{\tilde{W}}\bar{H}^2(\mathcal{H})).$$

If $\tilde{W}$ has restricted shift realization $[\psi, \Delta, \Lambda]$, then $\tilde{M}$ has a realization $[\psi, \alpha, \Lambda, \rho]$.

The function W has two realizations $[\psi^*, \Lambda^*, \Delta^*]$ and $[A, B, |Q|^{1/2}]$. A straightforward infinite-dimensional version of the state space isomorphism theorem [12, Thm. 3b.1] says that if $[A, B, |Q|^{1/2}]$ is approximately observable, there exists a 1–1 densely defined operator $\beta : \mathcal{H} \to X_{\tilde{W}}$ such that

$$\psi^*\beta = \beta A, \quad \Lambda^* = \beta B, \quad |Q|^{1/2} = \Delta^*\beta.$$

This implies that M has the realization $[A, B, \alpha^*\beta, \rho^*]$, that is, M has the representation (2.4) with $F = \alpha^*\beta$ and $G = \rho^*$.

To finish the theorem, we require some fine structure from Theorem 3b.1 of [12]. It is shown there under the assumption of continuous controllability and approximate observability that β is $\dot{\mathscr{C}}\mathscr{C}^{-1}$, where $\mathscr{C}$ (resp., $\dot{\mathscr{C}}$) is the controllability operator of $[A, B, |Q|^{1/2}]$ (resp., $[\psi^*, \Lambda^*, \Delta^*]$) and $\mathscr{C}^{-1}$ is a pseudoinverse of $\mathscr{C}$. These results extend immediately to the case at hand and validate this definition of β provided that it is interpreted as follows. If $y \in \mathscr{R}$, there is u such that $\mathscr{C}u = y$ define $\beta y = \dot{\mathscr{C}}u$. To check that this is not ambiguous, note that by the lemma in [12] null $\mathscr{C} \subset$ null $\mathscr{Q}^*\mathscr{C}$ = null Hankel$_W$ = null $\dot{\mathscr{Q}}^*\dot{\mathscr{C}}$ and since $\dot{\mathscr{Q}}^*$ is 1–1 this equals null $\dot{\mathscr{C}}$; thus if $\mathscr{C}u = 0$, then $\dot{\mathscr{C}}u = 0$. The construction in the theorem can be completed by setting $F = \alpha^*\beta$. Note that when one does not have approximate controllability, β will not be 1–1; in finite dimensions, for example, null β = (range $\mathscr{Q})^\perp$. Also observe that $F\mathscr{C} = \alpha^*\dot{\mathscr{C}}$ which is a continuous operator.

Now we must show that having the appropriate factorization for E implies that J_N can be put in reduced form. To see this we first observe

LEMMA 2.3. *The cost functional J_N can be written in reduced form* (2.1) *if and only if there exist appropriately defined F, G and $K(\cdot, \cdot)$ which satisfy*

$$\langle Gu_0, Gu_0\rangle = (u_0, Ru_0)_{\mathscr{U}} + K(Bu_0, Bu_0), \tag{2.5a}$$

$$\langle Fx, Gu_0\rangle = K(Ax, Bu_0), \tag{2.5b}$$

$$\langle Fx, Fy\rangle = (x, Qy)_{\mathscr{H}} + K(Ax, Ay) - K(x, y) \tag{2.5c}$$

for x, y in $\tilde{\mathscr{R}}$.

Proof. One simply substitutes (2.5) into the right side of (2.1) and observes, after using (1.1), that (1.2) the definition of J_N has been obtained.

As one might expect, the operators F, G, A, B and the space $\mathscr{H}_1$, appearing in the representation (2.4) for M, will turn out to be the operators required in the lemma. Here we let $\langle \cdot, \cdot \rangle$ denote the inner product on $\mathscr{H}_1$. The optimal cost form $K(\cdot, \cdot)$ is yet to be constructed. Formally, it is, for $x, y \in \tilde{\mathscr{R}}$,

$$K(x, y) = \sum_{j=0}^{\infty} (A^j x, [Q - F^*F]A^j y),$$

and it is not too difficult to check that this formally satisfies (2.5). For example, if this were a finite-dimensional problem, a very simple manipulation would finish the proof. However, our task is a bit tiresome.

Now we give the precise definition of $K(\cdot, \cdot)$. Set

$$\begin{aligned} L(e^{i\theta}) &= E(e^{i\theta}) - [M(e^{i\theta}) - G]^*[M(e^{i\theta}) - G] - R \\ &= G^*M(e^{i\theta}) + M(e^{i\theta})^*G - G^*G - R. \end{aligned} \tag{2.6}$$

Formally one should think of this as

$$L(e^{i\theta}) = B^*(e^{-i\theta} - A^*)^{-1}[Q - F^*F](e^{i\theta} - A)^{-1}B.$$

Define

$$K(A^l Bu_0, Bv_0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} (u_0, L(e^{i\theta})v_0)_{\mathcal{U}}\, e^{+il\theta}\, d\theta \tag{2.7a}$$

and

$$K(Bu_0, A^l Bv_0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} (u_0, L(e^{i\theta})v_0)_{\mathcal{U}}\, e^{-il\theta}\, d\theta \tag{2.7b}$$

for $l \geqq 0$. Next we use (2.5c) to define $K(,)$ inductively; it is

$$\begin{aligned} K(A^{l+1}Bu_0, A^{j+1}Bv_0) = {} & \langle FA^l Bu_0, FA^j Bv_0\rangle \\ & - (A^l Bu_0, QA^j Bv_0)_{\mathcal{H}} - K(A^l Bu_0, A^j Bv_0) \end{aligned} \tag{2.7c}$$

for u_0, v_0 in $\mathcal{U}$. Note that the term involving F is well-defined on $\tilde{\mathcal{R}}$, and so we have defined a function. After some tedious work, which we leave as an exercise, one can show that K is a consistently defined bilinear functional on $\tilde{\mathcal{R}}$.

By construction, $K(,)$ satisfies (2.5c). The identity (2.5a) follows by setting $l = 0$ and performing the integration on the right side of (2.7a) while observing that $1/2\pi \int_{-\pi}^{\pi} M(e^{i\theta})\, d\theta = G$. The identity (2.5b) follows from (2.7) and the fact that

$$\langle FA^l Bv_0, Gu_0\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} (v_0, L(e^{i\theta})u_0)_{\mathcal{U}}\, e^{(l+1)\theta}\, d\theta.$$

Only one property of K remains unverified; that is, $K(x_n, x_n) \to 0$. This follows because $J_\infty(0, u) = (1/(2\pi)) \int_{-\pi}^{\pi} (u, M^*Mu) = \sum_{n=0}^{\infty} \|Gu_i + Fx_i\|^2$, and so (2.1) implies that $K(x_n, x_n) \to 0$. The proof of Theorem 2.1 is finished.

The theorem just completed shows that an approximate control sequence always exists. During the remainder of the section, we describe ways for identifying approximate control sequences explicitly. By the discussion preceding Theorem 2.1, we are confronted with the problem: Given the outer factorization M of E and $y(e^{i\theta})$ in $H^2(X)$ (actually, we may take y to be a polynomial in $e^{i\theta}$), for each $\varepsilon > 0$, find $u_\varepsilon(e^{i\theta})$ in L^2 such that $\int_0^{2\pi} \|y - Mu_\varepsilon\|_X^2 < \varepsilon$. The Fourier transform $u(\varepsilon)$ of u_ε is an l^2-sequence which yields an ε-approximate optimizing control. We know that such u_ε must exist, but the problem is to give a method for finding them explicitly.

We begin by treating the case where E is scalar-valued. The function $u(z) \triangleq M(z)^{-1} y(z)$ is analytic on the disk, but can have fearsome boundary values $u(e^{i\theta})$. Standard ways to approximate $u(e^{i\theta})$ with L^2-functions are

(i) $[A_r u](e^{i\theta}) = u(re^{i\theta})$, Abel approximation,
(ii) $[P_N u](e^{i\theta}) = \sum_{j=0}^{N} u_j e^{ij\theta}$, Fourier approximation,
(iii) $[F_N u](e^{i\theta}) = (1/(N+1)) \sum_{K=0}^{N} P_K$, Cesaso approximation,

and the question before us is: Does $\int_0^{2\pi} \|y - MA_r u\|^2$,etc., go to zero? Since $y = Mu$ and $E = M^*M$, an alternative phrasing of this question is: Does $A_r u \to u$; $P_N u \to u$, *or* $F_N u \to u$ in $L^2(E\,d\theta)$? Fortunately, these are standard questions in harmonic analysis, and the answers are known.

Necessary and sufficient conditions on E for $P_N f \to f$ in $L^2(E\,d\theta)$ are

(a) (Helson and Szego [26]) E has the form $E = e^{g+\tilde{h}}$ where g and h are bounded functions with $\sup |h| < \pi/2$ and $\tilde{h}$ equals the harmonic conjugate of h;

or equivalently

(b) (Hunt, Muckenhaupt and Wheeden [27]) there is a constant C, independent of I such that for every interval I,

$$\frac{1}{|I|}\int_I E(e^{i\theta})\,d\theta \frac{1}{|I|}\int_I \frac{1}{E(e^{i\theta})}\,d\theta \leqq C.$$

Here $|I|$ is the length of I.

This settles approximation theory questions surrounding P_N. The first thing to note is that these conditions are extremely restrictive. They allow E to be singular or to vanish like $E(e^{i\theta}) = \theta^v$ only if $-1 < v < 1$; thus rational E with zeros or poles on the unit circle are eliminated. Traditionally, Cesaro or Abel summation is much more likely to converge than simple Fourier approximation, and our rather negative conclusion suggests that we turn to them as being more practical.

The first thing we mention is a theorem of Rosenblum [22, Thm. 2] which says that Cesaro summation converges on $L^2(E\,d\theta)$ if and only if Abel summation converges on $L^2(E\,d\theta)$. Thus we restrict attention to Abel summation. A necessary and sufficient condition [22, Thm. 1] for $A_r f$ to always converge in $L^2(E\,d\theta)$ is

$$\int_0^{2\pi} P_r(e^{i(\theta-\psi)}) \left|\frac{M(e^{i\theta})}{M(re^{i\theta})}\right|^2 d\theta < K$$

for all $0<r<1$ and ψ. Here $P_r(e^{i\theta})$ is the Poisson kernel. Thus a sufficient condition for the Abel approximation to always work is for

$$\sup_{r,\theta} \left|\frac{M(e^{i\theta})}{M(re^{i\theta})}\right| \leqq K^{1/2};$$

that is, M belongs to a class of functions discussed in Chap. 3, § 1.3 of [18]. This class does include the rational functions.

In the multi-input case where E is an operator, similar structure holds for Abel convergence, and we now derive the sufficiency condition, just used, directly.

$$\left\{\int_0^{2\pi} \|y - MA_r u\|^2\right\}^{1/2} \leqq \left\{\int_0^{2\pi} \|M(e^{i\theta})u(e^{i\theta}) - M(re^{i\theta})u(re^{i\theta})\|^2\right\}^{1/2}$$
$$+\left\{\int_0^{2\pi} \|[M(re^{i\theta}) - M(e^{i\theta})]u(re^{i\theta})\|^2\right\}^{1/2}.$$

The first majorizing term is $\int_0^{2\pi} \|y(e^{i\theta}) - y(re^{i\theta})\|^2$ which goes to zero since $y \in H^2(X)$. The second term is

$$\int_0^{2\pi} \|[I - M(e^{i\theta})M(re^{i\theta})^{-1}]y(re^{i\theta})\|^2$$

which goes to zero for y in $H^2(X)$ if

$$\operatorname*{ess\,sup}_{\theta} \|I - M(e^{i\theta})M(re^{i\theta})^{-1}\|$$

goes to zero or, for y in H^∞, if $\int_0^{2\pi} \|I - M(e^{i\theta})M(re^{i\theta})^{-1}\|^2 \to 0$. Since $M(re^{i\theta}) \to M(e^{i\theta})$ pointwise, the dominated convergence and uniform boundedness theorem imply that the first condition is equivalent to $M(e^{i\theta})M(re^{i\theta})^{-1}$ being uniformly bounded.

Now we turn to a formal question. Recall that the coefficients of the power series expansion for $u(z) = M(z)^{-1}y(z)$ give *formally* our control sequence. The following proposition gives a reasonable condition on the power spectrum E which guarantees that $M(z)^{-1}$ exists for $|z| < 1$ and consequently that this formal sequence exists.

PROPOSITION 2.4. *If M in $H^\infty(\mathcal{U}, \mathcal{H}_1)$ is outer, and if $M(e^{i\theta})^*M(e^{i\theta}) \geqq N(e^{i\theta})^*N(e^{i\theta})$, where N is also outer but with* Range $N(z) = \mathcal{H}_1$ *for some $|z| < 1$, then* Range $M(z) = \mathcal{H}_1$. *In particular, if $M(e^{i\theta})^*M(e^{i\theta}) \geqq \delta(e^{i\theta})I + T(e^{i\theta})$ where $\delta \geqq 0$,* $\log \delta(e^{i\theta})$ *is integrable and $T(e^{i\theta})$ is a trace class operator with* $\log \det [I + T(e^{i\theta})/\delta(e^{i\theta})]$ *integrable, then* Range $M(z) = \mathcal{H}_1$ *for any $|z| < 1$.*

Proof. The first statement follows immediately from the fact that $M(z)^*M(z) \geqq N(z)^*N(z)$ (see [18, Chap. V, Prop. 4.1]). The log integrability conditions imply that δ and $I + T(e^{i\theta})/\delta(e^{i\theta})$ have outer spectral factorizations $\varphi \in H^\infty(\mathbb{C})$ and φ analytic with a lenient growth condition (see [23, Thm. 3.8]). Since ψ is outer, $\det \psi$ is outer or identically zero. If it is identically zero, then we can write ψ as an infinite matrix with respect to a basis, one subset of which spans cl (Range ψ). The determinant of the minor derived from this basis is outer and so its value at the origin is not zero. Since the pseudoinverse $\psi(z)^{-1}$ can be constructed by Cramer's rule; this says that it is in fact bounded, and consequently Range $\psi(z)$ is closed. The function $N = \psi\varphi$ has closed range and satisfies the majorization hypothesis of the first part of this theorem. Consequently Range $M(z) = \mathcal{H}_1$.

We now give examples to show that Theorem 2.1 is in several senses the best possible. Theorem 2.1 says that J_N has reduced form if and only if E has a spectral factorization. Since E has the special form $R + W^* \operatorname{sgn} QW$, it is conceivable that a weak assumption such as $E \geqq 0$ actually forces E to have a spectral factorization. The following example shows that this is not the case. Take $\mathcal{U}$ to be one-dimensional $R = 1$, $\operatorname{sgn} Q = -1$ and set $W^*W = n \leqq 1$. By the realization theorem, any function W in $H(\mathbb{C})$ with $W(0) = 0$ comes from a system and so can arise in this context. By Theorem 18 [11], $\int_{-\pi}^{\pi} \log n(e^{i\theta})\, d\theta > -\infty$ if and only if n has a factorization $n = W^*W$ with W in $H^\infty(\mathbb{C})$. However, if $E = 1 - n \geqq 0$ has a spectral factorization, then $\int_{-\pi}^{\pi} \log (1 - n(e^{i\theta}))\, d\theta > -\infty$, and this is simply not guaranteed by the fact that $\log n$ is integrable.

The second example is of a system for which no exact optimal control law exists. Let $\mathcal{H}=\mathcal{U}=l^2(0,\infty,\mathbb{C})$, $R=0$ and $Q=1$. We shall take $W(e^{i\theta})$, and consequently $E(e^{i\theta})$, to be diagonal in the natural basis for $l^2(0,\infty,\mathbb{C})$ and denote the diagonal entries of $E(e^{i\theta})$ by $e_j(e^{i\theta})$. The outer factor M of E is diagonal with entries $m_j(e^{i\theta})$ each of which is an outer factor of e_j, i.e., $\bar{m}_j m_j = e_j$. The function M has the representation $G+zF(I-zA)^{-1}B$, and the system has an exact optimal control law only if Range G contains $F\tilde{\mathcal{R}}$. To see this, suppose that for each x_0' in $\tilde{\mathcal{R}}$ there is an input $v_0, v_1, \cdots$ which gives the optimal performance of the system. Let $u_0, u_1, \cdots, u_n$ be a control which drives the system from 0 to x_0' and set $u=(u_0, u_1, \cdots, u_n, v_0, v_1, \cdots)$. By optimality, $J(0,u)$ is finite. Thus $K(x_N, x_N) \to 0$ and $J(0,u)=\sum_{i=0}^n \|Gu_i+Fx_i\|^2+\sum_{j=0}^\infty \|Gv_j+Fx_{j+n+1}\|^2$. However, for each $\varepsilon>0$ we can find a control sequence so that the resulting cost is within ε of the $\sum^n$ term. Thus the $\sum^\infty$ term is 0, and so we can actually solve $Gv_0=Fx_{n+1}=Fx_0^*$, that is, $Fx_0'\in$ Range G.

Range G contains $F\tilde{\mathcal{R}}$ if and only if for $\{x_N\}$, any trajectory of the vector $\int_{-\pi}^{\pi} Fx(e^{i\theta})\, e^{-il\theta}\, d\theta$ for each $l=1,2,3,\cdots$ belongs to Range G. This is equivalent to the statement $\int_{-\pi}^{\pi} M(e^{i\theta})u(e^{i\theta})\, e^{-il\theta}\, d\theta$ belongs to Range $M(0)$ for each $u\in H^2(\mathcal{U})$ and $l>0$, and this in turn is equivalent to the statement that

$$\text{Range } M_n \subset \text{Range } M_0,$$

where $M(z)=\sum_{k=0}^\infty M_k z^k$. The operator M_n is multiplication by the sequence $\{(1/n!)(d^n/dz)n_j|_{z=0}\}_{j=0}^\infty$ on $l^2(0,\infty,\mathbb{C})$. Now Range $M_1\subset$ Range M_0 if and only if $M_1=M_0Y$ for Y a bounded operator (see [5]). Thus $(d/dz)m_j(0)/m_j(0)\equiv\delta_j$ must be a bounded sequence. The functions m_j are outer and consequently can be written

$$m_j(z)=\exp\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{e^{it}+z}{e^{it}-z}\log e_j(e^{it})\,dt.$$

Thus

$$m_j(0)=\exp\frac{1}{4\pi}\int_{-\pi}^{\pi}\log e_j$$

and

$$\frac{d}{dz}m_j(0)=\frac{-1}{2\pi}\int_{-\pi}^{\pi}e^{-it}\log e_j(e^{it})\,dt\, m_j(0).$$

To obtain the example, choose a sequence l_j of functions with $l_j(e^{it})\leqq 0$ with $\int_{-\pi}^{\pi} l_j>-\infty$ and $\int_{-\pi}^{\pi} e^{-it}l_j(e^{it})\,dt\to-\infty$. Set $e_j=\exp l_j$, let w_j be a spectral factorization of e_j which vanishes at $z=0$, and use the realizability theorem to determine a system which gives rise to W. By construction $\delta_j\to\infty$ and so Range $G\not\supset F\tilde{\mathcal{R}}$.

Remark 2.1. The case where J can be reduced with $K(x_N, x_N)\not\to 0$ is analyzed in § 4.

Remark 2.2. If Q and R are both nonnegative, then (1.2) and (2.1) together imply that $K(\cdot,\cdot)$ is a nonnegative bilinear form.

Remark 2.3. The feedback law we have obtained can frequently be expressed in terms of the optimal cost functional K. If Range $G = \mathcal{H}_1$ and G^{-1} denotes the standard pseudoinverse of G, then formally

$$u_i = -G^{-1}Fx_i = (G^*G)^{-1}G^*Fx_i, \tag{2.8}$$
$$u_i = (R + B^*KB)^{-1}B^*KAx_i.$$

This expression has a reasonable interpretation even when K is a bilinear functional. As a further aside, we note that the feedback law can be expressed in terms of L. Namely, if $x_i = \sum_{j=0}^{i} A^j B_j v_j$, then

$$u_i = \frac{1}{2\pi}\left[R + \frac{1}{2\pi}\int_{-\pi}^{\pi} L(e^{i\theta})\, d\theta\right]^{-1} 1 \sum_{j=0}^{N} \int_{-\pi}^{\pi} L(e^{i\theta})v_j e^{-i(j+1)\theta}\, d\theta.$$

3. Stability of feedback systems. In this section, we give some stability theorems which are suitable for analyzing the behavior of the control systems found in § 1. We shall not belabor this, since our results are near to existing results (see [4], [3]). Consider the system

$$x_{i+1} = Ax_i + Bu_i, \qquad y_i = Cx_i, \tag{3.1}$$

with feedback law $u_i = \Omega y_i$. The frequency response function is $R(z) = zC(I - zA)^{-1}B$. Define a function $\mathcal{I}(z) = I - \Omega R(z)$ and note that if M and $R(z)$ are scalars, then the classical Nyquist stability criterion (which we shall presently extend) is expressed in terms of the set $\{\mathcal{I}(e^{i\theta}) : \text{all } \theta\}$. Set $C = I$ and note that the formal feedback law $\Omega = -G^{-1}F$, obtained in § 2 from the spectral factor M, satisfies

$$G\mathcal{I}(z) = M(z). \tag{3.2}$$

If Range $G = \mathcal{H}_1$, then $\mathcal{I}$ is outer. Let G^{-1} denote the standard pseudoinverse for G.

The crux of this business is an easily verified identity

$$(I - z[A - B\Omega C])^{-1}B\mathcal{I}(z) = (I - zA)^{-1}B. \tag{3.3}$$

If u is an admissible input in $l^2[0, \infty, \mathcal{U}]$ with Fourier transform u in $H^2(\mathcal{U})$, then the Fourier transform of the trajectory associated with u is $x(z) = z(I - zA)^{-1}Bu(z)$. Thus $\mathcal{I}$ relates trajectories of the original system to those of the feedback system.

THEOREM 3.1. *Suppose that the power density function E for the system $[A, B]$ with absolutely finite J_N satisfies $E(e^{i\theta}) \geqq \delta I > 0$. Then the outer factorization M* of *E gives rise to an optimal feedback law Γ via* § 2 *with the property that the feedback system $[A, A + B\Gamma]$ has the same trajectories as the original system. Furthermore, the ranges of the controllability operators for the two systems are equal.*

Thus if all trajectories of the original system tend to zero, then all trajectories of the controlled system tend to zero. This will also guarantee a weak form of asymptotic stability; namely, if x is a state of the feedback system which is reachable in a finite amount of time, then $(A + B\Gamma)^n x \to 0$ as $n \to \infty$.

The first part of this theorem is an immediate consequence of the fact that G is invertible when $E \geqq \delta I > 0$ and of the following

PROPOSITION 3.2. *The function* $\mathcal{J}(e^{i\theta})$ *(resp.,* $\mathcal{J}(e^{i\theta})^{-1}$*) is in* $H^\infty(\mathcal{U}, \mathcal{U})$ *if and only if the trajectories of* $[A+B\Omega C, B]$ *are contained in (resp., contain) the trajectories of* $[A, B]$.

Proof. One side is obvious. To do the other side, suppose that the trajectories of the feedback system contain those of the original system. If u is a l^2-input sequence, let $x(u)$ be the corresponding trajectory of the original system. By assumption, there is an input Lu to the feedback system with trajectory $x(u)$, and clearly this determines Lu uniquely. So L is a map of $l^2(0, \infty, \mathcal{U})$ into itself. It is trivial to check that the graph of L is closed. Consequently, L is a bounded operator. However, (3.3) implies that L is just the operator "multiplication by $\mathcal{J}(e^{i\theta})$", and so $\mathcal{J}$ is in $H^\infty(\mathcal{U})$. The same type of argument applies to $\mathcal{J}^{-1}$.

Next we look at (3.3) in terms of controllability operators. Let $\mathcal{C}$ and $\mathcal{C}_f$ denote the controllability operators for the original and the feedback system. Let $P_{\bar{H}^2}$ and P_{H^2} denote the orthogonal projection of L^2 onto H^2 and $\bar{H}^2$. If $L \in L^\infty$, we define $\mathcal{T}_L : H^2 \to H^2$ by

$$\mathcal{T}_L f = P_{H^2} M_L f.$$

It is called the Toeplitz operator with generating function L. The best reference for scalar Toeplitz operators is [6]; for Hilbert Toeplitz operators see [19]. Let T_L denote the operator induced on l^2 by Fourier transforming $\mathfrak{T}_L$ on H^2. Let $\mathcal{J}^+(e^{i\theta}) = \mathcal{J}(e^{-i\theta})$. If $\mathcal{J} \in H^\infty$, then $\mathcal{J}^+ \in \bar{H}^\infty$. The second part of Theorem 3.1 follows from

PROPOSITION 3.3. *If* $\mathcal{J}(z)$ *is in* $H^\infty(\mathcal{U}, \mathcal{U})$, *then*

$$\mathcal{C}_f T_{\mathcal{J}^+} = \mathcal{C}$$

If $\mathcal{J}^{-1}(z)$ *is in* $H^\infty(\mathcal{U}, \mathcal{U})$, *then*

$$\mathfrak{C}_f = \mathcal{C} T_{(\mathcal{J}^+)^{-1}}.$$

The operator $T_{\mathcal{J}^+}$ *is invertible if both* $\mathcal{J}^+$ *and* $(\mathcal{J}^+)^{-1}$ *are in* $H^\infty(\mathcal{U}, \mathcal{U})$.

Proof. We do the second relationship first. Observe that

$$\mathcal{C}\{u_j\}_{j=0}^\infty = \sum_{j=0}^{\infty} A^j B u_j = \lim_{r\uparrow 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} (I - re^{-i\theta}A)^{-1} B u(e^{i\theta})\, d\theta.$$

Equation (3.2) implies

$$\mathcal{C}_f\{u_j\}_0^\infty = \lim_{r\uparrow 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} (I - re^{-i\theta}A)^{-1} B \mathcal{J}(re^{-i\theta})^{-1} u(e^{i\theta})\, d\theta.$$

Since $\mathcal{J}(e^{-i\theta})u(e^{i\theta})$ is in $L^2(\mathcal{U}, \mathcal{U})$, it can be written as the sum of its projection V onto $\bar{H}^2$ and its projection, $T_{\mathcal{J}^+}u$ on H^2. Since $(I - \bar{z}A)^{-1}Bv(z)$ is in $\bar{H}^2$, its integral over **T** is zero. This gives the desired result. The first part of the theorem follows similarly. The last statement in the theorem is a standard fact about Toeplitz operators.

Now that Theorem 3.1 is proved, we make a few remarks. The identity

$$\mathcal{J}_0(z) C(I - z[A - B\Omega C])^{-1} = C(I - zA)^{-1}, \tag{3.4}$$

where $\mathcal{S}_0(z) = I - R(z)\Omega$ is the analogue to (3.3) which allows one to connect all statements about trajectories and controllability in Propositions 3.1 and 3.2 to statements about observability. Another remark is that one can lift the hypothesis $E \geqq \delta I > 0$ and get a palatable theorem. Namely, if E has a spectral factorization and the feedback law Γ comes from an outer factor M, then there is a set of inputs u to the feedback system dense in $l^2(0, \infty, \mathcal{U})$ whose trajectories are precisely the trajectories of the original system. This is obtained by strengthening Proposition 3.1 in the obvious way.

Also note that continuous exact controllability and observability imply stability (see [8, Appendix] or [12, § 4, Remark]). The structure is

PROPOSITION 3.4. *If a system $[F, \varphi]$ is continuously exactly observable, then φ is asymptotically stable. If a system $[\varphi, D]$ is continuously exactly controllable, then φ^* is asymptotically stable.*

Remark. Suppose that $[A, B]$ is a finite-dimensional controllable system and that the eigenvalues of A lie inside $|z| < 1$. Then Theorem 3.1 can be strengthened because of these additional assumptions. One obtains that the state operator $A + B\Gamma$ for the feedback system has no eigenvalues on $|z| = 1$ if and only if $E(e^{i\theta}) \geqq \delta I > 0$. The eigenvalues of $A + B\Gamma$ always lie in $|z| \leqq 1$.

The last statement follows trivially from (3.2) since $M(z)^{-1}$ exists for $|z| < 1$. The absence of eigenvalues on $|z| = 1$ follows from Theorem 3.1. Conversely, if E has a zero on $|z| = 1$, then M^{-1} has a pole there. Since $(I - zA)^{-1} Bv_0 \neq 0$ for any v_0 or $|z| \leqq 1$, equation (3.3) implies that $(I - z[A + B\Gamma])^{-1} B$ has a pole on the circle, and so $A + B\Gamma$ has an eigenvalue there.

4. The algebraic Riccati equation. With the control problem we have studied (when R is invertible), one associates the formal linear fractional map

$$\mathcal{F}(P) = A^* P (I + CP)^{-1} A + Q, \tag{4.1}$$

where $C = BR^{-1}B^*$ and expects that the optimal cost "operator" K will be a fixed point $\mathcal{F}(K) = K$ of it. In this section, we give a fairly thorough study of when a fixed point exists. Although everything done is intimately linked with the original control problem, we try to present the forthcoming results as a study of the fixed-point problem for its own sake.

Throughout this section, we shall work with a slightly more general class of $\mathcal{F}$ than those given by (4.1). Any self-adjoint operator C can be written in the form $C = B^* R^{-1} B$, where R is an invertible self-adjoint operator. Provided that the appropriate inverses exist, a simple manipulation converts (4.1) to

$$\mathcal{F}(P) = A^*PA - A^*PB[R + B^*PB]^{-1} B^*PA + Q. \tag{4.2}$$

This formula is more symmetric than (4.1) and consequently easier to use. Also R need not be invertible in (4.2), so it is more general than (4.1). Henceforth, we work with $\mathcal{F}$ of (4.2). The self-adjoint operator $R + B^*PB$ plays an important role in the study of $\mathcal{F}$; we denote it by Λ_P and call it the *indicator* of P. When, for example, $\mathcal{H}$ is finite-dimensional, the natural domain of definition for $\mathcal{F}$ is precisely the set $\mathcal{P}_0$ of those matrices P satisfying Range $\Lambda_P \supset$ Range B^*PA since these are the matrices for which the second term of (4.2) is well-defined.

Not too surprisingly, spectral factorizations play as big a role in this section as they have previously. In fact, we shall require a type of signed factorization. A *signature operator* J is a self-adjoint operator with the property that $J^2 = I$. We say that the self-adjoint $(\mathcal{T}, \mathcal{H}, \mathcal{H})$-valued function E has a (outer) *signed spectral factorization* if and only if there is a signature operator $\mathcal{S}$ on a Hilbert space $\mathfrak{H}_1$ such that for each u in $l^2(\mathcal{U})$, the limit as $r\uparrow 1$ of $\int_0^{2\pi}(u(e^{i\theta}), M^*(re^{i\theta})M(e^{i\theta})u(e^{i\theta}))\,d\theta$ exists and is $\int_0^{2\pi}(u(e^{i\theta}), E(e^{i\theta})u(e^{i\theta}))\,d\theta$; here M is a (bounded outer) $\mathcal{L}(\mathcal{H}, \mathcal{H}_1)$-valued function analytic in the unit disk. The question of which functions have such factorizations was studied by Symeninco (cf. [10]), and he obtained that in many situations E has a signed spectral factorization if and only if $E = AB$ for some outer functions A in $\bar{H}^{\infty}$ and B in H^{∞}. This is consistent with the fact privately observed by A. Devinetz and R. G. Douglas that a uniformly invertible E has a signed spectral factorization if and only if the Toeplitz operator generated by E is invertible. Neither of these conditions are practical to apply, and it is fortunate for control theory purposes that only positive factorizations are interesting. Although the main theorems of this section concern the infinite-dimensional situation, the following corollary (of Theorem 4.7) is new in finite dimensions and describes the behavior there.

THEOREM 4.1. *Suppose that A, B, R, Q are finite-dimensional matrices with R, Q self-adjoint and all eigenvalues of A less than 1. Then the map $\mathcal{F}$ has a self-adjoint fixed point K in $\mathcal{P}_0$ with nonnegative indicator if and only if the function*

$$E(e^{i\theta}) = R + B^*(I - e^{i\theta}A)^{*-1}Q(I - e^{i\theta}A)^{-1}B$$

is nonnegative. The map $\mathcal{F}$ has a fixed point in $\mathcal{P}_0$ (if) and only if E has an (outer) signed spectral factorization.

By (2.5a) the optimal cost functionals from § 2 have positive indicator. These are the important ones and the author suspects without an improved theory of signed factorizations that the first part of Theorem 4.1 is the only part of real interest. It is analogous to the condition of Willems [25] for the continuous-time Riccati equation although here no controllability assumption is required.

4.1. Decomposition of a map into linear and quadratic parts. The fixed-point problem for $\mathcal{F}$ is, to a superficial glance, a quadratic problem, but it can also contain affine linear fixed-point problems of the form $K = NKD + Q$, one example being when $B = 0$. These problems have been studied [21] and can be treated by quite a different approach than a purely quadratic problem. Fortunately, the fixed-point problem decomposes neatly into what we may think of as purely quadratic and purely linear parts. This we now demonstrate.

Let $\mathcal{H}$ and $\mathcal{U}$ be Hilbert spaces and suppose that A, Q acting on $\mathcal{H}$, R acting on $\mathcal{U}$, and $B : \mathcal{H} \to \mathcal{H}$ are bounded operators with R and Q self-adjoint. Let $l_F^2(0, \infty, \mathcal{H})$ denote the $\mathcal{H}$-valued sequences of finite length and define $\mathcal{C} : l_F^2 \to \mathcal{H}$ by $\mathcal{C}\{x_j\} = \sum_{j=0}^{\infty} A^jBx_j$. Set $\tilde{\mathcal{R}} = \text{Range}\,\mathcal{C}$; denote its closure by $\mathcal{R}_1$ and its orthogonal complement by $\mathcal{R}_2$. If P is any self-adjoint operator on $\mathcal{H}$, then in the $\mathcal{R}_1$, $\mathcal{R}_2$ basis it can be written as a matrix $\begin{bmatrix} P_1 & P_2 \\ P_2^* & P_3 \end{bmatrix}$ with P_1 and P_3 self-adjoint. We would like to see how $\mathcal{F}$ acts on such 2×2 matrices. If Γ_1 and Γ_2 are the orthogonal projections of $\mathfrak{H}$ onto $\mathcal{R}_1$ and $\mathcal{R}_2$, respectively, then $A\Gamma_1 = \Gamma_1 A \Gamma_1$ and

$A^*\Gamma_2 = \Gamma_2 A^*\Gamma_2$. We write $A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}$ and $Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_2^* & Q_3 \end{bmatrix}$ and begin computing $\Gamma_i \mathcal{F}(P)\Gamma_j$:

$$\Gamma_1 \mathcal{F}(P)\Gamma_1 = A_1^* P_1 A_1 - A_1^* P_1 B(R + B^* P_1 B)^{-1} B^* P_1 A_1 + Q_1; \tag{4.3}$$

that is, $\mathcal{F}_1(P_1) = \Gamma_1 \mathcal{F}(P)\Gamma_1$, where $\mathcal{F}_1$ is the linear fractional map defined by (4.3),

$$\begin{aligned} \Gamma_1 \mathfrak{F}(P)\Gamma_2 = {} & [A_1^* - A_1^* P_1 B(R + B^* P_1 B)^{-1} B^*] P_2 A_3 \\ & + Q_2 + [A_1^* - A_1^* P_1 B(R + B^* P_1 B)^{-1} B^*] P_1 A_2, \end{aligned} \tag{4.4}$$

that is, $\Gamma_1 \mathcal{F}(P)\Gamma_2$ is an affine linear function of P_2,

$$\begin{aligned} \Gamma_2 \mathcal{F}(P)\Gamma_2 = {} & A_3^* P_3 A_3 + [A_2^* - A_2^* P_1 B(R + B^* P_1 B)^{-1} B^*] \\ & \cdot [P_1 + P_2] A \Gamma_2 \\ & + A_3^* P_2^* [B(R + B^* P_1 B)^{-1} B^* P_1 A_2 - A_2] \\ & - A_3^* P_2^* B[R + B^* P_1 B]^{-1} B^* P_2 A_3, \end{aligned} \tag{4.5}$$

which is an affine linear function of P_3. Thus we see that the only truly quadratic part of the fixed-point problem $\mathcal{F}(P) = P$ is the equation $\mathcal{F}_1(P_1) = P_1$. Also this is the only part of the problem which is interesting from the control theory point of view. We shall call the map $\mathcal{F}$ of (4.2) *purely quadratic* if and only if $\tilde{\mathcal{R}}$ is dense in $\mathcal{H}$. Such maps will take our main attention, and treatment of the linear maps is postponed to the end of this section.

4.2. Purely quadratic maps. Throughout this section, we assume that $\mathcal{F}$ is purely quadratic. The map $\mathcal{F}$ is clearly defined on all bounded operators with invertible indicator. It also will extend continuously to many unbounded operators, and so it is not clear offhand just what should be the natural domain of definition. However, the control problem strongly suggests that the natural space on which $\mathcal{F}$ should act is the space of all possible cost functionals. We formalize this: Let $\mathcal{P}$ denote the space of all symmetric bilinear forms P on $\tilde{\mathcal{R}}$ with the property that

$$P(\mathcal{C}\{u_j\}_{j=0}^N, v) = P(v, \mathcal{C}\{u_j\}_{j=0}^N) \tag{4.6}$$

is for fixed N continuous in u and v belonging to $l^2(\mathcal{U})$. If $P \in \mathcal{P}$, then bilinear form $\Lambda_P(x, y) = (x, Ry) + P(Bx, By)$ for x, y in $\mathcal{U}$ is actually continuous on $\mathcal{U}$, and so by the Riesz representation theorem, there is a bounded operator Λ_P such that $\Lambda_P(x, y) = (x, \Lambda_P y)$. Naturally, Λ_P will be called the *indicator of the bilinear form P*. Given P in $\mathcal{P}$, the bilinear form $P(Bx, y)$ for y in $\tilde{\mathcal{R}}$ is continuous in x, and, consequently, there is an operator E_P defined on $\mathcal{R}$ so that $(x, E_P y) = P(Bx, y)$. We want to have $\mathcal{F}$ defined on as big a subset of $\mathcal{P}$ as is reasonably possible. With this in mind define

$\mathcal{P}_0 = \{P$: there exists a decomposition $\Lambda_P = NSN$ with S a signature operator and N a nonnegative self-adjoint operator satisfying Range $N \supset$ Range $E_P A\}$.

The map $\mathcal{F}$ is defined on $\mathcal{P}_0$ by

$$\mathcal{F}(P)(x, y) = P(Ax, Ay) - (N^{-1} E_P Ax, SN^{-1} E_P Ay) + (x, Qy) \tag{4.7}$$

for x, y in $\tilde{\mathcal{R}}$. Here N^{-1} denotes the standard pseudoinverse of N. This is clearly consistent with (4.2) when P is a bounded operator, and it is straightforward to check that the definition of $\mathcal{F}$ depends only on $\Lambda_P{}^1$ and, consequently, is independent of which factorization NSN is used.

All results on fixed points will be given in terms of the function $W(z)=z|Q|^{1/2}(I-zA)^{-1}B$, which we henceforth assume to be in $H^\infty(\mathcal{H}, \mathcal{U})$, and, in particular, they will involve

$$E(e^{i\theta}) = R + W(e^{i\theta})^* \operatorname{sgn} QW(e^{i\theta}). \tag{4.8}$$

An operator A will be called *asymptotically stable* if $A^n x \to 0$ for each x. If B is an operator with one-dimensional range, then E is a real-valued function on the circle and we shall prove

THEOREM 4.2. *Consider a purely quadratic map $\mathcal{F}$ as in* (4.2) *with B a rank one operator and A asymptotically stable.*

(i) *If $\mathcal{H}$ is finite-dimensional, then $\mathcal{F}$ has a fixed-point in $\mathcal{P}_0$ if and only if E has one sign. The indicator for the fixed point has the same sign as E.*

(ii) *If $\mathcal{H}$ is not finite-dimensional, then $\mathcal{F}$ has many fixed points in $\mathcal{P}_0$. Some fixed points will have positive and some will have negative indicators.*

This theorem sets down the basic behavior of the purely quadratic fixed-point problem. The problem of higher-dimensional B is simply a mixture of these cases. In Theorem 4.6, we sort out this mixture to a large extent, and Theorem 4.2 will be an easy consequence of it.

It turns out that fixed points of $\mathcal{F}$ in $\mathcal{P}_0$ fall into two categories, those for which $P(x_N, y_N) \to 0$ along trajectories x_N, y_N of the system $[A, B]$, called *standard points*, and those which are not standard. Clearly, $P(A^n x, A^n y) \to 0$ for $x, y \in \tilde{\mathcal{R}}$ if P is standard, and one can show that up to terrible pathologies, fixed points for $\mathcal{F}$ of this type are standard. If P has positive indicator, this is always equivalent to being standard. Standard fixed points are the only ones of obvious control theoretic interest, and they correspond to signed spectral factorizations as the following theorem states.

THEOREM 4.3. *The map $\mathcal{F}$ has a standard fixed point* (*if and*) *only if E has a* (*outer*) *signed spectral factorization.*

In finite dimensions for asymptotically stable A, all points are standard, and so this theorem describes that situation completely. Before stating our most complete theorem on fixed points, we give the proof of this theorem since it is instructive.

Proof. We begin with the observation that the bilinear functional $\langle \cdot, \cdot \rangle$ defined on the space $\mathcal{H}_1$ which was used throughout § 2 (see (2.1) and (2.5), in particular) need not be nonnegative. In fact, had we assumed that $\mathcal{H}_1$ has inner product $[\cdot, \cdot]$ and that $\langle x, y\rangle = [\mathcal{S}, x, y]$ for some signature operator $\mathcal{S}$; then the proofs in § 2 would have gone through with the modification that $E = M^*\mathcal{S}M$. Then Lemma 2.3 and Theorem 2.1 combine to give

PROPOSITION 4.4. *There exist G, F and K satisfying* (2.5) *with signed $\langle \cdot, \cdot \rangle$ and having $K(A^N x, A^N y) \to 0$ if E has an outer signed factorization. Conversely, if such G, F, K exist, then E has a signed factorization. In the above statement, K satisfies $\Lambda_K \geqq 0$ if and only if "signed" is removed from the statements about factorizations.*

One thing which requires clarification is that the existence of G, F, K implies a signed factorization for E. The function $M(z)=G+zF(I-zA)^{-1}B$ is analytic inside the disk. If $u\in l^2(\mathcal{U})$, then set $y(re^{i\theta})=M(re^{i\theta})u(re^{i\theta})$. We wish to show that $\int_0^{2\pi}(\mathcal{S}y(re^{i\theta}), y(re^{i\theta}))\,d\theta$ converges to $\int_0^{2\pi}(u(e^{i\theta}), E(e^{i\theta})u(e^{i\theta}))\,d\theta$. If the power series for $y(z)$ is $\sum_{n=0}^{\infty} y_r z^n$, then the first integral is $\sum_{n=0}^{\infty} r^{2n}(\mathcal{S}y_n, y_n)$. Finiteness of the second integral forces the sequence $(x_i, Qx_i)+(u_iRu_i)$ to be summable, and since $K(x_N, x_N)\to 0$, this implies that $(\mathcal{S}Gu_i+Fx_i, Gu_i+Fx_i)$ is summable. However, $y_i=Gu_i+Fx_i$ and so $(\mathcal{S}y_i, y_i)$ is summable; its sum is $\int_0^{2\pi}(u(e^{i\theta}), E(e^{i\theta})u(e^{i\theta}))\,d\theta$. An Able summation argument (cf. [24, § 1.22]) gives $\sum_{n=0}^{\infty} r^{2n}(\mathcal{S}y_n, y_n)\to\sum_{n=0}^{\infty}(\mathcal{S}y_n, y_n)$ as $r\uparrow 1$.

The next thing to prove is that families of three objects G, F, K satisfying (2.5) correspond precisely to fixed points of $\mathcal{F}$.

PROPOSITION 4.5. *The bilinear functional K in $\mathcal{P}_0$ is a fixed point of $\mathcal{F}$ if and only if there exist G and F and possibly signed $\langle\cdot,\cdot\rangle$ so that* (2.5) *holds.*

Proof. Suppose that (2.5) holds. From (2.5a) you see $G^*\mathcal{S}G=\Lambda_K$. By (2.5b) the operator E_KA is $G^*\mathcal{S}F$. If $G=UN$ denotes the polar decomposition of G, then since Range G is dense, U^* is an isometry with Range $U^*\subset$ cl Range N and $u^*\mathcal{S}U-(I-U^*U)=\mathcal{S}'$ is a signature operator. Now $\Lambda_K=N\mathcal{S}'N$, but $E_KA=NU^*\mathcal{S}F$, and so Λ_K is in $\mathcal{P}_0$ and we have

$$\mathcal{F}(K)(x,y)=(N^{-1}NU^*\mathcal{S}Fx, \mathcal{S}N^{-1}NU^*\mathcal{S}Fy)+K(Ax,Ay)+(x,Qy)$$
$$=(Fx,Fy)+K(Ax,Ay)+(x,Qy).$$

By (2.5c) this is just $K(x,y)$.

If K in $\mathcal{P}_0$ is a fixed point of $\mathcal{F}$, then $\Lambda_K=N\mathcal{S}N$. Set $G=N$ and take N^{-1} to be the standard pseudoinverse of N. The bilinear functional $K(BN^{-1}\mathcal{S}x, Ay)$ for fixed y in $\tilde{\mathcal{R}}$ and a dense space of x's equals $(N^{-1}\mathcal{S}x, E_KAy)=(\mathcal{S}x, N^{-1}E_KAy)$ and so is continuous in x. Thus there is an operator F on $\mathcal{R}$ for which this equals (x, Fy). One can reverse the brief computations above and get that G, F, $\mathcal{S}$ and K satisfy (2.5).

The general situation is described by

THEOREM 4.6. *Suppose P in $\mathcal{P}_0$ is a fixed point of $\mathcal{F}$ for which* $\lim_{N\to 0}P(x_N, x_N)$ *exists for each trajectory of the system* $[A, B]$. (Note that for any fixed point with positive indicator this limit either exists or is infinite.) *Then there is a bilinear form $\lambda(,)$ defined on $\tilde{\mathcal{R}}$ so that* $\lim_{N\to\infty}P(A^Nx_1, A^Nx_2)=\lambda(x_1,x_2)$, *and there is a self-adjoint function $\lambda\in L^\infty(\mathcal{U},\mathcal{U})$ so that*

$$\lambda(\mathcal{C}\{u_j\}, \mathcal{C}\{v_j\})=\frac{1}{2\pi}\int_{-\pi}^{\pi}(u(e^{i\theta}), \lambda(e^{i\theta})v(e^{i\theta}))_{\mathcal{U}}\,d\theta. \tag{4.9}$$

The function λ has the properties

(4.10a) *if both $u(e^{i\theta})$ and $(I-e^{i\theta}A)^{-1}Bu(e^{i\theta})$ are vector-valued polynomials in $e^{i\theta}$, then $\lambda(e^{i\theta})u(e^{i\theta})=0$;*

(4.10b) *$E+\lambda$ has a signed spectral factorization;*

(4.10c) *if $u\in l_F^2(0,\infty)$ and $\mathcal{C}u=0$, then $\lambda(e^{-i\theta})u(e^{i\theta})\in H^2(\mathcal{U})$.*

Conversely, if λ *is an* $L^{\infty}(\mathcal{U}, \mathcal{U})$ *function for which*

(4.11a) *property* (4.10a) *holds,*

(4.11b) $E+\lambda$ *has an outer signed factorization,*

(4.11c) *if* $u(e^{i\theta}) \in H^2(\mathcal{U})$ *satisfies* $W(e^{-i\theta})u(e^{i\theta}) \in H^2(\mathcal{H})$, *then* $\lambda(e^{-i\theta})u(e^{i\theta}) \in H^2(\mathcal{U})$,

then $\mathcal{F}$ *has a fixed point P in* $\mathcal{P}_0$ *which satisfies*

$$P(A^N x_0, A^N y_0) \to \lambda(x_0, y_0),$$

where λ *is given by* (4.9) *and* $x_0, y_0 \in \tilde{\mathcal{R}}$.

Proof. Suppose $\lambda \in L^{\infty}(\mathcal{U}, \mathcal{U})$ is a function which satisfies (4.11). We shall give a construction for associating a fixed point of $\mathcal{F}$ with λ. The set of $u \in H^2$ such that $W(e^{-i\theta})u(e^{i\theta}) \in H^2$ is invariant under multiplication by $e^{i\theta}$, and so by the Lax–Beurling theorem, there is a $\varphi \in H^{\infty}(\mathcal{U}_1, \mathcal{U})$ for which $W(e^{-i\theta})\varphi(e^{i\theta}) \in H^{\infty}(\mathcal{U}_1, \mathcal{U})$ and $\varphi(e^{i\theta})^* W(e^{-i\theta})^* \in \bar{H}^{\infty}(\mathcal{H}, \mathcal{U}_1)$ is outer. Since (4.11c) is equivalent to $\varphi(e^{i\theta})^* \lambda(e^{-i\theta}) \in \bar{H}^{\infty}(\mathcal{U}, \mathcal{U}_1)$, we have $\mathrm{cl}\, P_{H^2(\mathcal{U})}(\mathcal{M}_{\lambda(e^{-i\theta})}\bar{H}^2(\mathcal{U})) \subset \mathrm{cl}\, P_{H^2(\mathcal{U})}(\mathcal{M}_{\tilde{W}}\bar{H}^2(\mathcal{H}))$. This along with (4.11b) is the crucial fact in the proof of Theorem 2.1 which yields that there are operators G and F so that the signed outer factor M of $E+\lambda$ has the representation $M(z) = G + zF(I-zA)^{-1}B$. Now we can follow the construction in Theorem 2.1 to obtain a bilinear functional K so that G, F, K reduces J_N. Consequently, K is a fixed point of $\mathcal{F}$.

To see this, we began by associating a bilinear functional $\lambda(\cdot, \cdot)$ on $\tilde{\mathcal{R}}$ with the function $\lambda(e^{i\theta})$ by equation (4.9). To see that this is well-defined we only need $\lambda(x_0, x_0) = 0$ whenever x_0 or $y_0 = 0$. That is, if $x_l = \int_{-\pi}^{\pi} (I - e^{i\theta}A)^{-1} Bu(e^{i\theta}) e^{-il\theta}\, d\theta = 0$ for $u(e^{i\theta})$ some polynomial of order $\leqq l$ in $e^{i\theta}$, then $\int_{-\pi}^{\pi} (V(e^{i\theta}), \lambda(e^{i\theta})u(e^{i\theta}))\, d\theta = 0$. This is equivalent to (4.11a). It is immediate from the definition that $\lambda(Ax, Ay) = \lambda(x, y)$ for $x, t \in \tilde{\mathcal{R}}$. Formally, if we set $K_1(x, y) = \sum_{n=0}^{\infty} (x, A^{*n}(Q - F^*F)A^n y) + \lambda(x, y)$, then the fact that K_1, etc., satisfies (2.5c) is a straightforward consequence of $E + \lambda = M^* \mathcal{S} M$. It is, however, unclear that such a K_1 can be rigorously defined. To see that K actually does exist define L by (2.6) and use $L + \lambda$ in (2.7) to define a function. With a bit of work one can check that this function actually has the properties (2.5) required of K.

Now we do the converse direction. Suppose K is a fixed point of $\mathcal{F}$. Associated with K we have operators $G, F, \mathcal{S}$ as in (2.5) and the function $M(z)$. From (2.2) we see that if $\Lambda_P \geqq 0$,

$$\lim_{N\to\infty} K(x_N, x_N) = -\frac{1}{2\pi}\int_{-p}^{\pi} (u(e^{i\theta}), E(e^{i\theta})u(e^{i\theta}))\, d\theta + \lim_{N\to\infty} \sum_{j=0}^{N-1} \|Gu_i + Fx_i\|^2,$$

where $\{x_N\}$ is the $[A, B]$ trajectory corresponding to input u, and consequently $\lim_{N\to\infty} K(x_N, x_N)$ exists or is $+\infty$, In general,

$$\lim_{N\to\infty} K(x_N, x_N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} (u(e^{i\theta}), [-E(e^{i\theta}) + M(e^{i\theta})^* \mathcal{S} M(e^{i\theta})]u(e^{i\theta}))\, d\theta$$

is finite for all trajectories if and only if the function $\lambda = M^*\mathcal{S}M - E$ is in $L^\infty(\mathcal{U}, \mathcal{U})$, thus establishing (4.10b). We also get $\lim_{N\to\infty} K(x_N, y_N)$ exists. From (2.1) one can also see that

$$\sum_{n=0}^{\infty} (A^n x_0, QA^n y_0) = \sum_{n=0}^{\infty} (FA^n x_0, FA^n y_0) + K(x_0, y_0) - \lim_{N\to\infty} K(A^N x_0, A^N y_0),$$

and so the last limit exists and serves to define $\lambda(\cdot, \cdot)$ *on* $\tilde{\mathcal{R}}$. Now $\lambda(\cdot, \cdot)$ will clearly be given by (4.9), and the fact that $\lambda(0, x_0) = \lambda(y_0, 0) = 0$ for x_0 in $\tilde{\mathcal{R}}$ is equivalent to (4.10a).

Finally we verify that λ satisfies (4.10c) and in the process show that (4.10c) and (4.11c) are closely related. Suppose that $u \in l_F^2(0, \infty)$ and $\mathcal{C}u = 0$. Then

$$\mathcal{C}u = \frac{1}{2\pi} \lim_{r\uparrow 1} \int_{-\pi}^{\pi} (I - re^{-i\theta}A)^{-1} Bu(e^{i\theta})\, d\theta = 0$$

and

$$A\mathcal{C}u = \frac{1}{2\pi} \lim_{r\uparrow 1} \int_{-\pi}^{\pi} e^{i\theta}[(I - re^{-i\theta}A)^{-1} - I] Bu(e^{i\theta})\, d\theta = 0.$$

Since $\int_{-\pi}^{\pi} e^{i\theta} Bu(e^{i\theta})\, d\theta = 0$, one has $\lim_{r\uparrow 1} \int_{-\pi}^{\pi} e^{i\theta}(I - re^{-i\theta}A)^{-1} Bu(e^{i\theta})\, d\theta = 0$, and by a similar manipulation of $A^n\mathcal{C}u = 0$, one obtains

$$\lim_{r\uparrow 1} \int_{-\pi}^{\pi} e^{in\theta}(I - re^{-i\theta}A)^{-1} Bu(e^{i\theta})\, d\theta = 0$$

for $n = 0, 1, 2, \cdots$. Thus if N is any operator defined on $\tilde{\mathcal{R}}$ for which $N(I - e^{-i\theta}A)^{-1}B \in \bar{H}^\infty$, then $N(I - e^{-i\theta}A)^{-1}Bu(e^{i\theta})$ is in H^2 and has the 0th Fourier coefficient equal to zero. In particular, by taking $N = |Q|^{1/2}$ or $N = F$ we get that $W(e^{-i\theta})u(e^{i\theta})$ or $M(e^{-i\theta})u(e^{i\theta})$ is in $H^2(\mathcal{U})$. Since $\lambda(e^{-i\theta}) = M(e^{-i\theta})^* \mathcal{S} M(e^{-i\theta}) - W(e^{-i\theta})^* \operatorname{sgn} Q W(e^{-i\theta}) - R$, the function $\lambda(e^{-i\theta})u(e^{i\theta})$ is in $H^2(\mathcal{U})$; thus (4.10c) holds. Also one sees that statement (4.10c) implies statement (4.11c) for $u \in l_F^2(0, \infty)$. The converse will be true under strong observability assumptions on the system $[|Q|^{1/2}, A]$ provided that $\mathcal{C}$ acts continuously on $l^2(0, \infty)$.

Proof of Theorem 4.2. If $\mathcal{H}$ is finite-dimensional, then E is rational and so has a signed spectral factorization if and only if E has one sign. Thus Theorem 4.3 applies. The statement about the indicator is clear from the proof of Theorem 4.3. If $\mathcal{F}$ has a fixed point P in $\mathcal{P}_0$, then since $\mathcal{H}$ is finite-dimensional, P is a continuous functional and the stability of A implies that $K(A^n x_0, A^n y_0) \to 0$. Thus Theorem 4.4 implies thut λ is 0; that is, E has a signed spectral factorization and consequently λ has one sign.

If $\mathfrak{H}$ is not finite-dimensional, then $(I-zA)^{-1}B$ cannot be rational. Thus $(I-zA)^{-1}Bp(z)$ can never be a polynomial in z when $p(z)$ is a polynomial in z. Thus property (4.11a) holds for any real-valued $\lambda \in L^{\infty}(\mathbb{C})$. Set $\lambda(e^{i\theta}) \equiv \sup_{\theta} |E(e^{i\theta})| + 1 = \lambda$. The function $E + \lambda \geqq \delta > 0$ certainly satisfies (4.11b). Since λ is a constant, (4.11c) is vacuously satisfied by λ. The resulting fixed point has a positive indicator. The function $E - \lambda$ gives a fixed point with negative indicator.

4.3. The linear part. Finally we shall consider maps which are not purely quadratic. In a formal sense, the linear and quadratic parts of the map $\mathcal{F}$ have a very nice relationship as we shall see. Technically speaking, the problems might be incompatible because the space $\tilde{\mathcal{R}}$ is crucial to the quadratic problem, and the operator $\Gamma_1 A \Gamma_2 = A_2$, which links the linear and quadratic parts of the equation, might have range very much disjoint from $\tilde{\mathcal{R}}$. Note that if P_1 satisfies (4.3) and Range A_2 does not strongly intersect the domain of definition of P_1, then the last term in (4.4) is not well-defined. Throughout most of this we shall assume that Range $A_2 \subset \tilde{\mathcal{R}}$ and call the linear and quadratic parts of $\mathcal{F}$ *compatible* whenever this happens. This assumption is certainly satisfied when $\mathcal{H}$ is finite-dimensional. Although weaker assumptions will do, we shall assume that B has finite-dimensional range in order to avoid annoying details. If $\|A^n\| \leqq Ka^n$ for some $a<1$, then A is called *exponentially stable.*

Suppose that A is exponentially stable, that E has an outer signed factorization $M\mathcal{S}^*M$ and that P_1 is the fixed point of $\mathcal{F}_1$ which corresponds to it. Clearly, P_1 satisfies (4.3). Next we seek a solution to (4.4). A formal solution is

$$P_2 = \sum_{k=2}^{\infty} N^{*k} P_1 A_2 A_3^{k-1} + \sum_{k=1}^{\infty} N^{*k} Q_2 A_3^k, \tag{4.12}$$

where $N^* = A_1^* - A_1^* P_1 B (R + B^* P_1 B)^{-1} B^*$. Note by (2.7) that N miraculously is $A_1^* + (G^{-1}F^*)B$ and so it is the adjoint of the operator which propagates the states of the feedback system. We have, in § 3, a stability analysis for this operator. If M were invertible outer, then by Theorem 3.1 $N^k x \to 0$ for $x \in \tilde{\mathcal{R}}$. Thus $|P_1(N^k x, A_3^k y)| \leqq C\|A_3^k\| \leqq C'a^n$, and so (4.12) makes good sense. If M is not invertible outer, then given $u(z)$ in H^2 there is a function $g(z)$ such that $M(z)g(z) = u(z)$ for $|z|<1$.

The identity (3.3) which underlies Theorem 3.1 implies that any trajectory $\{x_n\}$ of the feedback system satisfies $\|x_n\| \leqq Cr^n$ for any $r>1$. Thus $\|Nx^k\| \leqq C(1/(a+\varepsilon))^k$, and this is clearly enough to guarantee that (4.12) defines a bilinear functional on $\tilde{\mathcal{R}} \times \mathcal{R}_2$.

The final step is to obtain a solution for (4.5). The final solution is

$$P_3 = \sum_{k=1}^{\infty} A_3^{*k} T A_3^k,$$

where

$$T = \Gamma_2 N^* P_1 A_2 + \Gamma_2 N^* P_2 A_3 - A_3^* P_2^* N \Gamma_2 - A_3^* P_2^* B [R + B^* P_1 B]^{-1} B^* P_2 A_3.$$

The first three operators are, in fact, bounded primarily because range $N\Gamma_2 \subset \tilde{\mathcal{R}}$. Also $B^* P_2 A_3$ is bounded because of its construction and the $\tilde{\mathcal{R}}$ continuity of P_1. Thus P_3 is a well-defined bounded operator and we have

THEOREM 4.7. *Suppose the map $\mathcal{F}$ of* (4.2) *with finite rank B and exponentially stable A has compatible linear and quadratic parts. Then E has an* (*outer*) *signed spectral factorization* (*if and*) *only if $\mathcal{F}$ has a fixed point.*

The fixed point in this theorem is a bilinear functional on the obvious subspace of $\mathcal{H}$, namely, $(\tilde{\mathcal{R}}\oplus\tilde{\mathcal{R}}^{\perp})\times(\tilde{\mathcal{R}}\oplus\tilde{\mathcal{R}}^{\perp})$. It is continuous on $\tilde{\mathcal{R}}^{\perp}$ and has the continuity properties of $\mathcal{P}_0$ on $\tilde{\mathcal{R}}$.

Appendix. Symplectic maps. The linear fractional maps (4.1) we studied are close to the class of symplectic maps of C. L. Siegel [28] except we work with complex rather than real matrices. Complex symplectic maps have the form

$$\mathfrak{S}(K)=(\mathcal{B}+\mathcal{U}K)(\mathcal{D}+\mathcal{C}K)^{-1}, \tag{A.1}$$

where the coefficient matrix $\mathcal{M}=\begin{bmatrix}\mathcal{U} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{bmatrix}$ satisfies $\mathcal{M}^*\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\mathcal{M}=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$ or equivalently $\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}=\mathcal{M}\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\mathcal{M}^*$. These intertwining conditions are equivalent to $\mathcal{U}\mathcal{B}^*=\mathcal{B}\mathcal{U}^*$, $\mathcal{C}\mathcal{D}^*=\mathcal{D}\mathcal{C}^*$, $\mathcal{U}\mathcal{D}^*-\mathcal{B}\mathcal{C}^*=I$ which, in turn, are equivalent to $\mathcal{D}^*\mathcal{B}=\mathcal{B}^*\mathcal{D}$, $\mathcal{C}^*\mathcal{U}=\mathcal{U}\mathcal{C}$, $\mathcal{D}^*\mathcal{U}-\mathcal{B}^*\mathcal{C}=I$. If $\mathcal{D}$ is invertible, a straightforward computation shows that A is invertible; then $\mathcal{F}$ in (4.7) equals $\mathfrak{S}$ if and only if

$$A=\mathcal{D},\quad Q=\mathcal{B}\mathcal{D}^{-1},\quad C=\mathcal{D}^{-1}\mathcal{C}, \tag{A.2}$$

or equivalently, if

$$\mathcal{U}=QA^{-1}C+A^*,\quad \mathcal{B}=QA^{-1},\quad \mathcal{C}=A^{-1}\mathcal{C},\quad \mathcal{D}=A^{-1}. \tag{A.3}$$

This computation in fact shows

PROPOSITION A.1. *The map $\mathcal{F}$ of* (4.2) *is symplectic if and only if A is invertible. Any symplectic map with $\mathcal{D}$ invertible can be written in the form* (4.2) *with A invertible.*

The function E which determines the fixed-point behavior of $\mathfrak{S}$ is

$$E(e^{i\theta})=\operatorname{sgn}\mathcal{D}^{-1}\mathcal{C}+|\mathcal{D}^{-1}\mathcal{C}|^{1/2}(\mathcal{D}-e^{i\theta})^{-1*}\mathcal{D}^*\mathcal{B}(\mathcal{D}-e^{i\theta})^{-1}|\mathcal{D}^{-1}\mathcal{C}|^{1/2};$$

the *indicator* for a point K is $\Lambda_K=\operatorname{sgn}\mathcal{D}^{-1}\mathcal{C}+|\mathcal{D}^{-1}\mathcal{C}|^{1/2}K|\mathcal{D}^{-1}\mathcal{C}|^{1/2}$ and Theorem 4.1 translates to

THEOREM A.2. *If $\mathfrak{S}$ is the symplectic map* (A.1) *and the eigenvalues of $\mathcal{D}$ lie outside of $|z|=1$, then $\mathfrak{S}$ has a self-adjoint fixed point K in $\mathcal{P}_0$ with nonnegative indicator if and only if $E\geqq 0$. It has a self-adjoint fixed point* (*if and*) *only if E has an* (*outer*) *signed factorization.*

Even though the class of maps given by (4.2) is not the same as the symplectic maps, these maps do take the set of matrices K with $\operatorname{Im} K>0$ into those with $\operatorname{Im} K\geqq 0$. This is true because, formally,

$$\mathcal{F}(K)-\mathcal{F}(K)^*=A^*(1+CK)^{-1*}(K-K^*)(1+CK)^{-1}A,$$

and a glance at (4.2) reveals that $\mathcal{F}$ is well-defined when $\operatorname{sgn} C+|C|^{1/2}K|C|^{1/2}$ is invertible; $\operatorname{Im} K>0$ implies such invertibility.

REFERENCES

[1] B. D. O. ANDERSON AND J. B. MOORE, *Linear Optimal Control*, Prentice-Hall, Englewood Cliffs, N.J., 1971.

[2] R. BROCKETT, *Finite Dimensional Linear Systems*, John Wiley, New York, 1970.

[3] F. M. CALLIER AND C. A. DESOER, *Necessary and sufficient conditions for stability for n-input, n-output convolution feedback systems with a finite number of unstable poles*, IEEE Trans. Automatic Control, AC-18 (1973), no. 3.

[4] J. H. DAVIS, *Stability conditions derived from spectral theory: Discrete systems with periodic feedback*, this Journal, 10 (1972), pp. 1–13.

[5] R. G. DOUGLAS, *On majorization, factorization and range inclusion of operators on Hilbert space*, Proc. Amer. Math. Soc., 17 (1966), pp. 413–415.

[6] ———, *Banach algebra techniques in the theory of Toeplitz operators*, Regional Conf. Series, vol. 15, American Mathematical Society, Providence, R.I., 1972.

[7] R. G. DOUGLAS AND J. W. HELTON, *Inner dilations of analytic matrix functions and Darlington synthesis*, Acta Sci. Math. (Szeged), 34 (1973), pp. 61–67.

[8] P. FUHRMANN, *On realization of linear systems and applications to some questions of stability*, to appear.

[9] ———, *Exact controllability and observability and realization theory in a Hilbert space*, to appear.

[10] I. C. GOHBERG AND M. G. KREIN, *Systems of integral equations on a half-line with kernels depending on the difference of the arguments*, Amer. Math. Soc. Transl., 14 (1960), pp. 217–287.

[11] H. HELSON, *Lectures on Invariant Subspaces*, Academic Press, New York, 1964.

[12] J. W. HELTON, *Discrete time systems, operator models and scattering theory*, J. Functional Analysis, to appear.

[13] ———, *Operator techniques for distributed systems*, Proc. 11th Allerton Conf., October, 1973.

[14] K. Y. LEE, S. CHOW AND R. O. BARR, *On the control of discrete-time distributed parameter systems*, this Journal, 10 (1972), pp. 361–376.

[15] B. P. MOLLINARI, *The stabilizing solution of the algebraic Riccati equation*, this Journal, 11 (1973), pp. 262–271.

[16] ———, *Equivalence relations for the algebraic Riccati equation*, this Journal, 11 (1973), pp. 272–285.

[17] ———, *The stable regulator problem and its inverse*, IEEE Trans. Automatic Control, AC-18 (1973), no. 5.

[18] B. SZ.-NAGY AND C. FOIAS, *Harmonic Analysis of Operators on a Hilbert Space*, North-Holland, Amsterdam, 1970.

[19] M. RABINDRANATHAN, *On the inversion of Toeplitz operators*, J. Math. Mech., 19 (1969-70), pp. 195–206.

[20] M. REED AND B. SIMON, *Methods of Modern Mathematical Physics*, Academic Press, New York, 1972.

[21] M. ROSENBLUM, *On the operator equation $BX-XA=Q$*, Duke Math., J., 23 (1956), pp. 263–269.

[22] ———, *Summability of Fourier series in $L^p(d)$*, Trans. Amer. Math. Soc., 105 (1962), pp. 32–92.

[23] M. ROSENBLUM AND J. ROVNYAK, *The factorization problem for nonnegative operator valued functions*. Bull. Amer. Math. Soc., 77 (1971), pp. 287–318.

[24] E. L. TITCHMARSH, *The Theory of Functions*, 2nd ed., Oxford University Press, London, 1939.

[25] J. WILLEMS, *Least squares stationary optimal control and the algebraic Riccati equation*, IEEE Trans. Automatic Control, AC-16 (1971), no. 6.

[26] H. HELSON AND G. SZEGO, *A problem in production theory*, Ann. Mat. Pure Appl., 51 (1960), no. 4.

[27] R. HUNT, B. MUCKENHOUPT AND R. WHEEDEN, *Weighted norm inequalities for the conjugate function and Hilbert transform*, Trans. Amer. Math. Soc., 176 (1973), pp. 227–251.

[28] C. L. SIEGEL, *Symplectic Geometry*, Academic Press, New York, 1968.