OUTPUT-INDUCED SUBSPACES, INVARIANT DIRECTIONS AND INTERPOLATION IN LINEAR DISCRETE-TIME STOCHASTIC SYSTEMS *

ANDERS LINDQUIST † AND GYÖRGY MICHALETZKY ‡

Abstract. In this paper we analyze the structure of the class of discrete-time linear stochastic systems in terms of the geometric theory of stochastic realization. We discuss the role of invariant directions, zeros of spectral factors and output-induced subspaces in determining the systems-theoretical properties of the stochastic systems. A prototype interpolation problem for recovering lost state information is discussed and it is shown how it can be solved via Kalman filtering recursions tying together the state processes of a family of totally ordered splitting subspaces.

Key words. invariant directions, zero dynamics, discrete-time stochastic systems, splitting subspaces

AMS subject classifications. 93E03, 93B27, 60G10

1. Introduction.

It is a somewhat surprising fact that, in the discrete time case, the family of minimal state space representations

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \tag{1.1}$$

of a stationary stochastic vector process $\{y(t); t \in \mathbb{Z}\}$ with a rational spectral density exhibits a remarkably rich structure, affecting the implementation of most estimation algorithms, and that much of this structure is not present in the corresponding continuous-time setting. This diversity is also reflected in the structure of the corresponding family of matrix Riccati equations, studied in detail in [22] in the context of invariant directions of matrix Riccati equations [5, 25, 26], a phenomenon that is not present in the continuous-time case.

As usually, $\{u(t); t \in \mathbb{Z}\}$ is a vector-valued white noise process, which passed through a stable filter with transfer function

$$W(z) = C(zI - A)^{-1}B + D, \tag{1.2}$$

beginning at the remote past, produces the output process $\{y(t); t \in \mathbb{Z}\}$, say, of dimension m and with an $m \times m$ spectral density $\Phi(z) = W(z)W(1/z)'$. (Here $'$ denotes transpose and the white noise u is a zero-mean process such that $\mathrm{E}\{u(t)u(s)'\} = I\delta_{ts}$.)

The output process y is of course purely nondeterministic, and we assume that its spectral density Φ is full rank. The representation (1.1) is a minimal realization in the sense that the *state process* x has as few components as possible.

Obviously W is a rational spectral factor having all its poles strictly inside the unit circle, implying that the same holds for the eigenvalues of A. Note that we are

*This research was supported in part by grants from the Swedish Research Council for Engineering Sciences (TFR), the Hungarian National Science Foundation (OTKA, No 2042 and T15668), and the Göran Gustafsson Foundation.

†Division of Optimization and Systems Theory, Royal Institute of Technology, 100 44 Stockholm, Sweden

‡Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest, Hungary and Laboratory of Operations Research and Decision Systems, Computer and Automation Institute, Hungarian Academy of Sciences, Budapest, Hungary.

not confining ourselves to square spectral factors W, as the dimension p of the input noise process could be larger than m. In particular, this implies that the state vector may not be expressible in terms of the output process $y(t), t \in \mathbb{Z}$ alone but that it depends on some unobserved exogenous noise also. Another consequence is that the number of zeros of W may be fewer than the number of poles of W, even if D is full rank.

Part of this paper is devoted to the following prototype *interpolation problem*, which is of a somewhat different type than the interpolation problem considered in [23, 24]. Suppose, that we observe the state x as well as the output y on some finite or infinite interval except that there is a blackout of state information on some finite subinterval (t_0, t_1). Then the problem is to reconstruct the lost state information in the least squares sense, given the noisy output and the remaining states information. This problem provides a framework for studying many important questions concerning the structure of discrete-time linear stochastic systems.

This interpolation problem is preferably studied in the context of the geometric theory of stochastic realization (see [17, 18, 14, 15] and references therein), in which the properties of the state representation (1.1) are expressed in terms of the *minimal Markovian splitting subspace*

$$X = \{a'x(0) \mid a \in \mathbb{R}^n\}, \tag{1.3}$$

where $n := \dim X$ is the number of components of $x(0)$ so that $x(0)$ forms a basis in X. Due to stationarity, it is sufficient to study $\{x(t); t \in \mathbb{Z}\}$ at time $t = 0$. The family of such X corresponding to a given y will be denoted $\mathcal{X}$. It is known that $\mathcal{X}$ is endowed with a certain partial ordering. This ordering, reviewed in §2, will play an important role in this paper.

A basic tool in the analysis of the interpolation problem, and, more generally, the structural properties of the family $\mathcal{X}$ of state space representations, is a pair $(\sigma, \bar{\sigma})$ of shift operators on $\mathcal{X}$, which, given any $X \in \mathcal{X}$ produces a family $\{X^{(k)} \mid k \in \mathbb{Z}\}$ of totally ordered splitting subspaces. We show that these splitting subspaces are tied together by Kalman filtering recursions in the sense that we can pass from one state process $x^{(k)}$ to the next by (forward or backward) Kalman filtering, a remarkable fact that enables us actually to compute these spaces.

These sequences of splitting subspaces provide a deeper insight into the structure of the related discrete-time matrix Riccati difference equation. In fact, the corresponding sequence of state covariance matrices constitutes a solution of this Riccati equation. It is well known that the limits at $-\infty$ and at ∞ are solutions of the steady-state (algebraic) Riccati equation but our procedure also enables us to study the transient behavior of these equations. This should be compared with the corresponding continuous-time results in [13].

The interpolation estimate of $x(t)$ on the interval (t_0, t_1) turns out to be a linear combination of $x^{(t_0-t)}(t)$ and $x^{(t_1-t)}(t)$, the state processes of $X^{(t_0-t)}$ and $X^{(t_1-t)}$ respectively, in a certain uniform choice of coordinates, enabling us to use these Kalman recursions to determine the estimate. As $t_0 \to -\infty$ and $t_1 \to \infty$, we obtain the corresponding prototype *smoothing problem* and the structure of the solution is similar to those presented in [3] and in [18].

We show that the computational burden of determining the interpolation estimates depends on the dimension of the *internal subspace* $X \cap H_0$ of the splitting subspace, where H_0 is the closure, in the inner product $\langle \xi, \eta \rangle := \mathrm{E}\{\xi, \eta\}$, of all random variables

$$\{y_i(t) \mid i = 1, 2, \dots, m;\, t \in \mathbb{Z}\}$$

of the output process. In fact, we show that, if $\dim X \cap H_0 = n - \nu$, we only need to solve matrix Riccati equations of dimension at most $\nu \times \nu$ rather than $n \times n$, to compute the appropriate filter estimates. Sometimes, however, we need an initial number of time steps to achieve this reduction, and to understand this better we need to study the structure of the internal subspace $X \cap H_0$.

In this paper we show among other things that the internal subspace has the direct sum decomposition

$$X \cap H_0 = X \cap \{y(-1), \dots, y(-n)\} + Y^* + X \cap \{y(0), \dots, y(n-1)\},$$

where the subspace Y^* can be determined by algorithms akin to the one used to compute the maximal output-nulling subspace in geometric control theory [31]. This decomposition and the theoretical framework in which it is developed give a considerable amount of information about the structure of the discrete-time linear stochastic system (1.1).

First, if the *predictable subspace* $X \cap \{y(-1), \dots, y(-n)\}$ is nontrivial, there is an $a \in \mathbb{R}^n$ such that

$$a'x(t) \in \{y(t-1), y(t-2), \dots, y(t-n)\},$$

and consequently the *usual* Kalman filtering problem of estimating $x(t)$ given the data $y(t-1), y(t-2), \dots, y(0)$ reaches steady state in a finite number of steps in the direction a. An analogous statement holds for the initial point smoothing problem and the *smoothable subspace* $X \cap \{y(0), \dots, y(n-1)\}$. Nontrivial such directions a are known as *invariant directions* and have been studied extensively in the literature [5, 25, 26, 22], but the connections to the geometric theory of Markovian splitting subspaces are presented here for the first time.

Secondly, the basic reason why discrete-time models (1.1) are more complicated, and the study of them is more challenging, than in the continuous-time case is that DD' varies over $\mathcal{X}$. If $DD' > 0$ for all $X \in \mathcal{X}$, the results and the analysis of the (coercive) continuous-time case generally carry over verbatim. This is known as the *regular case*. In the regular case there are no invariant directions and $Y^* = X \cap H_0$. In this paper we give several geometric characterizations of regularity and investigate the fine structure of the nonregular case.

Thirdly, the zero structure of the transfer function (1.2) plays an important role in the analysis of the interpolation problem, and it can be studied in terms of *output-induced subspaces*, i.e. subspaces of $X \cap H_0$ with certain invariance properties to be specified below. The output-induced subspaces also provide a link between stochastic realization theory and geometric control theory [31, 4] (see Remark 7.3). This program was initiated in [18, 19] and was continued in [13] and [29], where, in particular, the connections to geometric control theory are discussed in great detail in continuous and discrete time, respectively. In this paper we introduce the concept of *strictly output-induced subspaces*, a refinement needed to study the discrete-time case. In particular, Y^* is the maximal strictly output-induced subspace, which plays the role of $X \cap H_0$ in the nonregular case. The zero structure also provides information about the possible reduction of the Riccati recursions in the interpolation problem.

The paper is organized as follows. Section 2 is devoted to preliminaries on the geometric theory of stochastic realization theory and to notations. In §3 we introduce the operators σ and $\bar{\sigma}$, characterize regularity in terms of these, and establish the properties of the family $\{X^{(k)} | k \in \mathbb{Z}\}$, and in §4 we introduce the interpolation problem and relate it to the results in §3. Section 5 is about the zero structure of $\{X^{(k)} | k \in \mathbb{Z}\}$ in the regular case. In §6 we discuss output-induced subspaces, and in §7 the role of invariant directions is investigated and the algorithm for determining Y^* is given. Finally, in §8, the change in zero structure when applying σ and $\bar{\sigma}$ is discussed and the connections to the zero dynamics operators and the reduction of the Riccati equations in the interpolation problem are explained.

2. Preliminaries and notations. Given a stationary purely nondeterministic m-dimensional stochastic process $\{y(t); t \in \mathbb{Z}\}$, any stochastic realization (1.1) of y may be represented in a coordinate-free manner by a triplet (X, H, U) where X is given by (1.3), underscoring the fact that two representations (1.1) are considered identical if they only differ by the choice of coordinates in X. Here H is the Hilbert space generated by the random variables

$$\{u_i(t) \mid i = 1, 2, \dots, p;\ t \in \mathbb{Z}\},$$

with inner product

$$\langle \xi, \eta \rangle = \mathrm{E}\{\xi, \eta\},$$

and the unitary operator $U : H \to H$ is the shift determined by

$$U u_i(t) = u_i(t+1).$$

Then U acts as the shift for all processes in the system, i.e., $U y_i(t) = y_i(t+1)$ and $U x_i(t) = x_i(t+1)$. We always assume that the matrix $\begin{bmatrix} B \\ D \end{bmatrix}$ has linearly independent columns so that H is generated also by

$$\{y_i(t), x_j(t) \mid i = 1, \dots, m;\ j = 1, \dots, n;\ t \in \mathbb{Z}\}.$$

The Hilbert space H so defined is called the *ambient space* of X.

For any subspace $Y \subset H$ we shall write $\mathrm{E}^Y \lambda$ to denote the orthogonal projection of $\lambda \in H$ onto Y. Occasionally we shall misuse notations somewhat by writing $\mathrm{E}^Y z$ when z is a random vector to denote the vector with components $\{\mathrm{E}^Y z_i\}$. By $\mathrm{E}^Y Z$ we shall mean the closure of $\{\mathrm{E}^Y \zeta \mid \zeta \in Z\}$. For any pair of subspaces Y and Z we write $Y + Z$ to denote direct sum (implying that $Y \cap Z = 0$), $Y \oplus Z$ for orthogonal direct sum, and $Y \vee Z$ for the vector sum in the general case, i.e., for closure$\{\eta + \zeta \mid \eta \in Y, \zeta \in Z\}$. Moreover, we write $Z^\perp$ to denote the orthogonal complement $H \ominus Z$ of Z in the ambient space H. Finally, we write $Z \perp Y \mid X$ to denote that Z and Y are *conditionally orthogonal* given X, i.e., that

$$\langle \eta - E^X \eta, \zeta - E^X \zeta \rangle = 0 \quad \text{for all } \eta \in Y,\ \zeta \in Z.$$

There are some important subspaces related to the given process y, which are subspaces of H for each representation (X, H, U), and which are considered fixed in this analysis. Define the *past space* H^- as the subspace generated by the random variables

$$\{y_i(t) \mid i = 1, 2, \dots, m;\ t = -1, -2, -3, \dots\},$$

and the *future space* H^+ as the subspace generated by

$$\{y_i(t) \mid i = 1, 2, \dots, m;\ t = 0, 1, 2, \dots\},$$

and let

$$H_0 := H^- \vee H^+ \subset H \tag{2.1}$$

be the space generated by all random variables in y. We shall also consider finite-dimensional subspaces $\{y(j), \dots, y(k)\}$ spanned by the components of the random vectors depicted inside the curly brackets. We shall also use the shorthand notation H^-_{t-1} and H^+_t for U^tH^- and U^tH^+ respectively, the past and future spaces shifted to time t. Then $H^-_{-1} = H^-$ and $H^+_0 = H^+$, which is consistent with the asymmetric definition of past and future.

It is well-known that X is a minimal Markovian splitting subspace [17, 18] and that it can be represented *uniquely* in terms of a pair $(S, \bar{S})$ of subspaces such that

$$S \supset H^- \quad \text{and} \quad \bar{S} \supset H^+, \tag{2.2}$$

$$U^{-1}S \subset S \quad \text{and} \quad U\bar{S} \subset \bar{S} \tag{2.3}$$

and

$$H = \bar{S}^\perp \oplus X \oplus S^\perp. \tag{2.4}$$

Consequently, S and $\bar{S}$ may be regarded as extensions of the past space H^- and future space H^+ respectively, inheriting their invariance properties, and they intersect perpendicularly so that

$$X = S \cap \bar{S} = \mathrm{E}^S\bar{S} = \mathrm{E}^{\bar{S}}S. \tag{2.5}$$

Conversely, S and $\bar{S}$ can be recovered from X in terms of

$$\begin{cases} S &= H^- \vee X^- \\ \bar{S} &= H^+ \vee X^+ \end{cases} \tag{2.6}$$

where $X^- := \bigvee_{t=-\infty}^{0} U^tX$ and $X^+ := \bigvee_{t=0}^{\infty} U^tX$. We shall write $X \sim (S, \bar{S})$ to exhibit the one-one correspondence between X and $(S, \bar{S})$.

Clearly, the ambient space has the representation

$$H = S \vee \bar{S}, \tag{2.7}$$

and $S \perp \bar{S} \mid X$, which is equivalent to

$$\mathrm{E}^S\lambda = \mathrm{E}^X\lambda \quad \text{for } \lambda \in \bar{S} \tag{2.8}$$

and to

$$\mathrm{E}^{\bar{S}}\lambda = \mathrm{E}^X\lambda \quad \text{for } \lambda \in S. \tag{2.9}$$

In particular, $H^- \perp H^+ \mid X$, i.e., X is a *splitting subspace.*

We recall that $X \sim (S, \bar{S})$ is *minimal* both in the sense of subspace inclusion and in the sense of dimension, two concepts of minimality which can be shown to be equivalent, if and only if

$$\bar{S} = H^+ \vee S^\perp \tag{2.10}$$

and

$$S = H^- \vee \bar{S}^\perp \tag{2.11}$$

[17, 18]. Condition (2.10) is equivalent to $X \cap (H^+)^\perp = 0$, i.e., to X being *observable*, and (2.11) to $X \cap (H^-)^\perp = 0$, i.e., to X being *constructible*. Therefore, in view of (2.5), we have

$$X = \mathrm{E}^S H^+ = \mathrm{E}^{\bar{S}} H^- \ , \tag{2.12}$$

whenever X in minimal.

The space S is actually identical to the subspace generated by the past of the driving white noise u in (1.1), so u can be constructed from S by Wold decomposition [14, 15]. Analogously, $\bar{S}$ corresponds to another white noise process $\{\bar{u}(t); t \in \mathbb{Z}\}$, the future space of which coincides with $\bar{S}$, and, passed through an antistable filter with transfer function

$$\bar{W}(z) = z\bar{C}(I - zA')^{-1}\bar{B} + \bar{D} \tag{2.13}$$

from the remote future, $\bar{u}$ produces a backward realization of y, namely

$$\begin{cases} \bar{x}(t-1) = A'\bar{x}(t) + \bar{B}\bar{u}(t-1) \\ y(t-1) = \bar{C}\bar{x}(t) + \bar{D}\bar{u}(t-1) \end{cases} \tag{2.14}$$

Here $\bar{x}(0)$ is just another basis in X such that

$$\bar{x}(t) = P^{-1}x(t), \tag{2.15}$$

where P is the state covariance

$$P = \mathrm{E}\{x(0)x(0)'\}. \tag{2.16}$$

The ambient space H will of course vary over the family $\mathcal{X}$ of minimal Markovian splitting subspaces. If $X \subset H_0$, then $H = H_0$ and we say that X is *internal.* We write $\mathcal{X}_0$ to denote the subclass of internal $X \in \mathcal{X}$.

The family $\mathcal{X}$ is endowed with a natural partial ordering [18]. We say that $X_1 \leq X_2$ if

$$\|\mathrm{E}^{X_1}\lambda\| \leq \|\mathrm{E}^{X_2}\lambda\| \quad \text{for all } \lambda \in H^+$$

or, equivalently,

$$\|\mathrm{E}^{X_2}\lambda\| \leq \|\mathrm{E}^{X_1}\lambda\| \quad \text{for all } \lambda \in H^-.$$

In this ordering the *predictor space* $X_- := \mathrm{E}^{H^-} H^+$ is the minimal element in $\mathcal{X}$ and $X_+ := \mathrm{E}^{H^+} H^-$ is the maximal element, i.e.,

$$X_- \leq X \leq X_+ \quad \text{for all } X \in \mathcal{X}. \tag{2.17}$$

Obviously, both X_- and X_+ are internal.

This ordering can be used to introduce a *uniform choice of bases* (or coordinates) in all $X \in \mathcal{X}$. In fact, let $x_+(0)$ be an arbitrary choice of basis in X_+ and define

$$x(0) = \mathrm{E}^X x_+(0) \tag{2.18}$$

for all $X \in \mathcal{X}$. This will insure the invariance of the matrices A and C over the class of forward minimal realizations (1.1). In the same way, we define

$$\bar{x}(0) = \mathrm{E}^X \bar{x}_-(0), \tag{2.19}$$

where $\bar{x}_-(0) = P_-^{-1} x_-(0)$, $x_-(0)$ being formed via (2.18) for $X = X_-$ and P_- being the corresponding state covariance (2.16). Then $\bar{C}$ will be invariant as well over the set of backward realizations (2.14).

Introducing coordinates in this uniform fashion, we can also parameterize the family $\mathcal{X}$ in terms of the corresponding class $\mathcal{P}$ of state covariances (2.16). The usual partial ordering of these positive definite matrices reflects the partial ordering of splitting subspaces in $\mathcal{X}$ introduced above. Consequently, in this parameterization (2.17) becomes

$$P_- \leq P \leq P_+ \quad \text{for all } P \in \mathcal{P}. \tag{2.20}$$

(Cf. [8].) In the same way, we can parameterize $\mathcal{X}$ in terms of the family $\bar{\mathcal{P}}$ of covariance matrices

$$\bar{P} := \mathrm{E}\{\bar{x}(0)\bar{x}(0)'\} = P^{-1} \tag{2.21}$$

of the backward realizations (2.14). Then (2.17) becomes

$$\bar{P}_+ \leq \bar{P} \leq \bar{P}_- \quad \text{for all } \bar{P} \in \bar{\mathcal{P}}. \tag{2.22}$$

3. An ordered family of splitting subspaces. A fact of central importance in this paper is that each splitting subspace $X \in \mathcal{X}$ can be naturally imbedded in a doubly infinite sequence of elements in $\mathcal{X}$, which contains finitely many different splitting subspaces if and only if $X \in \mathcal{X}_0$, i.e., X is internal. To see this, define operators σ and $\bar{\sigma}$ on $\mathcal{X}$ so that, for $X \sim (S, \bar{S})$,

$$\sigma X = \mathrm{E}^{H^- \vee (U^{-1}S)} X \tag{3.1}$$

$$\bar{\sigma} X = \mathrm{E}^{H^+ \vee (U\bar{S})} X \tag{3.2}$$

Observe that the operator σ is the geometric counterpart of a one-step-ahead state predictor given past output and state information. Our first result states, among other things, that σX is itself a splitting subspace so that $\sigma X \in \mathcal{X}$. Remarkably, as we shall see in §4, the states corresponding to $\{\sigma^k X\}$ are actually generated by a Kalman filter. Analogous statements hold for $\bar{\sigma}$ with respect to the backward setting.

Theorem 3.1. *Let $X \sim (S, \bar{S})$ be a minimal Markovian splitting subspace. Then*

(i) σX and $\bar{\sigma} X$ are minimal Markovian splitting subspaces and

$$\sigma X \leq X \leq \bar{\sigma} X. \tag{3.3}$$

Moreover, they have the same ambient spaces, namely $S \vee \bar{S}$.

(ii) $\sigma X = X$ *if and only if*

$$UX \subset X \vee \{y(0)\} \tag{3.4}$$

and $\bar{\sigma} X = X$ *if and only if*

$$U^{-1}X \subset X \vee \{y(-1)\}. \tag{3.5}$$

(iii) The fixed points of σ *and* $\bar{\sigma}$ *are internal minimal Markovian splitting subspaces.*

Proof. We prove all statements involving σ. Then those involving $\bar{\sigma}$ follow by symmetry replacing H^-, S and U^{-1} by H^+, $\bar{S}$ and U. Since $X \sim (S, \bar{S})$ is a minimal Markovian splitting subspace, S is U^{-1}–invariant and $X = \mathrm{E}^S H^+$. Obviously $S^{(-1)} := H^- \vee (U^{-1}S)$ is also U^{-1}–invariant and $S^{(-1)} \subset S$. Therefore

$$\sigma X = \mathrm{E}^{S^{(-1)}} \mathrm{E}^S H^+ = \mathrm{E}^{S^{(-1)}} H^+$$

is an observable Markovian splitting subspace. Since, in addition, $S^{(-1)} \subset S \perp H^+ \cap (H^-)^\perp$, σX is minimal [18, Theorem 4.10]. Since

$$\mathrm{E}^{S^{(-1)}} \mathrm{E}^S \lambda = \mathrm{E}^{S^{(-1)}} \lambda \quad \text{for each} \lambda \in H^+,$$

the splitting property (2.8) and the fact that $\|\mathrm{E}^{S^{(-1)}} \xi\| \leq \|\xi\|$ imply that $\sigma X \leq X$. Since X and σX are finite-dimensional and hence proper [17, 18], their ambient spaces are $\bigvee_{t=0}^{\infty} U^t S$ and $\bigvee_{t=0}^{\infty} U^t S^{(-1)}$ which must coincide in view of the fact that

$$U^{-1}S \subset S^{(-1)} \subset S.$$

In the same way we show that X and $\bar{\sigma} X$ have the same ambient space. This proves (i).

Next, we show that, if $\sigma X = X$, then $X \subset H_0$. Now, $\sigma X = X$ is equivalent to $X \subset H^- \vee U^{-1}S$, and hence to $UX \subset S \vee \{y(0)\}$. However, $S = X \oplus \bar{S}^\perp$ and $\{y(0)\} \subset H^+ \subset \bar{S} \perp \bar{S}^\perp$ so

$$UX \subset [X \vee \{y(0)\}] \oplus \bar{S}^\perp.$$

Since $UX \perp U\bar{S}^\perp \supset \bar{S}^\perp$ we have thus established that $\sigma X = X$ if and only

$$UX \subset X \vee \{y(0)\}, \tag{3.6}$$

which is the first part of (ii). A symmetric argument yields the second part.

Finally, to prove (iii), we note that (3.6) implies that

$$U\mathrm{E}^{H_0^\perp} X \subset \mathrm{E}^{H_0^\perp} X.$$

But the subspace $\mathrm{E}^{H_0^\perp} X$ is finite-dimensional. Since U is a bilateral shift, it has no eigenvalues [30] and hence cannot have a nontrivial finite-dimensional invariant subspace. Consequently we must have $X \subset H_0$. □

Corollary 3.2. *Let* $X \in \mathfrak{X}$ *and* $X \sim (S, \bar{S})$. *Then*

$$X^{(k)} = \begin{cases} \sigma^{-k} X & \text{for } k = 0, -1, -2, \dots \\ \bar{\sigma}^k X & \text{for } k = 0, 1, 2, \dots \end{cases} \tag{3.7}$$

defines a sequence $\{X^{(k)} \mid k \in \mathbb{Z}\}$ *of elements in* $\mathcal{X}$ *which have the same ambient space and which are totally ordered with* $X^{(0)} = X$. *More precisely,*

$$\cdots \leq X^{(-2)} \leq X^{(-1)} \leq X \leq X^{(1)} \leq X^{(2)} \leq \dots .$$

Moreover, for each $k \in \mathbb{Z}$, $X^{(k)} \sim (S^{(k)}, \bar{S}^{(k)})$ *where*

$$S^{(k)} = H^- \vee U^k S, \quad \bar{S}^{(k)} = H^+ \vee \left[S^{(k)}\right]^\perp \quad \textit{for } k \leq 0$$

and

$$\bar{S}^{(k)} = H^+ \vee U^k \bar{S}, \quad S^{(k)} = H^- \vee \left[\bar{S}^{(k)}\right]^\perp \quad \textit{for } k \geq 0.$$

Here the orthogonal complement $^\perp$ *is taken with respect to the common ambient space* $S \vee \bar{S}$.

Proof. This follows from the proof of Theorem 3.1, (2.10) and (2.11). In fact, it follows by induction that $\sigma^k X$ is also a minimal Markovian splitting subspace and that

$$\sigma^k X = \mathrm{E}^{S^{(-k)}} X \quad \text{for } k = 0, 1, 2, \dots,$$

where $S^{(-k)} := H^- \vee U^{-k} S$. The statement about $\bar{\sigma}$ follows by symmetry. □

Next we show that the sequence $\{X^{(k)} \mid k \in \mathbb{Z}\}$ can be extended to include limits $X^{(-\infty)}$ and $X^{(\infty)}$.

Theorem 3.3. *The limits* $\lim_{k\to-\infty} E^{S^{(k)}}\xi$ *and* $\lim_{k\to\infty} E^{\bar{S}^{(k)}}\xi$ *exist for all* $\xi \in X$ *and the spaces*

$$X^{(-\infty)} := \{\lim_{k\to-\infty} E^{S^{(k)}}\xi \mid \xi \in X\} \tag{3.8}$$

$$X^{(\infty)} := \{\lim_{k\to\infty} E^{\bar{S}^{(k)}}\xi \mid \xi \in X\} \tag{3.9}$$

are internal minimal Markovian splitting subspaces. Moreover, the sequences of splitting subspaces $\{X^{(-k)} \mid k = 0, 1, 2, \dots\}$ *and* $\{X^{(k)} \mid k = 0, 1, 2, \dots\}$ *converge in a finite number of steps if and only if* X *is internal. In that case the number of steps is no greater than* $\dim X$.

Proof. Since $\{S^{(-k)} \mid k = 0, 1, 2, \dots\}$ is a nonincreasing sequence of subspaces, i.e.,

$$S \supset S^{(-1)} \supset S^{(-2)} \supset S^{(-3)} \supset \dots, \tag{3.10}$$

it is well known [6, p. 24] that $\xi_{-\infty} = \lim_{k\to\infty} \mathrm{E}^{S^{(-k)}}\xi$ exists for all $\xi \in X$ and that $\xi_{-\infty} = \mathrm{E}^{S^{(-\infty)}}\xi$ where $S^{(-\infty)} = \bigcap_{k=0}^{\infty} S^{(-k)}$. Thus $X^{(-\infty)}$ is well-defined and, since $X = \mathrm{E}^S H^+$,

$$X^{(-\infty)} = \mathrm{E}^{S^{(-\infty)}} H^+.$$

Therefore, since $S^{(-\infty)}$ is U^*–invariant, $X^{(-\infty)}$ is an observable Markovian splitting subspace. But $S^{(-\infty)} \subset S \perp H^+ \cap (H^-)^\perp$, and hence $X^{(-\infty)}$ is minimal. It remains to show that $X^{(-\infty)}$ is internal. In view of Theorem 3.1, this would follow if $X^{(-\infty)}$ were a fixed point for σ. Next, we prove that this is in fact the case.

Consequently we want to prove that $\sigma X^{(-\infty)} = X^{(-\infty)}$, which follows from

$$H^- \vee U^{-1} S^{(-\infty)} = S^{(-\infty)}. \tag{3.11}$$

Let us prove (3.11). Since $S^{(-\infty)} = \bigcap_{k=0}^{\infty} S^{(-k)}$, $U^{-1}S^{(-k)} \subset S^{(-k)}$ and $H^- \subset S^{(-k)}$, it is trivial that

$$H^- \vee U^{-1} S^{(-\infty)} \subset S^{(-\infty)}.$$

It remains to prove the converse. To this end, note that

$$H^- \vee U^{-1} S^{(-k)} = \{y(-1)\} \vee U^{-1} S^{(-k)}.$$

This sum is in general not direct so we want to reformulate it into such a sum. Therefore, observe that $\{y(-1)\} \cap U^{-1}S^{(k)}$ is nonincreasing in k and finite-dimensional, and so there is a k_0 such that

$$\{y(-1)\} \cap U^{-1} S^{(k)} = \{y(-1)\} \cap U^{-1} S^{(k_0)} \quad \text{for } k \geq k_0.$$

Let V be a complement of $\{y(-1)\} \cap U^{-1}S^{(k_0)}$ in $\{y(-1)\}$. Then

$$H^- \vee U^{-1} S^{(-k)} = V + U^{-1} S^{(-k)} \tag{3.12}$$

is a direct sum. Now, if $\xi \in S^{(-\infty)} = \bigcap_{k=0}^{\infty}(H^- \vee U^{-1}S^{(-k)})$, then

$$\xi = v_k + \eta_k$$

with $v_k \in V$ and $\eta_k \in U^{-1}S^{(-k)}$ is a unique representation for each $k \geq k_0$. Therefore, since (3.12) is nonincreasing, $v_k = v \in V$ and $\eta_k = \eta \in \bigcap_{k=0}^{\infty} U^{-1}S^{(-k)} = U^{-1}S^{(-\infty)}$ for $k \geq k_0$. Hence

$$\xi = v + \eta \in V \vee U^{-1} S^{(-\infty)} \subset H^- \vee U^{-1} S^{(-\infty)},$$

proving that $X^{(-\infty)}$ is a fixed point.

Next, we assume that X is noninternal and prove that, in this case, all elements of the sequences $\{X^{(-k)} \mid k = 0, 1, 2, \dots\}$ and $\{X^{(k)} \mid k = 0, 1, 2, \dots\}$ are noninternal and that consequently these sequences cannot converge in a finite number of steps, the limits being internal. To see this, take a $\xi \in S$ such that $\xi \neq H_0$. Then $U^{-k}\xi \in S^{(-k)}$ but $U^{-k}\xi \notin H_0$, showing that $X^{(-k)}$ is noninternal for $k = 0, 1, 2, \dots$. A symmetric argument involving $\bar{S}$ shows that $X^{(k)}$ is also noninternal for $k = 0, 1, 2, \dots$.

To prove the converse, first recall that, for any internal $X \sim (S, \bar{S})$,

$$X = (X \cap X_-) \vee (X \cap X_+) \tag{3.13}$$

and that $S = H^- \vee X$. Relation (3.13) is proved in the same way as Lemma 2.9 in [13] and can also be found in [21]. Hence,

$$S = H^- \vee (X \cap X_+). \tag{3.14}$$

Since $X = \mathrm{E}^S H^+$, the subspace $X \cap X_+$ thus uniquely determines the internal splitting subspace X. In view of (3.14), (3.10) implies that the sequence $\{X^{(-k)} \cap X_+ \mid k = 0, 1, 2, \dots\}$ of finite-dimensional subspaces is nonincreasing. Therefore, it must converge in a finite number of steps which cannot be larger than $\dim X$, implying via (3.14) that the same holds for the sequence $\{X^{(-k)} \mid k = 0, 1, 2, \dots\}$. □

We remark that this proof also shows that, if X is internal, the whole sequence $\{X^{(k)} \mid k \in \mathbb{Z}\}$ cannot have more than $\dim X + 1$ different elements.

As pointed out in the end of the proof of Theorem 3.3, an internal X is completely characterized by its intersection $X \cap X_+ = X \cap H^+$ with the future via (3.14). In the same way,

$$\bar{S} = H^+ \vee (X \cap X_-), \tag{3.15}$$

so $X \in \mathcal{X}_0$ is also characterized by its intersection $X \cap X_- = X \cap H^-$ with the past. In particular, we have the following characterizations of σX_+ and $\bar{\sigma} X_-$.

Proposition 3.4. *The intersection of σX_+ with the past H^- is described by*

$$(\sigma X_+) \cap X_- = (X_+ \cap X_-) \vee (\{y(-1)\} \cap X_-) \tag{3.16}$$

and the intersection of $\bar{\sigma} X_-$ with the future H^+ is described by

$$(\bar{\sigma} X_-) \cap X_+ = (X_- \cap X_+) \vee (\{y(0)\} \cap X_+). \tag{3.17}$$

Proof. First observe that (3.16) is equivalent to

$$\bar{S}_+^{(-1)} = H^+ \vee (\{y(-1)\} \cap X_-). \tag{3.18}$$

In fact, that (3.18) implies (3.16) follows from the facts that

$$\bar{S}_+^{(-1)} \cap X_- = (\sigma X_+) \cap X_-$$

and

$$\left[H^+ \vee (\{y(-1)\} \cap X_-)\right] \cap X_- = (X_+ \cap X_-) \vee (\{y(-1)\} \cap X_-),$$

while the opposite implication follows from the fact that

$$\bar{S}_+^{(-1)} = H^+ \vee [(\sigma X_+) \cap X_-].$$

Next let us prove (3.18). We have

$$\begin{aligned}
\bar{S}_+^{(-1)} &= H^+ \vee (S_+^{(-1)})^\perp = H^+ \vee (H^- \vee U^{-1} S_+)^\perp \\
&= H^+ \vee \left[(H^-)^\perp \cap (U^{-1} H^+) \cap (U^{-1} H^-)^\perp\right] \\
&= H^+ \vee \left[(H^-)^\perp \cap (U^{-1} H^+)\right] \\
&= H^+ \vee \left[(H^-)^\perp \cap (H^+ \vee \{y(-1)\})\right]
\end{aligned}$$

where in the third step we have used the fact that $S_+^\perp = H^+ \cap (H^-)^\perp$. (See, for example, [18, Example 4.4].) Now suppose that $\xi \in (H^-)^\perp \cap (H^+ \vee \{y(-1)\})$. Then $\xi = \alpha + \beta$ where $\alpha \in H^+$ and $\beta \in \{y(-1)\} \subset H^-$. But $\alpha + \beta \perp H^-$, and consequently

$$\beta = -\mathrm{E}^{H^-} \alpha \in \mathrm{E}^{H^-} H^+ = X_-$$

so that $\beta \in \{y(-1)\} \cap X_-$. Hence $\bar{S}_+^{(-1)} \subset H^+ \vee (\{y(-1)\} \cap X_-)$.

Conversely, if $\beta \in \{y(-1)\} \cap X_-$, there is an $\alpha \in H^+$ such that $\beta = -\mathrm{E}^{H^-} \alpha$, which implies that $\alpha + \beta \perp H^-$ and that $\alpha + \beta \in (H^+ \vee \{y(-1)\})$. Hence

$$\{y(-1)\} \cap X_- \subset H^+ \vee \left[(H^-)^\perp \cap (H^+ \vee \{y(-1)\})\right]$$

which concludes the proof of (3.16) and (3.18).

The proof of the dual statement is completely symmetric, using the fact that

$$S_-^{(1)} = H^- \vee (\{y(0)\} \cap X_+) \tag{3.19}$$

is equivalent to (3.17). □

Since a minimal internal splitting subspace is completely characterized by its intersection with X_- via (3.15) and by its intersection with X_+ via (3.14), Proposition 3.4 has the following corollary, which we shall need later.

Corollary 3.5. *The splitting subspace X_+ is a fixed point of the operator σ if and only if $X_- \cap \{y(-1)\} = 0$. Likewise, X_- is a fixed point of the operator $\bar{\sigma}$ if and only if $X_+ \cap \{y(0)\} = 0$.*

The proof of Proposition 3.4 is easily modified to yield the following amplification, describing the chains of splitting subspaces $\{X_+^{(k)}\}$ and $\{X_-^{(k)}\}$.

Proposition 3.6. *For $k = 1, 2, 3, \dots,$ we have*

$$(\sigma^k X_+) \cap X_- = (X_+ \cap X_-) \vee (\{y(-1), \dots, y(-k)\} \cap X_-), \tag{3.20}$$

or, equivalently,

$$\bar{S}_+^{(-k)} = H^+ \vee (\{y(-1), \dots, y(-k)\} \cap X_-); \tag{3.21}$$

and

$$(\bar{\sigma}^k X_-) \cap X_+ = (X_- \cap X_+) \vee (\{y(0), \dots, y(k-1)\} \cap X_+), \tag{3.22}$$

or, equivalently,

$$S_-^{(k)} = H^- \vee (\{y(0), \dots, y(k-1)\} \cap X_+). \tag{3.23}$$

We shall now characterize the fixed points of the operators σ and $\bar{\sigma}$ in terms of the matrices D and $\bar{D}$ in equations (1.1) and (2.14), respectively.

Corollary 3.7. *Let $X \in \mathfrak{X}$ and let D and $\bar{D}$ be the corresponding matrices in the models* (1.1) *and* (2.14). *Then $\sigma X = X$ if and only if $\ker D' = 0$ and $\bar{\sigma} X = X$ if and only if $\ker \bar{D}' = 0$.*

Proof. Given (1.1) an elementary calculation yields

$$x(1) = Ax(0) + BD'(DD')^\sharp\, [y(0) - Cx(0)] + B_2 u(0),$$

where $B_2 := B - BD'(DD')^\sharp D$ and $(DD')^\sharp$ is a pseudoinverse of DD'. In particular, this implies that

$$\mathrm{E}\{B_2 u(0) y(0)'\} = BD' - BD'(DD')^\sharp DD' = 0.$$

Since therefore the components of $B_2 u(0)$ are orthogonal to both those of $x(0)$ and $y(0)$, (3.4) is equivalent to $B_2 = 0$, which in turn is equivalent to

$$\begin{bmatrix} B \\ D \end{bmatrix} [I - D'(DD')^\sharp D] = 0$$

But the columns of $\begin{bmatrix} B \\ D \end{bmatrix}$ are – according to our assumption – linearly independent so (3.4) is equivalent to $D'(DD')^\sharp D = I$, which holds if and only if DD' is full rank, i.e. if $(DD')^{-1}$ exists. Then the first statement follows from Theorem 3.1(ii). The second statement follows by symmetry. □

Remark 3.8. In view of Corollary 3.7 we have another proof of the fact that any fixed point of σ is internal. In fact, we established in the proof above that (3.4) is equivalent to $B_2 = 0$, which in the case when DD' is full rank implies that the transfer function (1.2) of (1.1) must be a square spectral factor and thus correspond to an internal realization [16, Theorem 5.2].

Theorem 3.1 and Corollary 3.7 give characterizations of precisely which internal X are fixed points of σ and $\bar{\sigma}$. It follows trivially from the definitions (3.1) and (3.2) that

$$\sigma X_- = X_- \quad \text{and} \quad \bar{\sigma} X_+ = X_+,$$

which, by Corollary 3.7, implies that D_- and $\bar{D}_+$ are always full rank, a well-known property of the innovations models. The following proposition together with Corollary 3.7 gives a more global picture on this question. (Also see [14].)

Proposition 3.9. *Let $X \in \mathfrak{X}$, and let D and $\bar{D}$ be the corresponding matrices in the models* (1.1) *and* (2.14). *Then*

$$\dim\ker D' = \dim(X \cap \{y(0)\}) \leq \dim(X_+ \cap \{y(0)\}) = \dim\ker D'_+ \tag{3.24}$$

and

$$\dim\ker \bar{D}' = \dim(X \cap \{y(-1)\}) \leq \dim(X_- \cap \{y(-1)\}) = \dim\ker \bar{D}'_-. \tag{3.25}$$

Proof. Let $a \in \ker D'$. Then $a'D = 0$ so that $a'y(0) \in X$. Conversely, suppose that $a'y(0) \in X$. Then, since $a'Cx(0) \in X$, we must have $a'Du(0) \in X \perp \{u(0)\}$, implying that $a'D = 0$. This proves the equalities in (3.24). To prove the inequality, note that

$$X \cap X_+ \cap \{y(0)\} = X \cap H^+ \cap \{y(0)\} = X \cap \{y(0)\}.$$

A symmetric argument yields (3.25). ☐

Corollary 3.10. *The splitting subspace X_+ is a fixed point of the operator σ if and only if $X_+ \cap \{y(0)\} = 0$. Likewise, X_- is a fixed point of the operator $\bar{\sigma}$ if and only if $X_- \cap \{y(-1)\} = 0$.*

Proof. In view of Corollary 3.7, this follows from the last equalities in (3.24) and (3.25) respectively. ☐

Comparing Corollaries 3.5 and 3.10, we can now see that the two conditions $X_+ \cap \{y(0)\} = 0$ and $X_- \cap \{y(-1)\} = 0$ are actually equivalent. We shall refer to the situation when they are satisfied as the *regular case*. From Proposition 3.9 and Corollary 3.7 it readily follows that, in the regular case, and only in the regular case, *all* $X \in \mathfrak{X}_0$ are fixed points of *both* σ and $\bar{\sigma}$. All this could also have been shown without using Corollary 3.5 by instead invoking the fact, proven in [14, Theorem 10.2], that D_+ has full rank if and only if $\bar{D}_-$ has.

The fact that $\sigma X_+ = X_+$ and $\bar{\sigma} X_- = X_-$ are the critical conditions in this analysis is also reflected in the ordering of covariances. In fact,

$$DD' = \Lambda_0 - CPC' \geq \Lambda_0 - CP_+C' = D_+D'_+$$

for all $P \leq P_+$ so that regularity is equivalent to

$$\Lambda_0 - CPC' > 0 \quad \text{for all } P \in \mathcal{P},$$

and analogous in the backward setting. This is also equivalent to all minimal spectral factors having zeros neither at zero nor at infinity.

We collect the regularity conditions in the following proposition. Some other characterizations can be found in [22, Theorem 3.2].

Proposition 3.11. *The following regularity conditions are equivalent.*

(i) $\Lambda_0 - CP_+C' > 0$
(ii) $\Lambda_0 - \bar{C}\bar{P}_-\bar{C}' > 0$
(iii) $X_+ \cap \{y(0)\} = 0$
(iv) $X_- \cap \{y(-1)\} = 0$
(v) $\sigma X_+ = X_+$
(vi) $\bar{\sigma} X_- = X_-$

Clearly regularity is a property of the output process y. Therefore, we introduce the following definition.

Definition 3.12. The process y is *regular* if the conditions of Proposition 3.11 are satisfied.

The regularity conditions can also be stated in terms of the whole family of minimal realizations.

Proposition 3.11$'$. *Each of the following regularity conditions is equivalent to those in Proposition 3.11.*

(i)$'$ $\Lambda_0 - CPC' > 0$ *for all* $P \in \mathcal{P}$
(ii)$'$ $\Lambda_0 - C\bar{P}C' > 0$ *for all* $\bar{P} \in \bar{\mathcal{P}}$
(iii)$'$ $X \cap \{y(0)\} = 0$ *for all* $X \in \mathcal{X}$
(iv)$'$ $X \cap \{y(-1)\} = 0$ *for all* $X \in \mathcal{X}$
(v)$'$ $\sigma X = X$ *for all* $X \in \mathcal{X}_0$
(vi)$'$ $\bar{\sigma} X = X$ *for all* $X \in \mathcal{X}_0$

We shall next prove that the operators σ and $\bar{\sigma}$ are invertible in the regular case and that $\bar{\sigma} = \sigma^{-1}$. In fact, as we shall see in Theorem 3.13 and Corollary 3.14 below, this property characterizes the regularity of the process y. In §6 we study the nonregular case and give a more complete description of the subspaces σX, $\bar{\sigma} X$ for any $X \in \mathcal{X}$.

In view of Corollary 3.4, a straight-forward calculation shows that

$$\bar{\sigma}\sigma X_+ = X_+ \quad \text{and} \quad \sigma\bar{\sigma} X_- = X_-. \tag{3.26}$$

A natural question is under what conditions these fixed point properties can be generalized to arbitrary $X \in \mathcal{X}$.

Theorem 3.13. *Let* $X \in \mathcal{X}$. *Then*

$$\sigma\bar{\sigma} X \le X \le \bar{\sigma}\sigma X, \tag{3.27}$$

and

$$\bar{\sigma}\sigma X = X \iff \{y(0)\} \cap X = \{y(0)\} \cap X_+. \tag{3.28}$$

Symmetrically,

$$\sigma\bar{\sigma} X = X \iff \{y(-1)\} \cap X = \{y(-1)\} \cap X_-. \tag{3.29}$$

Proof. We prove (3.29) and the first inequality in (3.27). Then, the rest follows by symmetry. First observe that, since $X \cap H^- = X \cap X_- \subset X_-$ and $\{y(-1)\} \subset H^-$, it always holds that

$$\{y(-1)\} \cap X \subset \{y(-1)\} \cap X_-.$$

In view of Corollary 3.2, $\bar{\sigma}X \sim (S^{(1)}, \bar{S}^{(1)})$ where

$$\begin{aligned}\bar{S}^{(1)} &= H^+ \vee U\bar{S}\\ S^{(1)} &= H^- \vee (\bar{S}^{(1)})^\perp = H^- \vee \left[(H^+)^\perp \cap (U\bar{S}^\perp)\right].\end{aligned}$$

Then apply σ to $\bar{\sigma}X$ to obtain

$$H^- \vee U^{-1}S^{(1)} = H^- \vee U^{-1}H^- \vee \left[(U^{-1}H^+)^\perp \cap \bar{S}^\perp\right].$$

Since $U^{-1}H^- \subset H^-$ and $U^{-1}H^+ = \{y(-1)\} \vee H^+$, we have

$$H^- \vee U^{-1}S^{(1)} = H^- \vee \left[\bar{S}^\perp \cap \{y(-1)\}^\perp\right] \subset S, \tag{3.30}$$

the last of which is a consequence of the condition $S = H^- \vee \bar{S}^\perp$. hence $\sigma\bar{\sigma}X \leq X$. To find a condition under which $\sigma\bar{\sigma}X = X$, we need to characterize the converse inequality. To this end, we consider the converse inclusion of (3.30) and take orthogonal complements in it to obtain

$$(H^-)^\perp \cap (\bar{S} \vee \{y(-1)\}) \subset (H^-)^\perp \cap \bar{S}. \tag{3.31}$$

Now, let ξ be an element in the subspace on the left side of (3.31). Then, $\xi = \alpha + \beta$, where $\alpha \in \bar{S}$ and $\beta \in \{y(-1)\} \subset H^-$, and $\xi \perp H^-$. Consequently,

$$\beta = -\mathrm{E}^{H^-}\alpha \in \mathrm{E}^{H^-}\bar{S} = X_-,$$

and hence $\beta \in \{y(-1)\} \cap X_-$. So, if $\{y(-1)\} \cap X = \{y(-1)\} \cap X_-$ holds, $\beta \in \{y(-1)\} \cap X \subset \bar{S}$. Therefore, since $\alpha \in \bar{S}$, we have $\beta \in \bar{S}$, and hence (3.31) holds.

Conversely, suppose that (3.31) holds. Consider a $\beta \in \{y(-1)\} \cap X_-$. Then, there is an $\alpha \in \bar{S}$ such that $\beta = -\mathrm{E}^{H^-}\alpha$ so that

$$\alpha + \beta \in (H^-)^\perp \cap (\bar{S} \vee \{y(-1)\}).$$

Using condition (3.31), we obtain $\alpha + \beta \in \bar{S}$. But $\alpha \in \bar{S}$, and hence $\beta \in \bar{S}$. In other words, (3.31) implies that $\{y(-1)\} \cap X_- \subset \bar{S}$, and consequently

$$\{y(-1)\} \cap X_- = \{y(-1)\} \cap X_- \cap \bar{S} \subset \{y(-1)\} \cap X,$$

since $X_- \cap \bar{S} = X_- \cap S \cap \bar{S} = X_- \cap X$. But the converse inclusion has already been proven above. Hence we have established (3.29). ☐

Corollary 3.14. *In the regular case the operators σ and $\bar{\sigma}$ are invertible and*

$$\bar{\sigma} = \sigma^{-1}. \tag{3.32}$$

Proof. This follows from regularity condition (iii) in Proposition 3.11 and (3.28) in Theorem 3.13. ☐

In particular Corollary 3.14 implies that relations (3.7) can be extended so that

$$X^{(j)} = \sigma^{j-k}X^{(k)} = \bar{\sigma}^{k-j}X^{(k)} \quad \text{for all } j, k \in \mathbb{Z}. \tag{3.33}$$

4. An interpolation problem. The ordered family of splitting subspaces introduced in §3 is intimately connected to the following estimation problem. Given a minimal stochastic system

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \tag{4.1}$$

of the type defined in §2 and integers t_0, t_1 such that $t_0 < t_1$, find, for each time t between t_0 and t_1, the linear least squares estimate[1]

$$\hat{x}(t \mid t_0, t_1) = \mathrm{E}\{x(t) \mid y(s), s \in \mathbb{Z}; x(\tau), \tau \in \mathbb{Z} \setminus \{t_0+1, \dots, t_1-1\}\} \tag{4.2}$$

of the state $x(t)$ given the whole output process y and the whole state process x except for times τ such that $t_0 < \tau < t_1$.

This interpolation problem is a prototype of an estimation problem of the following type. The state of a linear stochastic system is being observed both directly and through a noisy channel. During an interval of time (t_0, t_1) the direct state information is lost, and the problem is to estimate the lost states from the noisy observations and the remaining state information. Letting $t_0 \to -\infty$ and $t_1 \to \infty$, we obtain a *smoothing problem.* In a practical situation one would of course expect the information to be given on a finite interval containing $[t_0, t_1]$ and not on all of $\mathbb{Z}$ as here. However, as will be seen in Theorem 4.6 below, our solution will depend only on data from the interval $[t_0, t_1]$ and hence applies also to this situation, a remarkable fact that derives from the Markov property and allows us to use Kalman filtering. Nevertheless, it is convenient to formulate the problem in terms of infinite data.

In the more compact notation of §2, the interpolation estimate may be written

$$\hat{x}(t \mid t_0, t_1) = \mathrm{E}^{H_0 \vee (U^{t_0} X^-) \vee (U^{t_1} X^+)} x(t), \tag{4.3}$$

where X is the splitting subspace corresponding to (4.3), and X^- and X^+ are the past and future of X as defined after (2.6) . Now, one of the main results of this section is that the estimate (4.3) can be represented as a linear combination of the two estimates

$$x^{(t_0-t)}(t) = \mathrm{E}^{H_{t-1}^- \vee (U^{t_0} X^-)} x(t), \tag{4.4}$$

based on the past information, and

$$x^{(t_1-t)}(t) = \mathrm{E}^{H_t^+ \vee (U^{t_1} X^+)} x(t), \tag{4.5}$$

based on the future information. As we demonstrate below, this is due to the fact that $x^{(t_0-t)}$ and $x^{(t_1-t)}$ are state processes of minimal realizations of y, the splitting subspaces of which bound X from below and from above in the ordering defined in §2. In fact, $x^{(t_0-t)}(t) = U^t x^{(t_0-t)}(0)$ and $x^{(t_1-t)}(t) = U^t x^{(t_1-t)}(0)$, where

$$x^{(-k)}(0) = \mathrm{E}^{H^- \vee U^{-k} X^-} x(0) \tag{4.6}$$

and

$$x^{(k)}(0) = \mathrm{E}^{H^+ \vee U^k X^+} x(0) \tag{4.7}$$

are defined for $k = 0, 1, 2, \dots$. Obviously $x^{(0)} = x$ by both formulas. This relates the estimates (4.4) and (4.5) to the operators σ and $\bar{\sigma}$ defined in §3.

[1] Clearly $\mathrm{E}\{\cdot \mid \cdot\}$ denotes *wide sense* conditional expectation unless the system is assumed to be Gaussian.

Proposition 4.1. *The family of subspaces $\{X^{(k)} \mid k \in \mathbb{Z}\}$, defined in terms of (4.6) and (4.7) by*

$$X^{(k)} = \{a' x^{(k)}(0) \mid a \in \mathbb{R}^n\}, \tag{4.8}$$

is a family of minimal Markovian splitting subspaces such that

$$X^{(-k)} = \sigma^k X \quad \text{and} \quad X^{(k)} = \bar{\sigma}^k X \tag{4.9}$$

for $k = 0, 1, 2, \dots$, where $X = \{a'x(0) \mid a \in \mathbb{R}^n\}$ and $\sigma, \bar{\sigma}$ are the operators defined by (3.1) and (3.2). Moreover, for each $k \in \mathbb{Z}$, $x^{(k)}(0)$ is the basis in $X^{(k)}$ in the same uniform choice of coordinates as $x(0)$.

Proof. Let $X \sim (S, \bar{S})$. Then $S = H^- \vee X^-$ and $\bar{S} = H^+ \vee X^+$ and so, since $U^{-1}S \subset S$ and $U\bar{S} \subset \bar{S}$,

$$x^{(k)}(0) = \begin{cases} \mathrm{E}^{H^- \vee U^k S} x(0) = \mathrm{E}^{S^{(k)}} x(0) & \text{for } k \le 0 \\ \mathrm{E}^{H^+ \vee U^k \bar{S}} x(0) = \mathrm{E}^{\bar{S}^{(k)}} x(0) & \text{for } k \ge 0 \end{cases} \tag{4.10}$$

where $S^{(k)}$ and $\bar{S}^{(k)}$ are defined as in Corollary 3.2. Then the first statement is an immediate consequence of (4.10) and Corollary 3.2. Moreover, since $S^{(k)} \subset S$ for $k \le 0$,

$$x^{(k)}(0) = \mathrm{E}^{S^{(k)}} x(0) = \mathrm{E}^{S^{(k)}} \mathrm{E}^{S} x_+(0) = \mathrm{E}^{S^{(k)}} x_+(0)$$

for the appropriate choice of basis in X_+. Similarly, for $k \ge 0$, $\bar{S}^{(k)} \subset \bar{S}$ so that

$$\bar{x}^{(k)}(0) = \mathrm{E}^{\bar{S}^{(k)}} \bar{x}(0) = \mathrm{E}^{\bar{S}^{(k)}} \mathrm{E}^{\bar{S}} \bar{x}_-(0) = \mathrm{E}^{\bar{S}^{(k)}} \bar{x}_-(0),$$

where $\bar{x}_-(0) = P_-^{-1} x_-(0)$ and $x_-(0) = \mathrm{E}^{X_-} x_+(0)$. This proves the second statement. □

Consequently, we have established that

$$\hat{x}(0 \mid t_0, t_1) = \mathrm{E}^{S^{(t_0 - t)} \vee \bar{S}^{(t_1 - t)}} x(0), \tag{4.11}$$

where $X^{(t_0-t)} \sim (S^{(t_0-t)}, \bar{S}^{(t_0-t)})$ and $X^{(t_1-t)} \sim (S^{(t_1-t)}, \bar{S}^{(t_0-t)})$ are elements in $\mathfrak{X}$ such that

$$X^{(t_0-t)} \le X \le X^{(t_1-t)} \tag{4.12}$$

and such that $S^{(t_0-t)} \subset S$ and $\bar{S}^{(t_1-t)} \subset \bar{S}$. The following chain of lemmas deal with this setup and leads to the first main result of this section.

Lemma 4.2. *Let $X \sim (S, \bar{S})$, $X_1 \sim (S_1, \bar{S}_1)$ and $X_2 \sim (S_2, \bar{S}_2)$ be minimal Markovian splitting subspaces such that $S_1 \subset S$ and $\bar{S}_2 \subset \bar{S}$. Then $X_1 \le X \le X_2$ and*

$$x_1(0) = E^{X_1} x_2(0) \tag{4.13}$$

for any uniform choice $x_1(0), x_2(0)$ of bases in X_1 and X_2.

Proof. Since $\bar{S}_2 \subset \bar{S}$, we have

$$\bar{S}_2 \cap (H^-)^\perp \subset \bar{S} \cap (H^-)^\perp.$$

Since X and X_2 are minimal, this is equivalent to $S_2^\perp \subset S^\perp$ [18, Corollary 4.5], or, equivalently,

$$(\bar{S}_2 \vee S_2) \ominus S_2 \perp S. \tag{4.14}$$

In fact, $\bar{S}_2 \vee S_2$ is the ambient space of X_2. But (4.14) is equivalent to $\bar{S}_2 \perp S \mid S_2$ [18, Proposition 2.1] and also to

$$\mathrm{E}^{S}\lambda = \mathrm{E}^{S}\mathrm{E}^{S_2}\lambda \quad \text{for all } \lambda \in \bar{S}_2. \tag{4.15}$$

Apply E^{X_1} to this. Since $X_1 \subset S_1 \subset S$ and $H^+ \subset \bar{S}_2$,we obtain in particular

$$\mathrm{E}^{X_1}\lambda = \mathrm{E}^{X_1}\mathrm{E}^{S_2}\lambda \quad \text{for all } \lambda \in H^+.$$

But X_2 is a splitting subspace, so $\mathrm{E}^{S_2}\lambda = \mathrm{E}^{X_2}\lambda$ for all $\lambda \in H^+$ and $X_+ \subset H^+$. Consequently

$$\mathrm{E}^{X_1}\lambda = \mathrm{E}^{X_1}\mathrm{E}^{X_2}\lambda \quad \text{for all } \lambda \in X^+.$$

Hence, for an arbitrary choice of basis $x_+(0)$ in X_+, we have

$$\mathrm{E}^{X_1}x_+(0) = \mathrm{E}^{X_1}\mathrm{E}^{X_2}x_+(0),$$

which is equivalent to (4.13) with $x_1(0)$ and $x_2(0)$ being the corresponding bases in X_1 and X_2. The fact that $X_1 \leq X$ follows immediately from $\mathrm{E}^{S_1} = \mathrm{E}^{S_1}\mathrm{E}^{S}$ and [18, Lemma 6.7]. The relation $X \leq X_2$ follows analogously from $\bar{S}_2 \subset \bar{S}$. ◻

Lemma 4.3. *Let $X \sim (S, \bar{S})$, $X_1 \sim (S_1, \bar{S}_1)$ and $X_2 \sim (S_2, \bar{S}_2)$ be minimal Markovian splitting subspaces such that $S_1 \subset S$ and $\bar{S}_2 \subset \bar{S}$. Then*

$$E^{S_1 \vee \bar{S}_2} X \subset X_1 \vee X_2. \tag{4.16}$$

Proof. Applying the projection operator E^{S} to $H^+ \subset \bar{S}_2 \subset \bar{S}$ we obtain

$$\mathrm{E}^{S}H^+ \subset \mathrm{E}^{S}\bar{S}_2 \subset \mathrm{E}^{S}\bar{S} = X.$$

But, since X is observable, $X = \mathrm{E}^{S}H^+$, and therefore

$$X = \mathrm{E}^{S}\bar{S}_2. \tag{4.17}$$

Moreover, since $S_1 \subset S$, we have $\mathrm{E}^{S_1}H^+ = \mathrm{E}^{S_1}\mathrm{E}^{S}H^+$, and therefore, since X_1 and X are both observable,

$$X_1 = \mathrm{E}^{S_1}X, \tag{4.18}$$

which together with (4.17) yields

$$X_1 = \mathrm{E}^{S_1}\bar{S}_2. \tag{4.19}$$

Now, it is well-known and easy to check that the orthogonal decomposition

$$A = (\mathrm{E}^{A}B) \oplus (A \cap B^{\perp}) \tag{4.20}$$

holds for all pairs of subspaces A, B. Therefore, in view of (4.19), we have

$$S_1 = X_1 \oplus (S_1 \cap \bar{S}_2^{\perp}). \tag{4.21}$$

a completely symmetric argument yields

$$\bar{S}_2 = X_2 \oplus (\bar{S}_2 \cap S_1^{\perp}). \tag{4.22}$$

Therefore, since $X_1 \perp S_1^\perp$ and $X_2 \perp \bar{S}_2^\perp$, we have

$$S_1 \vee \bar{S}_2 = (S_1 \cap \bar{S}_2^\perp) \oplus (X_1 \vee X_2) \oplus (\bar{S}_2 \cap S_1^\perp). \tag{4.23}$$

To prove (4.16), take any $\xi \in X$. Then

$$\xi - \mathrm{E}^{S_1}\xi \perp S_1 \supset S_1 \cap \bar{S}_2^\perp$$

and, by (4.18) and (4.21),

$$\mathrm{E}^{S_1}\xi \in X_1 \perp S_1 \cap \bar{S}_2^\perp.$$

Consequently, $\xi \perp S_1 \cap \bar{S}_2^\perp$. In the same way we show that $\xi \perp \bar{S}_2 \cap S_1^\perp$, and therefore it follows from (4.23) that

$$\mathrm{E}^{S_1 \vee \bar{S}_2}\xi \in X_1 \vee X_2,$$

establishing (4.16). ◻

Lemma 4.4. *Let $X \sim (S, \bar{S})$, $X_1 \sim (S_1, \bar{S}_1)$ and $X_2 \sim (S_2, \bar{S}_2)$ be minimal Markovian splitting subspaces such that $S_1 \subset S$ and $\bar{S}_2 \subset \bar{S}$, and let $x(0)$, $x_1(0)$ and $x_2(0)$ be a uniform choice of bases in X, X_1 and X_2 with covariances P, P_1 and P_2, respectively. Then*

$$E^{X_1 \vee X_2} x(0) = (I - L)x_1(0) + Lx_2(0), \tag{4.24}$$

for any $n \times n$ matrix solution L of the linear system of equations

$$P - P_1 = L(P_2 - P_1). \tag{4.25}$$

Proof. Setting $\hat{x}(0) := \mathrm{E}^{X_1 \vee X_2} x(0)$, we have

$$\hat{x}(0) = Kx_1(0) + Lx_2(0) \tag{4.26}$$

for some $n \times n$ matrices K and L. By construction, $a'[x(0) - \hat{x}(0)] \perp X_1 \vee X_2$ for all $a \in \mathbb{R}^n$, which in particular implies that

(i) $a'[x(0) - \hat{x}(0)] \perp X_1$ for all $a \in \mathbb{R}^n$

(ii) $a'[x(0) - \hat{x}(0)] \perp X_2$ for all $a \in \mathbb{R}^n$.

Condition (i) together with (4.26) yields

$$\mathrm{E}\{x(0)x_1(0)'\} - KP_1 - L\mathrm{E}\{x_2(0)x_1(0)'\} = 0.$$

But, from Lemma 4.2 it follows that

$$\mathrm{E}\{x_2(0)x_1(0)'\} = P_1 \quad \text{and} \quad \mathrm{E}\{x(0)x_1(0)'\} = P_1,$$

and therefore, since P_1 is nonsingular,

$$K = I - L. \tag{4.27}$$

In the same way, Condition (ii) implies that

$$P = KP_1 + LP_2, \tag{4.28}$$

where again we have used Lemma 4.2 to see that

$$\mathrm{E}\{x(0)x_2(0)'\} = P \quad \text{and} \quad \mathrm{E}\{x_1(0)x_2(0)'\} = P_1.$$

Then (4.24) and (4.25) follow from (4.26)–(4.28).

To show that any solution L of (4.25) yields the same estimate $\hat{x}(0)$, let L_1 and L_2 be any two such solutions and let $\hat{x}_1(0)$ and $\hat{x}_2(0)$ be the corresponding estimates (4.26). Then

$$(L_1 - L_2)(P_2 - P_1) = 0 \tag{4.29}$$

and

$$\hat{x}_1(0) - \hat{x}_2(0) = (L_1 - L_2)\,[x_2(0) - x_1(0)]\,. \tag{4.30}$$

Equating the covariances of each sides in (4.30), equation (4.29) implies that $\hat{x}_1(0) = \hat{x}_2(0)$, as claimed. □

This immediately yields the following representation formula for the interpolation estimate.

Theorem 4.5. *Given the stochastic system* (4.1) *and* $t_0, t_1 \in \mathbb{Z}$ *such that* $t_0 < t_1$, *the state estimate*

$$\hat{x}(t \mid t_0, t_1) = E\{x(t) \mid y(s), s \in \mathbb{Z}; x(\tau), \tau \in (-\infty, t_0] \vee [t_1, \infty)\} \tag{4.31}$$

is given by

$$\hat{x}(t \mid t_0, t_1) = [I - L(t_0 - t, t_1 - t)]\, x^{(t_0 - t)}(t) + L(t_0 - t, t_1 - t) x^{(t_1 - t)}(t), \tag{4.32}$$

where $\{x^{(k)} \mid k \in \mathbb{Z}\}$ *is the estimation sequence* (4.6) – (4.7) *corresponding to* x *with covariances* $\{P^{(k)} \mid k \in \mathbb{Z}\}$ *and* $L(\tau, s)$ *is an arbitrary solution of*

$$P - P^{(\tau)} = L(\tau, s)\left[P^{(s)} - P^{(\tau)}\right]. \tag{4.33}$$

It remains to design a procedure for determining the estimation sequence $\{x^{(k)} \mid k \in \mathbb{Z}\}$. We shall address this question next. For this we need the following important consequence of the Markov property.

Theorem 4.6. *The state estimate* (4.31) *depends only on the data from the interval* $[t_0, t_1]$ *or, more precisely, on* $x(t_0)$, $x(t_1)$ *and* $y(t)$, $t = t_0, t_0 + 1, \dots, t_1$. *In particular,*

$$x^{(t_0 - t)}(t) := E\{x(t) \mid x(t_0), y(t_0), \dots, y(t-1)\} \quad \text{for } t > t_0 \tag{4.34}$$

$$x^{(t_1 - t)}(t) := E\{x(t) \mid y(t), \dots, y(t_1), x(t_1)\} \quad \text{for } t \le t_1\ , \tag{4.35}$$

where $\{x^{(k)}; k \in \mathbb{Z}\}$ *is the sequence of estimation processes defined by* (4.6) *and* (4.7).

Proof. Let $X \sim (S, \bar{S})$ be the splitting subspace corresponding to the state process x. In view of the definition (4.4), the first statement (4.34) is equivalent to

$$\mathrm{E}^{H^-_{k-1}(y) \vee X^-} \xi = \eta \quad \text{for all } \xi \in U^k X \text{ and } k \ge 0 \tag{4.36}$$

where

$$\eta := \mathrm{E}^{\{y(0), \dots, y(k-1)\} \vee X} \xi.$$

The original statement is obtained from (4.36) by merely applying the shift U^{t_0}. To prove (4.36) first note that, since $S = H^- \vee X^-$,

$$\begin{aligned} H^-_{k-1} \vee X^- &= \{y(0), \dots, y(k-1)\} \vee S \\ &= [\{y(0), \dots, y(k-1)\} \vee X] \oplus [S \ominus X]\,. \end{aligned}$$

To see this, note that $\{y(0), \dots, y(k-1)\} \subset H^+ \subset \bar{S} \perp S \ominus X$. Moreover, $\xi \in U^k X \subset U^k \bar{S} \subset \bar{S}$, and hence $\xi \perp S \ominus X$, which implies (4.36). A completely symmetric argument yields (4.35). □

Note that (4.34) and (4.35) are really forward and backward Kalman estimates initiated at $x(t_0)$ and $x(t_1)$ respectively, enabling us to use Kalman filtering techniques to generate them. Due to the fact that the initial conditions are states, these Kalman filters will have some remarkable properties, especially in the regular case when the reversibility condition (3.32) holds. This will be further discussed below.

The estimate (4.34) is generated by the recursion

$$\begin{cases} x^{(t_0-t)}(t) = Ax^{(t_0-t+1)}(t-1) + K^{(t_0-t)}\left[y(t-1) - Cx^{(t_0-t+1)}(t-1)\right] \\ x^{(0)}(t_0) = x(t_0) \end{cases} \tag{4.37}$$

where

$$K^{(-k)} = (\bar{C}' - AP^{(-k)}C')(\Lambda_0 - CP^{(-k)}C')^\sharp.$$

Here $^\sharp$ denotes pseudoinverse and the state covariance

$$P^{(-k)} = \mathrm{E}\{x^{(-k)}(0)x^{(-k)}(0)'\}$$

is given by the matrix Riccati equation

$$\begin{cases} P^{(-k-1)} = AP^{(-k)}A' + (\bar{C}' - AP^{(-k)}C')(\Lambda_0 - CP^{(-k)}C')^\sharp(\bar{C}' - AP^{(-k)}C')' \\ P^{(0)} = P \end{cases} \tag{4.38}$$

Note that this is the (invariant) formulation of the Kalman filter used in stochastic realization theory [1, 8, 16].

In the same way, the estimate (4.35) can be generated by a backward Kalman filter applied to the backward model

$$\begin{cases} \bar{x}(t-1) = A'\bar{x}(t) + \bar{B}\bar{u}(t-1) \\ y(t-1) = \bar{C}\bar{x}(t) + \bar{D}\bar{u}(t-1) \end{cases} \tag{4.39}$$

of X. Using a similar calculation as that in the forward direction, it is not hard to see that the process $\bar{x}^{(k)}(t) = [P^{(k)}]^{-1}x^{(k)}(t)$ is the solution of the backward Kalman filter

$$\begin{cases} \bar{x}^{(t_1-t)}(t) = A'\bar{x}^{(t_1-t-1)}(t+1) + \bar{K}^{(t_1-t-1)}\left[y(t) - \bar{C}\bar{x}^{(t_1-t-1)}(t)\right] \\ \bar{x}^{(0)}(t_1) = x(t_1) \end{cases} \tag{4.40}$$

where

$$\bar{K}^{(k)} = (C' - A'\bar{P}^{(k)}\bar{C}')(\Lambda_0 - \bar{C}\bar{P}^{(k)}\bar{C}')^\sharp$$

and the backward covariance matrix $\bar{P}^{(k)} = (P^{(k)})^{-1}$ is given by the matrix Riccati equation

$$\begin{cases} \bar{P}^{(k+1)} = A'\bar{P}^{(k)}A + (C' - A'\bar{P}^{(k)}\bar{C}')(\Lambda_0 - \bar{C}\bar{P}^{(k)}\bar{C}')^\sharp(C' - A'\bar{P}^{(k)}\bar{C}')' \\ \bar{P}^{(0)} = P^{-1}. \end{cases} \tag{4.41}$$

Then the process $x^{(t_1-t)}$ is given by

$$x^{(t_1-t)}(t) = (\bar{P}^{(t_1-t)})^{-1}\bar{x}^{(t_1-t)}(t),$$

defining $x^{(k)}$ for $k \geq 0$.

At least in the regular case, the inverse $(\Lambda_0 - CP^{(-k)}C')^{-1}$ will exist for all $k \in \mathbb{Z}$ (Proposition 3.11) and the pseudoinverses can be replaced with inverses. In the regular case we also have the reversibility property $\bar{\sigma} = \sigma^{-1}$ (Corollary 3.14) leading to (3.33). This useful property can be expressed in terms of estimation processes as

$$\mathrm{E}\{x(t) \mid H_{[t,t_1]}(y), x^{(t_0-t_1)}(t_1)\} = \mathrm{E}\{x(t) \mid H_{[t_0,t-1]}(y), x(t_0)\} = x^{(t_0-t)}(t), \quad (4.42)$$

i.e., tying together forward and backward estimation. This relation illustrates an important property of the Kalman recursions (4.37) and (4.40), namely that a consecutive application of forward and backward Kalman filtering brings us back through the same sequence of state processes of totally ordered stochastic realizations. This remarkable fact which is due to the invertibility of the operator σ, can also be justified by elementary calculations expressing $x^{(t_0-t+1)}(t-1)$ in terms of $x^{(t_0-t)}(t)$ and $y(t-1)$ in (4.37), leading to a backward Kalman filter which is an extension of (4.40) for negative $k = t_0 - t$. Similarly (4.40) can be reversed to give a forward Kalman filter identical to (4.37) for positive $k = t_1 - t$.

Given a stochastic realization (4.1) of y and a corresponding splitting subspace $X \sim (S, \bar{S})$, we have thus constructed a sequence of splitting subspaces $\{X^{(k)}; k \in \mathbb{Z}\}$ with bases

$$x^{(k)}(0) = \begin{cases} \mathrm{E}^{H^- \vee (U^k S)} x(0) & k \leq 0 \\ P^{(k)} P^{-1} \mathrm{E}^{H^+ \vee (U^k \bar{S})} x(0) & k \geq 0 \end{cases} \quad (4.43)$$

which are tied together by the Kalman filtering recursions (4.37) and (4.40). Each such basis vector defines a vector process

$$x^{(k)}(t) = U^k x^{(k)}(0)$$

which is the state process of a (forward) realization

$$\begin{cases} x^{(k)}(t+1) = Ax^{(k)}(t) + B^{(k)}u^{(k)}(t) \\ y(t) = Cx^{(k)}(t) + D^{(k)}u^{(k)}(t) \end{cases} \quad (4.44)$$

connected with a spectral factor

$$W^{(k)}(z) = C(zI - A)^{-1}B^{(k)} + D^{(k)}. \quad (4.45)$$

It is a manifestation of the fact that (4.43) is a uniform choice of bases for the splitting subspaces $\{X^{(k)} \mid k \in \mathbb{Z}\}$ in the sense defined in §2, and also easy to check, that the system matrices A and C remain constant for all $k \in \mathbb{Z}$ while $B^{(k)}$, $D^{(k)}$ and $P^{(k)}$ will vary. We shall not need to determine $\{B^{(k)}\}$ and $\{D^{(k)}\}$, but we note that this is easy to do either from the Riccati equation (4.37) or by means of a "fast algorithm" formulated directly in terms of $\{B^{(k)}, D^{(k)}\}$ as reported in Badawi [2].

Remark 4.7. Let us point out that the Riccati equation (4.38) can be written in the following form.

$$P^{(-k-1)} = P^{(-k)} - B^{(-k)}\{I - (D^{(-k)})'[D^{(-k)}(D^{(-k)})']^{\sharp}D^{(-k)}\}(B^{(-k)})'$$

The last term is nothing else than the covariance matrix of that part of the noise in the state space equation, i.e. of $B^{(-k)}u^{(-k)}$, which cannot be explained using the noise in the corresponding observation equation, i.e. via $D^{(-k)}u^{(-k)}$.

Similar statement can be formulated for the Riccati equation (4.41).

In the next section we show that, in the regular case, all spectral factors $\{W^{(k)} \mid k \in \mathbb{Z}\}$ have the same zeros, and in §8 we demonstrate that this is no longer the case in the nonregular case.

5. The zero structure of the estimation sequence in the regular case. Let us recall that $\lambda \in \mathbb{C}$ is an *(invariant) zero* of a spectral factor

$$W(z) = C(zI - A)^{-1}B + D$$

if there are row vectors a and b so that

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} = 0$$

or, in other words,

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \lambda a & 0 \end{bmatrix}. \tag{5.1}$$

Here a is called a *zero direction* (of order one) of W. In the regular case, when $DD' > 0$, we may eliminate b in these equations, to obtain

$$\begin{cases} a\Gamma = \lambda a \\ aB_2 = 0 \end{cases}$$

where

$$\Gamma := A - BD'(DD')^{-1}C$$
$$B_2 := B - BD'(DD')^{-1}D$$

showing that a is perpendicular to the reachability space

$$\langle \Gamma \mid B_2 \rangle = \operatorname{Im}(B_2, \Gamma B_2, \Gamma^2 B_2, \dots).$$

More generally, the zero directions (of any order) of W are defined using the Jordan structure of Γ. Then it can be proved that the orthogonal complement $\langle \Gamma \mid B_2 \rangle^{\perp}$ of this space in $\mathbb{R}^n$ is spanned by the *zero directions* of W. Hence, if Π is a matrix whose rows form a basis in $\langle \Gamma \mid B_2 \rangle^{\perp}$, i.e.,

$$\ker \Pi = \langle \Gamma \mid B_2 \rangle, \tag{5.2}$$

then there is a matrix Λ such that

$$\begin{cases} \Pi\Gamma = \Lambda\Pi \\ \Pi B_2 = 0. \end{cases} \tag{5.3}$$

Conversely, if Π is a matrix satisfying (5.3), then

$$\ker \Pi \supset \langle \Gamma \mid B_2 \rangle. \tag{5.4}$$

This fact can also be expressed in terms of a generalization of (5.1): Relation (5.4) is equivalent to the existence of matrices Λ and M so that

$$\begin{bmatrix}\Pi & -M\end{bmatrix}\begin{bmatrix}A & B\\ C & D\end{bmatrix} = \begin{bmatrix}\Lambda\Pi & 0\end{bmatrix}. \tag{5.5}$$

The row vectors of the maximal solution Π satisfying (5.2) (in the sense of having maximal rank) are the generalized zero directions, and the eigenvalues of the corresponding matrix Λ are of course precisely the finite zeros of W.

Remark 5.1. The matrix equation (5.5) is the appropriate generalization of (5.6) also in the nonregular case to be discussed in §§7 and 8; see [20]. Note, however, that W may have zeros at infinity in the nonregular case, so the eigenvalues of Λ corresponding to the maximal solution of (5.5) are here the *finite* zeros of W.

The following lemma enables us to characterize the zero directions in terms of a connection between the state x and the output y.

Lemma 5.2. *A matrix* Π *satisfies* (5.3) *if and only if there are matrices* Λ *and* M *such that*

$$\Pi x(t+1) = \Lambda\Pi x(t) + My(t) \tag{5.6}$$

We shall here give a proof which exhibits the connection between the Γ-matrix and the zero directions, and which works in the present regular case. In §8, we shall provide an alternative proof which also works in the nonregular case; in fact, even when the Γ-matrix cannot be defined.

Proof. As mentioned in the proof of Corollary 3.7, the state equation can be reformulated in the form

$$x(t+1) = \Gamma x(t) + BD'(DD')^{-1}y(t) + B_2u(t), \tag{5.7}$$

which, in the present regular case, is a unique decomposition of $x(t+1)$ in terms of $x(t)$, $y(t)$ and $B_2u(t)$. Hence

$$\Pi x(t+1) = \Pi\Gamma x(t) + \Pi BD'(DD')^{-1}y(t) + \Pi B_2u(t)$$

so that if Π satisfies (5.3) then (5.6) is also satisfied with $M = \Pi BD'(DD')^{-1}$. Conversely, if there are Λ and M so that (5.6) holds, then the uniqueness of decomposition (5.7) implies that (5.3) holds. □

Remark 5.3. Since Λ has no zero eigenvalues in the regular case, (5.6) may be written

$$\Pi P\bar{x}(t-1) = \Lambda^{-1}\Pi P\bar{x}(t) - \Lambda^{-1}My(t-1),$$

showing that the zeros of $\bar{W}(z^{-1})$ are precisely the eigenvalues of Λ^{-1}. Consequently, the forward and the backward models have the same zeros although the zero directions are transformed by the covariance matrix P. In fact, introducing the matrix

$$\bar{\Gamma} = A' - \bar{B}\bar{D}'(\bar{D}\bar{D}')^{-1}\bar{C},$$

the zeros of $\bar{W}$ are connected to the reciprocals of the eigenvalues of $\bar{\Gamma}$ in a manner analogous to (5.3).

Let us note the similarity between (5.6) and (3.4). In Theorem 3.1 we proved that (3.4) implies that $X \subset H_0$. In view of this, it is not surprising that (5.6) characterizes the subspace $X \cap H_0$. Recall that

$$X \cap H_0 = (X \cap X_-) \vee (X \cap X_+), \tag{5.8}$$

where the sum is direct if and only if $H^- \cap H^+ = 0$, i.e., if and only if

$$P_+ - P_- > 0 \tag{5.9}$$

[13, Lemma 2.9]. We note that $X \cap X_-$ is connected to the stable zeros of W (including the zeros on the unit circle) and that $X \cap X_+$ is connected to the antistable zeros (again including the zeros on the unit circle). If (5.9) holds, these sets of zeros are disjoint, there being no zeros on the unit circle.

As explained in [13], the subspaces $\ker(P - P_-)$ and $\ker(P_+ - P)$ are isomorphic to the subspaces $X \cap X_-$ and $X \cap X_+$ respectively under the bijection $a \mapsto a'x(0)$. Based on these observations it can be proved that the zeros of W form a subset of those of W_- and $\bar{W}_+$. Let us collect the statements about the zeros of W in the following theorem. Proofs can be found in [13, 19, 29].

Theorem 5.4. *The subspace* $\ker(P - P_-)$ *is invariant under* Γ'_- *and* Γ'. *Moreover,*

$$\Gamma'|_{\ker(P-P_-)} = \Gamma'_-|_{\ker(P-P_-)}. \tag{5.10}$$

The stable zeros of W *and* $\bar{W}$ *(including the ones on the unit circle) are the eigenvalues of* (5.10)*, and the corresponding zero directions of* W *span the subspace* $\ker(P - P_-)$. *Similarly,* $\ker(\bar{P} - \bar{P}_+)$ *is invariant under* $\bar{\Gamma}'_+$ *and* $\bar{\Gamma}'$. *Moreover,*

$$\bar{\Gamma}'|_{\ker(\bar{P}-\bar{P}_+)} = \bar{\Gamma}'_+|_{\ker(\bar{P}-\bar{P}_+)}. \tag{5.11}$$

The antistable zeros of W *and* $\bar{W}$ *(including the ones on the unit circle) are the reciprocals of the eigenvalues of* (5.11)*, and the corresponding zero directions of* $\bar{W}$ *span the subspace* $\ker(\bar{P} - \bar{P}_+)$.

Note that in the nonregular case, to be considered in §§7 and 8, the matrix Γ may not be well-defined for all X. Nevertheless all other statements of the theorem remain true.

To obtain coordinate-free versions of Γ' and $\bar{\Gamma}'$ we first observe that, in the regular case and with Π maximal so that $\ker \Pi = \langle \Gamma \mid B_2 \rangle$, (5.6) is equivalent to

$$U(X \cap H_0) \subset X \cap H_0 + \{y(0)\}, \tag{5.12}$$

where the sum is direct because of the regularity condition (iii)′ of Proposition 3.11. Similarly,

$$U^{-1}(X \cap H_0) \subset X \cap H_0 + \{y(-1)\}. \tag{5.13}$$

Now, following [13], let us introduce the *zero dynamics operators* in the regular case.

Definition 5.5 (regular case). Let the operators $G : X \cap H_0 \to X \cap H_0$ and $\bar{G} : X \cap H_0 \to X \cap H_0$ be defined as

$$G = \pi U|_{X \cap H_0} \tag{5.14}$$

and

$$\bar{G} = \bar{\pi} U^{-1}|_{X \cap H_0}, \tag{5.15}$$

where $\pi : (X \cap H_0) + \{y(0)\} \to X \cap H_0$ and $\bar{\pi} : (X \cap H_0) + \{y(-1)\} \to X \cap H_0$ are the oblique projectors projecting parallel to $\{y(0)\}$ and $\{y(-1)\}$ respectively.

In view of Definition 5.5, (5.10) and (5.11) may be written

$$G|_{X\cap X_-} = G_-|_{X\cap X_-}$$

and

$$\bar{G}|_{X\cap X_+} = \bar{G}_+|_{X\cap X_+}$$

respectively. Moreover, $X \cap X_-$ is invariant under both G and G_- and $X \cap X_+$ under both $\bar{G}$ and $\bar{G}_+$. In the nonregular case, the operators G and $\bar{G}$ may not be defined on all of $X \cap H_0$ but only on a subset of it, a circumstance manifested in the fact that Γ and $\bar{\Gamma}$ cannot be defined as above. However, G_- and $\bar{G}_+$ are always defined as in the regular case. This will be further discussed in §7.

Let us now return to the estimation sequence $\{x^{(k)} \mid k \in \mathbb{Z}\}$. The following theorem insures that no zeros are being lost when we move along the sequence $\{W^{(k)}\}$ from $k = 0$ through negative k.

Theorem 5.6. *If Π is a matrix of zero directions of $W^{(k)}$, it is also a matrix of zero directions for $W^{(k-j)}$ for $j = 0, 1, 2, \ldots$. Moreover, the zeros are preserved.*

Proof. Since Π is a zero direction of $W^{(k)}$, there is a matrix Λ such that

$$\Pi\Gamma^{(k)} = \Lambda\Pi,$$

and therefore, in view of (5.7),

$$\Pi x^{(k)}(t) - \Lambda\Pi x^{(k)}(t-1) - \Pi K^{(k)} y(t-1) = 0, \tag{5.16}$$

because $B^{(k)}(D^{(k)})'[D^{(k)}(D^{(k)})']^{-1} = K^{(k)}$. Consequently, by (4.37),

$$\begin{aligned}
&\Pi x^{(k-1)}(t+1) - \Lambda\Pi x^{(k-1)}(t) - \Pi K^{(k-1)} y(t) \\
&\quad = \Pi\Gamma^{(k)} x^{(k)}(t) - \Lambda\Pi\Gamma^{(k)} x^{(k)}(t-1) - \Lambda\Pi K^{(k)} y(t-1) \\
&\quad = \Lambda\left[\Pi x^{(k)}(t) - \Lambda K^{(k)}(t-1) - \Pi K^{(k)} y(t-1)\right]
\end{aligned}$$

which is zero by (5.16). This together with (5.16) establishes that not only the zero directions but also that the zeros are preserved, since the same matrix Λ can be used in each step. □

By symmetry we also have the following theorem.

Theorem 5.6′. *If $\bar{\Pi}$ is a matrix of zero directions of $W^{(k)}$, it is also a matrix of zero directions for $\bar{W}^{(k+j)}$ for $j = 0, 1, 2, \ldots$. Moreover, the zeros are preserved.*

We observe that $W^{(k)}$ and $\bar{W}^{(k)}$ have the same zeros in view of Remark 5.3. Theorems 5.6 and 5.6′ show that there is no loss of zeros when we apply a forward or backward Kalman filter step in (4.37) or (4.40). By the invertibility condition 3.32, all the elements in the sequence $\{W^{(k)} \mid k \in \mathbb{Z}\}$ must then have the *same* zeros. It is also easy to see that the zero directions are being preserved.

These results illustrate the fact that, in the regular case, all internal minimal splitting subspaces are fixed points of the operators σ and $\bar{\sigma}$. In fact, if X is internal, then so are σX and $\bar{\sigma} X$ by construction. Hence they have square spectral factors [16], which, by Theorems 5.6 and 5.6′, have the same zeros. Hence X, σX and $\bar{\sigma} X$ must be the same. This analysis and the fact that in general there may be $X \in \mathfrak{X}_0$ which are not fixed points, show that, in the nonregular case, the zeros may change as you move along the estimation sequence. The precise manner in which this happens is the topic of §8.

Remark 5.7. Note that Theorem 5.6 and 5.6$'$ imply that in the regular case the stable and unstable zero directions, i.e. the subspaces $\ker(P^{(-k)}-P_-)$, $\ker(P_+-P^{(-k)})$ and $\ker(\bar{P}^{(k)}-\bar{P}_+)$, $\ker(\bar{P}_+-\bar{P}^{(k)})$, remain unchanged as k tends to ∞ in the forward and backward Riccati equations (4.38) and (4.41). In other words, the solutions of the Riccati recursions remain constant in the zero directions, providing a possibility of reducing the size of the Riccati equation. In fact, choosing coordinates so that the last basis vectors span

$$\ker(P^{(-k)} - P_-) \vee \ker(P_+ - P^{(-k)}) \subset \mathbb{R}^n,$$

the matrices $\{P^{(-k)}\}$ in the solution of the Riccati recursion (4.38) take the form

$$P^{(-k)} = \begin{bmatrix} P_{11}^{(-k)} & P_{12} \\ P_{12}' & P_{22} \end{bmatrix}, \tag{5.17}$$

where only the upper left matrix block varies with k. Then, substituting (5.17) into (4.38) we obtain a reduced-order Riccati equation of dimension $\nu \times \nu$ where $\nu = n - \dim(X \cap H_0)$. A completely symmetric argument can be applied to the backward Riccati recursion (4.41).

6. Output-induced subspaces. We have just seen that the matrices Γ and $\bar{\Gamma}$ play an important role in the analysis of the estimation sequence $x^{(k)}$. We have also pointed out that they are only easily defined in the regular case. Therefore, in this section we shall only consider their coordinate-free versions, G and $\bar{G}$, which have natural definitions in the general case.

In the regular case, considered in §5, the zero dynamics operators G and $\bar{G}$, of a splitting subspace $X \in \mathcal{X}$, were defined on all of its internal subspace $X \cap H_0$. This is possible due to the direct sum decompositions (5.12) and (5.13). In the nonregular case these decompositions will fail to exist as we demonstrate in §7. Therefore, we must shrink the domains of the zero dynamics operators.

As demonstrated in [29], G can always be defined on $X \cap X_-$, yielding only the stable zeros (including those on the unit circle), and $\bar{G}$ can always be defined on $X \cap X_+$, producing only the antistable zeros (including those on the unit circle and those at infinity). In fact, this can also be seen from the following representations. (Also see [29, Lemma 5.1].)

Lemma 6.1. *Let $X \in \mathcal{X}$. Then*

$$U^{-1}(X \cap X_+) \subset (X \cap X_+) + \{y(-1)\} \tag{6.1}$$

and

$$U(X \cap X_-) \subset (X \cap X_-) + \{y(0)\}. \tag{6.2}$$

Proof. We prove (6.2). Then (6.1) follows by symmetry. Obviously,

$$U(X \cap X_-) \subset UX_- \subset H^- \vee \{y(0)\}.$$

Also $X \cap X_- \subset \bar{S} \cap \bar{S}_-$ which is U–invariant. Therefore,

$$U(X \cap X_-) \subset (H^- \vee \{y(0)\}) \vee (\bar{S} \cap \bar{S}_-).$$

But $\{y(0)\} \subset H^+ \subset \bar{S} \cap \bar{S}_-$ and $H^- \cap \bar{S} \cap \bar{S}_- = X \cap X_-$, implying (6.2). ◻

In this paper, however, we would like to define G and $\bar{G}$ on the largest possible spaces. We show that this can be done in such a way that the eigenvalues of G are precisely the finite zeros of X, and the eigenvalues of $\bar{G}$ are the reciprocals of the nonzero zeros of X (using the definition $1/\infty = 0$). Moreover, we want to know on which subspaces G and $\bar{G}$ are invertible so that they can be directly related to each other. This leads to the topic of *output-induced subspaces*, introduced in [13] in the in the continuous time setting. We now define it in the discrete-time case. Since, in the nonregular discrete-time case, the covariance matrix of the observation noise of the model (4.1) may be singular, the definition used in the continuous time case must be somewhat modified.

Definition 6.2. Let X be a Markovian splitting subspace. A subspace $Y \subset X$ is called *output-induced* if

(i) $Y \subset H_0$
(ii) $UY \subset Y \vee \{y(0), y(1), \dots, y(k)\}$ for some $k \geq 0$
(iii) $U^{-1}Y \subset Y \vee \{y(-1), y(-2), \dots, y(-k-1)\}$ for some $k \geq 0$.

We say that Y is *strictly output-induced* if it is output-induced and k can be chosen to be zero in (ii) and (iii).

The following proposition is an immediate consequence of the definition and the finite dimension of X.

Proposition 6.3. *The sum of two output-induced (strictly output-induced) subspaces is also output-induced (strictly output-induced). There exist a maximal output-induced (strictly output-induced) subspace in the sense of subspace inclusion.*

Since any output-induced subspace Y satisfies

$$Y \subset X \cap H_0 = (X \cap X_-) \vee (X \cap X_+),$$

let us first consider the subspaces $X \cap X_-$ and $X \cap X_+$. These, of course, trivially satisfy condition (i), and, by Lemma 6.1, they also satisfy one of the conditions (ii) and (iii) with $k = 0$, as required in the definition of being strictly output-induced. Next, we show that these subspaces also satisfy the remaining condition so that they are output-induced, and we investigate under what conditions they are actually strictly output-induced.

Theorem 6.4. *Let $X \in \mathfrak{X}$. Then the subspaces $X \cap X_+$ and $X \cap X_-$ are output-induced. Moreover, $X \cap X_+$ is strictly output-induced if and only if*

$$(\sigma X) \cap X_+ = X \cap X_+, \tag{6.3}$$

and $X \cap X_-$ is strictly output-induced if and only if

$$(\bar{\sigma} X) \cap X_- = X \cap X_-. \tag{6.4}$$

Proof. First we prove that $X \cap X_+$ is output-induced. To this end, in view of (6.1), it is enough to check that there exists a $k \leq \dim X$ such that

$$U(X \cap X_+) \subset (X \cap X_+) \vee \{y(0), y(1), \dots, y(k)\}. \tag{6.5}$$

Since

$$X \cap X_+ = X_{0-} \cap X_+$$

where X_{0-} is the tightest lower internal bound [18, 13], we may without loss of generality assume that X is internal. By Theorem 3.3, there is a $k \leq \dim X$ such that

$$\sigma^k X = \sigma^{k+1} X.$$

Consequently,

$$U^{-k}S \subset S^{(-k)} \subset S^{(-k-1)} = H^- \vee (U^{-(k+1)}S),$$

from which we have

$$US \subset S \vee U^{k+1}H^-.$$

Taking intersection with H^+ in both sides and noting that $U(S\cap H^+) \subset (US)\cap H^+$, we have

$$\begin{aligned} U(S\cap H^+) &\subset [S \vee \{y(0), \dots, y(k)\}] \cap H^+ \\ &= U(S\cap H^+) \vee \{y(0), \dots, y(k)\}. \end{aligned}$$

Then (6.5) follows from the fact that $X \cap X_+ = S \cap H^+$. In the same way, we prove that there is an $\ell \leq \dim X$ such that

$$U^{-1}(X\cap X_-) \subset (X \cap X_-) \vee \{y(-1), y(-2), \dots, y(-\ell-1)\}, \tag{6.6}$$

implying together with (6.2) that $X \cap X_-$ is output-induced.

To characterize the strictly output-induced property we prove that

$$(\sigma X) \cap X_+ = (X \cap X_+) \cap \left[\{y(-1)\} \vee U^{-1}(X \cap X_+)\right] \tag{6.7}$$

and that

$$(\bar\sigma X) \cap X_- = (X \cap X_-) \cap \left[\{y(0)\} \vee U(X \cap X_-)\right]. \tag{6.8}$$

To this end, let $X \sim (S, \bar S)$, and note that $S^{(-1)} := H^- \vee U^{-1}S \subset S$. Hence

$$\begin{aligned} (\sigma X) \cap X_+ &= S^{(-1)} \cap X_+ = S^{(-1)} \cap S \cap X_+ \\ &= S^{(-1)} \cap X \cap X_+ = \left[\{y(-1)\} \vee U^{-1}S\right] \cap X \cap X_+ \end{aligned}$$

But, since $U^{-1}S = U^{-1}X \oplus U^{-1}\bar S^\perp$ and $\{y(-1)\} \subset U^{-1}\bar S \perp U^{-1}\bar S^\perp$,

$$\begin{aligned} (\sigma X) \cap X_+ &= \left[(\{y(-1)\} \vee U^{-1}X) \oplus U^{-1}\bar S^\perp\right] \cap X \cap X_+ \\ &= (\{y(-1)\} \vee U^{-1}X) \cap X \cap X_+, \end{aligned}$$

because $X_+ \subset \bar S \subset U^{-1}\bar S \perp U^{-1}\bar S^\perp$. Moreover, if $\xi = (\{y(-1)\} \vee U^{-1}X) \cap X_+$, then $\xi = \alpha + \beta$ where $\alpha \in \{y(-1)\} \subset U^{-1}H^+$ and $\beta \in U^{-1}H^+$. Since $\xi \in H^+ \subset U^{-1}X$, we must have $\beta \in U^{-1}H^+$ so that $\beta \in U^{-1}(X \cap H^+) = U^{-1}(X \cap X_+)$. Therefore (6.7) follows. A symmetric argument yields (6.8).

Now, (6.7) and (6.8) immediately imply that

$$(\sigma X) \cap X_+ = X \cap X_+ \Longleftrightarrow U(X \cap X_+) \subset (X \cap X_+) \vee \{y(0)\} \tag{6.9}$$

and

$$(\bar\sigma X) \cap X_- = X \cap X_- \Longleftrightarrow U^{-1}(X \cap X_-) \subset (X \cap X_-) \vee \{y(-1)\} \tag{6.10}$$

concluding the proof. ∎

Corollary 6.5. *The subspace $X \cap H_0$ is the maximal output-induced subspace of $X \in \mathcal{X}$.*

Proof. This follows immediately from Theorem 6.4 and (5.8). □

Corollary 6.6. *The subspace $X_- \cap X_+$ is always strictly output-induced.*

Proof. This follows either from Lemma 6.1 or from (6.3) and the fact that $\sigma X_- = X_-$. □

We are now in a position to connect the concept strictly output-induced subspaces to fixed points of σ and $\bar{\sigma}$.

Corollary 6.7. *An $X \in \mathfrak{X}_0$ is a fixed point of σ if and only if $X \cap X_+$ is strictly output-induced. Likewise, $X \in \mathfrak{X}_0$ is a fixed point of $\bar{\sigma}$ if and only if $X \cap X_-$ is strictly output-induced.*

Proof. In the end of the proof of Theorem 3.3 we pointed out that the internal Markovian splitting subspaces are uniquely determined by $X \cap X_+$. Observe that, if $X \in \mathfrak{X}_0$, then $\sigma X \in \mathfrak{X}_0$. Consequently, Theorem 6.4 implies that $\sigma X = X$ if and only if $X \cap X_+$ is strictly output-induced. The rest follows by a symmetric argument. □

As we shall see in §8 these conditions can be formulated in terms of the stable and unstable zeros of the spectral factor (1.2) corresponding to the splitting subspace X.

The notion of strictly output induced subspaces enables us in some cases to characterize the limits $X^{(-\infty)}$ and $X^{(\infty)}$ of the sequence $\{X^{(k)} \mid k \in \mathbb{Z}\}$ defined in §3. To this end, let us recall [18] that the *tightest internal bounds,* X_{0-} and X_{0+}, are the closest internal X such that

$$X_{0-} \leq X \leq X_{0+}.$$

More precisely,

$$X_{0-} := \sup\{X_0 \in \mathfrak{X}_0 \mid X_0 \leq X\}$$

and

$$X_{0+} := \inf\{X_0 \in \mathfrak{X}_0 \mid X \leq X_0\}.$$

Corollary 6.8. *Let $X \in \mathfrak{X}$ and let X_{0-} and X_{0+} be its tightest internal bounds. Then*

$$X^{(-\infty)} = X_{0-}$$

if and only if $X \cap X_+$ is strictly output-induced, and

$$X^{(\infty)} = X_{0+}$$

if and only if $X \cap X_-$ is strictly output-induced.

Proof. Let us first recall that

$$S_{0-} = S \cap H_0 = H^- \vee (X \cap X_+).$$

(Cf. [18, Lemma 6.11] and [13].) Therefore Theorem 6.4 implies that X_{0-} is the lower tightest internal bound of σX also if and only if $X \cap X_+$ is strictly output-induced. By induction, we then have that $\sigma^{-k} X \geq X_{0-}$ for $k = 0, 1, 2, \dots$ and hence that $X^{(-\infty)} \geq X_{0-}$. But $X^{(-\infty)} \in \mathfrak{X}_0$ (Theorem 3.3), and consequently $X^{(-\infty)} = X_{0-}$ follows from the tightness of the bound. The proof for the upper bound is analogous. □

Another consequence of Theorem 6.4 is that the splitting subspaces in the sequence $\{X^{(k)} \mid k \in \mathbb{Z}\}$ have the same tightest local frame [18], if and only if the internal subspace $X \cap H_0$ is strictly output-induced. As we shall see in the next section, this is only true in the regular case.

Corollary 6.9. *A necessary and sufficient condition for all splitting subspaces in the family $\{X^{(k)} \mid k \in \mathbb{Z}\}$ to have the same tightest internal bounds is that $X \cap H_0$ is strictly output-induced.*

Proof. This follows immediately from Corollary 6.8, Proposition 6.3 and (5.8). □

Theorem 6.4 also yields the following alternative characterizations of regularity.

Corollary 6.10. *The following conditions are equivalent to the regularity conditions of Propositions 3.11 and 3.11$'$.*

(vii) X_+ *is strictly output-induced*
(viii) X_- *is strictly output-induced*
(vii)$'$ $X \cap X_+$ *is strictly output-induced for all* $X \in \mathcal{X}_0$
(viii)$'$ $X \cap X_-$ *is strictly output-induced for all* $X \in \mathcal{X}_0$
(ix)$'$ *All* $X \in \mathcal{X}_0$ *are strictly output-induced*
(x)$'$ *The internal subspace* $X \cap H_0$ *is strictly output-induced for all* $X \in \mathcal{X}$

Proof. By Corollary 6.7, (vii)$'$ and (viii)$'$ are equivalent to conditions (v)$'$ and (vi)$'$ of Proposition 3.11$'$, and (vii) and (viii) are equivalent to conditions (v) and (vi) of Proposition 3.11. In view of Proposition 6.3, (ix)$'$ follows from (vii)$'$, (viii)$'$ and (3.13), and (x)$'$ follows from (vii)$'$, (viii)$'$ and (5.8). Clearly either (vii) or (viii) imply (ix)$'$ and (x)$'$. □

7. Invariant directions and the maximal strictly output-induced subspace. Proposition 6.3 states that, to each $X \in \mathcal{X}$, there exists a maximal strictly output-induced subspace Y^*. In this section we construct Y^* explicitly. Let us recall that $Y \subset X \cap H_0$ is said to be strictly output-induced if

$$UY \subset Y \vee \{y(0)\} \tag{7.1}$$

and

$$U^{-1}Y \subset Y \vee \{y(-1)\}. \tag{7.2}$$

To determine Y^*, we first construct the subspaces $Y, \bar{Y} \subset X \cap H_0$ satisfying (7.1) and (7.2) respectively which are maximal in the sense of subspace inclusion and show that Y^* is precisely the intersection of these.

To this end, we design a procedure which is akin to the one used in geometric control theory [31] to construct the maximal output-nulling subspace. More precisely, define two sequences of subspaces $\{Y_0, Y_1, Y_2, \dots\}$ and $\{\bar{Y}_0, \bar{Y}_1, \bar{Y}_2, \dots\}$ by

$$Y_k = (\sigma^k X) \cap X \cap H_0 \tag{7.3}$$

and

$$\bar{Y}_k = (\bar{\sigma}^k X) \cap X \cap H_0 \tag{7.4}$$

and show that they converge monotonically to Y and $\bar{Y}$ respectively, in finitely many steps. As will be seen below these are precisely the largest spaces on which the zero dynamics operators may be defined. Obviously, $Y_0 = \bar{Y}_0 = X \cap H_0$. We now give alternative characterizations of these sequences and obtain iterative solutions of (7.1) and (7.2) respectively.

Lemma 7.1. *For each $k = 1, 2, 3, \dots$ the subspaces* (7.3) *and* (7.4) *can be written*

$$Y_k = \{\xi \in X \cap H_0 \mid U^k \xi \in (X \cap H_0) \vee \{y(0), \dots, y(k-1)\}\} \tag{7.5}$$

and

$$\bar{Y}_k = \{\xi \in X \cap H_0 \mid U^{-k} \xi \in (X \cap H_0) \vee \{y(-1), \dots, y(-k)\}\} \tag{7.6}$$

respectively. Moreover, the sequences $\{Y_k\}$ and $\{\bar{Y}_k\}$ satisfy the recursions

$$Y_{k+1} = \{\xi \in Y_k \mid U\xi \in Y_k \vee \{y(0)\}\} \tag{7.7}$$

and

$$\bar{Y}_{k+1} = \{\xi \in \bar{Y}_k \mid U^{-1}\xi \in \bar{Y}_k \vee \{y(-1)\}\} \tag{7.8}$$

for $k = 1, 2, 3, \dots$.

Proof. We prove only (7.5) and (7.7), (7.6) and (7.8) following by a symmetric argument.

To prove (7.5), observe that

$$\sigma^k X = \mathrm{E}^{(U^{-k}S) \vee H^-} X \subset U^{-k} X \vee \{y(-1), \dots, y(-k)\}, \tag{7.9}$$

in view of the decomposition

$$(U^{-k}S) \vee H^- = \left[(U^{-k}X) \vee \{y(-1), \dots, y(-k)\}\right] \oplus U^{-k}\bar{S}^\perp$$

and the fact that $U^{-k}\bar{S}^\perp \subset \bar{S}^\perp \perp X$.

Consequently, if $\xi \in Y_k$, then $\xi \in X \cap H_0$ and

$$U^k \xi \in [X \vee \{y(0), \dots, y(k-1)\}] \cap H_0 = (X \cap H_0) \vee \{y(0), \dots, y(k-1)\}.$$

Conversely, if $\xi \in X \cap H_0$ and $U^k \xi \in (X \cap H_0) \vee \{y(0), \dots, y(k-1)\}$, then

$$\mathrm{E}^{(U^{-k}S) \vee H^-} \xi = \mathrm{E}^{(U^{-k}X) \vee \{y(-1), \dots, y(-k)\}} \xi = \xi,$$

proving that $\xi \in \sigma^k X$ so that $\xi \in Y_k$.

Concerning the proof of (7.7), first consider a $\xi \in Y_k$ such that $U\xi \in Y_k \vee \{y(0)\}$. By (7.5) we have

$$\begin{aligned} U^{k+1}\xi = U^k(U\xi) &\in U^k Y_k \vee \{y(k)\} \\ &\subset (X \cap H_0) \vee \{y(0), \dots, y(k)\}, \end{aligned}$$

proving that $\xi \in Y_{k+1}$, as can be seen from (7.5).

Conversely, if $\xi \in Y_{k+1}$, then (7.5) implies that $U^{k+1}\xi$ has the representation

$$U^{k+1}\xi = \zeta + \lambda_0 + \lambda_1$$

where $\zeta \in X \cap H_0$, $\lambda_0 \in \{y(0), \dots, y(k-1)\}$ and $\lambda_1 \in \{y(k)\}$. We want to prove that $U\xi - U^{-k}\lambda_1 \in Y_k$, which implies (7.7). To this end, we note that

$$U\xi - U^{-k}\lambda_1 = U^{-k}\zeta + U^{-k}\lambda_0.$$

The left member of this belongs to $\bar{S}$, while the right member belongs to S, implying that they are in X and hence in $X \cap H_0$. Moreover, in view of (7.5), the identity

$$U^k(U\xi - U^{-k}\lambda_1) = \zeta + \lambda_0$$

implies that $U\xi - U^{-k}\lambda_1 \in Y_k$ concluding the proof of (7.7). ◻

An immediate consequence of Lemma 7.1 is that

$$X \cap H_0 = Y_0 \supset Y_1 \supset Y_2 \supset \dots \tag{7.10}$$

and that

$$UY_{k+1} \subset Y_k \vee \{y(0)\}. \tag{7.11}$$

Dually, we also have

$$X \cap H_0 = \bar{Y}_0 \supset \bar{Y}_1 \supset \bar{Y}_2 \supset \dots \tag{7.12}$$

and

$$U^{-1}\bar{Y}_{k+1} \subset \bar{Y}_k \vee \{y(-1)\}. \tag{7.13}$$

Since $X \cap H_0$ is finite-dimensional, the chain of inclusions (7.10) implies that there is a $k \leq \dim(X \cap H_0)$ such that $Y_{k+1} = Y_k$. Then (7.7) implies that $Y_\ell = Y_k$ for all $\ell \geq k$. Since $\dim(X \cap H_0) \leq \dim X := n$, we may refer to this subspace as Y_n. Clearly

$$UY_n \subset Y_n \vee \{y(0)\}$$

Similarly, $\bar{Y}_n$ is the limit of $\{\bar{Y}_k\}$ and satisfies

$$U^{-1}\bar{Y}_n \subset \bar{Y}_n \vee \{y(-1)\}.$$

Theorem 7.2. *The subspace Y_n is the maximal subspace of $X \cap H_0$ with the property*

$$UY \subset Y \vee \{y(0)\} \tag{7.14}$$

and $\bar{Y}_n$ is the maximal subspace in $X \cap H_0$ such that

$$U^{-1}\bar{Y} \subset \bar{Y} \vee \{y(-1)\}. \tag{7.15}$$

In the regular case, $Y_n = \bar{Y}_n = X \cap H_0$.

Proof. We have already proved that Y_n and $\bar{Y}_n$ satisfy (7.14)) and (7.15) respectively. To prove maximality, consider a $Y \subset X \cap H_0 = Y_0$ satisfying (7.14). We prove by induction that $Y \subset Y_k$ for $k = 0, 1, 2, \dots$. To this end, assume that $Y \subset Y_i$ and show that $Y \subset Y_{i+1}$. If $\xi \in Y$, then

$$U\xi \in Y \vee \{y(0)\} \subset Y_i \vee \{y(0)\}.$$

Consequently, in view of (7.7), $\xi \in Y_{i+1}$, as claimed. The maximality of $\bar{Y}_n$ is proved in the same way. The last statement follows from (5.12) and (5.13). ◻

Remark 7.3. Applying the orthogonal projection operator E^X to (7.14), we obtain

$$FY \subset Y \vee \mathrm{E}^X\{y(0)\}, \tag{7.16}$$

where F is the compressed shift operator $F := \mathrm{E}^X U|_X$. From the systems equations (1.1) one can infer that F has the matrix representations A' in the corresponding basis and that $\mathrm{E}^X\{y(0)\}$ has the matrix representations C'. Therefore, analogously to the continuous-time case [13], (7.16) is a stochastic version of (A', C')-invariance in geometric control theory [31, 4]. This connection to geometric control theory is elaborated upon in [29]. In this context, we note that a similar application of E^X to (7.11) yields

$$FY_{k+1} \subset Y_k \vee \mathrm{E}^X\{y(0)\},$$

which should be compared to the algorithm in geometric control theory to determine the maximal output-nulling subspace $\mathcal{V}^*$.

Now, referring back to the regular case and (5.12) and (5.13), we recall that, in this case, $X \cap H_0$ satisfies (7.14) and (7.15) with direct sum. This enabled us to define the operators G and $\bar{G}$. In the general case $X \cap H_0 \cap \{y(0)\}$ and $X \cap H_0 \cap \{y(-1)\}$ may be nontrivial subspaces. Nevertheless, as we will prove below, Y_n and $\bar{Y}_n$ satisfy (7.14) and (7.15) with direct sum decomposition in the right member. This requires a deeper analysis of so called *invariant directions* of a system representation (1.1) of X [5, 25, 26, 22].

More precisely, there are two kinds of invariant directions. An $a \in \mathbb{R}^n$ is a *predictable direction* if there is a positive integer k such that

$$a'x(0) \in \{y(-1), y(-2), \dots, y(-k)\}. \tag{7.17}$$

The smallest k with this property is called the *order* of the invariant direction a. If a satisfies (7.17), the Kalman filter estimate $\hat{x}$ takes the form

$$a'\hat{x}(t) = a'x(t) = \sum_{i=1}^{r} c_i' y(t-i)$$

in that direction so that the estimation error becomes zero. This manifests itself in that the filtering Riccati equation can be reduced in dimension after a finite number of steps. A similar reduction occurs in the fast filtering algorithm [11]; see in particular [12]. It can be shown [22] that a is a predictable direction if and only if, for some $k \geq 0$,

$$a \in \ker(\Gamma'_-)^k \cap \ker(P - P_-). \tag{7.18}$$

Dually, $a \in \mathbb{R}^n$ is a *smoothable direction* if there is a positive integer k such that

$$a'\bar{x}(0) \in \{y(0), y(1), \dots, y(k-1)\}, \tag{7.19}$$

causing a reduction in the backward Kalman filtering algorithms. Again the smallest k with this property is the order of the invariant direction a, and

$$a \in \ker(\bar{\Gamma}'_+)^k \cap \ker(P_+ - P), \tag{7.20}$$

for some k, is a necessary and sufficient condition condition for a to be a smoothable direction.

It can be seen from (7.18) and (7.20) that the order of an invariant direction cannot be larger than the dimension of X. Although the definition of invariant directions depends on the particular choice of coordinates in X, $a'x(0)$ and $a'\bar{x}(0)$ in the definitions (7.17) and (7.19) are independent of the coordinate system. Therefore we shall refer to these elements of X as the invariant directions of X.

Now, let H be the *frame space*

$$H \quad = X_- \vee X_+, \tag{7.21}$$

i.e., the closed linear hull of all internal subspaces $X \cap H_0$ as X ranges over $\mathfrak{X}$, and define the subspace

$$H_{0+} = H \quad \cap \{y(-n), \dots, y(n-1)\}.$$

In analogy with the continuous-time case [7], H_{0+} is called the *germ space* [22], since it contains all differences of y up to order n at $t = 0$.

Proposition 7.4. *The germ space has the direct sum decomposition*

$$H_{0+} = X_- \cap \{y(-1), \dots, y(-n)\} + X_+ \cap \{y(0), \dots, y(n-1)\}. \tag{7.22}$$

Moreover, X_- contains no smoothable and X_+ no predictable directions.

Proof. The inclusion $\supset$ is trivial. To prove the other direction, note that, since y is purely nondeterministic, the two terms in (7.22) has a zero intersection, and every $\xi \in H_{0+}$ has a unique representation $\xi = \xi_- + \xi_+$ such that $\xi_- \in \{y(-1), \dots, y(-n)\} \subset H^-$ and $\xi_+ \in \{y(0), \dots, y(n-1)\} \subset H^+$. But, in view of decomposition (4.20), ξ_- has an orthogonal decomposition $\xi_- = \hat{\xi}_- + \tilde{\xi}_-$ such that $\hat{\xi}_- \in X_-$ and $\tilde{\xi}_- \in H^- \cap (H^+)^\perp$ and ξ_+ can be written $\xi_+ = \hat{\xi}_+ + \tilde{\xi}_+$ where $\hat{\xi}_+ \in X_+$ and $\tilde{\xi}_+ \in H^+ \cap (H^-)^\perp$. Therefore, since

$$H_0 = \left[H^- \cap (H^+)^\perp\right] \oplus H \quad \oplus \left[H^+ \cap (H^-)^\perp\right],$$

the fact that $\xi = \tilde{\xi}_- + (\hat{\xi}_- + \hat{\xi}_+) + \tilde{\xi}_+ \in H$ shows that $\tilde{\xi}_- = \tilde{\xi}_+ = 0$. Hence $\xi_- \in X_-$ and $\xi_+ \in X_+$, establishing the inclusion $\subset$. □

Consequently, the germ space is spanned by the predictable invariant directions in X_- and the smoothable invariant directions in X_+. Moreover, y is regular if and only if it has a trivial germ space.

Proposition 7.5. *Let $X \in \mathfrak{X}$. Then*

$$X \cap H_{0+} = X \cap \{y(-1), \dots, y(-n)\} + X \cap \{y(0), \dots, y(n-1)\}, \tag{7.23}$$

i.e., $X \cap H_{0+}$ is spanned by the invariant directions of X. Moreover,

$$X \cap \{y(-1), \dots, y(-n)\} \subset X_- \cap \{y(-1), \dots, y(-n)\} \tag{7.24}$$

and

$$X \cap \{y(0), \dots, y(n-1)\} \subset X_+ \cap \{y(0), \dots, y(n-1)\}. \tag{7.25}$$

Proof. Let $X \sim (S, \bar{S})$. Relations (7.24) and (7.25) follow from the fact that $X \cap H^- = X \cap X_-$ and that $X \cap H^+ = X \cap X_+$ respectively. In view of this and Proposition 7.4, the inclusion $\supset$ in (7.23) is immediate. To prove $\subset$, take $\xi \in X \cap H_{0+}$. By Proposition 7.4, there is a unique decomposition $\xi = \xi_- + \xi_+$ such that $\xi_- \in X_- \cap \{y(-1), \dots, y(-n)\}$ and $\xi_+ \in X_+ \cap \{y(0), \dots, y(n-1)\}$. Hence it just remains to prove that $\xi_- \in X$ and $\xi_+ \in X$. To this end, note that $\xi_- \in H^- \subset S$ and that $\xi \in X \subset S$ so we must have $\xi_+ \in S$. But $\xi_+ \in H^+ \subset \bar{S}$, so $\xi_+ \in S \cap \bar{S} = X$. Since, $\xi \in X$, we must have $\xi_- \in X$ also. □

We have thus proved that all invariant directions of X_- are predictable and all the invariant directions of X_+ are smoothable, while an arbitrary X can have invariant directions of either kind. In view of (7.18), the predictable directions of X are also among the predictable directions of X_-. In the same way, (7.20) implies that the smoothable directions of X form a subspace of the smoothable directions of X_+ [22]. We call $X \cap \{y(-1), \dots, y(-n)\}$ the *predictable subspace* and $X \cap \{y(0), \dots, y(n-1)\}$ the *smoothable subspace* of X.

Theorem 7.6. *Let $X \in \mathfrak{X}$. Then,*

(i) The internal subspace $X \cap H_0$ of X has the direct-sum decomposition

$$X \cap H_0 = Y_n + X \cap \{y(0), \dots, y(n-1)\}. \tag{7.26}$$

Moreover,

$$X \cap X_- \subset Y_n. \tag{7.27}$$

In particular, Y_n contains the predictable directions $X \cap \{y(-1), \dots, y(-n)\}$ of X.

(ii) The internal subspace $X \cap H_0$ of X has the direct-sum decomposition

$$X \cap H_0 = \bar{Y}_n + X \cap \{y(-1), \dots, y(-n)\}. \tag{7.28}$$

Moreover,

$$X \cap X_+ \subset \bar{Y}_n. \tag{7.29}$$

In particular, $\bar{Y}_n$ contains the smoothable directions $X \cap \{y(0), \dots, y(n-1)\}$ of X.

(iii) The maximal strictly output-induced subspace of X is given by

$$Y^* = Y_n \cap \bar{Y}_n. \tag{7.30}$$

Moreover,

$$Y_n = Y^* + X \cap \{y(-1), \dots, y(-n)\} \tag{7.31}$$

and

$$\bar{Y}_n = Y^* + X \cap \{y(0), \dots, y(n-1)\}. \tag{7.32}$$

In the regular case, $Y^* = X \cap H_0$ for all $X \in \mathfrak{X}$.

This theorem, the proof of which we defer to the end of the section, shows, in particular, that the internal subspace $X \cap H_0$ can be decomposed as

$$X \cap H_0 = X \cap \{y(-1), \dots, y(-n)\} + Y^* + X \cap \{y(0), \dots, y(n-1)\}, \tag{7.33}$$

i.e., as the direct sum of the subspace of predictable directions, the maximal strictly output-induced subspace and the subspace of smoothable directions of X. In view of Proposition 7.5, $X \cap H_0$ is also the direct sum of Y^* and the germ subspace of X, i.e.,

$$X \cap H_0 = Y^* + X \cap H_{0+}. \tag{7.34}$$

This has the following consequence.

Corollary 7.7. *The process y is regular if and only if no $X \in \mathcal{X}$ has invariant directions.*
Remark 7.8. An immediate consequence of the definitions of Y_n, $\bar{Y}_n$ and Y^* is that

$$Y^* \subset Y_n \subset (\sigma^k X) \cap H_0 \quad \text{and} \quad Y^* \subset \bar{Y}_n \subset (\bar{\sigma}^k X) \cap H_0$$

for all k=0,1,2,. . . , showing that the maximal strictly output-induced subspace of any $X \in \mathcal{X}$ is contained in each of the internal subspaces of the corresponding sequence of splitting subspaces $\{X^{(k)}; k \in \mathbb{Z}\}$.

8. The change of zero dynamics under σ and $\bar{\sigma}$. In Definition **??** we assigned to each $X \in \mathcal{X}$ two operators G and $\bar{G}$, defined on the appropriate subspaces of X. Now, we will relate the eigenvalues of G and $\bar{G}$ to the zeros of W and $\bar{W}$, the spectral factors corresponding to X, justifying the name zero dynamics operators. Next we analyze the connections between the zero dynamics operators belonging to different splitting subspaces. Finally, using these operators, we describe completely the change in the zero structure when applying the prediction operators σ and $\bar{\sigma}$.

To this end, we recall from [20] that the finite zeros of W and the corresponding zero directions are characterized by the solutions of (5.5), i.e.,

$$\begin{bmatrix} \Pi & -M \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \Lambda\Pi & 0 \end{bmatrix}, \tag{8.1}$$

in the sense that the eigenvalues of Λ are zeros of W and the rows of Π span the subspace of the corresponding generalized zero directions. In order to describe all finite zeros we need to consider a maximal solution of (8.1) in the sense that Π has maximal rank, or in the sense that the subspace generated by the row vectors of Π is maximal.

We now give an alternative proof of a generalization of Lemma 5.2 which also works in the nonregular case.
Lemma 8.1. *A matrix Π satisfies* (8.1) *if and only if there are matrices Λ and M such that*

$$\Pi x(t+1) = \Lambda\Pi x(t) + My(t). \tag{8.2}$$

Proof. Equation (8.1) is equivalent to

$$\begin{bmatrix} \Pi & -M \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} \Lambda\Pi & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix},$$

where x is the state process and u is the driving noise of the stochastic model (1.1). This is seen by observing that the covariance matrix of $\begin{bmatrix} x(t) \\ u(t) \end{bmatrix}$ is full rank. Together with the systems equations (1.1) this yields (8.2). □

A coordinate-free version of (8.2) is provided by

$$UY \subset Y \vee \{y(0)\}, \tag{8.3}$$

where Y consists of the random variables of the form $b'\Pi x(0)$. This observation allows us to characterize the zeros of W in terms of the eigenvalues of G.
Proposition 8.2. *The eigenvalues of G are precisely the finite zeros of W. Similarly, the eigenvalues of $\bar{G}$ are the finite zeros of $\bar{W}(z^{-1})$.*

Proof. Consider the maximal solution of (8.1). Then the eigenvalues of Λ are the finite zeros of W. Moreover, since Y_n is the maximal subspace satisfying (8.3) (Theorem 7.2), $z := \Pi x(0)$ is a basis in Y_n, the space on which G is defined. It then follows from (8.2) that

$$Gz_i = \sum_j \Lambda_{ij} z_j,$$

and consequently Λ' is a matrix representation of G in the basis of z, thus having the same eigenvalues. This concludes the proof of the first statement. The second statement follows by symmetry. □

This together with Theorem **??** illustrates that we have zeros at zero and/or infinity precisely in the nonregular case and that such zeros are connected to invariant directions. More precisely, predictable directions correspond to zeros at zero and smoothable directions to zeros at infinity.

It was proved in [20] and in [29], independently and with different methods, that W and $\bar{W}$ have the same zeros also in the nonregular case. (A modification of the argument in Remark 5.3 could also be used to see this.) Therefore, any statement about the zeros of W also holds for $\bar{W}$ and vice versa.

Recall that

$$X \cap H_0 = (X \cap X_-) \vee (X \cap X_+), \tag{8.4}$$

where the sum is direct if and only if $X_- \cap X_+ = 0$, or, equivalently, $H^- \cap H^+ = 0$. Only in the regular case can we define G and $\bar{G}$ on all of $X \cap H_0$, but, in view of Theorem 7.6(i), $X \cap X_-$ is always contained in the domain of G and $X \cap X_+$ in the domain of $\bar{G}$.

Theorem 8.3. *Let G_- be the G-operator of X_- and $\bar{G}_+$ the $\bar{G}$-operator of X_+. Let $(W, \bar{W})$ be the spectral factors of $X \in \mathcal{X}$. Then*

$$G|_{X\cap X_-} = G_-|_{X\cap X_-} \tag{8.5}$$

and

$$\bar{G}|_{X\cap X_+} = \bar{G}_+|_{X\cap X_+}. \tag{8.6}$$

Consequently, the eigenvalues of $G|_{X\cap X_-}$ are the stable zeros of W (including those on the unit circle) and that the eigenvalues of $\bar{G}|_{X\cap X_+}$ are the reciprocals of the antistable zeros (including those on the unit circle and at infinity). Finally,

$$G|_{X_-\cap X_+} = \left[\bar{G}|_{X_-\cap X_+}\right]^{-1} \tag{8.7}$$

and its eigenvalues are precisely the zeros on the unit circle.

Proof. Referring to (6.2) we see that $X \cap X_-$ is invariant under both G and G_- and hence (8.5) follows. In the same way, (6.1) implies that $X \cap X_+$ is invariant under both $\bar{G}$ and $\bar{G}_+$ implying (8.6). This is in harmony with Theorem 5.4 and implies the statements on stable and unstable zeros. Finally, by Corollary 6.6, $X_- \cap X_+$ is strictly output-induced and is thus contained in Y^*, on which space G is invertible (Theorem 7.6). Consequently, the last statement follows. □

In particular, we have the following observation which has previously been reported in [10] and in [9]. (In the latter paper the proof is somewhat incomplete, since the multiplicities are not counted properly.)

Corollary 8.4. *All minimal spectral factors have the same number of zeros on the unit circle (counting multiplicity), namely* $\dim X_- \cap X_+ = \dim H^- \cap H^+$.

In §5 we showed that, in the regular case, the zeros, as well as the zero directions, are preserved as the operators σ and $\bar\sigma$ are applied. In general this is not true in the nonregular case. In view of Theorem 8.3, the following two theorems, relating σ and $\bar\sigma$ to the operators G and $\bar G$, show what happens.

Theorem 8.5. *Let* $X \in \mathfrak{X}$*. Then*

$$(\sigma X) \cap X_+ = \bar G_+(X \cap X_+) = \bar G(X \cap X_+) \tag{8.8}$$

and

$$(\bar\sigma X) \cap X_- = G_-(X \cap X_-) = G(X \cap X_-). \tag{8.9}$$

Proof. We prove only (8.8). Then a symmetric argument yields (8.9). First we show that

$$\sigma X \subset \{y(-1)\} \vee U^{-1}X. \tag{8.10}$$

To this end, observe that $H^- \vee U^{-1}S = \{y(-1)\} \vee U^{-1}S$, which in view of the fact that $\{y(-1)\} \subset U^{-1}H^+ \subset U^{-1}\bar S$, equals $(\{y(-1)\} \vee U^{-1}X) \oplus U^{-1}\bar S^\perp$. However, $X \perp \bar S^\perp \supset U^{-1}\bar S^\perp$, and consequently (8.10) follows from the definition (3.1). Now, consider $\zeta \in (\sigma X) \cap X_+$. In view of (8.10), we have the representation

$$\zeta = \eta + U^{-1}\xi,$$

where $\eta \in \{y(-1)\}$ and $\xi \in X$. On the other hand, since $U^{-1}\xi = \zeta - \eta \in U^{-1}H^+$, we see that $\xi \in H^+ \cap X = X \cap X_+$. From the definition of the operator $\bar G_+$, we have

$$\zeta = \bar G_+\xi \quad \text{where } \xi \in X \cap X_+.$$

Conversely, if $\xi \in X \cap X_+$, $\bar G_+\xi \in X_+$ and $\bar G_+\xi - U^{-1}\xi \in \{y(-1)\}$ implying that $\bar G_+\xi \in H^- \vee U^{-1}S = S^{(-1)}$. Hence,

$$\bar G_+\xi \in (\sigma X) \cap X_+,$$

which together with (8.6) concludes the proof of the theorem. □

Theorem 8.6. *Let* $X \in \mathfrak{X}$*. Then*

$$(\sigma X) \cap X_- = \{\xi \in X_- \mid G_-\xi \in X \cap X_-\} \tag{8.11}$$

and

$$(\bar\sigma X) \cap X_+ = \{\xi \in X_+ \mid \bar G_+\xi \in X \cap X_+\}. \tag{8.12}$$

Proof. We prove (8.11); then (8.12) follows by symmetry. Let $\xi \in X_-$. Since $\mathrm{E}^{H-}X = X_-$ [18, Lemma 4.6 and Theorem 4.10] and $\ker \mathrm{E}^{H-}|_X = X \cap (H^-)^\perp = 0$ (see §2), there is a unique $\zeta \in X$ such that $\xi = \mathrm{E}^{H-}\zeta$. By the definition of G_-,

$$U\xi = G_-\xi + \eta,$$

where $\eta \in \{y(0)\}$, and therefore

$$U(\zeta - \xi) = U\zeta - \eta - G_-\xi.$$

Since $G_-\xi \in X_- \subset S$, $U\zeta \in U\bar{S} \subset \bar{S}$ and $\eta \in H^+ \subset \bar{S}$, the splitting property (2.8) yields

$$\mathrm{E}^S U(\zeta - \xi) = \mathrm{E}^X (U\zeta - \eta) - G_-\xi. \tag{8.13}$$

Now, suppose $\xi \in (\sigma X) \cap X_-$. Then by definition (3.1), $\xi = \mathrm{E}^{H^- \vee U^{-1}S}\lambda$ for some $\lambda \in X$. But then, since $\xi \in H^-$, $\xi = \mathrm{E}^{H^-}\lambda$, so by uniqueness we must have $\lambda = \zeta$. Consequently, $\zeta - \xi \perp H^- \vee U^{-1}S$, which in particular implies that $U(\zeta - \xi) \perp S$. Hence, it follows from (8.13) that $G_-\xi \in X$, proving that $G_-\xi \in X \cap X_-$.

Conversely, suppose that $G_-\xi \in X \cap X_-$. Then, by (8.13),

$$\mathrm{E}^S U(\zeta - \xi) \in X. \tag{8.14}$$

But, since $\xi = \mathrm{E}^{H^-}\zeta$,

$$U(\zeta - \xi) \perp UH^- \supset H^-. \tag{8.15}$$

Therefore, since $S = H^- \oplus S \cap (H^-)^\perp$ by (4.20), we have

$$\mathrm{E}^S U(\zeta - \xi) = \mathrm{E}^{S \cap (H^-)^\perp} U(\zeta - \xi) \in S \cap (H^-)^\perp. \tag{8.16}$$

Since $X \cap S \cap (H^-)^\perp = X \cap (H^-)^\perp = 0$ (see §2), it follows from (8.14) and (8.16) that $\mathrm{E}^S U(\zeta - \xi) = 0$, and hence

$$U(\zeta - \xi) \perp S,$$

which together with (8.15) yields

$$\zeta - \xi \perp H^- \vee U^{-1}S.$$

Consequently, $\xi = \mathrm{E}^{H^- \vee U^{-1}S}\zeta \in \sigma X$ and so $\xi \in (\sigma X) \cap X_-$ as claimed. □

In particular, Theorems 8.5 and 8.6 show that

$$(\bar{\sigma} X) \cap X_- \subset X \cap X_- \subset (\sigma X) \cap X_-, \tag{8.17}$$

i.e., stable zeros may be lost as we apply $\bar{\sigma}$ and gained as we apply σ. In the same way,

$$(\sigma X) \cap X_+ \subset X \cap X_+ \subset (\bar{\sigma} X) \cap X_+, \tag{8.18}$$

showing that antistable zeros may be lost when applying σ and gained when applying $\bar{\sigma}$. This is in agreement with Theorem 3.6 and formulae (6.7) and (6.8).

To determine what zeros are being lost and gained under these operations, we observe from Theorems 8.5 and 8.6 that the subspaces being added or subtracted from $X \cap X_-$ and $X \cap X_+$ must be contained in the kernel of some G- or $\bar{G}$-operator. Consequently, by Theorem **??**, the corresponding zero directions are invariant directions.

We may therefore formulate an amplification of statement (8.17), namely that zeros at zero together with the corresponding predictable directions may be gained when applying σ and lost when applying $\bar{\sigma}$. In the same way, (8.18) and Theorem **??** show that zeros at infinity together with the corresponding smoothable directions may be lost when applying σ and gained when applying $\bar{\sigma}$.

The following corollary is an immediate consequence of Theorems 8.5 and 8.6.

Corollary 8.7. *Let $X \in \mathfrak{X}$ and let Y_n and $\bar{Y}_n$ be defined as in §7. Then*

$$\begin{aligned}(\sigma^k X) \cap H_0 &= Y_n \vee \{\textit{predictable directions in } X_-\}\\ &= Y^* + \{\textit{predictable directions in } X_-\}\end{aligned}$$

and

$$\begin{aligned}(\bar{\sigma}^k X) \cap H_0 &= \bar{Y}_n \vee \{\textit{smoothable directions in } X_+\}\\ &= Y^* + \{\textit{smoothable directions in } X_+\}\end{aligned}$$

for $k = n, n+1, \dots$.

Remark 8.8. Theorem **??**, Remark 7.8 and Corollary 8.7 enables us to generalize the statement in Remark 5.7 to the nonregular case. The same construction that was used in this remark to reduce the Riccati equations can be applied here with modifications which take into account that the fact that the internal subspace $X^{(k)} \cap H_0$ is no longer constant along the sequence of splitting subspaces $\{X^{(k)}\}$ in the nonregular case. In view of Remark 7.8, the solutions of the Riccati recursions are constant from the start in the zero directions of Y_n, while they become constant only after a finite number of steps in the remaining predictable directions by Corollary 8.7. Consequently, after a finite number of steps the size of the reduced Riccati equations is $\nu \times \nu$ where $\nu = n - \dim(\sigma^n X) \cap H_0$. In view of Corollary 8.7 and Theorem **??**, the backward Riccati equation can be reduced to the same size.

9. Conclusions. In this paper we discuss the very rich and intricate structure of discrete-time linear stochastic systems in the context of an interpolation type problem, namely to reconstruct lost state information on a finite interval using the whole history of the output process and the remaining state information. We show that, at each time, the (least-squares) state estimate can be written as a linear combination of two filter estimates, which are generated by (forward respectively backward) Kalman filtering type recursions with the initial condition being itself a state. Remarkably, these Kalman filtering recursions generate sequences of state processes from different stochastic realizations which are totally ordered. When $k \to \infty$ and when $k \to -\infty$, the sequence of splitting subspaces $X^{(k)}$ converge to limits which are internal splitting subspaces. These limits are determined by the zero structure of the spectral factor (1.2).

In the regular case, when there are no zeros at the origin and at infinity, the set of zeros and zero directions of the spectral factors $W^{(k)}$ corresponding to the splitting subspaces $X^{(k)}$ remain invariant during these recursions giving a full set of invariants. We show that in the nonregular case the whole set of zeros is no longer invariant but the finite zeros with finite reciprocals still are.

We have shown that the computational burden of determining the interpolation estimate depends on the dimension of the internal subspace $X \cap H_0$, i.e., on the number of zeros. This leads to the study of output-induced and strictly output-induced subspaces and zero dynamics operators. In particular, if $a'x(0) \in X \cap H_0$, then, in the regular case, the solutions of the Riccati equations (4.38) and (4.41) for the interpolation problem becomes constant in the direction a, allowing for a reduction in size of the Riccati equations. In fact, if $\dim X \cap H_0 = n - \nu$, we only need to solve Riccati equations of dimension $\nu \times \nu$ rather than $n \times n$ in the regular case (Remark 5.7). In the nonregular case the reduction may be even larger after a finite number of steps (Remark 8.8).

What makes the discrete-time case more complicated than the continuous-time case is the possibility that the *predictable subspace* $X \cap \{y(-1), \dots, y(-n)\}$ and the *smoothable subspace* $X \cap \{y(0), \dots, y(n-1)\}$ are nontrivial. In fact, if these spaces are zero spaces (the regular case), the structure of the problem is very much like the continuous-time coercive case, studied in [13], and $X \cap H_0$ is itself strictly output-induced. If they are not, the matrix D will loose rank, and the matrix Riccati equations of forward and backward Kalman filtering will become constant in the directions a for which $a'x(0)$ is an element of these spaces, thus influencing the implementation of the filtering algorithms, as explained above. These a are called *invariant directions.* Nonregularity, and hence invariant directions, are connected with zeros at zero and at infinity.

In particular, we have demonstrated that $X \cap H_0$ can be decomposed as a direct sum of the predictable subspace, the smoothable subspace and the maximal strictly output-induced subspace, corresponding to the zeros at zero, the zeros at infinity, and the remaining zeros respectively. The maximal strictly output-induced subspace Y^* equals $X \cap H_0$ in the regular case and plays the role of $X \cap H_0$ in the nonregular case. We have given several geometric characterizations of regularity (Propositions 3.11 and 3.11′ and Corollaries 6.10 and 7.7). We have also shown that Y^* can be determined by algorithms akin to that used in geometric control theory for determining the maximal output-nulling subspace.

On the maximal strictly output-induced subspace Y^* the forward and backward zero dynamics operators G and $\bar{G}$, respectively, are inverses of each other. The eigenvalues are the finite zeros with finite reciprocals. The operators G and $\bar{G}$ can be separately extended to a larger subspace. On these subspaces (in the nonregular case) these operators are in general singular and the invariant directions determine the kernel of these operators.

Appendix A. Proof of Theorem ??.

Let us denote by $I_p(X)$ the predictable directions in $X \in \mathfrak{X}$ under the natural isomorphism $a \mapsto a'x(0)$ and by $I_s(X)$ the smoothable directions under the same isomorphism. Then Theorem **??** implies that

$$I_p(X) = \ker(P - P_-) \cap \ker(\Gamma'_-)^n \quad \text{and} \quad I_s(X) = \ker(P_+ - P) \cap \ker P_+^{-1}(\bar{\Gamma}'_+)^n P_+.$$

In particular,

$$I_p(X_-) = \ker(\Gamma'_-)^n \quad \text{and} \quad I_s(X_+) = \ker P_+^{-1}(\bar{\Gamma}'_+)^n P_+.$$

As in [28, p. 53], a straight-forward but somewhat tedious calculation yields the identity

$$\bar{\Gamma}_+ P_+^{-1}(P_+ - P_-) = P_+^{-1}(P_+ - P_-)\Gamma'_-. \tag{A.1}$$

First, we prove that

$$\dim I_p(X_-) = \dim I_s(X_+).$$

To this end, observe that (A.1) implies that

$$(P_+ \bar{\Gamma}_+^n P_+^{-1})(P_+ - P_-) = (P_+ - P_-)(\Gamma'_-)^n. \tag{A.2}$$

Since it follows from Theorem 8.3 that $\ker(P_+ - P_-) \cap \ker \Gamma'_- = 0$, we obtain that if $\xi \in \ker(\Gamma'_-)^n$ is nonzero then $(P_+ - P_-)\xi \neq 0$. Thus

$$(P_+ - P_-) I_p(X_-) \subset \ker P_+^{-1}(\bar{\Gamma}_+)^n P_+,$$

implying that

$$\dim I_p(X_-) \leq \dim I_s(X_+).$$

A symmetric argument yields the reverse inequality proving the first statement in the theorem and also that

$$\ker P_+^{-1}(\bar{\Gamma}_+)^n P_+ = (P_+ - P_-)I_p(X_-).$$

Consequently, $Y := \operatorname{Im} P_+^{-1}(\bar{\Gamma}'_+)^n P_+$, the counterpart of the maximal strictly output-induced subspace Y^* of X_+ under the natural isomorphism, is the orthogonal complement of $(P_+ - P_-)I_p(X_-)$, i.e.,

$$Y = \{a \in \mathbb{R}^n \mid a'(P_+ - P_-)b = 0 \quad \text{for} \quad b \in I_p(X_-)\}.$$

Thus, again invoking that $(P_+ - P_-)$ is nonsingular on $I_p(X_-)$, we obtain the direct sum decomposition

$$I_p(X_-) + Y = \mathbb{R}^n,$$

where the two summands are "orthogonal" in the inner product defined by $(P_+ - P_-)$. In view of the direct sum decomposition

$$I_s(X_+) + Y = \mathbb{R}^n \tag{A.3}$$

we see that both the predictable directions and the smoothable directions under the natural isomorphisms are mapped to subspaces which are complementary to Y.

Now observe that if $a \in \ker(P - P_-)$ and $b \in \ker(P_+ - P)$ then $a'(P_+ - P_-)b = 0$, i.e. $\ker(P - P_-)$ and $\ker(P_+ - P)$ are also orthogonal in the inner product determined by $(P_+ - P_-)$. This inner product is nonsingular on $I_p(X_-)$, so we can consider a $(P_+ - P_-)$-orthogonal complement Z of $I_p(X)$ in $I_p(X_-)$, i.e.,

$$I_p(X) + Z = I_p(X_-). \tag{A.4}$$

Then the $(P_+ - P_-)$-orthogonal complement of $I_p(X)$ in $\mathbb{R}^n$ is $Z + Y$, the latter obviously containing $\ker(P_+ - P)$. Consequently

$$I_s(X) \subset (Z + Y) \cap I_s(X_+). \tag{A.5}$$

In the internal case we have equality in this inclusion. Now, the identity (A.3) implies that

$$\dim(Z + Y) \cap I_s(X_+) = \dim Z,$$

which together with (A.4) and (A.5) yields

$$\dim I_p(X) + \dim I_s(X) \leq \dim I_p(X_-) = \mu$$

with equality in the internal case, concluding the proof of the theorem.

Appendix B. Zero direction of σX and $\bar{\sigma} X$.

Theorems 8.5 and 8.6 can be reformulated in terms of (generalized) zero directions.

Theorem B.1. *The antistable zero directions of σX are described by*

$$(\sigma X) \cap X_+ = \{a'\bar{x}_+(0) \mid a = \bar{\Gamma}'_+ b \text{ where } b \in \ker(\bar{P} - \bar{P}_+)\}. \tag{B.1}$$

Similarly, the stable zero directions of $\bar{\sigma} X$ are given by

$$(\bar{\sigma} X) \cap X_- = \{a'\bar{x}_-(0) \mid a = \bar{\Gamma}'_- b \text{ where } b \in \ker(P - P_-)\}. \tag{B.2}$$

Theorem B.2. *The stable zero directions of σX are described by*

$$(\sigma X) \cap X_- = \{a' x_-(0) \mid a \in \mathbb{R}^n, \Gamma'_- a \in \ker(P - P_-)\}. \tag{B.3}$$

Similarly, the antistable zero directions of $\bar{\sigma} X$ are given by

$$(\bar{\sigma} X) \cap X_+ = \{a'\bar{x}_+(0) \mid a \in \mathbb{R}^n, \bar{\Gamma}'_+ a \in \ker(\bar{P} - \bar{P}_+)\}. \tag{B.4}$$

These theorems follow directly from Theorems 8.5 and 8.6, identifying G_- and $\bar{G}_+$ with Γ'_- and $\bar{\Gamma}'_+$ and $X \cap X_-$ and $X \cap X_+$ with $\ker(P - P_-)$ and $\ker(\bar{P} - \bar{P}_+)$ respectively. (Also see [13].) However, we also have the following independent coordinate-dependent proofs.

Proof of Theorem B.1. We prove (B.1). Then (B.2) follows by symmetry. The proof of this theorem runs parallel with that of Theorem 8.6. In view of the definition (3.1) of σX, we need to characterize all $\xi \in X$ such that

$$\mathrm{E}^{H^- \vee U^{-1} S} \xi = \mathrm{E}^{\{y(-1)\} \vee U^{-1} X} \xi \in X_+,$$

or, in other words, $\xi = d'\bar{x}_+(0)$ such that

$$\mathrm{E}^{\{y(-1)\} \vee U^{-1} X} d'\bar{x}(0) = c' y(-1) + b'\bar{x}(-1) = a'\bar{x}_+(0)$$

for appropriate vectors a, b, c, d. The equations connecting a, b, c, d are

$$d' \begin{bmatrix} \bar{P}\bar{C}' & \bar{P}A \end{bmatrix} = \begin{bmatrix} c' & b' \end{bmatrix} \begin{bmatrix} \Lambda_0 & C \\ C' & P \end{bmatrix}, \tag{B.5}$$

$$\begin{bmatrix} c' & b' \end{bmatrix} \begin{bmatrix} \bar{C} \\ A' \end{bmatrix} = a' \tag{B.6}$$

and

$$\begin{bmatrix} c' & b' \end{bmatrix} \begin{bmatrix} \Lambda_0 & C \\ C' & P \end{bmatrix} = a' \bar{P}_+ a. \tag{B.7}$$

Here the first two equations are projection formulas projecting ξ onto $\{y(-1)\} \vee U^{-1} X$ and $c' y(-1) + b'\bar{x}(-1)$ onto X_+ respectively, and the third equation expresses that in the latter projection the error is zero. Now, insert (B.6) into (B.7) and rearrange terms to obtain

$$\begin{bmatrix} c' & b' \end{bmatrix} \begin{bmatrix} \Lambda_0 - \bar{C}\bar{P}_+\bar{C}' & C - \bar{C}\bar{P}_+ A \\ C' - A'\bar{P}_+\bar{C}' & \bar{P} - A'\bar{P}_+ A \end{bmatrix} \begin{bmatrix} c \\ b \end{bmatrix} = 0.$$

Using the facts that $\Lambda_0 - \bar{C}\bar{P}_+\bar{C}' = \bar{D}_+\bar{D}'_+$, $C - \bar{C}\bar{P}_+ A = \bar{D}_+\bar{B}'_+$, and also $\bar{P} - A'\bar{P}_+ A = \bar{B}_+\bar{B}'_+ + (\bar{P} - \bar{P}_+)$, we obtain

$$\begin{bmatrix} c' & b' \end{bmatrix} \begin{bmatrix} \bar{D}_+\bar{D}'_+ & \bar{D}_+\bar{B}'_+ \\ \bar{B}_+\bar{D}'_+ & \bar{B}_+\bar{B}'_+ + (\bar{P} - \bar{P}_+) \end{bmatrix} \begin{bmatrix} c \\ b \end{bmatrix} = 0.$$

Since $\bar{P} \geq \bar{P}_+$, this is clearly a positive semidefinite quadratic form, and therefore

$$\begin{bmatrix} \bar{D}_+\bar{D}_+' & \bar{D}_+\bar{B}_+' \\ \bar{B}_+\bar{D}_+' & \bar{B}_+\bar{B}_+' + (\bar{P} - \bar{P}_+) \end{bmatrix} \begin{bmatrix} c \\ b \end{bmatrix} = 0.$$

The first block equation together with the fact that $\bar{D}_+\bar{D}_+'$ is invertible yields

$$c = -(\bar{D}_+\bar{D}_+')^{-1}\bar{D}_+\bar{B}_+'b. \tag{B.8}$$

Inserting this into the second block equation we get

$$(\bar{P} - \bar{P}_+)b = 0.$$

This shows that, if $a'\bar{x}_+(0) \in (\sigma X) \cap X_+$, then there is a $b \in \mathbb{R}^n$ such that $b \in \ker(\bar{P} - \bar{P}_+)$ and $a = \bar{\Gamma}_+'b$. Conversely, assume that this is satisfied, define c by (B.8) and set $d := a$. Now straight-forward calculations show that (B.5), (B.6) and (B.7) are satisfied. □

Proof of Theorem B.2. We prove (B.3); then (B.4) follows by symmetry. Recall that $\sigma X = \mathrm{E}^{H^- \vee U^{-1}S} X$. Therefore, if $a'x_-(0) \in (\sigma X) \cap X_-$, there exists a $\xi \in X$ such that

$$\mathrm{E}^{H^- \vee U^{-1}S}\xi = a'x_-(0).$$

Apply E^{H^-} to this to see that $\mathrm{E}^{H^-}\xi = a'x_-(0)$. Hence, by uniqueness of the uniform choice of bases, $\xi = a'x(0)$. On the other hand,

$$H^- \vee U^{-1}S = \{y(-1)\} \vee U^{-1}S = (\{y(-1)\} \vee U^{-1}X) \oplus U^{-1}\bar{S}^\perp,$$

since $\{y(-1)\} \in U^{-1}H^+ \subset U^{-1}\bar{S} \perp U^{-1}\bar{S}^\perp$ and $S = X \oplus \bar{S}^\perp$. Hence, since $\xi \in X \perp \bar{S}^\perp \supset U^{-1}\bar{S}^\perp$,

$$\mathrm{E}^{H^- \vee U^{-1}S}\xi = \mathrm{E}^{\{y(-1)\} \vee U^{-1}X}\xi.$$

But $a'x(0) - a'x_-(0) \perp H^- \supset \{y(-1)\}$, so the space $(\sigma X) \cap X_-$ is completely characterized by the condition

$$a'x(0) - a'x_-(0) \perp \{x(-1)\},$$

or, in other words,

$$\mathrm{E}\{[a'x(0) - a'x_-(0)]x(-1)'\} = 0. \tag{B.9}$$

To compute this covariance, note that the error process $x(t) - x_-(t)$ satisfies the forward state equation

$$x(t+1) - x_-(t+1) = \Gamma_-[x(t) - x_-(t)] + (B - B_-D_-^{-1}D)u(t)$$

so that

$$\mathrm{E}\{[x(0) - x_-(0)][x(-1) - x_-(-1)]'\} = \Gamma_-(P - P_-).$$

However, since $x_-(0) = \mathrm{E}^{H^-}x(0)$ and $a'x_-(-1) \in H^-$, (B.9) yields $a'\Gamma_-(P - P_-) = 0$ as claimed. □

REFERENCES

[1] B. D. O. Anderson, *The inverse problem of stationary convariance generation*, J. Statistical Physics, 1 (1969), pp. 133–147.

[2] F. Badawi, *PhD Thesis*, PhD Disertation, University of Kentucky, 1980.

[3] F.Badawi, A.Lindquist, and M.Pavon, *A stochastic realization approach to the smoothing problem*, IEEE Trans. Automatic Control, AC-24 (1979), pp. 878–888.

[4] G. Basile and G. Marro, *Controlled and conditioned invariant subspaces in linear system theory,* J. Optimization Theory and Appl., 3 (1973), pp. 306–316.

[5] R. S. Bucy, D. Rappaport and L. M. Silverman, *Correlated noise filtering and invariant directions for the Riccati equation*, IEEE Trans. Automatic Control, AC-15 (1970), pp. 535–540.

[6] P. E. Caines, *Linear Stochastic Systems*, Wiley, 1988.

[7] H.Dym and H.P.McKean, *Gaussian Processes, Function Theory and the Inverse Spectral Problem*, Academic Press, 1976.

[8] P. Faurre, M. Clerget, and F. Germain, *Opérateurs Rationnels Positifs*, Dunod, 1979.

[9] M. Green, *Balanced stochastic realizations,* Linear Algebra and Appl., 98 (1988), pp. 211–247.

[10] E.J. Hannan and D.S. Poskitt, *Unit canonical correlations between future and past*, Annals of Statistics, 16 (1988), pp. 784–790.

[11] A. Lindquist, *A new algorithm for optimal filtering of discrete-time stationary processes*, SIAM J. Control Optim., 12 (1974), pp. 736–746.

[12] ———, *Some reduced-order non-Riccati equations for linear least-squares estimation: the stationary, single-output case*, Int. J. Control, 24 (1976), pp. 821–842.

[13] A. Lindquist, Gy. Michaletzky and G. Picci, *Zeros of spectral factors, the geometry of splitting subspaces, and the algebraic Riccati inequality*, SIAM J. Control Optim., 33 (1995), pp. 365–401.

[14] A. Lindquist and M. Pavon, *On the structure of state-space models for discrete-time stochastic vector processes,* IEEE Trans. Automatic Control, AC-29 (1984), pp. 418–432.

[15] A. Lindquist, M. Pavon, and G. Picci, *Recent trends in stochastic realization theory,* in Prediction Theory and Harmonic Analysis: the Pesi Masani Volume, V.Mandrekar and H.Salehi (eds), North Holland, (1983).

[16] A. Lindquist and G. Picci, *On the stochastic realization problem,* SIAM J. Control Optim., 17 (1979), pp. 365–389.

[17] A. Lindquist and G. Picci, *Realization theory for multivariate stationary Gaussian processes,* SIAM J. Control Optim., 23 (1985), pp. 809–857.

[18] A. Lindquist and G. Picci, *A geometric approach to modelling and estimation of linear stochastic systems,* Journal of Mathematical Systems, Estimation and Control, 1 (1990), pp. 241–333.

[19] Gy. Michaletzky, *Zeros of (non-square) spectral factors and canonical correlations,* Proc. 11th IFAC World Congress, Tallinn, Estonia, (1992), pp. 167 – 172.

[20] Gy. Michaletzky and A. Ferrante, *Splitting subspaces and acausal spectral factors,* J. Mathematical Systems, Estimation and Control, 5 (1995), pp. 363–366, (summary; long paper to appear).

[21] Gy. Michaletzky and G. Tusnády, *Representation of inner products and stochastic realization,* in Topics in Stochastic Systems: Modelling, Estimation and Adaptive Control, L. Gerencsér and P. E. Caines (eds.), Springer-Verlag, 1991.

[22] M. Pavon, *Stochastic realization and invariant directions of the matrix Riccati equation,* SIAM J. Control Optim., 18 (1980), pp. 155–180.

[23] M. Pavon, *New results on the interpolation problem for continuous-time stationary increments processes,* SIAM J. Control Optim., 22 (1984), pp. 133–142.

[24] M. Pavon, *A new algorithm for optimal interpolation of discrete-time stationary processes,* in Analysis and Optimization of Systems, A. Bensoussan and J. L. Lions, eds., Springer-Verlag, New York, 1982, pp. 701–718.

[25] D. Rappaport, *Constant directions of the Riccati equation*, Automatica, 8 (1972), pp. 175–186.

[26] D. Rappaport and L. M. Silverman, *Structure and stability of discrete-time optimal systems*, IEEE Trans. Automatic Control, AC-16 (1971), pp. 227–232.

[27] N. I. Rozanov, *Stationary Random Processes*, Holden Day, 1963.

[28] G.Ruckebuch, *Représentations Markoviennes de processus Gaussiens Stationaires*, Thesis for Doctorat de 3ème cycle, L'Université de Paris VI, 1975.

[29] J.-Å. Sand, *Zeros of discrete-time spectral factors and the internal part of a Markovian splitting subspace,* SIAM J. Control Optim., to be published.

[30] B. Sz.– Nagy and C. Foias, *Harmonic Analysis of Operators on Hilbert Space,* North-Holland,

1970.
[31] W. M. Wonham, *Linear Multivariable Control: a Geometric Approach*, Springer-Verlag, 1977.