SIAM J. CONTROL AND OPTIMIZATION
Vol. 31, No. 4, pp. 942-959, July 1993

006

KALMAN FILTERING WITH RANDOM COEFFICIENTS AND CONTRACTIONS*

PHILIPPE BOUGEROL†

Abstract. The Riccati transformation of linear filtering/control theory is shown to be a contraction on the space of positive symmetric matrices. This is used to describe the asymptotic behavior of the filter for systems with stochastic stationary parameters.

Key words. Kalman filter, stochastic parameters, Riccati equation, stationary process

AMS subject classifications. 93C05, 93E11, 60G35, 34F05

Introduction. In this paper we study the asymptotic properties of the Kalman filter in a random stationary environment, under weak controllability and observability conditions. We show that the covariance matrix of the conditional error converges in law and that the filter is exponentially stable. This is a direct generalization of Kalman's classical results.

Our main tool is the following. Consider the classical Riccati transformation that associates the error covariance matrix at time $n+1$ to the error covariance matrix at time n. We show that this transformation is a contraction with respect to the Riemannian metric on the set of positive symmetric matrices. This important fact does not seem to have been noticed before. For instance, it leads to a straightforward proof of Kalman's results on the asymptotic behavior of the filter. It can also be useful in other parts of filtering or control theory.

For the convenience of the reader who is not interested in random environment, we present our results in three parts. Section 1 is devoted to the study of the above-mentioned contraction property in the classical set-up. Filtering with random parameters is considered in § 2. Section 3 is an appendix that proves the general results on iteration of random Lipschitz contractions (needed in § 2).

Let us describe the main results of this paper. We consider the linear system

$$(1)\qquad \begin{aligned} X_n &= A_n X_{n-1} + F_n \varepsilon_n, \qquad n \geqq 1, \\ Y_n &= C_n X_n + \eta_n, \end{aligned}$$

where $X_n \in \mathbb{R}^d$, $\varepsilon_n \in \mathbb{R}^p$, η_n, $Y_n \in \mathbb{R}^q$. In § 1, the parameters A_n, F_n, and C_n are *deterministic* matrices with size $d \times d$, $d \times p$ and $q \times d$, respectively. The random vectors $\{(\varepsilon_n, \eta_n), n \in \mathbb{N}\}$ are independent; they have the same Gaussian law with mean 0 and covariance matrix equal to the identity. We assume that X_0 has a Gaussian law with mean $\hat{X}_0$ and covariance matrix P_0. We always suppose that the matrices A_n are nonsingular (our approach does not apply in the singular case).

For any $n \geqq 1$, let $\mathscr{F}_n$ be the sigma-algebra generated by the random vectors $Y_1, Y_2, \ldots, Y_n$, and

$$(2)\qquad \hat{X}_n := \mathbb{E}(X_n/\mathscr{F}_n),$$

$$(3)\qquad P_n := \mathbb{E}((X_n - \hat{X}_n)(X_n - \hat{X}_n)^*/\mathscr{F}_n).$$

* Received by the editors January 2, 1991; accepted for publication (in revised form) November 1, 1991.
† Laboratoire de Probabilités, Université Paris VI, 4, Place Jussieu, 75252 Paris Cedex 05, France.

When $\mathcal{F}_n$ is known, $\hat{X}_n$ is the best estimate of X_n, and P_n is the conditional error covariance matrix. Let $\mathcal{P}$ (respectively, $\mathcal{P}_0$) denote the set of $d \times d$ nonnegative (respectively, positive) symmetric matrices. For any n in $\mathbb{N}$ and $P \in \mathcal{P}$, we set

$$\Phi_n(P) = (A_n P A_n^* + S_n)(I + R_n S_n + R_n A_n P A_n^*)^{-1}, \tag{4}$$

where $R_n = C_n^* C_n$ and $S_n = F_n F_n^*$; then $\Phi_n(P) \in \mathcal{P}$, and Φ_n maps $\mathcal{P}_0$ in $\mathcal{P}_0$. The classical Kalman's recursive equations can be written as

$$P_n = \Phi_n(P_{n-1}) \tag{5}$$

$$\hat{X}_n = (A_n - P_n R_n A_n)\hat{X}_{n-1} + P_n C_n^* Y_n \tag{6}$$

(see, e.g., Balakrishnan [3, Relations 4-1-19, -27, -31, -34]). The main result of § 1 is that the maps Φ_n are contractions on $\mathcal{P}_0$, if we equip this set with the Riemannian metric δ, which is invariant under conjugacy (see Theorem 1.7). Moreover, the fact that these contractions are strict and/or uniform depends on the observability and the controllability properties of the linear system (1). These results are proved in their natural context, namely, by looking at the symplectic matrices that act on the set of symmetric matrices by preserving $\mathcal{P}_0$. In that setting, they can be seen as generalizations of the Perron–Frobenius theorem. We also could have considered the recursion associated with $\mathbb{E}(X_n/\mathcal{F}_{n-1})$ for which analogous results hold true (see [8]).

Section 2 is devoted to filtering in a random stationary environment. We consider again the linear equation (1), but we now suppose that the parameters A_n, F_n, and C_n are *stochastic* and that $\{(A_n, F_n, C_n),\ n \geqq 1\}$ is a stationary ergodic process. Under suitable hypotheses (see Hypothesis in § 2), system (1) is conditionally Gaussian, and $\hat{X}_n$ and P_n are also given by the recursive equations (5) and (6) of Kalman. These hypotheses hold, for instance, when the parameters are independent of the noises. We first describe in § 2.1 some actual situations that can be described by such systems. Then, in § 2.2 we introduce weak controllability and observability assumptions (in contrast with the uniform conditions of Kalman). These conditions can hold for systems that are usually neither controllable nor observable. Fault-tolerant systems usually have this property. Under these assumptions, we describe the asymptotic behavior of the conditional error covariance matrices P_n. Our main result is Theorem 2.4. We show that there exists a stationary $\mathcal{P}_0$-valued process $\{\bar{P}_n,\ n \in \mathbb{N}\}$ with the following universal property: "Almost surely, for any solution P_n of (5), $\|P_n - \bar{P}_n\|$ converges to 0 as $n \to +\infty$." In particular, P_n converges in law. In § 2.3 we prove that the filter (6) is exponentially stable. These results are deduced from properties of processes that are defined by iterations under stationary Lipschitz maps (Relation (5) and Theorem 1.7 show that the process P_n is of this type). These properties are interesting for their own sake; they are established in § 3.

This paper is self-contained and in some sense elementary. Some of its ideas are already in the literature. The trick of studying the filtering of Riccati's equation through the action of symplectic matrices is, of course, well known (see, e.g., Hermann [15], Shayman [24], and their references). Our semigroup $\mathcal{H}$ has been introduced already by Wojtkowski [31], [32] in a different context. The contraction property of the Riccati transformation is a generalization of the contraction property of matrices with nonnegative elements for the Hilbert metric, due to G. Birkhoff [5]. It's also related with the contraction properties of product of random matrices on boundaries. In Bougerol [7] we recover some of the results obtained here by making use of the Osseledets theorem and the Lyapunov exponents of the associated Hamiltonian matrices.

All our results can be generalized easily to the continuous-time case, either by a direct study or by a reduction to discrete-time.

1. Contraction properties of Riccati equation.

1.1. The semigroup of Hamiltonian matrices. We consider the classical linear system (1) with deterministic parameters (A_n, F_n, C_n), $n \geqq 1$,

$$X_n = A_n X_{n-1} + F_n \varepsilon_n,$$

$$Y_n = C_n X_n + \eta_n,$$

defined in the Introduction. We always suppose that the matrices A_n are invertible. We associate to this system the so-called Hamiltonian matrices M_n of order $2d$ written in block form as

$$M_n = \begin{pmatrix} A_n & S_n A_n^{*-1} \\ R_n A_n & (I + R_n S_n) A_n^{*-1} \end{pmatrix}, \tag{7}$$

where $R_n = C_n^* C_n$ and $S_n = F_n F_n^*$. These matrices are in the symplectic group $\mathrm{Sp}(d, \mathbb{R})$. This group is defined as the set of all the matrices M of order $2d$ such that $M^*JM = J$, where $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, ($I$ is the identity matrix of order d and M^* is the transpose of M). This relation can be written as $M^{-1} = JM^*J$, thus we see that M^* is also in $\mathrm{Sp}(d, \mathbb{R})$. If we write

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where the entries are $d \times d$ matrices, then BA^* and A^*C are symmetric and $A^*D - C^*B = I$.

Let $\mathscr{P}$ (respectively, $\mathscr{P}_0$) be the set of $d \times d$ nonnegative (respectively, positive) symmetric matrices (we recall that a matrix M is nonnegative, respectively, positive, when for all $x \neq 0$, $x^*Mx \geqq 0$, respectively, >0). The set of all Hamiltonian matrices is

$$\mathscr{H} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}(d, \mathbb{R});\ A \text{ is invertible, } BA^* \in \mathscr{P},\ A^*C \in \mathscr{P} \right\}.$$

Indeed, every matrix M_n defined previously is in $\mathscr{H}$, and every matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ in $\mathscr{H}$ is the Hamiltonian matrix of the linear system (1) with the constant parameters $A_n = A$, $F_n = \sqrt{BA^*}$, $C_n = \sqrt{CA^{-1}}$, and dimensions $p = q = d$. We define three subsets $\mathscr{H}_1$, $\mathscr{H}_2$, and $\mathscr{H}_0$ of $\mathscr{H}$ by

$$\mathscr{H}_1 = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathscr{H};\ BA^* \in \mathscr{P},\ A^*C \in \mathscr{P}_0 \right\},$$

$$\mathscr{H}_2 = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathscr{H};\ BA^* \in \mathscr{P}_0,\ A^*C \in \mathscr{P} \right\},$$

$$\mathscr{H}_0 = \mathscr{H}_1 \cap \mathscr{H}_2.$$

We remark that $\mathscr{H}_2$ is the dual of $\mathscr{H}_1$ (in the sense that $M \in \mathscr{H}_1$ if and only if $M^* \in \mathscr{H}_2$). The following semigroup property of $\mathscr{H}$ already appeared in a different context, implicitly in Ol'shanskii [22] and explicitly in Wojtkowski [31] and [32].

PROPOSITION 1.1. *The product of matrices in $\mathscr{H}$ is in $\mathscr{H}$. The product of a matrix in $\mathscr{H}$ with a matrix in $\mathscr{H}_1$ (respectively, $\mathscr{H}_2$, $\mathscr{H}_0$) is in $\mathscr{H}_1$ (respectively, $\mathscr{H}_2$, $\mathscr{H}_0$). In other words, $\mathscr{H}$ is a semigroup of matrices, and $\mathscr{H}_1$, $\mathscr{H}_2$, and $\mathscr{H}_0$ are two-sided ideals of $\mathscr{H}$.*

We will need the following well-known lemma.

LEMMA 1.2. *When P and Q are in $\mathcal{P}$, then $I+PQ$ is invertible.*

Proof. If P is positive, we can find an invertible matrix M such that $P^{-1}=M^*M$ and $Q=M^*DM$, where D is diagonal with nonnegative entries. Then $I+PQ=M^{-1}(I+D)M$. Thus, the eigenvalues of $I+PQ$ are greater than one. By density, this remains true even if P is only nonnegative. □

Proof of Proposition 1.1. Let M_1 and M_2 be two matrices in $\mathcal{H}$, and $M_3=M_2M_1$. For $i=1$, 2, or 3 we write $M_i=\left(\begin{smallmatrix} A_i & B_i \\ C_i & D_i \end{smallmatrix}\right)$, and we set $Q_1=C_1A_1^{-1}$ and $P_2=A_2^{-1}B_2$. Since $M_1\in\mathcal{H}$, the matrix $A_1^*C_1$ is in $\mathcal{P}$ and the fact that $Q_1=A_1^{*-1}(A_1^*C_1)A_1^{-1}$ yields that Q_1 is in $\mathcal{P}$. Similarly, P_2 is also in $\mathcal{P}$. We have

$$A_3=A_2A_1+B_2C_1=A_2(I+A_2^{-1}B_2C_1A_1^{-1})A_1=A_2(I+P_2Q_1)A_1.$$

Hence, it follows from the previous lemma that A_3 is invertible. We will make use of the relation

$$A_2^*D_2=C_2^*B_2+I=C_2^*A_2P_2+I=A_2^*C_2P_2+I$$

in the next computation. Since $C_3=C_2A_1+D_2C_1$, one has

$$\begin{aligned} A_3^*C_3&=A_1^*(I+Q_1P_2)A_2^*(C_2A_1+D_2C_1)\\ &=A_1^*(I+Q_1P_2)(A_2^*C_2A_1+A_2^*C_2P_2C_1+C_1)\\ &=A_1^*(I+Q_1P_2)A_2^*C_2(A_1+P_2C_1)+A_1^*C_1+C_1^*P_2C_1\\ &=A_1^*(I+Q_1P_2)A_2^*C_2(I+P_2Q_1)A_1+A_1^*C_1+C_1^*P_2C_1. \end{aligned} \tag{8}$$

This shows that $A_3^*C_3$ is a nonnegative symmetric matrix. Similarly (or using transpositions) we see that $A_3B_3^*$ is also nonnegative. This proves that M_3 is in $\mathcal{H}$. Thus, $\mathcal{H}$ is a semigroup. If, moreover, $A_2^*C_2$ or $A_1^*C_1$ is invertible, then (8) shows that $A_3^*C_3$ is positive definite. Hence, $\mathcal{H}_2$ is an ideal of $\mathcal{H}$. Similarly, $\mathcal{H}_1$, and thus $\mathcal{H}_0$, is also an ideal of $\mathcal{H}$.

The following result will be useful later to link the fact that a linear system is observable or controllable with the fact that the associated Hamiltonian matrices are in $\mathcal{H}_1$ or in $\mathcal{H}_2$.

PROPOSITION 1.3. *Let $M_n=\left(\begin{smallmatrix} A_n & B_n \\ C_n & D_n \end{smallmatrix}\right)$, $n\in\mathbb{N}$, be matrices in $\mathcal{H}$. Then $M_nM_{n-1}\ldots M_1$ is in $\mathcal{H}_1$ if and only if*

$$\operatorname{Det}(A_1^*C_1+A_1^*A_2^*C_2A_1+\cdots+A_1^*\ldots A_{n-1}^*A_n^*C_nA_{n-1}\ldots A_1)\neq 0,$$

and $M_nM_{n-1}\ldots M_1$ is in $\mathcal{H}_2$ if and only if

$$\operatorname{Det}(B_nA_n^*+A_nB_{n-1}A_{n-1}^*A_n^*+\cdots+A_n\ldots A_2B_1A_1^*A_2^*\ldots A_n^*)\neq 0.$$

Proof. For any $M=\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ in $\mathcal{H}$, let $\xi(M)=A^*C$ and $\alpha(M)=A$. We first show that, if M_1 and M_2 are in $\mathcal{H}$, then for any x in $\mathbb{R}^d$,

$$\xi(M_2M_1)x=0 \tag{9}$$

if and only if

$$\xi(M_1)x=0 \quad\text{and}\quad \xi(M_2)A_1x=0. \tag{10}$$

We use the following two properties:

(i) When P, $Q\in\mathcal{P}$ and $(P+Q)x=0$, then $Px=Qx=0$;

(ii) For any matrix M, if $P\in\mathcal{P}$ and $M^*PM\,x=0$, then $PMx=0$.

We know by (8) that if $Q_1=C_1A_1^{-1}$ and $P_2=A_2^{-1}B_2$, then

$$\xi(M_2M_1)=A_1^*(I+Q_1P_2)A_2^*C_2(I+P_2Q_1)A_1+A_1^*C_1+C_1^*P_2C_1. \tag{11}$$

The right-hand side is a sum of nonnegative symmetric matrices. Therefore, we see that if (9) holds, then $A_1^* C_1 x = 0$ and $A_2^* C_2(I + P_2 Q_1) A_1 x = 0$. This implies that $C_1 x = 0$ (since A_1 is invertible) and $A_2^* C_2 A_1 x = 0$ (since $Q_1 A_1 x = C_1 x = 0$); thus, (10) holds. The converse also follows easily from (11). Using the equivalence between (9) and (10) we see that the following statements are equivalent:

$$\begin{aligned} &\xi(M_n \dots M_2 M_1)x = 0, \\ &\Leftrightarrow \xi(M_1)x = 0 \quad \text{and} \quad \xi(M_n \dots M_2)\alpha(M_1)x = 0, \\ &\Leftrightarrow \dots, \\ &\Leftrightarrow \xi(M_1)x = 0 \quad \text{and} \quad \xi(M_2)\alpha(M_1)x = 0, \dots, \\ &\text{and} \quad \xi(M_n)\alpha(M_{n-1}) \dots \alpha(M_1)x = 0. \end{aligned}$$

Since the matrices $\xi(M_n)$ are nonnegative and $\alpha(M_n)$ invertible, this is also equivalent to

$$\{\xi(M_1) + \alpha(M_1)^* \xi(M_2)\alpha(M_1) + \cdots + \alpha(M_1)^* \dots \alpha(M_{n-1})^* \xi(M_n) \cdot \alpha(M_{n-1}) \dots \alpha(M_1)\} x = 0.$$

This proves the first claim. The second claim is obtained by duality.

1.2. Contraction property. For any matrix $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ in $\mathscr{H}$ we define a map $\Phi_M : \mathscr{P}_0 \to \mathscr{P}_0$ by

$$\Phi_M(T) = (AT + B)(CT + D)^{-1}, \quad T \in \mathscr{P}_0. \tag{12}$$

The fact that the right-hand side is a well-defined element of $\mathscr{P}_0$ will be shown in Proposition 1.5. A straightforward computation shows that the map $M \to \Phi_M$ defines an action of the semigroup $\mathscr{H}$ on $\mathscr{P}_0$, in the sense that, for any M, N in $\mathscr{H}$,

$$\Phi_{MN} = \Phi_M \circ \Phi_N \tag{13}$$

(in fact this action is induced by the linear action of symplectic matrices on d-dimensional linear subspaces of $\mathbb{R}^{2d}$). We remark that the relation (4), which defines the maps Φ_n, can be written as

$$\Phi_n(P) = (A_n P + S_n A_n^{*-1})(R_n A_n P + (I + R_n S_n) A_n^{*-1})^{-1}.$$

This shows that $\Phi_n = \Phi_{M_n}$. Therefore, by (5), the error covariance matrix P_n satisfies

$$P_n = \Phi_{M_n}(P_{n-1}). \tag{14}$$

This equation is called the *discrete Riccati equation.* The classical continuous matrix Riccati equation on $\mathscr{P}_0$ is similar. Indeed, under some mild regularity and boundedness assumptions, if P_t, $t \in \mathbb{R}^+$, is the solution of the matrix-valued differential equation

$$\dot{P}_t = A_t P_t + P_t A_t^* - P_t R_t P_t + S_t, \quad P_0 \in \mathscr{P}_0,$$

where R_t, S_t, $t \geqq 0$, are in $\mathscr{P}$, then there exists a family N_t, $t \geqq 0$, of matrices in $\mathscr{H}$ such that $P_t = \Phi_{N_t}(P_0)$ (cf., Hermann [15]).

DEFINITION 1.4. The Riemannian distance δ on $\mathscr{P}_0$ is defined by: for any $P, Q \in \mathscr{P}_0$,

$$\delta(P, Q) = \left\{ \sum_{i=1}^{d} \operatorname{Log}^2 \lambda_i \right\}^{1/2},$$

where $\lambda_1, \dots, \lambda_d$ are the eigenvalues of the matrix PQ^{-1}.

It is shown in Maass [19, Thm. p. 27] (see also Terras [27]), that δ is the usual Riemannian distance on $\mathscr{P}_0$ when this set is considered as the Riemannian symmetric space $\mathrm{Gl}(d, \mathbb{R})/O(d)$ (this metric is associated to the arc length $ds^2 = \mathrm{tr}\{(P^{-1}\, dP)^2\}$, which is invariant under conjugacy and coincides with the Euclidean arc length on the logarithms of the diagonal matrices in $\mathcal{P}_0$). In particular, $(\mathcal{P}_0, \delta)$ is complete and δ induces the usual topology. The main property of this distance is its invariance under conjugacy and under inversion. For any invertible matrix A and for all P, Q in $\mathcal{P}_0$,

$$\delta(APA^*, AQA^*) = \delta(P, Q) = \delta(P^{-1}, Q^{-1}).$$

We next prove that the transformations Φ_M are contractions of $(\mathcal{P}_0, \delta)$ when $M \in \mathcal{H}$.

PROPOSITION 1.5. *Let* $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *be a matrix in* $\mathcal{H}$. *Then*

(i) *For any* T *in* $\mathcal{P}$ (*respectively*, $\mathcal{P}_0$), $CT + D$ *is invertible and* $(AT+B)(CT+D)^{-1}$ *is in* $\mathcal{P}$ (*respectively*, $\mathcal{P}_0$).

(ii) *If* $M \in \mathcal{H}_2$, *then for any* T *in* $\mathcal{P}$, $(AT+B)(CT+D)^{-1}$ *is in* $\mathcal{P}_0$.

Proof. Let $T \in \mathcal{P}$. The matrices $P = A^{-1}B$, $Q = CA^{-1}$, and $S = A\,(T+P)A^*$ are in $\mathcal{P}$. Since

$$CT + D = QAT + QAP + A^{*-1} = (QS + I)A^{*-1},$$

it follows from Lemma 1.2 that $CT + D$ is invertible. Now, the relation

$$(AT+B)(CT+D)^{-1} = (AT + AP)A^*(QS+I)^{-1} = S(QS+I)^{-1}$$

easily implies the proposition (we note that this is equal to $(S^{-1}+Q)^{-1}$ when S is invertible). □

We always use the Euclidean norm on $\mathbb{R}^d$ and the associated operator norm on the set of matrices: if M is a matrix of order d, we let $\|M\| = \mathrm{Sup}\{\|Mx\|;\, x \in \mathbb{R}^d,\, \|x\| = 1\}$.

PROPOSITION 1.6. *Let* T, S *be matrices in* $\mathcal{P}_0$ *and* $\alpha = \mathrm{Max}(\|T\|, \|S\|)$. *Then for all* $P \in \mathcal{P}$,

$$\delta(T+P, S+P) \leqq \frac{\alpha}{\alpha+\beta}\, \delta(T, S)$$

where $\beta = \mathrm{Inf}\{\langle Px, x\rangle;\, \|x\| = 1\}$.

Proof. The mean value theorem yields that, when $0 < a$, $b \leqq m$ and $r > 0$, then

$$\mathrm{Log}\,\frac{a+r}{b+r} \leqq \frac{m}{m+r}\, \mathrm{Log}^+ \frac{a}{b} \tag{15}$$

(where $\mathrm{Log}^+ x = \mathrm{Max}(\mathrm{Log}\, x, 0)$). It is known and not difficult to prove (see Gantmacher [14, Ch. 10, § 7]) that the eigenvalues of TS^{-1} are real, positive, and that they have the following Min-Max representation. Let

$$\lambda_1(T, S) \leqq \lambda_2(T, S) \leqq \cdots \leqq \lambda_d(T, S)$$

be the eigenvalues of TS^{-1} written in ascending order. Then

$$\lambda_k(T, S) = \mathrm{Min}\left\{\mathrm{Max}\left\{\frac{\langle Tx, x\rangle}{\langle Sx, x\rangle};\, x \in V\right\};\, V \in \Gamma(k)\right\},$$

where $\Gamma(k)$ is the set of k-dimensional linear subspaces of $\mathbb{R}^d$. We prove that

$$|\mathrm{Log}\, \lambda_k(T+P, S+P)| \leqq \frac{\alpha}{\alpha+\beta}\, |\mathrm{Log}\, \lambda_k(T, S)|, \tag{16}$$

for any $1 \leqq k \leqq d$. We first suppose that $\lambda_k(T+P, S+P) > 1$. Relation (15) entails that

$$\mathrm{Log}\,\frac{\langle Tx, x\rangle + \langle Px, x\rangle}{\langle Sx, x\rangle + \langle Px, x\rangle} \leqq \frac{\alpha}{\alpha+\beta}\, \mathrm{Log}^+ \frac{\langle Tx, x\rangle}{\langle Sx, x\rangle}.$$

Thus,

$$\begin{aligned}|\mathrm{Log}\,\lambda_k(T+P, S+P)| &= \mathrm{Log}\,\lambda_k(T+P, S+P)\\ &= \mathrm{Log}\,\mathrm{Min}\left\{\mathrm{Max}\left\{\frac{\langle (T+P)x, x\rangle}{\langle (S+P)x, x\rangle};\ x\in V\right\};\ V\in\Gamma(k)\right\}\\ &= \mathrm{Min}\left\{\mathrm{Max}\left\{\mathrm{Log}\,\frac{\langle Tx, x\rangle+\langle Px, x\rangle}{\langle Sx, x\rangle+\langle Px, x\rangle};\ x\in V\right\};\ V\in\Gamma(k)\right\}\\ &\leqq \frac{\alpha}{\alpha+\beta}\,\mathrm{Min}\left\{\mathrm{Max}\left\{\mathrm{Log}^+\frac{\langle Tx, x\rangle}{\langle Sx, x\rangle};\ x\in V\right\};\ V\in\Gamma(k)\right\}\\ &\leqq \frac{\alpha}{\alpha+\beta}\,\mathrm{Log}^+\,\lambda_k(T, S).\end{aligned}$$

Since the left-hand term is positive, this first gives that $\mathrm{Log}^+\,\lambda_k(T, S) = |\mathrm{Log}\,\lambda_k(T, S)|$, and then that (16) holds. When $\lambda_k(T+P, S+P) = 1$, (16) is obvious. When $\lambda_k(T+P, S+P) < 1$, we use the relation $\lambda_k(T+P, S+P) = 1/\lambda_{d-k+1}(S+P, T+P)$, and we apply (16) to $\lambda_{d-k+1}(S+P, T+P)$. Finally, (16) implies immediately the proposition. □

THEOREM 1.7. *The following properties hold*:

(i) *For any M in $\mathscr{H}$, and T, S in $\mathscr{P}_0$,*

$$\delta(\Phi_M(T), \Phi_M(S)) \leqq \delta(T, S).$$

(ii) *For any M in $\mathscr{H}_1$ or in $\mathscr{H}_2$, and T, S in $\mathscr{P}_0$,*

$$\delta(\Phi_M(T), \Phi_M(S)) < \delta(T, S).$$

(iii) *For any M in $\mathscr{H}_0$, there exists $\rho(M)$, $0 < \rho(M) < 1$, such that, for all T, S in $\mathscr{P}_0$,*

$$\delta(\Phi_M(T), \Phi_M(S)) \leqq \rho(M)\delta(T, S).$$

Proof. Let $M = \begin{pmatrix} A & B\\ C & D\end{pmatrix}$ be a matrix in $\mathscr{H}$. The matrices $P = A^{-1}B$ and $Q = CA^{-1}$ are in $\mathscr{P}$ and

$$M = \begin{pmatrix} I & 0\\ Q & I\end{pmatrix}\begin{pmatrix} A & 0\\ 0 & A^{*-1}\end{pmatrix}\begin{pmatrix} I & P\\ 0 & I\end{pmatrix}.$$

We consider the transformations $\tau_P(T) = T+P$, $\tau_Q(T) = T+Q$, $\gamma_A(T) = ATA^*$ and $\sigma(T) = T^{-1}$ defined on $\mathscr{P}_0$. By making use of (13) we obtain

$$\Phi_M(T) = (\sigma\circ\tau_Q\circ\sigma\circ\gamma_A\circ\tau_P)(T). \tag{17}$$

We have already noticed that γ_A and σ are isometries of the metric space $(\mathscr{P}_0, \delta)$. It follows from Proposition 1.6 that τ_P and τ_Q are contractions. This and (17) prove (i). If M is in $\mathscr{H}_1$, then Q is invertible. Hence, τ_Q is a strict contraction by Proposition 1.6. Similarly, when M is in $\mathscr{H}_2$, P is invertible and τ_P is a strict contraction. Thus, (ii) follows from (17). Let us prove (iii). We consider an M in $\mathscr{H}_0$. Then both P and Q are invertible. Moreover, for any T in $\mathscr{P}_0$, $\tau_P(T) \geqq P$ (in the sense that $\tau_P(T) - P \in \mathscr{P}$), which implies that $(\gamma_A\circ\tau_P)(T) \geqq APA^*$, and thus,

$$(\sigma\circ\gamma_A\circ\tau_P)(T) \leqq (APA^*)^{-1}.$$

Let $\zeta = \|(APA^*)^{-1}\|$ and $\varepsilon = \mathrm{Inf}\,\{\langle Qx, x\rangle;\ \|x\| = 1\}$. It follows from Proposition 1.6 that for all T_1 and T_2 in $\mathcal{P}_0$,

$$\begin{aligned}\delta(\tau_Q[(\sigma\circ\gamma_A\circ\tau_P)(T_1)], \tau_Q[(\sigma\circ\gamma_A\circ\tau_P)(T_2)])\\ \leqq \frac{\zeta}{\zeta+\varepsilon}\,\delta((\sigma\circ\gamma_A\circ\tau_P)(T_1), (\sigma\circ\gamma_A\circ\tau_P)(T_2))\\ \leqq \frac{\zeta}{\zeta+\varepsilon}\,\delta(T_1, T_2).\end{aligned}$$

Since σ is an isometry, this relation and (16) yield that (iii) holds with $\rho(M)=\zeta/(\zeta+\varepsilon)$. □

As an application, let us outline a short proof of a classical result of Kalman. We suppose that the linear system with constant coefficients

$$X_n = AX_{n-1} + F\varepsilon_n, \qquad Y_n = CX_n + \eta_n,$$

is controllable and observable. Let M be the Hamiltonian matrix (here independent of n) defined by (7). It follows from Proposition 1.3 that there is an integer $p>0$ such that M^p is in $\mathcal{H}_0$. Therefore, by Theorem 1.7, Φ_M has a power that is a uniform contraction. This implies by the fixed point theorem that there exists a matrix P in $\mathcal{P}_0$ such that all the solutions P_n of (5) converge to P when $n\to+\infty$, as soon as P_0 is in $\mathcal{P}_0$.

2. Filtering with random parameters. We now study the asymptotic behavior of linear filtering in a random environment. We consider the case where the parameters A_n, F_n, C_n of the linear equation (1) are stochastic. More precisely, we suppose that the following hypothesis holds.

Hypothesis H. For all $n\geqq 1$, the quantities A_n, F_n, C_n, ε_n, η_n are random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\{(A_n, F_n, C_n), n\geqq 1\}$ is a strictly stationary ergodic process. There is a σ-algebra $\mathcal{F}_0$ contained in $\mathcal{F}$ such that, if $\mathcal{F}_n=\sigma(\mathcal{F}_0, Y_1, \ldots, Y_n)$ is the σ-algebra generated by $\mathcal{F}_0$ and by $Y_1, Y_2, \ldots, Y_n$, then for all $n\geqq 1$,

(i) A_n, F_n, and C_n are $\mathcal{F}_{n-1}$ measurable.

(ii) The random vector (ε_n, η_n) has a Gaussian law with mean zero and covariance matrix equal to the identity matrix. It is independent of X_{n-1} and $\mathcal{F}_{n-1}$.

(iii) Conditionally on $\mathcal{F}_0$, the random vector X_0 has a Gaussian law with mean $\hat{X}_0$ and covariance matrix P_0.

This set-up is called *conditionally Gaussian.* The conditional expectations $\hat{X}_n=\mathbb{E}(X_n/\mathcal{F}_n)$ and the conditional error covariance matrices $P_n=\mathbb{E}((X_n-\hat{X}_n)(X_n-\hat{X}_n)^*/\mathcal{F}_n)$ are given by the recursions (5) and (6) (see, e.g., Whittle [29, p. 260]). Work on such systems with stochastic parameters goes back to Kalman [17], [18]. A recent reference is De Koning [12] (see also Nahi [21]). An important example is the following: suppose that $\{(\varepsilon_n, \eta_n), n\geqq 1\}$ is a sequence of independent normalized Gaussian random vectors, independent of a stationary ergodic process (A_n, F_n, C_n), $n\geqq 1$. Then, if $\mathcal{F}_0=\sigma\{(A_n, F_n, C_n), n\geqq 1\}$, the hypothesis (H) holds.

In § 2.1, we present some examples of real situations that can be modeled by these equations. In § 2.2, we describe the asymptotic behavior of P_n as $n\to+\infty$. The exponential stability of the filter is proved in § 2.3. We will always suppose that the matrices A_n are nonsingular. Without loss of generality, we can and will suppose that the stationary process (A_n, F_n, C_n) is defined on $(\Omega, \mathcal{F}, \mathbb{P})$ for all $n\in\mathbb{Z}$.

2.1. Examples.

2.1.1. Filter with periodic parameters. We suppose that there exist functions A, B, C on $\Omega_1=\mathbb{Z}/p\mathbb{Z}$ such that, for all $\omega\in\Omega_1$,

$$(A_n(\omega), B_n(\omega), C_n(\omega)) = (A(\omega+n), B(\omega+n), C(\omega+n)),$$

where $\omega + n$ is the sum modulo p. Let $\mathcal{F}_1$ be the set of all the subsets of Ω_1, and let $\mathbb{P}_1$ be the uniform measure on Ω_1. We also consider a sequence of noises (ε_n, η_n), $n \geqq 1$, defined on some $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$. Then these coefficients define a linear system on the probability space $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$, for which (H) holds. These systems with periodic parameters have been studied recently, for instance, by De Souza and Goodwin [13], and Bittanti, Colaneri, and Di Nicolao [6].

2.1.2. Random sampling. In several situations, a linear system can be observed only at random times $T_0 < T_1 < \cdots$. This so-called stochastic sampling phenomenon can occur because of technical imperfections in the instrumentation. It may also be applied intentionally, for instance, when a digital computer is time shared in a stochastic manner as suggested by Kalman [17]. In Snyder and Fishman [25], the tracking of fireflies, which can be observed only by their flashes, is studied (we can easily imagine some more realistic examples); see also Chang [11]. These systems are used in modeling of ARMA processes with missing data.

The basic model is the usual time-invariant system

$$X_n = AX_{n-1} + F\varepsilon_n, \qquad Y_n = CX_n + \eta_n. \tag{18}$$

We suppose that the state is observed only at random times T_n, $n \geqq 0$, independent of this system, and that $\{T_{n+1} - T_n,\ n \geqq 0\}$ is a stationary ergodic process with values in $\mathbb{N}^*$. If we let $Z_n = X_{T_n}$ and $W_n = Y_{T_n}$ then

$$Z_n = A^{T_n - T_{n-1}} Z_{n-1} + \sum_{k=0}^{T_n - T_{n-1} - 1} A^k F \varepsilon_{T_n - k}, \qquad W_n = CZ_n + \eta_{T_n}.$$

For each $n \geqq 1$, let $A_n = A^{T_n - T_{n-1}}$ and let F_n be a symmetric matrix such that

$$F_n^2 = \sum_{k=0}^{T_n - T_{n-1} - 1} A^k F F^* A^{*k}.$$

Using if necessary a generalized inverse of F_n, it is easy to see that there exists a sequence of independent Gaussian random variables $\alpha_n \in \mathbb{R}^p$, $\beta_n \in \mathbb{R}^q$, with mean 0 and covariance matrix equal to the identity, independent of the sequence $\{(A_n, F_n),\ n \geqq 1\}$, such that

$$\sum_{k=0}^{T_n - T_{n-1} - 1} A^k F \varepsilon_{T_n - k} = F_n \alpha_n, \qquad \eta_{T_n} = \beta_n.$$

We obtain that

$$Z_n = A_n Z_{n-1} + F_n \alpha_n, \qquad W_n = CZ_n + \beta_n. \tag{19}$$

This is a system with stochastic parameters for which (H) holds. In this setting, the asymptotic properties of the filter have been studied by Viano [28] under the additional assumption that the matrix A is stable and that C^*C is invertible. We are able to treat the case where the system (18) is only controllable and observable. We remark that no uniform controllability property of (19) can be expected when the T_n's are not bounded. When the random variables $T_{n+1} - T_n$ are independent and identically distributed, the error covariance P_n, $n \in \mathbb{N}$, is a Markov chain on $\mathcal{P}_0$. But this process is singular (it does not satisfy the Harris irreducibility condition). Even in that case, it does not seem easy to study its asymptotic behavior without recourse to contractions properties.

2.1.3. Fault-tolerant filtering. Consider a failure-prone linear system. It can be, for instance, a manufacturing plant or a space-station under the bombardment of meteorites. We can assume that the plant state has two equations: at time n, either $X_n = MX_{n-1} + \varepsilon_n$ if the system is operational, or $X_n = NX_{n-1} + \varepsilon_n$ if the system is in a state of failure and undergoing repair. The failure/repair process may be modeled by a stationary sequence A_n, $n \geqq 1$, of random matrices such that $A_n \in \{M, N\}$. Such systems are considered, for instance, in Akella and Kumar [1] and in Mariton [20] (see also Willems and Willems [30]). If $Y_n = CX_n + \eta_n$, the associated filtering system will satisfy (H).

We may also consider a filtering system with a failure-prone observation process. This can be due to the instrumentation or to the fact that at some unexpected times, the state cannot be determined. For instance, we can think of the tracking of a plane, which is sometimes hidden by clouds. A model for this situation can be

$$X_n = AX_{n-1} + F\varepsilon_n, \qquad Y_n = C_nX_n + \eta_n,$$

where C_n is equal to some matrix C when the observation process is operational and some other matrix D, otherwise. Notice that it is natural to assume that under failure, i.e., when $C_n = D$, the system is not observable.

It follows from the results of the next section that the filter has very good asymptotic properties. This shows that in some sense, Kalman's filtering is fault tolerant. Of course, users are already aware of this fact.

2.1.4. Estimation of AR processes with AR parameters. Suppose that we observe an univariate autoregressive (AR) process $Z_n = \rho_n Z_{n-1} + \eta_n$, where the parameters ρ_n satisfy $\rho_n = a\rho_{n-1} + \varepsilon_n$. Here, $\{(\varepsilon_n, \eta_n), n \geqq 1\}$ is a sequence of independent normalized Gaussian random variables. These models occur, for instance, in stochastic adaptative control (see, e.g., Caines and Meyn [10]). If we want to estimate the parameter a, it is useful to compute the conditional law of ρ_n, once $Z_1, \ldots, Z_n$ are observed. Let $X_n = \rho_n$, $Y_n = Z_n$, $A_n = a$, $F_n = 1$, and $C_n = Z_{n-1}$. This system can be written as (1). If $a \in (-1, 1)$, then the parameter sequence is stationary and Hypothesis (H) holds.

2.2. Asymptotic properties of the error covariance matrix. We consider a linear system (1) with random parameters for which (H) holds. In particular, the process (A_n, F_n, C_n) is stationary and ergodic. In the sequel, δ is the distance on $\mathcal{P}_0$ introduced in Definition 1.4. We recall that $R_n = C_n^* C_n$ and $S_n = F_n F_n^*$. For any $n \geqq 1$, let

$$\Omega_n = \{\omega \in \Omega;\ \text{Det}\,(A_1^* R_1 A_1 + A_1^* A_2^* R_2 A_2 A_1 + \cdots + A_1^* \ldots A_n^* R_n A_n \ldots A_1) \neq 0\},$$

and

$$\Xi_n = \{\omega \in \Omega;\ \text{Det}\,(S_n + A_n S_{n-1} A_n^* + \cdots + A_n \ldots A_2 S_1 A_2^* \ldots A_n^*) \neq 0\}.$$

DEFINITION 2.1. The system (1) is called *weakly observable* if for some $n > 0$, $\mathbb{P}(\Omega_n) > 0$; it is called *weakly controllable* if for some $n > 0$, $\mathbb{P}(\Xi_n) > 0$.

When the parameters are deterministic, we recover the usual observability and controllability conditions. But these notions are much weaker than the one commonly used in the study of time-dependent systems (see, e.g., Jazwinski [16], Anderson and Moore [2]). In some of the examples given in § 2.1, only these weak conditions were natural. We will need the following lemmas.

LEMMA 2.2. *Let M_n, $n \in \mathbb{N}$, be the sequence of Hamiltonian matrices associated with a linear system* (1) *satisfying* (H). *If the system is weakly observable* (*respectively, weakly controllable*), *then, almost surely, $M_n \ldots M_1$ is in $\mathcal{H}_1$* (*respectively, $\mathcal{H}_2$*) *for all $n \in \mathbb{N}$ large enough.*

Proof. If the system is weakly observable, then $\mathbb{P}(\Omega_k) > 0$ for some $k \geqq 1$. Proposition 1.3 yields that $M_k(\omega) M_{k-1}(\omega) \dots M_1(\omega)$ is in $\mathcal{H}_1$ when $\omega \in \Omega_k$. Thus, $\mathbb{P}(M_k \dots M_1 \in \mathcal{H}_1) > 0$. It follows from the ergodic theorem that for almost all $\omega \in \Omega$ there exists an integer p, depending on ω, such that $M_{p+k}(\omega) \dots M_{p+1}(\omega) \in \mathcal{H}_1$. Since $\mathcal{H}_1$ is an ideal in $\mathcal{H}$ (cf. Proposition 1.1) this shows that, almost surely, for n large enough, $M_n \dots M_1 \in \mathcal{H}_1$. When the system is weakly controllable, the proof is similar. □

LEMMA 2.3. *For any* $Q \in \mathcal{P}_0$,

$$\text{(20)} \quad \operatorname{Max}(\operatorname{Log}^2 \|Q\|, \operatorname{Log}^2 \|Q^{-1}\|) \leqq \delta(Q, I)^2 \leqq d \operatorname{Max}(\operatorname{Log}^2 \|Q\|, \operatorname{Log}^2 \|Q^{-1}\|).$$

Proof. If $\lambda_1 \leqq \cdots \leqq \lambda_d$ are the eigenvalues of Q, then $\lambda_1 = 1/\|Q^{-1}\|$ and $\lambda_d = \|Q\|$. Since $\delta(Q, I)^2 = \sum_{i=1}^d \operatorname{Log}^2 \lambda_i$, the conclusion of the lemma is clear. □

Our main result is the following theorem. It implies that:

(i) The filtering process is successful, since the conditional error covariance matrix P_n does not explode. This error is asymptotically stationary. For instance, it converges in law (see Corollary 3.3).

(ii) Even for a fixed $\omega \in \Omega$ (outside an exceptional subset of measure 0), there is no optimal choice of the initial condition P_0, since all the sequences P_n have the same asymptotic behavior. An analogous result for the usual distance on $\mathcal{P}$ is shown in Proposition 2.5.

THEOREM 2.4. *We consider a linear system* (1) *with stochastic parameters for which*

(a) *Condition* (H) *holds.*

(b) *The system is weakly observable and weakly controllable.*

(c) *The random variables* $\operatorname{LogLog}^+ \|A_1\|$, $\operatorname{LogLog}^+ \|A_1^{-1}\|$, $\operatorname{LogLog}^+ \|C_1\|$, $\operatorname{LogLog}^+ \|F_1\|$ *are integrable.*

Then, there exists an ergodic stationary $\mathcal{P}_0$*-valued process* $\{\bar{P}_n, n \in \mathbb{Z}\}$, *that is solution of* (5). *Furthermore, there is a negative real number* $\alpha < 0$ *such that, almost surely, for any solution* P_n *of* (5) *for which* $P_0 \in \mathcal{P}_0$,

$$\overline{\lim_{n \to +\infty}} \frac{1}{n} \operatorname{Log} \delta(P_n, \bar{P}_n) \leqq \alpha < 0.$$

Proof. We are going to apply Theorem 3.1, proved in the Appendix, to the sequence $\{\Phi_n, n \in \mathbb{Z}\}$ of random contractions of the metric space $(\mathcal{P}_0, \delta)$ defined by (4). We first check the condition (C1) of this theorem, namely that for some P in $\mathcal{P}_0$, $\mathbb{E}[\operatorname{Log} \delta(\Phi_1(P), P)]$ is finite. Actually, we choose P equal to the identity matrix I. Let $T = A_1 A_1^* + S_1$. We get $\Phi_1(I) = T(I + R_1 T)^{-1} = (T^{-1} + R_1)^{-1}$. Since $T - (T^{-1} + R_1)^{-1}$ is a nonnegative matrix, we have

$$\|\Phi_1(I)\| \leqq \|(T^{-1} + R_1)^{-1}\| \leqq \|T\| \leqq \|A_1 A_1^*\| + \|S_1\| \leqq \|A_1\|^2 + \|F_1\|^2$$

and

$$\|\Phi_1(I)^{-1}\| \leqq \|T^{-1} + R_1\| \leqq \|T^{-1}\| + \|R_1\| \leqq \|(A_1 A_1^*)^{-1}\| + \|R_1\| \leqq \|A_1^{-1}\|^2 + \|C_1\|^2.$$

By Lemma 2.3,

$$\text{(21)} \quad \delta(\Phi_1(I), I)^2 \leqq d \operatorname{Max}(\operatorname{Log}^2 \|\Phi_1(I)\|, \operatorname{Log}^2 \|\Phi_1(I)^{-1}\|).$$

Using these inequalities and hypothesis (c) from Theorem 2.4, we see that $\mathbb{E}[\operatorname{Log} \delta(\Phi_1(I), I)]$ is finite. Now we check that Condition (C2) holds. By Theorem 1.7, Φ_n is a contraction. Thus, it suffices to show that the coefficient of contraction $\rho(\Phi_p \circ \cdots \circ \Phi_1)$ is smaller than 1 for some $p > 0$, with positive probability. Let M_n, $n \in \mathbb{N}$, be the Hamiltonian matrices associated to (1). It follows from Lemma 2.2 that, almost surely, for all $p \in \mathbb{N}$ large enough, $M_p \dots M_1$ is both in $\mathcal{H}_1$ and in $\mathcal{H}_2$. Since

$\mathcal{H}_0 = \mathcal{H}_1 \cap \mathcal{H}_2$, this yields that for some p, $\mathbb{P}(M_p \dots M_1 \in \mathcal{H}_0) \neq 0$. By (12), $\Phi_p \circ \cdots \circ \Phi_1 = \Phi_{M_p \dots M_1}$, and therefore, $\mathbb{P}(\rho(\Phi_p \circ \cdots \circ \Phi_1) < 1) \neq 0$ by Theorem 1.7 (iii). Thus, Condition (C2) of Theorem 3.1 holds. This theorem implies the result. □

PROPOSITION 2.5. *We suppose that the hypotheses of Theorem* 2.4 *hold and that* $\operatorname{Log}^+ \|A_1\|$, $\operatorname{Log}^+ \|A_1^{-1}\|$, $\operatorname{Log}^+ \|C_1\|$, *and* $\operatorname{Log}^+ \|F_1\|$ *are integrable. Then, almost surely, for any solution* P_n *of* (5) *for which* $P_0 \in \mathcal{P}_0$,

$$\overline{\lim_{n\to+\infty}} \frac{1}{n} \operatorname{Log} \|P_n - \bar{P}_n\| \leqq \alpha < 0. \tag{22}$$

Proof. Since P_n and $\bar{P}_n$ are symmetric positive matrices, we can find matrices K_n and D_n such that $\bar{P}_n = K_n^* K_n$, $P_n = K_n^* D_n K_n$, and such that D_n is a diagonal matrix with positive entries $\lambda_1^{(n)}, \dots, \lambda_d^{(n)}$. We have $\delta(P_n, \bar{P}_n) = \{\sum_{i=1}^d \operatorname{Log}^2 \lambda_i^{(n)}\}^{1/2}$. It follows from Theorem 2.4 that

$$\overline{\lim_{n\to+\infty}} \frac{1}{n} \operatorname{Log} |\operatorname{Log} \lambda_i^{(n)}| \leqq \alpha,$$

for $i = 1, \dots, d$. This implies that

$$\overline{\lim_{n\to+\infty}} \frac{1}{n} \operatorname{Log} |1 - \lambda_i^{(n)}| \leqq \alpha. \tag{23}$$

As in the proof above, we see that $\mathbb{E}[\delta(\Phi_1(I), I)]$ is finite by (21). Moreover, by Lemma 2.3, $\operatorname{Log} \|\bar{P}_n\| \leqq \delta(I, \bar{P}_n)$. Thus, it follows from Proposition 3.4 that, almost surely,

$$\overline{\lim_{n\to+\infty}} \frac{1}{n} \operatorname{Log} \|\bar{P}_n\| \leqq 0. \tag{24}$$

For each $x \in \mathbb{R}^d$, $\|K_n x\|^2 = \langle K_n^* K_n x, x\rangle \leqq \|\bar{P}_n x\|$, so that $\|K_n\|^2 \leqq \|\bar{P}_n\|$. This yields that

$$\|P_n - \bar{P}_n\| \leqq \|K_n^*(D_n - I) K_n\| \leqq \|K_n\|^2 \|D_n - I\| \leqq \|\bar{P}_n\|^2 \max_{1 \leqq i \leqq d} |1 - \lambda_i^{(n)}|.$$

It is clear that (22) is a consequence of (23), (24), and of this inequality. □

Remark. It is not difficult to see that the conclusion of the theorem also holds when P_0 is only in $\mathcal{P}$. (The main point is to note that since the system is weakly controllable, there is almost surely an integer k such that $M_k \dots M_1$ is in $\mathcal{H}_2$, by Lemma 2.2, which implies that $P_k = \Phi_{M_k \dots M_1}(P_0)$ is in $\mathcal{P}_0$, by Proposition 1.5 (ii).)

2.3. Stability of the filter. We show that the linear equation (6) of the filter is exponentially stable. We make use of the classical method of Lyapunov, as in Anderson and Moore [2], for instance.

THEOREM 2.6. *We consider a system* (1) *with stochastic parameters for which*:

(i) *Hypothesis* (H) *holds.*

(ii) *The system is weakly observable and weakly controllable.*

(iii) *The random variables* $\operatorname{Log}^+ \|A_1\|$, $\operatorname{Log}^+ \|A_1^{-1}\|$, $\operatorname{Log}^+ \|C_1\|$, $\operatorname{Log}^+ \|F_1\|$ *are integrable. Then the equation* (6) *of the filter is exponentially stable: namely, there is a real number* $\gamma > 0$ *such that, almost surely,*

$$\overline{\lim_{n\to+\infty}} \frac{1}{n} \operatorname{Log} \|(A_n - P_n R_n A_n) \dots (A_1 - P_1 R_1 A_1)\| \leqq -\gamma < 0,$$

for any solution $\{P_n, n \in \mathbb{N}\}$ *of* (5) *such that* $P_0 \in \mathcal{P}_0$.

We need the following classical lemma. It is an immediate consequence of the relation

$$\hat{X}_n - X_n = (A_n - P_n R_n A_n)(\hat{X}_{n-1} - X_{n-1}) + P_n C_n^* \eta_n + (P_n R_n - I) F_n \varepsilon_n,$$

which itself results from (1) and (6).

LEMMA 2.7. *Let* $B_n = A_n - P_n R_n A_n$ *and* $T_n = P_n R_n P_n + (I - P_n R_n) S_n (I - P_n R_n)^*$. *Then,* $P_n = B_n P_{n-1} B_n^* + T_n$.

LEMMA 2.8. *Let* $G_n = T_n + B_n T_{n-1} B_n^* + \cdots + B_n \dots B_2 T_1 B_2^* \dots B_n^*$. *Under the hypotheses of Theorem* 2.6, *there exists* $n \in \mathbb{N}$ *such that* $\mathbb{P}(\text{Det}(G_n) \neq 0) > 0$.

Proof. We follow an argument in Anderson and Moore [2], where a similar result is proved. For each integer $i \geqq 1$, let $K_i = P_i C_i^*$ and $H_i = (I - P_i R_i)$. We consider the two $d \times (p+q)n$ matrices

$$W_n = (K_n, F_n, A_n K_{n-1}, A_n F_{n-1}, \dots, A_n A_{n-1} \dots A_2 K_1, A_n A_{n-1} \dots A_2 F_1)$$

$$V_n = (K_n, H_n F_n, B_n K_{n-1}, B_n H_{n-1} F_{n-1}, \dots, B_n B_{n-1} \dots B_2 K_1, B_n B_{n-1} \dots B_2 H_1 F_1).$$

It is easy to see by straightforward manipulations that there exists a $(p+q)n \times (p+q)n$ upper triangular matrix U_n, with all diagonal terms equal to 1, such that $V_n = W_n U_n$. We remark that

$$W_n W_n^* \geqq S_n + A_n S_{n-1} A_n^* + \cdots + A_n \dots A_2 S_1 A_2^* \dots A_n^*.$$

Since the system is weakly controllable, there exists a positive integer n such that the subset Ξ_n, where the right-hand side is invertible, is of positive measure. On Ξ_n, the rank of W_n is d. The same property holds for V_n since $V_n = W_n U_n$ and since U_n is invertible. Then the lemma results from the fact that $G_n = V_n V_n^*$. □

Proof of Theorem 2.6. For notational simplicity we suppose that $\mathbb{P}(\text{Det}(G_1) \neq 0) > 0$. The general case is treated in the same way (by looking at the sequence P_{nk}, $k \in \mathbb{N}$, where n is given by Lemma 2.8). For any $n \in \mathbb{N}$, let $\lambda_n = \|T_n^{-1}\|^{-1}$, $\sigma_n = \|P_n\|$, $\alpha = \|P_0^{-1}\|$, with the convention that $\lambda_n = 0$, if T_n is not invertible. Let p be a positive integer. For a fixed $x_p \in \mathbb{R}^d$, we define a finite sequence $x_0, x_1, \dots, x_p$, by the backward recursion $x_n = B_{n+1}^* x_{n+1}$. Let $V_n = x_n^* P_n x_n$. We have

$$V_{n+1} - V_n = x_{n+1}^* T_{n+1} x_{n+1} \geqq \lambda_{n+1} \|x_{n+1}\|^2 \geqq \frac{\lambda_{n+1}}{\sigma_{n+1}} V_{n+1}$$

so that, if $\tau_n = (1 - \lambda_n / \sigma_n)$, then $V_n \leqq \tau_{n+1} V_{n+1}$. Therefore,

$$\|x_0\|^2 \leqq \|P_0^{-1}\| V_0 \leqq \|P_0^{-1}\| \tau_1 \dots \tau_p V_p \leqq \|P_0^{-1}\| \tau_1 \dots \tau_p \|P_p\| \, \|x_p\|^2.$$

Since $x_0 = B_1^* \dots B_p^* x_p$, this implies that

$$\|B_p \dots B_1\|^2 \leqq \alpha \tau_1 \dots \tau_p \sigma_p$$

and

$$\frac{1}{p} \text{Log} \|B_p \dots B_1\|^2 \leqq \frac{1}{p} \text{Log}\, \alpha + \frac{1}{p} \text{Log}\, \sigma_p + \frac{1}{p} \sum_{i=1}^{p} \text{Log}\, \tau_i.$$

As in the proof of (24) we can apply Proposition 3.4 to see that, almost surely,

$$\overline{\lim_{p \to +\infty}} \frac{1}{p} \text{Log}\, \sigma_p \leqq 0.$$

Therefore, it follows from Birkhoff's ergodic theorem (see the proof of Corollary 3.2) that, almost surely,

$$\overline{\lim_{p\to+\infty}}\frac{1}{p}\operatorname{Log}\|B_p\dots B_1\|^2\leqq \mathbb{E}(\operatorname{Log}\tau_1).$$

Since we have supposed that $\mathbb{P}(\operatorname{Det}(G_1)\neq 0)>0$, we know that $\mathbb{E}(\operatorname{Log}\tau_1)<0$. □

Remark 1. The exponential rate γ can be chosen to be equal to the smallest positive Lyapunov exponent of the associated Hamiltonian matrices. This is shown in Bougerol [7]. 2. It is not difficult to see that the theorem also holds if we only suppose that $P_0\in\mathcal{P}$.

3. Appendix. Iteration of stationary Lipschitz functions. In this appendix, we establish some general properties of the processes that are obtained by iteration of random Lipschitz functions. At least in particular cases, similar results are already known. But we think that our set-up and formulation can be useful in several situations; we applied them in § 2.

Let (E,δ) be a complete separable metric space. A Lipschitz map $\phi: E\to E$ is a map for which

$$\rho(\phi):=\operatorname{Sup}\left\{\frac{\delta(\phi(x),\phi(y))}{\delta(x,y)};\ x,y\in E, x\neq y\right\}$$

is finite. If ϕ and φ are such Lipschitz maps, then

$$\rho(\phi\circ\varphi)\leqq\rho(\phi)\rho(\varphi). \tag{25}$$

When $\rho(\phi)\leqq 1$, the map ϕ is called a *contraction.* It is called a *uniform contraction* when $\rho(\phi)<1$. We consider a stationary ergodic process $\{\phi_n, n\in\mathbb{Z}\}$ defined on a probability space $(\Omega,\mathcal{A},\mathbb{P})$, where each $\phi_n: E\to E$ is a random Lipschitz map (we suppose that the maps $(\omega,x)\in\Omega\times E\to\phi_n^\omega(x)\in E$ are measurable when E is equipped with its Borel σ-algebra; for notational convenience, we do not write ω explicitly). We consider the processes X_n, $n\in\mathbb{N}$, on E for which the following difference equation holds:

$$X_n=\phi_n(X_{n-1}). \tag{26}$$

The following theorem is more or less known. It generalizes results of Sunyach [26], Brandt [9], and Barnsley and Elton [4]. We recall that $\operatorname{Log}^+ x=\operatorname{Max}(\operatorname{Log}x,0)$. If E' is a countable dense subset of E, then $\rho(\phi_1)$ is the supremum of the countable set $\{\delta(\phi_1(x),\phi_1(y))/\delta(x,y);\ x,y\in E', x\neq y\}$; thus, $\rho(\phi_1)$ is measurable.

THEOREM 3.1. *Let $\{\phi_n, n\in\mathbb{Z}\}$ be a stationary ergodic sequence of Lipschitz maps from E into E. We suppose that the following conditions hold*:

(C1) *For some x in E, $\mathbb{E}[\operatorname{Log}^+\delta(\phi_1(x),x)]$ is finite.*

(C2) *The random variable $\operatorname{Log}^+\rho(\phi_1)$ is integrable, and for some integer $p>0$, the real number*

$$\alpha=\frac{1}{p}\mathbb{E}[\operatorname{Log}\rho(\phi_p\circ\phi_{p-1}\circ\cdots\circ\phi_1)]$$

is strictly negative.

Then there exists an ergodic stationary process $\{\tilde{X}_n, n\in\mathbb{Z}\}$ with values in E, solution of (26), *such that, almost surely,*

$$\overline{\lim_{n\to+\infty}}\frac{1}{n}\operatorname{Log}\delta(X_n,\tilde{X}_n)\leqq\alpha<0$$

for any process $\{X_n, n\geqq 0\}$, such that $X_n=\phi_n(X_{n-1})$ for all $n>0$.

Proof. Let us first show that, almost surely,

$$\overline{\lim_{k\to+\infty}} \frac{1}{k} \operatorname{Log} \rho(\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k}) \leqq \alpha. \tag{27}$$

By (25),

$$\overline{\lim_{k\to+\infty}} \frac{1}{k} \operatorname{Log} \rho(\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k}) \leqq \overline{\lim_{k\to+\infty}} \frac{1}{k} \operatorname{Log} \rho(\phi_{-1} \circ \cdots \circ \phi_{-k}),$$

thus, the left-hand side of (27) is a subinvariant function. By ergodicity, it is constant, almost surely. Let β be this constant. We know that for any nonnegative integrable random variable Z, $\sum_{k=0}^{+\infty} \mathbb{P}(Z > k)$ is finite. Therefore, the integrability condition in (C2) entails that, for any $\varepsilon > 0$, $\sum_{k=0}^{\infty} \mathbb{P}(\operatorname{Log}^+ \rho(\phi_1) > k\varepsilon) < +\infty$. Since all the ϕ_n's have the same law, this entails that $\sum_{k=0}^{+\infty} \mathbb{P}(\operatorname{Log}^+ \rho(\phi_{-k}) > k\varepsilon) < +\infty$, so that by the Borel-Cantelli lemma, almost surely,

$$\overline{\lim_{k\to+\infty}} \frac{1}{k} \operatorname{Log} \rho(\phi_{-k}) \leqq 0.$$

By making use of this inequality, and of the fact that if $k = mp + r$, where r is an integer in $[0, p)$,

$$\operatorname{Log} \rho(\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k}) \leqq \left\{ \sum_{i=0}^{m-1} \operatorname{Log} \rho(\phi_{-ip} \circ \phi_{-ip-1} \circ \cdots \circ \phi_{-(i+1)p+1}) \right\} + \operatorname{Log} \rho(\phi_{-mp} \circ \cdots \circ \phi_{-k}),$$

we see that

$$\beta \leqq \overline{\lim_{k\to+\infty}} \frac{1}{mp} \sum_{i=0}^{m-1} \operatorname{Log} \rho(\phi_{-ip} \circ \phi_{-ip-1} \circ \cdots \circ \phi_{-(i+1)p+1}).$$

It follows from Birkhoff's ergodic theorem that the expectation of the right-hand side is equal to α, proving (27). On the other hand, it follows as above from the integrability condition (C1) and from the Borel-Cantelli lemma that for some fixed x in E, almost surely,

$$\overline{\lim_{k\to+\infty}} \frac{1}{k} \operatorname{Log} \delta(\phi_{-k}(x), x) \leqq 0. \tag{28}$$

Now

$$\delta((\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k})(x), (\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k} \circ \phi_{-k-1})(x)) \leqq \rho(\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k}) \delta(x, \phi_{-k-1}(x)),$$

thus, (27) and (28) imply that, almost surely,

$$\overline{\lim_{k\to+\infty}} \frac{1}{k} \operatorname{Log} \delta((\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k})(x), (\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k} \circ \phi_{-k-1})(x)) \leqq \alpha.$$

Since $\alpha < 0$, this shows that $\{(\phi_0 \circ \phi_{-1} \circ \cdots \circ \phi_{-k})(x), k \in \mathbb{N}\}$ is a Cauchy sequence for almost all $\omega \in \Omega$. We suppose that E is complete, thus, this sequence converges. In the same way, we see that $(\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_{n-k})(x)$ converges, almost surely, for each fixed $n \in \mathbb{Z}$, when $k \to +\infty$. Let

$$\tilde{X}_n = \lim_{k\to+\infty} (\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_{n-k})(x). \tag{29}$$

Since $\{\phi_k, k \in \mathbb{Z}\}$ is a stationary ergodic process, and since we can write $\tilde{X}_n = F(\phi_k, k \leqq n)$ for some measurable function F independent of $n \in \mathbb{Z}$, we see that $\tilde{X}_n$ is itself a stationary ergodic process. It is clear that it is a solution of (26). Finally, let $\{X_n, n \geqq 0\}$ be any process satisfying (26). Since

$$\delta(X_n, \tilde{X}_n) = \delta((\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_1)(X_0), (\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_1)(\tilde{X}_0))$$
$$\leqq \rho(\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_1) \delta(X_0, \tilde{X}_0),$$

we see that, almost surely,

$$\overline{\lim_{n \to +\infty}} \frac{1}{n} \operatorname{Log} \delta(X_n, \tilde{X}_n) \leqq \frac{1}{p} \mathbb{E}(\operatorname{Log} \rho(\phi_p \circ \phi_{p-1} \circ \cdots \circ \phi_1)) = \alpha.$$

This concludes the proof of the theorem. □

Remark. Suppose that $\operatorname{Log}^+ \rho(\phi_1)$ is integrable, then if $\mathbb{E}\{\operatorname{Log} \delta(\phi_1(x_0), x_0)\}$ is finite for some x_0 in E, then it is finite for all $x \in E$ since

$$\delta(\phi_1(x), x) \leqq \delta(\phi_1(x), \phi_1(x_0)) + \delta(\phi_1(x_0), x_0) + \delta(x_0, x)$$
$$\leqq \delta(\phi_1(x_0), x_0) + (\rho(\phi_1) + 1)\delta(x_0, x).$$

COROLLARY 3.2. *Let π be the common law of the $\tilde{X}_n$'s. Under the hypotheses above, almost surely, for any sequence X_n satisfying* (26), *and for any bounded continuous function $f: E \to \mathbb{R}$,*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} f(X_i) = \int f \, d\pi.$$

Proof. The hypotheses on E imply that there exists a countable set D of bounded continuous functions on E such that any sequence of probability measures μ_n on E converges weakly to a probability measure μ if and only if $\int f \, d\mu_n \to \int f \, d\mu$ for any f in D (see, e.g., Parthasarathy [23, Thm. II.6.6]). Let $\Omega(f)$ be the subset of Ω, where $1/n \sum_{i=1}^{n} f(\tilde{X}_i)$ converges to $\int f \, d\pi$. Since $(\tilde{X}_n)$ is stationary and ergodic, $\mathbb{P}(\Omega(f)) = 1$ by Birkhoff's ergodic theorem. Let Ω_0 be the intersection of the set, where $\delta(X_n, \tilde{X}_n)$ converges to 0 and of all the sets $\Omega(f)$, $f \in D$. It follows from Theorem 3.1 that $\mathbb{P}(\Omega_0) = 1$. We fix an ω in Ω_0. Let m_n be the empirical measure of the sequence $\{\tilde{X}_n(\omega), n \geqq 1\}$, defined by $\int f \, dm_n = 1/n \sum_{i=1}^{n} f(\tilde{X}_i(\omega))$, when $f: E \to \mathbb{R}$ is bounded and continuous. The sequence $\{m_n, n \geqq 1\}$ converges weakly to π. Moreover, since $\delta(X_n(\omega), \tilde{X}_n(\omega)) \to 0$,

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} f(X_i(\omega)) = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} f(\tilde{X}_i(\omega)) = \int f \, d\pi \tag{30}$$

when f is uniformly continuous. This, in turn, yields that the empirical measure of the sequence $\{X_n(\omega), n \geqq 1\}$ converges weakly to π, i.e., that (30) holds for any function f that is bounded and continuous (see Parthasarathy [23, Thm. II.6.1]). □

COROLLARY 3.3. *Under the hypotheses of the previous theorem, all the solutions X_n of* (26) *converge in law to the same limit π.*

This corollary is an immediate consequence of the theorem. The following technical proposition has been used in § 2.

PROPOSITION 3.4. *Suppose that the random maps ϕ_n are contractions, that $\mathbb{E}[\delta(\phi_1(x), x)]$ is finite for some x in E, and that for some $p > 0$, $\mathbb{E}[\operatorname{Log}(\rho(\phi_p \circ \cdots \circ \phi_1))] < 0$. Then, almost surely,*

$$\lim_{n \to +\infty} \frac{1}{n} \delta(x, X_n) = 0$$

for any sequence $\{X_n, n \geqq 0\}$ for which (26) *holds.*

Proof. For any $n \geqq 1$,

$$\delta(x, (\phi_n \circ \phi_{n-1} \circ \cdots \circ \phi_1)(x)) \leqq \delta(x, (\phi_n \circ \cdots \circ \phi_2)(x))$$
$$+ \delta((\phi_n \circ \cdots \circ \phi_2)(x), (\phi_n \circ \cdots \circ \phi_1)(x))$$
$$\leqq \delta(x, (\phi_n \circ \cdots \circ \phi_2)(x)) + \rho(\phi_n \circ \cdots \circ \phi_2) \delta(x, \phi_1(x)).$$

So that, by induction,

$$(31)\qquad \delta(x,(\phi_n\circ\phi_{n-1}\circ\cdots\circ\phi_1)(x))\leqq\sum_{i=1}^{n}\rho(\phi_n\circ\cdots\circ\phi_{i+1})\delta(x,\phi_i(x)).$$

Since the ϕ_n's are contractions, this implies in particular, using (25), that

$$\delta(x,(\phi_n\circ\phi_{n-1}\circ\cdots\circ\phi_1)(x))\leqq\delta(x,\phi_n(x))+\sum_{i=1}^{n-1}\rho(\phi_{i+1})\delta(x,\phi_i(x)).$$

Thus, by Birkhoff's ergodic theorem, almost surely,

$$\overline{\lim_{n\to+\infty}}\frac{1}{n}\delta(x,(\phi_n\circ\cdots\circ\phi_1)(x))\leqq\overline{\lim_{n\to+\infty}}\frac{1}{n}\sum_{i=1}^{n-1}\rho(\phi_{i+1})\delta(x,\phi_i(x))\leqq\mathbb{E}[\rho(\phi_1)\delta(x,\phi_0(x))].$$

Similarly, for any fixed k in $\mathbb{N}$, we obtain from (31) that, when $n>k$,

$$\delta(x,(\phi_n\circ\cdots\circ\phi_1)(x))\leqq\sum_{i=n-k+1}^{n}\rho(\phi_n\circ\cdots\circ\phi_{i+1})\delta(x,\phi_i(x))$$
$$+\sum_{i=1}^{n-k}\rho(\phi_{i+k}\circ\cdots\circ\phi_{i+1})\delta(x,\phi_i(x))$$

and by the ergodic theorem,

$$(32)\qquad \begin{aligned}\overline{\lim_{n\to+\infty}}\frac{1}{n}\delta(x,(\phi_n\circ\cdots\circ\phi_1)(x))&\leqq\overline{\lim_{n\to+\infty}}\frac{1}{n}\sum_{i=1}^{n-1}\rho(\phi_{i+k}\circ\cdots\circ\phi_{i+1})\delta(x,\phi_i(x))\\&\leqq\mathbb{E}[\rho(\phi_k\circ\cdots\circ\phi_1)\delta(x,\phi_0(x))].\end{aligned}$$

Now, $\rho(\phi_k\circ\cdots\circ\phi_1)\delta(x,\phi_0(x))$ converges, almost surely, to 0 as $k\to+\infty$ and is dominated by the integrable function $\delta(x,\phi_0(x))$. Thus, its expectation goes to 0 by Lebesgue's theorem; from (32) we obtain

$$\overline{\lim_{n\to+\infty}}\frac{1}{n}\delta(x,(\phi_n\circ\phi_{n-1}\circ\cdots\circ\phi_1)(x))=0$$

almost surely. Finally, since $X_n=(\phi_n\circ\cdots\circ\phi_1)(X_0)$, we see that

$$\begin{aligned}\delta(x,X_n)&\leqq\delta(x,(\phi_n\circ\cdots\circ\phi_1)(x))+\delta((\phi_n\circ\cdots\circ\phi_1)(x),(\phi_n\circ\cdots\circ\phi_1)(X_0))\\&\leqq\delta(x,(\phi_n\circ\cdots\circ\phi_1)(x))+\delta(x,X_0),\end{aligned}$$

so that

$$\delta(x,X_n)/n\to 0\qquad\text{as}\qquad n\to+\infty.\qquad\square$$

Remark. When the random maps ϕ_n are independent and identically distributed, then $\mathbb{E}(\delta(\tilde{X}_n,x))$ is finite under the hypothesis of this proposition. In that case, its conclusion follows directly from the Borel–Cantelli lemma.

REFERENCES

[1] R. AKELLA AND P. R. KUMAR, *Optimal control production rate in a failure prone manufacturing system*, IEEE Trans. Automat. Control, AC-31 (1985), pp. 116–126.

[2] B. D. O. ANDERSON AND J. B. MOORE, *Detectability and stabilizability of time-varying discrete-time linear systems*, SIAM J. Control Optim., 19 (1981), pp. 20–32.

[3] A. V. BALAKRISHNAN, *Kalman filtering theory*, Optimization Software, Inc., New York, 1984.

[4] M. F. BARNSLEY AND J. H. ELTON, *A new class of Markov processes for image encoding*, Adv. Appl. Probab., 20 (1988), pp. 14–32.

[5] G. BIRKHOFF, *Extensions of Jentzsch's theorem*, Trans. Amer. Math. Soc., 85 (1957), pp. 219–227.

[6] S. BITTANTI, P. COLANERI, AND G. DE NICOLAO, *The difference periodic Riccati equation for the periodic prediction problem*, IEEE Trans Automat. Control, AC-33 (1988), pp. 706–712.

[7] PH. BOUGEROL, *Filtre de Kalman Bucy et exposants de Lyapounov*, in Lyapunov Exponents, Proceedings, Oberwolfach 1990, L. Arnold, H. Crauel, and J. P. Eckmann, eds., Lecture Notes in Math., Springer-Verlag, Berlin, Heidelberg, New York, 1486 (1991), pp. 112-122.

[8] ———, *Some results on the filtering equation with random parameters*, in Applied Stochastic Analysis, I. Karatzas and D. Ocone, eds., Lecture Notes in Control, Springer-Verlag, Berlin, Heidelberg, New York, 177 (1992), pp. 30-37.

[9] A. BRANDT, *The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients*, Adv. Appl. Probab., 18 (1986), pp. 211-220.

[10] P. E. CAINES AND S. P. MEYN, *A new approach to stochastic adaptative control.*, IEEE Trans. Automat. Control, AC-32 (1987), pp. 220-226.

[11] S. S. CHANG, *Optimum filtering and control of randomly sampled systems*, IEEE Trans. Automat. Control, AC-12 (1967), pp. 537-546.

[12] W. L. DE KONING, *Optimal estimation of linear discrete-time systems with stochastic parameters*, Automatica, 20 (1984), pp. 113-115.

[13] C. E. DE SOUZA AND G. C. GOODWIN, *Periodic solutions of matrix Riccati equation in optimal filtering of nonstabilizable periodic systems*, Proc. 10th World IFAC Congress, R. Ishermann, ed., Pergamon Press, IX (1987), pp. 243-248.

[14] F. R. GANTMACHER, *The theory of matrices*, Vol. 1, Chelsea Press, New York, 1959.

[15] R. HERMANN, *Cartanian geometry, nonlinear waves, and control theory*, Part A, Interdisciplinary Mathematics, Vol. 20, Math Sci. Press, Brookline, MA, 1979.

[16] A. H. JAZWINSKI, *Stochastic processes and filtering theory*, Academic Press, New York, 1970.

[17] R. E. KALMAN, *Analysis and synthesis of linear systems operating on randomly sampled data*, Ph.D. thesis, Columbia University, New York, 1957.

[18] ———, *Control of randomly varying linear dynamical systems*, Proc. Sympos. Appl. Math., 13 (1961), pp. 287-298.

[19] H. MAASS, *Siegel's modular forms and Dirichlet series*, Lecture Notes in Math., 216, Springer-Verlag, Berlin, Heidelberg, New York, 1971.

[20] M. MARITON, *Jump linear quadratic control with random state discontinuities*, Automatica—J. IFAC, 23 (1987), pp. 237-240.

[21] N. E. NAHI, *Optimal recursive estimation with uncertain observation*, IEEE Trans. Inform. Theory, IT-15 (1969), pp. 457-462.

[22] G. I. OL'SHANSKII, *Invariant cones in Lie algebras, Lie semigroups, and the holomorphic discrete series*, Functional Anal. Appl., 15 (1981), pp. 275-285.

[23] K. R. PARTHASARATHY, *Probability measures on metric spaces*, Academic Press, New York, London, 1967.

[24] M. A. SHAYMAN, *Phase portrait of the matrix Riccati equation*, SIAM J. Control Optim., 24 (1986), pp. 1-65.

[25] D. L. SNYDER AND P. M. FISHMAN, *How to track a swarm of fireflies by observing their flashes*, IEEE Trans. Inform. Theory, IT-21 (1975), pp. 692-695.

[26] C. SUNYACH, *Une classe de chaines de Markov récurrentes sur un espace métrique complet*, Ann. Inst. H. Poincaré, 11 (1975), pp. 329-343.

[27] A. TERRAS, *Harmonic analysis on symmetric spaces and applications* II, Springer-Verlag, Berlin, Heidelberg, New York, 1988.

[28] M. C. VIANO, *Iterations aléatoires et filtrage de Kalman*, C.R.A.S., 305 (1987), Série 1, pp. 831-834.

[29] P. WHITTLE, *Optimization over time*, Vol. 1, Wiley, Chichester, New York, 1982.

[30] J. L. WILLEMS AND J. C. WILLEMS, *Robust stabilization of uncertain systems*, SIAM J. Control Optim., 21 (1983), pp. 352-374.

[31] M. WOJTKOWSKI, *Invariant families of cones and Lyapunov exponents*, Ergodic Theory Dynamic Systems, 5 (1985), pp. 145-161.

[32] ———, *Measure theoretic entropy of the system of hard spheres*, Ergodic Theory Dynamic Systems, 8 (1988), pp. 133-153.