A Basic Optimization Problem in Linear Systems*

by

WILLIAM A. PORTER[†]

1. Introduction. Consider the simple system in Figure 1 which satisfies the equation

$$(1) \qquad\qquad x = Tu + N\xi.$$

Here $T: H_1 \to H_3$ and $N: H_2 \to H_3$ are bounded linear transformations between Hilbert spaces. Our initial concern is with the following optimization problem: For a fixed element, $\xi \in H_2$, find the control $u \in H_1$ which minimizes the functional

$$(2) \qquad\qquad J(u) = \|u\|^2 + \|Tu + N\xi\|^2.$$

This problem is extensively considered in the literature and includes the familiar optimal regulator and optimal tracking problems for linear systems with a quadratic performance criterion (see [1]). The solution may be obtained by the following rather direct method (see [2], Section 4.4). Observe that $H_1 \times H_3$, equipped with the usual inner product, is a Hilbert space whose norm is computed by $|(u, x)|^2 = \|u\|^2 + \|x\|^2$ for $(u, x) \in H_1 \times H_3$. Note also that the infimum of J over H_1 is the distance of $(0, N\xi) \in H_1 \times H_3$ from the graph of T. Since the graph of T is a closed linear subspace, this infimum is attained uniquely by the orthogonal projection of $(0, N\xi)$ on this subspace. It is then easily established that the optimal control is given by

$$(3) \qquad\qquad u_0 = -(I + T^*T)^{-1}T^*N\xi.$$

This solution although concise and elegant can lead to difficulties in its physical implementation. To illustrate, let H_1 be $L_2(-\infty, \infty)$ while T is the delay line: $(Tu)(t) = u(t - t_0)$, $t \in (-\infty \ \infty)$. It is easily shown then that $(T^*x)(t) = x(t + t_0)$, $t \in (-\infty, \infty)$. In other words T^* is a pure predictor. Similarly there are many familiar examples involving continuous, discrete, and distributive phenomena where nonanticipatory systems have anticipatory adjoints. Thus the implementation of Equation (3), particularly in feedback form, may well be physically impossible.

From these remarks it is evident that the problem posed above should be augmented with a physical realizability constraint. The analytic description of

physical realizability, however, is not obvious in the setting of abstract spaces. The idea of a causal system, for instance, seems intimately related to concrete properties of the system as a mapping between spaces composed of functions defined on an underlying ordered set.

Consider, however, a bounded linear transformation $u \to x$ determined by

$$x(t) = \int_0^t w(t-\beta)\, u(\beta)\, d\beta, \qquad t \in (0, \infty)$$

where w, u, $x \in L_2(0, \infty)$. In Laplace transform notation this system may be represented in multiplicative form,

$$\hat{x}(s) = \hat{w}(s)\hat{u}(s), \qquad s = \sigma + j\omega.$$

Now bounded causal convolutions can be identified with transfer functions which are analytic in $\sigma > 0$. Indeed for the causal map $u \to x$ on $L_2(0, \infty)$ the functions $\hat{u}$ $\hat{w}$ and hence $\hat{x}$ are all members of the closed linear subspace of functions analytic in $\sigma > 0$. Thus it would seem that one useful (tentative) constraint on the optimization problem would be that the optimal control must lie in an appropriate closed linear subspace.

With this as motivation a sequence of abstract constrained optimization problems are posed and solved, and the solutions are characterized. While examples are given to motivate and illustrate the development, emphasis is placed on the method itself. In this regard the following questions are dealt with. (i) Can the Wiener-Hopf methodology be generalized to a form which is independent of analytic function theory? (ii) Can this same generalization encompass multivariate cases and related time-frequency domain results which utilize the algebraic Riccati equation? (iii) Can nonstationary systems and finite-time-interval problems be solved using the generalized solution?

To illustrate that no hard results have been lost in the abstraction, some well-known results on the optimal regulator problem are derived. Some new results related to system sensitivity are also derived.

2. The Basic Problem. In this section we consider in some detail the *Problem* 1: Characterize the element $u_0 \in L$ which minimizes J of Equation (2).

It is noted that this constrained minimization problem has a *unique* solution for arbitrary L. Indeed the remarks of Section 1 hold for the constrained case where the element $(0, N\xi) \in L \times H_3$ is projected orthogonally on the graph of the restriction of T to L. Our emphasis in this section therefore lies in characterizing and identifying this solution.

The functional J has a Fréchet derivative (see [3], Chapter 6) which will be denoted by δJ. It is easily verified that

$$\delta J(u, \delta u) = 2\mathrm{Re}\,\{\langle u + T^*Tu + T^*N\xi, \delta u \rangle\}$$

where $\langle\ ,\ \rangle$ denotes the inner product on H_1. This expansion leads to

LEMMA 1. $u_0 \in L$ *is minimal for J on L if and only if*

$$u_0 + T^*Tu_0 + T^*N\xi \in L^\perp.$$

Proof. If $u_0 + T^*Tu_0 + T^*N\xi \in L^\perp$, then clearly $\delta J(u_0; \delta u) = 0$ for all $\delta u \in L$. Conversely, if for some u_0, $J(u_0; \delta u) = 0$ for all $\delta u \in L$, then clearly Re $\langle u_0 + T^*Tu_0 + T^*N\xi, \delta u \rangle = 0$ for all $\delta u \in L$. However, if we use the identity Im $\langle x, y \rangle = $ Re $\langle x, iy \rangle$ for complex spaces and the fact that L is a linear subspace, it follows immediately that if $J(u_0; \delta u) = 0$ for all $\delta u \in L$ then $u_0 + T^*Tu_0 + T^*N\xi \in L^\perp$. The lemma now follows immediately from the expansion

$$J(u + \delta u) = J(u) + \delta J(u; \delta u) + \|\delta u\|^2 + \|T\delta u\|^2.$$

Consider the function V defined on H_1 by

$$Vu = (I + T^*T)u + T^*N\xi, \qquad u \in H_1.$$

This function is everywhere defined, continuous and strongly monotone:

$$\langle u_1 - u_2, Vu_1 - Vu_2 \rangle = \langle (u_1 - u_2), (I + T^*T)(u_1 - u_2) \rangle \geq \|u_1 - u_2\|^2.$$

In view of Lemma 1 a critical question is where a $u \in L$ exists such that $Vu \in L^\perp$. This question is concisely answered by the theorem of Minty (see [4], Theorem 2.4).

THEOREM 1. *Let V be an everywhere defined, continuous strongly monotone function. Let L be any closed subspace of H_1. Then there is a unique point in L which is mapped by V into a point of $L^\perp$.*

To paraphase the above development: For every $\xi \in H_2$ and any closed subspace $L \subset H_1$ a unique $u_0 \in L$ exists satisfying (4) and minimizing J over L. As we have noted, this result can be obtained directly by considering the graph of the restriction of T to L in $L \times H_3$. The formulation given in (4) and Theorem 1 is more convenient for later developments.

While Theorem 1 establishes the existence of a unique solution, it does not lend much insight into the form of this solution. With some (temporary) assumptions it is possible to further characterize this solution.

LEMMA 2. *Let F be an invertible operator on H_1 which maps L onto L. The set $\{u : F^*Fu + x \in L^\perp\}$ has a unique element $u_0 \in L$ which is given by*

(5) $$u_0 = -F^{-1}P(F^*)^{-1}x,$$

where P is the orthogonal projection on L.

Proof. The hypothesis of the lemma suffices for the mapping $u \to F^*Fu + x$ to satisfy the requirements of Theorem 1; thus u_0 exists and is unique. The hypothesis $F(L) = L$ implies $F^*(L^\perp) = L^\perp$ which implies $F^{*-1}(L^\perp) = L^\perp$. Thus it suffices to consider the condition $F^*[Fu + (F^*)^{-1}x] \in L^\perp$ which implies that $Fu + (F^*)^{-1}x \in L^\perp$. Since P annihilates $L^\perp$, $PFu + P(F^*)^{-1}x = 0$. Now since $PFu = Fu$ for $u \in L$ and choosing $u = u_0$, the solution supplied by Theorem 1, we see that $u_0 = -F^{-1}P(F^*)^{-1}x$, which implies directly the lemma.

Consider now the case where $M = F(L)$ has no particular relationship to L other than the defining relationship. If F is one-to-one and onto, then M inherits closedness from L and the orthogonal projection Q on M may be defined. In Lemma 2 the condition $F(L) = L$ was used in an essential way. The following corollary removes this condition altogether.

COROLLARY. *Let F be an invertible operator on H_1, and let Q denote the orthogonal projection on $M = F(L)$. Then the minimal element $u_0 \in L$ of Lemma 2 is given by $u_0 = -F^{-1}QF^{*-1}x$.*

Proof. It suffices to find $u_0 \in L$ such that

$$0 = \langle F^*Fu_0 + x, v \rangle = \langle Fu_0 + F^{*-1}x, Fv \rangle \qquad \text{for all } v \in L.$$

By definition Fv spans M as v spans L. It is therefore necessary and sufficient to solve $u_0 \in L$, $Fu_0 + F^{*-1}x \in M^\perp$. Since Q projects on M, $QFu_0 = Fu_0$, $u_0 \in L$; moreover, Q annihilates $M^\perp$ and hence the corollary follows.

Our next theorem summarizes these results in the context of the original problem.

THEOREM 2. *Let $F^*F = I + T^*T$ and let Q be the orthogonal projection on $M = F(L)$. Then Problem 1 has the solution*

$$(6) \qquad\qquad u_0 = -F^{-1}QF^{*-1}T^*N\xi$$

which is unique in L. If $F(L) = L$ then $Q = P$.

Proof. It is necessary only to note that the identity $F^*F = I + T^*T$ automatically requires that F is one-to-one and onto.

Remark 1. In Theorem 2 there are two distinct possibilities. The first is to find a factorization of $I + T^*T$ such that $F(L) = L$. This problem is dealt with in the next section and in Appendix B. The second possibility is to use any factorization of $I + T^*T$, for instance, the positive square root and then determine the projection Q on M. If the projection P on L is computationally convenient, then all the difficulties in the first case lie with the factorization problem. In the second case, however, some of the difficulties are transferred to the determination of Q. This latter problem is dealt with in Appendix C.

Suppose now that T is an operator on a Hilbert space H. Let L be a closed subspace of H. The following two definitions are offered for consideration.[†]

Definitions (1) T is *L-causal* if $T(L) \subset L$;
            (2) T is *L-miniphase* if $T(L) = L$.

These two definitions are temporary and will be improved upon at a later point. However, with this notation a portion of our above results may be restated as follows. If F is an *L*-miniphase partition of $I + T^*T$, then Equation (6) with $P = Q$ characterizes the solution of Problem 1.

Finally we note that under some circumstances it is possible to choose $Q = P$ in Theorem 2 although $F(L) \neq L$. If $F(L) \subset L$ and $PF^{*-1}x \in F(L)$, then, for the u_0 in (5),

    (i) $F^*Fu_0 + x \in L^\perp$
    (ii) $Fu_0 \in L$
    (iii) $PF^{*-1}x \in F(L)$ if and only if $u_0 \in L$.

[†]These hyphenated terms are contractions for causal with respect to L and minimum phase with respect to L

Indeed since $Fu_0 = -PF^{*-1}x$, (ii) and (iii) are evident. To establish (i) we note that $F^*Fu_0 + x = (I - F^*PF^{*-1})x \in L$ if and only if $\langle l, x \rangle - \langle l, F^*PF^{*-1}x \rangle = 0$ for all $l \in L$. However, if we use $\langle l, F^*PF^{*-1}x \rangle = \langle F^{-1}PFl, x \rangle = \langle l, x \rangle$, $l, \in L$, this latter identity follows. Thus (5) supplies the proper $u_0 \in L$.

Example 1. The above development bears a definite relationship with the classical Wiener-Hopf methodology. To bring this connection to the surface we shall consider in this example a stationary system T on $L_2(-\infty\ \infty)$. Once more the summary given in Appendix A will be relied on. Because this example takes place entirely in the transform domain the notation x rather than $\hat{x}$, for instance, will be used. In Fourier transform notation our example system is modeled as

$$x(\omega) = (u(\omega) + \xi)/(j\omega + a), \qquad \omega \in (-\infty, \infty),$$

where $a > 0$, which is also recognizable as an equivalent model to that of a first-order stationary differential system with initial condition ξ at $t = 0$. The following identifications will be made: $(Tu)(\omega) = (j\omega + a)^{-1}u(\omega)$, $(T^*z)(\omega) = (-j\omega + a)^{-1}z(\omega)$, and $(N\xi)(\omega) = (j\omega + a)^{-1}\xi$. It follows immediately that

$$(6) \qquad (I + T^*T) \sim \frac{\omega^2 + b^2}{\omega^2 + a^2} = \frac{-j\omega + b\ j\omega + b}{-j\omega + a\ j\omega + a} \sim F^*F,$$

where $b^2 = a^2 + 1$ and that

$$T^*N\xi \sim \frac{1}{-j\omega + a} \cdot \frac{\xi}{j\omega + a}.$$

In Appendix A the set E of all L_2 functions with Laplace transforms analytic in the right half-plane is introduced. It is known that the convolution T is causal if and only if the transfer function associated with T is in E. Similarly, since our system is at rest for $t \leq 0$, all controls must be contained in E which is a closed linear subspace. In our optimization problem we take $L = E$ in which case the projection P can be viewed as the familiar operation of taking the left half-plane part of the transform in question. It is clear then that (6) has identified the proper factorization of $I + T^*T$.

To continue with the solution we have, for the left half part,

$$PF^{*-1}T^*N\xi = \left\{ \frac{\xi}{(-j\omega + b)(j\omega + a)} \right\} = \frac{\xi}{a + b} \cdot \frac{1}{j\omega + a},$$

which results in the optimal control

$$u_0(\omega) = \frac{\xi}{a + b} \cdot \frac{1}{j\omega + b}.$$

The system response to this control is given by $x_0(\omega) = \xi/(j\omega + b)$ and hence $u_0(\omega) = -x_0(\omega)/(a + b)$ defines the optimal control in feedback form.

With regard to the above example it is noted that L-causal mappings include all transfer functions with no poles in the right-half plane. Similarly, L-miniphase mappings include all transfer functions with no poles or zeros in the right-half plane. For instance, the mapping F of Example 1 is L-miniphase.

It is of interest at this point to consider the projection P (i.e. taking left-half plane parts) as an operator on $L_2(-\infty, \infty)$. A moment's reflection reveals that P is L-miniphase. Also P is obviously not invertible. Thus the concepts of L-causality and L-miniphase are independent of the invertibility of the operator in question. However any invertible L-causal operator with an L-causal inverse is L-miniphase.

3. Multivariate Differential Systems. In this section we consider systems described by families of linear differential equations acting between cartesian products of $L_2(t_0, t_1)$. The notation to be used in this study is embodied in the equations

(7)
$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ \dot{y}(t) = C(t)x(t), \qquad t \in [t_0, t_1]. \end{cases}$$

Here x, y, u are n-, r- and m-tuples of functions from $L_2[t_0, t_1]$. The matrices A, B and C are of compatible dimensions and, without loss of generality, they are assumed to consist of piecewise continuous functions on $[t_0, t_1]$.

Our attention turns first to the stationary case on the interval $(-\infty, \infty)$. In doing so we shall assume that A has only eigenvalues with negative real parts. In the Fourier domain, this system may be identified with the matrix

(8)
$$T \sim C\Phi(\omega)B, \qquad \omega \in (-\infty, \infty),$$

where $\Phi(\omega) = (j\omega I - A)^{-1}$, $\omega \in (-\infty, \infty)$.

The present study is an obvious generalization of Example 1 and we shall continue the use of material from Appendix A. Of particular interest is the closed linear subspace E of functions analytic in $\sigma \geq 0$ where $s = \sigma + j\omega$ is the Laplace variable. The subspace L is taken as the obvious multivariate generalization of E.

It is evident from Section 2 that the solution of the posed optimization problem hinges on the factorization problem. More generally for an arbitrary subspace L, does an operator F exist such that $F(L) = L$ and $F^*F = I + T^*T$? This question is the subject of Appendix B where the affirmative answer of the Theorem B.1 is established.

THEOREM B.1. *For every bounded transformation T and every closed subspace L there exists an operator F such that $F(L) = L$ and $F^*F = I + T^*T$.*

We note that F is by no means unique for if μ is a unitary operator such that L is a reducing subspace, then $(\mu F)(L) = L$ and $(\mu F)^*(\mu F) = F^*F$.

In the context of (8) the factorization problem takes the following form. For the A, B, C and Φ above, find an invertible matrix F such that F, F^{-1} are analytic in $\sigma \geq 0$ and

$$F^*(\omega)F(\omega) = I + B^*\Phi^*(\omega)C^*C\Phi(\omega)B, \qquad \omega \in (-\infty, \infty)$$

Here Φ^* is the conjugate transpose of Φ. For emphasis we note that Theorem B.1 removes any doubt as to the existence of such an F nor does this existence depend in any way on the fact that Φ is a rational matrix.

In our quest for the operator F the contents of Lemma B.4 and its corollary are directly applicable. These are summarized in

LEMMA 3. *Let K be a constant matrix. If K is selfadjoint and satisfies*

$$(9) \qquad -KA - A^*K = C^*C - KBB^*K,$$

then

$$(10) \qquad I + B^*\Phi^*C^*C\Phi B = (I + B^*K\Phi B)^*(I + B^*K\Phi B).$$

Accordingly our attention focuses on the solution of (9), which is called the algebraic Ricatti equation. First some known results concerning (9). The matrix tuplet $\{A, B, C\}$ is called *regular* if (1) no eigenvector of A with a non-negative real part is a null vector of C, and (2) no eigenvector of A^* with a non-negative real part is a null vector of B^*. It can be shown that this definition is equivalent to the assumption that all unstable open loop modes of (7) are completely controllable and observable. The following lemma is due to Potter [12].

LEMMA 4. *If $\{A, B, C\}$ is regular then (9) has a unique solution $K \geq 0$; moreover, $A - BB^*K$ is stable.*

We shall need also the

LEMMA 5. *Whenever the inverses $(sI - A)^{-1}$ and $(sI - A + BB^*K)^{-1}$ exist, the identity*

$$(I + B^*K\Phi B)^{-1} = I + B^*K[sI - A + BB^*K]^{-1}B$$

holds.

In this lemma it is understood that $\Phi = (sI - A)^{-1}$. A proof can be constructed by direct expansion. The details are left to reader with the suggestion that the two intermediate identities

$$(sI - A + BB^*K)^{-1} = [I + (sI - A)^{-1}BB^*K]^{-1}(sI - A)^{-1}$$

$$B^*K(sI - A + BB^*K)^{-1}B = [I + B^*K(sI - A)^{-1}B]B^*K(sI - A)^{-1}B$$

be verified first and used with the identity $I + (I - \beta)^{-1}\beta = (I - \beta)^{-1}$ which is valid for any linear β such that the right-hand side exists.

A corollary which follows by inspection of the lemma is the following.

COROLLARY. *If the spectrum of $A - BB^*K$ lies in $\sigma < 0$ then $(I + B^*K\Phi B)^{-1}$ is analytic in $\sigma > 0$.*

Using this result with Lemma 4 we have the key result,

THEOREM 3. *If the matrices $\{A, B, C\}$ are regular, then the corollary of Lemma 5 provides a splitting of $I + T^*T$ such that $(I + B^*K\Phi B)(L) = L$.*

In terms of Definitions 1 and 2 with $L = E$ the above development can be interpreted as follows. A necessary and sufficient condition for the matrix $F = I + B^*K\Phi B$ to represent an L-causal operator is that F be analytic in $\sigma > 0$. If, in addition, F^{-1} is analytic in $\sigma > 0$, then F is L-miniphase. If the spectra of A and $A - BB^*K$ lie in $\sigma < 0$, then F and F^{-1} are both L-miniphase and are bounded.

With regard to the above development attention is called to the work of Youla [9] and Anderson [10] who have considered related problems. The tools used here however are quite different and the present treatment extends readily to nonstationary systems.

Nonstationary Systems. Our attention turns now to nonstationary systems. Let ν denote an ordered set and $\{(L_\beta, P_\beta): \beta \in \nu\}$ a family of closed linear subspaces and associated orthogonal projections respectively. It is assumed that $L_{t_0} = \{0\}$, $L_{t_1} = H$, and that $L_\beta \subset L_\gamma$ whenever $\beta \le \gamma$ in ν. This induces a natural ordering on the orthogonal projections; $P_\beta \subset P_\gamma$ if $L_\beta \subset L_\gamma$.

For $u \in H$ let $\{u_\beta = P_\beta u : \beta \in \nu\}$ denote a parameterization of u by projection on the subspaces $\{L_\beta, \beta \in \nu\}$. Here we view u as a function and u_β as a truncation, in the sense indicated, of that function.

We shall have also need for the following definitions.

Definitions (3) The function $u \in H$ is called $\{L_\beta\}$-minimal for J if μ_β minimizes J over L_β, all $\beta \in \nu$.

(4) The operator T is called $\{L_\beta\}$-causal if $T(L_\beta) \subset L_\beta$, all $\beta \in \nu$.

(5) The operator T is called $\{L_\beta\}$-miniphase if $T(L_\beta) = L_\beta$, all $\beta \in \nu$.

A natural questions is that of the existence of $\{L_\beta\}$-minimal functions. This is answered concisely by the following lemma.

LEMMA 6. *For every collection of closed linear subspaces $\{L_\beta : \beta \in \nu\}$, ordered by containment there is a unique $\{L_\beta\}$-minimal element which is specified by $u_\beta = -F^{-1}Q_\beta F^{*-1}T^*N\xi$, $\beta \in \nu$, where Q_β is the orthogonal projection on $M_\beta = F(L_\beta)$.*

Proof. Theorem 2 establishes the existence of the unique u_β for each specific β. It suffices to note that $L_\beta \subset L_\gamma$ implies that $M_\beta \subset M_\gamma$ and hence that $\{u_\beta : \beta \in \nu\}$ is a well-defined parameterization which satisfies the necessary conditions.

An obvious corollary is the following.

COROLLARY. *If F is $\{L_\beta\}$-miniphase, then the result of Lemma 6 simplifies to $u_\beta = -F^{-1}P_\beta F^{*-1}T^*N\xi$, $\beta \in \nu$.*

It is claimed that Lemma 6 solves the optimization problem for large classes of nonstationary systems. The next example will serve to clarify this assertion.

Example 2. Let $\nu = [t_0, t_1]$ and T be the mapping $u \to y$ computed by (7). In this case A, B, C are piecewise continuous time-varying matrices. We may, of course, model T in explicit form by the equation

$$(Tu)(t) = \int_{t_0}^t C(t)\Phi(t, s)B(s)u(s)ds, \qquad t \in \nu.$$

In our discussion of this example we shall need some specific tools which may be found in [2], Section 3.3, and are summarized in the following. First we write $T = C\Phi B$, where B, C are the mappings computed by multiplication with the obvious time-varying matrix, i.e., $(Bu)(t) = B(t)u(t)$ and Φ is the integral map: $(\Phi f)(t) = \int_{t_0}^t \Phi(t, s)f(s)ds$. We note that Φ has a densely defined inverse given by $(\Phi^{-1}x)(t) = \dot{x}(t) - A(t)x(t)$, $t \in \nu$. It is also known that

$$(\Phi^*z)(t) = \int_t^{t_1} \Phi^*(s, t)z(s)ds,$$

and that Φ^{-1} has the formal adjoint

$$(\Phi^{-1*}x)(t) = -x(t) - A^*(t)x(t), \qquad t \in \nu$$

where, in terms of this formal adjoint,

(11) $$\langle v, \Phi^{-1}x \rangle - \langle \Phi^{-1*}v, x \rangle = \sum_j v_i(t_1)x_i(t_1).$$

Use has been made of the fact that $x(t_0) = 0$. Of course, $T^* = B^*\Phi^*C^*$ holds with the natural interpretation, where for instance, $(C^*z)(t) = C^*(t)z(t)$. It is evident that P_β is the orthogonal projection of $L_2^m(\nu)$ onto L_β.

We shall use Lemma B.3 of Appendix B and since Φ is one-to-one we write part (i) of that lemma in the form

(12) $$C^*C = K\Phi^{-1} + \Phi^{*-1}K^* + K^*BB^*K.$$

Consider the equation $z = \Phi^{*-1}K^*x$ for some suitable x. From the above remark it is clear that

$$z(t) = -\left\{\frac{d}{d_t} + A^*(t)\right\}K^*(t)x(t), \qquad K^*(t_1)x(t_1) = 0.$$

Here the boundary condition is necessary to make the right-hand side of (11) zero. If we let both sides of (12) operate on a typical function x, the identity

$$0 = \{-C^*(t)C(t) - K(t)A(t) - A^*(t)K^*(t) - \dot{K}^*(t) + K^*(t)B(t)B^*(t)K(t)\}x(t)$$
$$+ \{K(t) - K^*(t)\}\dot{x}(t),$$

plus the condition $K^*(t_1)x(t_1) = 0$, results. Since this must hold for all x, obviously K must be selfadjoint and satisfy

(13)
$$-K(t) = A^*(t)K(t) + K(t)A(t) - K(t)B(t)B^*(t)K(t) + C^*(t)C(t)$$
$$K(t_1) = 0.$$

Our results thus far are summarized in the

LEMMA 7. *If K satisfies (13) and if F is defined by*

$$(Fu)(t) = u(t) + B^*(t)K(t)\int_{t_0}^t \Phi(t, s)B(s)u(s)ds,$$

then $F^*F = I + T^*T$.

Proof. The lemma follows by inspection from Lemma B.3 and the above development.

LEMMA 8. *If $z = Fv$, then F^{-1} may be computed by the equations*

$$v(t) = z(t) - B^*(t)K(t)p(t)$$
$$\dot{p}(t) = [A(t) - B(t)B^*(t)K(t)]p(t) + B(t)z(t), \qquad p(t_0) = 0.$$

Proof. For p satisfying $\dot{p}(t) = A(t)p(t) + B(t)v(t)$ it is clear that $z(t) = v(t) + B^*(t)K(t)p(t)$. If we solve this expression for v and substitute this in the differential equation for p, the lemma follows.

A direct consequence is the

COROLLARY. *The adjoint of F is determined by*

$$(F^*v)(t) = v(t) + B^*(t)\int_t^{t_1} \Phi^*(s, t)K(s)B(s)v(s)ds.$$

The inverse of F^* may be identified with the equations

$$v(t) = x(t) - B^*(t)q(t),$$
$$\dot{q}(t) = -[A^*(t) + K(t)B(t)B^*(t)]q(t) + K(t)B(t)x(t), \qquad q(t_1) = 0,$$

where $v = F^{*-1}x$.

Proof. The integral form of F^* follows without difficulty from Lemma 7. To arrive at the differential equation representation for F^{*-1} it is helpful to first note the identity

$$\int_t^{t_1} \Phi^*(s, t)K(s)B(s)x(s)ds = \Phi^*(t_0, t) \int_{t_0}^{t_1} \Phi^*(s, t_0)K(s)B(s)x(s)ds$$

$$-\int_{t_0}^{t} \Phi^*(s, t)K(s)B(s)x(s)ds.$$

Note also that $\Phi^*(t_0, t)$ is the transition matrix for the formal adjoint system. With these two facts the proof proceeds exactly as in the proof of the lemma.

In Lemmas 7 and 8 a factorization of $I + T^*T$ is given and it is demonstrated that F, F^* and their respective inverses can be computed by direct means. The optimization problem of interest is the minimization of the functional

$$J(u) = \int_{t_0}^{t_1} [u(t), u(t)]dt + \int_{t_0}^{t_1} [y(t) + (N\xi)(t), y(t + (N\xi)(t)]dt,$$

where $[,]$ denotes the inner product on E^n. The subspaces $\{L_\beta\}$ are defined by

$$L_\beta = \{u \in L_2^m(\nu): u(t) = 0, t \le \beta\}, \qquad \beta \in \nu.$$

The orthogonal projections of $L_2^m(\nu)$ onto the $\{L_\beta\}$ are evidently given by

$$(P_\beta u)(t) = \begin{cases} u(t) & \beta \le t \le t_1 \\ 0 & t_0 \le t < \beta. \end{cases}$$

From Lemma 7 it is readily shown that F is $\{L_\beta\}$-causal. Indeed we have that $(I - P_\beta)FP_\beta \equiv 0$ for all $\beta \in \nu$. To see that F is $\{L_\beta\}$-miniphase, consider Lemma 8. For arbitrary $z \in L_\beta$ clearly $p \in L_\beta$ and hence $v \in L_\beta$. This shows that F^{-1} is $\{L_\beta\}$-causal and hence both F and F^{-1} are $\{L_\beta\}$-miniphase.

To summarize, the element $u \in L_2^m(\nu)$ is $\{L_\beta\}$-minimal for J if and only if u has the parameterization $\{u_\beta\}$, where u_β is computed by Lemma 6. The mappings F^{*-1} and F^{-1} are computed as specified by Lemma 8 and its corollary.

Equation (13) above is known as the matrix Riccati differential equation. There are several interesting relationships between the algebraic and differential equations, however, we shall not explore these here. The interested reader is referred to [3] and [19] which will serve as entries to this literature.

4. The Optimal Regulator Problem. One specific optimization problem which is included by the present theory is the optimal regulator problem. This class of special problems arises when the function $N\xi$ is induced by the free response of the plant. A well-known result for both systems treated in Section 3 is that the optimal regulator can be realized by matrix feedback of the plant state vector. In view of the importance of this basic result it is of interest to see how it falls out of the present formulation.

In the following it is convenient to use the notation $(N\xi)(t) = C(t)(\pi\xi)(t)$, $t \in \nu$, where $(\pi\xi)(t) = \Phi(t, t_0)\xi$, $t \in \nu$. In the stationary case we have that $\pi\xi = \Phi\xi$. In the time-varying case, however, Φ has been used to denote an integral operator and the use of the symbol π as above is a strategem to avoid delta functions.

A useful tool in the development is the content of the following result.

LEMMA 8. $T^*C\pi\xi = F^*B^*K\pi\xi$.

Proof. In the stationary case it follows from Equation (B.2) of Appendix B that

$$B^*\Phi^*C^*C\Phi = B^*\Phi^*K + (I + B^*\Phi^*KB)B^*K^*\Phi.$$

Now since $T^*C\pi\xi = B^*\Phi^*C^*C\Phi\xi$ and in view of the fact that

$$F^*B^*K\pi\xi = (I + B^*K\Phi B)^*B^*K\Phi\xi,$$

the lemma follows by inspection.

For the time-varying case we begin with (13) which implies that

$$\Phi^*(t, t_0)C^*(t)C(t)\Phi(t, t_0) = -\frac{d}{dt}\{\Phi^*(t, t_0)K(t)\Phi(t, t_0)\}$$

$$+ \Phi^*(t, t_0)K(t)B(t)B^*(t)K(t)\Phi(t, t_0),$$

where the defining equation for Φ has been used in an obvious manner. Integrating both sides and using $K(t_1) = 0$ results in

$$\int_t^{t_1} \Phi^*(s, t_0)C^*(s)C(s)\Phi(s, t_0)ds = \Phi^*(t, t_0)K(t)\Phi(t, t_0)$$

$$+ \int_t^{t_1} \Phi^*(s, t_0)K(s)B(s)B^*(s)K(s)\Phi(s, t_0)ds.$$

Now in view of the defining equation for T^*, namely

$$(T^*x)(t) = B^*(t)\Phi^*(t_0, t)\int_t^{t_1} \Phi^*(s, t_0)C^*(s)x(s)ds, \qquad t \in \nu,$$

and the assumed form of $\pi\xi$ it follows from the above identity that

$$(T^*C\pi\xi)(t) = B^*(t)\{K(t)\Phi(t, t_0)$$

$$+ \Phi^*(t_0, t)\int_t^{t_1} \Phi^*(s, t_0)K(s)B(s)B^*(s)K(s)\Phi(s, t_0)ds\}\xi.$$

Consider now the computation of $F^*B^*K\pi\xi$. In view of the corollary to Lemma 8 we have that

$$(F^*B^*K\pi\xi)(t) = \{B^*(t)K(t)\Phi(t, t_0)$$

$$+ B^*(t)\int_t^{t_1} \Phi^*(s, t)K(s)B(s)B^*(s)K(s)\Phi(s, t_0)ds\}\xi$$

$$= B^*(t)\{K(t)\Phi(t, t_0)$$

$$+ \Phi^*(t_0, t)\int_t^{t_1} \Phi^*(s, t_0)K(s)B(s)B^*(s)K(s)\Phi(s, t_0)ds\}\xi$$

By direct comparison the lemma is established.

A direct consequence of the last lemma is that $B^*K\pi\xi = F^{-1}T^*C\pi\xi$ which implies that $u_0 = -F^{-1}PB^*K\pi\xi$. Here, in the time-varying case we have suppressed the subscript β on both u_0 and P. With this understanding the equality $x_0 = \Phi P_0 u_0 + \pi\xi$ holds, which, together with the form of u_0, results in $x_0 = \{I - \Phi BF^{-1}PB^*K\}\pi\xi$; hence

$$-B^*Kx_0 = -\{(I - P)B^*K + F^{-1}PB^*K\}\pi\xi$$

$$= -(I - P)B^*K\pi\xi + u_0.$$

In the stationary case the assumption that the spectrum of A lies in $\sigma < 0$ implies that $\pi\xi$, and hence $B^*K\pi\xi$, lies in L. Thus $u_0(t) = -B^*Kx_0(t)$, which is the desired result. In the time-varying case, operating with P we see that it suffices to consider $u_{0,\beta} = -P_\beta B^* Kx_{0,\beta}$. Now B^*K is a matrix multiplier and in view of the definition of P_β it follows that $(P_\beta z)(\beta) = z(\beta)$ holds for all z and β. Consequently

$$u_{0,\beta}(\beta) = u_0(\beta) = -B^*(\beta)K(\beta)x_0(\beta), \qquad \beta \in \nu,$$

which establishes the desired result.

It is sometimes possible to implement the optimal control in output feedback form. In this regard let $y_0 = Tu_0 + N\xi$ with the optimal control, u_0, for example of (5). By a simple substitution we then have

$$(I - F^{-1}PF^{*-1}T^*T)u_0 = -F^{-1}PF^{*-1}T^*y_0.$$

Thus when the indicated inverses exist

(14)
$$\begin{aligned} u_0 &= -(I - F^{-1}PF^{*-1}T^*T)^{-1}F^{-1}PF^{*-1}T^*y_0 \\ &= -F^{-1}PF^{*-1}T^*(I - TF^{-1}PF^{*-1}T^*)^{-1}y_0. \end{aligned}$$

In many cases this solution has dubious practical value. The trouble lies with the fact that this relationship is really implicit in P and unwieldy for solution.

5. The System Sensitivity Problem. In the study of system sensitivity for negative feedback closed loop systems the operator $S = (I + TM)^{-1}$, where T is the forward gain and M is the feedback gain, plays an important role (see [13], [14], [16]). Several authors ([14], [15], [17]) have addressed themselves to the question: does the optimal feedback compensator reduce system sensitivity. In Hilbert spaces the basic requirement is that $0 \le R - S^*RS$, where R is an appropriate weighting operator. The results of the studies cited can be easily obtained from the above development as we shall now show.

The following development utilizes the context of Lemma B.3 of Appendix B. Here $T = C\Phi B$ and $u_0 = -B^*Kx_0$ where K satisfies Lemma B.3. It is noted for later use that we do *not* require any properties from the frequency domain, that is, the development is carried out in an abstract operator setting.

Attention is called first to Figure 3, where the plant and the controller have been depicted as compositions of the operators $\{B, C, \Phi, K, B^*\}$. It is of interest to consider disturbances in each and/or combinations of these component transformations. In this regard one can consider any two distinct nodes of Figure 3 as an input/output position and thereby define a feedback system. For example, to consider the sensitivity to changes in K let e_3 be viewed as an input and e_4 as an output. The feed forward transmission then is K while the feedback transmission is $-\Phi BB^*$.

As a sample calculation, consider the case where $e_1 \rightarrow e_5$ is the forward transmission, namely $B^*K\Phi B$, and the feedback is unity, thus making the identifications $T = B^*K\Phi B$ and $M = I$. It follows that

$$S = (I + TM)^{-1} = (I + B^*K\Phi B)^{-1} = F^{-1}.$$

By direct calculation we have that

$$I - S^*S = S^*\{S^{*-1}S^{-1} - I\}S = S^*T^*TS \geq 0$$

and the desired sensitivity property is established.

In considering the sensitivity question between other node pairs it is convenient to present first a generalization of the problem under study. In this regard $\{T_1, \cdots, T_n\}$ will denote a collection of linear operators. With no loss of generality we assume that this collection is finite and that all operators act between a single Hilbert space H. In Figure 4 this operator family is used as branch transmissions on a closed loop signal flow graph. The inputs and outputs on the graph have been suppressed; however, it is assumed that an input-ouput pair can be attached to any two distinct nodes.

For instance, suppose an output is attached to the node e_n and an input to the node e_j. We have then a closed loop system with forward transmission $T = T_n T_{n-1} \cdots T_j$ and feedback transmission $M = T_{j-1} \cdots T_2 T_1$. The pair $\{T, M\}$ is said to be an *insensitive partition* of the loop if the sensitivity operator $S = (I + TM)^{-1}$ exists (that is, the indicated inverse exists) and if, for some self-adjoint non-negative operator R, the condition

(15) $0 \leq R - S^*RS$

holds.

LEMMA 12. *If an insensitive partition of the loop exists, all partitions are insensitive partitions.*

Proof. First let V_a and V_b be any two operators on H such that $I + V_a$ and $I + V_b$ are invertible and define $S_a = (I + V_a)^{-1}$ and $S_b = (I + V_b)^{-1}$ respectively. Suppose that Γ is an operator such that $\Gamma V_a = V_b \Gamma$. It then follows that $\Gamma(I + V_a) = (I + V_b)\Gamma$ if and only if $S_b \Gamma = \Gamma S_a$. If R_b is an arbitrary operator on H then

$$\Gamma[R_b - S_b^* R_b S_b]\Gamma = \Gamma^* R_b \Gamma - S_a^* \Gamma^* R_b \Gamma S_a$$
$$= R_a - S_a^* R_a S_a,$$

where $R_a = \Gamma^* R_b \Gamma$.

Suppose now that an insensitive partition of the loop has been found. Since the subscripts are arbitrary, let the output node be labeled e_n and the input node e_a. We have that $T = T_n \cdots T_a$, $M = T_{a-1} \cdots T$, $TM = T_n \cdots T_a$ while Equation (a) holds for some $R \geq 0$, and $S = (I + TM)^{-1}$.

Let e_i, e_j be any two distinct nodes. If we take e_i as the input and e_j as the output, it follows that the forward loop transmission is $T_{ji} = T_{j-1} T_{j-2} \cdots T_i$ while the feedback transmission is $M_{ij} = T_{i-1} \cdots T_1 T_n \cdots T_j$, and consequently $T_{ji} M_{ij} = T_{j-1} \cdots T_1 T_n \cdots T_j$. Now let $\Gamma = T_n \cdots T_j$ and note that $\Gamma(T_{ji}M_{ij}) = (TM)\Gamma$. If we use the above result, then it follows that $0 \leq R_j - S_j^* R_j S_j$ holds, where $\Gamma_j = \Gamma^* R \Gamma$ and $S_j = (I + T_{ji}M_{ij})^{-1}$.

In the above proof we have tacitly assumed the invertibility of the operators $I + T_{ji}M_{ij}$, $i, j = 1, \cdots, n$. From a physical point of view this is closely related to the stability (that is input/output boundedness) between the respective nodes. Since all physical systems have saturation the basic assumption of linearity is enough to justify the invertibility assumptions.

From a mathematical viewpoint, it is not necessary that $I + TM$ be invertible. In short, it can be shown that the sensitivity condition,

$$(16) \qquad (I+TM)^*R(I+TM) - R \geq 0,$$

is valid for all $I + TM$ and is equivalent to (15) when the subject inverse exists. If we use (16) in place of (15), it can easily be seen that the last lemma is still valid.

Returning now to the optimal system of Figure 3, if we combine the above results we have the

COROLLARY. *If K satisfies Lemma B.3, then a Γ_a exists such that $0 \leq \Gamma_a^* \Gamma_a - S_a^* \Gamma_a^* \Gamma_a S_a$, where S_a is the sensitivity operator between any two distinct loop node pairs. In particular this relation holds for*

$$
\begin{array}{llll}
(i) & \Gamma_a = I & T_a = B^*K\Phi B & M_a = I \\
(ii) & \Gamma_a = B^* & T_a = K\Phi B & M_a = B^* \\
(iii) & \Gamma_a = B^*K & T_a = \Phi B & M_a = B^*K \\
(iv) & \Gamma_a = B^*KB & T_a = \Phi & M_a = B^*KB
\end{array}
$$

Proof. It remains only to verify that $\Gamma_a T_a M_a = B^*K\Phi B\Gamma_a$ in each case.

The corollary has the following interpretation in terms of Figure 2. Part (i) deals with the forward transmission $e_1 \to e_5$. Part (ii) deals with the forward transmission $e_1 \to e_4$ with feedback transmission $e_4 \to e_1$. Part (iii) deals with the forward transmission $e_1 \to e_3$ and the feedback transmission $e_3 \to e_1$. Part (iv) deals with the forward transmission $e_2 \to e_3$ and the feedback transmission $e_3 \to e_2$. In all cases (15) is satisfied by an appropriate choice of a positive self-adjoint operator R.

It is known that, except in case (i), R may not necessarily be chosen arbitrarily if the sensitivity criterion is to be satisfied. In particular, in case (iii), (18) holds with $R = I$ if and only if K commutes with BB^* (see [17]).

The implementation of the optimal control in output feedback form, rather than state feedback form, suggests similar problems in sensitivity analysis. To illustrate, consider the forward gain T and the optimal feedback law $u_0 = -My_0$ where M is specified in (14). Consequently

$$I + TM = I + TF^{-1}PF^{*-1}T^*(I - TF^{-1}PF^{*-1}T^*)^{-1} = (I - TF^{-1}PF^{*-1}T^*)^{-1},$$

and hence

$$(19) \qquad S = I - TF^{-1}PF^{*-1}T^*.$$

Now for the present case

$$
\begin{aligned}
I - S^*S &= I - (I - TF^{-1}PF^{*-1}T^*)(I - TF^{-1}PF^{*-1}T^*) \\
&= TF^{-1}P(2I - F^*T^*TF^{-1})PF^{*-1}T^* \\
&= TF^{-1}PF^{*-1}(2F^*F - T^*T)F^{-1}PF^{*-1}T^*,
\end{aligned}
$$

where $P = P^2$ was used. Noting that $I + T^*T = F^*F$ we then have

$$I - S^*S = TF^{-1}PF^{*-1}(2I + T^*T)F^{-1}PF^{*-1}T^*,$$

the right-hand side of this expression being positive.

6. Comments and Conclusions. In the body of the paper the concepts of causality and minimum phase were introduced in conjunction with the optimization problem. These concepts are of interest in their own right and we shall first comment on them.

First it is apparent that the Hilbert-space structure is extraneous to the definitions. It suffices to have two sets X, Y with parameterizations $\{X_\beta, \beta \in \nu\}$, $\{Y_\beta, \beta \in \nu\}$, where ν is an ordered set and the parameterizations are ordered by containment. A function $f: X \to Y$ is called ν-causal if $f(X_\beta) \subset Y_\beta$ for all $\beta \in \nu$ and/or ν-miniphase if $f(X_\beta) = Y_\beta$ for all $\beta \in \nu$. Such definitions generalize in an obvious way the definition offered in Section 3. However, it should be noted that these latter two definitions can be used in Section 3 without disturbing that development. Indeed the Hilbert-space structure, and the linearity and closedness of L_β all bear on the optimization problem but not on the concepts of causality and minimum phase. Finally these latter two concepts do not require boundedness, invertibility or linearity of the mapping in question.

While the concepts of causality and minimum phase do not require the optimization problem for support, the converse is also true. Namely, it is not necessary, only convenient, to construct a minimum phase factorization of $I + T^*T$. This was pointed out for instance in Lemma 6. It can be shown, however, with only slight modifications to the proof of Theorem B. 1, that for every family of ordered (by containment) subspaces $\{L_\beta\}$ there exists an L_β-miniphase operator F such that $F^*F = I + T^*T$.

It is interesting to note that Theorem 1 does not require linearity of the function V. This allows certain classes of nonlinear problems to be brought within the framework of this study.

Suppose that all function spaces are real and the J is of the form $J(u) = \|u\|^2 + \|f(u) + q\|^2$. Here f is a nonlinear mapping from H_1 to H_2 which is everywhere Fréchet differentiable. That is, for small $\|\delta u\|$ the expansion $f(u + \delta u) = f(u) + T(u)\delta u$ holds to a first order where $T(u)$ is linear in δu. By direct methods it follows, to first order in $\|\delta u\|$, that

$$J(u + \delta u) - J(u) = 2\langle u + T^*(u)q + T^*(u)f(u), \delta u \rangle$$

Now if the mapping V defined by

$$V(u) = u + T^*(u)q + T^*(u)f(u)$$

satisfies the hypotheses of Theorem 3, then there exists a unique $u_0 \in L$ which minimizes J over L; moreover $V(u_0) \in L^\perp$. The solution of the subsequent nonlinear equations are more difficult than in the linear case and we shall not consider these matters here.

While the examples given here all arise from finite families of ordinary differential equations, the results apply as well to discrete time and distributive systems. We shall not consider these applications. However, the interested reader will find, in [2] (Appendix 9) and [20] an entry to this literature.
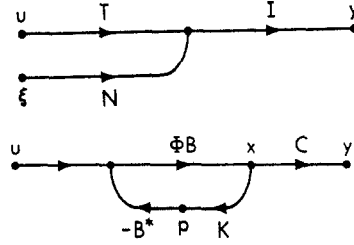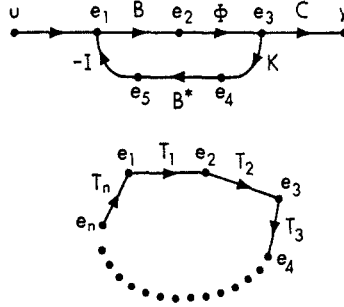
Figure 1. A Typical Linear Plant

Figure 2. A Single Loop System

Appendix A. A Summary of Fourier Analysis

Throughout this appendix L_p will denote the usual Lebesgue space on the infinite interval $(-\infty, \infty)$. That is, L_p consists of all measurable complex-valued functions f of a real variable such that

$$\|f\| = \left[\int_{-\infty}^{\infty} |f(t)|^p \, dt \right]^{1/p} < \infty$$

holds (with integration in the Lebesgue sense). The cases $p = 1, 2$ will be of primary interest. The Hilbert space L_2 is equipped with the inner product $(,)$ where

$$(x, y) = \int_{-\infty}^{\infty} x(s)\bar{y}(s) \, ds, \qquad x, y \in L_2.$$

The following discussion extends easily to finite products of L_2. The notation L_2^n will denote the space of all tuplets

$$f = (f_1, \cdots, f_n), \qquad f_i \in L_2$$

being finite with respect to the norm

$$\|f\| = \left[\sum_{i=1}^{n} \int_{-\infty}^{\infty} |f_i(t)|^2 \, dt \right]^{1/2}.$$

Again L_2^n is a Hilbert space with the inner product $\langle , \rangle$, where

$$\langle f, g \rangle = \sum_{i=1}^{n} \int_{-\infty}^{\infty} f_i(t)\bar{g}_i(t)dt \qquad f, g \in L_2^n.$$

L_2^n is equipped with the usual algebraic operations.

In dealing with physical systems an important notion is that of stationarity. Let $D(T)$ denote the domain of T and let K_a, for real a, denote the time translation operator

$$(K_a x)(t) = x(t-a), \qquad x \in L_2^n.$$

Then T is *stationary* whenever

(1) $D(T) = K_a D(T)$ for all real a;

(2) $TK_a = K_a T$ for all real a.

In other words if $x \in D(T)$ and if $y = Tx$ then $K_a x \in D(T)$ and $K_a y = TK_a x$.

In dealing with stationary systems Fourier transforms play an important role. The Fourier transform $\mathbf{F}$ of a function $x \in L_2$ is to be defined by the expression

$$(\text{A.1}) \qquad (\mathbf{F}x)(\omega) = \underset{N \to \infty}{\text{l.i.m.}} \left(\frac{1}{2\pi}\right)^{1/2} \int_{-N}^{N} e^{j\omega t} x(t)\, dt.$$

Here l.i.m. denotes limit in the mean. Some salient properties of $\mathbf{F}$ are summarized in the following theorem of Plancherel (see [5], page 51).

THEOREM A.1. *As an operator on L_2, $\mathbf{F}$ is one-to-one, onto, and norm-preserving. The inverse of $\mathbf{F}$ is determined by*

$$(\text{A.2}) \qquad (\mathbf{F}^{-1}y)(t) = \underset{N \to \infty}{\text{l.i.m.}} \left(\frac{1}{2\pi}\right)^{1/2} \int_{-N}^{N} e^{-j\omega t} y(\omega)\, d\omega.$$

In other words $\mathbf{F}$ is an isometric isomorphism of L_2 onto itself. A consequence of this is that for $x, y \in L_2$ $(\mathbf{F}x, \mathbf{F}y) = (x, y)$; moreover $\mathbf{F}^* = \mathbf{F}^{-1}$. The Fourier transform is also extended to the space L_2^n in the natural way, the Fourier transform of a tuplet being the tuplet of component Fourier transforms. We remark only that the inner product relationship takes the form

$$\langle \mathbf{F}x, \mathbf{F}y \rangle = \sum_{i=1}^{n} \int_{-\infty}^{\infty} (\mathbf{F}x)_i(\omega)\, \overline{(\mathbf{F}y)_i(\omega)}\, d\omega$$

$$= \sum_{i=1}^{n} \int_{-\infty}^{\infty} x_i(t)\bar{y}_i(t)\, dt = \langle x, y \rangle$$

The next result which is apparently due to Bochner (see [6]) is a cornerstone in the study of stationary systems. In this theorem $y = Tx$ is mapping from L_2^m into L_2^n while $\hat{y}$ and $\hat{x}$ denote the L_2 Fourier transform of y and x respectively. $\hat{T}$ denotes an $n \times m$ matrix of measurable functions.

THEOREM A.2. *A necessary and sufficient condition for T to be linear, bounded and stationary is that $\hat{y}(\omega) = \hat{T}(\omega)\hat{x}(\omega)$, $\omega \in (-\infty, \infty)$ where each component of the matrix $\hat{T}$ is a uniformly bounded measurable function.*

This theorem may be paraphrased as stating that the transformation T from L_2^m into L_2^n is linear, bounded and stationary if and only if T has a multiplicative form. Since $y = Tx$ if and only if $\hat{y} = \hat{T}\hat{x}$, the matrix $\hat{T}$ which represents T in its multiplicative form can be represented as $\hat{T} = \mathbf{F}T\mathbf{F}^{-1}$, as we see from the equality chain $\hat{y} = \mathbf{F}y = \mathbf{F}Tx = \mathbf{F}T\mathbf{F}^{-1}\mathbf{F}x$.

Convolutions. The examples to be presented later come from the class of systems which may be identified with convolution operators. The convolution operator $f \otimes$ is defined by

$$(A.3) \qquad (f \otimes g)(t) = \int_{-\infty}^{\infty} f(t-s)g(s) \, ds, \qquad t \in (-\infty, \infty).$$

The well-known operational properties: $f \otimes g = g \otimes k$ and $f \otimes (g \otimes k) = (f \otimes g) \otimes k$ are easily verified. The convolution of an $n \times m$ matrix and an $m \times k$ matrix is defined in the obvious manner. In particular, the case $m \times 1$ is of interest for, as we shall see, such convolutions can define the type of linear transformation under discussion.

The next theorem (see [7], p. 951) isolates several important facts concerning convolutions on L_1 and L_2.

THEOREM A.3. *For f, $x \in L_1$ the convolution $f \otimes x$ is well defined and satisfies $\|f \otimes x\|_1 \leq \|f\|_1 \cdot \|x\|_1$. For $f \in L_1$, $x \in L_2$ the convolution $f \otimes x$ exists in L_2 and satisfies $\|f \otimes x\|_2 \leq \|f\|_1 \cdot \|x\|_2$. If f, $x \in L_2$ the convolution $f \otimes x$ defines a continuous function with norm (sup norm) at most $\|f\|_2 \cdot \|x\|_2$.*

In view of Equation (A.3) and this theorem it is clear that for $f \in L_1$ the convolution $f \otimes$ defines a bounded linear transformation on both L_1 and L_2. With the domain of $f \otimes$ being the entire space L_2, it is easily verified that this operator is stationary. The theorem also generalizes easily to the multivariate setting. For instance, if W is an $n \times m$ matrix of functions $W_{ij} \in L_1$ such that $|W| \in L_1$, where $|W|(t) = |W(t)|$ denotes the norm of $W(t)$ as a mapping from $l_1(m)$ into $l_1(m)$, then $W \otimes$ is a bounded linear stationary transformation sending L_2^m into L_2^n with norm satisfying $\|W \otimes\| \leq \int_{-\infty}^{\infty} |W(t)| \, dt$. Finally it is noted that $f \otimes$, as an operator on L_2, has a Hilbert space adjoint $(f \otimes)^*$. This adjoint is itself a convolution, namely $(f \otimes)^* = \tilde{f} \otimes$, where $\tilde{f}(t) = f(-t)$, $t \in (-\infty, \infty)$. More generally, for any bounded linear stationary system T, acting between finite products of L_2 with $\hat{T}$ the matrix multiplicative representation T, the identity chain $\langle T^*z, x \rangle = \langle z, Tx \rangle = \langle \hat{z}, \hat{T}\hat{x} \rangle = \langle (\hat{T})^*\hat{z}, \hat{x} \rangle$ shows that $(\hat{T})^*$, the conjugate transpose of T, is the multiplicative matrix representation of T^*.

Nonanticipatory Systems. Heuristically a nonanticipatory system is one in which present values of the output are not influenced by future values of the input. To sharpen this somewhat let P_τ, for real τ, denote the projection operator defined by

$$(P_\tau x)(t) = \begin{cases} x(t) & -\infty < t \leq \tau \\ 0 & \tau < t. \end{cases}$$

In other words, P_τ is computed by multiplication with the characteristic function of the interval $(-\infty, \tau)$. A function T is said to be nonanticipatory if for every x_1, $x_2 \in D(T)$ such that $P_\tau x_1 = P_\tau x_2$, for any real τ, then $P_\tau T x_1 = P_\tau T x_2$.

A convolution $f \otimes$ is non anticipatory if and only if $f(t) = 0$ for $t \leq 0$. In this case with $y = f \otimes x$ we have

$$y(t) = \int_{-\infty}^{t} f(t-s)x(s) \, ds = \int_0^{\infty} f(s)x(t-s) \, ds, \qquad t \in (-\infty, \infty),$$

and the output y at any $t \in (-\infty, \infty)$ clearly depends only on past values of the input. Similarly a convolution $f \otimes$, where $f(t) = 0$ for $t \geq 0$, is called *purely anticipatory*. In this case evidently

$$(f \otimes x)(t) = \int_t^\infty f(t-s)x(s)\,ds, \qquad t \in (-\infty, \infty),$$

and hence present output values depend only on future values of the input.

The Fourier transform representation of a nonanticipatory convolution has a certain familiar and important property. To state this result we introduce the complex Fourier transform on L_2 by means of definition

$$\hat{x}(s) = (\mathbf{F}x)(s) = \left(\frac{1}{2\pi}\right)^{1/2} \int_{-\infty}^\infty x(t)\,e^{-st}dt, \qquad s = \sigma + j\omega.$$

From the definition it is not difficult to show that if $x(t) = 0$, $t \leq 0$, then $\hat{x}(s)$ is analytic in the right half (i.e., $\sigma > 0$) of the complex plane. Furthermore, for $\sigma > 0$

$$\int_{-\infty}^\infty |x(\sigma + j\omega)|^2 d\omega = \int_0^\infty |x(t)|^2 e^{-2\sigma t}dt \leq \|x\|^2 < \infty.$$

The converse of this result is also true as is stated in the next theorem (see [8], Section 1.4).

THEOREM A.4. *The two subsets of $L_2(-\infty, \infty)$, $E' = \{x: x(t) = 0, t \leq 0\}$ and $E = \{x: \hat{x}(s) \text{ is analytic in } \sigma > 0, \|\hat{x}(\sigma)\| \leq \|x\|, \sigma > 0\}$, are identical. Moreover, for $x \in E$, $\hat{x}(\omega) = \text{l.i.m.}_{\sigma \to 0+} \hat{x}(s)$, where the limit exists for almost all $\omega \in (-\infty, \infty)$.*

Theorem 4 extends to L_2^n with norms replacing absolute values where necessary.

Suppose now that $x \in L_2$ satisfies $x(t) = 0$ for $t \geq 0$ and that y is the function defined by $y(t) = x(-t)$. Then $y \in L_2$ and $y(t) = 0$ for $t \leq 0$. The equality chain

$$\hat{x}(s) = \int_0^\infty e^{-st}x(t)\,dt = -\int_0^{-\infty} e^{-st}x(-t)\,dt$$

$$= \int_{-\infty}^0 e^{-st}y(t)\,dt = \hat{y}(-s)$$

then shows that the properties set forth in Theorem A.4 for $\hat{y}$ in the half-space $\sigma > 0$ hold also for $\hat{x}$ in the half-space $\sigma < 0$. This observation establishes the corollary: the conjugate of a nonanticipatory convolution is a purely anticipatory convolution. The proof of this statement follows from the above remarks and the fact that if $\Phi(s) \leftrightarrow \varphi \otimes$ then $\Phi(s)^* \leftrightarrow (\varphi \otimes)^*$.

Appendix B. Splitting $I + T^*T$

In this appendix the following theorem is proved.

THEOREM B.1. *For every bounded linear operator T and every closed linear subspace L there exists an operator F such that $F(L) = L$ and $F^*F = I + T^+T$.*

Following the proof of this theorem the operator F is characterized. In

establishing Theorem B.1 some standard notation will be helpful. First let $H = H_1 \times H_3$, equipped with the usual inner product. For fixed linear bounded T sending H_1 into H_3 the graph of T is defined to be the set $\mathscr{T}(T) = \{(x, Tx): x \in H_1\}$. $\mathscr{T}(T)$ is known to be a closed linear subspace[†] of H. Two other linear subspaces, namely, $\mathscr{L}(T) = \{(x, Tx): x \in L\}$ and $\mathscr{M}(T) = \{(x, Tx): x \in L^\perp\}$, will also be of interest in the following. The subspaces $\mathscr{L}(T)$ and $\mathscr{M}(T)$ are closed in $H_1 \times H_3$ and although they are not in general orthogonal they do yield a direct-sum decomposition of $\mathscr{T}(T)$.

The transformation of $M: H_1 \to \mathscr{T}(T)$ is now defined by $Mx = (x, Tx)$, $x \in H_1$. M is linear, bounded, one-to-one and onto $\mathscr{T}(T)$. Moreover, it is readily verified that $M^*: \mathscr{T}(T) \to H_1$ is computed by the rule $M^*(x, Tx) = x + T^*Tx$ and hence $M^*M = I + T^*T$. Thus, in this sense, M provides a type of factorization for $I + T^*T$.

Now, since M is invertible, the subspaces L and $\mathscr{L}$ are closed, in one-to-one correspondence and are in fact congruent. Thus a norm-preserving transformation U_1 exists such that $U_1(\mathscr{L}) = L$.

We shall note that $L^\perp$ and $\mathscr{M}$ are congruent and since $\mathscr{L}$, $\mathscr{M}$ constitute a direct-sum decomposition of $\mathscr{T}(T)$ it follows that U_1 may be extended to all of $\mathscr{T}(T)$ while preserving its unitary properties (see [18]). That is, a transformation U exists such that $U^*U = I$ and $U(\mathscr{L}) = L$. The operator F on H_1 is then defined by $F = UM$. Since $M(L) = \mathscr{L}$, the operator F has the properties $F^*F = M^*U^*UM = M^*M = I + T^*T$ and $F(L) = L$, which establishes Theorem B.1.

Our attention turns now to characterizing unitary transformations from H to H_1. Suppose that S is a linear transformation from H to H_1. Then linear transformations $G: H_1 \to H_1$ and $D: H_3 \to H_1$ exist such that $S(x, y) = Gx + Dy$, $(x, y) \in H$. The adjoint of S is easily determined, being given by $S^*z = (G^*z, D^*z)$, $z \in H_1$, from which it follows that

$$S^*S(x, y) = (G^*Gx + G^*Dy, D^*Gx + D^*Dy), \qquad (x, y) \in H.$$

In the following we shall consider the restriction of S^*S to $\mathscr{T}(T)$ for which we have

(B.1) $(S^*S)(x, Tx) = (G^*Gx + G^*DTx, D^*Gx + D^*DTx, \quad x \in H_1.$

In our first lemma, the notation K is used to denote a square root of $I + T^*T$ (see [3], p. 157), that is, $K = K^*, K^2 = I + T^*T$.

LEMMA B.1. *The following statements are equivalent*: (a) $S^*S = I$; (b) $G^*G + G^*DT = I$ and $D^*G + D^*DT = T$; (c) $D = GT^*$ and $G^*G = K^{-2}$; (d) $G = VK^{-1}$, where $V^*V = I$ and $D = GT^*$.

Proof. In view of equation (B.1), $S^*S = I$ if and only if condition (b) holds; thus (a) and (b) are equivalent. The second part of (b) implies that $T^*T = T^*D^*DT$. Using this with the first part of (b) shows that $I + T^*T = (G + DT)^*(G + DT)$, from which it follows that $(G + DT)$ is invertible. Thus condition (b) may be written as $G^* = (G + DT)^{-1}$ and $D^* = T(G + DT)^{-1}$ from which the first

[†] That $\mathscr{T}(T)$, $\mathscr{L}(T)$ and $\mathscr{M}(T)$ are all closed follows from the boundedness of T (see [2], p. 432).

part of (c) is apparent. Using $D = GT^*$ in either part of (b) completes the proof of (c). Now $G^*G = K^{-2}$ implies that $KG^*GK = I$, which means that $V = GK$ satisfies part (d). Finally for G and D as in (d), it is readily verified that (b), which is equivalent to (a), holds, which completes the proof.

LEMMA B.2. *The following statements are equivalent. There exists a* Q *such that* (e) $T^*T = T^*Q^* + QT + T^*Q^*QT$; (f) $I + T^*T = (I+QT)^*(I+QT)$; (g) $(I+QT)K^{-1} = W$ *where* $W^*W = I$.

Proof. The equivalence of (e) and (f) is made apparent by expanding the right-hand side of (f). If (g) holds, then clearly $K^*W^*WK = (I+QT)^*(I+QT)$. Since $K = K^*$, $K^2 = I + T^*T$ and $W^*W = I$, we see that (g) implies (f). Since K is invertible, (f) implies that $K^{-1}(I+QT)^*(I+QT)K^{-1} = I$, which means that $W = (I+QT)K^{-1}$ satisfies condition (g).

Since Lemmas B.1 and B.2 both deal with conditions for factoring the operator $I + T^*T$, it is not surprising to find that the operators G and Q are related. Using parts (d) and (g), we see that for every $I+QT$ satisfying Lemma B.2 the operator $G = (I+QT)K^2$ satisfies Lemma B.1, and conversely. Thus conditions (a) through (g) are all equivalent.

In some applications, a slight generalization of Lemma B.2 is useful. Assume now that T is the composition $T = C\Phi B$ of linear transformations between compatible spaces.

LEMMA B.3. *Let* K *be a linear transformation. Then the identity*

$$(i) \quad \Phi^*C^*C\Phi = \Phi^*K + K^*\Phi + \Phi^*K^*BB^*K\Phi$$

implies that

$$(ii) \quad I + T^*T = (I + B^*K\Phi B)^*(I + B^*K\Phi B).$$

Proof. The proof of this lemma is similar to our previous considerations. Using (i) we have

$$B^*\Phi^*C^*C\Phi B = B^*\Phi^*KB + B^*K^*\Phi B + B^*\Phi^*KBB^*K^*\Phi B,$$

which, if we add the identity to both sides and inspect, implies identity (ii). When B is nonsingular the reverse implication is also obviously true.

Of special interest in the applications is a factorization of $I + T^*T$ where C, B, K are constant matrices and $\Phi(\omega) = (j\omega I - A)^{-1}$. For this further specialization we have

LEMMA B.4. *If* K *is a constant matrix, then the identity*

(B.2) $$\Phi^*K + K^*\Phi = \Phi^*C^*C\Phi - \Phi^*KBB^*K^*\Phi$$

holds if and only if $K = K^*$ *and*

(B.3) $$-KA - A^*K = C^*C - KBB^*K.$$

Proof. Since Φ is invertible, Equation (B.2) may be rewritten as

$$K\Phi^{-1} + \Phi^{*-1}K^* = C^*C - KBB^*K^*.$$

The right-hand side of this equality is a constant matrix. The left-hand side has the form

$$K\Phi^{-1}+\Phi^{*-1}K^* = j\omega(K-K^*)-(AK+A^*K^*).$$

Thus for Equation (B.2) to hold, it is both necessary and sufficient that $K = K^*$ and that Equation (B.3) holds.

In view of Lemma B.3, an obvious corollary is the following.

COROLLARY. *If K is a selfadjoint matrix satisfying equation (B.3), then*

$$I+T^*T = (I+B^*K\Phi B)^*(I+B^*K\Phi B).$$

If B is nonsingular, then the converse is true.

Remark. Attention is called to [11] which contains the following development. Suppose that β denotes a commutative Banach algebra with identity I and norm ρ. The operator P on β is called a projection if $P = P^2$ and if $xy \in R(P)$ whenever $x, y \in R(P)$. The operators P_+ and $P_- = I-P_+$ are two such projections with ranges β_+ and β_- respectively. Also $\tilde{\beta}_+$ and $\tilde{\beta}_-$ denote subspaces spanned by $\{I, \beta_+\}$ and $\{I, \beta_-\}$ respectively. The result of [11] of interest here is the following.

LEMMA. *For every $0 \neq Z \in \beta$ with $\rho(Z) < 1$ elements $Z_+ \in \tilde{\beta}_+$ and $Z_- \in \tilde{\beta}_-$ exist such that (a) $I+Z = Z_-Z_+$ and (b) $Z_+^{-1} \in \tilde{\beta}_+$, $Z_-^{-1} \in \tilde{\beta}_-$. Moreover, $Z_+ = \exp \{P_+[\log (I+Z)]\}$, $Z_- = \exp \{P_-[\log (I+Z)]\}$, $Z_+^{-1} = \exp \{-P_+[\log (I+Z)]\}$ and $Z_-^{-1} = \exp \{-P_-[\log (I+Z)]\}$.*

The proof of the lemma uses in an essential way the commutative property of β.

Suppose we specialize the context of this appendix with the assumptions (1) $\|T\| < 1$, and (2) the class of operators (of the form T^*T) in question is a commutative Banach algebra. Then this lemma can be used to sharpen somewhat the basic result. To do this it is necessary only to choose $\beta_+ = \{X \in \beta: X(L) = L\}$. Conversely, when Z is selfadjoint, the development of the appendix removes the assumption of commutativity in the factorization of [11].

Appendix C. The Projection Q

Let $h \in H$ be arbitrary and L, M, P, Q be as defined in Section 2. We wish to compute the orthogonal projection Q, that is, to find elements $h_1 \in M$, $h_2 \in M^\perp$ such that $h = h_1+h_2$. It is apparent that $h_1 = Fu$ for some $u \in L$. It is also true that $h_2 = F^{*-1}v$ for some $v \in L^\perp$. This latter statement follows from the implication chain: $h_2 \in M^\perp \Leftrightarrow \langle h_2, FPu \rangle = 0$ for all $u \Leftrightarrow \langle PF^*h_2, u \rangle = 0$ for all $u \Leftrightarrow PF^*h_2 = 0 \Leftrightarrow F^*h_2 \in L^\perp \Leftrightarrow h_2 \in F^{*-1}(L^\perp)$, where the fact that F is invertible has been used in an obvious way. The computation of h_1 is therefore equivalent to finding the (unique) solution to

(C.1) $$h = Fu+F^{*-1}v, \qquad u \in L, \qquad v \in L^\perp.$$

It is convenient to introduce the mapping K from H^2 into H defined by

$$K(u, v) = FPu+F^{*-1}(I-P)v, \qquad (u, v) \in H^2.$$

From the above remarks it follows that K is onto. Moreover, the tuplet (u, v)

which satisfies Equation (C.1) is the minimum norm pre-image of h under K. To establish this suppose that $(\hat{u}, \hat{v})$ is the (unique) minimum norm pre-image of h under K; then $\hat{u} \in L$. Indeed if $\hat{u} = u_1 + u_2$, $u_1 \in L$, $0 \neq u_2 \in L^\perp$, then $FPu_1 = FP\hat{u}$ while $\|u\| > \|u_1\|$ and hence $\|(\hat{u}, v)\| > \|(u_1, v)\|$, which is a contradiction. Similarly, it follows that $\hat{u} \in L^\perp$, and consequently that $(\hat{u}, \hat{v})$ is a solution, and hence the unique solution.

Now $(\hat{u}, \hat{v})$ may be determined by direct means (see [2], Chap. 4). Indeed, $(\hat{u}, \hat{v}) = K^*(KK^*)^{-1}h$, where the invertibility of KK^* follows from the property that K is onto. It is only a moment's calculation to show that $KK^* = FPF^* + F^{*-1}(I-P)F^{-1}$ and consequently that

$$\hat{u} = PF^*[FPF^* + F^{*-1}(I-P)F^{-1}]^{-1}h$$

$$\hat{v} = (I-P)F^{-1}[FPF^* + F^{*-1}(I-P)F^{-1}]^{-1}h.$$

As a check on this result the reader should note that Equation (C.1) is satisfied.

LEMMA C.1. *If Q is the orthogonal projection on $M = F(L)$, then*

$$Q = FPF^*[FPF^* + F^{*-1}(I-P)F^{-1}]^{-1}$$

where P is the orthogonal projection on L.

Proof. It is necessary only to note that $Qh = F\hat{u}$.

To obtain the optimal control let $h = F^{*-1}T^*N\xi$ and compute $-F^{-1}Qh$. However, since $Qh = F\hat{u}$ it follows that $u_0 = -\hat{u}$ and hence we have

COROLLARY 1. *The minimizing element u_0 for J on L is given by*

$$u_0 = -PF^*[FPF^* + F^{*-1}(I-P)F^{-1}]^{-1}F^{*-1}T^*N\xi.$$

We have noted above that when $F(L) = L$ then $Q = P$. This can be shown directly as is done in proving

COROLLARY 2. *$F(L) = L$ implies that $Q = P$.*

Proof. First note that $F(L) = L$ implies that $FP = PFP$ and that $F^{-1}P = PF^{-1}P$. Taking adjoints we have also that $PF^{*-1} = PF^{*-1}P$. Now using the explicit form of Q we have the implication chain

$$Q = P \Leftrightarrow FPF^* = P[FPF^* + F^{*-1}(I-P)F^{-1}]$$

$$\Leftrightarrow (FP - PFP)F^* = (PF^{*-1} - PF^{*-1}P)F^{-1}.$$

The lemma then follows by inspection.

The optimal control u_0 and the projection Q may also be written explicitly in terms of T and P as we see in

COROLLARY 3. *For Q of Lemma C.1 and u_0 of Corollary 1,*

$$Q = FD(I + T^*T)[(I + PT^*T)(I + T^*TP)]^{-1}F^*$$

$$u_0 = -P(I + T^*T)[(I + PT^*T)(I + T^*TP)]^{-1}T^*N\zeta.$$

Proof. With some elementary manipulations the equality chain

$$F^*[FPF^* + F^{*-1}(I-P)F^{-1}]^{-1}F^{*-1} = F^*[I + FF^*(FPF^* - F^{*-1}PF^{-1})]^{-1}F$$
$$= F^*F[I + F^*FPF^*F - P]^{-1}$$

is established. If we use $F^*F = I + T^*T$, this chain may be continued in the form

$$= (I + T^*T)[I + T^*TP + PT^*T$$
$$\qquad + T^*TPT^*T]^{-1}$$
$$= (I + T^*T)(I + PT^*T)^{-1}(I + T^*TP)^{-1}.$$

The corollary then follows from this identity and the lemma.

In this last proof an identity of the form $[I + \Gamma\Delta]^{-1} = \Gamma[I + \Delta\Gamma]^{-1}$ is used, where $[I + \Gamma\Delta]$ and Γ are known to be invertible. If $x = [I + \Gamma\Delta]^{-1}\Gamma y$, then $x = \Gamma y - \Gamma\Delta x$ and hence the range of $[I + \Gamma\Delta]^{-1}\Gamma$ is contained in the range of Γ and the equality $[I + \Gamma\Delta]^{-1}\Gamma = \Gamma\mu$ therefore holds for some linear operator μ. Clearing fractions we have $\Gamma = [I + \Gamma\Delta]\Gamma\mu = \Gamma[I + \Delta\Gamma]\mu$ which implies $[I + \Delta\Gamma]\mu = I$ and in turn the identity.

It is noted also that the identity

$$I + F^*FPF^*F - P = (I + T^*TP)(I + PT^*T)$$

and the invertibility of the left-hand side implies the invertibility of the right.

REFERENCES

[1] L. Markus and E. B. Lee, *Foundations of Optimal Control Theory*, Wiley, New York, 1967.
[2] W. A. Porter, *Modern Foundations of Systems Engineering*, Macmillan, New York, 1966.
[3] L. Liusternik and V. Sobolev, *Elements of Functional Analysis*, Frederick Ungar, New York, 1961.
[4] G. J. Minty, Monotone operators and certain systems of nonlinear ordinary differential equations, *Proceedings of a Symposium on System Theory*, Polytechnic Institute of Brooklyn, 1965.
[5] R. R. Goldberg, *Fourier Transforms*, Cambridge Univ. Press, 1962.
[6] S. Bochner and K. Chandrasekeharan, *Fourier Transforms*, Annals of Mathematics Studies, No. 19, Princeton, 1949.
[7] N. Dunford and J. T. Schwartz, *Linear Operators II*, Wiley, New York, 1963.
[8] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, Amer. Math. Soc. Colloq. Publ. Vol. 19, New York, 1934.
[9] D. C. Youla, On the factorization of rational matrices, *IRE Trans.* IT-7 (1961), 172–189.
[10] B. D. O. Anderson, "Spectral Factorization by Algebra", Stanford Electronic Laboratory Report No. 6560–6, Stanford Univ., Stanford, Calif., Sept. 1966.
[11] G. Zames and P. Falb, Stability conditions for systems with monotone and slope-restricted nonlinearities, *SIAM J. Control* 6 (1968).
[12] J. E. Potter, "A Matrix Equation Arising in Statistical Filter Theory", NASA Report CR-270, August, 1965.
[13] W. A. Porter, Some theoretical limitations of system sensitivity reduction, *Proceedings Allerton Conference on Circuit and System Theory*, Urbana, Illinois, October 1965.
[14] E. Kreindler, "Synthesis of Flight Control Systems Subject to Vehicle Parameter Variations", Technical Report AF-FDL-TR-66-209, December, 1966.
[15] R. E. Kalman, When is a linear control system optimal?, *ASMA J. Basic Engr.*, March, 1964.

[16] W. A. PORTER, On Sensitivity in Multivariate Nonstationary Systems, *Internat. J. Control*
      7 (1968), 481–491.
[17] C. ZAHM, "Structure of Sensitivity Reduction", Ph.D. Thesis, University of Michigan,
      1969.
[18] P. R. HALMOS, Normal dilations and extensions of operators, *Summa Brasil. Math.* 2
      (1950), 125–134.
[19] W. A. PORTER, On the matrix Riccati equation, *IEEE Trans. Automatic Control* AC-12
      (1967), 746–749.
[20] W. A. PORTER, Sensitivity problems in distributive systems, *Internat. J. Control* 5 (1967),
      393–412.