AFOSR-TR- 80-0458

LEVEL II

B.S. 4B

January 1980

Research Supported by:

Contract AFOSR-78-3519

See 1473

ADA086465

STRUCTURES AND ALGORITHMS IN STOCHASTIC REALIZATION THEORY AND THE SMOOTHING PROBLEM

Faris A. Badawi

Department of Mathematics
UNIVERSITY OF KENTUCKY, LEXINGTON, KENTUCKY 40506

80 7 2 110

January 1980

STRUCTURES AND ALGORITHMS IN STOCHASTIC REALIZATION THEORY AND THE SMOOTHING PROBLEM

by

Faris Abdul-Jabbar Badawi

This report is based on the unaltered thesis of Faris A. Badawi, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics at the University of Kentucky in January 1980. The research was conducted with partial support by Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-78-3519. The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.

Department of Mathematics
University of Kentucky
Lexington, Kentucky 40506

STRUCTURES AND ALGORITHMS IN STOCHASTIC REALIZATION

THEORY AND THE SMOOTHING PROBLEM

By

Faris Badawi

Anders Lindquist
Director of Dissertation

Raymond W Cox
Director of Graduate Studies

January 25, 1980
Date

This thesis contains two main topics, each of which is connected to the stochastic realization problem. First, we consider some structural and algorithmic problems in wide sense stochastic realization theory which also have applicability to many problems outside the realm of stochastic realization theory but are here formulated in that framework. We consider some geometric questions concerning the solution set of the positive real lemma and provide a Hamiltonian framework for the non-Riccati algorithms of Kailath and Lindquist; these are then applied to the stochastic realization problem. Secondly, we apply the basic techniques and concepts of the strict sense (proper) stochastic realization theory of Lindquist and Picci and Ruckebusch to the discrete-time smoothing problem. This provides a *natural* interpretation of the Mayne-Fraser two-point formula as well as many other smoothing results, the interpretations of which have hitherto been quite unclear from a probabilistic point of view. Hence we have laid the ground work for a theory of smoothing which has so far been lacking.

F. BADAWI

FARIS A. BADAWI

JANUARY 25, 1980

DATE

ACKNOWLEDGEMENTS

The author would like to express his deep gratitude to his dissertation advisor, Professor Anders Lindquist, without whose assistance, support and encouragement, this dissertation would not have materialized. Thanks are also due to Professors Roger Wets, Danny Sorensen and Georgio Picci and to Dr. Michele Pavon, with whom the author had rewarding discussions concerning this research.

My deepest thanks go to my wife, whose patience has no bounds. Her moral support throughout has been invaluable to an extent that cannot be estimated (filtered, predicted or smoothed).

The author is also grateful for financial support from the Department of Mathematics of the University of Kentucky and the U. S. Air Force Office of Scientific Research.

Finally, but not least, the author would like to thank Mrs. Beth Kaufman and Mrs. Virginia Gandy for their expert typing.

TABLE OF CONTENTS

INTRODUCTION

The *stochastic realization problem* can be simply stated as follows: Given an m-dimensional stochastic process $\{y(t);\ t \in I\}$, where the index set I may be either an interval of the real line or a set of integers, find all linear stochastic systems (in some suitable class) having the process y as its output process. These stochastic systems are called stochastic realizations of y. This problem is of considerable importance in stochastic systems theory and has applications in and connections to many fields of study, among which are network theory [8], spectral factorization [3,8], optimal control theory [6,11], stability theory [4] and the smoothing problem [64,65].

The early contributions to this problem are due to B.D.O. Anderson [3] and Faurre [11], the first of whom called it the "inverse problem of covariance generation." In these early papers, the stochastic realization problem was studied from a deterministic point of view, the objective being to determine the parameters of the stochastic systems rather than to clarify their probabilistic structures. These early results have been extended by Clerget [71] and Germain [18]. Following [2], we shall term these aspects of the stochastic realization problem *wide sense*. These problems are intimately connected to spectral factorization [12] and the *positive real lemma* [75-77] (and its nonstationary extensions), the set of all state covariance matrices P of the stochastic realizations of a stationary process being the solution set of the positive real lemma. The set P is also the solution set of the Quadratic Matrix Inequality [12,16], and a certain subset P_0 of P contains the solutions of the corresponding Algebraic Riccati Equation [16].

More recently the probabilistic aspects of stochastic realizations and their relation to Markovian representations have been studied in various aspects and degrees of completeness by Akaike[78,79], Picci [80], Lindquist and Picci [2,53-56], Ruckebusch [1,10,58-60], Lindquist, Picci and Ruckebusch [57], Pavon [9,62] and Willems and van Schuppen[81]. Here one is interested in a complete probabilistic description of the stochastic realizations; such a realization will be named *proper* [2].

In this thesis, we study certain aspects connected with the stochastic realization problem: We consider some structural and algorithmic problems in wide sense stochastic realization theory and the applications of proper stochastic realization theory to the smoothing problem. However, some of these results, in particular those related to non-Riccati algorithms and the structure of the set P, are not only part of stochastic realization theory, but have wider applicability.

In Chapter 1, we consider wide sense stochastic realization theory for stationary processes with rational spectral densities in both discrete- and continuous-time. The structures of the sets P and P_0 mentioned above are studied. A parametric representation for P is given and its boundary points are characterized. These results are generalizations of some found in [2,11,18]. Then we show that the elements of P_0 are extreme points of P. This seems to be a well-known result; however, we have been unable to find a proof of this anywhere in the literature. In Sections 1.4 and 1.5, we apply the theory of Hamiltonian systems to obtain a new derivation of the non-Riccati algorithms of Kailath and his coworkers [21,29] and Lindquist [22,23,27,83]. The basic idea of this proof was suggested to us by L.E. Zachrisson. The continuous-time version of this result, presented in Section 1.4, is quite straight-forward and our derivation follows [82] closely. As expected the discrete-time version is considerably more complicated; it is presented in Section 1.5. We obtain these results for the special type of Riccati equations that arises in the context of stochastic realizations; our results on the general case will be presented elsewhere. Finally, the factorization of the discrete-time Riccati equation presented in Section 1.5 is applied to generate realizations of y. These results are the continuous-time counterpart of Section 6 in [2].

While the study of the *proper* stochastic realization problem is of interest in itself, its concepts and techniques can be applied to other problems, an example of which is smoothing. In Chapter 2, a smoothing theory for discrete-time nonstationary systems is developed, much in analogy with our continuous-time papers [64,65]. It is shown that the smoothing estimate is contained in a finite-dimensional space $H_t^{\square}$, the frame space [53-56]. Unlike the situation in continuous-time, in the discrete setting, $H_t^{\square}$ is *not* of constant dimension, contributing to the fact that discrete-time problems are not just trivial modifications of their continuous-time counterparts. In the continuous-time setting [64,65], the invertibility of certain covariance matrices is essential. In discrete-time, this invertibility does not hold on the whole interval. Hence we apply the generalized Moore-Penrose pseudo-inverse [13,66], thereby introducing some further structure in the theory. In Section 2.6, we derive a two-filter formula of the Mayne-Fraser type for the smoothing estimate. This formula is in terms of two estimates: the first (x_*) is generated by the usual forward Kalman-Bucy filter and the second (x^*) by the forward counterpart of the backward Kalman-Bucy filter. This provides further insight into the classical theory of smoothing. In the final Section 2.7 we use a different technique to derive the smoothing formulas of Bryson and Frazier [35] and Rauch, Tung and Striebel [36], which does not employ the frame space, but an orthogonal decomposition of the closed linear span of $\{y(t);\ t \in I\}$ much along the same lines as the procedure used in [9] to solve the stationary stochastic realization problem. Unfortunately, due to time constraints, we have not had time to tie up all the loose ends of this theory, and as explained in the text, some problems have been left open, which we feel otherwise would have been resolved with a moderate amount of extra work since all the ingredients needed are at hand. We shall have to return to this in a subsequent paper.

Chapter 3 is devoted to a study of some topics on the stochastic realization problem for continuous-time nonstationary processes defined on the real line. One aim is to generalize the finite interval theory presented in [64,65]. As a by-product we obtain a natural stochastic interpretation of an algorithm due to Clerget [71] for the minimum and maximum variance realizations. Then, the non-Riccati algorithm of [2] (and Section 1.6) is generalized to this nonstationary setting. All the results of the thesis are for the so-called *regular* case (i.e., there is a full-rank observation noise), but many of them can be generalized to the nonregular setting. To aid the reader in doing the necessary conversions for this, we have included a section (Section 3.6) providing the necessary transformations. Our results here are generalizations of the stationary counterparts in [18], in which paper, a control theory approach is taken. Our results should be compared with those of Anderson *et al* [74].

Sections 1.4 and 1.5, and Chapter 2 are based on joint work with Professor Anders Lindquist.

We shall adopt the following notations in the sequel. The transpose of a matrix is denoted by ('). $E\{\cdot\}$ is the mathematical expectation, I is the unit matrix. All vectors without prime are column vectors. If R is a symmetric matrix, $R > 0$ ($R \geq 0$) means that R is positive (nonnegative) definite. If $R \geq 0$, $R^{\frac{1}{2}}$ is the unique nonnegative square root of R. δ_{st} is the Kronecker symbol. The set of integers will be denoted by Z; Z^+ will denote $\{0, 1, 2, \ldots\}$. Finally, the set of real numbers will be denoted by $\mathbb{R}$.

CHAPTER 1

RICCATI AND NON-RICCATI METHODS IN STOCHASTIC REALIZATION THEORY OF STATIONARY PROCESSES

1.1. Stochastic Realization Theory: A Review

In this section, we shall review certain facts from stochastic realization theory for both discrete - and continuous - time stationary processes. The discrete-time case will be presented first.

Let $\{y(t);\ t \in Z\}$, where Z is the set of integers, be an m-dimensional centered stationary and purely nondeterministic stochastic process defined on an underlying probability space. The process y is said to have an n-dimensional *Markovian representation* [1] if there exists an n-dimensional stochastic process $\{x(t);\ t \in Z\}$, which together with y satisfy a linear stochastic system

$$x(t+1) = Ax(t) + Bw(t) \tag{1.1a}$$

$$y(t) = Cx(t) + Dw(t) \tag{1.1b}$$

where A, B, C and D are constant matrices of dimensions $n \times n$, $n \times p$, $m \times n$ and $m \times p$ respectively, A is a stability matrix i.e. all its eigenvalues are strictly inside the unit circle (for short, $|\lambda(A)| < 1$), and $\{w(t);\ t \in Z\}$ is a p-dimensional white noise, i.e. a zero-mean stochastic process of covariance $E\{w(t)w(s)'\} = I\delta_{st}$. The process x is called the *state* of the system; it is stationary with constant covariance

$$E\{x(t)\ x(t)'\} = P = P' \geq 0 \tag{1.2}$$

which clearly satisfies the Liapunov-type equation

$$P = APA' + BB'. \tag{1.3}$$

The process y is called the *output* of the system (1.1) and w is the *input*.

The covariance function $K_y(s) := E\{y(t+s)y(s)'\}$ of the output process y of (1.1) is easily seen to be

$$K_y(s) = CA^{s-1}G\, 1_s + G'A'^{-s-1}\, C'1_{-s} + (CPC' + DD')\delta_{so}, \tag{1.4}$$

where $G = APC' + BD'$ and $1_s = 1$ if $s > 0$ and 0 otherwise. Then the *spectral density* function $\Phi(z)$ of y is given by

$$\Phi(z) = \sum_{s=-\infty}^{\infty} K_y(s)z^{-s} = C(zI-A)^{-1}G + G'(z^{-1}I-A')^{-1}C' + (CPC' + DD'), \tag{1.5}$$

which is a rational function in z. It is easy to see that $\Phi(z)$ has the following properties

(i) each element of Φ is analytic on the unit disc: $|z| \leq 1$,

(ii) $\Phi(z) = \Phi(z^{-1})'$ and

(iii) $\Phi(e^{i\omega}) \geq 0$ for all $\omega \in \mathbb{R}$.

The *stochastic realization problem* is the inverse of the above construction i.e. from the knowledge of the spectral density of a process, we wish to determine *all* representations (1.1) of the process.

More specifically, let $\{y(t); t \in Z\}$ be as above. Let the spectral density $\Phi(z)$ of y be given. Assume $\Phi(z)$ has the above properties (i.e. it is rational and satisfies (i)-(iii) above.) In addition, assume that $\Phi(e^{i\omega}) > 0$ for all $\omega \in \mathbb{R}$ and that $0 < \Phi(\infty) < \infty$ (the significance of these assumptions will be made clear when the need arises.) The problem is to find *all* Markovian representations (1.1) with $n = \dim A$ minimal and whose outputs have the same spectral density Φ as that of the given process y. Such a representation will be called *a wide sense stochastic realization* [2] of y, although it might be more descriptive to call it a realization of Φ. In fact, this is a deterministic problem which requires determining all quadruplets $[A, B, C, D]$ from the knowledge of Φ. The probabalistic problem of finding all *proper* [2] stochastic realizations (i.e. all systems (1.1) whose outputs not merely have the same covariance properties as the given process, but are equal to it a.s.) will be discussed in Chapter 2 in the nonstationary setting.

The wide sense stochastic realization problem is equivalent to the classical *spectral factorization problem* [3]: given $\Phi(z)$, find all minimal stable spectral factors of Φ i.e. all matrices $W(z)$ of proper real rational functions of minimal McMillan degree [4] with all poles inside the unit circle and satisfying

$$\Phi(z) = W(z)\, W(z^{-1})' . \tag{1.6}$$

To see that this is the case, first observe that if $[A, B, C, D]$ is a minimal realization of y, then

$$W(z) = C(zI-A)^{-1}B + D \tag{1.7}$$

is a minimal stable spectral factor of Φ. Conversely, any such minimal spectral factor W, gives rise to a whole class of wide sense stochastic realizations of the form

$$[T^{-1}AT,\ T^{-1}B,\ CT, D] \tag{1.8}$$

where [A, B, C, D] is a minimal realization [4] of W and T is an arbitrary nonsingular $n \times n$ matrix.

Using the method of partial fractions, $\Phi(z)$ can be written

$$\Phi(z) = S(z) + S(z^{-1})' \tag{1.9}$$

where S is a discrete positive real rational function [5]. Since S is proper(i.e. $S(\infty) < \infty$), it has a minimal realization [F, G, H, J], i.e.

$$S(z) = H(zI-F)^{-1}G + J \tag{1.10}$$

for some constant matrices F, G, H and J of dimensions $n \times n$, $n \times m$, $m \times n$ and $m \times m$ respectively, where n is the McMillan degree [4] of S. Hence, $|\lambda(F)| < 1$, (F, G) is controllable and (H, F) is observable. Several procedures are available for determining [F, G, H, J] [6,7], which is unique up to the equivalence (1.8). Using the fact that S is discrete positive real and the Positive Real Lemma, Anderson [5,8] has shown that all wide sense stochastic realizations of y are given by

$$[A, B, C, D] = [T^{-1}FT,\ T^{-1}(B_1, B_2)\,V,\ HT,\ (R(P)^{1/2}, 0)V] \tag{1.11}$$

where T is as above, V is a $p \times p$ constant orthogonal matrix, B_1 is $n \times m$ and B_2 is $n \times (p-m)$, P is $n \times n$ symmetric positive definite matrix which together with B_1, B_2 and R(P) satisfy

$$P = FPF' + B_1B_1' + B_2B_2' \tag{1.12a}$$

$$G = FPH' + B_1R(P)^{\frac{1}{2}} \tag{1.12b}$$

$$R(P) = J + J' - HPH' . \tag{1.12c}$$

Since $0 < \Phi(\infty) < \infty$, and assuming dim $F = n \geq 1$, it is easy to show that F is nonsingular [9], $R(P) > 0$ [9,10] and that

$$R(P) = G'F'^{-1}H' + \Phi(\infty) - HPH' . \tag{1.13}$$

It is no restriction to take $T = V = I$ i.e. to consider realizations of the form

$$x(t+1) = F\, x(t) + B_1 u(t) + B_2 v(t) \tag{1.14a}$$

$$y(t) = H\, x(t) + R(P)^{\frac{1}{2}} u(t) , \tag{1.14b}$$

where $w = \begin{bmatrix} u \\ v \end{bmatrix}$.

Let $\mathcal{P} = \{P \mid P \text{ solves } (1.12)\}$. For each $P \in \mathcal{P}$, define

$$\Lambda(P) = -P + FPF' + (G-FPH')R(P)^{-1}(G-FPH')' , \tag{1.15}$$

and let $\mathcal{P}_o = \{P \in \mathcal{P} \mid \Lambda(P) = 0\}$

In the following preposition, we collect some facts from Anderson [5], Faurre [11] and Pavon [9].

Proposition 1.1. *The set* $\mathcal{P}$ *is convex and compact and there are two elements* P_* *and* P^* *in* $\mathcal{P}_o$ *such that* $P_* \leq P \leq P^*$ *for all* $P \in \mathcal{P}$. *Moreover,* $\mathcal{P} = \{P \mid \Lambda(P) \leq 0\}$. *Finally,* $\mathcal{P}_o$ *is the set of all solutions of* (1.12) *for which* $B_2 = 0$.

Remark. In the proper stochastic realization setting (to be discussed in Chapter 2), P_0 has an interesting interpretation; these realizations are the *internal* ones (i.e. those which can be constructed in terms of the given process without introducing exogeneous noise).

The minimum P_* and the maximum P^* are of particular interest. The following matrix Riccati equations, given in [11], may be used to calculate them.

Proposition 1.2. *Let* $\{\Pi(t);\ t \in Z^+\}$ *and* $\{\bar{\Pi}(t);\ t \in Z^+\}$ *be the solutions of the* $n \times n$*-matrix difference equations*

$$\Pi(t+1) - \Pi(t) = \Lambda(\Pi(t)) \quad ; \quad \Pi(0) = 0\ , \tag{1.16a}$$

$$\bar{\Pi}(t+1) - \bar{\Pi}(t) = \bar{\Lambda}(\bar{\Pi}(t)) \quad ; \quad \bar{\Pi}(0) = 0 \tag{1.16b}$$

respectively, where Λ *is given by* (1.15) *and* $\bar{\Lambda}$ *by*

$$\bar{\Lambda}(P) = -P + F'PF + (H' - F'PG)(J + J' - G'PG)^{-1}(H' - F'PG)'. \tag{1.16c}$$

Then $\Pi(t) \to P_*$ *and* $\bar{\Pi}(t)^{-1} \to P^*$ *as* $t \to \infty$.

Remark. Equation (1.16a) has an immediate stochastic interpretation in terms of the *Kalman filter*. Consider an arbitrary realization of the form (1.14) with state covariance P. The *linear least squares estimate* $\hat{x}(t)$ of $x(t)$ given the data $\{y(0), y(1),\dots,y(t-1)\}$ is generated by the Kalman filter

$$\hat{x}(t+1) = F\hat{x}(t) + K(t)R(t)^{-\frac{1}{2}}[y(t) - H\hat{x}(t)] \quad ; \ \hat{x}(0) = 0\ , \tag{1.17}$$

where K is the *Kalman gain* and is given by

$$K = [F\Sigma H' + B_1 R(P)^{\frac{1}{2}}]R^{-\frac{1}{2}}, \tag{1.18b}$$

$$R = H\Sigma H' + R(P) \tag{1.18c}$$

and Σ is the error covariance $\Sigma(t) = E\{[x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]'\}$ which satisfies the matrix Riccati difference equation

$$\Sigma(t+1) = F\Sigma(t)F' - K(t)\ K(t)' + BB' \quad ; \quad \Sigma(0) = P\ . \tag{1.18d}$$

Set $\Pi(t) := E\{\hat{x}(t)\hat{x}(t)'\}$. Then $\Pi(t) = P - \Sigma(t)$ which inserted in (1.18d) implies that Π satisfies (1.16a) and that

$$K = (G - F\Pi H')R(\Pi)^{-\frac{1}{2}}. \tag{1.18e}$$

Hence, by the above proposition $\Pi(t) \to P_*$ and $K(t) \to B_*$ as $t \to \infty$, i.e. P_* and B_* can be regarded as the state covariance and the gain of the steady-state Kalman-Bucy filter.

In introducing the continuous-time version of the stochastic realization problem, we shall closely follow the presentation in [2]. Let $\{y(t);\ t \in \mathbb{R}\}$ be a mean-square and purely nondeterministic m-dimensional stochastic process with stationary increments and zero mean. Then there exists an orthogonal stochastic measure $d\hat{y}$ such that $y(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\, d\hat{y}(\omega)$ and $E\{d\hat{y}(\omega)\ d\hat{y}(\omega)^+\} = \Phi(i\omega)d\omega$. (Here + denotes conjugation and transposition.) The $m \times m$-matrix of real functions Φ is the *spectral density* satisfying (i) each element of Φ is analytic on the imaginary axis, (ii) $\Phi(s) = \Phi(-s)'$, (iii) $\Phi(i\omega) \geq 0$ for all $\omega \in \mathbb{R}$ and (iv) $\Phi(\infty) < \infty$. Furthermore, Φ is assumed to enjoy the additional properties that $R := \Phi(\infty)$ is positive definite (the singular case will be studied in Chapter 3) and that $\Phi(i\omega) > 0$ for all $\omega \in \mathbb{R}$. Analogously to the discrete-time case, Φ can be written

$$\Phi(s) = Z(s) + Z(-s)' , \tag{1.19}$$

where Z is a positive real function [8]. Let [F, G, H, R] be a minimal realization [4] of Z. Then $\mathrm{Re}\{\lambda(F)\} < 0$, (F, G) is controllable, and (H, F) is observable.

Now, the problem is to find *all* representations of the type

$$dx = Ax\,dt + Bdw$$

$$dy = Cx\,dt + Ddw ,$$

such that the output y has spectral density Φ, $n = \dim A$ minimal and $\mathrm{Re}\{\lambda(A)\} < 0$.

Modulo a trivial coordinate transformation in the state space, all solutions to this problem are of the form

$$dx = F\,x\,dt + B_1du + B_2dv \tag{1.20a}$$

$$dy = H\,x\,dt + R^{\frac{1}{2}}du , \tag{1,20b}$$

where $B = (B_1,B_2)$ and $P = E\{x(t)\,x(t)'\}$ satisfy the Positive Real Lemma Equations

$$FP + PF' + B_1B_1' + B_2B_2' = 0 \tag{1.21a}$$

$$G = PH' + B_1R^{\frac{1}{2}} \tag{1.21b}$$

$$P = P' > 0 \quad (n \times n\text{-matrix}) . \tag{1.21c}$$

Let $\mathcal{P} = \{P \mid P \text{ solves } (1.21)\}$. For each $P \in \mathcal{P}$, define

$$\Lambda(P) = FP + PF' + (G - PH')R^{-1}(G - PH')' . \tag{1.22}$$

Then $\mathcal{P} = \{P = P' > 0 \mid \Lambda(P) \leq 0\}$[2]. Let $\mathcal{P}_0 = \{P \in \mathcal{P} \mid \Lambda(P) = 0\}$. Then Proposition 1.1 holds for this setting also [11]. Moreover, in analogy with the discrete-time case, this problem can be seen to be equivalent to finding all minimal stable spectral factors of the type

$$W(s)\,W(-s)' = \Phi(s)\ . \qquad (1.23)$$

The continuous-time counterpart of Proposition 1.2 is

Proposition 1.3. *Let* Π *and* $\bar{\Pi}$ *be the unique solutions of the* $n \times n$*-matrix differential equations*

$$\dot{\Pi}(t) = \Lambda(\Pi(t)) \quad , \quad \Pi(0) = 0 \qquad (1.24a)$$

and

$$\dot{\bar{\Pi}}(t) = \bar{\Lambda}(\bar{\Pi}(t)) \quad , \quad \bar{\Pi}(0) = 0 \qquad (1.24b)$$

respectively, where Λ *is given by* (1.22) *and* $\bar{\Lambda}$ *by*

$$\bar{\Lambda}(P) = F'P + PF + (H' - PG)R^{-1}(H' - PG)'\ .$$

Then $\Pi(t) \to P_*$ *and* $\bar{\Pi}(t)^{-1} \to P^*$ *as* $t \to \infty$.

Remark. As was remarked (after Proposition 1.2), equation (1.24a) has the interpretation that $\Pi(t) = E\{\hat{x}(t)\,\hat{x}(t)'\}$, where $\hat{x}(t)$ is the *linear least squares estimate* of the state process $x(t)$ of the system (1.20) given the record $\{y(s);\ 0 \le s \le t\}$ which is generated by the *Kalman-Bucy filter*

$$d\hat{x} = F\,\hat{x}\,dt + K(t)R^{-\frac{1}{2}}[dy - H\,\hat{x}\,dt];\ \hat{x}(0) = 0\ ,$$

where the *Kalman gain* K is given by

$$K = (G - \Pi H')R^{-\frac{1}{2}}\ . \qquad (1.24c)$$

1.2. Structure of the Set $\mathcal{P}$: Continuous-Time

Since each wide sense stochastic realization is determined by its covariance matrix P, an investigation of the structure of the set $\mathcal{P}$ of all such matrices is deemed necessary. In this section, we shall exploit the role played by a Hamiltonian matrix to be defined below to provide some new links between the solutions of what is known as the Algebraic Riccati Equation (ARE) (the solution set of which is $\mathcal{P}_o$) and those of a Quadratric Matrix Inequality (QMI) with solution set $\mathcal{P}$. The boundary and extreme points of $\mathcal{P}$ will be studied.

Let $\mathcal{P}_+ = \{P \in \mathcal{P} \mid P > P_*\}$ and $\mathcal{P}_- = \{P \in \mathcal{P} \mid P < P^*\}$, where P_* and P^* are the minimum and maximum elements of $\mathcal{P}$ defined in Section 1.1. Since $\Phi(i\omega) > 0$ for all real $\omega, P^* - P_* > 0$ [12; p. 360], and consequently $\mathcal{P}_+$ and $\mathcal{P}_-$ are both nonempty. For each $P \in \mathcal{P}$, define the *feedback matrix*

$$\Gamma = F - (G - PH')R^{-1}H . \tag{1.25}$$

Let the feedback matrices corresponding to P_* and P^* be denoted Γ_* and Γ^* respectively. It can be shown that $\mathrm{Re}\{\lambda(\Gamma_*)\} < 0$ and $\mathrm{Re}\{\lambda(\Gamma^*)\} > 0$ [12; p. 360], [11; p. 53]. Finally, from the given matrices F, G, H and R, construct the $2n \times 2n$-matrix

$$\mathcal{F} = \begin{bmatrix} -(F - GR^{-1}H)' & -H'R^{-1}H \\ GR^{-1}G' & (F - GR^{-1}H) \end{bmatrix}, \tag{1.26}$$

(the significance of which will be clear shortly.) It is trivial to see that $\mathcal{F}$ is a *Hamiltonian matrix* i.e. $\mathcal{F} = \mathcal{I}\mathcal{F}\mathcal{I}'$, where $\mathcal{I} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$.

Consequently, if λ_i ; $i = 1,2,\ldots,n$ is an eigenvalue of $\mathcal{F}$, so is $-\lambda_i$

[13]. It can also be shown that F must have no purely imaginary eigenvalues [14].

In the following proposition, we collect some facts from Brockett [4], Faurre [11], MacFarlane [14], Mårtensson [15] and Willems [16].

Proposition 1.4. *There is one and only one* $P \in P_o$ *with* $\mathrm{Re}\{\lambda(\Gamma)\} < 0$ $(\mathrm{Re}\{\lambda(\Gamma)\} > 0)$, *namely* $P_*(P^*)$. *Moreover, the eigenvalues of the corresponding feedback matrix* $\Gamma_*(\Gamma^*)$ *are the* n *eigenvalues of* F *with negative (positive) real parts.*

The following lemma will be needed in this section.

Lemma 1.5. *Let* P_1 *and* P_2 *be arbitrary elements of* P_o *and let* Γ_1 *and* Γ_2 *be the corresponding feedback matrices* (1.25). *Then*

$$\Gamma_1 \Delta P + \Delta P \Gamma_2' = 0 , \tag{1.27}$$

where $\Delta P = P_1 - P_2$.

Proof: Since P_1 and P_2 belong to P_o, $\Lambda(P_1) = 0$ and $\Lambda(P_2) = 0$. Subtracting the second from the first and adding and subtracting the quantity $P_1 H' R^{-1} H P_2$, we obtain (1.27). □

As a first corollary to the above, we can easily prove some of the statements of the previous proposition.

Corollary 1.6. *The feedback matrices* Γ_* *and* $-\Gamma^{*\prime}$ *are similar. (Consequently, if* λ_i ; $i = 1,2,\ldots,n$ *are the eigenvalues of* Γ_*, *then* $-\lambda_i$; $i = 1,2,\ldots,n$ *are the eigenvalues of* Γ^*.*)*

Proof. Since $\Phi(i\omega) > 0$ for all $\omega \in \mathbb{R}$, $P^* - P_* > 0$. By the above lemma, $\Gamma_*(P^* - P_*) + (P^* - P_*)\Gamma^{*\prime} = 0$. Hence $\Gamma_* = -(P^* - P_*)\Gamma^{*\prime}(P^* - P_*)^{-1}$. □

Now, we turn our attention to the other solutions of the Algebraic Riccati Equation : $\Lambda(P) = 0$

For an arbitrary $n \times n$-matrix M with n^+ eigenvalues with positive real parts and n^- eigenvalues with negative real parts, let $L^+(M)$ and $L^-(M)$ denote the invariant subspaces spanned by the corresponding (generalized) eigenvectors.

Lemma 1.7 (J. C. Willems [16]). *Let* $P \in P_0$ *and* Γ *the corresponding feedback matrix* (1.25). *Then*

$$\Gamma a = \Gamma_* a \quad , \quad \textit{for } a \in L^-(\Gamma) \tag{1.28a}$$

and

$$\Gamma b = \Gamma^* b \quad , \quad \textit{for } b \in L^+(\Gamma) \tag{1.28b}$$

The following corollary is a trivial consequence of the above lemma.

Corollary 1.8. *Let* P *and* Γ *be as in* Lemma 1.7. *Then the eigenvalues of* Γ *are among those of* Γ_* *and* Γ^* *(*i.e. *among those of* ⁏ *(In particular, the feedback matrix corresponding to any solution* $P \in P_0$ *has no purely imaginary eigenvalues.)*

The above corollary is in agreement with the well known result of Potter [17] (generalized in [15] to the case of nondistinct eigenvalues) that all solutions of the (ARE) may be obtained from the eigenvectors corresponding to the eigenvalues of the Hamiltonian matrix F.

The next corollary, which holds under a natural and standard assumption that F can be transformed to Jordan form, provides some more information about the feedback matrix Γ corresponding to any solution $P \in P_o$.

Corollary 1.9. *Assume the Hamiltonian matrix F has distinct eigenvalues. Let λ be an eigenvalue of the feedback matrix Γ corresponding to an arbitrary element P of P_o. Then $-\lambda$ cannot be an eigenvalue of Γ.*

Proof. Let $[F, B, H, R^{\frac{1}{2}}]$ be the (unique) realization corresponding to P (since $P \in P_o$, $B_2 = 0$), which gives rise to the spectral factor

$$W(s)^{-1} = -R^{-\frac{1}{2}}H(sI - \Gamma)^{-1} B_1 R^{-\frac{1}{2}} + R^{-\frac{1}{2}} ,$$

which can be written $\frac{1}{\chi_\Gamma(s)}M(s)$, where χ_Γ is the characteristic polynomial of Γ and M is a matrix polynomial. Therefore, by Cramer's rule, χ_Γ equals the numerator of det W, and consequently, in view of (1.23), $\chi_\Gamma(s)\ \chi_\Gamma(-s) = \phi(s)$, where ϕ is the numerator polynomial of det Φ. But, since, in particular, this relation holds for $\Gamma = \Gamma_*$ and since $\chi_{\Gamma_*}(-s) = \chi_{\Gamma*}(s)$ (Corollarly 1.6), ϕ must be the characteristic polynomial χ_F of F (Proposition 1.4), i.e.

$$\chi_\Gamma(s)\ \chi_\Gamma(-s) = \chi_F(s).$$

Now suppose λ and $-\lambda$ are eigenvalues of Γ. The $(s - \lambda)(s + \lambda)$ is a factor of both $\chi_\Gamma(s)$ and $\chi_\Gamma(-s)$. Consequently, $(s - \lambda)^2(s + \lambda)^2$ is a factor of $\chi_F(s)$, which is clearly a contradiction to the assumption that F has distinct eigenvalues. □

In particular, we are able to see that the sets $\mathcal{P}_0^+ = \{P \in \mathcal{P}_0 \mid P > P_*\}$ and $\mathcal{P}_0^- = \{P \in \mathcal{P}_0 \mid P < P^*\}$ are singletons.

Corollary 1.10. *An* $n \times n$ *symmetric matrix* P *belongs to* $\mathcal{P}_0^+$ $(\mathcal{P}_0^-)$ *if and only if the feedback matrix* Γ *corresponding to* P *is similar to* $-\Gamma_*'$ (Γ_*).

Proof. Let $P \in \mathcal{P}_0^+$. Then $P > P_*$. Hence, by (1.27), $\Gamma = -(P - P_*)\Gamma_*' (P - P_*)^{-1}$. Conversely, if Γ is similar to $-\Gamma_*'$, then Γ has all eigenvalues in the right half plane. But, by Proposition 1.4, there is only one such feedback matrix, which is Γ^*: the one corresponding to P^*. By the assumption $\Phi(i\omega) > 0$, $P^* - P_* > 0$ i.e. $P^* \in \mathcal{P}_0^+$. The proof of the other part is analogous. □

Indeed, Corollary 1.10 may be reformulated as: $\mathcal{P}_0^+ = \{P^*\}$ and $\mathcal{P}_0^- = \{P_*\}$.

Next, we shall discuss the relationships between the solutions of (ARE) : $\Lambda(P) = 0$ and $(\overline{\text{QMI}})$: $\Lambda(P) \leq 0$.

For any $\varepsilon > 0$ and any matrix M, define the ball $\mathcal{U}(M,\varepsilon) = \{L : L = M + N, \|N\| < \varepsilon\}$, where $\|\cdot\|$ is the usual matrix norm associated to the vector Euclidean norm.

Definition 1.11. An $n \times n$ symmetrix matrix P belongs to the *boundary* of $\mathcal{P}$ (denoted by $\partial\mathcal{P}$) if, for all $\varepsilon > 0$, there exist two matrices P_1 and P_2 belonging to $\mathcal{U}(P,\varepsilon)$ such that $P_1 \in \mathcal{P}$ and $P_2 \notin \mathcal{P}$. (Since $\mathcal{P}$ is closed, $\partial\mathcal{P} \subset \mathcal{P}$.)

The following theorem, which provides a complete characterization of the boundary points of $\mathcal{P}$, is due to Germain [18].

Theorem 1.12. *Let* $P_o \in \mathcal{P}$ *and set* $-Q_o = FP_o + P_oF'$ *and* $S_o = G - P_oH'$. *Then* $P_o \in \partial\mathcal{P}$ *if and only if the matrix* $M_o = \begin{bmatrix} Q_o & S_o \\ S_o' & R \end{bmatrix}$ *is singular.*

As a corollary, we obtain the following result linking $\mathcal{P}_o$ with $\mathcal{P}$. A sharper result will be given later in this section.

Corollary 1.13. $\mathcal{P}_o \subset \partial\mathcal{P}$.

Proof. Let $P_o \in \mathcal{P}_o$. Then $\Lambda(P_o) = 0$, consequently, the matrix M_o (see Theorem 1.12) is singular. □

In fact, if $m < n$ (which is usually the case in application), we have a stronger result, namely

Corollary 1.14. *Let* $m < n$ *and let* P_1 *and* P_2 *be arbitrary elements of* $\mathcal{P}_o$. *Then the segment* $[P_1, P_2]$ *is contained in* $\partial\mathcal{P}$. *(In particular,* $[P_*, P^*] \subset \partial\mathcal{P}$.)

Proof. Let $\alpha \in [0,1]$ and define $P(\alpha) = \alpha P_1 + (1 - \alpha)P_2$. Then $\Lambda(P(\alpha))$ may be written

$$\Lambda(P(\alpha)) = \alpha\Lambda(P_1) + (1 - \alpha)\Lambda(P_2) - \alpha(1 - \alpha)(P_1 - P_2)H'R^{-1}H(P_1 - P_2). \quad (1.29)$$

Since $P_1, P_2 \in \mathcal{P}_o$, $\Lambda(P_1) = \Lambda(P_2) = 0$. Let $Q(\alpha) = -FP(\alpha) - P(\alpha)F'$ and $S(\alpha) = G - P(\alpha)H'$. Then it is easy to see that

$-\Lambda(P(\alpha) = Q(\alpha) - S(\alpha)R^{-1}S(\alpha)' = \alpha(1-\alpha)(P_1 - P_2)\, H'R^{-1}H(P_1 - P_2).$

If $m < n$, $H'R^{-1}H$ is not full rank and hence the matric $M(\alpha) = \begin{bmatrix} Q(\alpha) & S(\alpha) \\ S(\alpha)' & R \end{bmatrix}$ is singular. Then $P(\alpha) \in \partial\mathcal{P}\ \forall\, \alpha \in [0,1]$. □

The final task of this section is to prove that solutions of the (ARE) are extreme points of the set of solutions of the (QMI). This is a much stronger result than Corollary 1.13. (The extreme points of a set are contained in its boundary.) To this end, we shall need the following lemma.

Lemma 1.15. *Let* P *be an arbitrary element of* $\mathcal{P}_o$. *Suppose there exist two elements* P_1 *and* P_2 *belonging to* $\mathcal{P}$ *such that* $P = \alpha P_1 + (1-\alpha)P_2$ *for some* $\alpha \in (0,1)$. *Then,* $P_1 \in \mathcal{P}_o$, $P_2 \in \mathcal{P}_o$ *and* $\Delta P H'R^{-1}H\Delta P = 0$, *where* $\Delta P = P_1 - P_2$.

Proof. Let P, P_1 and P_2 be as in the lemma. Then, by (1.29), $\Lambda(P) = \alpha\Lambda(P_1) + (1-\alpha)\Lambda(P_2) - \alpha(1-\alpha)\Delta P H'R^{-1}H\Delta P$. Since $\alpha \in (0,1)$, the last term is ≤ 0. On the other hand, since $\Lambda(P) = 0$ (for $P \in \mathcal{P}_o$), $\Lambda(P_1) \leq 0$ and $\Lambda(P_2) \leq 0$ (for both belong to $\mathcal{P}$), the last term must be ≥ 0. As $\alpha > 0$, $\Delta P H'R^{-1}H\Delta P = 0$. Consequently, $\Lambda(P_1) = \Lambda(P_2) = 0$. which implies P_1 and P_2 belong to $\mathcal{P}_o$. □

The above lemma says in essence that elements of $\mathcal{P}_o$ cannot be written as the convex combination of elements other than those of $\mathcal{P}_o$.

The next theorem is the main result of this section. The basic idea of its proof was suggested to us by Professor D. Sorensen.

Theorem 1.16. *Let* $P \in \mathcal{P}_o$. *Then* P *is an extreme point of* $\mathcal{P}$.

Proof. Let $P \in \mathcal{P}_o$ and assume there exist P_1 and $P_2 \in \mathcal{P}$ such that $P = \alpha P_1 + (1-\alpha)P_2$ for $\alpha \in (0,1)$. We shall show that $P = P_1 = P_2$. By Lemma 1.15, $P_1 \in \mathcal{P}_o$, $P_2 \in \mathcal{P}_o$ and $\Delta PH'R^{-1}H\Delta P = 0$. The last of these facts implies $\Delta PH'R^{-\frac{1}{2}} = 0$, i.e. $P_1H'R^{-\frac{1}{2}} = P_2H'R^{-\frac{1}{2}}$. This in turn implies that $(G - P_1H')R^{-1}(G - P_1H')' = (G - P_2H')R^{-1}(G - P_2H')' := E$. However, P_1 and P_2 are in $\mathcal{P}_o$ implies $FP_1 + P_1F' = -E = FP_2 + P_2F'$. Hence $F\Delta P + \Delta PF' = 0$. But F is a stability matrix i.e. $\mathrm{Re}\lambda(F) < 0$. Then F and $-F'$ have no eigenvalue in common, which implies $\Delta P = 0$ (see e.g. [19]). Hence $P_1 = P_2 = P$. □

If the eigenvalues of F were distinct, the above result may alternatively be proved by the following.

Proposition 1.17. *Let* P_1 *and* P_2 *be two elements of* $\mathcal{P}_o$ *such that* $\Delta PH'R^{-1}H\Delta P = 0$, *where* $\Delta P = P_1 - P_2$. *Then,* $\Gamma_1 = \Gamma_2$ *and*

$$\Gamma_2\Delta P + \Delta P\Gamma_2' = 0 , \qquad (1.30)$$

where Γ_1 *and* Γ_2 *are the feedback matrices corresponding to* P_1 *and* P_2 *respectively.*

Proof. Recall that $\Gamma_1 = F - GR^{-1}H + P_1H'R^{-1}H$. As was indicated in the proof of Theorem 1.16, $\Delta PH'R^{-1}H\Delta P = 0$ implies $P_1H'R^{-\frac{1}{2}} = P_2H'R^{-\frac{1}{2}}$. Then $\Gamma_1 = \Gamma_2$. The rest of the result then follows by (1.27). □

Therefore, if the eigenvalues of F are distinct, so are those of Γ_2 (Corollary 1.8). Then, by Corollary 1.9, Γ_2 has no opposite eigenvalues, and consequently ΔP in (1.30) will be zero [19]. Hence, in view of Lemma 1.15 any $P \in P_o$ is an exteme point of $\mathcal{P}$.

Of course, if m = n, Theorem 1.16 would follow trivially from Lemma 1.15 since then $H'R^{-1}H$ is full rank and hence the condition $\Delta PH'R^{-1}H\Delta P = 0$ implies $\Delta P = 0$.

1.3 Structure of the Set $\mathcal{P}$: Discrete-Time

In this section, we shall present the discrete-time versions of some of the results of the previous section; the purpose being to facilitate easy comparisons. At times, we shall need to resort to a well-known equivalence between dynamic systems in continuous- and discrete-time [8, 11, 18] to prove some of these results. Further properties of the set $\mathcal{P}$ in this discrete setting will be studied in Section 1.6.

Let P_+ and P_- be defined as in the previous section. Here again, since $\Phi(e^{i\omega}) > 0$ for all real ω, $P^* - P_* > 0$ [18]. Analogously with (1.25), in this case we define the *feedback matrix* to be

$$\Gamma = F - (G - FPH')R(P)^{-1}H. \tag{1.31}$$

The feedback matrices Γ_* and Γ^* corresponding to P_* and P^* satisfy the properties: $|\lambda(\Gamma_*)| < 1$ and $|\lambda(\Gamma^*)| > 1$ [11]. Furthermore, it can be shown [9] that $\Phi(\infty) > 0$ implies that Γ_* is nonsingular.

The following is a slight modification of a result in [18]. It will be useful in what follows. Our proof will shed more light on the structure of $\mathcal{P}$. Unlike [18], we do not rely upon the corresponding continuous-time result.

Lemma 1.18 (a) *Set* $R_* = J + J' - HP_*H'$. *Let* $P \in \mathcal{P}_+$ *and set* $M_* = (P - P_*)^{-1}$. *Then* M_* *satisifies*

$$-M_* + \Gamma_*' M_* \Gamma_* + H'R_*^{-1}H + N = 0 \tag{1.32a}$$

for some nonnegative definite matrix N.

(b) *Set* $R^* = J + J' - HP^*H'$. *Let* $P \in \mathcal{P}_-$ *and set* $M^* = (P^* - P)^{-1}$. *Then* M^* *satisfies*

$$M^* - \Gamma^{*\prime} M^* \Gamma^* + H'R^{*-1}H + N = 0 \tag{1.32b}$$

for some nonnegative definite matrix N.

Proof. (a) Let $P \in \mathcal{P}_+$. In view of (1.12) and (1.16a), we have

$$P = FPF' + (G - FPH')R(P)^{-1}(G - FPH')' + B_2B_2', \text{ and}$$

$$P_* = FP_*F' + (G - FP_*H')R_*^{-1}(G - FP_*H')' .$$

Upon subtracting the second of these two relations from the first, the following is obtained

$$M_*^{-1} = FM_*^{-1}F' + KR(P)K' - K_*R_*K_*' + B_2B_2' , \tag{1.33}$$

where K and K_* are defined by $KR(P) = G - FPH'$ and $K_*R_* = G - FP_*H'$. It is easy to see that $R_* = R(P) + HM_*^{-1}H'$ and that $K = [K_*R_* - FM_*^{-1}H']R(P)^{-1}$. Let $\Delta K = K - K_*$. Then $\Delta K = (K_*H - F)M_*^{-1}H'R(P)^{-1}$ and $K = K_* + \Delta K$ and (1.33) becomes

$$M_*^{-1} = FM_*^{-1}F' + (K_* + \Delta K)R(P)(K_* + \Delta K)' - K_*(R(P) + HM_*^{-1}H')K_*' + B_2B_2' .$$

After long, but simple calculation, we get

$$M_*^{-1} = \Gamma_*M_*^{-1}\Gamma_*' + \Gamma_*M_*^{-1}H'R(P)^{-1}HM_*^{-1}\Gamma_*' + B_2B_2'$$

or

$$M_*^{-1} \geq \Gamma_* [M_*^{-1} + M_*^{-1}H'R(P)^{-1}HM_*^{-1}]\Gamma_*' .$$

Taking the inverse of both sides, we obtain

$$M_* \le \Gamma_*^{\prime -1}[M_*^{-1} + M_*^{-1}H'R(P)^{-1}HM_*^{-1}]^{-1}\Gamma_*^{-1}$$

from which, the following relation is obtained

$$M_* \le \Gamma_*^{\prime -1}M_*\Gamma_*^{-1} - \Gamma_*^{\prime -1}H'R_*^{-1}H\Gamma_*^{-1} ,$$

using the matrix inversion lemma

$$[A + BD^{-1}C]^{-1} = A^{-1} - A^{-1}B[D + CA^{-1}B]^{-1}CA^{-1} . \qquad (1.34)$$

Then, premultiply the last inequality by Γ_*' and postmultiply by Γ_* to get

$$\Gamma_*'M_*\Gamma_* \le M_* - H'R_*^{-1}H ,$$

which yields (1.32a) for some nonnegative definite matrix N. This proves (a). The proof of (b) is analogous. □

As a first application of this lemma, in the next theorem, we give a parametric representation for the set $\mathcal{P}$. The formulation of the result is analogous to that in [2] for the continuous-time case.

Theorem 1.19 *Let* $M_*(N)$ *and* $M^*(N)$ *be the solutions of* (1.32a) *and* (1.32b) *respectively. Then*

(a) the matrix $P = P_* + [M_*(N)]^{-1}$ *belongs to* P_+ *if and only if* $N \ge 0$,

(b) the matrix $P = P^* - [M^*(N)]^{-1}$ *belongs to* P_- *if and only if* $N \ge 0$, *and*

(c) $P^* - P_* = [M_*(0)]^{-1} = [M^*(0)]^{-1}$.

Proof. (a) We have to prove the "if" part; the "only if" was proved in Lemma 1.18. Let $M_*(N)$ be a solution of (1.32a) with $N \geq 0$. The pair (Γ_*, H) is observable for (F, H) is [20]. Recalling that Γ_* is a stability matrix, a standard result in stability theory (see e.g. [13; p. 86]) implies $M_*(N) > 0$. Consequently $P_* + [M_*(N)]^{-1} \in \mathcal{P}_+$. The proof of (b) is analogous and that of (c) is immediate. □

In addition to its significance in parametrizing the set $\mathcal{P}$, Theorem 1.19, together with Proposition 1.2 provide us with a procedure to generate stochastic realizations of y corresponding to an arbitrary element $P \in \mathcal{P}_+ \cup \mathcal{P}_-$: First use (1.16a) to compute P_*; P^* will be obtained from Theorem 1.19(c) and varying N over the nonnegative cone will generate the other elements of $\mathcal{P}_+ \cup \mathcal{P}_-$. The realization $[F, B, H, (R(P)^{\frac{1}{2}},0)]$ corresponding to $P \in \mathcal{P}_+ \cup \mathcal{P}_-$ can be computed via

$$B_1 = (G - FPH')R(P)^{-\frac{1}{2}} \tag{1.35a}$$

$$B_2B_2' = -\Lambda(P) \ . \tag{1.35b}$$

This procedure for generating stochastic realizations requires solving a matrix Riccati equation (1.16a) in order to determine P_* in addition to the burden of determining $P \in \mathcal{P}_+ \cup \mathcal{P}_-$. In Section 1.6 another procedure that eliminates the intermediate step of computing P will be given.

The second aspect of the structure of $\mathcal{P}$ that will be discussed in this section is its boundary.

As was pointed out in Proposition 1.1, $\mathcal{P}$ is bounded. A complete characterization of the boundary points of the corresponding set $\mathcal{P}^c$ in the continuous-time setting was given in Theorem 1.12. Here, by associating the discrete-time quadruplet [F, G, H, J] with a continuous-time one $[F_c, G_c, H_c, R_c]$(an idea that is well known [8, 11, 18]), we shall give complete characterization of these boundary points. We shall also gain more insight into the relationship between the set P and the set P_o.

In order to distinguish between the continuous and discrete settings, we shall adjoin the letter c(d) (as a subscript or superscript) to the matrices and sets of the continuous (discrete) setting whenever there is a need for distinction.

Definition 1.20. The discrete-time quadruplet [F, G, H, J] and the continuous-time one $[F_c, G_c, H_c, R_c]$ are said to be *equivalent* if $P = P^c$.

The following proposition shows how to construct a quadruplet in one setting from one in the other setting. The proof can be found in [11].

Proposition 1.21. *Every discrete-time quadruplet* [F, G, H, J] *is equivalent to a continuous-time one* $[F_c, G_c, H_c, R_c]$, *where*

$$F_c := (F + I)^{-1}(F - I), \tag{1.36a}$$

$$G_c := \sqrt{2}\,(F + I)^{-1}G, \tag{1.36b}$$

$$H_c := \sqrt{2}\,H(F + I)^{-1}, \tag{1.36c}$$

$$R_c := J + J' - H(F + I)^{-1}G - G'(F' + I)^{-1}H' \tag{1.36d}$$

Conversely, every continuous-time quadruplet [F, G, H, R] *is equivalent to a discrete-time one* $[F_d, G_d, H_d, J_d]$, *where*

$$F_d := (I - F)^{-1}(I + F), \tag{1.37a}$$

$$G_d := \sqrt{2}\,(I - F)^{-1}G, \tag{1.37b}$$

$$H_d := \sqrt{2}\,H(I - F)^{-1}, \tag{1.37c}$$

$$J_d := \frac{1}{2}R + H(I - F)^{-1}G\ . \tag{1.37d}$$

Now, we are ready to state the discrete-time version of Theorem 1.12.

Theorem 1.22. *Let* $P_o \in P$ *and set* $Q_o = P_o - FP_oF'$ *and* $S_o = G - FP_oH'$. *Then* $P_o \in \partial P$ *if and only if the matrix* $M_o = \begin{bmatrix} Q_o & S_o \\ S_o' & R(P_o) \end{bmatrix}$ *is singular.*

Proof. Let the quadruplet $[F_c, G_c, H_c, R_c]$ be the one given by (1.36) corresponding to $[F, G, H, J]$. Let $\Phi_c(\cdot)$ be its corresponding spectral density. It is not hard to see that $\Phi(e^{i\omega}) > 0$ for all $\omega \in \mathbb{R}$ implies $\Phi_c^{\cdot}(i\omega) > 0$; hence $R_c > 0$. By Theorem 1.12, $P_o \in \partial P^c$ if and only if the matrix $M_c = \begin{bmatrix} Q_c & S_c \\ S_c & R_c \end{bmatrix}$ is singular. However, it is easily seen that [11]

$$M_c = \begin{bmatrix} \sqrt{2}\,(F + I)^{-1} & 0 \\ -H(F + I)^{-1} & I \end{bmatrix} M_o \begin{bmatrix} \sqrt{2}\,(F' + I)^{-1} & -(F + I)^{-1}H' \\ 0 & I \end{bmatrix}$$

Hence M_c and M_o have the same rank and one is singular if and only if the other is. □

Corollary 1.23. $P_o \subset \partial P$

Proof. If $P_o \in P_o$, then M_o in Theorem 1.22 is signular. □

Corollary 1.24. *Let* $m < n$ *and let* P_1 and P_2 *be arbitrary elements of* P_o. *Then the segment* $[P_1, P_2] \subset \partial P$. *(In particular* $[P_*, P^*] \subset \partial P$.*)*

Proof. It is not hard to see that $P \in P_o$ if and only if $P \in P_o^c$. Hence, by Corollary 1.14, $P(\alpha) = \alpha P_1 + (1 - \alpha)P_2 \in \partial P^c = \partial P$, for all $\alpha \in [0,1]$. □

Theorem 1.25. *Let* $P \in P_o$. *Then* P *is an extreme point of* P.

Proof. By Theorem 1.16, P is an extreme point of $P^c = P$. □

1.4 A Hamiltonian Approach to the Factorization of the Matrix Riccati Differential Equation

Consider the matrix Riccati differential equation

$$\dot{P} = \Lambda(P) \quad ; \quad P(0) = P_o \tag{1.38}$$

which is of the type encountered in Section 1.1. For convenience, we recall that

$$\Lambda(P) := FP + PF' + (G - PH')R^{-1}(G - PH')',$$

where the quadruplet [F, G, H, R] is a minimal realization of the positive real matrix function Z(s) defined by (1.19). Hence (1.38) has a unique bounded solution on $[0,\infty)$ for all $P_o \le P^*$ [18]. In this section, we shall present a new approach for factorizing the above Riccati equation based on Pontryagin's Maximum Principle. In this way, we shall obtain a new derivation of the non-Riccati algorithms due to Kailath [21] and Lindquist [22-24]. Our derivation will shed more light on these algorithms and will provide new links to the Hamiltonian formulation of (1.38). The basic idea of this procedure (which was indirectly suggested to us by Professor L. E. Zachrisson) is to consider the above Riccati equation as arising from an optimal control problem, with which a Hamiltonian function is associated. The factorization will then be a direct consequence of the fact that the Hamiltonian function is constant along the optimal trajectory. Another procedure based on the Hamiltonian formulation can be found in Bucy and Joseph [25].

It is worth noting that (1.38) is not the most general type of Riccati equations that one might encounter (e.g. one might have an extra constant term added to the right hand side of (1.38)). However, our aim here is to convey the basic ideas of the method, and we are therefore using the form (1.38) which arises in stochastic realization theory.

Consider the following control problem. Find a square integrable control function $u(\cdot)$ so as to minimize

$$J(u;t_1,a) = -\frac{1}{2}x'(0)P_o\,x(0) + \frac{1}{2}\int_0^{t_1}[u(t)'Ru(t) + 2\,x(t)'G\,u(t)]dt, \tag{1.39}$$

subject to

$$\dot{x}(t) = -F'x(t) - H'u(t) \quad ; \quad x(t_1) = a \tag{1.40}$$

Note that since $Z(s)$ is positive real, the function $J(u;.,.)$ is bounded from below [8; pp. 231-232]. Hence $\xi := \inf_u J(u;.,.) > -\infty$. Let $\{u_k\ ;\ k \in Z^+\}$ (where $Z^+ = \{0,1,2,\ldots\}$) be a control sequence such that $J(u_k;.,.) \to \xi$ as $k \to \infty$. Then, using the parallelgram identity in Hilbert space and a completeness argument, it is seen that there is a control u^o

such that $J(u^o;.,.) = \xi$. In fact, this can also be seen from the proof of

Proposition 1.26. *There exists a unique square integrable function* u^o *minimizing* $J(u;t_1,a)$. *Moreover,* $J(u^o;t_1,a) = -\frac{1}{2}a'P(t_1)a$, *where* P *is the unique solution of* (1.38).

Proof. Consider the function $\frac{1}{2}x(t)'P(t)\,x(t)$. Upon differentiating this quantity and integrating between 0 and t_1, we obtain

$$J(u;t_1,a) + \frac{1}{2}a'P(t_1)a = \frac{1}{2}x(0)'[P(0) - P_0]x(0) + \frac{1}{2}\int_0^{t_1}[\dot{P}(t) - \Lambda(P(t))]$$
$$+ \frac{1}{2}\int_0^{t_1}\|u(t) + R^{-1}G'x(t) - R^{-1}HP(t)x(t)\|_R dt,$$

where $\Lambda(P)$ is given by (1.22) and $\|x\|_R = x'Rx$. Then the result follows by noting that P satisfies (1.38) and that it can be choosen to make the last term zero. □

The optimal control u^o can be obtained as a corollary to this proposition. However, since the Hamiltonian function will play an important role in what follows, we shall instead apply the Maximum Principle. First, let x^o be the solution of (1.40) corresponding to u^o. To apply the Pontryagin Maximum Principle, define the *Hamiltonian function*

$$H(t,x(t),u(t),y(t)) = \frac{1}{2}u(t)'Ru(t) + x'(t)Gu(t) + y(t)'[-F'x(t) - H'u(t)]. \qquad (1.41)$$

Then the Maximum Principle requires that for a control u^o to be optimal, $H(t,x^o(t),u^o(t),y(t))$ must be minimal i.e.

$$\frac{\partial H}{\partial u}(t,x^o(t),u^o(t),y(t)) = 0 = Ru^o(t) + G'x^o(t) - Hy(t), \qquad (1.42a)$$

where the adjoint function y is given by

$$\frac{\partial H}{\partial x}(t,x^o(t),u^o(t),y(t)) = -\dot{y} = Gu^o(t) - Fy(t) \tag{1.42b}$$

with initial condition $y(0) = P_o x^o(0)$. Hence the optimal control is

$$u^o(t) = -R^{-1}[G'x^o(t) - Hy(t)]. \tag{1.43}$$

Using (1.40), (1.42) and (1.43), it is easy to see that the 2n-vector $\begin{bmatrix} x^o \\ y \end{bmatrix}$ satisfies

$$\begin{bmatrix} \dot{x}^o \\ \dot{y} \end{bmatrix} = \mathcal{F} \begin{bmatrix} x^o \\ y \end{bmatrix} \quad ; \quad \begin{bmatrix} x^o(t_1) \\ y(0) \end{bmatrix} = \begin{bmatrix} a \\ P_o x^o(0) \end{bmatrix}, \tag{1.44}$$

where $\mathcal{F}$ is defined by (1.26). Then

$$x^o(t) = X(t)x_o \tag{1.45a}$$

$$y(t) = Y(t)x_o \tag{1.45b}$$

where $x_o = x^o(0)$ and X and Y are $n \times n$ matrix functions satisfying

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = \mathcal{F} \begin{bmatrix} X \\ Y \end{bmatrix} \quad ; \quad \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} = \begin{bmatrix} I \\ P_o \end{bmatrix}. \tag{1.45c}$$

We recall the following well-known fact:

Proposition 1.27. *Let* X *and* Y *be as in* (1.45). *Then* X(t) *is non-singular for all* t *and the matrix function* $P(t) := Y(t)\,X(t)^{-1}$ *is the unique solution of* (1.38).

As a corollary to the above proposition and using (1.45a,b), it can be seen that

$$y(t) = P(t)\,x^o(t). \tag{1.46}$$

Now, let the function $H^o : [0,t_1] \to \mathbb{R}$ be defined by

$$H^o(t) := H(t,x^o(t),u^o(t),y(t)). \tag{1.47}$$

Lemma 1.28. *Let* H^o *be defined by* (1.47). *Then*

$$H^o(t) = -\frac{1}{2}x^o(t)'\dot{P}(t)x^o(t). \tag{1.48}$$

Moreover, $\frac{\partial H^o}{\partial t}(t) = 0$, i.e. H *is constant along the optimal trajectory.*

Proof. Clearly, for each $t \in [0,t_1]$, the problem to minimize $J(u;t,x^o(t))$ has the optimal solution $\{u^o(s);s \in [0,t]\}$; we shall misuse notations somewhat by calling this restricted function u^o also. Then, by Proposition 1.26, $J(u^o;t,x^o(t)) = -\frac{1}{2}x^o(t)'P(t)x^o(t)$. Hence, since $H^o(t) = \frac{dJ}{dt}(u^o;t,x^o(t)) + y(t)'\dot{x}^o(t)$, where $y(t)$ is given by (1.46), (1.48) follows. The fact that $\frac{\partial H^o}{\partial t}(t) = 0$ follows from elementary calculus. □

Lemma 1.29. *Let* H^o *be defined by* (1.47). *Then*

$$H^o(t) = -\frac{1}{2}x_o'M(t)x_o\ , \tag{1.49}$$

where $x_o = x^o(0)$ *and the* $n \times n$ *matrix function* M *is defined by*

$$M(t) := X(t)'\dot{P}(t)X(t). \tag{1.50}$$

Proof. The result is an immediate consequence of (1.45a) and (1.48) □

Lemma 1.30. *Let* M *and* Λ *be given by* (1.50) *and* (1.22) *respectively. Then*

$$M(t) = \Lambda(P_o) \tag{1.51}$$

(*i.e.* M *is constant.*)

Proof. Using (1.44) and (1.46), it is easy to see that $\dot{x}^o = -\Gamma' x^o$, where Γ is the feedback matrix (1.25). Let Ψ be the transition matrix of Γ. Then $x_o = \Psi(0,t_1)a$. Hence, $H^o(t) = -\frac{1}{2} a'\Psi(0,t_1)'M(t)\Psi(0,t_1)a$, which is constant for all $a \in \mathbb{R}^n$. Consequently, $\Psi(0,t_1)'M(t)\Psi(0,t_1)$ is a constant matrix, hence the same is true for M. But, by definition $M(0) = \dot{P}(0)$, which, in view of (1.38) is the same as $\Lambda(P_o)$; and consequently (1.51) follows. □

Lemma 1.31. *Let* X *be as in* (1.45). *Then*

$$\frac{d}{dt}(X(t)')^{-1} = \Gamma(t)(X(t)')^{-1},$$

where Γ *is given by* (1.25).

Proof. Using $\frac{d}{dt}(X'^{-1}) = -X'^{-1}\dot{X}X'^{-1}$, (1.45) and $Y(t) = P(t)X(t)$, the result follows. □

We are now ready to state the main result. First observe that, since the $n \times n$-matrix $\Lambda(P_o)$ is symmetric, there exist two constant matrices N and S such that $\Lambda(P_o) = NSN'$, where N is $n \times r$, S is $r \times r$ and r is the rank of $\Lambda(P_o)$. (*For example,* S can be choosen as the signature matrix although we shall use a different S below.)

Theorem 1.32 ([21]). *Let* P *be the unique solution of* (1.38). *Then*

$$\begin{cases} \dot{P}(t) = Q(t)SQ(t) & ; \quad P(0) = P_o \\ \dot{Q}(t) = \Gamma(t)Q(t) & ; \quad Q(0) = N, \end{cases} \qquad \begin{matrix} (1.52a) \\ (1.52b) \end{matrix}$$

where $Q(t) = (X'(t))^{-1}N$ *and* N *and* S *are given as above.*

Proof. From (1.50) and (1.51), we have $\dot{P} = X'^{-1}\Lambda(P_o)X^{-1}$, which by the preceding discussion and Lemma 1.31, is (1.52). □

As an application of the factorization (1.52), consider the problem of determining the Kalman gain K given by (1.24c), where Π is the solution of (1.24a), in which case $\Pi_o = 0$ and $\Lambda(0) = GR^{-1}G'$. Hence choosing $N = GR^{-\frac{1}{2}}$ and $S = I$ in (1.52), we obtain the following non-Riccati algorithm for K

$$\dot{K} = -QQ'H'R^{-\frac{1}{2}} \quad ; \quad K(0) = GR^{-\frac{1}{2}} \qquad (1.53a)$$

$$\dot{Q} = (F - KR^{-\frac{1}{2}}H)Q \quad ; \quad Q(0) = GR^{-\frac{1}{2}}, \qquad (1.53b)$$

which was first obtained independently by Kailath [21], who used the factorization above, and Lindquist [22], who derived it from basic principles using backward innovations.

As another application of interest in realization theory, consider the case where P_o is any element of $\mathcal{P}$ as defined in Section 1.1. In that case $\Lambda(P_o) = -B_2B_2'$, where B_2 is given by (1.20a), and we may choose N and S to be B_2 and -I respectively to obtain

$$\begin{cases} \dot{P} = -QQ' & ; \quad P(0) = P_o \\ \dot{Q} = \Gamma Q & ; \quad Q(0) = B_2. \end{cases} \qquad \begin{matrix} (1.54a) \\ (1.54b) \end{matrix}$$

Then it can be shown [2] that $P(t) \in \mathcal{P}$ for each $t \in \mathbb{R}$ and therefore $P(t)$ is the state covariance of a realization for which (in view of (1.21))

$$B_1 = (G - P(t)H')R^{-\frac{1}{2}} \quad \text{and} \quad B_2 = Q(t) .$$

Hence, we have the following non-Riccati algorithm generating a family of realizations

$$\dot{B}_1 = B_2 B_2' H' R^{-\frac{1}{2}} \quad ; \quad B_1(0) = (B_o)_1 \tag{1.54c}$$

$$\dot{B}_2 = (F - B_1 R^{-\frac{1}{2}} H) B_2 \quad ; \quad B_2(0) = (B_o)_2 \tag{1.54d}$$

for a given initial matrix $B_o = [(B_1)_o, (B_2)_o]$. (Note the parameter t is *not* time now.)

This algorithm was first presented in [2]. In Section 1.6, we are going to derive its discrete-time version, which is, as expected, more complicated. Also, in Chapter 3, we shall derive the nonstationary version of (1.54).

1.5. A Hamiltonian Approach to the Factorization of the Matrix Riccati Difference Equation

In this section, the matrix Riccati difference equation

$$P(t + 1) = FP(t)F' + (G - FP(t)H')R(t)^{-1}(G - FP(t)H')'; \; P(0) = P_o. \tag{1.55}$$

[which is the same as $P(t + 1) - P(t) = \Lambda(P(t))$, where Λ is given by (1.15)] will be considered, where the quadruplet $[F, G, H, J]$ is a minimal realization of $S(z)$ defined by (1.9) and $R(t) = J + J' - HP(t)H'$. The aim is to obtain a factorization of this matrix Riccati equation, analogous to the one obtained in the previous section. This will not only facilitate comparisons with the continuous setting, but will also be the basis for the next section. Moreover, as will be seen shortly, the lack of symmetry between the two settings will again be illustrated.

Before stating the control problem which gives rise to the above Riccati equation, we note that the matrix $T := J + J'$ is nonsingular since $T = R(t) + HP(t)H'$, and $R(t)$ is assumed to be positive definite for all t. Again the problem is to find a control $u(\cdot)$ which minimizes

$$J(u;t_1,a) = -\frac{1}{2}x(0)'P_0x(0) + \frac{1}{2}\sum_{t=0}^{t_1-1}[u(t+1)'Tu(t+1) + 2x(t+1)'Gu(t+1)], \tag{1.56}$$

subject to

$$x(t) = F'x(t+1) + H'u(t+1) \quad ; \quad x(t_1) = a. \tag{1.57}$$

As in the continuous-time setting, the assumption that $S(z)$ is positive real insures the boundedness of the functional J and the existence of the optimal control u^o. Also, using an argument similar to that of Proposition 1.26, it is not hard to check that

$$J(u^o;t,x^o(t)) = -\frac{1}{2}x^o(t)P(t)x^o(t), \tag{1.58}$$

where x^o is the solution of (1.57) corresponding to u^o and $J(u;t,x(t))$ is the *value function* defined by

$$J(u;t,x(t)) = -\frac{1}{2}x(0)'P_0x(0) + \frac{1}{2}\sum_{k=0}^{t-1}[u(k+1)'Tu(k+1) + 2x(k+1)'Gx(k+1)]. \tag{1.59}$$

As in Section 1.4 we are misusing notations somewhat by denoting u^o restricted to $[0,t]$ u^o also. Again, the optimal control u^o can be obtained from the derivation of (1.58); however, we shall resort to the Hamiltonian. To exploit the analogy with the continuous-time problem, we shall use the Maximum Principle of [26] with the *Hamiltonian function*

$$H(t,x(t+1),\, u(t+1),\, y(t)) = \frac{1}{2}\, u(t+1)'Tu(t+1) + x'(t+1)Gu(t+1) + y(t)'[x(t+1) - F'x(t+1) - H'u(t+1)], \tag{1.60}$$

where $y(t) - y(t+1) = \frac{\partial H}{\partial x(t+1)}(t,x^o(t+1),\, u^o(t+1),\, y(t))$, i.e.

$$y(t+1) = Fy(t) - G\,u^o(t+1) \quad ; \quad y(0) = P_o x^o(0), \tag{1.61a}$$

u^o and x^o are as above. Note that then

$$x^o(t+1) - x^o(t) = \frac{\partial H}{\partial y(t)}(t,x^o(t+1),\, u^o(t+1),\, y(t)).$$

Hence, with this formulation, there is a complete analogy between the discrete- and the continuous-time settings just exchanging derivatives for differences.

Now the Maximum Principle states that $H(t,x^o(t+1),\, u^o(t+1),\, y(t))$ has a minimum for $u = u^o(t)$ i.e. by differentiation

$$Tu^o(t+1) + G'x^o(t+1) - Hy(t) = 0.$$

which implies

$$u^o(t+1) = -T^{-1}[G'x^o(t+1) - Hy(t)]. \tag{1.61b}$$

Using relations (1.57) and (1.61), some straight-forward algebraic manipulations yield the result that the 2n-vector $\begin{bmatrix} x^o \\ y \end{bmatrix}$ satisfies

$$\begin{bmatrix} x^o(t) \\ y(t+1) \end{bmatrix} = \mathcal{F} \begin{bmatrix} x^o(t+1) \\ y(t) \end{bmatrix} \quad ; \quad \begin{bmatrix} x^o(t_1) \\ y(0) \end{bmatrix} = \begin{bmatrix} a \\ P_o x^o(0) \end{bmatrix}, \tag{1.62}$$

where the matrix F is given by

$$F = \begin{bmatrix} A' & H'T^{-1}H \\ GT^{-1}G' & A \end{bmatrix} \tag{1.63a}$$

and

$$A = F - GT^{-1}H. \tag{1.63b}$$

Then, as in the continuous-time setting

$$x^o(t) = X(t)x_o \tag{1.64a}$$

$$y(t) = Y(t)x_o \tag{1.64b}$$

where $x_o = x^o(0)$ and X and Y are $n \times n$ matrix functions satisfying

$$\begin{cases} A'X(t+1) = X(t) - H'T^{-1}HY(t) \quad ; \quad X(0) = I & (1.65a) \\ Y(t+1) = AY(t) + GT^{-1}G'X(t+1) \quad ; \quad Y(0) = P_o. & (1.65b) \end{cases}$$

To exploit the analogy with the continuous-time setting, the following lemma is needed.

Lemma 1.33. *Let* P *be any* $n \times n$ *symmetric matrix and let* $R := T - HPH'$ *be nonsingular. Then, the matrix* $[I - H'T^{-1}HP]$ *is full rank and its inverse is* $[I + H'R^{-1}HP]$.

Proof. It is easy to check that

$$[I - H'T^{-1}HP][I + H'R^{-1}HP] = I.$$

Hence, the two matrices in the left hand side are full rank. □

Lemma 1.34. *Let* $\left[\begin{smallmatrix}X\\Y\end{smallmatrix}\right]$ *be the solution of* (1.65) *and let* P *be the solution of the Riccati equation* (1.55). *Then* $Y(t) = P(t)X(t)$.

Proof. Replace Y by PX in (1.65b) to obtain

$$P(t+1)X(t+1) = AP(t)x(t) + GT^{-1}G'X(t+1)$$

Then, use Lemma 1.33 to obtain X(t) from (1.65a) with Y(t) set equal to P(t)X(t). Then inserting this into the above equation, we obtain

$$[P(t+1) - P(t) - \Lambda(P(t))]X(t+1) = 0$$

which, in view of (1.55) is an identity. Hence the lemma follows. □

As a first corollary to the above two lemmas, we have

Corollary 1.35. *Let* X *and* P *be as in Lemma* 1.34. *Then the matrix* X(t) *is nonsingular for all* $t \in Z^+$. *Moreover, the matrix* A *is nonsingular.*

Proof. From (1.65a), we have $A'X(1) = [I - H'T^{-1}HP_o]$, which by Lemma 1.33 is full rank since $R(0) := T - HP_oH'$ is nonsingular. Hence A and X(1) are also. The result now follows by repeating this argument for t=1,2,... . □

As another corollary, we obtain the counterpart of Proposition 1.27.

Proposition 1.36. *Let* $\left[\begin{smallmatrix}X\\Y\end{smallmatrix}\right]$ *be the solution of the system*

$$\begin{bmatrix} X(t+1) \\ Y(t+1) \end{bmatrix} = \begin{bmatrix} A'^{-1} & -A'^{-1}H'T^{-1}H \\ GT^{-1}G'A'^{-1} & A-GT^{-1}G'A'^{-1}H'T^{-1}H \end{bmatrix} \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}; \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} = \begin{bmatrix} I \\ P_o \end{bmatrix} \tag{1.66a}$$

Then $P(t) := Y(t)X(t)^{-1}$ *is the solution of* (1.55) *and*

$$y(t) = P(t)x^o(t). \tag{1.66b}$$

In analogy with the continuous-time setting, define the $n \times n$-matrix

$$M(t) = X(t+1)'\delta P_{t+1}X(t+1) \tag{1.67}$$

where $X(t)$ is given by (1.64) and (1.66a) and $\delta P_{t+1} := P(t+1) - P(t)$. Just as in Section 1.4, we want to express the optimal sequence $\{H^o(t);\ t = 0,1,\ldots,t_1\}$ defined by

$$H^o(t) := H(t, x^o(t+1), u^o(t+1), y(t)) \tag{1.68}$$

in terms of M.

Proposition 1.37. *Let* H^o *and* M *be given by* (1.68) *and* (1.67) *respectively. Then*

$$H^o(t) = -\frac{1}{2}\, x_o' M(t) x_o - \frac{1}{2}\, \delta x^{o\prime}(t+1)P(t)\delta x^o(t+1) \tag{1.69}$$

where $x^o = x^o(0)$ *and* $\delta x^o(t+1) = x^o(t+1) - x^o(t)$.

Proof. It is easy to see that

$$H^o(t) = J(u^o; t+1, x^o(t+1)) - J(u^o; t, x^o(t)) + y(t)'(x^o(t+1) - x^o(t)),$$

where $J(u;t,x(t))$ is given by (1.59). Remember that u^o restricted to $[0,t]$ minimizes $J(u;t,x^o(t))$. Hence, by (1.58) and (1.66b)

$$\begin{aligned} H^o(t) = {} & \frac{1}{2}\, x^o(t)'P(t)x^o(t) - \frac{1}{2}\, x^o(t+1)'P(t+1)x^o(t+1) - x^o(t)'P(t)x^o(t) \\ & + x^o(t)'P(t)x^o(t+1), \end{aligned}$$

which, upon adding and subtracting the quantity $\frac{1}{2}\, x^o(t+1)'P(t)x^o(t+1)$ and using (1.64) and (1.67), yields (1.69). □

However, unlike the continuous-time setting, the sequence $H^o(t)$ is not constant, nor is the matrix function M. We shall now use the Riccati equation (1.55) to get an alternative expression for the Hamiltonian sequence $H^o(t)$.

Lemma 1.38. *Let* H^o *and* M *be given by* (1.68) *and* (1.67). *Then*

$$H^o(t) = -\frac{1}{2} x_o'[M(t-1) - M(t-1)X(t)^{-1}H'R(t-1)^{-1}H(X(t)')^{-1}M(t-1)]x_o'$$
$$- \frac{1}{2}\delta x^o(t+1)'P(t)\delta x^o(t+1). \tag{1.70}$$

Proof. First, using (1.60), we may write

$$H^o(t) = \frac{1}{2} u^o(t+1)'Tu^o(t+1) + x^o(t+1)'Gu^o(t+1) - y(t)'x^o(t)$$
$$+ y(t)'x^o(t+1).$$

Upon inserting (1.61b) and (1.66b) into this equation, we obtain

$$H^o(t) = -\frac{1}{2} x^o(t)'P(t)x^o(t) - \frac{1}{2} x^o(t+1)'GT^{-1}G'x^o(t+1)$$
$$+ \frac{1}{2} x^o(t)'P(t)H'T^{-1}HP(t)x^o(t) - \frac{1}{2}x^o(t)'P(t)x^o(t) + x^o(t)'P(t)x^o(t+1).$$

Next, it is not hard to check that the Riccati equation (1.55) can be reformulated to read

$$P(t) - A[P(t-1) + P(t-1)H'R(t-1)^{-1}HP(t-1)]A' = GT^{-1}G',$$

where A is given by (1.63b). Inserting this value for $GT^{-1}G'$ into the above, we get

$$H^o(t) = -\frac{1}{2} x^o(t)'P(t)x^o(t) + \frac{1}{2} x^o(t+1)'A[P(t-1)$$
$$+ P(t-1)H'R(t-1)^{-1}HP(t)]A'x^o(t+1)' + \frac{1}{2} x^o(t)'P(t)H'T^{-1}HP(t)x^o(t)$$
$$- \frac{1}{2}\delta x^o(t+1)'P(t)\delta x^o(t+1).$$

Finally, it is not difficult to check, using (1.61b) and (1.57) that

$$x^o(t+1) = A'^{-1}[I - H'T^{-1}HP(t)]x^o(t) \tag{1.71}$$

which inserted into the above equation for $H^o(t)$ yields (1.70). □

The above lemma, together with Proposition 1.37, provide us with a recursion for M.

Lemma 1.39. *Let* M *be given by* (1.67). *Then* M *satisfies*

$$M(t) = M(t-1) - M(t-1)X(t)^{-1}H'R(t-1)^{-1}H(X(t)')^{-1}M(t-1). \tag{1.72}$$

Proof. Using the same argument as in the proof of Lemma 1.31, this follows from the fact that $\Gamma(t)$ is nonsingular for all t, which follows from Lemma 1.40 below. □

Before we state the main results, we shall need the following

Lemma 1.40. *Let* X(t) *be as in* (1.64) *and* (1.65). *Then*

$$(X(t+1)')^{-1} = \Gamma(t)(X(t)')^{-1}, \tag{1.73}$$

where Γ *is the feedback matrix* (1.31).

Proof. Upon applying the matrix inversion lemma (1.34) to (1.65a), in view of Proposition 1.36, one obtains

$$(X(t+1)')^{-1} = [F + FP(t)H'R(t)^{-1}H - G(T^{-1} + T^{-1}HP(t)H'R(t)^{-1})H](X(t)')^{-1}. \tag{1.74}$$

The last term in (1.74) can be written

$$-G(T^{-1}R(t) + T^{-1}HP(t)H')R(t)^{-1}H(X(t)')^{-1},$$

which is $-GR(t)^{-1}H(X(t)')^{-1}$. □

Analogously to the continuous-time case, let $r := \operatorname{rank} M(0)$; then $M(0) = X(1)'\delta P_1 X(1)$ can be written NSN', where N is $n \times r$ and S is $r \times r$.

Theorem 1.41. *Let* $\{P(t);\ t \in Z^+\}$ *be the solution of* (1.55) *and let* N *and* S *be as above. Then* P *can be determined from the system*

$$P(t+1) = P(t) - Q(t)Z(t)Q(t)' \quad ; \quad P(0) = P_0 \qquad (1.75a)$$

where the matrix sequences $\{Q(t);\ t \in Z^+\}$, $\{Z(t);\ t \in Z^+\}$ *are generated by*

$$Q(t+1) = [F - U(t+1)R(t+1)^{-1}H]Q(t) \quad ; \quad Q(0) = \Gamma(0)N \qquad (1.75b)$$

$$U(t+1) = U(t) + FQ(t)Z(t)Q(t)'H' \quad ; \quad U(0) = G - FP_0H' \qquad (1.75c)$$

$$R(t+1) = R(t) + HQ(t)Z(t)Q(t)'H' \quad ; \quad R(0) = J + J' - HP_0H' \qquad (1.75d)$$

$$Z(t+1) = Z(t) + Z(t)Q(t)'H'R(t)^{-1}HQ(t)Z(t) \quad ; \quad Z(0) = -S. \qquad (1.75e)$$

Proof. Let $Q(t) := (X(t+1)')^{-1}N$ and $U(t) := G - FP(t)H'$. Then by (1.73), $Q(t+1) = \Gamma(t+1)Q(t)$ which, by (1.31) and the definition of U, yields (1.75b). Next, in view of (1.72) and the fact that $M(0) = NSN'$, it can be easily seen that $M(t) = -NZ(t)N'$. Then (1.75a) follows from (1.67). Finally, (1.75c) and (1.75d) follow from (1.75a) and the definitions of U and R. □

As an application of the factorization (1.75), consider the problem of determining the Kalman gain K given by (1.18e). In this case $\Pi_0 = 0$

and $\Lambda(0) = GT^{-1}G'$. Hence, we may choose $N = X'(1)GT^{-\frac{1}{2}}$ and $S = I$, in which case,

$$K(t) = U(t)R(t)^{-\frac{1}{2}} \quad ; \quad K(0) = GT^{-\frac{1}{2}}, \tag{1.76}$$

where $U(t)$, $R(t)$ are given by (1.75) with initial conditions $Q(0) = GR^{-\frac{1}{2}}$, $U(0) = G$, $R(0) = T$ and $Z(0) = -I$.

This version of the algorithm is the one originally presented by Lindquist [27, 83]. The general case can be found in [29], where the following factorization was used

$$\delta P_{t+1} = \Gamma(t)[\delta P_t - \delta P_t H' R(t-1)^{-1} H \delta P_t]\Gamma(t)'. \tag{1.77}$$

Relation (1.77) can be obtained from (1.72) by inserting (1.67) and noting that $(X(t+1)')^{-1}X(t)' = \Gamma(t)$.

1.6 Non-Riccati Algorithms Inside the Set B.

Each $P \in \mathcal{P}$ can be interpreted as the state covariance matrix (1.2) of the corresponding realization (1.14) [5]. Consequently, there is a minimum-variance (P_*) and a maximum-variance (P^*) realization for each of which $B_2 = 0$.

Faurre's Algorithms (1.16) show that the solutions $\Pi(t)$ and $\bar{\Pi}(t)^{-1}$ converge to P_* and P^* respectively as $t \to \infty$. However, these solutions start outside the set $\mathcal{P}$ (for $0 \notin \mathcal{P}$). In this section, P_* and P^* will be approached from inside $\mathcal{P}$. In particular, for a given $P_o \in \mathcal{P}_-$ ($P_o \in \mathcal{P}_+$), we shall construct a trajectory extending from P_o to P_* (from P_o to P^*) so that this trajectory is a totally ordered set of matrices satisfying (1.12). Such a result will enable us to construct a countable family of realizations of y, the state covariances of which are totally ordered, yielding a procedure to obtain a countable family of realizations without resort to the intermediate step of determining the auxilliary quantity P. Indeed, the (non-Riccati) factorization of the previous section will be the basis of this procedure.

The work presented in the rest of this chapter is the discrete-time version of Section 6 in [2]. The procedure is more complicated than its continuous-time counterpart. However, this is natural and is largely due to the fact that the matrix R(P) depends on P, while its counterpart in the continuous case does not.

First, let us start with the following

Lemma 1.42. *Let* $\Lambda(P)$ *be defined by* (1.15). *Then, for each* $P_o \in \mathcal{P}$, *the solution* $\{P(i);\ i \in Z^+\}$ *of the matrix difference equation*

$$P(i+1) - P(i) = \Lambda(P(i)) \quad ; \quad P(0) = P_o \tag{1.78}$$

satisfies (*i*) $P(i) \in \mathcal{P}$ *for all* $i \in Z^+$, (*ii*) $P(i_2) \le P(i_1)$ *for* $i_1 \le i_2$ *and* (*iii*) *if* $P_o \in \mathcal{P}_-$, $P(i) \to P_*$ *as* $i \to \infty$.

Proof. Since $P_o \in$ `, $\Lambda(P_o) = -B_2B_2'$ [see (1.35)]. Then P(i) satisfies (1.75) with $N = X(1)'B_2$ and $S = -I$. Then $Z(0) = I$ and consequently it follows that $Z(i) \ge 0$ for all $i \in Z^+$. Therefore, in view of (1.75a), $\delta P_{i+1} = \Lambda(P(i)) \le 0$. Hence $P(i) \in \mathcal{P}$ for all $i \in Z^+$. This proves (i) and (ii).

To prove (iii), let $M_i = P^* - P(i)$. An argument similar to that used in proving Lemma 1.18 yields that

$$M_{i+1} = \Gamma^* M_i \Gamma^{*\prime} - \Gamma^* M_i H' R_i^{-1} H M_i \Gamma^{*\prime}; \; M_o = P^* - P_o.$$

Since $M_o > 0$ (for $P_o \in P_-$) and $M_{i+1} - M_i \geq 0$ by (1.75a), $M_i > 0$ $\forall\, i \in Z^+$. Consequently M_i^{-1} exists. Let $M^*(N)$ be as defined in Theorem 1.19. Define $V_i := M^*(0) - M_i^{-1}$. Then $V_{i+1} - V_i = -(M_{i+1}^{-1} - M_i^{-1})$. By (1.34), we have

$$M_{i+1}^{-1} = \Gamma^{*\prime -1} M_i^{-1} \Gamma^{*-1} + \Gamma^{*\prime -1} H'(R_i - HM_i H')^{-1} H\Gamma^{*-1}$$

$$= \Gamma^{*\prime -1} M_i^{-1} \Gamma^{*-1} + \Gamma^{*\prime -1} H' R^{*-1} H\Gamma^{*-1}.$$

Consequently,

$$V_{i+1} - V_i = M_i^{-1} - \Gamma^{*\prime -1} M_i^{-1} \Gamma^{*-1} - \Gamma^{*\prime -1} H' R^{*-1} H\Gamma^{*-1}.$$

But by (1.32b), $M^*(0)$ satisfies

$$-M^*(0) + \Gamma^{*\prime -1} M^*(0) \Gamma^{*-1} + \Gamma^{*\prime -1} H' R^{*-1} H\Gamma^{*-1} = 0.$$

Therefore

$$V_{i+1} - V_i = -V_i + \Gamma^{*\prime -1} V_i \Gamma^{*-1} \quad ; \quad i \in Z^+.$$

Since $|\lambda\{(\Gamma^*)^{-1}\}| < 1$, $V_i \to 0$ as $i \to \infty$ and consequently $M_i \to [M^*(0)]^{-1} = P^* - P_*$. Hence, $P(i) \to P_*$ as $i \to \infty$. □

Now, we are ready to state the first main result of this section: the non-Riccati algorithm. Since the realizations are determined by the matrix B, the algorithm will be given in terms of this parameter.

Let B be the set of all $B = (B_1, B_2)$ given by (1.35) with $P \in P$. Let B_o, B_- and B_+ be defined analogously in terms of P_o, P_- and P_+. It is clear that $B_o = \{B \in B \mid B_2 = 0\}$. In particular, let B_* and B^* denote those elements of B_o corresponding to P_* and P^* respectively.

Theorem 1.43. *Let* $[F, B_o, H, (R_o^{\frac{1}{2}}, 0)]$ *be an arbitrary realization of* y, *and, for each* $i \epsilon Z^+$, *let* $B(i) = [B_1(i), B_2(i)]$ *be given by*

$$B_1(i) = U(i)R(i)^{-\frac{1}{2}} \tag{1.79a}$$

$$B_2(i) = Q(i)Z(i)^{\frac{1}{2}}, \tag{1.79b}$$

where the matrix sequences $U(i)$, $Q(i)$, $Z(i)$ and $R(i)$ *are generated by* (1.75) *with initial conditions* $U(0) = (B_o)_1 R_o^{\frac{1}{2}}$, $R_o = R_o$, $Q(0) = (B_o)_2$ *and* $Z(0) = I$. *For each* $i \epsilon Z^+$ *let* $P(i)$ *be the solution of*

$$- P + FPF' + B(i)B(i)' = 0 \tag{1.80}$$

Then, for all $i \epsilon Z^+$, $[F, B(i), H, (R^{\frac{1}{2}}(i), 0]$ *is a realization of* y, *with state covariance* $P(i)$. *Moreover, if* $B_o \epsilon B_-$, $B(i) \to (B_*, 0)$ *as* $i \to \infty$. *Finally, the sequence* $\{P(i); i \epsilon Z^+\}$ *satisfies conditions* (i)-(iii) *of Lemma* 1.42 *and the difference equation*

$$P(i + 1) - P(i) = - B_2(i)B_2(i)'. \tag{1.81}$$

Proof. Let P_o be the state covariance of the initial realization $[F, B_o, H, (R_o^{\frac{1}{2}}, 0)]$, and let $\{P(i); i \epsilon Z^+\}$ be the trajectory through P_o defined by Lemma 1.42. Then $P(i) \epsilon \mathcal{P}$ for all $i \epsilon Z^+$. Define $B_1(i)$ and $B_2(i)$ by

$$B_1(i) := [G - FP(i)H']R(i)^{-\frac{1}{2}}$$

and

$$B_2(i) := Q(i)Z(i)^{\frac{1}{2}}.$$

Then, since $P(i + 1) - P(i) = - Q(i)Z(i)Q(i)'$, (1.81) follows. From the proof of Theorem 1.41 we see that $U(i) = G - FP(i)H'$. Hence

(1.79a) follows. Equations (1.78) and (1.81) imply $B_2(i)B_2(i)' = -\Lambda(P(i))$, which together with the above definition of $B_1(i)$ yield (1.80). Since $|\lambda(F)| < 1$ and $(F, B(i))$ is controllable (for (F, B_o) is), the solution of (1.80) is unique, symmetric and positive definite. This fact, together with (1.80) and the definition of $B_1(i)$ insure that $(P(i), B(i))$ satisfies (1.12). Hence $[F, B(i), H, (R(i)^{\frac{1}{2}}, 0)]$ is a realization of y with state covariance $P(i)$. By Lemma 1.42, $P(i)$ satisfies conditions (i) - (iii). Finally, by the same lemma, $P(i) \rightarrow P_*$ as $i \rightarrow \infty$ if $P_o \in P_-$. Hence, if $B_o \in B_-$, $B_1(i) \rightarrow B_*$ as $i \rightarrow \infty$ and, in view of (1.81), $B_2(i)B_2(i) \rightarrow 0$ i.e. $B_2(i) \rightarrow 0$. □

Remark. Throughout this section, we have used the parameter i rather than t to stress the fact that this quantity has nothing to do with time.

The next task is to construct a sequence belonging to set P which is increasing (rather than decreasing) in i and which converges to P*. In the continuous-time case, this can be done using the same Riccati equation; the analogue of (1.78). Here, unfortunately, to achieve this, we shall have to follow an indirect procedure through a "backward" approach. To this end, let us review certain facts about backward realizations.

We would like to consider realizations of y that evolve backward in time of the form

$$\bar{x}(t-1) = \bar{A}\bar{x}(t) + \bar{B}\bar{w}(t) \tag{1.82a}$$

$$y(t) = \bar{C}\bar{x}(t) + \bar{D}\bar{w}(t). \tag{1.82b}$$

where $\bar{w}$ is a normalized white noise sequence such that, for each t, $\bar{w}(t)$ is uncorrelated to *future* (rather than past) values of x.

Hence, we can state the *backward wide sense stochastic realization problem* as follows: Given the spectral density Φ of y (or equivalently, the quadruplet [F, G, H, J]), determine all quadruplets $[\bar{A}, \bar{B}, \bar{C}, \bar{D}]$ with $\dim \bar{A} = n$ minimal and $|\lambda(\bar{A})| > 1$ such that the output y of (1.82b) has spectral density Φ.

The problem has been studied by Pavon [9] and earlier in [10, 11, 18]. We shall outline some of the results of [9, 11] here, since we shall need them below.

The backward realization problem is deterministic in nature and is equivalent to the *dual spectral factorization problem* considered by Anderson [30] and Faurre [11], which requires determining all minimal unstable factors $\bar{W}(z)$ of $\Phi(z)$. Since $\bar{W}(z^{-1})\bar{W}(z)' = \Phi(z)'$, this problem is equivalent to finding all minimal stable factors $\bar{W}(z^{-1})$ of $\Phi(z)'$. Consequently, we have reduced the problem to the one considered before. In fact all solutions to this problem are given by

$$[\bar{A}, \bar{B}, \bar{C}, \bar{D}] = [T^{-1}F'T,\ T^{-1}(\bar{B}_1, \bar{B}_2)V,\ G'T,\ (\bar{R}(\bar{P})^{\frac{1}{2}}, 0)V]$$

where T and V are as before, $\bar{B}_1$, $\bar{B}_2$, $\bar{P}$ and $\bar{R}(\bar{P})$ satisfy

$$\bar{P} = F'\bar{P}F + \bar{B}_1\bar{B}_1' + \bar{B}_2\bar{B}_2' \tag{1.83a}$$

$$H' = F'\bar{P}G + \bar{B}_1\bar{R}(\bar{P})^{\frac{1}{2}} \tag{1.83b}$$

$$\bar{R}(\bar{P}) = J + J' - G'\bar{P}G\ , \tag{1.83c}$$

and $\bar{P}$ is an $n \times n$ symmetric positive definite matrix.

Again, here it is no restriction to take T = V = I i.e. to consider backward realizations of the form

$$\bar{x}(t-1) = F'\bar{x}(t) + \bar{B}_1\bar{u}(t) + \bar{B}_2\bar{v}(t) \tag{1.84a}$$

$$y(t) = G'\bar{x}(t) + \bar{R}(\bar{P})^{\frac{1}{2}}\bar{u}(t) \tag{1.84b}$$

where $\bar{w} = [\begin{smallmatrix}\bar{u}\\ \bar{v}\end{smallmatrix}]$.

Analogous to what we have done before, let $\bar{\mathcal{P}}$ be the set of all solutions of (1.83), define the map

$$\bar{\Lambda}(\bar{P}) = -\bar{P} + F'\bar{P}F + (H' - F'\bar{P}G)\bar{R}(\bar{P})^{-1}(H' - F'\bar{P}G)', \tag{1.85}$$

and let $\bar{\mathcal{P}}_o = \{\bar{P} \mid \bar{\Lambda}(\bar{P}) = 0\}$. Then $\bar{\mathcal{P}} = \{\bar{P} \mid \bar{\Lambda}(\bar{P}) \le 0\}$ and it has the same properties as $\mathcal{P}$. Hence there exist two elements $\bar{P}_*$ and $\bar{P}^*$ in $\bar{\mathcal{P}}_o$ such that $\bar{P}_* \le \bar{P} \le \bar{P}^*$ for every $\bar{P} \in \bar{\mathcal{P}}$. It is well known [11, 18] that $\bar{\mathcal{P}}$ is related to $\mathcal{P}$ by $\bar{\mathcal{P}} = \{P^{-1} \mid P \in \mathcal{P}\}$. Thuse $\bar{P}_* = (P^*)^{-1}$ and $\bar{P}^* = (P_*^{-1})$. This also explains the choice of (1.16b, c) by Faurre.

In fact there is a one-one correspondence between forward and backward realizations. It was shown in [9] how to compute the backward elements $\bar{B}$ and $\bar{R}(\bar{P})$ from the knowledge of the forward ones. In the following proposition, for convenience, we also include the converse statement.

Proposition 1.44 (a)([9]). *Assume the quadruplet* $[F, B, H, (R(P)^{\frac{1}{2}}, 0)]$ *solves the forward problem. Set*

$$\bar{B} = -P^{-1}F^{-1}B(I - B'P^{-1}B)^{\frac{1}{2}}, \quad \textit{and} \tag{1.86a}$$

$$(\bar{R}(\bar{P})^{\frac{1}{2}}, 0) = [(R(P)^{\frac{1}{2}}, 0) - HF^{-1}B][I - B'P^{-1}B]^{\frac{1}{2}}. \tag{1.86b}$$

Then, $[F', \bar{B}, G', (\bar{R}(\bar{P})^{\frac{1}{2}}, 0)]$ *solves the backward one.*

(b) *Assume the quadruplet* $[F', \bar{B}, G', (\bar{R}(\bar{P})^{\frac{1}{2}}, 0)]$ *solves the backward problem. Set*

$$B = -\bar{P}^{-1}F'^{-1}\bar{B}(I - \bar{B}'\bar{P}^{-1}\bar{B})^{\frac{1}{2}} \quad \textit{and} \tag{1.86c}$$

$$(R(P)^{\frac{1}{2}}, 0) = [(\bar{R}(\bar{P})^{\frac{1}{2}}, 0) - G'F'^{-1}\bar{B}][I - \bar{B}'\bar{P}^{-1}\bar{B}]^{\frac{1}{2}} \tag{1.86d}$$

Then, $[F, B, H, (R(P)^{\frac{1}{2}}, 0)]$ *solves the forward one.*

The following lemma is the backward counterpart of Lemma 1.42 and Theorem 1.43.

Lemma 1.45. (a) *Let* $\bar{\Lambda}(\bar{P})$ *be defined by* (1.85). *Then, for each* $\bar{P}_o \in \bar{\mathcal{P}}$, *the solution* $\{\bar{P}(i);\ i \in Z^+\}$ *of the matrix difference equation*

$$\bar{P}(i+1) - \bar{P}(i) = \bar{\Lambda}(\bar{P}(i)) \quad ; \quad \bar{P}(0) = \bar{P}_o \tag{1.87}$$

satisfies (i) $\bar{P}(i) \in \bar{\mathcal{P}}$ *for all* $i \in Z^+$ *, (ii)* $\bar{P}(i_2) \le \bar{P}(i_1)$ *for* $i_1 \le i_2$ *and (iii) if* $\bar{P}_o \in \bar{\mathcal{P}}_-$, $\bar{P}(i) \to \bar{P}_*$ *as* $i \to \infty$.

(b) *Let* $[F', \bar{B}_o, G', (\bar{R}_o^{\frac{1}{2}}, 0)]$ *be an arbitrary backward realization of* y, *and, for each* $i \in Z^+$, *let* $\bar{B}(i) = [\bar{B}_1(i), \bar{B}_2(i)]$ *be given by*

$$\bar{B}_1(i) = \bar{U}(i)\bar{R}(i)^{-\frac{1}{2}} \tag{1.88a}$$

$$\bar{B}_2(i) = \bar{Q}(i)\bar{Z}(i)^{\frac{1}{2}} \tag{1.88b}$$

where the matrix sequences $\bar{U}(i)$, $\bar{Q}(i)$, $\bar{Z}(i)$ and $\bar{R}(i)$ *are generated by*

$$\bar{Q}(i+1) = [F' - \bar{U}(i+1)\bar{R}(i+1)^{-1}G']\bar{Q}(i) \ ; \ \bar{Q}(0) = (\bar{B}_o)_2 \tag{1.88c}$$

$$\bar{U}(i+1) = \bar{U}(i) + F'\bar{Q}(i)\bar{Z}(i)\bar{Q}(i)'G \quad ; \ \bar{U}(0) = (\bar{B}_o)_1\bar{R}_o^{\frac{1}{2}} \tag{1.88d}$$

$$\bar{R}(i+1) = \bar{R}(i) + G'\bar{Q}(i)\bar{Z}(i)\bar{Q}(i)'G \quad ; \ \bar{R}(0) = \bar{R}_o \tag{1.88e}$$

$$\bar{Z}(i+1) = \bar{Z}(i) + \bar{Z}(i)\bar{Q}(i)'G\bar{R}(i)^{-1}G'\bar{Q}(i)\bar{Z}(i) \ ; \ \bar{Z}(0) = I \tag{1.88f}$$

For each $i \in Z^+$, *let* $\bar{P}(i)$ *be the solution of*

$$-\bar{P} + F'\bar{P}F + \bar{B}(i)\bar{B}(i)' = 0. \tag{1.89}$$

Then, for all $i \in Z^+$, $[F', \bar{B}(i), G', (\bar{R}(i)^{\frac{1}{2}}, 0)]$ *is a backward realization of* y, *with state covariance* $\bar{P}(i)$. *Moreover, if* $\bar{B}_0 \in \bar{\mathcal{B}}_-$, $\bar{B}(i) \to (\bar{B}_*, 0)$ *as* $i \to \infty$. *The sequence* $\{\bar{P}(i); i \in Z^+\}$ *is the same as in (a) and it also satisfies the difference equation*

$$\bar{P}(i+1) - \bar{P}(i) = -\bar{B}_2(i)\bar{B}_2(i)' . \tag{1.90}$$

Proof. (a) By an argument similar to that of Theorem 1.41, it can be shown that

$$\bar{P}(i+1) - \bar{P}(i) = -\bar{Q}(i)\bar{Z}(i)\bar{Q}(i)', \tag{1.91}$$

where $\bar{Q}$, $\bar{Z}$, together with $\bar{R}$ and $\bar{U}$ are given by (1.88). The rest of the proof is analogous to that of Lemma 1.42.

(b) Let $\bar{P}_0$ be the state covariance of the initial backward realization $[F', \bar{B}_0, G', (\bar{R}_0^{\frac{1}{2}}, 0)]$, and let $\{\bar{P}(i); i \in Z^+\}$ be the trajectory through P_0 defined by part (a). Then the proof of (b) follows upon defining $\bar{B}_1(i) := (H' - F'\bar{P}(i)G)\bar{R}(i)^{-\frac{1}{2}}$ and $\bar{B}_2(i) := \bar{Q}(i)\bar{Z}(i)^{\frac{1}{2}}$. □

Now, we are ready to state the second main result of this section. It provides us with a trajectory inside $\mathcal{P}$ converging to P^* and a non-Riccati procedure to generate a family of realizations, the state covariances of which are increasing.

Theorem 1.46. (a) *Let* $P_o \in P$ *and set* $\bar{P}_o := P_o^{-1}$. *Let* $\{P(i);\ i \in Z^+\}$ *be the solution of*

$$P(i+1) - P(i) = P(i)N(i)P(i) \quad ; \quad P(0) = P_o. \tag{1.92}$$

where

$$N(i) = \bar{Q}(i)[\bar{\bar{Z}}(i) - \bar{Q}(i)'P(i)\bar{Q}(i)]^{-1}\bar{Q}(i)', \tag{1.93}$$

and $\bar{Q}(i)$ *and* $\bar{\bar{Z}}(i)$ *are given by* (1.88). *Then* (*i*) $P(i) \in P\ \forall i \in Z^+$, (*ii*) $P(i_1) \le P(i_2)$ *for* $i_1 \le i_2$ *and* (*iii*) *if* $P_o \in P_+$, $P(i) \to P^*$ as $i \to \infty$.

(b) *Let* $[F,\ B_o,\ H\ (R_o^{\frac{1}{2}},\ 0)]$ *be an arbitrary realization of* y, *and for each* $i \in Z^+$, *let* $[F',\ \bar{B}(i),\ G',\ (\bar{R}(i)^{\frac{1}{2}},0)]$ *be the family of realizations of Lemma 1.45, having state covariance* $\bar{P}(i)$. *Let*

$$B(i) := -\bar{P}(i)^{-1}F'^{-1}\bar{B}(i)(I - \bar{B}(i)'\bar{P}(i)^{-1}\bar{B}(i))^{\frac{1}{2}}, \tag{1.94a}$$

$$(R(i)^{\frac{1}{2}},\ 0) := [(\bar{R}(i)^{\frac{1}{2}},0) - G'F'^{-1}\bar{B}(i)][I - \bar{B}(i)'\bar{P}(i)^{-1}\bar{B}(i)]^{\frac{1}{2}}. \tag{1.94b}$$

Then, $[F,\ B(i),\ H,(R(i)^{\frac{1}{2}},0)]$ *is a realization of* y *with state covariance* $P(i) = \bar{P}(i)^{-1}$. *Moreover, if* $B_o \in B_+$, $B(i) \to (B^*,\ 0)$ *as* $i \to \infty$. *Finally, the sequence* $\{P(i);\ i \in Z^+\}$ *is the same as in (a) and it also satisfies the difference equation*

$$P(i+1) - P(i) = -B_2(i)B_2(i)'.$$

Proof. (a) Let $P_o \in P$. Then $\bar{P}_o := P_o^{-1} \in \bar{P}$. Let $\{\bar{P}(i);\ i \in Z^+\}$ be the sequence generated by (1.91) and (1.88) corresponding to $\bar{P}_o$. Set $P(i) := \bar{P}(i)^{-1}$ for each $i \in Z^+$. Then, by (1.91), $P(i+1)^{-1} = P(i)^{-1} - \bar{Q}(i)\bar{Z}(i)\bar{Q}(i)'$. Using (1.34), we obtain (1.92) and (1.93). Since $\bar{P}(i) \in \bar{P} = \{P^{-1} \mid P \in P\}$, $P(i) \in P$ for all $i \in Z^+$. This proves (i).

By condition (ii) of Lemma 1.45, $\bar{P}(i_2) \le \bar{P}(i_1)$ for $i_1 \le i_2$, from which $P(i_1) \le P(i_2)$ for $i_1 \le i_2$ follows. Finally, if $P_o \in P_+$, Then $\bar{P}_o \in \bar{P}_-$. Therefore, by condition (iii) of Lemma 1.45, $\bar{P}(i) \to \bar{P}_*$ i.e. $P(i)^{-1} \to P*^{-1}$ as $i \to \infty$.

(b) The proof of this part is an immediate consequence of Proposition 1.44. □

Remarks. (1) Theorems 1.43 and 1.46 have the following interpretation.

Let

$$x_o(t+1) = Fx_o(t) + (B_o)_1u(t) + (B_o)_2v(t) \tag{1.95a}$$

$$y(t) = Hx_o(t) + R_o^{\frac{1}{2}}u(t) \tag{1.95b}$$

be an arbitrary (wide sense) realization of y with state covariance P_o.

(a) Let $B(i) = [B(i), B_2(i)]$ and $R(i)$ be given by (1.79) and (1.75). Then for each $i \in Z^+$,

$$x_i(t+1) = Fx_i(t) + B_1(i)u(t) + B_2(i)v(t) \tag{1.96a}$$

$$y(t) = Hx_i(t) + R(i)^{\frac{1}{2}}u(t) \tag{1.96b}$$

is a realization of y with state covariance $P(i) = E\{x_i(t)x_i(t)'\}$ given by

$$P(i+1) - P(i) = -B_2(i)B_2(i)' \quad ; \quad P(0) = P_o. \tag{1.97}$$

Furthermore, $\{P(i); i \in Z^+\}$ is a *decreasing* sequence in i such that, if $B_o \in B_-$, $P(i) \to P_*$ and $B(i) \to (B_*, 0)$ as $i \to \infty$, where B_* is the Kalman gain of the steady-state Kalman-Bucy filter (1.17).

(b) Let $B(i) = [B_1(i), B_2(i)]$ and $R(i)$ be given by (1.94). Then (1.96) is a realization of y with state covariance $P(i)$ given by (1.97). But now, the sequence $\{P(i);\ i \in Z^+\}$ is *increasing* in i such that, if $B_o \in \mathcal{B}_+$, $P(i) \to P^*$ and $B(i) \to (B^*, 0)$ as $i \to \infty$. In fact, B^* is the forward counterpart (in the sense of Proposition 1.44) of the gain of the steady-state *backward* Kalman-Bucy filter.

(2) As an application of Theorems 1.43 and 1.46, we can exploit the equivalence between dynamical systems in the discrete- and the continuous-time settings summarized in Proposition 1.21, to construct a discrete trajectory inside the "continuous" set $\mathcal{P}$ as defined in Section 1.1 and to generate families of realizations for the continuous-time problem via difference equations rather than differential ones.

To this end, suppose we are given the *continuous-time quadruplet* $[F, G, H, R]$. Let $[F_d, G_d, H_d, J_a]$ be defined by (1.37). Then, as we have seen in Section 1.3, $\mathcal{P}^d = \mathcal{P}$. For each $P \in \mathcal{P}$ define

$$\Lambda_d(P) = -P + F_d P F_d' + (G_d - F_d P H_d') R_d^{-1} (G_d - F_d P H_d')',$$

where $R_d = J_d + J_d' - H_d P H_d'$.

For each $P_o \in \mathcal{P}$, the solution $\{P(i);\ i \in Z^+\}$ of

$$P(i+1) - P(i) = \Lambda_d(P(i)) \quad ; \quad P(0) = P_o$$

satisfies conditions (i) - (iii) of Lemma 1.42. Also, the analogue of Theorem 1.43 holds using the quadruplet $[F_d, G_d, H_d, J_d]$.

CHAPTER 2

SMOOTHING FOR LINEAR DISCRETE-TIME STOCHASTIC SYSTEMS IN THE CONTEXT OF STOCHASTIC REALIZATION THEORY

2.1. Introduction

The linear least-squares estimation problem is of great importance in stochastic systems theory. The classical results on this subject, which were started in frequency domain language are primarily due to Kolmogrov [31] and Wiener [32]. However, here we shall be concerned with the state space formulations introduced by Kalman [33], Kalman and Bucy [34], and others.

The problem deals with estimating the state of a given system from noisy measurements. It can be classified into three categories: given a past record of data, estimating current values of the state (filtering), future values (prediction) and past values (smoothing). In this chapter, it is the last category that we are interested in.

The smoothing problem has received considerable attention in the literature in the last few years [35-51,61]. (See the survey paper [52] for further references.) Originally, our interest in this problem was caused by the well-known two-filter formula due to Mayne [38] and Fraser [39], on which topic a large number of papers had been written [40-42, 45-51], and which, nevertheless, we think had not received a satisfactory stochastic interpretation. This led to our continuous-time papers [64,65]. Here, we shall give the discrete-time version of this theory.

Let $\{x(t);\ t \in [0,T+1]\}$ and $\{y(t);\ t \in [0,T]\}$ (here $[0,T] := 0, 1, \ldots T$) be two stochastic processes of dimensions n and m respectively, defined as the solution of the linear stochastic system

$$(S)\ \begin{cases} x(t+1) = F(t)x(t) + B(t)w(t) \quad ; \quad x(0) = \xi & (2.1a) \\ \quad\ \ y(t) = H(t)x(t) + D(t)w(t) \quad ; \quad y(0) = 0\ , & (2.1b) \end{cases}$$

where w is a p-dimensional $(p \geq m)$, zero mean white noise sequence satisfying

$$E\{w(t)\} = 0 \quad \text{and} \quad E\{w(t)w(s)'\} = I\delta_{ts}\ , \tag{2.2}$$

and ξ is an n-dimensional, zero mean random vector with finite covariance matrix $N := E\{\xi\xi'\}$ and uncorrelated with w. The matrix $R(t) := D(t)D(t)'$ is positive definite for all $t \in [0,T]$, and F, B, H and D are time-varying matrices of dimensions compatible with x, y and w. Finally, $F(t)^{-1}$ exists for all $t \in [0,T]$.

The model S will be called a *linear stochastic system*; x is its *state* process, y its *output* process and w its *input* process.

The state covariance function $P(t) := E\{x(t)x(t)'\}$ clearly satisfies the Liapunov-type equation

$$P(t+1) = F(t)P(t)F(t)' + B(t)B(t)' \quad ; \quad P(0) = N\ . \tag{2.3}$$

The *fixed-interval smoothing problem* can now be stated as follows: for an arbitrary $t \in [0,T]$, find the linear least-squares estimate $\hat{x}(t)$ of $x(t)$ given $\{y(s);\ s \in [0,T]\}$ i.e., the wide-sense conditional expectation [70]

$$\hat{x}(t) = \hat{E}\{x(t) \mid y(s) \quad ; \quad s \in [0,T]\} \ . \tag{2.4}$$

In this chapter, we shall study this problem from a new angle, our aim being to develop a *unified theory* which, we feel, the literature is still lacking. It is true that some authors [49] have attempted to do so; nevertheless, we feel those attempts are not satisfactory. Many problems of interpretation of the existing solutions have remained unresolved. The approach we follow to provide such a theory employs concepts and techniques from the *stochastic realization theory* developed in [2,9,10] and in [1,53-60]. The basic idea is to embed the given model S in a class of models $\mathcal{S}$ all having the same output process (not only the covariances are the same as in Chapter 1, but also the processes are equal for each t a.s.) and the same Kalman-Bucy filter. Such a representation is called a *proper* [2] *stochastic realization* of y to distinguish it from the wide sense realizations of Chapter 1. The class $\mathcal{S}$ will be shown to contain an element (S_*), which together with another element (S^*) which does not belong to $\mathcal{S}$ in general, contains all the information on y needed to estimate x.

We note that the model S we are considering is more general than the one usually encountered in the literature in that Bw and Dw may be correlated, i.e., $BD' \neq 0$. Secondly, as will be shown later, one of the major obstacles in developing this theory is the fact that P(t) is in general not positive definite. We could assume that the model S is *minimal* (see Section 2.3); however, this assumption will not guarantee the invertibility of P on the whole interval [0,T]. In our continuous-time papers [64,65], this obstacle was removed upon imposing the extra assumption that the system matrices are analytic functions, which implies that S is totally controllable. It is worth noting that in the smoothing literature, conditions to insure invertibility of P are either ignored all together [61] or mistakenly assumed to hold on the entire interval as a consequence of complete controllability (minimality) of S. This is clearly incorrect (see e.g. [63]). Hence, at times we shall apply the generalized Moore-Penrose pseudo-inverse $P^{\#}$ of P wherever $P \nsucc 0$, leading to certain nontrivial complications.

The organization of this chapter goes as follows. Section 2.2 is devoted to preliminaries. A strict sense version of some results on backward representations developed in Section 1.6 is presented. Our results are generalizations of those of [9] obtained in the stationary setting. In Section 2.3, the stochastic realization theory concepts will be developed and in Section 2.4, we present a discussion about the frame space, the importance of which is that it contains the smoothing estimate. Section 2.5 will be devoted to the model S^* mentioned above. In Section 2.6, we give a general formula for the smoothing estimate in terms of S_* and S^*. Section 2.7 is devoted to deriving some previously known formulas for the smoothing estimate and to interpret them in terms of our realization theory.

2.2. Preliminaries

Let H be the space of all centered stochastic variables with finite second order moments; H is a Hilbert space when endowed with the inner product $(\xi,\eta) = E\{\xi\eta\}$. If u is a p-dimensional stochastic vector process with components in H, define H(u) to be the subspace spanned by $\{u_1, u_2, \ldots, u_p\}$. Then, for each stochastic vector process $\{z(t); t \in [t_o,t_1]\}$ and $t \in [t_o,t_1]$, define $H_t(z)$ to be $H(z(t))$; $H(z)$, $H_t^-(z)$ and $H_t^+(z)$ will denote the closed linear hulls in H of all subspaces $H_s(z)$ such that $s \in [t_o,t_1]$, $[t_o,t]$ and $[t,t_1]$ respectively. Given $\eta \in H$ and a subspace $H_1 \subset H$, $\hat{E}\{\eta \mid H_1\}$ will be the orthogonal projection of η onto H_1, i.e., the *wide sense conditional mean* [70]. We shall write $\hat{E}\{\eta \mid u\}$ in place of $\hat{E}\{\eta \mid H(u)\}$. The process η is called a *wide sense Markov process* if $\hat{E}\{\eta(t) \mid H_s^-(\eta)\} = \hat{E}\{\eta(t) \mid \eta(s)\}$ for $s \le t$ or equivalently, $\hat{E}\{\eta(s) \mid H_t^+(\eta)\} = \hat{E}\{\eta(s) \mid \eta(t)\}$, i.e., the Markov property is independent of the time direction.

Let Φ be the transition function of F, i.e.,

$$\Phi(t+1,s) = F(t)\Phi(t,s) \quad ; \quad \Phi(s,s) = I \, .$$

Since the state process x defined by (2.1a) satisfies

$$x(t) = \Phi(t,s)x(s) + \sum_{j=s}^{t-1} \Phi(t,j+1)B(j)w(j) \tag{2.5a}$$

for $s \in [0,t-1]$ and, consequently, $H_s^+(w) \perp H_o(x) \oplus H_s^-(w) \supset H_s^-(x)$, it easily follows that $\hat{E}\{x(t) \mid H_s^-(x)\} = \hat{E}\{x(t) \mid x(s)\}$. Hence, x is a Markov process. However, the difference equation (2.1a) is not symmetric with time: the two terms in the right-hand member of (2.5a) are orthogonal if and only if $t > s$.

Of great importance in this chapter is a backward counterpart of (2.1a). By the above argument, such an equation cannot be readily obtained by merely reversing the time direction. In this section, we shall construct the backward version of (2.1a) under the assumption that $N > 0$; later in Section 2.5, we shall remove this assumption. Since the covariance matrix function P is given by

$$P(t) = \Phi(t,0)N\Phi(t,0)' + \sum_{j=0}^{t-1} \Phi(t,j+1)B(j)B(j)'\Phi(t,j+1)' , \tag{2.5b}$$

it follows that $P(t) > 0$ for all $t \in [0,T]$ if and only if $N > 0$.

In this case, the process

$$\bar{x}(t-1) = P(t)^{-1}x(t) \tag{2.6}$$

is well-defined for all $t \in [-1,T]$, with components in H. Let $\bar{P}$ denote its covariance function

$$\bar{P}(t-1) := E\{\bar{x}(t-1)\bar{x}(t-1)'\} . \tag{2.7}$$

Then, the backward version of (2.1a) is given by

Lemma 2.1. *Let* x *be the state process of the linear stochastic system* S *, and let* $N > 0$. *Then, the process* $\{\bar{x}(t);\ t \in [-1,T]\}$ *defined by* (2.6) *satisfies the backward recursion*

$$\bar{x}(t-1) = F(t)'\bar{x}(t) + \bar{B}(t)\bar{w}(t) \quad ; \quad \bar{x}(T) = \bar{\xi} \tag{2.8}$$

for $t \in [0,T]$, *where* $\bar{\xi} = P(T+1)^{-1}x(T+1)$,

$$\bar{B}(t) = -P(t)^{-1}F(t)^{-1}B(t)[I - B(t)'P(t+1)^{-1}B(t)]^{\frac{1}{2}} \tag{2.9}$$

and $\bar{w}$ *is a* p-*dimensional normalized white noise sequence satisfying* (2.2) *and the condition* $H_t^-(\bar{w}) \perp H_t^+(\bar{x})$ *for* $t \in [0,T]$ *and is given by*

$$\bar{w}(t) = [I - B(t)'P(t+1)^{-1}B(t)]^{\frac{1}{2}}[w(t) - B(t)'F'(t)^{-1}P(t)^{-1}x(t)].\quad(2.10)$$

The covariance function (2.7) *is given by*

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

and satisfies the Liapunov equation

$$\bar{P}(t-1) = F'(t)\bar{P}(t)F(t) + \bar{B}(t)\bar{B}(t)' \quad ; \quad \bar{P}(T) = P(T+1)^{-1} \qquad (2.11)$$

If $N \nsucc 0$, *equations* (2.8)-(2.11) *are defined for all* $t \in [0,T]$ *for which* $P(t) > 0$.

This lemma is a generalization of the results of Section 1.6 which were obtained in [9] for the stationary case. This is a strict sense version of the wide sense results presented in [48,50] which are insufficient for our purposes since they are deterministic rather than probabilistic in nature. Moreover, we have chosen to write the backward version (2.8) in terms of $\bar{x}$ rather than x, since this choice will yield a backward Kalman-Bucy filter which is invariant over the class S.

The proof of this lemma is based on the observation that, as the orthogonal decomposition

$$x(t+1) = \hat{E}\{x(t+1) \mid H_t^-(x)\} + [x(t+1) - \hat{E}\{x(t+1) \mid H_t^-(x)\}] \quad (2.12)$$

yields (2.1a), the orthogonal decomposition

$$x(t) = \hat{E}\{x(t) \mid H_{t+1}^+(x)\} + [x(t) - \hat{E}\{x(t) \mid H_{t+1}^+(x)\}] \qquad (2.13)$$

yields the backward version (2.8). (The basic idea of this proof first appeared in [2].) We shall need the following lemma, the proof of which can be found in most standard books on estimation theory.

Lemma 2.2. *Let* u *and* v *be two stochastic vectors with components in* H, *and assume* $E\{vv'\} > 0$. *Then,*

$$\hat{E}\{u \mid v\} = E\{uv'\}(E\{vv'\})^{-1}v \ .$$

Proof of Lemma 2.1. We shall prove the lemma for the case $N > 0$; if $N \not> 0$, everything will be the same for $t \in [0,T]$ such that $P(t)^{-1}$ exists. Since x is Markov,

$$\hat{E}\{x(t) \mid H^{+}_{t+1}(x)\} = \hat{E}\{x(t) \mid x(t+1)\} \ .$$

Upon using Lemma 2.2, the right-hand side is $P(t)F(t)'P(t+1)^{-1}x(t+1)$. Inserting this into (2.13) and multiplying by $P(t)^{-1}$ yields

$$\bar{x}(t-1) = F(t)'\bar{x}(t) + P(t)^{-1}\eta(t) \tag{2.14a}$$

where $\eta(t) := x(t) - \hat{E}\{x(t) \mid x(t+1)\}$. But, $x(t) = F(t)^{-1}x(t+1) - F(t)^{-1}B(t)w(t)$, and consequently

$$\begin{aligned} \eta(t) &= -F(t)^{-1}B(t)[w(t) - \hat{E}\{w(t) \mid x(t+1)\} \\ &= -F(t)^{-1}B(t)[w(t) - B(t)'P(t+1)^{-1}x(t+1)] \end{aligned} \tag{2.14b}$$

where we have used Lemma 2.2 to obtain the last relation. Now, from (2.3), it is not hard to see that

$$B(t)'P(t+1)^{-1}F(t) = [I - B(t)'P(t+1)^{-1}B(t)]B(t)'F(t)'^{-1}P(t) \ .$$

Then inserting this and (2.1a) into (2.14b), we obtain

$$\eta(t) = -F(t)^{-1}B(t)[I - B(t)'P(t+1)^{-1}B(t)][w(t) - B(t)'F(t)'^{-1}P(t)x(t)] \ .$$

This together with (2.14a) yields (2.8) with $\bar{B}$ and $\bar{w}$ given by (2.9) and (2.10). From the above discussion, it follows that

$$[I - B(t)'P(t+1)^{-1}B(t)]^{\frac{1}{2}}\bar{w}(t) = w(t) - \hat{E}\{w(t) \mid H_{t+1}^{+}(x)\}$$

(cf. [9]), which relation implies that $\bar{w}$ is a white noise such that $H_t^-(\bar{w}) \perp H_{t+1}^+(x) = H_t^+(\bar{x})$; the factor in front of $\bar{w}(t)$ is the appropriate normalization factor so that $\bar{w}$ satisfies (2.2), as can be easily checked. Finally, $\bar{P}(t-1) = E\{\bar{x}(t-1)\bar{x}(t-1)'\} = P(t)^{-1}E\{x(t)x(t)'\}P(t)^{-1} = P(t)^{-1}$. Equation (2.11) is obtained from (2.8) precisely as (2.3) is obtained from (2.1a). □

2.3. Forward and Backward Realizations

Let the output process y be defined as in Section 2.1. Any system of type (2.1) (with $\xi_i \in H$, for $i = 1, 2, \ldots, n$, w satisfying (2.2) and $\xi_i \perp H(w)$ for all i) having y as its output is called a *realization* of y. Clearly, the components of x, y and w belong to H.

The purpose of this section is to introduce the two models S_* and $\bar{S}_*$, the knowledge of which determines the frame space (to be defined in Section 2.4), which in turn contains the smoothing estimate.

As we have seen in Section 1.1, the linear least-squares estimate

$$x_*(t) = \hat{E}\{x(t) \mid H_{t-1}^-(y)\} \tag{2.15}$$

of the state process x of S is generated on [0,T] by the *Kalman-Bucy filter*

$$x_*(t+1) = F(t)x_*(t) + B_*(t)R_*(t)^{-\frac{1}{2}}[y(t) - H(t)x_*(t)] \;;\; x_*(0) = 0, \tag{2.16a}$$

where the *Kalman gain* function B_* is given by

$$B_* = [FQ_*H' + BD']R_*^{-\frac{1}{2}}, \tag{2.16b}$$

$$R_* = HQ_*H' + DD' \tag{2.16c}$$

and the *error covariance matrix*

$$Q_*(t) = E\{[x(t) - x_*(t)][x(t) - x_*(t)]'\} \tag{2.16d}$$

is the solution of

$$Q_*(t+1) = F(t)Q_*(t)F(t)' - B_*(t)B_*(t)' + B(t)B(t)' \; ; \; Q_*(0) = N \tag{2.16e}$$

which is a matrix Riccati difference equation when (2.16b) is inserted.

As we shall see shortly, there are other realizations of y which have (2.16a) as their Kalman-Bucy filter. Hence define $\mathcal{S}$ to be the class of all realizations S of y such that $R(t) := D(t)D(t)' > 0 \;\forall\, t \in [0,T]$ and such that the corresponding Kalman-Bucy filter is given by (2.16a), that is, it has the same matrix functions F, H and $K_* := B_*R_*^{-\frac{1}{2}}$ as those of (2.16a) and consequently the same estimates $\{x_*(t);\ t \in [0,T+1]\}$.

The sequence $\{w_*(t);\ t \in [0,T]\}$ defined by

$$w_*(t) = R_*(t)^{-\frac{1}{2}}[y(t) - H(t)x_*(t)] \tag{2.17}$$

is called the *innovation process*. It is a normalized white noise satisfying (2.2) and characterized by the property $H_t^-(w_*) = H_t^-(y)$ for all $t \in [0,T]$. Combining (2.16a) and (2.17), we obtain the model

$$(S_*)\begin{cases} x_*(t+1) = F(t)x_*(t) + B_*(t)w_*(t) \quad ; \quad x_*(0) = 0 \\ \quad y(t) = H(t)x_*(t) + R_*(t)^{\frac{1}{2}}w_*(t) \;, \end{cases} \tag{2.18a}$$

which clearly belongs to $\mathcal{S}$. It can be immediately seen that the covariance matrix $P_*(t) := E\{x_*(t)x_*(t)'\}$ of $x_*(t)$ satisfies

$$P_*(t+1) = F(t)P_*(t)F(t)' + B_*(t)B_*(t)' \; ; \; P_*(0) = 0 \;, \tag{2.18b}$$

and that

$$Q_* = P - P_* . \qquad (2.18c)$$

It is essential at this point to show that S_* is uniquely defined regardless of the choice of the $S \in \mathcal{S}$ from which S_* was formed, i.e., that the matrices B_* and R_* are both invariants for the class $\mathcal{S}$ (by definition, F and H clearly are). To this end, we need to define the $n \times m$-matrix function

$$G = FPH' + BD' \qquad (2.19a)$$

for each realization $S \in \mathcal{S}$; P is its state covariance function.

Lemma 2.3. *Let* G, R_* *and* B_* *be defined by* (2.19a), (2.16c) *and* (2.16b) *respectively. Then* G, R_* *and* B_* *are invariants for the class* $\mathcal{S}$.

Proof. Let $S \in \mathcal{S}$ be arbitrary and let G be as in (2.19a). Then $P = Q_* + P_*$. Also, by (2.16b), $BD' = K_* - FQ_*H'$. Inserting these two relations into (2.19a), we obtain

$$G = FP_*H' + B_*R_*^{-\frac{1}{2}} , \qquad (2.19b)$$

which, by the definition of $\mathcal{S}$, is invariant over $\mathcal{S}$. Next, since $\Lambda_0(t) := E\{y(t)y(t)'\} = H(t)P(t)H(t)' + D(t)D(t)'$, $R_* = \Lambda_0 - HP_*H'$, which does not depend on the choice of S, for Λ_0 and P_* do not. Finally, since $B_* = K_*R_*^{\frac{1}{2}}$ (or $B_* = (G - FP_*H')R_*^{-\frac{1}{2}}$), B_* is also invariant over $\mathcal{S}$. □

Consequently, F, H, G, B_* and R_* are invariants for $\mathcal{S}$, whereas B, D, P, w and x will vary with different realizations $S \in \mathcal{S}$. Actually, even the dimension p of w will vary. However, since R is of full rank, we will always have $p \geq m$.

The model S_* belongs to a class of realizations for which p is minimal, i.e., $p = m$. Define $\mathcal{S}_0$ to be the subclass of all $S \in \mathcal{S}$ such that $p = m$ and $x(0) \in H(y)$. Let

$$(S_0)\ \begin{cases} x_0(t+1) = Fx_0(t) + B_0 w_0(t) \quad ; \quad x_0(0) = \xi_0 \\ \quad y(t) = Hx_0(t) + D_0 w_0(t) \end{cases} \tag{2.20}$$

be a realization in $\mathcal{S}_0$ with state covariance P_0. As D_0 is invertible,

$$x_0(t+1) = Fx_0(t) + B_0 D_0^{-1}[y(t) - Hx_0(t)] \ ; \ x_0(0) = \xi_0 \ . \tag{2.21a}$$

Let (2.1) be an arbitrary realization in $\mathcal{S}$ and define

$$Q_0 = P - P_0 \ . \tag{2.21b}$$

Then, by (2.20) and Lemma 2.3,

$$R_0 = D_0 D_0' = \Lambda_0 - HP_0 H' = HQ_0 H' + DD' \tag{2.21c}$$

and

$$B_0 = (G - FP_0 H')R_0^{-\frac{1}{2}} = (FQ_0 H' + BD')R_0^{-\frac{1}{2}} \ . \tag{2.21d}$$

Inserting (2.21d) into the equation (2.3) for P_0 and subtracting from (2.3), we conclude that Q_0 satisfies

$$Q_0(t+1) = F(t)Q_0(t)F(t)' - B_0(t)B_0(t)' + B(t)B(t)' \ ; \ Q_0(0) = N - N_0 . \tag{2.21e}$$

Equations (2.21) look formally like the filtering equations (2.16), only the initial conditions are different. In view of the assumption that $\xi_0 \in H(y)$, (2.21a) implies $H(x_0) \subset H(y)$. We shall call a realization $S \in \mathcal{S}$ *internal* if it satisfies the condition $H(x) \subset H(y)$ and *external* otherwise [2].

Therefore, by the above discussion, we have shown that all $S \in \mathcal{S}_0$ are internal.

Our next task is to derive a backward realization $\bar{S}$ for each $S \in \mathcal{S}$. We shall begin by restricting our attention to the subclass $\mathcal{S}_+$ consisting of all $S \in \mathcal{S}$ for which $N > 0$. The class $\mathcal{S}_+$ is nonempty. This can be seen by using an argument similar to that in the continuous-time case [64]. The basic idea is that the stochastic process y can be extended to an interval $[\tilde{t},T]$ where $\tilde{t} < 0$ so that the covariance matrix $\tilde{P}(t)$ of a realization $\tilde{S}$ of y on $[\tilde{t},T]$ is positive definite for $t \in [0,T]$. Hence, the restriction of $\tilde{S}$ to $[0,T]$ belongs to $\mathcal{S}_+$. (See [64] and [73] for details.)

Let $S \in \mathcal{S}_+$. Then, by Lemma 2.1, $\bar{x}(t-1) = P(t)^{-1}x(t)$ is defined for every $t \in [0,T]$ and satisfies (2.8). It remains to obtain a "backward" equation for y.

Lemma 2.4. *Let y be given by* (2.1b). *Then, y can be written*

$$y(t) = G'(t)\bar{x}(t) + \bar{D}(t)\bar{w}(t) \tag{2.22}$$

where G *is given by* (2.19) *and the* m × p-*matrix function* $\bar{D}$ *is given by*

$$\bar{D}(t) = [D(t) - H(t)F(t)^{-1}B(t)][I - B(t)'P(t+1)^{-1}B(t)]^{\frac{1}{2}} \tag{2.23}$$

Proof. Inserting (2.19a) into (2.1b), we get

$$y(t) = [G'(t)F'(t)^{-1}P(t)^{-1} - D(t)B(t)'F(t)'^{-1}P(t)^{-1}]x(t) + D(t)w(t) .$$

From (2.8),

$$x(t) = P(t)F(t)'P(t+1)^{-1}x(t+1) + P(t)\tilde{B}(t)\tilde{w}(t) .$$

Using this expression for x(t) in the above equation for y, we obtain after some lengthy algebraic manipulations

$$y(t) = G'(t)P(t+1)^{-1}x(t+1) - G'(t)P(t+1)^{-1}B(t)[w(t) - B(t)F'(t)^{-1}P(t)^{-1}x(t)]$$

$$+ D(t)w(t) - D(t)B(t)'F'(t)^{-1}P(t)^{-1}x(t).$$

$$= G'(t)P(t+1)^{-1}x(t+1)$$

$$+ [D(t) - G'(t)P(t+1)^{-1}B(t)][w(t) - B'(t)F'(t)^{-1}P(t)^{-1}x(t)] .$$

Using (2.19a) and adding and subtracting the quantity $H(t)F(t)^{-1}B(t)$, we get

$$D(t) - G(t)'P(t+1)^{-1}B(t) =$$

$$= D(t) - D(t)B(t)'P(t+1)^{-1}B(t) - H(t)F(t)^{-1}B(t) + H(t)F(t)^{-1}B(t)$$

$$- H(t)P(t)F(t)'P(t+1)^{-1}B(t) .$$

$$= D(t) - D(t)B(t)'P(t+1)^{-1}B(t) - H(t)F(t)^{-1}B(t)$$

$$+ H(t)^{-1}F(t)^{-1}[P(t+1) - F(t)P(t)F(t)']P(t+1)^{-1}B(t) .$$

$$= [D(t) - H(t)F(t)^{-1}B(t)][I - B(t)'P(t+1)^{-1}B(t)]$$

where in the last step, we employed (2.3). Then, using the definition (2.10) for $\tilde{w}$, the desired result follows. □

Combining (2.8) and (2.22), we obtain the following backward model

$$(\tilde{S}) \begin{cases} \tilde{x}(t-1) = F(t)'\tilde{x}(t) + \tilde{B}(t)\tilde{w}(t) \quad ; \quad \tilde{x}(T) = \tilde{\xi} \\ y(t) = G(t)'\tilde{x}(t) + \tilde{D}(t)\tilde{w}(t) \end{cases} \tag{2.24}$$

where $\bar{B}$, $\bar{w}$ and $\bar{D}$ are given by (2.9), (2.10) and (2.23) respectively and $\bar{\xi} = P(T+1)^{-1}x(T+1) \perp H(\bar{w})$. The state covariance function $\bar{P}(t) = P(t+1)^{-1}$ satisfies (2.11). We shall call any model of type (2.24) with y as its output, $\bar{\xi}_i \in H$, for $i = 1, 2, \ldots, n$, $\bar{w}$ satisfying (2.2) and $\bar{\xi}_i \perp H(\bar{w})$ for all i, a *backward realization* of y. Note that S and $\bar{S}$ have the same *state spaces*, i.e.,

$$H_t(x) = H_{t-1}(\bar{x}) \quad . \tag{2.25}$$

for each $t \in [0,T+1]$.

It is essential at this point to show that the matrix $\bar{R}(t) := \bar{D}(t)\bar{D}(t)'$ is positive definite wherever the matrix $R(t) := D(t)D(t)'$ is. To this end, define the $m \times m$-matrix functions

$$A = DD' - DB'F'^{-1}H' \tag{2.26}$$

$$\bar{A} = \bar{D}\bar{D}' - \bar{D}\bar{B}'F^{-1}G \quad . \tag{2.27}$$

Lemma 2.5. *Let* A *and* $\bar{A}$ *be defined by* (2.26) *and* (2.27) *respectively. Then* $\bar{A} = A'$. *Furthermore,* A *and* $\bar{A}$ *are both invariants for the class* S.

Proof. Inserting $DD' = \Lambda_0 - HPH'$ (by (2.16)) and $DB' = G' - HPF'$ (by (2.19a)) into (2.26), we obtain

$$A = \Lambda_0 - G'F'^{-1}H' \quad , \tag{2.28}$$

which is invariant for S. Using (2.19a), (2.9) and (2.23), it can be seen that the matrix function H can be written

$$H' = F'\bar{P}G + \bar{B}\bar{D}' \quad . \tag{2.29}$$

Now, by the above argument and (2.24), it easily follows that

$$\bar{A} = \Lambda_0 - HF^{-1}G \tag{2.30}$$

from which the lemma follows. □

Proposition 2.6. *Let* A *and* $\bar{A}$ *be defined by* (2.26) *and* (2.27) *respectively. Then, for each* $t \in [0,T]$, *the following statements are equivalent:* *(i)* $D(t)D(t)' > 0$, *(ii)* $A(t)$ *is nonsingular,* *(iii)* $\bar{A}(t)$ *is nonsingular, and* *(iv)* $\bar{D}(t)\bar{D}(t)' > 0$.

Proof. The equivalence between (i) and (ii) is proved by Pavon ([9;Theorem 3.2], [62]); this proof does not require stationarity. The same argument can be used to prove the equivalence between (iii) and (iv). Finally, the equivalence between (ii) and (iii) follows trivially from Lemma 2.5. □

In analogy with the forward setting, the least-squares estimate

$$\bar{x}_*(t) = \hat{E}\{\bar{x}(t) \mid H^+_{t+1}(y)\} \tag{2.31}$$

is generated by the *backward Kalman-Bucy filter:*

$$\bar{x}_*(t-1) = F(t)'\bar{x}_*(t) + \bar{B}_*(t)\bar{R}_*(t)^{-\frac{1}{2}}[y(t) - G(t)'\bar{x}_*(t)] \; ; \; \bar{x}_*(T) = 0, \tag{2.32a}$$

where

$$\bar{B}_* = [F'\bar{Q}_*G + \bar{B}\bar{D}']\bar{R}_*^{-\frac{1}{2}}, \tag{2.32b}$$

$$\bar{R}_* = G'\bar{Q}_*G + \bar{D}\bar{D}', \tag{2.32c}$$

(which, by Proposition 2.6, is positive definite for all $t \in [0,T]$) and the *error covariance matrix function*

$$\bar{Q}_*(t) = E\{[\bar{x}(t) - \bar{x}_*(t)][\bar{x}(t) - \bar{x}_*(t)]'\} \qquad (2.32d)$$

being the solution of the dual matrix Riccati equation

$$\bar{Q}_*(t-1) = F(t)'\bar{Q}_*(t)F(t) - \bar{B}_*(t)\bar{B}_*(t)' + \bar{B}(t)\bar{B}(t)' ; \bar{Q}_*(T) = P(T+1)^{-1}, \qquad (2.32e)$$

with $\bar{B}_*$ given by (2.32b).

The *backward innovation process* $\bar{w}_*$ defined by

$$\bar{w}_*(t) = \bar{R}_*(t)^{-\frac{1}{2}}[y(t) - G'(t)\bar{x}_*(t)] \qquad (2.33)$$

is a normalized white noise satisfying (2.2) and the condition $H_t^+(\bar{w}_*) = H_t^+(y)$. The covariance matrix $\bar{P}_*(t) := E\{\bar{x}_*(t)\bar{x}_*(t)'\}$ satisfies the backward Liapunov equation

$$\bar{P}_*(t-1) = F'(t)\bar{P}_*(t)F(t) + \bar{B}_*(t)\bar{B}_*(t)' \; ; \; \bar{P}_*(T) = 0 \qquad (2.34)$$

Again, we need to show the invariance of the backward filter.

Lemma 2.7. *Let* $\bar{R}_*$ *and* $\bar{B}_*$ *be given by* (2.32c) *and* (2.32b) *respectively. Then,*

$$\bar{R}_* = \Lambda_0 - G'\bar{P}_*G \qquad (2.35)$$

and

$$\bar{B}_* = (H' - F'\bar{P}_*G)\bar{R}_*^{-\frac{1}{2}} , \qquad (2.36)$$

i.e., $\bar{R}_*$, $\bar{B}_*$, *and hence the model* $\bar{S}_*$, *defined by* (2.37) *below, are all invariants for the class S.*

Proof. It is easy to see that (2.35) holds. As for $\bar{B}_*$, observe that it can be written

$$\bar{B}_* = (F'\bar{P}G + \bar{B}\bar{D}' - F'\bar{P}_*G)\bar{R}_*^{-\frac{1}{2}} ,$$

from which (2.36) follows, upon using (2.29). Hence $\bar{S}_*$ does not depend on $\bar{S}$ (consequently on the choice of $S \in S_+$.) □

Now, in the same way as above, define $\bar{\mathcal{S}}$ to be the class of all backward realizations $\bar{S}$ having (2.32) as their backward Kalman-Bucy filter, and let $\bar{\mathcal{S}}_+$ be the subclass consisting of those $\bar{S} \in \bar{\mathcal{S}}$ for which $\bar{N} := E\{\bar{\xi}\bar{\xi}'\} > 0$. In the same way as in the forward setting, it is seen that the realization

$$(\bar{S}_*) \begin{cases} \bar{x}_*(t-1) = F'(t)\bar{x}_*(t) + \bar{B}_*(t)\bar{w}_*(t) \ ; \ \bar{x}_*(T) = 0 \\ y(t) = G'(t)\bar{x}_*(t) + \bar{R}_*(t)^{\frac{1}{2}}\bar{w}_*(t) \end{cases} \tag{2.37}$$

belongs to $\bar{\mathcal{S}}$. By Lemma 2.7, the class $\bar{\mathcal{S}}$ is uniquely defined in terms of the invariants F, H, G and Λ_0, and therefore, the backward counterpart $\bar{S}$ of any $S \in S_+$ belongs to $\bar{\mathcal{S}}$. In particular, since P(T+1) is positive definite and since $\bar{P}(T) = P(T+1)^{-1}$, $\bar{S} \in \bar{\mathcal{S}}_+$. Also, note that, by Proposition 2.6, $\bar{D}\bar{D}' > 0$ for all $\bar{S} \in \bar{\mathcal{S}}_+$.

It is clear that there is a complete symmetry between forward and backward realizations. In particular, the subclasses S_+ and $\bar{\mathcal{S}}_+$ are in one-one correspondence. Therefore, in the following lemma, we summarize the procedure for constructing a forward realization corresponding to a backward one in $\bar{\mathcal{S}}_+$. This lemma is the counterpart of Lemma 2.1.

Lemma 2.8. *Let (2.24) be an arbitrary backward realization in $\bar{\mathcal{S}}_+$ with state process $\bar{x}$ and state covariance function $\bar{P}$. Then, the process x, defined by*

$$x(t) = \bar{P}(t-1)^{-1}\bar{x}(t-1) \tag{2.38}$$

for $t \in [0,T]$, *satisfies the forward recursion*

$$x(t+1) = F(t)x(t) + B(t)w(t) \ ; \ x(0) = \bar{P}(-1)^{-1}\bar{x}(-1), \tag{2.39a}$$

where

$$B(t) = -\bar{P}(t)^{-1}F(t)'^{-1}\bar{B}(t)[I - \bar{B}(t)'\bar{P}(t-1)^{-1}\bar{B}(t)]^{\frac{1}{2}} \tag{2.40}$$

and w *is a* p*-dimensional normalized white noise sequence satisfying* (2.2) *and the condition* $H_t^+(w) \perp H_t^-(x)$ *for all* $t \in [0,T]$ *and is given by*

$$w(t) = [I - \bar{B}(t)'\bar{P}(t-1)^{-1}\bar{B}(t)]^{\frac{1}{2}}[\bar{w}(t) - \bar{B}(t)F(t)'^{-1}\bar{P}(t)^{-1}\bar{x}(t)]. \tag{2.41}$$

Moreover, the process y *satisfies the recursion*

$$y(t) = H(t)x(t) + D(t)w(t) \ , \tag{2.39b}$$

where the matrix function D *is given by*

$$D(t) = [\bar{D}(t) - G(t)'F(t)'^{-1}\bar{B}(t)][I - \bar{B}(t)'\bar{P}(t-1)^{-1}\bar{B}(t)] \ . \tag{2.42}$$

Finally, if $\bar{S} \notin \bar{S}_+$, *relations* (2.38)-(2.43) *hold for all* $t \in [0,T]$, *for which* $\bar{P}(t-1)^{-1}$ *exists.*

In the sequel, we shall be interested in obtaining the backward (forward) counterpart of S_* ($\bar{S}_*$). As is clear by now, for this we need to invert the matrices $P_*(t)$ in (2.18b) and $\bar{P}_*(t-1)$ in (2.34). However, since $P_*(0) = 0$ and $\bar{P}_*(T) = 0$, P_*^{-1} and $\bar{P}_*^{-1}$ will not exist on the whole interval [0,T]; they may not even exist on part of it for that matter. Therefore, we introduce the following

Definition 2.9. Let t_* be the smallest t such that $P_*(t) > 0$; if there is no such t on [0,T], set $t_* := T+1$. Similarly, let $\bar{t}_*$ be the largest t such that $\bar{P}_*(t-1) > 0$; if there is no such t, set $\bar{t}_* = -1$.

As is clear from the definition above, t_* and $\bar{t}_*$ might lie outside the interval [0,T]. However, if we impose some more conditions on the class S, we can guarantee that t_* and $\bar{t}_*$ belong to [0,T].

Definition 2.10. The class S is said to be *minimal* if there is no realization of type (2.1), the state process of which has dimension smaller than n.

Lemma 2.11. *S is minimal if and only if $\bar{S}$ is minimal.*

Proof. Let S be minimal. Assume that there is a backward realization $\bar{S} \in \bar{S}$ such that $\bar{S}$ is not minimal. Then all $\bar{S}$ in the (nonempty) subclass $\bar{S}_+$ are also nonminimal, and consequently, by Lemma 2.8, we could construct a nonminimal forward realization from such an $\bar{S}$, contradicting the minimality assumption of S. The converse follows analogously. □

Lemma 2.12. *Let S be minimal. Then* $t_* \le T$ *and* $\bar{t}_* \ge 0$.

Proof. Since $S_* \in S$, S_* is minimal. Hence, the pair (F,B_*) is completely controllable on the interval [0,T] [63]. In fact, were this not the case, the input-output map of S_* could be reduced [4] contradicting minimality. Therefore, the controllability gramian

$$W_*(0,t_*) = \sum_{j=0}^{t_*-1} \Phi(t_*,j+1)B_*(j)B_*(j)'\Phi(t_*,j+1)' \qquad (2.43)$$

is positive definite for some $t_* \in [0,T]$. Consequently, upon writing the solution P_* of (2.18b) as in (2.5b), we see that $P_*(t) > 0$ for all $t \in [t_*,T]$. Now, if S is minimal, $\bar{S}$ is minimal also by Lemma 2.11. Since $\bar{S}_* \in \bar{S}$, we can analogously show that $\bar{t}_* \geq 0$. □

Of course, there is no guarantee that there is a $t \in [0,T]$ for which *both* $P_*(t)$ and $\bar{P}_*(t-1)$ are positive definite.

Definition 2.13. The class S is said to be *regular* if $t_* \leq \bar{t}_*+1$, i.e., for each $t \in [0,T]$, either $P_*(t) > 0$ or $\bar{P}_*(t-1) > 0$ (or both), and is said to be *irregular* otherwise.

In fact, the regularity property of the class S depends to a certain extent on the length of the interval $[0,T]$; for if T is sufficiently large compared with n, S will be regular, since then the controllability gramian (2.43) will eventually become positive definite; the same holds in the backward direction for the controllability gramian $\bar{W}_*(T,\bar{t}_*)$ of $\bar{S}_*$. If $T < n$, we will encounter irregularity; then we do not have minimality either.

2.4. The Frame Space

Now, we are ready to justify introducing the two processes x_* and $\bar{x}_*$. It follows from (2.15) that

$$H_t(x_*) = \hat{E}\{H_t(x) \mid H_{t-1}(y)\} \subset H^-_{t-1}(y) .$$

Also, by (2.25) and (2.31), we have

$$H_{t-1}(\bar{x}_*) = \hat{E}\{H_t(x) \mid H_t^+(y)\} \subset H_t^+(y) ,$$

for all $t \in [0,T]$. Define the orthogonal complements

$$N_t^- := H_{t-1}^-(y) \ominus H_t(x_*) \quad \text{and} \quad N_t^+ := H_t^+(y) \ominus H_{t-1}(\bar{x}_*) .$$

Then, $H(y)$ can be decomposed as

$$H(y) = N_t^- \oplus H_t^{\square} \oplus N_t^+ , \tag{2.44}$$

where $H_t^{\square}$ is the *frame space* [53,54,56]

$$H_t^{\square} = H_t(x_*) \vee H_{t-1}(\bar{x}_*) \tag{2.45}$$

for all $t \in [0,T]$.

Lemma 2.14. *Let* x *be the state process of a realization* $S \in \mathcal{S}$. *Then*

$$H_t(x) \subset H_t^{\square} \oplus [H(y)]^{\perp} \tag{2.46}$$

for all $t \in [0,T]$.

Proof. See the proof of Lemma 3.7 in [64]. □

Notice that therefore, the smoothing estimate (2.4) will always be contained in the frame space, hence its importance.

In the continuous-time setting of our papers [64,65], the frame space has the constant dimension 2n on the open interval $(0,T)$. This however will not be the case here, and this contributes to the fact that the discrete-time results are nontrivial modifications of the continuous-time ones. To see this, first note that the dimensions of $H_t(x_*)$ and $H_{t-1}(\bar{x}_*)$ can be related to the ranks of P_* and $\bar{P}_*$, and hence to the condition of regularity through the following

Lemma 2.15. *Let* x *be a stochastic vector with covariance* P *and let* H(x) *be the span in* H *of its components. Then*

$$\dim H(x) = \text{rank } P . \tag{2.47}$$

Proof. Set $r := \dim H(x)$ and let $\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_r \end{bmatrix}$ be an orthonormal basis in H(x). Then $x = \ell_1\xi_1 + \ell_2\xi_2 + \ldots + \ell_r\xi_r$ for some $\ell_1, \ell_2, \ldots, \ell_r \in \mathbb{R}^n$. Define the $n \times r$-matrix $L = (\ell_1, \ell_2, \ldots, \ell_r)$. Then $x = L\xi$, and hence $P = LL'$. Consequently, rank $P \leq r$, with equality if and only if L is full rank. But L must be full rank, because otherwise $\ell_r = \Sigma_{k=1}^{r-1} \alpha_k\ell_k$ for some $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}$. Then $x = \Sigma_{i=1}^{r-1} \ell_i\hat{\xi}_i$, where $\hat{\xi}_i = \xi_i + \alpha_i$; $i = 1, 2, \ldots, r-1$. Hence, H(x) is the span of the r-1 random variables $\{\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_{r-1}\}$, which contradicts the fact that $\dim H(x) = r$. □

Consequently, the dimensions of $H_t(x_*)$ and $H_{t-1}(\bar{x}_*)$ vary between 0 and n. When S is regular, the lemma implies that, for each t, at least one of these spaces has dimension n and that there are some t's for which both of them have dimension n. Consequently, for each $t \in [0,T]$, $n \leq \dim H_t^{\square} \leq 2n$, where each limit is attained for some t. (To see this, we note that under the given conditions, it can be shown that $H_t(x_*) \cap H_{t-1}(\bar{x}_*) = \{0\}$.) On the other hand, if S is irregular, $\dim H_t^{\square} < n$ on the whole interval.

It will be more convenient in the sequel to express the frame space $H_t^{\square}$ in a somewhat more symmetric form. To this end, we shall construct the forward counterpart of $\bar{S}_*$. However, since $\bar{P}_*(t-1)$ is not positive definite for every $t \in [0,T]$, the previous argument of constructing a backward realization from a forward one cannot be reversed on the whole interval. Therefore, we apply the *generalized Moore-Penrose pseudo-inverse* (see e.g. [13]).

Definition 2.16. Let P be any matrix. The *generalized Moore-Penrose pseudo-inverse* $P^{\#}$ of P is the unique matrix satisfying

$$\text{(i) } PP^{\#}P = P, \quad \text{(ii) } P^{\#}PP^{\#} = P^{\#}, \quad \text{(iii) } (PP^{\#})' = PP^{\#}, \text{ and}$$
$$\text{(iv) } (P^{\#}P)' = P^{\#}P \,. \tag{2.48}$$

(Hence, if P is nonsingular, $P^{\#} = P^{-1}$.) In the sequel, we shall need the following

Lemma 2.17. *Let* x *be a stochastic vector with covariance matrix* P. *Then*

$$PP^{\#}x = x \,. \tag{2.49}$$

Proof. If $x = 0$, the statement is trivial; hence assume that $x \neq 0$. By the Singular Value Decomposition Theorem [66], there exists an orthogonal matrix V (i.e., $VV' = I$) such that $P = V\begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix}V'$ where $P_1 > 0$. Then x can be written $V\begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ where x_1 is stochastic vector with covariance P_1. Also, it is easy to see that $P^{\#} = V\begin{bmatrix} P_1^{-1} & 0 \\ 0 & 0 \end{bmatrix}V'$. Relation (2.49) then follows by direct multiplication. □

Now, define

$$x^*(t) = \bar{P}_*(t-1)^{\#}\bar{x}_*(t-1) \tag{2.50}$$

for each $t \in [0,T]$. Then, premultiplying both sides of (2.50) by $\bar{P}_*(t-1)$ and applying Lemma 2.17, we obtain

$$\bar{x}_*(t-1) = \bar{P}_*(t-1)x^*(t) \quad . \tag{2.51}$$

Hence $H_t(x^*) = H_{t-1}(\bar{x}_*)$ and consequently, we have the following symmetric expression

$$H_t^{\square} = H_t(x_*) \vee H_t(x^*) \tag{2.52}$$

for the frame space.

The next section will be devoted to finding a forward recursion for x^*.

Of course, we can equally well have a symmetric expression for $H_t^{\square}$ involving two processes that are the states of backward realizations via

$$H_t^{\square} = H_{t-1}(\bar{x}_*) \vee H_{t-1}(\bar{x}^*) \tag{2.53}$$

where $\bar{x}^*$ is defined by

$$\bar{x}^*(t-1) = P_*(t)^{\#}x_*(t) \tag{2.54}$$

for all $t \in [0,T]$.

2.5. The Model S*

The aim of this section is to construct and study the properties of the model S^* whose state process is x^* defined by (2.50). As we mentioned earlier, since $\bar{S}_* \nleq \bar{S}_+$, the forward-backward construction of Lemmas 2.1 and 2.8 cannot be applied here. However, the basic idea of the derivation is the same, but the results will be somewhat different.

Although for our purposes we only need to construct the forward model (S^*) corresponding to ($\bar{S}_*$), nevertheless, and for completeness only, we shall do the reverse procedure, i.e., also construct a backward model corresponding to any $S \in \mathcal{S}$ (note not in $\mathcal{S}_+$).

To do this, we need the following three lemmas. The first of these is the natural generalization of Lemma 2.2 and can be found in most standard texts. The other two are generalizations of results in [9,67].

Lemma 2.18. *Let* u *and* v *be two stochastic vectors with components in* H. *Then*

$$\hat{E}\{u \mid v\} = E\{uv'\}(E\{vv'\})^{\#}v \ . \qquad (2.55)$$

Lemma 2.19. (a) *Let* P *be the state covariance function of any* $S \in \mathcal{S}$. *Then*

$$P(t)F(t)' = P(t)F(t)'P(t+1)^{\#}P(t+1) \qquad (2.56)$$

for all $t \in [0,T]$.

(b) *Let* $\bar{P}$ *be the state covariance function of any* $\bar{S} \in \bar{\mathcal{S}}$. *Then*

$$\bar{P}(t)F(t) = \bar{P}(t)F(t)\bar{P}(t-1)^{\#}\bar{P}(t-1) \qquad (2.57)$$

for all $t \in [0,T]$.

Proof. (a) Postmultiply both sides of (2.12) by $x(t+1)$ and taking covariances, Lemma 2.18 and the fact that the components of $x(t+1)$ are orthogonal to those of the second term of the right-hand side of (2.12) yield (2.56).

(b) The proof is analogous to that of (a). □

Lemma 2.20. *Let* P *be as in the previous lemma. Then*

$$P(t+1)P(t+1)^{\#}B(t) = B(t) \tag{2.58}$$

for all $t \in [0,T]$.

Proof. To prove this lemma, we follow [67]. First, premultiply (2.1a) by $P(t+1)P(t+1)^{\#}$ to obtain

$$P(t+1)P(t+1)^{\#}x(t+1) = P(t+1)P(t+1)^{\#}F(t)x(t) + P(t+1)P(t+1)^{\#}B(t)w(t), \tag{2.59}$$

and observe that the left-hand side of (2.59) is $x(t+1)$ (by Lemma 2.17). Next, reformulate (2.56) to read

$$P(t+1)P(t+1)^{\#}F(t)P(t) = F(t)P(t) .$$

Postmultiplying this by $P(t)^{\#}x(t)$ and using Lemma 2.17 again, we see that the first term of the right-hand side of (2.59) is $F(t)x(t)$. Comparing (2.59) with (2.1a), one obtains

$$B(t)w(t) = P(t+1)P(t+1)^{\#}B(t)w(t) ,$$

which postmultiplied by $w(t)'$ and taking expectations yields (2.58). □

Now, we are ready to state the first main result of this section, which is the analogue of Lemma 2.1 and which provides us with a backward counterpart of any $S \in \mathcal{S}$ on the whole interval [0,T].

Proposition 2.21. *Let* S *be an arbitrary realization in* $\mathcal{S}$ *and let the matrix function* Θ *be given by*

$$\Theta(t) = I + B(t)'F(t)'^{-1}P(t)^{\#}F(t)^{-1}B(t) \; . \tag{2.60}$$

Then S *has the following backward counterpart*

$$(\bar{S}) \begin{cases} x(t) = P(t)F(t)'P(t+1)^{\#}x(t+1) + P(t)\tilde{B}(t)\tilde{w}(t) \\ y(t) = \tilde{G}(t)'x(t+1) + \tilde{D}(t)\tilde{w}(t) \end{cases} \tag{2.61}$$

on $[0,T]$, *where*

$$\tilde{B}(t) = -P(t)^{\#}F(t)^{-1}B(t)[I - B(t)'P(t+1)^{\#}B(t)]\Theta(t)^{\frac{1}{2}} \tag{2.62a}$$

$$\tilde{D}(t) = [D(t) - \tilde{G}(t)'P(t+1)^{\#}B(t)]\Theta(t)^{\frac{1}{2}} \tag{2.62b}$$

$$\tilde{G}(t) = F(t)P(t)^{\#}P(t)F(t)^{-1}G(t) \tag{2.62c}$$

and $\tilde{w}$ *is a white noise satisfying* (2.2) *such that* $H_t^-(\tilde{w}) \perp H_t^+(x)$ *given by,*

$$\tilde{w}(t) = \Theta(t)^{-\frac{1}{2}}[w(t) - B(t)'F(t)'^{-1}P(t)^{\#}x(t)] \; . \tag{2.63}$$

Proof. Applying the orthogonal decomposition (2.12) to (2.1a), we obtain

$$x(t) = P(t)F(t)'P(t+1)^{\#}x(t+1) + [x(t) - P(t)F(t)'P(t+1)^{\#}x(t+1)] \, . \tag{2.64}$$

Many, but straight-forward algebraic manipulations, applying Lemmas 2.17, 2.19 and 2.20, yield the first of relations (2.61). The argument of Lemma 2.4 can now be used to prove the second relation in (2.61). $\square$

Remark. If we define $\bar{x}(t-1) = P(t)^{\#}x(t)$, we could, using (2.61), obtain a backward model whose state is $\bar{x}$; but the model obtained will not, in general, belong to $\tilde{S}$ since $P(t)^{\#}P(t)F(t)'$ and $\tilde{G}$ are not equal to $F(t)'$ and G. If $P(t) > 0$, (2.61) premultiplied by $P(t)^{-1}$ will be (2.24).

Using an analogous argument in the backward setting, the following proposition can be easily proved.

Proposition 2.22. *Let* $\bar{S}_*$ *and* x^* *be given by* (2.37) *and* (2.50) *respectively and define* Θ^* *as*

$$\Theta^*(t) = I + \bar{B}_*(t)'F(t)^{-1}\bar{P}_*(t)^{\#}F(t)'^{-1}\bar{B}_*(t) \ . \qquad (2.65)$$

Then x^* *is the state process of the following model, which is the forward counterpart of* $\bar{S}_*$.

$$(S^*)\begin{cases} x^*(t+1) = \Pi(t+1)F(t)x^*(t) + \Pi(t+1)B^*(t)w^*(t) \ ; \ x^*(0) = \bar{P}_*(-1)^{\#}\bar{x}_*(-1) \\ \quad y(t) = \tilde{H}(t)x^*(t) + R^*(t)^{\frac{1}{2}}w^*(t)^{\frac{1}{2}}, \end{cases} \qquad (2.66)$$

where

$$B^*(t) = -\bar{P}_*(t)^{\#}F(t)'^{-1}\bar{B}_*(t)[I - \bar{B}_*(t)'\bar{P}_*(t-1)^{\#}\bar{B}_*(t)]\Theta^*(t)^{\frac{1}{2}}, \qquad (2.67a)$$

$$R^*(t)^{\frac{1}{2}} = [\bar{R}_*(t)^{\frac{1}{2}} - \tilde{H}(t)\bar{P}_*(t-1)^{\#}\bar{B}_*(t)]\Theta^*(t)^{\frac{1}{2}} \ , \qquad (2.67b)$$

$$\tilde{H}(t) = H(t)F(t)^{-1}\Pi(t+1)F(t) \ , \qquad (2.67c)$$

w^* *is a white noise satisfying* (2.2) *and the forward property* $H_t^+(w^*) \perp H_t^-(x^*)$ *given by*

$$w^*(t) = \Theta^*(t)^{-\frac{1}{2}}[\bar{w}_*(t) - \bar{B}_*(t)'F(t)^{-1}\bar{P}_*(t)^{\#}\bar{x}_*(t) \ , \qquad (2.68)$$

and the $n \times n$*-matrix function* Π *is defined by*

$$\Pi(t) := P^*(t)P^*(t)^{\#} \ , \qquad (2.69)$$

where P^* *is the covariance function of* x^*, *i.e.*, $P^*(t) = E\{x^*(t)x^*(t)'\}$, *which satisfies the Liapunov type equation*

$$P^*(t+1) = \Pi(t+1)F(t)P^*(t)F(t)'\Pi(t+1) + B^*(t)B^*(t)' \; ; \; P^*(0) = \bar{P}_*(-1)^{\#}. \quad (2.70)$$

Proof. First, in a manner analogous to that of Proposition 2.21, it can easily be seen that $\bar{x}_*$ is the state process of the forward realization

$$\begin{cases} \bar{x}_*(t) = \bar{P}_*(t)F(t)\bar{P}_*(t-1)^{\#}\bar{x}_*(t-1) + \bar{P}_*(t)B^*(t)w^*(t) \\ y(t) = \tilde{H}(t)\bar{P}_*(t-1)^{\#}\bar{x}_*(t-1) + R^*(t)^{\frac{1}{2}}w^*(t) \; , \end{cases} \qquad \begin{array}{r} (2.71) \\ (2.71) \end{array}$$

where B^*, w^*, $\tilde{H}$ and R^* are given above. Premultiply the first equation in (2.71) by $\bar{P}_*(t)^{\#}$, observe that $\bar{P}_*(t)^{\#} = P^*(t+1)$ and use (2.50), (2.48), and (2.69) to obtain (2.66). Finally, (2.70) follows from the state equation in (2.66). □

The matrix function Π defined by (2.69) will play an important role in what follows. Relations (2.48) imply that $\Pi^2 = \Pi$ and that $\Pi = \Pi'$, hence Π is an *orthogonal projection.* Also, Lemma 2.17 yields

$$\Pi(t)x^*(t) = x^*(t) \qquad (2.72)$$

for all $t \in [0,T+1]$. Finally, if S is minimal, $\bar{S}$ is also minimal (Lemma 2.11), in which case Lemma 2.12 guarantees the existence of a $\bar{t}_* \in [0,T]$ such that $P^*(t) > 0$ for all $t \in [0,\bar{t}_*]$. Then $\Pi \equiv I$ on $[0,\bar{t}_*]$; in this case, x^* will satisfy a recursion of type (2.1a), and the following relation

$$P_*(t) \le P(t) \le P^*(t) \qquad (2.73)$$

holds for each such t. (In fact, $P_*(t) \le P(t)$ holds for all $t \in [0,T]$.)

Finally, the following three lemmas will be needed in the sequel.

Lemma 2.23. *Let* x *be the state process and* P *the state covariance function of any* $S \in \mathcal{S}_+$ *and let* x_* *and* x^* *be defined by* (2.15) *and* (2.50) *respectively. Then*

$$E\{x(t)x_*(t)'\} = P_*(t) \tag{2.74}$$

and

$$E\{x(t)x^*(t)'\} = P(t)\Pi(t) \ . \tag{2.75}$$

for all $t \in [0,T+1]$.

Proof. In view of the definition (2.15), $H_t(x - x_*) \perp H_t(x_*)$ and therefore, (2.74) follows. Since $S \in \mathcal{S}_+$, $\bar{S} \in \bar{\mathcal{S}}_+$, and hence the backward counterpart of (2.74) reads $E\{\bar{x}(t)\bar{x}_*(t)'\} = \bar{P}_*(t)$. But $x^*(t) = P^*(t)\bar{x}_*(t-1)$, and hence $E\{x(t)x^*(t)'\} = P(t)E\{\bar{x}(t-1)\bar{x}_*(t-1)'\}P^*(t)$, which yields (2.75). □

Lemma 2.24. *Assume that* S *is regular. Let* x_* *and* x^* *be defined by* (2.16) *and* (2.66) *respectively. Then, for all* $t \in [0,T+1]$,

$$\hat{E}\{x^*(t) \mid x_*(t)\} = \Pi(t)x_*(t) \tag{2.76}$$

and

$$E\{x_*(t)x^*(t)'\} = P_*(t)\Pi(t) \ . \tag{2.77}$$

Proof. Since S is regular, the two intervals $[0,\bar{t}_*]$ and $[t_*,T+1]$ cover the whole interval $[0,T+1]$. On $[0,\bar{t}_*]$, $\bar{P}_*(t-1) > 0$, and consequently S^* has all the properties of realizations in $\mathcal{S}_+$ *on that interval.* Since $\Pi \equiv I$ on $[0,\bar{t}_*]$, this implies that (2.76) and (2.77) hold there.

Similarly, on $[t_*, T+1]$, $P_*(t) > 0$, and therefore, analogously with the above, $E\{\bar{x}_*(t-1)\bar{x}^*(t-1)'\} = \bar{P}_*(t-1)$ on this interval. Premultiply this by $P^*(t)$ and postmultiply by $P_*(t)$ and remembering that $\bar{P}_*(t-1) = P^*(t)^{\#}$, (2.77) is seen to hold on $[t_*, T+1]$. Then Lemma 2.18 provides the requested formula (2.76) on this interval. □

Lemma 2.25. *Let* $S \in S_+$ *and let* Q^* *be the covariance function of* $\Pi(x^* - x)$. *Then*

$$Q^* = \Pi(P^* - P)\Pi \ . \tag{2.78}$$

If in addition S is regular, then

$$E\{[x - x_*][x^* - x]'\Pi\} = 0 \ . \tag{2.79}$$

Proof. Relation (2.78) follows from (2.75). Relation (2.79) follows from (2.74), (2.75) and (2.77). □

Remark. Since Π is projection, $\Pi Q^* \Pi = Q^*$.

2.6. A Mayne-Fraser-Type Smoothing Formula

Given the state process x and the output process y of a model (2.1), the smoothing problem consists in determining the *smoothing estimate*

$$\hat{x}(t) = \hat{E}\{x(t) \mid H(y)\} \tag{2.80}$$

of x and the smoothing error covariance

$$\Sigma(t) = E\{[x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]'\} \tag{2.81}$$

for all $t \in [0,T]$. In this section, we shall derive a smoothing formula for the case that S is regular. At the end of the section, we shall present a conjecture which, if true, would allow us to remove this regularity assumption. Our approach will utilize an orthogonal decomposition of the frame space $H_t^{\square}$ to be given below.

In view of Lemma 2.14, $\hat{x}(t) \in H_t^{\square}$. In order to obtain a formula for $\hat{x}$, we need to decompose the frame space $H_t^{\square}$ perpendicularly. To this end, we introduce the process z defined by

$$z(t) = x^*(t) - \Pi(t)x_*(t) \tag{2.82}$$

for all $t \in [0,T+1]$.

Lemma 2.26. *Assume that the class S is regular. Let z be defined by* (2.82), *and let* $Q(t) := E\{z(t)z(t)'\}$. *Then*

$$H_t^{\square} = H_t(x_*) \oplus H_t(z) \tag{2.83}$$

for all $t \in [0,T+1]$. *Moreover,*

$$Q = \Pi(P^* - P_*)\Pi \tag{2.84}$$

and

$$Q = \Pi(Q_* + Q^*)\Pi \tag{2.85}$$

where Q_* *and* Q^* *are given by* (2.18c) *and* (2.78) *respectively and* P *is the state covariance of any* $S \in S$.

Proof. Since S is regular, Lemma 2.24 implies that the components of z are orthogonal to those of x_*, and therefore (2.83) follows from (2.52). Next, since $x^* = \Pi x^*$, $z = \Pi(x^* - x_*)$. Then (2.84) is a direct consequence of (2.77). Finally, write $Q = \Pi(P^* - P + P - P_*)\Pi$ to obtain (2.85). □

Lemma 2.27. *Let* z *be given by* (2.82) *and* Q *be its covariance function. Then*

$$Q = \Pi Q = Q\Pi = \Pi Q \Pi \ , \tag{2.86a}$$

$$Q^{\#} = \Pi Q^{\#} = Q^{\#}\Pi = \Pi Q^{\#}\Pi \tag{2.86b}$$

and

$$QQ^{\#} = Q^{\#}Q = \Pi \ . \tag{2.86c}$$

Proof. We shall first show that $Q = \Pi Q \Pi$ and $Q^{\#} = \Pi Q^{\#}\Pi$, the first of which follows trivially from (2.84) and the fact that Π is a projection. For the second relation, let $t \in [0,T]$ be fixed. The case $P^*(t) = 0$ is trivial (for then $Q(t) = 0$ and $\Pi(t) = 0$); hence we shall assume that $P^*(t) \neq 0$. As mentioned in Lemma 2.17, there exists an orthogonal matrix V such that $P^*(t) = V\begin{bmatrix}\tilde{P}^*(t) & 0\\ 0 & 0\end{bmatrix}V'$ where $\tilde{P}^* > 0$. Then $P^*(t)^{\#} = V\begin{bmatrix}\tilde{P}^*(t)^{-1} & 0\\ 0 & 0\end{bmatrix}V'$ and $\Pi(t) = V\begin{bmatrix}I & 0\\ 0 & 0\end{bmatrix}V'$. In view of the fact that $Q = \Pi Q \Pi$, $Q(t)$ can be written $Q(t) = V\begin{bmatrix}\tilde{Q}(t) & 0\\ 0 & 0\end{bmatrix}V'$. We want to show that $\tilde{Q}(t) > 0$. But in view of (2.85), this must be the case, because if we choose S in S_+, $Q_*(t) > 0$ and $Q^*(t) \geq 0$. Hence $Q(t)^{\#} = V\begin{bmatrix}\tilde{Q}(t)^{-1} & 0\\ 0 & 0\end{bmatrix}V'$ and consequently, $Q^{\#} = \Pi Q^{\#}\Pi$. From the above discussion it is clear that (2.86c) holds. The rest of relations (2.86) follow trivially by remembering that Π is a projection. □

Now, we are ready to present the main result of this chapter.

Theorem 2.28. *Assume that the class S is regular. Let* x *be the state process of any* S *of class* S_+. *Then the smoothing estimate* (2.80) *is given by*

$$\hat{x}(t) = [I - Q_*(t)Q(t)^{\#}]x_*(t) + Q_*(t)Q^{\#}(t)x^*(t) \tag{2.87}$$

and the error covariance (2.81) *by*

$$\Sigma(t) = Q_*(t) - Q_*(t)Q(t)^{\#}Q_*(t) \tag{2.88}$$

for all $t \in [0,T+1]$.

Proof. Since $\hat{x}(t) = \hat{E}\{x(t) \mid H_t^{\square}\}$ (Lemma 2.14), (2.83) yields

$$\hat{x}(t) = \hat{E}\{x(t) \mid x_*(t)\} + \hat{E}\{x(t) \mid z(t)\} ,$$

which upon using (2.15) and Lemma 2.18 can be written

$$\hat{x}(t) = x_*(t) + E\{x(t)z(t)'\}Q(t)^{\#}z(t) .$$

In view of (2.74) and (2.75), $E\{x(t)z(t)'\} = Q_*(t)\Pi(t)$. Since $z = \Pi z$, (2.87) follows from the above relation noting (2.86). To prove (2.88), observe that

$$x - \hat{x} = (I - Q_*Q^{\#})(x - x_*) + Q_*Q^{\#}(x - x^*) .$$

Replacing $Q^{\#}$ by $\Pi Q^{\#}\Pi$, and noting that the two terms above are, in view of (2.79), orthogonal, we obtain

$$\Sigma = Q_* - Q_*\Pi Q^{\#}\Pi Q_* - Q_*\Pi Q^{\#}\Pi Q_*$$
$$+ Q_*\Pi Q^{\#}\Pi Q_*\Pi Q^{\#}\Pi Q_* + Q_*\Pi Q^{\#}\Pi Q^*\Pi Q^{\#}\Pi Q_* ,$$

which, in view of (2.85) and (2.86), yields (2.88). □

To obtain the discrete-time version of the Mayne-Fraser formula [38,39], the following lemma is needed.

Lemma 2.29. *Let* $S \in \mathcal{S}_+$. *Then, for each* $t \in [0,T+1]$, $Q_*(t)$ *and* $\Sigma(t)$ *are positive definite and satisfy*

$$\Sigma(t)^{-1} = Q_*(t)^{-1} + Q^*(t)^\# \qquad (2.89)$$

where Q^* *is given by* (2.78).

Proof. Let P be the covariance function of S. Since $S \in \mathcal{S}_+$, $Q_*(0) = P(0) - P_*(0) = N > 0$, consequently $Q_*(t) > 0$ for all $t \in [0,T+1]$. To see this, observe that the Riccati equation (2.16e) can be reformulated to read

$$\begin{aligned} Q_*(t+1) &= \Gamma_*(t)Q_*(t)\Gamma_*(t)' \\ &\quad + (B_*(t)R_*(t)^{-\frac{1}{2}}D(t) - B(t))(B_*(t)R_*(t)^{-\frac{1}{2}}D(t) - B(t))', \end{aligned} \qquad (2.90)$$

where Γ_* is the *feedback matrix*

$$\Gamma_* = F - B_*R_*^{-\frac{1}{2}}H \ . \qquad (2.91)$$

The Liapunov-type equation (2.90) can be written in the form (2.5b) to yield $Q_*(t) > 0$ for all $t \in [0,T+1]$, since $Q_*(0) > 0$. The same argu-emtn can be used to prove that $\Sigma > 0$: first determine Σ from the Liapunov-type equation corresponding to the *backward* representation (2.115) below (the proof of which, of course, does not depend on this lemma); then, as above, note that for all $t \in [0,T]$, $\Sigma(t) \geq \Sigma(T+1) = Q_*(T+1) > 0$. This proves the positivity of Σ. Next, it follows from (2.88) that

$$I - Q_*Q^{\#} = \Sigma Q_*^{-1} \tag{2.92}$$

which is nonsingular for all $t \in [0,T+1]$. Then applying the matrix inversion lemma (1.34) to (2.88), it is seen that $\Sigma^{-1} = Q_*^{-1} + Z$, where

$$Z = Q^{\#}(I - Q_*Q^{\#})^{-1} = (I - Q^{\#}Q_*)^{-1}Q^{\#} . \tag{2.93}$$

Therefore, it just remains to show that $Z = Q^{*\#}$. Since Π is a projection, (2.85) yields

$$Q^* = Q - \Pi Q_* \Pi \quad . \tag{2.94}$$

In view of Lemma 2.27, (2.94) and the first of relations (2.93) yield $Q^*Z = \Pi$, which is symmetric. Likewise, using the second of relations (2.93), we obtain $ZQ^* = \Pi$. Then, Lemma 2.27 implies that $Q^*ZQ^* = Q^*$ and $ZQ^*Z = Z$ also. Consequently, by Definition 2.16, $Z = Q^{*\#}$. □

Now, we are in a position to state the *Mayne-Fraser two-filter formula.*

Theorem 2.30. *Let* S *be regular and let* $S \in S_+$. *Then the smoothing estimate* $\hat{x}$ *of the state* x *of* S *is given by*

$$\hat{x}(t) = \Sigma(t)[Q_*(t)^{-1}x_*(t) + Q^*(t)^{\#}x^*(t)] \tag{2.95}$$

where x_*, x^* *and* Σ *are given by* (2.16), (2.66) *and* (2.89) *respectively.*

Proof. It follows from (2.89) and (2.92) that $Q_*Q^{\#} = I - \Sigma Q_*^{-1} = \Sigma(\Sigma^{-1} - Q_*^{-1}) = \Sigma Q^{*\#}$. This together with (2.92) yields (2.95) when inserted in (2.88). □

Although the regularity assumption imposed on the class S is not a very stringent one (especially if T is large enough in comparison with n), it would be interesting to see if it can be removed. Actually, this assumption was introduced *only* to obtain relation (2.76). We believe that (2.76) holds without the regularity assumption; as expressed in the following

Conjecture. *Let* x_* *and* x^* *be defined by* (2.15) *and* (2.50) *respectively. Then*

$$\hat{E}\{x^*(t) \mid x_*(t)\} = \Pi(t)x_*(t) \ . \tag{2.96}$$

Justification. First, note that $\hat{E}\{x^*(t) \mid x_*(t)\} = \hat{E}\{x^*(t) \mid H^-_{t-1}(y)\}$. To see this, note that $H^-_{t-1}(y) = H_t(x_*) \oplus N^-_t$ where $N^-_t \perp H_t(x^*)$, for $N^-_t \perp H^{\square}_t$. Now, let x^+ be defined by

$$(S^+) \begin{cases} x^+(t+1) = F(t)x^+(t) + B^*(t)w^*(t) \\ y(t) = \tilde{H}(t)x^+(t) + R^*(t)^{\frac{1}{2}}w^*(t) \ . \end{cases} \tag{2.97}$$

Then, noting that (2.57) implies $\Pi(t+1)F(t)\Pi(t) = \Pi(t+1)F(t)$ and $\tilde{H}(t) = \tilde{H}(t)\Pi(t)$, we obtain

$$x^*(t) = \Pi(t)x^+(t) \ . \tag{2.98}$$

Hence, (2.96) is equivalent to showing that $\hat{x}^+(t) :=$ $\hat{E}\{x^+(t) \mid H^-_{t-1}(y)\} = x_*(t)$. But $\hat{x}^+(t)$ is generated by the Kalman-Bucy filter

$$\hat{x}^+(t+1) = F(t)\hat{x}^+(t) + K(t)R^+(t)^{-\frac{1}{2}}[y(t) - \tilde{H}(t)\hat{x}^+(t)] \ ; \ \hat{x}^+(0) = 0 \ , \tag{2.99}$$

where K is the gain function and R^+ is given by

$$R^+ = \Lambda_0 - \tilde{H}\hat{P}\tilde{H}' \ , \tag{2.100}$$

$\Lambda_o(t) = E\{y(t)y(t)'\}$ and $\hat{P}(t) = E\{\hat{x}^+(t)\hat{x}^+(t)'\}$. Using the output equations of (2.18a) and (2.91) alternatively to compute $\hat{E}\{y(t) \mid H^-_{t-1}(y)\}$, we easily see that

$$Hx_* = \tilde{H}\hat{x}^+ , \tag{2.101}$$

which implies that $HP_*H' = \tilde{H}\hat{P}\tilde{H}'$. Hence $R^+ = R_*$ and the process $R^+(t)^{-\frac{1}{2}}[y(t) - \tilde{H}(t)\hat{x}^+(t)]$ is the innovation process $w_*(t)$ defined by (2.17). To justify the conjecture, it then only remains to show that the gain functions K and B_* are the same. Due to time limitations, we shall leave this open. □

2.7. The Bryson-Frazier Formulation

In this section we shall derive the discrete-time version of the Bryson-Frazier smoothing formula [35]. This will be done by using a procedure, based on an orthogonal decomposition of $H(y)$, which does not require that S be regular. Then the smoothing formula of Rauch, Tung and Striebel [36] will be obtained as a corollary.

Since $H(y) = H(w_*) = H^-_{t-1}(w_*) \oplus H^+_t(w_*)$ (for w_* is a white noise) and $H^-_{t-1}(w_*) = H^-_{t-1}(y)$

$$H(y) = H^-_{t-1}(y) \oplus H^+_t(w_*) , \tag{2.102}$$

and consequently, (2.80) yields

$$\hat{x}(t) = \hat{E}\{x(t) \mid H^-_{t-1}(y)\} + \hat{E}\{x(t) \mid H^+_t(w_*)\} ,$$

which in view of (2.15) and the orthogonality between $x_*(t)$ and $H_t^+(w_*)$ can be written

$$\hat{x}(t) = x_*(t) + \hat{E}\{z_*(t) \mid H_t^+(w_*)\} \tag{2.103}$$

where

$$z_*(t) = x(t) - x_*(t) \ . \tag{2.104}$$

This stochastic process satisfies the forward recursion

$$z_*(t+1) = \Gamma_*(t)z_*(t) + [B(t) - B_*(t)R_*(t)^{-\frac{1}{2}}D(t)]w(t), \tag{2.105}$$

where Γ_* is the feedback matrix function (2.91). To see this, first note that (2.16a) can be written

$$x_*(t+1) = \Gamma_*(t)x_*(t) + B_*(t)R_*(t)^{-\frac{1}{2}}y(t) \ ;$$

then insert (2.1b) into this and subtract from (2.1a) to obtain (2.105). Moreover, we see that the covariance function of z_* is precisely Q_* as defined in Section 2.3.

However, to evaluate the second term of (2.103) we shall need the backward counterpart of (2.105), in the sense of Section 2.5. Modulo a complete description of the exogeneous noise ν, such a backward representation was provided by Pavon in [9].

Lemma 2.31. ([9]) *Let* x *be the state process of any realization in* S *and let* P *be its covariance function. Then the process* z_* *defined by* (2.104) *satisfies the backward recursion*

$$z_*(t) = Q_*(t)\Gamma_*(t)'Q_*(t+1)^{\#}z_*(t+1) + Q_*(t)H(t)'R_*(t)^{-\frac{1}{2}}w_*(t)+\nu(t) \tag{2.106}$$

where Γ_* *and* Q_* *are given by* (2.91) *and* (2.18b) *respectively and* ν *is an (unnormalized) white noise whose components are contained in* $[H(y)]^{\perp}$.

Relation (2.106) *is a backward representation in the sense that* $H_t^-(w_*,\nu) \perp H_t^+(z_*)$ *for all* $t \in [0,T]$.

Equation (2.106) was derived in [9] by first noting that it is no restriction to assume that the basic Hilbert space H can be written $H = H(y) \oplus H(\eta)$, where η is an n-dimensional white noise process of type (2.2) such that $H_{t-1}^-(\eta) \perp H_t^+(z_*)$ for all $t \in [0,T]$. In fact, such a framework is sufficient for representing the state, output and input processes. Next, it was seen that w_* and η could be regarded as outputs of a forward realization with (2.105) as its state equation; the white noise character of the output modifies the construction of a backward representation.

An alternative derivation of Lemma 2.31, which in addition provides a complete characterization of the process ν, can be obtained along the lines of Theorem 4.3 in our continuous-time paper [64]: First note that, for each $t \in [0,T]$, there is an orthogonal $p \times p$-matrix $V(t)$ such that

$$\begin{bmatrix} \bar{B}(t) \\ D(t) \end{bmatrix} = \begin{bmatrix} \bar{B}_1(t) & B_2(t) \\ R(t)^{\frac{1}{2}} & 0 \end{bmatrix} V(t) \tag{2.106}$$

where B_1 is $n \times m$ and B_2 is $n \times (p - m)$. Let

$$\begin{bmatrix} u \\ v \end{bmatrix} = Vw \tag{2.107}$$

define a pair of mutually uncorrelated white noise processes u and v of dimensions m and p - m respectively. Then $Bw = B_1u + B_2v$ and $Dw = R^{\frac{1}{2}}u$, and (2.105) can be written

$$z_*(t+1) = \Gamma_*(t)z_*(t) - \Gamma_*(t)Q_*(t)H(t)'R(t)^{-\frac{1}{2}}u(t) + B_2(t)v(t). \tag{2.108}$$

To see this, first use (2.16c) to see that $R_*^{-1}R = I - R_*^{-1}HQ_*H'$ and (2.16b) to see that $B_1R^{\frac{1}{2}} - B_*R_*^{\frac{1}{2}} = -FQ_*H'$; from this it is easy to see that $B_1 - B_*R_*^{-\frac{1}{2}}R^{\frac{1}{2}} = -\Gamma_*Q_*H'$. Given (2.108), Lemma 2.31 follows from the appropriate modifications of Proposition 2.21 and some tedious calculation. This also provides an expression for ν in terms of v and z_*, which is useful in obtaining a representation for the smoothing error, such as the one in Theorem 4.3 in [64].

Now, we are ready to state the main result of this section. To simplify notations, we introduce the process $\{r(t)\ ;\ t \in [0,T+1]\}$ defined by

$$r(t) := \hat{E}\{z_*(t) \mid H_t^+(w_*)\} \quad . \tag{2.109}$$

Since w_* is defined only on $[0,T]$, set $r(T+1) = 0$.

Theorem 2.32. *Let* x *be the state process of any* $S \in \mathcal{S}$ *with state covariance function* P *and let* $\hat{x}$ *be the corresponding smoothing estimate* (2.80). *Then, for all* $t \in [0,T]$,

$$\hat{x}(t) = x_*(t) + r(t) \tag{2.110}$$

where r *is defined by* (2.109) *and satisfies the backward recursion*

$$\begin{cases} r(t) = Q_*(t)\Gamma_*(t)'Q_*(t+1)^{\#}r(t+1) + Q_*(t)H(t)'R_*(t)^{-\frac{1}{2}}w_*(t) \\ r(T+1) = 0\ . \end{cases} \tag{2.111}$$

Proof. In view of the definition (2.109), (2.110) is the same as (2.103). Next, since the components of ν are orthogonal to $H(w_*)$, (2.111) follows trivially from Lemma 2.31. □

Remark. Relations (2.110) and (2.111) imply that the covariance matrix $\Sigma(T+1) = Q_*(T+1)$, which, in view of Lemma 2.29, is positive definite if $S \in S_+$.

The *Bryson-Frazier* formula [35] can now be obtained from the above theorem.

Corollary 2.33. *Let* x *be the state process of a realization* $S \in S_+$. *Then the smoothing estimate* $\hat{x}$ *satisfies*

$$\hat{x}(t) = x_*(t) + Q_*(t)\bar{z}(t-1) \tag{2.112}$$

where x_* *and* Q_* *are given by* (2.16) *and* $\bar{z}$ *satisfies*

$$\bar{z}(t-1) = \Gamma_*(t)'\bar{z}(t) + H(t)'R_*(t)^{-\frac{1}{2}}w_*(t) \; ; \; \bar{z}(T) = 0 \; . \tag{2.113}$$

The process $\bar{z}$ *is related to* r *through*

$$\bar{z}(t-1) = Q_*(t)^{-1}r(t) \; . \tag{2.114}$$

Proof. Since $S \in S_+$, $Q_*(t) > 0$ for all $t \in [0,T+1]$ (Lemma 2.29) and consequently, the process $\bar{z}$ given by (2.114) is well-defined. Then the result follows from the theorem. □

Finally, to obtain the result of Rauch *et al* [36], the following proposition is needed.

Proposition 2.34. *Let $S \in S_+$ and let x be its state process. Then the smoothing estimate $\hat{x}$ satisfies the difference equation*

$$\begin{cases} \hat{x}(t+1) = F(t)\hat{x}(t) + B(t)(I - D(t)'R_*(t)^{-1}D(t))B(t)'\bar{z}(t) \\ \qquad\qquad + B(t)D(t)'R_*(t)^{-1}[y(t) - H(t)x_*(t) - H(t)Q_*(t)F(t)'\bar{z}(t)] \\ \hat{x}(T+1) = x_*(T+1) \ , \end{cases} \tag{2.115}$$

where

$$\bar{z}(t) = Q_*(t+1)^{-1}[\hat{x}(t+1) - x_*(t+1)] \ .$$

Proof. Writing (2.112) as

$$\hat{x}(t+1) = x_*(t+1) + Q_*(t+1)\bar{z}(t) \ , \tag{2.116}$$

using the recursions (2.18a) and (2.16e) for x_* and Q_* and adding and subtracting $F(t)Q_*(t)\bar{z}(t-1)$ in (2.116), we obtain

$$\hat{x}(t+1) = F(t)\hat{x}(t) + B(t)B(t)'\bar{z}(t) + \mu(t) \tag{2.117}$$

where

$$\begin{aligned} \mu(t) = & -F(t)Q_*(t)\bar{z}(t-1) + F(t)Q_*(t)F(t)'\bar{z}(t) \\ & - B_*(t)B_*(t)'\bar{z}(t) + B_*(t)w_*(t) \end{aligned} \tag{2.118}$$

Now inserting (2.113) into (2.118) and using (2.91) to eliminate Γ_*, we get

$$\begin{aligned} \mu(t) = & F(t)Q_*(t)H(t)'R_*(t)^{-\frac{1}{2}}B_*(t)'\bar{z}(t) - F(t)Q_*(t)H(t)'R_*(t)^{-\frac{1}{2}}w_*(t) \\ & - B_*(t)B_*(t)'\bar{z}(t) + B_*(t)w_*(t) \ . \end{aligned}$$

Inserting (2.16b) and (2.17) into the last two terms of this expression and cancelling similar terms, we see that

$$\begin{aligned} \mu(t) = & -B(t)D(t)'R_*(t)^{-1}D(t)B(t)'\bar{z}(t) \\ & + B(t)D(t)'R_*(t)^{-1}[y(t) - H(t)x_*(t) - H(t)Q_*(t)F(t)'\bar{z}(t)] \end{aligned} \tag{2.119}$$

which immediately yields (2.115). □

Now, let $BD' = 0$; this is a basic assumption in [36]. Then (2.115) reduces to

$$\hat{x}(t+1) = F(t)\hat{x}(t) + B(t)B(t)'\bar{z}(t) \; ; \; \hat{x}(T+1) = x_*(T+1) \; ,$$

which, in view of (2.116), becomes

$$\hat{x}(t+1)=F(t)\hat{x}(t)+B(t)B(t)'Q_*(t+1)^{-1}[\hat{x}(t+1)-x_*(t+1)] \; ; \; \hat{x}(T+1)=x_*(T+1) \; . \tag{2.120}$$

Corollary 2.35. *Let* x *be the state process of any* $S \in S_+$ *and let* $BD' = 0$. *Then, the smoothing estimate* $\hat{x}$ *can be written*

$$\begin{cases} \hat{x}(t)=x(t|t)+P(t|t)F(t)'[F(t)P(t|t)F(t)'+B(t)B(t)']^{-1}[\hat{x}(t+1)-F(t)x(t|t)] \\ \hat{x}(T+1) = x_*(T+1) \end{cases} \tag{2.121}$$

where $x(t|t)$ *is the filter* $\hat{E}\{x(t) \mid H_t^-(y)\}$ *and* $P(t|t)$ *is its error covariance matrix, i.e.,* $P(t|t) = E\{[x(t) - x(t|t)][x(t) - x(t|t)]'\}$.

Proof. First, solve (2.120) for $\hat{x}(t)$ in terms of $\hat{x}(t+1)$, add and subtract $F(t)^{-1}x_*(t+1)$ and rearrange terms to obtain

$$\hat{x}(t)=F(t)^{-1}x_*(t+1)+[F(t)^{-1}-F(t)^{-1}B(t)B(t)'Q_*(t+1)^{-1}][\hat{x}(t+1)-x_*(t+1)].$$

This can be rewritten as

$$\begin{aligned} \hat{x}(t) = F(t)^{-1}x_*(t+1) + F(t)^{-1}[Q_*(t+1) - B(t)B(t)']F(t)'^{-1}F(t)' \cdot \\ \cdot [(Q_*(t+1) - B(t)B(t)') + B(t)B(t)']^{-1}[\hat{x}(t+1) - x_*(t+1)]. \end{aligned} \tag{2.122}$$

It is well-known and easy to see that the one step predictor $x_*(t+1)$ and the filter $x(t|t)$ are related by

$$x_*(t+1) = F(t)x(t|t) \tag{2.124}$$

and that the corresponding error covariances satisfy the relation

$$Q_*(t+1) = F(t)P(t|t)F(t)' + B(t)B(t)' . \tag{2.125}$$

Finally, note that (2.124) and (2.125) may be reformulated as

$$x(t|t) = F(t)^{-1}x_*(t+1)$$

and

$$P(t|t) = F(t)^{-1}[Q_*(t+1) - B(t)B(t)']F(t)'^{-1} ,$$

which, inserted into (2.122), yields (2.121). □

Relation (2.121) is the formula of Rauch, Tung and Striebel presented in [36].

It remains to clarify the connections between the results of Sections 2.6 and 2.7. Note that the two-filter formula (2.87) can be written

$$\hat{x}(t) = x_*(t) + Q_*(t)Q(t)^{\#}z(t) ,$$

where z is defined by (2.82). Comparing this with (2.112), it is seen that we need to prove that

$$\bar{z}(t-1) = Q(t)^{\#}z(t) ,$$

analogously with the continuous-time setting [64,65]. The problems encountered in trying to show this are similar to those of proving the conjecture in Section 2.6, and due to time limitations, we are leaving this question for a future paper.

CHAPTER 3

TOPICS ON THE STOCHASTIC REALIZATION PROBLEM FOR CONTINUOUS-TIME NONSTATIONARY STOCHASTIC PROCESSES

3.1. Introduction

Let the Hilbert space H be as defined in Section 2.2. In this chapter, the following notations will be adopted. For any n-dimensional stochastic process z, $H_t(z)$ will denote the (closed) subspace spanned by the random variables $\{z_1(t), z_2(t), \ldots, z_n(t)\}$. Let $H(z)$ and the past spaces $H_t^-(z)$ and $H_{[t_0,t]}(z)$ be defined as $\bigvee_{\tau\in I} H_\tau(z)$, where the interval I is $(-\infty, \infty)$, $(-\infty,t]$ and $[t_0,t]$ respectively. The future spaces $H_{[t,t_1]}(z)$ and $H_t^+(z)$ are defined analogously. Sometimes, we shall be interested in spaces spanned by the *increments* of z. Hence, we define $H(dz)$, $H_t^-(dz)$ and $H_t^+(dz)$ to be the closed linear hulls in H of $\{z(s) - z(r);\ s, r \in I\}$, where the interval I is $(-\infty,\infty)$, $(-\infty,t]$ and $[t,\infty)$ respectively.

Let $\{x(t);\ t \in \mathbb{R}\}$ and $\{y(t);\ t \in \mathbb{R}\}$ be two stochastic processes of dimensions n and m respectively, defined as the solution of the linear stochastic system

$$(S)\begin{cases} dx = F(t)x(t)dt + B(t)dw \\ dy = H(t)x(t)dt + D(t)dw \end{cases} \tag{3.1a}$$

where w is a vector process, of dimension $p \geq m$, with orthogonal increments such that

$$E\{dw\} = 0 \quad ; \quad E\{dwdw'\} = I\,dt \tag{3.2}$$

and $H_t^+(dw) \perp H_t^-(x)$ for all $t \in \mathbb{R}$. The matrix $R(t) := D(t)D(t)'$ is positive definite on $\mathbb{R}$, the matrix $F(t)$ is uniformly asymptotically stable on $\mathbb{R}$ and F, B, H, D and R^{-1} are matrices of bounded and analytic functions. As before, the process x is called the *state* of the model S, y is the *output* and w is the *input*. We shall assume the model S to be *minimal* in the sense that there is no other model of the form (3.1) with the process y as its output and with a state process x of smaller dimension than n. *The stochastic realization problem* consists of finding all possible stochastic systems (3.1) (belonging to a class S to be prescribed below) having the process $\{y(t); t \in \mathbb{R}\}$ as their output. Each such model S will be called a *stochastic realization* of y. In particular, S is minimal and analytic, i.e., F, B, H, D and R^{-1} are analytic on $\mathbb{R}$.

For each $t \in \mathbb{R}$, there exists an orthogonal matrix $V(t)$ such that

$$\begin{bmatrix} B(t) \\ D(t) \end{bmatrix} = \begin{bmatrix} B_1(t) & B_2(t) \\ R(t)^{1/2} & 0 \end{bmatrix} V(t) \tag{3.3}$$

where B_1 is $n \times m$ and B_2 is $n \times (p - m)$. It is no restriction to take $V(t) \equiv I$. Next, let

$$\begin{bmatrix} du \\ dv \end{bmatrix} = dw \tag{3.4}$$

define a pair of orthogonal increment processes u and v, of dimensions m and p - m respectively. It is obvious that (3.4) satisfies (3.2). Accordingly, we have reduced the problem to that of finding all admissable models S of the type

$$(S)\quad \begin{cases} dx = F(t)x(t)dt + B_1(t)du + B_2(t)dv & (3.5a) \\ dy = H(t)x(t)dt + R(t)^{\frac{1}{2}}\, du\ , & (3.5b) \end{cases}$$

having the process $\{y(t);\ t \in \mathbb{R}\}$ as an output. Clearly, the matrix function $P(t) := E\{x(t)x(t)'\}$ satisfies the differential equation

$$\dot{P} = FP + PF' + B_1B_1' + B_2B_2' \tag{3.6}$$

on $\mathbb{R}$. We shall call P the *state covariance function* of S.

As we have seen in the previous chapters, it is sometimes more convenient to use a backward representation for the state process x. To do this, we need to invert the matrix function P. Since S is minimal, (F,B) must be completely controllable [4,63]. Since in addition, F and B are analytic, (F,B) must be totally controllable [68,69]. With this condition satisfied, it is not hard to prove that P(t) must be positive definite for all $t \in \mathbb{R}$ (see [71; p.28]). Then, the process

$$\bar{x}(t) = P(t)^{-1}\, x(t) \tag{3.7}$$

is well-defined with components in H. Let $\bar{P}$ be its covariance function

$$\bar{P}(t) = E\{\bar{x}(t)\bar{x}(t)'\}\ . \tag{3.8}$$

The following lemma, the proof of which can be found in [64], is the analogue of Lemma 2.1.

Lemma 3.1. *Let* x *be the state process of an arbitrary realization* (3.5) *with state covariance* P. *Then the process* $\bar{x}$ *defined by* (3.7) *satisfies the backward model*

$$d\bar{x} = -F(t)'\,\bar{x}(t)dt + \bar{B}(t)d\bar{w} \tag{3.9a}$$

for all $t \in \mathbb{R}$, *where* $\bar{B} = P^{-1}B$ *and* $\bar{w}$ *is a* p*-dimensional orthogonal increments process satisfying* (3.2) *and the condition* $H_t^-(d\bar{w}) \perp H_t^+(\bar{x})$ *for all* $t \in \mathbb{R}$. *The increments of* $\bar{w} = \begin{bmatrix}\bar{u}\\ \bar{v}\end{bmatrix}$ *are given by*

$$d\bar{w} = \begin{bmatrix} d\bar{u} \\ d\bar{v} \end{bmatrix} = \begin{bmatrix} du - B_1'P^{-1}\;x\;dt \\ dv - B_2'P^{-1}\;x\;dt \end{bmatrix} \tag{3.9b}$$

and the covariance function (3.8) *by* $\bar{P} = P^{-1}$; *it satisfies*

$$\dot{\bar{P}} = -F'\bar{P} - \bar{P}F - \bar{B}\bar{B}' \quad . \tag{3.10}$$

To obtain a backward realization for y, insert (3.9) in (3.5b) to get $dy = (HP + R^{\frac{1}{2}}B_1')\bar{x} + R^{\frac{1}{2}}d\bar{u}$. Then, defining $\bar{B}_1$ and $\bar{B}_2$ to be $P^{-1}B_1$ and $P^{-1}B_2$ respectively, Lemma 3.1 yields

$$(\bar{S})\;\begin{cases} d\bar{x} = -F'(t)\bar{x}(t)dt + \bar{B}_1(t)d\bar{u} + \bar{B}_2(t)d\bar{v} \\ dy = G'(t)\bar{x}(t)dt + R^{\frac{1}{2}}(t)d\bar{u} \quad , \end{cases} \tag{3.11}$$

where the $n \times m$-matrix function G is defined by

$$G = PH' + B_1R^{\frac{1}{2}} \quad . \tag{3.12}$$

Note that (3.5) and (3.11) have the same state space, i.e.,

$$H_t(x) = H_t(\bar{x}) \tag{3.13}$$

for each $t \in \mathbb{R}$. We shall call any model of type (3.11) with y as its output, and $H_t^-(d\bar{w}) \perp H_t^+(\bar{x})$ for each $t \in \mathbb{R}$ a *backward realization* of $\{y(t);\ t \in \mathbb{R}\}$. Note that $\bar{S}$ is also analytic.

3.2. The Finite Interval Case: A Review

In order to complete the statement of the realization problem indicated in Section 3.1, we shall have to specify the classes S and $\bar{S}$ to which the realizations (3.5) and (3.11) respectively belong. To this end, we shall first restrict ourselves to the finite interval $[t_o,t_1]$, where t_o and t_1 are arbitrary elements of $\mathbb{R}$. Realization theory for processes defined on a finite interval was developed in our papers [64,65] (this is the continuous-time version of the theory of Chapter 2) and next we shall briefly review some facts from it.

Hence we shall consider models

$$(S)\begin{cases} dx = F(t)x(t)dt + B_1(t)du + B_2(t)dv \quad ; \quad x(t_o) = \xi \\ dy = H(t)x(t)dt + R(t)^{\frac{1}{2}}du \end{cases} \tag{3.14}$$

of type (3.5) but defined on the finite interval $[t_o,t_1]$. We shall call such representations *stochastic realizations* of $\{y(t);\ t \in [t_o,t_1]\}$. Note that it is now necessary to specify the initial condition ξ; different ξ will define different realizations S. If $N := E\{\xi\xi'\}$ is positive definite, $x(t_o)$ can be thought of as being generated by a model (3.5) on the interval $(-\infty,t_o]$; consequently, (3.14) is merely a restriction of (3.5) to the interval $[t_o,t_1]$. However, note that the class (3.14) also contains realizations S such that N is singular (even zero); then there is no such interpretation, and we shall have to take some care in defining the corresponding backward model.

The linear least-squares estimate

$$x_*(t;t_o) = \hat{E}\{x(t) \mid H_{[t_o,t]}(y)\} \tag{3.15}$$

of the state process x of S is generated on $[t_0,t_1]$ by the *Kalman-Bucy filter*

$$dx_*(t;t_0) = F(t)x_*(t,t_0)dt + B_*(t,t_0)du_*(t,t_0) \quad ; \quad x_*(t_0;t_0) = 0 \quad (3.16a)$$

where

$$du_*(t,t_0) = R(t)^{-\frac{1}{2}}[dy - H(t)x_*(t;t_0)dt] \ . \qquad (3.16b)$$

The matrix function B_*, called the Kalman gain is given by

$$B_*(t,t_0) = Q_*(t,t_0)H(t)'R(t)^{-\frac{1}{2}} + B_1(t) \ , \qquad (3.16c)$$

the error covariance matrix

$$Q_*(t,t_0) := E\{[x(t) - x_*(t;t_0)][x(t) - x_*(t;t_0)]'\} \qquad (3.16d)$$

being the solution of the matrix Riccati equation

$$\begin{cases} \frac{dQ_*}{dt}(t,t_0) = F(t)Q_*(t,t_0) + Q_*(t,t_0)F(t)' - B_*(t,t_0)B_*(t,t_0)' + B(t)B(t)' \\ Q_*(t_0,t_0) = P(t_0) \ , \end{cases} \qquad (3.16e)$$

with B_* given by (3.16c).

Let $P_*(t,t_0) := E\{x_*(t;t_0)x_*(t;t_0)'\}$. Then, it is easy to see that

$$P_*(t,t_0) = P(t) - Q_*(t,t_0) \qquad (3.17)$$

for all $t \in [t_0,t_1]$ and that

$$\frac{dP_*}{dt}(t,t_0) = F(t)P_*(t,t_0) + P_*(t,t_0)F(t)' + B_*(t,t_0)B_*(t,t_0)' \ ; \ P_*(t_0,t_0) = 0 \ . \qquad (3.18)$$

The representation (3.14) is not the only model defined on $[t_0,t_1]$ that has (3.16a) as its Kalman-Bucy filter. Hence, define $S[t_0,t_1]$ to be the class of all analytic realizations S of $\{y(t);\ t \in [t_0,t_1]\}$ which has (3.16a) as its Kalman-Bucy filter on $[t_0,t_1]$. Since we are only considering *proper* [2] realizations, not only F, H, R and B_*, but also the estimates $x_*(t;t_o)$ will be the same. Of course, different $S \in \mathcal{S}[t_o,t_1]$ will have different state processes x and different Q_*, B, w and P. Rewriting (3.16b) to obtain an expression for dy, it is easy to see that (3.16a) and (3.16b) together define a realization in $\mathcal{S}[t_o,t_1]$; we shall call it $S_*(t_o,t_1)$.

Now for any realization $S \in \mathcal{S}[t_o,t_1]$, the $n \times m$-matrix function G defined by (3.12) is invariant for the class $\mathcal{S}$. To see this, note that (3.16c) and (3.17) yield that

$$G(t) = P_*(t,t_o)H(t)' + B_*(t,t_o)R(t)^{\frac{1}{2}}, \tag{3.19}$$

and that

$$B_*(t,t_o) = (G(t) - P_*(t,t_o)H(t)')R(t)^{-\frac{1}{2}} . \tag{3.20}$$

Therefore, (3.18) may be written

$$\frac{dP_*}{dt}(t,t_o) = \Lambda(t,P_*(t,t_o)) \quad ; \quad P_*(t_o,t_o) = 0, \tag{3.21}$$

where Λ is defined by

$$\Lambda(t,P) = F(t)P + PF(t)' + (G(t) - PH(t)')R(t)^{-1}(G(t) - PH(t)')'. \tag{3.22}$$

Let $\mathcal{S}_+[t_o,t_1]$ denote the subclass of all realizations in $\mathcal{S}[t_o,t_1]$ such that $\bar{N} > 0$. For such an S the construction of backward realizations presented in Section 3.1 remains valid, for $\bar{x}$ will be well-defined on the whole interval $[t_o,t_1]$. Then, by symmetry with the forward setting, we can now see that, given an arbitrary backward realization $S \in \mathcal{S}_+[t_o,t_1]$, the estimate

$$\bar{x}_*(t;t_1) = \hat{E}\{x(t) \mid H_{[t,t_1]}(y)\} \tag{3.23}$$

of the process $\bar{x}$ is generated on $[t_o,t_1]$ by the *backward Kalman-Bucy filter*

$$d\bar{x}_*(t;t_1) = -F(t)'\bar{x}_*(t;t_1)dt + \bar{B}_*(t,t_1)d\bar{u}_*(t,t_1) \ ; \ \bar{x}_*(t_1;t_1) = 0, \quad (3.24a)$$

where

$$d\bar{u}_*(t,t_1) = R(t)^{-\frac{1}{2}}[dy - G(t)'\bar{x}_*(t;t_1)dt] \ . \quad (3.24b)$$

The matrix function $\bar{B}_*$ is given by

$$\bar{B}_*(t,t_1) = -\bar{Q}_*(t,t_1)G(t)R(t)^{-\frac{1}{2}} + \bar{B}_1(t) \ , \quad (3.24c)$$

where

$$\begin{cases} \frac{d\bar{Q}_*}{dt}(t,t_1) = -F(t)'\bar{Q}_*(t,t_1) - \bar{Q}_*(t,t_1)F(t) + \bar{B}_*(t,t_1)\bar{B}_*(t,t_1)' - \bar{B}(t)\bar{B}(t)' \\ \bar{Q}_*(t_1,t_1) = \bar{P}(t_1) = P(t_1)^{-1} \ . \end{cases} \quad (3.24d)$$

Here, of course, $\bar{Q}_*$ is the error covariance matrix

$$\bar{Q}_*(t,t_1) := E\{[\bar{x}(t) - \bar{x}_*(t;t_1)][\bar{x}(t) - \bar{x}_*(t;t_1)]'\} \ . \quad (3.24e)$$

Let $\bar{P}_*(t,t_1) := E\{\bar{x}_*(t;t_1)\bar{x}_*(t;t_1)'\}$. Then

$$\bar{P}_*(t,t_1) = \bar{P}(t) - \bar{Q}_*(t,t_1) \quad (3.25)$$

for all $t \in [t_o,t_1]$ and

$$\frac{d\bar{P}_*}{dt}(t,t_1) = \bar{\Lambda}(t,\bar{P}_*(t,t_1)) \ ; \ \bar{P}_*(t_1,t_1) = 0 \ , \quad (3.26)$$

where $\bar{\Lambda}$ is given by

$$\bar{\Lambda}(t,P) = F(t)'P + PF(t) + (H(t)' - PG(t))R(t)^{-1}(H(t)' - PG(t))' . \quad (3.27)$$

It can be shown that the matrix function $\bar{B}_*$ is invariant over the class $S[t_o,t_1]$ (cf. Lemma 2.5). Hence the backward filter (3.24a) is invariant also. Now, define $\bar{S}[t_o,t_1]$ to be the class of all analytic backward realizations $\bar{S}$ of $\{y(t);\ t \in [t_o,t_1]\}$ which has (3.24a) as its Kalman-Bucy filter.

Analogously to the forward setting, it is seen that (3.24a) and (3.24b) properly reformulated, constitute a realization in $\bar{S}[t_o,t_1]$; it will here be called $\bar{S}_*(t_o,t_1)$.

We wish to construct the forward counterpart $S^*(t_o,t_1)$ of $\bar{S}_*(t_o,t_1)$. However, we now run into problems, because $\bar{P}_*(t_1) = 0$ so that the (reverse of the) construction in Section 3.1 is no longer valid. But, as shown in [64,65], the minimality and analyticity of $\bar{S}_*$ implies that $P^* := \bar{P}_*^{-1}$ is well-defined on $[t_o,t_1-\varepsilon]$ for all $\varepsilon > 0$, and hence so is $x^* := \bar{P}_*^{-1}\bar{x}_*$. Consequently, we can define the model

$$(S^*(t_o,t_1))\begin{cases} dx^*(t;t_1) = F(t)x^*(t;t_1)dt + B^*(t,t_1)R(t)^{-\frac{1}{2}}[dy - H(t)x^*(t;t_1)dt] \\ x^*(t_o;t_1) = \bar{P}_*(t_o,t_1)^{-1}\bar{x}_*(t_o;t_1) \end{cases} \tag{3.28a}$$

on any such interval; such a realization is called a *generalized* realization of $\{y(t);\ t \in [t_o,t_1]\}$. (See [64,65].) Here

$$B^*(t,t_1) = -(Q^*(t,t_1)H(t)'R(t)^{-\frac{1}{2}} - B_1(t)) , \tag{3.28b}$$

with Q^* satisfying

$$\begin{cases} \frac{dQ^*}{dt}(t,t_1) = F(t)Q^*(t,t_1) + Q^*(t,t_1)F(t)' + B^*(t,t_1)B^*(t,t_1)' - B(t)B(t)' \\ Q^*(t_o,t_1) = \bar{P}_*(t_o,t_1)^{-1} - P(t) . \end{cases} \tag{3.28c}$$

Let $P^*(t,t_1) := E\{x^*(t;t_1)x^*(t;t_1)'\}$. Then, it can be seen that

$$P^*(t,t_1) = P(t) + Q^*(t,t_1) \tag{3.29}$$

and that

$$\frac{dP^*}{dt}(t,t_1) = \hat{\Lambda}(t,P^*(t,t_1)) \; ; \; P^*(t_o,t_1) = \bar{P}_*(t_o,t_1)^{-1}. \quad (3.30)$$

As we have seen in Chapter 2, since $Q_*(t,t_o) \geq 0$ and $\bar{Q}_*(t,t_1) \geq 0$, it can be easily seen that for all $t \in [t_o,t_1]$,

$$P_*(t,t_o) \leq P(t) \leq P^*(t,t_1) \; . \quad (3.31)$$

Hence the models S_* and S^* are called the minimum-variance and the maximum-variance respectively (cf. [64,65]).

Finally, we recall that these two models S_* and S^*, contain all the information on y that is needed to estimate x. Consequently, as was done in Chapter 2 and in [64,65], it is seen that the *smoothing estimate*

$$\hat{x}(t;t_o,t_1) = \hat{E}\{x(t) \mid H_{[t_o,t_1]}(y)\} \quad (3.32)$$

of the state process x of any realization $S \in \mathcal{S}_+[t_o,t_1]$ is given by

$$\begin{aligned} x(t;t_o t_1) = {} & [I - Q_*(t,t_o)Q(t,t_o,t_1)^{-1}]x_*(t;t_o) + \\ & + Q_*(t,t_o)Q(t,t_o,t_1)^{-1}x^*(t;t_1) \end{aligned} \quad (3.33)$$

on $[t_o,t_1)$, where

$$Q(t,t_o,t_1) = P^*(t,t_1) - P_*(t,t_o) \; . \quad (3.34)$$

3.3. Stochastic Realizations on $\mathbb{R}$

In this section, we shall let $t_o \to -\infty$ and $t_1 \to \infty$. Consequently, we shall extend the discussion of the previous section to the infinite interval setting of Section 3.1.

Since, for each fixed $t \in \mathbb{R}$, the process $\{x_*(t;-\tau); \tau \geq -t\}$ is a uniformly integrable wide sense martingale [70], $x_*(t;t_o)$ tends to a limit $x_*(t) := \hat{E}\{x(t) \mid \bigvee_{t_0 \le t} H_{[t_0,t]}(y)\}$ in mean square as $t_0 \to -\infty$ (cf. [2; p.378]). But $\bigvee_{t_0 \le t} H_{[t_0,t]}(y) = H_t^-(y)$ and consequently

$$x_*(t) = \hat{E}\{x(t) \mid H_t^-(y)\} \tag{3.35}$$

for all $t \in \mathbb{R}$. Then $u_*(t,t_0)$ tends to a limit process $\{u_*(t);\ t \in \mathbb{R}\}$ which satisfies (3.2), for $u_*(t,t_0)$ satisfies (3.2). Since $x_*(t;t_0) \to x_*(t)$, $P_*(t,t_0)$ and $B_*(t,t_0)$ as given by (3.17) and (3.20) tend to the limits $P_*(t)$ and $B_*(t)$ respectively. Consequently, x_* and u_* must satisfy

$$dx_*(t) = F(t)x_*(t)dt + B_*(t)R(t)^{-\frac{1}{2}}[dy - H(t)x_*(t)\,dt] \tag{3.36}$$

for each $t \in \mathbb{R}$. Now, we define S to be the class of all analytic realizations (3.5) of $\{y(t);\ t \in \mathbb{R}\}$ whose Kalman-Bucy filter on any interval $[t_0,t_1]$ tends to (3.36) (in the obvious sense) as $t_0 \to -\infty$ and as $t_1 \to \infty$. It is easy to see that (3.36) may be reformulated to yield the model

$$(S_*)\ \begin{cases} dx_* = Fx_*dt + B_*du_* \\ dy\ = Hx_*dt + R^{\frac{1}{2}}du_* \end{cases} \tag{3.37}$$

Let $\hat{E}\{x_*(t) \mid H_{[t_0,t]}(y)\}$ be denoted $\hat{x}_*(t;t_0)$. Then, by a similar argument to that leading to (3.35), the limit in mean square of $\hat{x}_*(t;t_0)$ as $t_0 \to -\infty$ is seen to be $\hat{E}\{x_*(t) \mid H_t^-(y)\}$, which is $x_*(t)$. Since, in addition S_* is minimal and analytic, this implies the model S_* belongs to S.

The following proposition summarizes the pertinent facts about the model S_*. The results are obtained from the corresponding finite interval results by merely taking the appropriate limits as explained above. It is the nonstationary version of Theorem 4.1 in [2].

Proposition 3.2. *There is one and only one realization* (3.5) *in* $\mathcal{S}$*, namely* (3.37) *having any of the following properties:*

(i) $x_*(t)$, $u_*(t)$, $B_*(t)$ *and the state covariance* $P_*(t)$ *of* S_* *are the limits (the first two in the mean square) of* $x_*(t;t_0)$, $u_*(t,t_0)$, $B_*(t,t_0)$ *and* $P_*(t,t_0)$ *respectively as* $t_0 \to -\infty$,

(ii) the covariance matrix function P_* *satisfies*

$$\dot{P}_*(t) = \Lambda(t,P_*(t)) \ , \tag{3.38}$$

where Λ *is given by* (3.22)*, and it is minimum in the sense that*

$$P_*(t) \le P(t) \ (\textit{for each } t \in \mathbb{R}) \ , \tag{3.39}$$

(iii) the innovation process u_* *satisfies*

$$H_t^-(du_*) = H_t^-(y) \tag{3.40}$$

for all $t \in \mathbb{R}$*, and*

(iv) for any realization $S \in \mathcal{S}$*, with state process* x*, the process* x_* *is the estimate*

$$\hat{E}\{x(t) \mid H_t^-(y)\} = x_*(t) \tag{3.41}$$

(i.e., x_* *is invariant with respect to the particular realization* $S \in \mathcal{S}$*.).*

In analogy with the forward setting, we see that the process $\bar{x}_*(t;t_1)$ tends to a limit $\bar{x}_*(t)$ in mean square as $t_1 \to \infty$. Consequently, $\bar{u}_*(t,t_1)$, $\bar{B}_*(t,t_1)$ and $\bar{P}_*(t,t_1)$ tends to the limits $\bar{u}_*(t)$, $\bar{B}_*(t)$ and $\bar{P}_*(t)$ as $t_1 \to \infty$. Hence, we obtain a representation analogous to (3.36) for $\bar{x}_*$; we shall call it the *steady-state backward Kalman-Bucy filter*. Therefore, we define $\bar{S}$ to be the class of all analytic backward realizations (3.11) of $\{y(t);\ t \in \mathbb{R}\}$ whose Kalman-Bucy estimate on the interval $[t_o,t_1]$ tends to the steady-state backward Kalman-Bucy filter as $t_1 \to \infty$. Precisely as in the forward setting, we can see that $\bar{x}_*$ is the state process of the model

$$(\bar{S}_*) \begin{cases} d\bar{x}_* = -F'\bar{x}_* dt + \bar{B}_* d\bar{u}_* \\ dy \ = G'\bar{x}_* dt + R^{\frac{1}{2}} d\bar{u}_* \end{cases} \tag{3.42}$$

and that $\bar{S}_* \in \bar{S}$. The state covariance function $\bar{P}_*$ of $\bar{S}_*$ satisfies

$$\dot{\bar{P}}_*(t) = \bar{\Lambda}(t,\bar{P}_*(t)) \tag{3.43}$$

where $\bar{\Lambda}$ is given by (3.27).

Since $\bar{P}_*(t) > 0$ for all $t \in \mathbb{R}$, we see that the process

$$x^*(t) = \bar{P}_*(t)^{-1}\bar{x}_*(t) \tag{3.44}$$

is well-defined, and is the state of the model

$$(S^*) \begin{cases} dx^* = Fx^* dt + B^* du^* \\ dy \ = Hx^* dt + R^{\frac{1}{2}} du^* \ , \end{cases} \tag{3.45}$$

where B^* and u^* are the limits of $B^*(t,t_1)$ and $u^*(t,t_1)$ as $t_1 \to \infty$. The state covariance function P^* of (3.45) is $\bar{P}_*^{-1} = (\lim_{t_1\to\infty} \bar{P}_*(t,t_1))^{-1} = \lim_{t_1\to\infty} (\bar{P}_*(t,t_1))^{-1} = \lim_{t_1\to\infty} P^*(t,t_1)$. It is easy to see that P^* satisfies

$$\dot{P}^*(t) = \Lambda(t,P^*(t)) \tag{3.46}$$

and that $Q^*(t) := E\{[x(t) - x^*(t)][x(t) - x^*(t)]'\}$ is given by

$$Q^*(t) = P^*(t) - P(t) \ . \tag{3.47}$$

Since $x^*(t) = \lim_{t_1 \to \infty} x^*(t;t_1)$ and $\hat{E}\{x^*(t;t_1) \mid H_{[t_o,t]}(y)\} = x_*(t;t_o)$, we see that the Kalman-Bucy filter of S^* on $[t_o,t_1]$ tends to x_* in mean square as $t_o \to -\infty$. Hence $S^* \in \mathcal{S}$.

As a corollary to Proposition 3.2 and the above discussion, it is now clear that

$$P_*(t) \le P(t) \le P^*(t) \tag{3.48}$$

for all $t \in \mathbb{R}$.

As another corollary, we shall obtain the following algorithms to calculate P_* and P^*. They were originally obtained by Clerget [71] through an argument first used in [11]. The proof in [71] relies on control theory techniques and gives little insight into the stochastic nature of the problem. Our proof is not only simpler, but it also provides a stochastic interpretation for these equations.

Corollary 3.3. *Let* Π *and* $\bar{\Pi}$ *be the unique solutions of the* $n \times n$-*matrix partial differential equations*

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right)\Pi(t,s) = \Lambda(t,\Pi(t,s)) \ ; \ \Pi(t,0) = 0 \tag{3.49}$$

and

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right)\bar{\Pi}(t,s) = \Lambda(t,\bar{\Pi}(t,s)) \ ; \ \bar{\Pi}(t,0) = 0 \tag{3.50}$$

respectively, where Λ *and* $\bar{\Lambda}$ *are given by* (3.22) *and* (3.27) *respectively. Then* $\Pi(t,s) \to P_*(t)$ *and* $\bar{\Pi}(t,s) \to P^*(t)^{-1}$ *as* $s \to \infty$.

Proof: Let $t_o \in \mathbb{R}$. Set $s = t - t_o$ and $\Pi(t,s) := P_*(t,t - s)$. Then $P_*(t,t_o) = \Pi(t,t - t_o)$. Hence $\frac{dP_*}{dt}(t,t_o) = \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right)\Pi(t,t - t_o)$.

But, by (3.22), $\frac{dP_*}{dt}(t,t_o) = \Lambda(t,P_*(t,t_o)) = \Lambda(t,\Pi(t,t-t_o))$. Since $P_*(t,t) = 0$, $\Pi(t,0) = 0$. Hence (3.49) follows. To see that $\Pi(t,s) \to P_*(t)$ as $s \to \infty$, just observe that, by Proposition 3.2, $P_*(t,t_o) \to P_*(t)$ as $t_o \to -\infty$. Finally, (3.50) follows from (3.43) and an argument analogous to the above. □

Finally, we shall solve the smoothing problem for this infinite interval setting.

Let $S \in \mathcal{S}$ be arbitrary. The smoothing problem requires determining the estimate

$$\hat{x}(t) := \hat{E}\{x(t) \mid H(y)\} \tag{3.51}$$

of the state x of S. Recall that the smoothing estimate $x(t;t_o,t_1)$ on the finite interval $[t_o,t_1]$ is defined by (3.32). Since, by the argument used above

$$\hat{x}(t;t_o,t_1) = \hat{E}\{x(t) \mid H_{[t_o,t_1]}(y)\} \to \hat{E}\{x(t) \mid H(y)\} = \hat{x}(t)$$

as $t_o \to -\infty$ and as $t_1 \to \infty$, we see that (3.33) yields the following formula ((3.53)), which holds under the assumption that the covariance function K_y of y is *coercive*, i.e., there exists a positive constant ζ such that

$$\int_{t_o}^{t}\int_{t_o}^{t} u'(r)K_y(r,s)u(s)\,dr\,ds > \zeta\,\|u(t)\|^2 \tag{3.52}$$

for all square integrable functions u. It was shown by Clerget [71; p.29] that this assumption implies $P^* - P_* > 0$.

Proposition 3.4. *Assume that the covariance function* K_y *of* y *is coercive. Let* x *be the state process of an arbitrary realization* $S \in \mathcal{S}$, *and let* $\hat{x}$ *be the corresponding smoothing estimate* (3.51). *Then*

$$\hat{x}(t) = [I - Q_*(t)Q(t)^{-1}]x_*(t) + Q_*(t)Q(t)^{-1}x^*(t) , \qquad (3.53)$$

where $Q_* = P - P_*$ *and* $Q = P^* - P_*$.

Remark. Formula (3.53) suggests a representation for the state process of an *internal* realization (i.e., a realization satisfying $H(x) \subset H(dy)$.) If $S \in \mathcal{S}$ is internal, then, by (3.51), $x(t) = \hat{x}(t)$ for each $t \in \mathbb{R}$.

Corollary 3.5. *Assume that the covariance function* K_y *of* y *is coercive. Let* x *be the state process of an internal realization* $S \in \mathcal{S}$. *Then* x *can be written*

$$x(t) = \Pi(t)x_*(t) + [I - \Pi(t)]x^*(t) , \qquad (3.54)$$

where Π *is a projection given by* $\Pi = Q^*Q^{-1}$ *and* $Q^* = P^* - P$.

Proof. In view of Proposition 3.4, it only remains to show that $\Pi = Q^*Q^{-1}$ is a projection. Let $\Sigma(t) := E\{[x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]'\}$. Then, it can be shown (Chapter 2) that $\Sigma = Q_* - Q_*Q^{-1}Q_*$. Then, since $\hat{x} = x$, $\Sigma = 0$. Consequently, $Q_*Q^{-1} = Q_*Q^{-1}Q_*Q^{-1}$. But $I - Q_*Q^{-1} = Q^*Q^{-1}$, hence $Q^*Q^{-1} = Q^*Q^{-1}Q^*Q^{-1}$, i.e., $\Pi^2 = \Pi$. $\square$

3.4. The Set $\mathcal{P}$

Let $\mathcal{P}$ be the set of all state covariance functions P given by (3.6) as S ranges over $\mathcal{S}$. Any $P \in \mathcal{P}$ must also satisfy (3.12); (3.6) and (3.12) constitute the equations of the nonstationary version of the Positive Real Lemma. Since, in addition, the state covariance P of any $S \in \mathcal{S}$ is positive definite, we may write $\mathcal{P} = \{P = P' > 0 \mid P$ solves (3.6) and (3.12)$\}$. It can easily be checked that $\mathcal{P}$ is bounded and convex [71]. Some straightforward algebraic manipulations yield

$$\mathcal{P} = \{P = P' \mid \Lambda(t,P(t)) - \dot{P}(t) \leq 0 \text{ for all } t \in \mathbb{R}\}. \tag{3.55}$$

For each $P \in \mathcal{P}$, define the *feedback matrix*

$$\Gamma = F - (G - PH')R^{-1}H \ . \tag{3.56}$$

Let the feedback matrices corresponding to P_* and P^* be denoted Γ_* and Γ^* respectively. Let $\mathcal{P}_+ = \{P \mid P > P_*\}$ and $\mathcal{P}_- = \{P \mid P < P^*\}$. If the covariance function of y is coercive, i.e., satisfies (3.52), $P^* - P_* > 0$. Consequently $\mathcal{P}_+$ and $\mathcal{P}_-$ are both nonempty.

Now let $\mathcal{P}_0$ be the subset of $\mathcal{P}$ defined by

$$\mathcal{P}_0 = \{P \in \mathcal{P} \mid \Lambda(t,P(t)) - \dot{P}(t) \equiv 0\} \ . \tag{3.57}$$

Then it is immediately seen that $\mathcal{P}_0 = \{P \in \mathcal{P} \mid B_2 = 0\}$.

Lemma 3.6. *Let $S \in \mathcal{S}$ with state covariance function P. Then S is internal if and only if $P \in \mathcal{P}_0$.*

Proof. Let $P \in \mathcal{P}_0$. Then $B_2 = 0$, hence S has the form

$$\begin{cases} dx = Fxdt + B_1 du \\ dy = Hxdt + R^{\frac{1}{2}} du \end{cases} .$$

Consequently, x can be solved for in terms of y as follows

$$dx = Fxdt + B_1 R^{-\frac{1}{2}}[dy - Hxdt]$$

which clearly implies that S is internal. Since the smoothing formula (3.53) is the same as (3.33), the argument in proving Theorem 4.4 in [64] can be used to prove the converse. $\square$

Corollary 3.7. *P_* and P^* belong to P_0.*

Proof. This follows from (3.38) and (3.46). $\square$

It is worth noting that once the covariance function P of a realization $S \in \mathcal{S}$ is known, the quadruplet $[F, B, H, (R^{\frac{1}{2}}, 0)]$ is determined upon observing that

$$B_1 = (G - PH')R^{-\frac{1}{2}} \tag{3.58a}$$

$$B_2 B_2' = \dot{P} - \Lambda(P) . \tag{3.58b}$$

Determining the quadruplets $[F, B, H, (R^{\frac{1}{2}}, 0)]$ solves what we have called before the *wide sense stochastic realization problem,* i.e., is equivalent to finding all realizations whose outputs have the same covariance properties. In the next section, we shall present an algorithm to generate such quadruplets.

Finally, it is not hard to see that Proposition 1.12 and Corollaries 1.13 and 1.14 hold for this setting also with obvious modifications.

3.5. Non-Riccati Algorithm Inside $\mathcal{B}$

In this section, we shall present an algorithm to generate (wide sense) realizations without resort to the intermediate step of solving for P. This algorithm is the nonstationary, continuous-time version of (1.54) and Theorem 6.2 in [2]. In order for the idea of proof used in [2] to be applicable in our nonstationary setting, we shall have to introduce an additional condition on the given process $\{y(t);\ t \in \mathbb{R}\}$: Assume that y is generated as the output of a stochastic system (3.5) such that B_2 is *constant* and *nonzero*. Of course this does not imply that *all* $S \in \mathcal{S}$ have such a B_2, merely that there are nontrivial ($B_2 \neq 0$), elements in the subclass $\mathcal{S}_c := \{S \in \mathcal{S} \mid B_2 \text{ constant}\}$. Let $\mathcal{P}_c$ be the subset of $\mathcal{P}$ corresponding to realizations in $\mathcal{S}_c$.

Lemma 3.8. *Let* $P \in \mathcal{P}_c$. *Then* $\Lambda(t,P(t)) - \dot{P}(t)$ *is constant on the real line.*

To develop the algorithm, we shall first construct, for a given matrix function $P_o \in \mathcal{P}_c$, a trajectory of matrix functions in $\mathcal{P}$ extending from P_* through P_o to P^*, so that these functions are totally ordered in a sense to be defined below.

Theorem 3.9. *Assume that the covariance function* K_y *of* y *is coercive. Let* Λ *be defined by* (3.22). *Let* P_o *be an arbitrary function in* $\mathcal{P}_c$. *Let* P *be the unique solution in the* (t,s) *plane of the matrix partial differential equation*

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right) P(t,s) = \Lambda(t,P(t,s)) \ ; \ P(t,0) = P_o(t) \ . \tag{3.59}$$

Then, (i) $P(\cdot,s) \in \mathcal{P}$ *for all* $s \in (-\infty,\infty)$, *(ii) for each* $t \in \mathbb{R}$, $P(t,s_2) \le P(t,s_1)$ *for* $s_1 \le s_2$, *(iii) if* $P_o \in \mathcal{P}_-$, *for each* $t \in \mathbb{R}$, $P(t,s) \to P_*(t)$ *as* $s \to \infty$ *and (iv) if* $P_o \in \mathcal{P}_+$, *for each* $t \in \mathbb{R}$, $P(t,s) \to P^*(t)$ *as* $s \to -\infty$.

Remark. Before proving this theorem, it is worth noting that the partial differential equation (3.59) can be trivially transformed to an (infinite) family of ordinary differential equations (3.18) with different initial conditions. To see this, set $\sigma = t - s$ and let $\tilde{P}(t,\sigma) := P(t,t-\sigma)$. Then, it is easy to see that the left-hand side of (3.59) is $\frac{d\tilde{P}}{dt}(t,\sigma)$. Finally, $P(\sigma,0) = \tilde{P}(\sigma,\sigma)$. Hence, (3.59) is the ordinary differential equation

$$\frac{d\tilde{P}}{dt}(t,\sigma) = \Lambda(t,\tilde{P}(t,\sigma)) \ ; \ \tilde{P}(\sigma,\sigma) = P_o(\sigma) \ . \tag{3.60}$$

The following two lemmas will be needed for the proof of Theorem 3.9.

Lemma 3.10. *The matrix partial differential equation* (3.59) *has a unique solution* $P(t,s)$ *which is a matrix of analytic functions in the two real variables* t *and* s.

Proof. It is well known that the ordinary differential equation (3.60) has a unique solution [4; p.156] which is also analytic (for the parameter matrices are). Then the same is true for (3.59). □

Lemma 3.11. *Let* $P_o \in \mathcal{P}_c$. *Then the matrix partial differential equation* (3.59) *can be replaced (in the sense that it has the same solution* P) *by the system*

$$\begin{cases} \dfrac{\partial P}{\partial s}(t,s) = U(t,s)[\Lambda(t,P_o(t)) - \dot{P}_o(t)]U(t,s)' \; ; \; P(t,0) = P_o(t) & (3.61a) \\ \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial s}\right)U(t,s) = \Gamma(t,s)U(t,s) \; ; \; U(t,0) = I & (3.61b) \end{cases}$$

where $\Gamma(t,s)$ *is the feedback matrix* (3.56) *corresponding to* $P(t,s)$.

Proof. We shall use the differentiation technique of [21]. The reason why this method works in this nonstationary setting is of course that the coefficient matrices do not depend on s, which is the dynamic variable of the algorithm. First, reformulate relation (3.59) to read

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial s}\right)P(t,s) = [F(t) - G(t)R(t)^{-1}H(t)]P(t,s) + P(t,s)[F(t) - G(t)R(t)^{-1}H(t)]'$$
$$+ P(t,s)H(t)'R(t)^{-1}H(t)P(t,s) + G(t)R(t)^{-1}G(t)' \; . \qquad (3.62)$$

Since $P(t,s)$ is a matrix of analytic functions, the mixed partial derivatives of $P(t,s)$ with respect to t and s are identical. Using this fact, differentiating (3.62) with respect to s and setting $N(t,s) := \frac{\partial P}{\partial s}(t,s)$, it can be seen that

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right)N(t,s) = \Gamma(t,s)N(t,s) + N(t,s)\Gamma(t,s)' \; . \qquad (3.63a)$$

In view of (3.59), this partial differential equation has the boundary condition

$$N(t,0) = \Lambda(t,P_o(t)) - \dot{P}_o(t) \; , \qquad (3.63b)$$

which, by Lemma 3.8, is constant for $P_o \in \mathcal{P}_c$. Consequently, (3.63) can be integrated to yield

$$N(t,s) = U(t,s)N(t,0)U(t,s)' , \tag{3.64}$$

where U is given by (3.61b). Consequently, P satisfies (3.61). But (3.61) clearly has a unique solution, and therefore the lemma follows. □

Proof of Theorem 3.9. Let $P_o \in \mathcal{P}_c$. Then $\Lambda(t,P_o(t)) - \dot{P}_o(t) \le 0$, and consequently, by (3.61a),

$$\frac{\partial P}{\partial s}(t,s) \le 0 , \tag{3.65}$$

which, in view of (3.59) and (3.55), implies that $P(\cdot,s) \in \mathcal{P}$ for all $s \in (-\infty,\infty)$, i.e., (i) holds. Property (ii) is an immediate consequence of (3.65). To prove (iii), we follow [72]. First note that (i), (ii), and (3.39) imply that, for each $t \in \mathbb{R}$, $s \mapsto P(t,s)$ is a nonincreasing function bounded from below by $P_*(t)$, and consequently $P(t,s)$ tends to a limit $\tilde{P}(t)$ as $s \to \infty$. It remains to show that $\tilde{P} = P_*$. Keeping t fixed and letting $s \to \infty$ in (3.59), it is not hard to see that $\tilde{P}$ satisfies the differential equation $\frac{d\tilde{P}}{dt} = \Lambda(t,\tilde{P}(t))$; hence $\tilde{P} \in \mathcal{P}_o$. Then, by Lemma 3.6, $\tilde{P}$ is the state covariance function of an internal realization; let $\tilde{x}$ denote its state process. But, then $\tilde{x}$ satisfies (3.54) for some family $\{\Pi(t);\ t \in \mathbb{R}\}$ of projections, i.e., $x^* - \tilde{x} = \Pi(x^* - x_*)$, which yields

$$P^*(t) - \tilde{P}(t) = \Pi(t)[P^*(t) - P_*(t)]\Pi(t)' . \tag{3.66}$$

To see this, use the orthogonality relations of Lemma 3.5 in [64]. Now, use the fact that $P_o \in \mathcal{P}_-$ and $\tilde{P}(t) \le P_o(t)$ for all $t \in \mathbb{R}$ (for $P(t,\cdot)$ is a nonincreasing function), to see that $\tilde{P}(t) < P^*(t)$ for all $t \in \mathbb{R}$. Consequently, for each $t \in \mathbb{R}$,

$$\Pi(t)[P^*(t) - P_*(t)]\Pi(t)' > 0$$

which can hold only if $\Pi(t)$ is nonsingular. But $\Pi(t)$ is a projection, i.e., $\Pi(t)^2 = \Pi(t)$, which together with nonsingularity implies that $\Pi(t) \equiv I$. Hence, it follows from (3.66) that $\tilde{P} = P_*$, as required. The proof of (iv) is analogous. □

We are now ready to formulate the non-Riccati algorithm to generate families of (wide sense) realizations. As all realizations are determined by the matrix B, the algorithm will be given in terms of this parameter only.

Let $\mathcal{B} = \{B = (B_1, B_2) \mid B_1$ and B_2 are given by (3.58) as P ranges over $\mathcal{P}\}$. Let $\mathcal{B}_o$, $\mathcal{B}_-$ and $\mathcal{B}_+$ be defined analogously in terms of $\mathcal{P}_o$, $\mathcal{P}_-$ and $\mathcal{P}_+$ respectively. It is clear that $\mathcal{B}_o$ consists of those $B \in \mathcal{B}$ for which $B_2 = 0$. In particular, let B_* and B^* denote the elements of $\mathcal{B}_o$ corresponding to P_* and P^*.

Theorem 3.12. *Assume that the covariance function* K_y *of* y *is coercive. Let* $[F, B^o, H, (R^{1/2}, 0)]$ *be an arbitrary (wide sense) realization of* y *such that* B_2^o *is constant on* $(-\infty, \infty)$ *and let* $s \mapsto B(t,s) = [B_1(t,s), B_2(t,s)]$ *be the unique solution of the system*

$$\begin{cases} \dfrac{\partial B_1}{\partial s}(t,s) = B_2(t,s)B_2(t,s)'H(t)'R(t)^{-1/2} & ; \ B_1(t,0) = B_1^o(t) \quad (3.67a) \\ \left(\dfrac{\partial}{\partial s} + \dfrac{\partial}{\partial t}\right)B_2(t,s) = [F(t) - B_1(t,s)R(t)^{-1/2}H(t)]B_2(t,s) & ; \ B_2(t,0) = B_2^o. \quad (3.67b) \end{cases}$$

For each $s \in (-\infty, \infty)$, *let* P(t,s) *be the unique solution of*

$$\frac{\partial P}{\partial t}(t,s) = F(t)P(t,s) + P(t,s)F(t)' + B(t,s)B(t,s)' \ . \qquad (3.68)$$

Then, for each $s \in (-\infty,\infty)$*,* $[F, B(\cdot,s), H, (R^{1/2},0)]$ *is a (wide sense) realization of* y*, with state covariance function* $P(\cdot,s)$*. If* $B_o \in \mathcal{B}_-$*, for each* $t \in \mathbb{R}$*,* $B(t,s) \to B_*(t)$ *as* $s \to \infty$*, and if* $B^o \in \mathcal{B}_+$*, for each* $t \in \mathbb{R}$*,* $B(t,s) \to B^*(t)$ *as* $s \to -\infty$*. Finally, the function* P *satisfies conditions (i) - (iv) of Theorem 3.9 and the equation*

$$\frac{\partial P}{\partial s}(t,s) = -B_2(t,s)B_2(t,s)' \ . \tag{3.69}$$

Proof. Let P_o be the state covariance function of the initial real-realization $[F, B^o, H, (R^{1/2},0)]$ and let $s \mapsto P(t,s)$ be the trajectory through P_o defined by Theorem 3.9. Define

$$B_1(t,s) = [G(t) - P(t,s)H(t)']R(t)^{-1/2} \tag{3.70a}$$

and

$$B_2(t,s) = U(t,s)B_2^o \ , \tag{3.70b}$$

where $U(t,s)$ is given by (3.61b). Then, $\frac{\partial P}{\partial s}(t,s) = U(t,s)N(t,0)U(t,s)' = -U(t,s)B_2^oB_2^{o\prime}\ U(t,s)' = -B_2(t,s)B_2(t,s)'$. This proves (3.69). Differentiate (3.70a) with respect to s and use (3.69) to get (3.67a). To prove (3.67b), differentiate (3.70b) with respect to s and use (3.61b). In view of (3.59) and (3.69), $\frac{\partial P}{\partial s}(t,s) = -B_2(t,s)B_2(t,s)' = \Lambda(t,P(t,s)) - \frac{\partial P}{\partial t}(t,s)$. Hence $\frac{\partial P}{\partial t}(t,s) = \Lambda(t,P(t,s)) + B_2(t,s)B_2(t,s)'$ which is (3.68). Hence $(P(t,s), B(t,s))$ satisfy (3.6) and (3.12), and consequently $[F(t), B(t,s), H(t), (R(t)^{1/2},0]$ is a realization of y with state covariance $P(t,s)$ which satisfies conditions (i)-(iv) of Theorem 3.9. Finally, the fact that $B_1(t,s) \to B_*(t)$ $(B^*(t))$ under the stated conditions follows from condition (iii) ((iv)) of Theorem 3.9. Since $\frac{\partial P}{\partial s}(t,s) \to 0$ as $s \to +\infty(-\infty)$, $B_2(t,s) \to 0$ as $s \to +\infty(-\infty)$. Hence $B(t,s) \to (B_*(t),0)$ $((B^*(t),0))$ as $s \to +\infty(-\infty)$. $\square$

Remark. This theorem has the following interpretation. Each realization (3.5) in S_c gives rise to a family of realizations indexed by $s \in (-\infty,\infty)$

$$(S_s)\begin{cases} dx_s = F(t)x_s(t)dt + B_1(t,s)du_s + B_2(t,s)dv_s & (3.71a) \\ dy = H(t)x_s(t)dt + R(t)^{\frac{1}{2}}du_s & (3.71b) \end{cases}$$

which are totally ordered in the sense that the state covariances $P(t,s) = E\{x_s(t)x_s(t)'\}$ of S_s satisfy $P(t,s_2) \leq P(t,s_1)$ for $s_1 \leq s_2$.

If $B^o \in \mathcal{B}_-$, this family will contain the minimum-variance realization S_*, and if $B^o \in \mathcal{B}_+$, the family contains the maximum-variance realization S^*. Finally, if $B^o \in \mathcal{B}_0$, (3.71) will contain only one realization: (3.5) itself.

3.6. The Singular and The Mixed Cases

The stochastic realization problem may be classified into three categories:

(i) the *regular* case, for which $R(t)$ is positive definite for all $t \in \mathbb{R}$,

(ii) the *singular* case, for which $R(t) \equiv 0$, and

(iii) the *mixed* case, for which $R(t) \neq 0$, but $\det R(t) = 0$ for all $t \in \mathbb{R}$.

(In fact, there is a fourth case, which we shall not deal with here, i.e., the case for which $\det R(t) = 0$ for *some, but not all* $t \in \mathbb{R}$.)

By the assumptions made in Section 3.1, we have just studied the first of these cases. The analysis used depends heavily on the fact that $R(t)$ is invertible for all $t \in \mathbb{R}$. To solve the problems of the last two categories, we shall first have to reformulate them in the following way: the given realization [F, G, H, R] will be converted to an equivalent realization (in the sense of the definition given below) $[F_a, G_a, H_a, R_a]$ where $R_a(t)$ is positive definite for all $t \in \mathbb{R}$. The results of this section are generalizations of similar results presented in [18].

Definition 3.13. ([18]). *The two realizations* $[F_1, G_1, H_1, R_1]$ *and* $[F_2, G_2, H_2, R_2]$ *are said to be equivalent if they have the same set* P.

Let us start with the singular case, i.e., let $R(t) = 0$ for all $t \in \mathbb{R}$. Then the equations (3.6) and (3.12) of the Positive Real Lemma become

$$\dot{P} = FP + PF' + BB' \tag{3.72a}$$

$$G - PH' = 0 \tag{3.72b}$$

$$P = P' > 0 . \tag{3.72c}$$

Proposition 3.14. *Let the entries of* G *and* H *be differentiable at least* n *times. Then, the realizations* [F, G, H, 0] *and* $[F, G_a, H_a, R_a]$, *where*

$$G_a = \dot{G} - FG , \tag{3.73a}$$

$$H_a = HF + \dot{H} , \tag{3.73b}$$

and

$$R_a = H\dot{G} - G'\dot{H}' - HFG - G'F'H' \qquad (3.73c)$$

are equivalent (in the sense of Definition 3.13).

Proof. All realizations arising from the quadruplet [F, G, H, 0] have the form

$$dx = Fxdt + Bdw \qquad (3.74a)$$

$$\dot{y} = Hx \qquad (3.74b)$$

with state covariance function P, which, together with F, G, and H satisfy equations (3.72). Let $y_a = \dot{y}$. Then, using relation (3.74b), the following is a realization of the process y_a

$$dx = Fxdt + Bdw \qquad (3.75a)$$

$$dy_a = H_a xdt + HBdw\ . \qquad (3.75b)$$

where H_a is given by (3.73b). Observe that (3.74) and (3.75) have the same state process x and hence the same state covariance P. Hence, these realizations of y and y_a have the same set $\mathcal{P}$. We shall show that all realizations obtained using the quadruplet $[F, G_a, H_a, R_a]$ are of the form (3.75). First, observe that G_a as given by (3.73a) can be written as $G_a = PH_a' + BB'H'$. To see this, we use the following sequence of equalities: $G_a = \dot{G} - FG = P\dot{H}' + \dot{P}H' - FPH' = P[F'H' + \dot{H}'] + \dot{P}H' - FPH' - PF'H' = PH_a' + BB'H'$, which is the G that corresponds to (3.75), obtained from the Positive Real Lemma. It remains to show that $R_a = D_aD_a'$ where $D_a = HB$. This is shown by the following sequence of identities:

$$\begin{aligned} R_a &= H\dot{G} - G'\dot{H}' - HFG - G'F'H' \\ &= H\dot{P}H' - HFPH' - HPF'H' \\ &= H[\dot{P} - FP - PF']H' \\ &= HBB'H' = D_a D_a' . \quad \square \end{aligned}$$

Consequently, by the above proposition, instead of studying the quadruplet [F, G, H, 0], we can study $[F, G_a, H_a, R_a]$. If $R_a = 0$, the above procedure of differentiating the output may be repeated until a positive definite R_a is obtained. If $R_a \neq 0$, but $\det R_a \equiv 0$, the procedure of the third category, which will be discussed shortly, may be used.

With this setup, all the results of the previous sections can be carried over using G_a, H_a and R_a instead of G, H and 0.

Next, we consider case (iii) which is the most general. Here we need more assumptions than in case (ii). Assume R(t) has constant rank for all $t \in \mathbb{R}$ and the entries of $R(\cdot)$ are differentiable at least n times. Set k := rank R(t).

Since $R(t) \neq 0$ but $\det R(t) = 0$, by Dolezal's Theorem (see e.g. [73]), there exists a nonsingular matrix function S, the entries of which are differentiable at least n times, such that $\hat{R} = SRS = \begin{bmatrix} R_1 & 0 \\ 0 & 0 \end{bmatrix}$ with R_1 a positive definite $k \times k$ matrix function. This transformation corresponds to a change of basis in the subspace spanned by y: instead of examining the process y, we examine another process $\hat{y} = Sy$. Without loss of generality, we shall assume $R = \begin{bmatrix} R_1 & 0 \\ 0 & 0 \end{bmatrix}$ with R_1 as above.

Partition the matrices H, G and D in the following manner:

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \text{ with } H_1 \text{ is } k \times n \text{ and } H_2 \text{ is } (m - k) \times n,$$

$$G = [G_1, G_2] \text{ with } G_1 \text{ is } n \times k \text{ and } G_2 \text{ is } n \times (m - k) \text{ and}$$

$$D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} \text{ with } D_1 \text{ is } k \times p \text{ and } D_2 \text{ is } (m - k) \times p.$$

Consequently, equations (3.6) and (3.12) become

$$\dot{P} = FP + PF' + BB' \tag{3.76a}$$

$$G_1 - PH_1' = BD_1' \tag{3.76b}$$

$$D_1D_1' = R_1 > 0 \tag{3.76c}$$

$$D_2 = 0 \tag{3.76d}$$

$$G_2 - PH_2' = 0 \tag{3.76e}$$

$$P = P' > 0 \tag{3.76f}$$

The following is the generalized version of Proposition 3.14.

Proposition 3.15. *Let the entries of* G, H *and* R *be differentiable at least* n *times and let* R(t) *be of constant rank for all* $t \in \mathbb{R}$. *Let* G_1, G_2, H_1, H_2 *and* R_1 *be as defined above. Then the realizations* [F, G, H, R] *and* [F, G_a, H_a, R_a], *where*

$$\dot{G}_a = [G_1 \dot{G}_2 - FG_2] \ , \tag{3.77a}$$

$$H_a = \begin{bmatrix} H_1 \\ H_2F + \dot{H}_2 \end{bmatrix} \tag{3.77b}$$

and

$$R_a = \begin{bmatrix} R_1 & G_1'H_2' - H_1G_2 \\ H_2G_1 - G_2'H_1' & H_2\dot{G}_2 - G_2\dot{H}_2' - H_2FG_2 - G_2'F'H_2' \end{bmatrix} \ , \tag{3.77c}$$

are equivalent (in the sense of Definition 3.13).

Proof. The proof is an immediate consequence of that of Proposition 3.14 and follows upon defining $y_a = \begin{bmatrix} y_1 \\ \dot{y}_2 \end{bmatrix}$, where y has been partitioned compatibly with the rest of the matrices, as $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. □

If R_a obtained above is not full rank, we repeat the above procedure of changing the basis and differentiating the component of the output that does not contain a white noise until we arrive at a nonsingular R_a. The natural question that arises is whether this procedure terminates in a finite number of steps. Silverman [74] has shown that, subject to some extra regularity conditions, the answer is yes.

As a special case, we may quite easily obtain all the results of Germain [18] for the stationary singular and mixed cases. These results are summarized in the following two corollaries.

Corollary 3.16. *Let* F, G *and* H *be constant and let* $R = 0$. *Then the quadruplets* $[F, G, H, 0]$ *and* $[F, -FG, HF, -HFG - G'F'H']$ *are equivalent.*

Corollary 3.17. *Let* F, G, H *and* R *be constant and let* $R \neq 0$ *but* $\det R = 0$. *Then the quadruplets* $[F, G, H, R]$ *and* $[F, G_a, H_a, R_a]$, *where*

$$G_a = [G_1 - FG_2] \;,\; H_a = \begin{bmatrix} H_1 \\ H_2F \end{bmatrix} \text{ and}$$

$$R_a = \begin{bmatrix} R_1 & G_1'H_2' - H_1G_2 \\ H_2G_1 - G_2'H_1' & -H_2FG_2 - G_2'F'H_2' \end{bmatrix}$$

are equivalent.

Remark. Germain [18] has obtained the following values for G_a, H_a and R_a; a choice that we are not able to explain or understand:

$$G_a = [G_1, C_1] \ , \ H_a = [H_1, -C_2] \ , \ R_a = \begin{bmatrix} R_1 & 0 \\ 0 & B \end{bmatrix} \text{ where}$$

$$C_1 = -FG_2 - G_1R_1^{-1}[G_1'H_1' - H_1G_2]$$

$$C_2 = F'H_2' - H_1'R_1^{-1}[G_1'H_1' - H_1G_2]$$

$$B = -[H_2FG_2 + G_2'F'H_2] - [H_2G_1 - G_2'H_1']R_1^{-1}[H_2G_1 - G_2'H_1']' \ .$$

REFERENCES

1. G. Ruckebusch, A state space approach to the stochastic realization problem, *Proc. 1978 Intern. Symp. Circuits and Systems*, New York.

2. A. Lindquist and G. Picci, On the stochastic realization problem, *SIAM J. Control*, 17(1979), 365-389.

3. B.D.O. Anderson, The inverse problem of stationary covariance generation, *J. Statistical Physics*, 1(1969), 133-147.

4. R.W. Brockett, Finite Dimensional Linear Systems, Wiley, 1970.

5. L. Hitz and B.D.O. Anderson, Discrete positive real functions and their application to system stability, *Proc. IEEE*, Vol.116, No.1(1969), 153-155.

6. B.L. Ho and R.E. Kalman, Effective construction of linear state variable models from input/output data, *Proc. 3rd Allerton Conference*, 1965, 449-459.

7. J. Rissanen, Recursive identification of linear systems, *SIAM J. Control*, Vol.9, No.3(August 1971), 420-430.

8. B.D.O. Anderson and S. Vongpanitlerd, Network Analysis and Synthesis, Prentice-Hall, 1973.

9. M. Pavon, Stochastic realization and invariant directions of the matrix Riccati equation, *SIAM J. Control* (to appear).

10. G. Ruckebusch, Representations Markoviennes de processus Gaussiens stationnaires, Ph.D. Thesis, Univ. of Paris, VI, 1975.

11. P. Faurre, Réalisations Markoviennes de processus stationnaires, Research Report No.13, March 1973, IRIA(LABORIA), Le Chesnay, France.

12. J.C. Willems, Dissipative dynamical systems, Part II: Linear systems with quadratic supply rates, *Archive for Rational Mechanics and Analysis*, 45(1972), 352-393.

13. S. Barnett, Matrices in Control Theory, Van Nostrand Reinhold Co., London, 1971.

14. A.G.J. MacFarlane, An eigenvector solution of the optimal linear regulator problem, *J. Electron. Control*, 14(1963), 643-654.

15. K. Mårtensson, On the matrix Riccati equation, *Inform. Sci.*, 3(1971), 17-49.

16. J.C. Willems, Least-squares stationary optimal control and the algebraic Riccati equation, *IEEE Trans. Automatic Control*, AC-16(1971), 621-634.

17. J.E. Potter, Matrix quadratic solutions, *SIAM J. Appl. Math.*, 14(1966), 496-501.

18. F. Germain, Algoritmes de calcul de réalisations stationnaires, Research Report No.66, April 1974, IRIA(LABORIA), Le Chesnay, France.

19. F.R. Ganmacher, Theory of Matrices, Vol.1, Chelsea Publishing Company, New York, 1960.

20. W.M. Wonham, On a matrix Riccati equation of stochastic control, *SIAM J. Control*, 6(1968), 681-697.

21. T. Kailath, Some new algorithms for recursive estimation in constant linear systems, *IEEE Trans. Information Theory*, IT-19(1973), 750-760.

22. A. Lindquist, Optimal filtering of continuous-time stationary processes by means of the backward innovation process, *SIAM J. Control*, 4(1974), 747-754.

23. __________, On Fredholm integral equations, Toeplitz equations and Kalman-Bucy filtering, *Appl. Math. & Optimization*, 1(1975), 355-373.

24. __________, Some reduced order non-Riccati equations for linear least-squares estimation: The stationary single-output case, *Int. J. Control*, 24(1976), 821-842.

25. R.S. Bucy and P.D. Joseph, Filtering for Stochastic Processes with Applications to Guidance, Interscience Publishers, 1968.

26. M.D. Canon, C.D. Collum, Jr., and E. Polak, Theory of Optimal Control and Mathematical Programming, McGraw-Hill Co., New York, 1970.

27. A. Lindquist, A new algorithm for optimal filtering of discrete-time stationary processes, *SIAM J. Control*, 12(1974), 736-746.

28. __________, On integrals in multi-output discrete-time Kalman-Bucy filtering, *Math. Prog. Study*, 5(1976), 145-168.

29. M. Morf, G.S. Sidhu, and T. Kailath, Some new algorithms for recursive estimation in constant, linear discrete-time systems, *IEEE Trans. Automatic Control*, AC-19(1974), 315-323.

30. B.D.O. Anderson, Dual form of a positive real lemma, *Proc. IEEE*, 55(1969), 1749-1750.

31. A.N. Kolmogorov, Interpolation and extrapolation of stationary random sequences, *Bull. Acad. Sci. USSR, Math. Ser.*, Vol.5, 1941.

32. N. Wiener, The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications, New York, Wiley, 1949.

33. R.E. Kalman, A new approach to linear filtering and prediction problems, *Trans. ASME, J. Basic Eng'g.*, Vol.82(1960), 34-45.

34. R.E. Kalman and R.S. Bucy, New results in linear filtering and prediction theory, *Trans. ASME, J. Basic Eng'g.*, Ser.D, Vol.83(1961), 95-107.

35. A.E. Bryson and M. Frazier, Smoothing for linear and nonlinear dynamic systems, TDR 63-119, Aero. Sys. Div., (1963), Wright-Patterson Air Force Base, Ohio, 353-364.

36. H.E. Rauch, F. Tung, and C.T. Striebel, Maximum likelihood estimates of linear dynamic systems, *AIAA J.*, Vol.3, No.8, August 1965, 1445-1450.

37. J.S. Meditch, On optimal fixed point smoothing, *Int. J. Control*, Vol.6, No.2(1976), 189-199.

38. D.Q. Mayne, A solution of the smoothing problem for linear dynamic systems, *Automatica*, Vol.4(1966), 73-92.

39. D.C. Fraser, A new technique for optimal smoothing of data, Sc.D. Dissertation, MIT, Cambridge, MA (1967).

40. R.K. Mehra, On optimal and suboptimal linear smoothing, *Proc. Nat. Electr. Conf.*, XXIV(1968), 119-124.

41. D.C. Fraser and J.E. Potter, The optimum linear smoother as a combination of two optimum linear filters, *IEEE Trans. Automatic Control*, AC-14(1969), 387-390.

42. T. Kailath and P.A. Frost, An innovations approach to least-squares estimation -- Part II: linear smoothing in additive noise, *IEEE Trans. Automatic Control*, AC-13(1968), 655-660.

43. L.E. Zachrisson, On optimal smoothing of continuous-time Kalman processes, *Information Sciences*, 1(1969), 143-172.

44. A. Lindquist, A theorem on duality between estimation and control for linear stochastic systems with time delay, *J. Mathematical Analysis and Applications*, Vol.37, No.2(1972), 516-536.

45. D.G. Lainiotis, Partitioned linear estimation algorithms: discrete case, *IEEE Trans. Automatic Control*, AC-20(1975), 255-257.

46. D.G. Lainiotis and K.S. Govindaraj, A unifying approach to linear estimation via the partitioned algorithms, II: discrete models, *Proc. of the 1975 IEEE Decision and Control Conf.*

47. T. Kailath and L. Ljung, Backwards Markovian models for second order stochastic process, *IEEE Trans. Information Theory*, IT-22(1976), 488-491.

48. B. Friedlander, T. Kailath, and L. Ljung, Scattering theory and linear least squares estimation - II: discrete-time problems, *J. Franklin Inst.*, Vol.301(1976), 71-82.

49. L. Ljung and T. Kailath, A unified approach to smoothing formulas, *Automatica*, Vol.12(1976), 147-157.

50. G.S. Sidhu and U.B. Desai, New smoothing algorithms based on reversed-time lumped models, *IEEE Trans. Automatic Control*, AC-21(1976), 538-541.

51. ______________________, Reversed-time lumped models, smoothing algorithms and Fredholm resolvents, *Proc. of the JHU Conf. on Information Science and Systems*, Johns Hopkins Univ., Baltimore, MD, March-April, 1976.

52. J.S. Meditch, A survey of data smoothing for linear and non-linear dynamic systems, *Automatica*, Vol.9(1973), 151-162.

53. A. Lindquist and G. Picci, On the structure of minimal splitting subspaces in stochastic realization theory, *Proc. 1977 Conf. Decision and Control*, New Orleans, 1977, 42-48.

54. ______________________, A state-space theory for stationary stochastic processes, *Proc. 21st Midwest Symposium on Circuits and Systems*, Ames, Iowa, August, 1978.

55. ______________________, A Hardy space approach to the stochastic realization problem, *Proc. 1978 Conf. Decision and Control*, San Diego, January, 1979.

56. ______________________, Realization theory for multivariate stationary Gaussian process I: State space construction, 4th International Symposium on the Mathematical Theory of Networks and Systems, July, 1979, Delft, Holland.

57. A. Lindquist, G. Picci, and G. Ruckebusch, On minimal splitting subspaces and Markovian representations, *Mathematical Systems Theory*, Vol.12(1979), 271-279.

58. G. Ruckebusch, Représentations Markoviennes de processus Gaussiens stationnaires, *C.R. Acad. Sc. Paris*, Series A, 282(1976), 649-651.

59. G. Ruckebusch, Représentations Markoviennes de processus Gaussien stationnaires et applications statistiques, *Tournées de Statistiques des Processus Stochastiques*, Grenoble, June 1977, Springer-Verlag Lecture Notes in Mathematics 636.

60. ____________, A geometric approach to the stochastic realization problem, Center de Mathématiques Appliques, Ecole Polytechnique, Internal Report No.41, November 1978.

61. J.E. Wall, Jr., Control and estimation for large-scale systems having spatial symmetry, Ph.D. Thesis, MIT, Cambridge, M.A. (1978).

62. M. Pavon, Constant directions and singular stochastic realization: the scalar case, *Proc. 1re Colleque AFCET-SMF*, September 1978, Palaiseau, France.

63. R.E. Kalman, Contributions to the theory of optimal control, *Bol. Soc. Mat. Mexicana*, 5(1960), 102-119.

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

65. ____________________________, On the Mayne-Fraser Formula and stochastic realization theory for nonstationary stochastic systems, *Proc. 1979 Conf. Decision and Control*, Fort Lauderdale, Florida, December 1979.

66. G.W. Stewart, Introduction to Matrix Computations, Academic Press, 1973.

67. M. Pavon, Private communication.

68. L.M. Silverman and H.E. Meadows, Controllability and observability in time-variable linear systems, *SIAM J. Control*, Vol.5, No.1(1967), 64-73.

69. D.M. Wiberg, State Space and Linear Systems, Shaum's Outline Series, McGraw-Hill Book Co., 1971.

70. J.L. Doob, Stochastic Processes, Wiley, 1953.

71. M. Clerget, Systemes linéaires positifs non stationnaires, Research Report No.65, April 1974, IRIA(LABORIA), Le Chesnay, France.

72. A. Lindquist, Private communication.

73. B.D.O. Anderson and P.J. Moylan, Spectral factorization of a finite-dimensional nonstationary matrix covariance, *IEEE Trans. Automatic Control*, AC-19(1974), 680-692.

74. L.M. Silverman, Inversion of multivariable linear systems, *IEEE Trans. Automatic Control*, AC-14(1969), 226-234.

75. V.A. Yakubovich, The solution of certain matrix inequalities in automatic control theory, *Doklady Akademii Nauk. SSSR*, 143(1962), 1304-1307.

76. R.E. Kalman, Lyaponov functions for the problem of Lur'e in automatic control, *Proc. National Academy of Sciences*, 49(1963), 201-205.

77. V.M. Popov, Hyperstability and optimality of automatic systems with several control functions, *Rev. Roumaine Sci. Tech.-Elektrokchn. et Energ.*, 9(1964), 629-690.

78. H. Akaike, Markovian representation of stochastic processes by canonical variables, *SIAM J. Control*, 13(1975), 162-173.

79. ________, Stochastic theory of minimal realization, *IEEE Trans. Automatic Control*, A-12(1974), 667-674.

80. G. Picci, Stochastic realization for Gaussian processes, *Proc. IEEE*, 64(1976), 112-122.

81. J.C. Willems and J.H. van Schuppen, Stochastic systems and the problem of state space realization, (will be submitted for publication).

82. L.E. Zachrisson, Private communication.

83. A. Lindquist, Linear least squares estimation of discrete-time stationary processes by means of backward innovations, *Proc. International Symp. on Control Theory, Numerical Methods and Computer Systems Modeling*, Rocquencourt, France, June 1974, Lecture Notes in Economics and Mathematical Systems, Vol.107 (Springer,Berlin, 1975), 44-63.

VITA

Faris A. Badawi

He received the Teacher's Training Diploma from Al-Hussein Agricultural Institute, Tulkarm, West Bank of Jordan, the B.S. and M.S. degrees in Mathematics from the American University of Beirut, Beirut, Lebanon in 1967, 1971 and 1975, respectively.

From 1971 to 1975, he was with the Training Department at the Arabian American Oil Company (ARAMCO), Dhahran, Saudi Arabia as an Advanced Bilingual Teacher. Since 1975, he has been either a teaching assistant or a research fellow at the Department of Mathematics, University of Kentucky, Lexington, Kentucky.

His research interests lie primarily in the fields of stochastic systems theory, estimation and control.

He has two published papers [64] and [65], together with A. Lindquist and M. Pavon.

F. BADAWI

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 18 19 AFOSR-TR-80-0455	2. GOVT ACCESSION NO. AD-A086465
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) 6 Structures and Algorithms in Stochastic Realization Theory and the Smoothing Problem.	5. TYPE OF REPORT & PERIOD COVERED Technical Report Interim
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Faris A. Badawi	8. CONTRACT OR GRANT NUMBER(s) 15 AFOSR-78-3519
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 16 2304/A1 61102F
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research, Air Force Systems Command - USAF, Bolling Air Force Base, D.C. 20332	12. REPORT DATE 11 Jan[redacted]80
	13. NUMBER OF PAGES 147
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 17 A1	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

12 149

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

9 Interim technical rept.,

18. SUPPLEMENTARY NOTES

10 Faris Abdul-Jabbar/Badawi

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Stochastic realization, Hamiltonian Systems, non-Riccati algorithms, smoothing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report contains two main topics, each of which is connected to the stochastic realization problem. First, ~~we~~ IT considers some structural and algorithmic problems in wide sense stochastic realization theory which also have applicability to many problems outside the realm of stochastic realization theory but are here formulated in that framework. ~~We~~ IT considers some geometric

DD FORM 1 JAN 73 1473

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

410906

questions concerning the solution set of the positive real lemma and provide a Hamiltonian framework for the non-Riccati algorithms of Kailath and Lindquist; these are then applied to the stochastic realization problem. Secondly, IT APPLIES ~~we apply~~ the basic techniques and concepts of the strict sense (proper) stochastic realization theory of Lindquist and Picci and Ruckebusch to the discrete-time smoothing problem. This provides a *natural* interpretation of the Mayne-Fraser two-point formula as well as many other smoothing results, the interpretations of which have hitherto been quite unclear from a probabilistic point of view. Hence we have laid the ground work for a theory of smoothing which has so far been lacking.