JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **131**, 127–156 (1988)

A Factorization on the Semi-Infinite Interval I: General Theory

MARK H. MILMAN

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109

Submitted by Dr. A. Schumitzky

Received October 22, 1986

This paper extends the methods of special factorization to treat a class of factorization problems on the half-line. Factorizations involving integral operators with stationary and nonstationary kernels are presented. A simple "time domain" connection between the factorization problem and the stable regulator problem in Hilbert space is developed. © 1988 Academic Press, Inc.

1. INTRODUCTION

This paper is concerned with the factorization problem for a class of self-adjoint operators on a Hilbert space H with respect to a given chain of orthoprojectors C on H. It is assumed that these orthoprojectors can be parameterized by a mapping P: $[0, \infty] \to C$ with the properties

(i) P is onto

(ii) $P(0)=0$, $P(\infty)=I$

(iii) $P(t_1) < P(t_2)$ if and only if $t_1 < t_2$

(iv) $P(\cdot)$ is strongly continuous, i.e., $t \to P(t)x$ is continuous for each $x \in H$.

In this context the factorization problem is posed: Given a self-adjoint operator $K \in B(H)$, find $X_+, X_- \in B(H)$ such that

$$I + K = (I + X_-)(I + X_+), \tag{1.1}$$

where $P(t)\, X_+ P(t) = P(t) X_+$ and $P(t)\, X_- P(t) = X_- P(t)$ for all t.

The basic approach to the factorization problem defined above is due to Gohberg and Krein [6]. Definitive results for a large class of compact perturbations of the identity were obtained using projection integral methods. In [7] it is shown that these methods are also applicable when the pertur-

0022-247X/88 $3.00

409/131/1-9

bation of the identity enjoys a certain bounded variation property with respect to the chain C. Applications of the factorization include filtering and smoothing of nonstationary processes over a finite interval [9], a unifying treatment of methods for solving two-point boundary value problems [12], solutions to matrix Riccati equations [18], and inverse problems in the spectral theory of differential operators [11].

Milman and Schumitzky [15] motivated by problems in optimal control introduced the notion of an operator dominated by a (finite) measure—the conditions for which are similar to the bounded variation condition in [7]. The formalism that results from the additional structure induced by the dominating measure permits (what is loosely speaking) the "Lebesgue" analog of the projection integral. This formalism in turn leads to factorization theorems, a simple proof of Volterra inversion, and applications to operator Riccati equations and problems in control and filtering of infinite dimensional systems [13–15].

The classical Wiener–Hopf factorization problem can be placed into the "factorization with respect to a chain" framework by identifying the Hilbert space H with $L_2(0, \infty)$, the projections $P(t)$ with the truncation projections

$$P(t)x : s \to \begin{cases} x(s) & s \leqslant t \\ 0 & s > t, \end{cases}$$

and the operator K with an integral operator with difference kernel $k(t-s)$, $k(\cdot) \in L_1(-\infty, \infty)$. A complete theory exists for the Wiener–Hopf factorization in both the scalar and matrix cases [5]. An important and much studied generalization of the Wiener–Hopf problem involves factoring a nonnegative matrix or operator-valued function $M(\lambda) = \Phi^*(\lambda)\, \Phi(\lambda)$, where Φ is the boundary value function of an operator function which is analytic in the upper-half plane (see, for example, [2, 4, 6, 17]). Applications of the Wiener–Hopf factorization and its generalizations include filtering and prediction of stationary processes [19], optimal control problems on the semi-infinite interval [1, 8], and problems in transport theory [10].

A certain gap exists between the factorization problems that can be treated by the methods that are applicable to the Wiener–Hopf type problems and those problems that are amenable to projection integral methods. One set of methods relies heavily on the time-invariance properties of the system while the other set requires the operator to be close to the identity in some fashion. In this paper we will narrow this gap somewhat by extending the projection integral formalism of [15] to include operators that are dominated by Legesgue measure on the semi-infinite interval. This will enable a unified treatment of the factorization problem for a wide class of both nonstationary and time-invariant systems.

The projection integral approaches in [6, 7, 15], all lead to factorizations in which the factors are quasinilpotent. Thus it is seen that these ideas do not directly extend to the Wiener–Hopf problems, since the factors there cannot be quasinilpotent. However, it will be shown that the "Lebesgue" theory in [15] extended to the semi-infinite interval, together with the following condition on the operator K in (1.1),

$$\sup_{\substack{x \in H \\ |x|=1}} \int_0^\infty |(I-P(t))\, KP(t)\, x|^2\, dt < \infty. \tag{1.2}$$

leads to a viable factorization theory that includes the Wiener–Hopf factorization. Importantly, (1.2) is also precisely the condition that enables an extension of the control approach in [14] to infinite time problems.

It is interesting to note that in the classical Wiener–Hopf factorization [5], the compactness of $(I-P(t))\, KP(t))$ proved to be crucial. Condition (1.2) represents a trade of compactness for a norm condition on the operator K with respect to the chain C. Also, with regards to the control applications, when K is causal the family of operators $(I-P(t))\, KP(t)$ can be interpreted as determining the free response of the system after time t to inputs that terminate at time t. In this context (1.2) is a sort of stability condition.

Although one motivation for extending the projection integral methods is to present this unified treatment, our primary motivation is rooted in applications to optimal control problems on the semi-infinite interval, and to develop a theory for these problems that parallels the development in [14]. These applications will be developed more fully in a future paper; however, a control application is presented in Section 4. We now briefly review the organization of the paper.

The second section introduces the class of operators that serves as the focal point of the paper. An example (Example 2.5) is given to demonstrate the nontriviality of the class. Section 3 begins with a couple of general remarks on causal invertibility (similar to Feintuch [3]), and then the main factorization theorem with some extensions are proved. The fourth section is devoted to applications. The first application recovers the Wiener–Hopf factorization in the matrix case. A couple of straightforward extensions involving operator valued kernels are also presented. A second application is to the particular factorization problem that arises in the infinite time regulator problem for infinite-dimensional systems. The feedback solution to the regulator problem is then derived using the methods of the paper.

2. Background and Preliminary Results

In this section we introduce the class of operators that will be the focal point of the paper. We will also discuss some of its elementary properties. Before doing so, it is necessary to set some notations and groundwork for further discussion.

Let Σ denote the class of Borel subsets of $[0, \infty)$ and let λ denote Lebesgue measure. For each $s \leqslant \infty$, the measure λ_s is defined by $\lambda_s(\omega) = \lambda(\omega \cap (0, s))$, $\omega \in \Sigma$. Now let H be a separable Hilbert space and let $E: \Sigma \to B(H)$ denote a resolution of the identity. We will use the notations $P^t = E([0, t])$ and $P_t = I - P^t$. Furthermore, we assume that the projections P^t are strongly continuous, i.e., $P^{t_n}x \to P^t x$ whenever $t_n \to t$, and we adopt the conventions that $P^\infty = I$ and $P^0 = 0$. Given two Hilbert spaces H_1 and H_2 with resolutions of the identity E_1 and E_2, respectively, a map $T \in B(H_1, H_2)$ is said to be causal if $P_2^t T P_1^t = P_2^t T$ and memoryless if $E_2(\omega) T = T E_1(\omega)$ for all $\omega \in \Sigma$. The subspace of memoryless maps in $B(H_1, H_2)$ will be denoted $M(H_1, H_2)$ in the sequel.

Let H and E be defined as in the paragraph above. A map $T \in B(H)$ is said to be dominated by λ_s (written $T < \lambda_s$) if there exists a constant α such that $|E(\omega) T| < \alpha \sqrt{\lambda_s(\omega)}$ for all $\omega \in \Sigma$. We also define $L^{\lambda_s} \subset B(H)$ by $L^{\lambda_s} = \{T: T < \lambda_s\}$, $s \leqslant \infty$. With each $s \leqslant \infty$ we associate with H the Hilbert space $H_s = L_2((0, s), H)$ and resolution of the identity $\tilde{E}$, $[\tilde{E}(\omega)x](t) = \chi(\omega)(t)\, x(t)$. Then if $T < \lambda$ we can define the mapping $F(T) \in M(H_s, H)$ by its action on simple functions in H_s (cf. [15]),

$$F(T)x = \sum_{i=1}^{n} E(\omega_i)\, Tx_i; \qquad x(t) = \sum_{i=1}^{n} \chi(\omega_i)(t) x_i. \tag{2.1}$$

In addition to the mapping above we will have occasion to deal with several other mappings on an between the spaces H and H_s. For simplicity we will gather some of these here: For $s < \infty$ define $G^- \in B(H, H_s)$ by

$$\begin{aligned} [G^+ x](t) &= P^t x \\ [G^- x](t) &= P_t x. \end{aligned} \tag{2.2}$$

For $K \in \in B(H)$ define the mappings

$$\bar{K} \in M(H_s); \qquad [\bar{K}x](t) = Kx(t), \qquad 0 \leqslant s \leqslant \infty \tag{2.3}$$

$$K^s \in B(H); \qquad K^s = P^s K P^s, \qquad 0 \leqslant s \leqslant \infty \tag{2.4}$$

$$h(K) \in B(H, C((0, \infty), H)); \qquad h(K)x : t \to P_t K P^t x \tag{2.5}$$

$$a(K) \in M(H_s); \qquad a(K)x : t \to (I + P_t K P_t)^{-1} x(t) \qquad \text{(when defined)}. \tag{2.6}$$

The map defined in (2.5) provides the link between the factorization and control theory developed for operators dominated by a finite measure in [14, 15] and the extension which is developed in the present paper. The following two results from [15] will be called upon repeatedly in the sequel.

THEOREM 2.1. *Suppose K, $K^* < \lambda_s$ for $s < \infty$ and $a(k) \in M(H_s)$. Then $I + K$ has the unique (right) factorization*

$$I + K = (I + X_-)(I + X_+)$$

with $X_\pm$ respectively causal and anticausal and $X_\pm$, $X_\pm^* < \lambda_s$. Furthermore $W_- = (I + X_-)^{-1} - I$ has the representation

$$W_- = -F(K)\, a(K) G^-.$$

THEOREM 2.2. *If $s < \infty$ and $X < \lambda_s$ with X causal (anticausal) then X is quasinilpotent.*

If $K < \lambda$ then clearly $K^s < \lambda_s$ for each $s < \infty$. It follows from results in [15] that $F(K^s) G^\pm$ are projections on L^{λ_s} such that

$$\begin{aligned} &\text{(i)} \quad F(K^s) G^\pm \text{ is causal (anticausal)} \\ &\text{(ii)} \quad K^s = F(K^s) G^+ + F(K^s) G^- \\ &\text{(iii)} \quad F(K^s) G^\pm < \lambda_s. \end{aligned} \tag{2.7}$$

Now define the subset $S \subset B(H)$,

$$S = \{K: K < \lambda \text{ and } \lim_{s \to \infty} F(K^s) G^+ x \text{ exists for all } x \in H\}, \tag{2.8}$$

and the mapping $p^+ : S \to B(H)$ by

$$p^+(K)x = \lim_{s \to \infty} F(K^s)\, G^+ x.$$

We note that when $K \in S$, $p^+(K) \in B(H)$ by virtue of the Banach–Steinhaus theorem, and that $p^+(K)$ is causal since it is the strong limit of causal maps. For each $K \in S$ we also define $p^-(K) = I - p^+(K)$.

With this bit of background we define the class R as

$$R = \{K \in S: K^* \in S, \text{ and } h(K), h(K^*) \in B(H, H_\infty)\}. \tag{2.9}$$

The following proposition collects some of the elementary properties of R.

PROPOSITION 2.3. *The following hold:*

(i) *R is a vector space*

(ii) *If $K \in R$, then $p^-(K)$ is anticausal*

(iii) *$p^{\pm}$ are projections on R and $R = p^+(R) \oplus p^-(R)$*

Proof. (i) The proof of this is straightforward and is omitted.

(ii) Using (2.7) we have for each $s < \infty$, $p^-(K^s) = K^s - p^+(K^s) = F(K^s)G^-$. Thus $p^-(K^s)$ is anticausal and consequently so is $p^-(K)$, since it is the strong limit of $p^-(K^s)$ as $s \to \infty$.

(iii) Noting that we can write each $K \in R$ as $K = p^+(K) + p^-(K)$, using (i) it suffices to show that p^+ is a projection on R. Since $p^-(K)$ is anticausal,

$$P_t\, p^+(K) P^t = P_t K P^t \qquad \text{for all } t.$$

Hence, $h(p^+(K))$, $h([p^+(K)]^*) \in B(H, H_\infty)$. It remains then to show that $p^+(K)$, $[p^+(K)]^* < \lambda$. So assume $|E(\omega)K| < \alpha\sqrt{\lambda}(\omega)$ for all $\omega \in \Sigma$. Now fix $\omega \in \Sigma$ and choose $\varepsilon > 0$. Then we can find an $x \in H$ with $|x| = 1$ such that $|E(\omega)\, p^+(K)| \leqslant |E(\omega)\, p^+(K)x| + \varepsilon/3$. Also there exists t such that $s \geqslant t$ implies

$$|E(\omega)[p^+(K)x - p^+(K^s)x]| < \varepsilon/3.$$

And for a suitably fine partition of $[0, s]$, say $\{t_i\}_{i=0}^n$,

$$\left| E(\omega)\, p^+(K^s)x - E(\omega) \sum_{i=0}^{n-1} E(\omega_i)\, K^s P^{t_i} x \right| < \varepsilon/3 \qquad (\omega = [t_i, t_{i+1}]).$$

But,

$$\begin{aligned} \Sigma\, |E(\omega_i)\, E(\omega)\, p^+(K^s) P^{t_i} x|^2 &= \Sigma\, |E(\omega \cap \omega_i)\, K^s P^{t_i} x|^2 \\ &\leqslant \Sigma\, |E(\omega \cap \omega_i) K|^2\, |P^{t_i} x|^2 \\ &\leqslant \alpha^2 \lambda(\omega). \end{aligned}$$

Thus, $|E(\omega)\, p^+(K)| \leqslant \varepsilon + \alpha\sqrt{\lambda}(\omega)$. We can verify that $[p^+(K)]^* < \lambda$ by using essentially the same argument above together with the identity $[p^+(K)]^* = p^-(K^*)$. ∎

The proposition above states that R is a vector space. The following result shows that right (left) multiplication in R is defined for causal (anticausal) elements.

PROPOSITION 2.4. *Let $X, K \in R$ with X causal. Then $KX \in R$. Also if X is anticausal, then $XK \in R$.*

Proof. Assume that X is causal. It is trivial that KX, $X^*K^* < \lambda$. Next we show $h(KX)$, $h(X^*K^*) \in B(H, H_\infty)$. To see this note for each t the identities,

$$P_t KXP^t = P_t KP^t X + P_t KP_t XP^t,$$
$$P_t X^*K^*P^t = P_t X^*P_t K^*P^t.$$

From these it follows readily that $|h(KX))| < [|h(K)|^2 + 2|K|\,|h(K)|\,|h(X)| + |K|^2 |h(X)|^2]^{1/2}$ and $h(X^*K^*) \leqslant |X|\,|h(K^*)|$. It remains to verify that KX, $X^*K^* \in S$ (cf. (2.8)). This can be proved using a result in [13, Lemma 4.1]. The extension of this result we use reads as: If $A, B \in R$ with A anticausal, then $p^+(AB) = F(A)\,h(B)$. Thus we obtain

$$p^+(KX) = p^+(K)X + F(p^-(K))\,h(X),$$

and

$$p^+(X^*K^*) = F(X^*)\,h(K^*).$$

Thus the result is proved for X causal. We note that the anticausal case also been argued (i.e., X^* is anticausal. ∎

The following example describes a class of operators satisfying the condition (2.9). This example will be referred to later in Section 4.

EXAMPLE 2.5. Let H_0 denote a separable Hilbert space and let $H = L_2((0, \infty), H_0)$ with the resolution of the identity E, $[E(\omega)x](t) = \chi(\omega)(t)\,x(t)$. Define $K \in B(H)$ by

$$Kx\colon t \to \int_0^\infty K(t, s)\,x(s)\,ds,$$

where $K(t, s) \in B(H_0)$ for each t, s, $K(t, s)v$ is jointly measurable for each $v \in H_0$, and $|K(t, s)| \leqslant k(t-s)$, where $k \in L_1(-\infty, \infty) \cap L_2(-\infty, \infty)$. Further assume that $|t|^{1/2} \cdot |k(t)| \in L_1(-\infty, \infty)$. We claim that under these assumptions $K \in R$. First we verify that K, $K^* \in S$. Let $x \in H$ with $|x| = 1$. Then for a Borel set ω with finite measure,

$$|E(\omega)\,Kx|^2 \leqslant \int_\omega^\infty \left| \int_0 K(t, s)\,x(s)\,dx \right|^2 dt$$
$$\leqslant |k|_2^2\,\lambda(\omega),$$

where $|k|_2$ denotes the L_2-norm. Since $|K^*(t, s)| \leqslant |k(t-s)|$, the same bound holds for $|E(\omega)K^*|$. Thus K, $K^* < \lambda$. Now it is straightforward to verify that

$$p^+(K^s)x: t \to \begin{cases} \int_0^t K(t, \sigma)\, x(\sigma)\, d\sigma & t \leqslant s \\ 0 & t \leqslant s, \end{cases}$$

and that $p^+(K^s)$ converges strongly to the operator $p^+(K)$,

$$p^+(K)x: t \to \int_0^t K(t, \sigma)\, x(\sigma)\, d\sigma.$$

The analogous result is readily obtained for K^*. Therefore we have shown that $K \in S$. It remains to demonstrate that $h(K)$, $h(K^*) \in B(H, H_\infty)$. To this end note that

$$\begin{aligned} |[h(K)x](t)|^2 &\leqslant \int_t^\infty \left| \int_0^t k(\tau - s)\, x(s)\, ds \right|^2 d\tau \\ &\leqslant \int_t^\infty \left\{ \int_0^t k(\tau - s)\, ds \right\} \left\{ \int_0^t k(\tau - s)\, |x(s)|^2\, ds \right\} d\tau. \end{aligned} \tag{2.10}$$

And after defining

$$\gamma(t, \tau) = \int_0^t k(\tau - s)\, ds,$$

an interchange in the order of integration yields

$$|[h(K)x](t)|^2 \leqslant \int_0^t z(t, s)\, |x(s)|^2\, ds, \tag{2.11}$$

where

$$z(t, s) = \int_t^\infty \gamma(\tau - s)\, k(\tau - s)\, d\tau. \tag{2.12}$$

Fubini's theorem now implies that $h(K) \in B(H, H_\infty)$ if the function $\phi(s)$,

$$\phi(s) = \int_s^\infty z(t, s)\, dt$$

is uniformily bounded on $[0, \infty)$, since $|h(K)x|^2 \leqslant \int_0^\infty |x(s)|^2\, \phi(s)\, ds$. Now we write

$$\phi(s) = \int_s^\infty \int_s^\tau \gamma(t, \tau)\, dt\, k(\tau - s)\, d\tau, \tag{2.13}$$

and define the function

$$K(\xi) = \int_{\xi}^{\infty} k(\sigma)\, d\sigma.$$

After a little manipulation we have

$$\int_{s}^{\tau} \gamma(t, \tau)\, dt = -(\tau - s)\, K(\tau) + \int_{0}^{\tau - s} K(\xi)\, d\xi.$$

Hence, combining this with (2.13) we obtain,

$$0 \leqslant \phi(s) \leqslant \int_{s}^{\infty} (s - \tau)\, K(\tau)\, k(\tau - s)\, d\tau + \int_{s}^{\infty} \left\{ \int_{0}^{\tau - s} K(\xi)\, d\xi \right\} k(\tau - s)\, d\tau.$$

Since the first integral above is negative,

$$\phi(s) \leqslant \int_{0}^{\infty} \left\{ \int_{0}^{u} K(\xi)\, d\xi \right\} k(u)\, du.$$

But

$$\int_{0}^{u} K(\xi)\, d\xi = uK(u) + \int_{0}^{u} \xi k(\xi)\, d\xi.$$

Thus,

$$\begin{aligned} \phi(s) &\leqslant \int_{0}^{\infty} \sigma k(\sigma)\, K(\sigma)\, d\sigma + \int_{0}^{\infty} \int_{0}^{u} \xi k(\xi)\, d\xi\, k(u)\, du \\ &= 2 \int_{0}^{\infty} \sigma k(\sigma)\, K(\sigma)\, d\sigma. \end{aligned}$$

Defining $g(t) = (1 + t^{1/2})\, k(t)$, $t \geqslant 0$, it follows that

$$|h(K)| \leqslant \sqrt{2} \int_{0}^{\infty} g(t)\, dt. \tag{2.14}$$

A similar argument establishes $h(K^*) \in B(H, H_{\infty})$. Thus $K \in R$.

The applications of Section 4 will rely on this example as well as the estimate (2.14).

3. Main Results

In this section we prove our main factorization results concerning operators of the form $I + K$, $K \in R$. Before doing so, it is necessary to establish some preliminary propositions.

The first few statements concern some general properties of causal invertibility.

PROPOSITION 3.1. *Suppose* $X \in B(H)$ *is causal (anticausal) and* $(I + P^t X P^t)$ *is invertible for each* $t \in [0, \infty)$. *Then if* $I + X$ *is invertible,*

(i) $W = (I + X)^{-1} - I$ *is causal (anticausal),*

(ii) $(I + P^t X P^t)^{-1} - I = P^t W P^t$ *for each* t,

(iii) $(I + P_t X P_t)^{-1} - I = P_t W P_t$ *for each* t.

Proof. We shall prove the proposition for the causal case. The anticausal case follows by taking adjoints.

(i) The proof of this is essentially contained in [15].

(ii) Noting that W satisfies the identities

$$W + X + XW = 0 \tag{3.1}$$

and

$$W + X + WX = 0, \tag{3.2}$$

for each t, we then have

$$P^t W P^t + P^t X P^t + P^t X W P^t = 0,$$

$$P^t W P^t + P^t X P^t + P^t W X P^t = 0.$$

And since X and W are both causal,

$$P^t W P^t + P^t W P^t + P^t X P^t W P^t = 0,$$

$$P^t W P^t + P^t X P^t + P^t W P^t X P^t = 0.$$

Adding the identity to the above,

$$(I + P^t X P^t)(I + P^t W P^t) = (I + P^t W P^t)(I + P^t X P^t) = I.$$

Thus, (ii) holds.

(iii) This is proved in an analogous manner using $P_t W P_t = W P_t$ and $P_t X P_t = X P_t$ for causal X and W. ∎

Remarks. 1. The proof of conclusions (ii) and (iii) of the proposition rely only on (i). Thus (ii) and (iii) are valid under the hypotheses that $I + X$ is causal (anticausal) and causally (anticausally) invertible. Also note that these hypotheses imply $\sup |(I + P^t X P^t)^{-1}| < \infty$. The converse of this statement is also true, i.e., if $\sup|(I+P^tXP)^t|^{-1}<\infty$, then $I+X$ is causally invertible ([3]).

2. If it is assumed that $X\in R$, the following argument shows that $W\in R$: First we note that W, $W^*<\lambda$ by using the resolvent identities (3.1)–(3.2) and the ideal properties of L^λ. Now (3.1) implies

$$\begin{aligned} P_t WP^t &= -P_tXP^t - P_tXWP^t \\ &= -P_tXP^t - P_tXP^tWP^t - P_tXP_tWP^t. \end{aligned}$$

Thus,

$$(I+P_tXP_t)\,P_tWP^t = -P_tXP^t(I+W),$$

so that using (iii) of the proposition,

$$P_tWP^t = -(I+P_tWP_t)\,P_tXP^t(I+W).$$

And consequently,

$$\begin{aligned} \sup_{|v|=1} \int |P_tWP^tv|^2\,dt &\leqslant \{1+|W|\}^2 \sup_{|v|=1} \int |P_tXP^tv|^2\,dt \\ &< \infty. \end{aligned}$$

The proof of the main result is based on certain limiting arguments using Theorem 2.1. The following two lemmas are the main tools of the arguments. The first result concerns a convergence property of the operator $F(\cdot)$. This result is essentially an "infinite-time" version of a result in [13].

LEMMA 3.2. *Suppose K, $K_n\in R$, $n=1,2,\ldots$. If $K_n\to K$ strongly and there exists a constant α such that $|E(\omega)K_n|<\alpha\sqrt{\lambda}(\omega)$ for all n, ω, then $F(K_n)\to F(K)$ strongly.*

Proof. The proof is essentially the same as the one in [13], and is omitted.

The next lemma concerns analogous convergence properties of the operators $a(\cdot)$ and $h(\cdot)$ (cf. (2.5)–(2.6)).

LEMMA 3.3. *Suppose $K\in R$ and $\sup_{t,s}|(I+P_tK^sP_t)^{-1}|<\infty$. Then*

(i) *$a(K^s)\to a(K)$ strongly,*

(ii) *$h(K^s)\to h(K)$ strongly.*

Proof. (i) Fix $x\in H$. Then

$$[a(K-K^s)x](t)=[(I+P_tKP_t)^{-1}x(t)-(I+P_tK^sP_t)^{-1}x(t)],$$

so that

$$|a(K-K^s)x|^2 \leqslant \int |(I+P_tK^sP_t)^{-1}P_t(K^s-K)\,P_t(I+P_tKP_t)^{-1}\,x(t)|^2\,dt$$

$$\leqslant \alpha^2 \int |(K^s-K)\,P_t(I+P_tKP_t)^{-1}\,x(t)|^2\,dt,$$

where $\alpha = \sup_{t,s}|(I+P_tK^sP_t)^{-1}|$. Note that the integrand above is bounded by $[2\alpha|K|\cdot|x(t)|]^2$, which is integrable. Also for a.e. t, $|(K^s-K)\,P_t(I+P_tKP_t)^{-1}\,x(t)| \to 0$ as $s\to\infty$, since $K^s\to K$ strongly. The result follows from the dominated convergence theorem.

(ii) Since,

$$[h(K-K^s)x](t) = \begin{cases} P_sKP^tx & t\leqslant s \\ P_tKP^tx & t>s, \end{cases}$$

it follows that

$$|h(K-K^s)x|^2 = \int_0^s |P_sKP^tx|^2\,dt + \int_s^\infty |P_tKP^tx|^2\,dt.$$

For $t\leqslant s$, $|P_sKP^tx| \leqslant |P_tKP^tx|$ so the first integrand is bounded by $|P_tKP^t|^2$ for all s. And since for each t, $|P_sKP^tx| \to 0$ as $s\to\infty$, by dominated convergence the first integral tends to zero. The second integral trivially tends to zero as $s\to\infty$. The lemma is proved. ∎

With these preliminary results established we can now prove our major result on factorization in R.

THEOREM 3.4. *Assume $K\in R$ and $I+K>0$ with $I+K$ invertible. Then there exists a unique causal $X\in R$ with $I+X$ causally invertible such that*

$$(I+K) = (I+X^*)(I+X). \tag{3.3}$$

Proof. (a) We first prove the existence of a causal X such that (3.3) holds. Since $I+K$ is invertible and $I+K>0$ it folows that for some $\varepsilon>0$, $\inf\sigma(K) = -1+\varepsilon$. Thus for any $s>0$,

$$-1+\varepsilon = \inf_{\substack{|x|=1\\ x\in H}} \langle Kx, x\rangle \leqslant \inf_{\substack{|x|=1\\ x\in P^sH}} \langle Kx, x\rangle = \inf\sigma(K^s).$$

Consequently, $I+K^s>0$ for all s. Furthermore, since $K<\lambda$ it is evident that $K^s<\lambda_s$. Theorem 2.1 now implies the existence of a unique causal $X_s\in L^{\lambda_s}\cap L^{*\lambda_s}$ such that

$$I+K^s = (I+X_s^*)(I+X_s). \tag{3.4}$$

We claim that $X_s = K^s_+ - F(K^s)\, a(K^s)\, h(K^s)$. To see this we first use (3.4) to obtain

$$X_s = K^s_+ + F(W^*_s)\, \bar{K}^s G^+, \tag{3.5}$$

where $W_s = (I + X_s)^{-1} - I$. Next we compute $F(W^*_s)x$ for $x \in H_\infty$ simple; say $x(t) = \Sigma \chi(\omega_i)(t) x_i$. Using Theorem 2.1,

$$\begin{aligned} F(W^*_s)x &= \Sigma E(\omega_i)\, W^*_s x_i \\ &= -\Sigma E(\omega_i)[F(K^s)\, a(K^s) G^-] x_i \\ &= -F(K^s)\{\Sigma \tilde{E}(\omega_i)\, a(K^s)\, G^- x_i\}. \end{aligned} \tag{3.6}$$

Define the mapping $\chi^s \in M(H_\infty)$,

$$[\chi^s x](t) = [I + P_t K^s P_t]^{-1}\, P_t x(t).$$

It is evident from (3.6) that $F(W^*_s)$ and $-F(K^s)\chi^s$ agree on the simple functions in H_∞. Hence, by continuity $F(W^*_s) = -F(K^s)\chi^s$. Therefore we can write (3.5) as

$$X_s = K^s_+ - F(K^s)\, \chi^s \bar{K}^s G^+.$$

But,

$$\chi^s \bar{K}^s G^+ x : t \to (I + P_t K^s P_t)^{-1}\, P_t K^s P^t x.$$

And indeed,

$$X_s = K^s_+ - F(K^s)\, a(K^s)\, h(K^s). \tag{3.7}$$

Since $K \in R$, Lemmas 3.2 and 3.3 allow us to take a strong limit (as $s \to \infty$) above. Then taking weak limits in (3.4) we obtain

$$I + K = (I + X^*)(I + X); \quad X = K_+ - F(K)\, a(K)\, h(K). \tag{3.8}$$

Also note that X is causal since it is the strong limit of the causal operators X_s.

(b) Next we establish the invertibility of $I + X$. Using (3.4) we write

$$I + W^*_s = (I + X_s)(I + K^s)^{-1}. \tag{3.9}$$

Note that $|X_s|$ and $|(I + K^s)^{-1}|$ are bounded independently of s. Thus the right side above converges strongly as $s \to \infty$, and consequently W^*_s converges strongly to an anticausal operator, say W^*. Thus we obtain

$$(I + W^*)(I + K) = (I + X). \tag{3.10}$$

Then multiplying the above by $I+X^*$ and using the factorization (3.8) we have the identities

$$(I+X^*)(I+W^*)=I \tag{3.11}$$

and

$$(I+W)(I+X)=I. \tag{3.12}$$

Now let ω be a Borel subset with finite Lebesgue measure. From the definition of X (cf. (3.8)) it follows

$$\sup_{|v|=1} |E(\omega)\,Xv| \leqslant \sup_{|v|=1} |E(\omega)\,K_+ v| + |F(K)\,a(K)| \sup_{|v|=1} |E(\omega)\,h(K)v|.$$

But,

$$\begin{aligned} \sup_{|v|=1} |E(\omega)\,h(K)v|^2 &\leqslant \int_\omega |P_t K P^t|^2\,dt \\ &< \lambda(\omega)\,|K|^2. \end{aligned}$$

Thus $X<\lambda$. Defining $R=(I+K)^{-1}-I$, we obtain from (3.10)

$$W^*=X+R+XR.$$

Now since $R<\lambda$, it follows that $W^*<\lambda$. Now suppose there exists $v\in H$ such that $(I+W)v=0$. Then for all $s\in(0,\infty)$, $P^s(I+W)v=0$. Hence, $(I+P^sWP^s)\,P^s v=0$. But by Theorem 2.2, P^sWP^s is quasinilpotent. Thus $v=0$ and $(I+W)$ is 1–1. From (3.12) it then follows that $I+X$ is invertible.

(c) Now we show $X\in R$. We have already established that $X<\lambda$ and $W^*<\lambda$. But then it follows $X^*<\lambda$. Finally, from (3.8) we have

$$\begin{aligned} P_t K P^t &= P_t X P_t + P_t X^* X P^t \\ &= P_t X P^t + P_t X^* P_t X P^t \\ &= (I+P_t X^* P_t)\,P_t X P^t. \end{aligned}$$

Then,

$$P_t X P^t = (I+P_t X^* P_t)^{-1}\,P_t K P^t$$

and

$$\sup_{|v|=1} \int |P_t X P^t v|^2 \leqslant c \sup_{|v|=1} \int |P_t K P^t v|^2,$$

where $c = \sup|(I + P_t X^* P_t)^{-1}|^2 < \infty$ by the first remark following Proposition 3.1.

(d) To prove uniqueness first recall that $R = (I + K)^{-1} - I < \lambda$. Now let X_i, $i = 1, 2$, be two solutions of the factorization problem. Then by what we have already proved

$$I + R = (I + W_i)(I + W_i^*), \qquad i = 1, 2,$$

where $W_i = (I + X_i)^{-1} - I \in R$. Thus for each s,

$$I + P^s R P^s = (I + P^s W_i P^s)(I + P^s W_i^* P^s).$$

Since $P^s R P^s < \lambda_s$, it follows from the uniqueness of the (left) factorization in $L^{\lambda_s} \cap L^{*\lambda_s}$ (see [15]) that $P^s W_1 P^s = P^s W_2 P^s$ for all s. Hence, $X_1 = X_2$. ∎

The result above requires the operator to be self-adjoint. Although we have not topologized the space R, in the nonself-adjoint case we may expect some type of "small-norm" result to hold. We have the following.

THEOREM 3.5. *Let $K \in R$. Then there exists $\delta > 0$ such that for any $\mu \in C$ with $|\mu| < \delta$, $I + \mu K$ has the unique factorization in R,*

$$I + \mu K = (I + X_-(\mu))(I + X_+(\mu))$$

with $(I + X_\pm(\mu))$ causally (anticasually) invertible. Furthermore, the mappings $\mu \to X_\pm(\mu)$ are analytic with respect to the $B(H)$ topology.

Proof. First note that if $|\mu| < 1/|K|$, there exists a constant c such that $\sup|(I + \mu P_t K^s P_t)^{-1}| < c$. Theorem 2.1 then yields the factorization

$$I + \mu K^s = (I + X_-(\mu, s))(I + X_+(\mu, s)). \tag{3.13}$$

From the same theorem, $W_-(\mu, s) = (I + X_-(\mu, s))^{-1} - I$ has the representation

$$W_-(\mu, s) = -F(\mu K^s)\, a(\mu K^s) G^-.$$

Arguing as before, $X_+(\mu, s) \to X_+(\mu)$ strongly where

$$X_+(\mu) = \mu\{K_+ - F(K)\, a(\mu K)\, h(\mu K)\} \tag{3.14}$$

and

$$X_+(\mu, s) = \mu\{K_+^s - F(K^s)\, a(\mu K^s)\, h(\mu K^s)\}. \tag{3.15}$$

Next note that

$$|[h(K^s)x](t)| = |P_t K^s P^t x| \leqslant \begin{cases} |P_t K P^t x| & t < s, \\ 0 & t \geqslant s. \end{cases}$$

Thus $\sup|h(K^s)| < \infty$. Using this and (3.14)–(3.15), it follows that for $|\mu|$ sufficiently small we obtain $|X_+(\mu)| < 1$ and $\sup|X_+(\mu, s)| < 1$. Hence,

$$(I + X_+(\mu, s))^{-1} \to (I + X_+(\mu))^{-1}$$

strongly as $s \to \infty$. Consequently by (3.13), $I + X_-(\mu, s)$ converges strongly to the (necessarily anticausal) operator $(I + \mu K)(I + X_+(\mu))^{-1}$. The one-sided ideal properties of the L^λ-spaces (see [15]), yield $X_\pm(\mu) < \mu$. Now for any bounded Borel subset $\omega \subset [0, \infty)$,

$$KE(\omega) = X_-(\mu)\, E(\omega) + X_-(\mu)\, X_+(\mu)\, E(\omega) + X_+(\mu)\, E(\omega). \tag{3.16}$$

Thus,

$$X_+(\mu)\, E(\omega) = (I + X_-(\mu))^- [K - X_-(\mu)]\, E(\omega).$$

Let $W_-(\mu) = (I + X_-(\mu))^{-1} - I$. Then applying p^+ to the above and using Proposition 2.3 we obtain

$$|X_+(\mu)\, E(\omega)| \leqslant |K_+ E(\omega)| + |[W_-(\mu) K]_+ E(\omega)|.$$

Therefore $X_+^*(\mu) < \lambda$. And now it routinely follows from (3.16) that $X_-^*(\mu) < \lambda$. The remainder of the argument to show that $X_\pm(\mu) \in R$ and are unique follows in the same manner as in the proof of Theorem 3.4. Now since $\mu \to a(\mu K)$ is analytic, noting from (3.14) that

$$X_+(\mu) = \mu K_+ - \mu^2 F(K)\, a(\mu K)\, h(K),$$

we also have that $\mu \to X_+(\mu)$ is analytic. Finally, $\mu \to X_-(\mu)$ is analytic since $X_-(\mu) = (I + \mu K)(I + X_+(\mu))^{-1}$. ∎

This small norm result, together with Proposition 2.4, yields the following corollary.

COROLLARY 3.6. *Suppose $A \in R$ and $I + A$ has the factorization*

$$I + A = (I + X_-)(I + X_+)$$

with $X_{\pm} \in R$ respectively causal and anticausal. Then given any $K \in R$, for $\mu \in C$ with sufficiently small modulus, $I + A + \mu K$ has the factorization

$$I + A + \mu K = (I + X_{-}(\mu))(I + X_{+}(\mu))$$

with $(I + X_{\pm}(\mu))$ causally (anticausally) invertible. Furthermore, the mappings $\mu \to X_{\pm}(\mu)$ are analytic.

Proof. Write

$$I + A + \mu K = (I + K_{-})\{I + \mu(I + W_{-})\, K(I + W_{+})\}(I + X_{+}),$$

where $W_{\pm} = (I + X_{\pm})^{-1} - I$. Proposition 2.4 implies $(I + W_{-})\, K(I + W_{+}) \in R$. The theorem then implies that we can factor $I + \mu(I + W_{-})\, K(I + W_{+})$ for sufficiently small $|\mu|$. And the result follows upon multiplication and Proposition 2.4. ∎

This result will be useful in the next section when we apply the theory to the classical Wiener-Hopf factorization problem.

4. Applications

In this section we will apply the preceding theory to two well-known examples in which the operator K is an integral operator with difference kernel (possibly operator-valued). In the first example, we derive the classical Wiener–Hopf factorization using the "projection-integral" methods of the preceding section. And in the second example, we derive the factorization for what is essentially the infinite-time version of a control problem considered in [15].

In the following, $C^{n \times n}$ will denote the $n \times n$ complex valued matrices and $\mathscr{K}$ will denote the class of functions

$$\mathscr{K} = \{k \in L_2((-\infty, \infty), C^{n \times n}) \colon (1 + |t|^{1/2})\, k(t) \in L_1((-\infty, \infty), C^{n \times n})\}.$$

The norm on $\mathscr{K}$ is defined

$$|k| = \int_{-\infty}^{\infty} \{|k(t)|^2 + (1 + |t|^{1/2})\, k(t)\}\, dt.$$

The Wiener–Hopf factorization problem can be placed into the framework we have developed by identifying:

(i) the underlying Hilbert space H with $L_2((0, \infty), C^n)$,

(ii) the resolution of the identity E with multiplication by the characteristic function, i.e., $[E(\omega)x](t) = \chi(\omega)\, x(t)$, and

(iii) the operator K with the integral operator on H,

$$Kx: \to \int_0^{\infty} k(t-s)\ x(s)\, ds,\ k(\cdot) \in L_1((-\infty, \infty)\ C^{n\times n}). \tag{4.1}$$

The classical result [5] states that if $I+K>0$ then there exists a unique function $x(\cdot) \in L_1((\infty, \infty), C^{n\times n})$ vanishing on $(-\infty, 0)$ such that

$$I+K=(I+X^*)(I+X), \tag{4.2}$$

where X is the integral operator with kernel $x(t-s)$. Now if the matrix function k of (4.1) is an element of $\mathscr{K}$, then Example 2.5 and Theorem 3.4 immediately imply the existence of a factorization of the type (4.2). However, the particular form of the factor (i.e., integral operator with difference kernel) is clearly not evident. Since the factor X has an explicit representation (cf. 3.8)), the first order of business is to give the expressions comprising this representation a more concrete meaning in the present context. This is the concern of the next two lemmas.

LEMMA 4.1. *Assume $k \in \mathscr{K}$ and let K represent the associated integral operator on H (cf. (4.1)). Then given $x \in H_\infty = L_2((0, \infty), H)$,*

$$F(K)x: t \to \int_0^{\infty} k(t-\sigma)\ x(t)(\sigma)\, d\sigma \qquad \text{a.e.}$$

Proof. This proof follows along identical lines as one given in [13] for a similar proposition. ∎

We note that each $g(\cdot) \in \mathscr{K}$ induces a mapping $\tilde{g} \in B(H_\infty)$ by

$$[\tilde{g}x](t) = Gx(t), \tag{4.3}$$

where $G \in B(H)$ is the integral operator

$$Gv: t \to \int_0^{\infty} g(t-s)\ v(s)\, ds.$$

LEMMA 4.2. *Let k and K be as in the lemma above and let $g(\cdot) \in \mathscr{K}$. Define*

$$v(t) = \int_t^{\infty} k(t-s)\ g(s)\, ds, \qquad -\infty < t < \infty, \tag{4.4}$$

and let V denote the integral operator with kernel $v(t-s)$. Then $v(\cdot) \in \mathscr{K}$ and

$F(V)=F(K)P^{-}\tilde{g}$, where $\tilde{g}$ is the mapping induced by $g(\cdot)$ as in (4.3) *and $P^{-}\in B(H_\infty)$ is defined as $[P^{-}x](t)=P_t x(t)$.*

Proof. Let $x\in H_\infty$. From Lemma 4.1 we have the representation

$$[F(K)P-\tilde{g}x](t)=\int_0^\infty k(t-s)[P^{-}\tilde{g}x](t)(s)\,ds. \tag{4.5}$$

From the definitions of $\tilde{g}$ and P^{-},

$$\begin{aligned}[P^{-}\tilde{g}x](t)\colon s\to &\begin{cases}[\tilde{g}x](t)(s), & s\geqslant t,\\ 0, & s<t,\end{cases}\\ =&\begin{cases}\int_0^\infty g(s-u)\,x(t)(u)\,du, & s\geqslant t,\\ 0, & s<t.\end{cases}\end{aligned}$$

This, together with Fubini's theorem in (4.4) gives

$$[F(K)\,P^{-}\tilde{g}x](t)=\int_0^\infty\left\{\int_t^\infty k(t-s)\,g(s-u)\,ds\right\}x(t)(u)\,du. \tag{4.6}$$

Now since

$$v(t-u)=\int_t^\infty k(t-s)\,g(s-u)\,ds,$$

Lemma 4.1 implies that the right side of (4.6) is $F(V)$. It is straightforward to verify that $v(\cdot)\in\mathscr{K}$. ∎

THEOREM 4.3 (Wiener–Hopf). *Let the operator K be defined as in* (4.1) *and assume $I+K>0$. Then there exists a unique $x(\cdot)\in L_1((0,\infty),\,C^{n\times n})$ such that*

$$I+K=(I+X^*)(I+X),$$

where the operator X is defined,

$$Xu\colon t\to\int_0^t x(t-s)\,u(s)\,ds.$$

Furthermore, $I+X$ is causally invertible and $(I+X)^{-1}-I$ is an integral operator with L_1 difference kernel $w(t-s)$.

Proof. First we assume $k(\cdot)\in\mathscr{K}$. From Theorem 3.4 we have for each $\mu\in[0,1]$ a factorization of the operator $I+\mu K$. Thus Corollary 3.6 implies the existence of an open set $\Theta \subset C$ such that $[0, 1] \subset \Theta$ and for each $\lambda \in \Theta$, $I + \lambda K$ has the factorization, say

$$I + \lambda K = (I + X^*(\lambda))(I + X(\lambda)), \qquad \lambda \in \Theta. \tag{4.7}$$

Furthermore, for $|\lambda|$ sufficiently small (say $|\lambda| < \delta'$) from Theorem 3.5

$$X(\lambda) = \lambda K_+ + \lambda^2 F(K)\, a(\lambda K)\, h(K).$$

Now for $y \in H_\infty$ (cf. Lemma 4.1), $a(\lambda K)\, y$: $t \to (I + \lambda P_t K P_t)^{-1}\, y(t)$. Thus for $|\lambda| < \delta$, where $\delta = \min\{\delta', |k|_{\mathscr{K}}\}$, the following expansion is valid,

$$a(\lambda K)\, h(K) = \left[\sum_{n=0}^{\infty} \tilde{a}^n(\lambda K) \right] h(K),$$

where $\tilde{a}(\lambda K) \in B(H_\infty)$, $\tilde{a}(\lambda K)\, y$: $t \to P_t \lambda K y(t)$. Hence we may write

$$X(\lambda) = \lambda K_+ + \lambda^2 \left\{ \sum_{n=0}^{\infty} F(K)\, \tilde{a}(\lambda K) \right\} h(K).$$

Define the sequence of functions $\{v_i(\cdot)\}_{i=0}^{\infty} \subset \mathscr{K}$ by the recursion

$$v_n(t) = \lambda \int_t^\infty v_{n-1}(t-s)\; k(s)\, ds, \tag{4.8}$$

with $v_0(t) = k(t)$. Also, let V_n denote the integral operator with kernel $v_n(t-s)$. Noting that $F(K)\, \tilde{a}^n(\lambda K) = \lambda F(K)\, \tilde{a}^{n-1}(\lambda K)\, P^- \tilde{k}$ (cf. Lemma 4.2), it follows that $F(K)\, \tilde{a}^n(\lambda K) = F(V_n)$. Since $|\lambda| < \delta$, from (4.8) we have

$$\sum_{n=0}^{\infty} v_n(t) = v(t) \in \mathscr{K},$$

where the convergence is understood in the sense of the $\mathscr{K}$-metric. If we define the operator V with kernel $v(t-s)$, the L_1-convergence implies

$$\sum_{n=0}^{M} V_n \to V$$

in the operator norm, and the L_2-convergence implies the existence of a constant α such that

$$\left| E(\omega) \sum_{n=0}^{m} V_n \right| \leqslant \alpha \sqrt{\lambda(\omega)}$$

for all m. Thus Lemma 3.2 applies and we obtain the representation

$$X(\lambda) = \lambda K_+ + \lambda^2 F(V)\, h(K).$$

Defining $g \in \mathscr{K}$ by

$$g(t) = \begin{cases} \int_t^\infty v(t-s)\, k(s)\, ds, & t \geqslant 0, \\ 0, & t < 0, \end{cases}$$

it follows from Lemma 4.2 that $X(\lambda)$ is an integral operator with kernel $x_\lambda(t-s)$, where $x_\lambda(\cdot) \in \mathscr{K}$,

$$x_\lambda(t) = \begin{cases} \lambda k(t) + \lambda^2 g(t), & t \geqslant 0, \\ 0, & t < 0. \end{cases}$$

It is evident that a suitable restriction of the magnitude of $|\lambda|$ results in $|x_\lambda(\cdot)|_{\mathscr{K}} < 1$. Thus there exists unique $w_\lambda \in \mathscr{K}$ such that $w_\lambda(t) = 0$ for $t < 0$ and

$$w_\lambda(t) + x_\lambda(t) + \int_0^t x_\lambda(t-s)\, w_\lambda(s)\, ds = 0. \tag{4.9}$$

It is also clear from (4.9) that the mapping $I + W(\lambda)$, where

$$W(\lambda)u\colon t \to \int_0^t w_\lambda(t-s)\, u(s)\, ds$$

is the inverse of the mapping $(I + X(\lambda))$.

Now returning to the factorization (4.7) we have (for $|\lambda|$ such that (4.9) holds),

$$X(\lambda) = \lambda K_+ + \lambda F(W^*(\lambda))\, h(K).$$

Thus from Lemma 4.1,

$$x_\lambda(t) = \lambda k(t) + \lambda \int_t^\infty w_\lambda^*(s-t)\, k(s)\, ds, \qquad t \geqslant 0.$$

Or equivalently,

$$x_\lambda(\cdot) = \lambda(I + W^*(\lambda))\, k_+(\cdot),$$

where $k_+(t) = \chi[0, \infty](t)\, k(t)$. Applying $(I + W(\lambda))$ to the above results in

$$(I + W(\lambda))\, x_\lambda(\cdot) = (I + W(\lambda))(I + W^*(\lambda))\, \lambda k_+(\cdot).$$

Thus noting (4.7),

$$(I + W(\lambda)) x_\lambda(\cdot) = (I + \lambda K)^{-1} \lambda k_+(\cdot).$$

And noting (4.9),

$$w_\lambda(\cdot) = (I + \lambda K)^{-1} \lambda k_+(\cdot). \tag{4.10}$$

Since $I + \lambda K$ is invertible on $L_1((0, \infty), C^n)$ (see [5]), it follows that the right side of the above defines an analytic function on Θ with values in $L_1((0, \infty), C^{n \times n})$ so we may assume $w_\lambda(\cdot)$ is defined for $\lambda \in \Theta$. Now define the analytic $L_1((0, \infty), C^{n \times n})$-valued function

$$\tilde{x}_\lambda(t) = \lambda k(t) + \lambda \int_t^\infty w_\lambda^*(s - t)\, k(s)\, ds. \tag{4.11}$$

It is evident that the $B(H)$-valued operator function on Θ, $\tilde{X}(\lambda)$ defined

$$\tilde{X}(\lambda) u : t \to \int_0^t \tilde{x}_\lambda(t - s)\, u(s)\, ds, \tag{4.12}$$

is also analytic. From Corollary 3.6 we have that $\lambda \to X(\lambda)$ (in the factorization (4.7)) is analytic. Furthermore for $|\lambda|$ sufficiently small we have already shown that $\tilde{X}(\lambda) = X(\lambda)$. Thus $\tilde{X}(\lambda) = X(\lambda)$ in Θ. In particular, $\tilde{X}(1) = X(1)$.

To remove the restriction that $k \in \mathscr{K}$, first note that $\mathscr{K}$ is dense in $L_1((-\infty, \infty), C^{n \times n})$. Therefore, there exists a sequence $\{k_n\} \subset \mathscr{K}$ such that $k_n \to k$ (in the L_1-topology) and the associated operators $I + K_n$ have the factorizations

$$I + K_n = (I + X_n^*)(I + X_n). \tag{4.13}$$

Now it is evident that (4.10)–(4.12) depend continuously on $k(\cdot)$ (with respect to the L_1-topology). Thus it follows that the choice $x(\cdot) \in L_1((0, \infty), C^{n \times n})$ for the kernel of the operator X defined by

$$x(t) = k(t) + \int_t^\infty w^*(s - t)\, k(s)\, ds,$$

where

$$w(\cdot) = (I + K)^{-1}\, k_+(\cdot)$$

leads to the factorization

$$I + K = (I + X^*)(I + X) \tag{4.14}$$

in the general case.

Finally we show that $I + X$ is invertible. Using (4.13) and the fact that

$X_n \to X$ in $B(H)$, the invertibility argument of the proof of Theorem 3.4 leads to the identity

$$(I+W)(I+X)=I, \tag{4.15}$$

where W is the causal map, $W=(I+K)^{-1}(I+X^*)-I$ (take limits in (4.13)). Thus it suffices to show that $(I+W)$ is one-to-one. Assume there exists $v\in H$ such that $(I+W)v=0$. Then for any $s>0$ it follows from the causality of W that $(I+P^sWP^s)\,P^sv=0$. But (4.15) implies for any $s>0$,

$$(I+P^sWP^s)(I+P^sXP^s)=I.$$

But P^sXP^s is a Volterra operator with L_1-kernel on the interval $[0,s]$, and $P^s+P^sXP^s$ is therefore invertible on P^sH for any s. Hence $v=0$. The theorem is completely proved. ∎

Before we move on to the next example, some remarks concerning a couple of easy extensions of this proof will be given.

The first extension is to replace the underlying finite dimensional space C^n with any separable Hilbert space H_0 and regard K as an integral operator on the space $L_2((0,\infty),H_0)$ with strongly measurable kernel $k(t)\in B(H_0)$ satisfying

$$\int_{-\infty}^{\infty} |k(t)|_{B(H_0)}\,dt<1. \tag{4.16}$$

This assumption is sufficient to guarantee that (4.10) holds in some neighborhood of $[0,1]$. The only place in the proof of the theorem the finite dimensionality of C^n was invoked was to use a result of Gohberg and Krein [5] to obtain (4.11) from (4.10). Thus (4.16) is enough to allow the proof to be valid in the infinite dimensional setting as well. Similar results appear in [4, 10], for example, when $k(t)$ is compact-valued and H_0 is relaxed to a Banach space. Gohberg's [4] results are most comprehensive here.

If we assume $k(t)$ to be compact valued and $|k(t)|$ to satisfy the integrability conditions defining the space $\mathscr{K}$ as in the beginning of this section, the general theory of Section 3 can be used to remove the condition (4.16) and obtain a variation of Theorem 4.3. So now let P_N be a sequence of finite-dimensional projections on H_0 converging strongly to the identity, and define for each N, $k^N(t)=P_Nk(t)P_N$ and let K_N denote the operator with difference kernel $k_N(t-s)$. Then $k_N\to k$ a.e. in the uniform topology, and by dominated convergence

$$\lim_N \int_{-\infty}^{\infty} \{|k_N(t)-k(t)|^2+(1+|t|^{1/2})\,|k_N(t)-k(t)|\}\,dt=0.$$

Example 2.5 and Theorem 3.4 imply $I+K$ and $I+K_N$ have factorizations

$$I+K=(I+X^*)(I+X), \qquad (I+K_N)=(I+X_N^*)(I+X_N)$$

with $(I+X_N)$ causal and causally invertible. We claim that $X_N \to X$ in the uniform topology. To see this we return to (3.8) and obtain

$$|X_N - X| \leqslant |F(K-K_N)\, a(K)\, h(K)| + |F(K_N)|\, |a(K)\, h(K) - a(K_N)\, h(K_n)|. \tag{4.17}$$

Let $x \in L_2((0,\infty), H_0)$ with $|x|=1$. From Lemma 4.1 we have

$$\begin{aligned} &|F(K-K_N)\, a(K)\, h(K)x| \\ &\quad \leqslant \int_0^\infty \left| \int_0^\infty (k_N(t-s)) - k(t-s))[a(K)\, h(K)x](t)(s)\, ds \right|^2 dt \\ &\quad \leqslant \int_0^\infty \left\{ \int_0^\infty |k_N(t-s)-k(t-s)|^2\, ds \int_0^\infty |[a(K)\, h(K)x](t)(\sigma)|^2\, d\sigma \right\} dt \end{aligned}$$

But for each t,

$$\lim_N \int_0^\infty |k_N(t-s)-k(t-s)|^2\, ds = 0.$$

So by dominated convergence it follows that the first term in (4.17) tends to zero, since

$$\int_0^\infty \int_0^\infty |[a(K)\, h(K)x](t)(\sigma)|^2\, d\sigma\, dt = |a(K)\, h(K)x|^2 < \infty.$$

To show that the second term in (4.17) also tends to zero, we note that since $K_N \to K$ uniformly, $a(K_N) \to a(K)$ uniformly and $h(K_N) \to h(K)$ uniformly (cf. (2.14)). Thus

$$\lim_N |F(K_N)| \;\; |a(K)\, h(K) - a(K_N)\, h(K_N)| = 0$$

follows from the triangle inequality and the boundedness of $\{|F(K_N)|\}$. Hence, we have $X_N \to X$ uniformly. Now note that since $P_N H_0$ is finite dimensional, the theorem implies that X_N is an integral operator with (matrix) kernel $x_N(t-s)$. Therefore,

$$|X_N| = \sup_{-\infty<\lambda<\infty} |\hat{x}_N(\lambda)|,$$

where

$$\hat{x}_N(\lambda) = \int_0^\infty x_N(t)\, e^{i\lambda t}\, dt$$

and the norm is taken as the operator (matrix) norm over $P_N H_0$. Note also that $\hat{x}_N(\lambda)$ can be continued analytically in the upper half plane π_+. Now since $X_N \to X$ uniformly, it follows that $\hat{x}_N$ converges in the Hardy space $H^\infty(\pi_+, B(H_0))$ of bounded analytic $B(H_0)$-valued functions on Π_+ to an element $\hat{x}$ such that for any $u \in L_2((0, \infty), H_0)$,

$$\widehat{Xu}(\lambda) = \hat{x}(\lambda)\, \hat{u}(\lambda),$$

where ^ denotes the Fourier transform. Thus, denoting

$$\hat{k}(\lambda) = \int_{-\infty}^\infty k(t)\, e^{i\lambda t},$$

we obtain the factorization

$$I + \hat{k}(\lambda) = (I + \hat{x}^*(\lambda))(I + \hat{x}(\lambda))$$

with $\hat{x} \in H^\infty(\Pi_+, B(H_0))$. Furthermore, noting that the argument above also applies to $(I + X)^{-1} - I$, we also have

$$(I + \hat{x}(\lambda))^{-1} - I = \hat{w}(\lambda)$$

with $\hat{w} \in H^\infty(\Pi_+, B(H_0))$. More complete results along these lines can be found in [4, 16, 17].

Our next example concerns the factorization of an operator that arises in connection with the infinite time horizon linear regulator problem. The causal factor here has a particularly nice representation as well shall see.

Let H_1 and H_2 denote separable Hilbert spaces and let $U = L_2((0, \infty), H_1)$ and $X = L_2((0, \infty), H_2)$. We shall consider the factorization of the operator $I + T^*T \in B(U)$, where $T \in B(U, X)$,

$$Tu\colon t \to Q \int_0^t S(t - \sigma)\, Bu(\sigma)\, d\sigma, \tag{4.18}$$

$B \in B(H_1, H_2)$, $Q \in B(H_2)$, and $S(\cdot)$ is an exponentially stable C_0 semigroup on H_2 (i.e., there exist $\alpha, \beta > 0$ such that $|S(t)| \leqslant \alpha e^{-\beta t}$). It follows routinely from (4.18) that T^*T satisfies the hypotheses of Example 2.5. Hence, Theorem 3.4 applies and we have the factorization

$$I + T^*T = (I + V^*)(I + V). \tag{4.19}$$

We will show that there exists $K \in B(H_2, H_1)$ such that V has the representation

$$Vu: t \to \int_0^t KS(t-\sigma)\, Bu(\sigma)\, d\sigma. \tag{4.20}$$

Introducing the notation $W = (I+V)^{-1} - I$, we can write from (4.19) (see proof of Proposition 2.4),

$$V = F((I+W^*)\, T^*)\, h(T). \tag{4.21}$$

As in [15], the following decomposition holds:

$$h(T) = MS, \tag{4.22}$$

where $S \in B(U, X)$,

$$Su: t \to \int_0^t S(t-\sigma)\, Bu(\sigma)\, d\sigma, \tag{4.23}$$

and $M \in M(X, L_2((0, \infty), X))$,

$$[Mx](t): \sigma \to \begin{cases} QS(\sigma - t)\, x(t), & \sigma \geqslant t, \\ 0, & \sigma < t. \end{cases} \tag{4.24}$$

Thus, $V = F((I+W^*)T^*)\, MS$. But $F((I+W^*)T^*)M \in M(X, U)$. Hence (from the straightforward extension of [13, Proposition 3.1] to the infinite interval case), we conclude there exists a strongly measurable essentially bounded $B(H_2, H_1)$-valued function $K(\cdot)$ on $(0, \infty)$ such that for any $x \in X$,

$$[F((I+W^*)T^*)\, Mx](t) = K(t)\, x(t) \qquad \text{a.e.} \tag{4.25}$$

It remains to show that $K(t)$ is constant. We note that if H_1 is finite dimensional, Theorem 4.3, together with Lemma 4.1, implies the result. To deduce the result in the infinite dimensional case we let P_N denote a sequence of finite dimensional orthoprojectors on H_1 converging strongly to I and define the sequence $\{T_N\} \subset B(U, X)$,

$$T_N u: t \to Q \int_0^t S(t-\sigma)\, BP_N u(\sigma)\, d\sigma.$$

Clearly $T_N^* T_N \to T^*T$ strongly. And from Theorem 3.4 we have the factorization

$$I + T_N^* T_N = (I + V_N^*)(I + V_N),$$

where

$$V_N = F(T_N^*)\, h(T_n) - F(T_N^* T_N)\, a(T_N^* T_N)\, h(T_N^* T_N).$$

Now,

$$h(T_N^* T_N)x\colon t \to P_t T_N^* T_N P^t x = P_t T_N^* P_t T_N P^t x.$$

Thus we can write

$$h(T_N^* T_N) = b(T_N^*)\, h(T_N),$$

where $b(T_N^*) \in M(L_2((0, \infty), X),\ \ L_2((0, \infty), U))$, $b(T_N^*)x\colon t \to P_t T_N^* x(t)$. Therefore

$$V_N = [F(T_N^*) - F(T_N^* T_N)\, a(T_N^* T_N)\, b(T_N^*)]\, h(T_N). \tag{4.26}$$

From the definition of T_N and Lemma 3.2 it is straightforward to verify that the term in brackets in (4.26) converges strongly as $N \to \infty$. Next we show that $h(T_N) \to h(T)$ strongly. To see this we first note that

$$|P_t(T - T_N)\, P^t u|^2 \leqslant \int_t^\infty |Q|^2 \left\{ \int_0^t |S(\tau - \sigma)\, B[I - P_N]\, u(\sigma)|\, d\sigma \right\}^2 d\tau.$$

So for each t, as $N \to \infty$, $|P_t(T - T_N)\, P^t u|^2 \to 0$ by the dominated convergence theorem, since $P_N \to I$ strongly. We can now argue as in Example 2.5 (note the argument following (2.10) with $K = T - T_N$), to establish that the sequence of real-value functions $|P_t(T - T_N)\, P^t u|^2$ can be uniformly dominated by an integrable function. Thus by what was just shown, dominated convergence implies that $h(T_N) \to h(T)$ strongly. Therefore we can assert that $V_N \to V$ strongly. And since

$$(I + W_N^*) = (I + V_N)(I + T_N^* T_N)^{-1},$$

we also have $W_N^* \to W^*$ strongly (where $W_N = (I + V_N)^{-1} - I$). Further, the construction of V_N from Theorem 3.4 yields a constant α independent of N such that $|E(\omega)\, W_N^*| < \alpha \sqrt{\lambda(\omega)}$ for all $\omega \in \Sigma$. It then follows from Lemma 3.2 that $K_N \to K$ strongly where

$$K_N = F((I + W_N^*)\, T_N^*)\, M$$

and K is defined via the function $K(t)$ (cf. (4.25)), i.e., $[Kx](t) = K(t)\, x(t)$. Now because $P_N H_1$ is finite dimensional, Lemma 4.1 and Theorem 4.3 imply that K_N has the representation $[K_N x](t) = K_N x(t)$, where $K_N \in B(H_2, H_1)$. And since $K_N \to K(t)$ strongly for a.e. t, $K(\cdot)$ is necessarily constant.

With the spaces H_1, H_2, U, and X defined as above, the foregoing discussion can be summarized as follows.

THEOREM 4.4. *Let T be defined as in* (4.18). *Then $I+T^*T$ has the factorization*

$$I+T^*T=(I+V^*)(I+V),$$

where the factor V has the form

$$Vu: t \to \int_0^t KS(t-\sigma)\, Bu(\sigma)\, d\sigma$$

with $K \in B(H_2, H_1)$ defined via (4.24)–(4.25).

Remark. The nonstationary analog of the theorem above also holds. Specifically, suppose now that $T \in B(U, X)$ is defined

$$Tu: t \to \int_0^t Q(t)\, S(t, \sigma)\, B(\sigma)\, u(\sigma)\, d\sigma$$

with $B(\cdot)$ and $Q(\cdot)$ strongly measurable and essentially bounded on $[0, \infty]$, and $S(\cdot, \cdot)$ is a strongly continuous evolution operator such that $|S(t, \sigma)| \leqslant \alpha e^{-\beta(t-\sigma)}$; $\alpha, \beta > 0$. Then after making the appropriate substitutions into (4.21)–(4.25), it follows that $I+T^*T$ has the factorization

$$I+T^*T=(I+V^*)(I+V),$$

where V has the form

$$Vu: t \to \int_0^t K(t)\, S(t, \sigma)\, B(\sigma)\, u(\sigma)\, d\sigma$$

with $K(\cdot)$ strongly measurable and essentially bounded on $[0, \infty)$.

We can apply this result to the infinite time linear regulator quadratic cost problem. The argument here is essentially an extrapolation of its finite time counterpart in [15].

Using the notations above let $\tau \in (0, \infty)$ and consider the regulator problem with dynamics

$$x(t)=w(t)+\int_\tau^t S(t-\sigma)\, Bu(\sigma)\, d\sigma$$

and cost

$$J(u, x)=\int_0^\infty |Q^*Qx(t)|^2+|u(t)|^2\, dt.$$

We will assume that $w(\cdot) \in L_2((0, \infty), H_2)$,

$$w(t) = \int_0^\tau S(t-\sigma)\, Bv(\sigma)\, d\sigma$$

for some $v(\cdot) \in L_2((0, \infty), H_1)$. Using the definition of T in (4.18) we can pose this optimization problem as

$$\min_{\substack{u \in U \\ y \in X}} |y|^2 + |u|^2$$

subject to the constraint

$$y = TP^\tau v + TP_\tau u.$$

The "open-loop" solution is easily obtained:

$$\hat{u} = (I + P_\tau T^* T P_\tau)^{-1} P_\tau T^* T P^\tau v.$$

Now the factorization in Theorem 4.4 implies

$$\begin{aligned} P_\tau T^* T P^\tau &= P_\tau [V + V^* + V^* V] P^\tau \\ &= P_\tau (I + V^*) V P^\tau \\ &= (I + P_\tau V^* P_\tau) P_\tau V P^\tau. \end{aligned}$$

Thus, noting that (using Theorem 3.4 and Proposition 3.1)

$$(I + P_\tau T^* T P_\tau)^{-1} = (I + P_\tau V P_\tau)^{-1} (I + P_\tau V^* P_\tau)^{-1},$$

it follows that

$$\hat{u} = -(I + P_\tau V P_\tau)^{-1} P_\tau V P^\tau v.$$

Hence,

$$\hat{u} = -P_\tau V [P_\tau \hat{u} + P^\tau v].$$

But from the theorem

$$Vz\colon t \to \int_0^t KS(t-\sigma)\, Bz(\sigma)\, d\sigma.$$

In particular,

$$\begin{aligned} \hat{u} &= -K \left\{ \int_0^\tau S(t-\sigma)\, Vv(\sigma)\, d\sigma + \int_\tau^t S(t-\sigma)\, Bu(\sigma)\, d\sigma \right\} \\ &= -K\hat{x}. \end{aligned}$$

References

1. J. H. Davis, Wiener–Hopf methods for open-loop unstable distributed systems, *SIAM J. Control. Opt.* **17** (1979), 713–728.
2. A. Devinatz, The factorization of operator values functions, *Ann. of Math.* **73** (1961), 458–495.
3. A. Feintuch, Strong causality conditions and causal invertibility, *SIAM J. Control Opt.* **18** (1980), 317–324.
4. I. C. Gohberg, The factorization problem for operator functions, *Amer. Math. Soc. Transl.* **49** (1966), 130–161.
5. I. C. Gohberg and M. G. Krein, Systems of integral equations on a half-line with kernels depending on the difference of arguments, *Amer. Math. Soc. Transl.* **14** (1960), 217–288.
6. I. C. Gohberg and M. G. Krein, "Theory of Volterra Operators in Hilbert Space and its Applications," Amer. Math. Soc., Providence, RI, 1970.
7. I. C. Gohberg and M. G. Krein, Description of contraction operators which are similar to unitary operators, *Functional Anal. Appl.* **1** (1967), 33–52.
8. J. W. Helton, A spectral factorization approach to the distributed stable regulator problem: The algebraic Riccati equation, *SIAM J. Control Opt.* **14** (1976), 639–661.
9. T. Kailath, A note on least squares estimation by the innovations method, *SIAM J. Control* **10** (1972), 447–486.
10. C. T. Kelley, Convolution and H-equations for operator-valued functions with applications to neutron transport theory, *J. Math. Phys.* **18** (1977), 764–769.
11. M. G. Krein, On inverse problems for a nonhomogeneous cord, *Dokl. Akad. Nauk SSSR* **82** (1952).
12. A. McNabb and A. Schumitzky, Factorization of operators—III: Initial value methods for linear two-point boundary value problems, *J. Math. Anal. Appl.* **31** (1970), 391–405.
13. M. Milman, Special factorization and Riccati integral operations, *J. Math. Anal. Appl.* **100** (1984), 155–187.
14. M. Milman, J. Foster, and A. Schumitzky, Optimal feedback control of infinite dimensional linear systems, *J. Math. Anal. Appl.*, in press.
15. M. Milman and A. Schumitzky, On a class of operators on Hilbert space with applications to factorization and systems theory, *J. Math. Anal. Appl.* **99** (1984), 494–512.
16. B. Sz.-Nagy and C. Foias, "Harmonic Analysis of Operators on Hilbert Space," North-Holland, Amsterdam, 1970.
17. M. Rosenblum and J. Rovynak, The factorization problem for nonnegative operator valued functions, *Bull. Amer. Math. Soc.* **77** (1971), 287–318.
18. A. Schumitzky, On the equivalence between matrix Riccati equations and Fredholm resolvents, *J. Comput. System Sci.* **2** (1968), 76–87.
19. N. Wiener, "Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications," Wiley, New York, 1949.