An Approach, Via Entropy, to the Stability of Random Large-Scale Sampled-Data Systems Under Structural Perturbations[1]

Guy Jumarie
Department of Mathematics,
Université du Québec à Montréal,
Montreal, Quebec, H3C 3P8
Canada
Mem. ASME

The concept of entropy in information theory is used to investigate the sensitivity and the stability of sampled-data systems in the presence of random perturbations. After a brief background on the definition, the practical meaning and the main properties of the entropy, its relations with asymptotic insensitiveness are exhibited and then some new results on the sensitivity and the stochastic stability of linear and nonlinear multivariable sampled data systems are derived. A new concept of stochastic conditional asymptotic stability is obtained which seems to be of direct application in the analysis of large-scale systems. Sufficient conditions for stability are stated. This approach provides a new look over stochastic stability. In addition, variable transformations act additively on entropy, via Jacobian determinant, and as a result the corresponding calculus is very simple.

1 Introduction

When a control system is designed to follow a given trajectory, referred to as optimal by the designer, a problem of importance is to analyze what happens when, because of some accidental reasons, the system deviates from this optimal trajectory. After the external disturbance vanishes, will the deviation increase with time, or on the contrary, will it decrease in such a manner that the system will again follow the optimal trajectory?

This problem, of course, is well known. When the external disturbance affects directly the state of the system, we have a problem of stability; when it affects the parameter of the system, we then deal with structural stability, or else rough stabilty, sensitivity, robustness, etc. Although these problems have slightly a different meaning from each other, their essential content is the same. Considerable amount of effort has been devoted to these questions in the deterministic framework and the corresponding literature is by now standard.

The problem is quite different in the stochastic framework, for instance in large-scale computerized systems subject to random disturbances, which are of increasing importance in practical applications. At the present state-of-art it seems that the most direct approach would be to combine the stability theory for large-scale systems (see for instance Šiljak [1]) with the stochastic stability theory as described by Kushner [2]. Both approaches extensively use Lyapunov functions so that the synthesis should be direct. In the same way, it seems that the stability of large-scale systems with respect to random structural perturbations could be tackled by combining some results obtained by Siljak [3] in the deterministic case with the theory of stochastic Lyapunov functions. Another approach could be to consider this problem as a problem of sensitivity in a statistical framework and to optimize a sensitivity criterion which depends explicitly upon the parameters of the system. Such approaches generally work with covariance matrices and they necessarily involve advanced results in the theory of stochastic differential equations.

In this paper an alternative approach is developed by using information theory, and more especially the so-called entropy concept for a random variable. This entropy is defined by means of the probability density function of the variable only, and furthermore, it evolves monotonically with the variance, in such a manner that it can be valuably considered as a measure for the concentration or the randomness of the variable. The advantage of this approach is four-fold. (i) Transformation of variables merely increases the entropy via the corresponding Jacobian determinant, so that the calculus as well as the theoretical framework is very simple; (ii) it provides a new look at the questions of stochastic sensitivity and stochastic stability; (iii) it seems of valuable utilization in the study of large-scale systems; (iv) and finally it exhibits some new relations between information theory and control systems.

Due to the prospect of a bright future for distributed computer networks in control systems, we herein restrict ourselves to the study of sampled data control systems. First, for the reader who is not acquainted with information theory, we recall the definition and the main properties for the entropy, which are sufficient to read the paper. Then we analyze the sensitivity of linear sampled data systems with respect to disturbances on their state and their control and applications of the results to stability are considered. We then arrive at the new concept of stochastic conditional asymptotic stability which appears to be of direct application in connective stability. Nonlinear systems are considered.

The proofs of the propositions are given in the main text as

[1] Research supported by the National Research Council of Canada.
Contributed by the Dynamic Systems and Control Division for publication in the Journal of Dynamic Systems, Measurement, and Control. Manuscript received by the Dynamic Systems and Control Division, April 4, 1981.

they may help to the meaning of the results, but the reader may drop them at first reading without loss of continuity.

2 Background on Entropy

In this section, we shall recall only those definitions and properties of entropy, which we shall need for our purpose.

2.1 Uncertainty and Entropy. Consider a random vector $X \in \mathbb{R}^n$ with the probability density function $p(x)$; and assume that X is the output of a certain random experiment E. Assume just prior to perform E, an observer R tries to guess the value of the output which he will obtain after making E. R is thus facing with a certain amount of uncertainty so involved by X, and the problem is to define a possible measure for this uncertainty.

In mathematical statistics, it is common to use the standard deviations of X or the covariance matrix of X as such a measure. In the framework of information theory, Shannon [4, 5] proposed to define this uncertainty by means of the so-called *entropy* $H(X)$ defined as

$$H(X) := -\int_{\mathbb{R}^n} p(x)\ln p(x)\,dx \tag{2.1}$$

(the Algol-Pascal-like symbol := means that the left-hand side term is defined by the right-hand side one). We herein consider the natural logarithm, but any other basis may be used since it is then equivalent to multiply (2.1) by a positive scaling factor.

In a more general framework, see for instance Aczel and Daroczy [6], Renyi obtained a quantitative measure for uncertainty in the form of the *α-entropy* $H_\alpha(X)$ defined as

$$H_\alpha(X) := \frac{1}{1-\alpha}\ln\int_{\mathbb{R}^n} p^\alpha(x)\,dx \tag{2.2}$$

where $\alpha \in \mathbb{R}$ is such that $\alpha > 0$ and $\alpha \neq 1$. α may take on any positive value different from the unity, but it has been shown [7] that only $0 < \alpha < 1$ is meaningful from a practical standpoint, since then $H_\alpha(X)$ represents a loss of information with regard to $H(X)$.

2.2 Some Properties of the Entropy.

(i) One has

$$\lim_{\alpha \to 1} H_\alpha(X) = H(X), \tag{2.3}$$

in other words, the Shannon entropy is the limiting form of the Renyi entropy. As a result, if we define

$$H_1(X) := H(X), \tag{2.4}$$

we can then work with $H_\alpha(X)$ only, in the following.

(ii) $H_\alpha(X)$ may be positive or negative. The higher $H_\alpha(X)$ is, the more X is distributed over $\mathbb{R}^n$; the lower $H_\alpha(X)$ is, the more X tends to be concentrated in $\mathbb{R}^n$. $H_\alpha(X)$ is an index for the homogeneity, or again the uniformness of the probability distribution of X.

(iii) Assume that $H_\alpha(X)$ is defined with any logarithm denoted by log, and assume that X is a one-dimensional Gaussian variable X_N with the mean value μ and the standard deviation σ; then one has

$$H_\alpha(X_N) = \log_a \sigma\sqrt{2\pi} + (\log_a \sqrt{\alpha})/(\alpha - 1), \alpha \neq 1 \tag{2.5}$$

$$H(X_N) = \log_a \sigma\sqrt{2\pi e}. \tag{2.6}$$

If we now assume that X is an n-dimensional Gaussian vector $X_{N,n}$ with the probability density

$$g(x) := (2\pi)^{-\frac{n}{2}} \left|\Sigma\right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) \tag{2.7}$$

where $|\Sigma|$ denotes the absolute value of the determinant of the covariance matrix Σ and where the superscript denotes the transpose, we have

$$H_\alpha(X_{N,n}) = \log_a \left|\Sigma\right|^{\frac{1}{2}} (2\pi)^{\frac{n}{2}} + (\log_a(\alpha)^{\frac{n}{2}})/(\alpha-1), \alpha \neq 1 \tag{2.8}$$

$$H(X_{N,n}) = \log_a\left(\left|\Sigma\right|^{\frac{1}{2}} (2\pi e)^{\frac{n}{2}}\right) \tag{2.9}$$

These equations make clear, in the special case of Gaussian variables, how the entropy $H_\alpha(X_{N,n})$ is related to the variances.

(iv) If X and Y denote two random vectors, one has

$$H_\alpha(X+Y) \le H_\alpha(X) + H_\alpha(Y). \tag{2.10}$$

Nomenclature

The numbers in parentheses refer to the equation where the symbol first appears.

x	=	$\in \mathbb{R}^n$, deterministic state of the system, (2.1)
X	=	$\in \mathbb{R}^n$, stochastic state of the system, (2.1)
$H(X)$	=	Shannon entropy of the random variable X, (2.1)
$H_\alpha(X)$	=	Renyi entropy of X, (2.2)
$X_{N,n}$	=	n-dimensional Gaussian variable, (2.7)
X_N	=	Gaussian variable (2.5)
σ	=	variance, (2.5)
Σ	=	covariance matrix, (2.7)
$\|\Sigma\|$	=	absolute value of the determinant of Σ, (2.7)
$\|D(x)/D(y)\|$	=	Jacobian of the transformation $x \to y$, (2.13)
$E(\cdot)$	=	mathematical expectation of $(\cdot)$, (3.1)
r_x	=	component of x in $\mathbb{R}^r$, $r < n$
r_X	=	component of X in $\mathbb{R}^r$, $r < n$, (3.5)
k	=	holds for kT, (3.1)
T	=	constant sampling period
$u(k)$	=	$\in \mathbb{R}^m$, control of the system, (4.1)
$A(k)$	=	(n, n) matrix, (4.1)
$B(k)$	=	(n, m) matrix, (4.1)
$\Phi(k_2, k_1)$	=	transition matrix, (4.2)
$\det(\cdot)$	=	determinant of the matrix $(\cdot)$, (4.6)
$C(k)$	=	(n, n)-matrix, (4.9)
A^T	=	matrix-transpose, (4.14)
$x^*(k)$	=	nominal trajectory
$y(k)$	=	costate of the system, (4.14)
$\tilde{A}(k)$	=	(n, n) random matrix associated with $A(k)$, (7.1)
$\tilde{a}_{ij}(k)$	=	coefficients of $\tilde{A}(k)$, (7.2)
$tr(\cdot)$	=	trace of the matrix $(\cdot)$
$f(x)$	=	nonlinear vector $(f_1, f_2, \ldots f_n)^T$, (9.1)

Likewise, by substituting integral for summation, if X_t is a time dependent random vector, one has

$$H_\alpha\left(\int_0^t X_\tau d\tau\right) \leq \int_0^t H_\alpha(X_\tau)\,d\tau \tag{2.11}$$

(v) If X is a random vector and z a deterministic one, one has

$$H_\alpha(X+z) = H_\alpha(X) \tag{2.12}$$

(vi) The α-entropy is related to the variance, and it increases from $-\infty$ to $+\infty$ as $|\Sigma|$ varies from 0 to ∞.

2.3 Substitution of Variable. The following result will be useful in the following.

Result. Assume that the random n-vector X is itself derived from the random n-vector Y by the change of variable $X = g(Y)$; where $g: \mathbb{R}^n \to \mathbb{R}^n$ is a continuous mapping. Let $q(y)$ and $D(x)/D(y)$ denote the probability density function of Y and the Jacobian of the transformation, respectively. Then one has

$$H_\alpha(X) = \frac{1}{1-\alpha}\ln\int_{\mathbb{R}^n} q^\alpha(y)\left|\frac{D(x)}{D(y)}\right|^{1-\alpha} dy. \tag{2.13}$$

Proof. (i) Let $p(x)$ denote the probability density of X. Then the equality

$$\int_{\mathbb{R}^n} p(x)\,dx = \int_{\mathbb{R}^n} p(g(y))\left|\frac{D(x)}{D(y)}\right| dy$$
$$= 1$$

directly yields

$$q(y) = p(g(y))\,|(D(x)/D(y))|$$

Now we make the substitution $x = g(y)$ in the expression of $H_\alpha(X)$ to get

$$H_\alpha(X) = \frac{1}{1-\alpha}\ln\int_{\mathbb{R}^n} p^\alpha(x)\,dx$$
$$= \frac{1}{1-\alpha}\ln\int_{\mathbb{R}^n} p^\alpha(g(y))\left|\frac{D(x)}{D(y)}\right| dy$$
$$= \frac{1}{1-\alpha}\ln\int_{\mathbb{R}^n} p^\alpha(g(y))\left|\frac{D(x)}{D(y)}\right|^{\alpha+1-\alpha} dy$$
$$= \frac{1}{1-\alpha}\ln\int_{\mathbb{R}^n} q^\alpha(y)\left|\frac{D(x)}{D(y)}\right|^{1-\alpha} dy$$

2.4 On the Practical Meaning of the Entropy. In the special case of the one dimensional Gaussian variable, equation (2.5) makes clear how the α-entropy is related to the variance σ^2. For non-Gaussian variables, this dependence is not so much explicit, but it nevertheless remains transparent: $H_\alpha(X)$ and σ vary in the same way, and $H_\alpha(X)$ decreases to $-\infty$ as σ decreases to zero.

The same remark holds for n-dimensional random vectors and their covariance matrix Σ, but here a word of caution is in order. The condition $H_\alpha(X)\downarrow -\infty$ means that X converges to a random vector whose *distribution is concentrated on a set of measure zero in* $\mathbb{R}^n$; in other words, this limit can be a random vector, but distributed in a subspace $\mathbb{R}^n$, $1 \leq r \leq n$, of $\mathbb{R}^n$. As a result, we are led to introduce the following definition.

Definition 2.1. The n-random vector $X \in \mathbb{R}^n$ is said to have a concentration of order r, (r-concentration in the following), if it is distributed in a subspace $\mathbb{R}^r$ of $\mathbb{R}^n$.

H_α can be taken as a measure of the n-concentration of X. This measure is not absolute since it is not defined for $\sigma = 0$, but it can be considered as a relative measure which provides comparisons between the concentrations of two distinct variables X and Y. Likewise, if we are interested in the r-concentration of X related to a given r-subspace, then we shall consider the α-entropy $H_\alpha(X_r)$ of the corresponding r-component of X.

2.5 A Few Comments. On a theoretical standpoint, it suffices to modify the probability density $p(x)$ by adding a Dirac delta distribution to get $H(X) = -\infty$, so that $H(X)$ may appear not satisfactory to characterize stochastic convergence in any way.

Fortunately, such a situation does not occur in practical situations. Indeed, in most cases, dynamical systems are such that the corresponding density $p(x, t)$ does satisfy functional equations which themselves imply relatively stringent continuity conditions for $p(x, t)$, therefore strong continuity conditions for any variation $\delta p(x, t)$. For instance, if $x_t \in \mathbb{R}$ is defined by an Ito stochastic differential equation, $p(x, t)$ is then the solution of a Fokker-Planck-Kolmogorov equation. From a physical point of view, this somewhat insensitiveness merely means that the system does not react to disturbances concentrated on a set with zero measure.

These are exactly similar considerations which justify the utilization of differential geometry to analyze the structural stability of dynamical systems.

It is rare to get a closed form expression for the entropy of a random variable, except for some specific distributions such as the Gaussian law, the Poisson law, or the uniform law, for instance. But in fact we need only the explicit expression for the variation of the entropy, and this can be obtained easily. Indeed, a transformation on the random variable X results into an increment in $H(X)$, and this increment is simply defined by means of the Jacobian of the transformation in question.

In other words, even though we do not have explicit expression for $H(X)$, we can analyze its variations, and it is the basic remark which supports the approach.

We shall further point out that the same remark applies to the utilization of the thermodynamical entropy in physics: we deal only with its variations!

3 Asymptotic Insensitiveness

Assume that a *deterministic* sampled data system is designed to yield a certain nominal trajectory $x^*(k)$ (k denotes the discrete time) which is considered as being optimal, in some sense to the designer. Assume now that, when it runs, this system may be subject to some accidental disturbances either in the form of variations of its structural parameters or in the form of external inputs. If these disturbances are random, then $x^*(k)$ becomes the random variable $X(k) := x^*(k) + \delta X(k)$ where $\delta X(k)$ denotes the deviation of $X(k)$ from its reference value $x^*(k)$. Well obviously, it is desirable that the effects of this disturbance vanish with time. In a statistical framework, it is common to express this requirement with the following equations

$$\lim E(X(k)) = x^*(k) \quad ,k\to\infty \tag{3.1}$$

$$\lim E((X(k)-x^*(k)))^2 = 0 \quad ,k\to\infty \tag{3.2}$$

where the symbol $E(\cdot)$ holds for the mathematical expectation. This is the *convergence in the mean square sense.*

In information theory, consider the requirements

$$\lim E(X(k)) = x^*(k) \quad ,k\to\infty \tag{3.3}$$

$$H(X(k))\downarrow -\infty \quad ,k\to\infty; \tag{3.4}$$

according to the above remark, $X(k)$ will converge to a random variable which is concentrated in $\mathbb{R}^n$. More generally, if we consider condition (3.4) applied to the component of $X(k)$ in a given subspace $\mathbb{R}^r$, we shall have the r-concentration of $X(k)$ in $\mathbb{R}^r$, therefore the following definition.

Definition 3.1. Assume that the random vector $X(k)$ is the result of a random disturbance D_n applied to the deterministic vector $x^*(k)$. Let $\mathcal{R}_j^r$ denote a given subspace of $\mathcal{R}^n$, and let $r_X(k)$ and $r_x{}^*(k)$ denote the corresponding components of $X(k)$ and $x^*(k)$, respectively. We shall say that $X(k)$ has the property of *asymptotic insensitiveness* in $\mathcal{R}_j^r$ (A.I. in the following) w.r.t. the nominal value $x^*(k)$ and the disturbance D_n, whenever the following conditions are satisfied:

$$\lim E(r_X(k)) = {}^r x(k) \quad , k \to \infty \tag{3.5}$$

$$H(r_X(k)) \downarrow = -\infty \quad , k \to \infty \tag{3.6}$$

Definition 3.2. Framework of Definition 3.1. X_k is said to have the property of *asymptotic insensitiveness of order r*, r-A.I. in the following, whenever it is asymptotically insensitive in every subspace $\mathcal{R}_j^r$ of $\mathcal{R}^n$.

Remark of Importance. Assume that the vector $X := (X_1, x_2, \ldots, x_n)^T$ is random by its first component X_1 only, then the conditions of A.I. in $\mathcal{R}^n$ guarantees the convergence of X to a deterministic variable. Assume now that $X := (X_1, x_2, \ldots, X_n)$ is random on the two components X_1 and X_n only. Then we can partition X as $X := ({}^1X, {}^2X) := ((X_1, \ldots, x_j), (x_{j+1}, \ldots, X_n))^T$ and the conditions of A.I. for ${}^1X$ and ${}^2X$, respectively, will ensure the convergence of X to a deterministic variable.

This remark is of practical interest for the applications, since in distributed computerized systems, it may happen that only a few lines are subject to external disturbances.

4 Analysis of Linear Sampled Data Systems

In this section, we shall analyze the sensitivity of linear sampled data systems w.r.t. their initial states and w.r.t. additive disturbing noises.

4.1 Definition of the System.

The system is described by the equations

$$x(k+1) = A(k)x(k) + B(k)u(k) \tag{4.1}$$

$$x(0) = x_0, \text{ given initial state}$$

with the following definitions. $x(k)$ holds for $x(kT)$ where k denotes an integer and T denotes the sampling period; $x(k) \in \mathcal{R}^n$ is the state of the system; $u(k) := u(kT) \in \mathcal{R}^m$ is the input of the system; $A(k) := A(kT)$ is the (n, n) transition matrix, $B(k) := B(kT)$ is an (n, m) matrix. $A(k)$ and $B(k)$ may depend explicitly upon k or they may not. $x(k+m)$ is then related to $x(k)$ by the equation

$$x(k+m) = \Phi(k+m,k)x(k) + \sum_{i=0}^{m-1} \Phi(k+m,k+i+1)B(k+i)u(k+i) \tag{4.2}$$

with

$$\Phi(k_2,k_1) := A(k_2-1)\cdot A(k_2-2) \ldots A(k_1+1)A(k_1), k_2 > k_1 \tag{4.3}$$

Equations (4.1)–(4.3) *hold even when the sampling period T is not constant.* In the special case where the sampling period is constant and the system is time invariant, then $A(k) = A$, $B(k) = B$ and one has

$$x(k+m) = A^m x(k) + \sum_{i=1}^{m} A^{i-1}B\,u(k+m-i) \tag{4.4}$$

or else, by putting $k_1 := k + m$ and $j := k + m - u$,

$$x(k_1) = A^{k_1-k}x(k) + \sum_{j=k}^{k_1-1} A^{k_1-j-1}B\,u(j) \tag{4.5}$$

4.2 Sensitivity w.r.t. Disturbances on the State. We now consider the following problem.

Problem 4.1. Assume that the system (4.1) is designed to follow a certain trajectory $x^*(k)$ considered as a reference. Assume further that at a given instant p it is subject to a random disturbance which moves the system state from the reference $x^*(p)$. The question we ask is under which conditions the effects of this disturbance will vanish, and the system will tend to follow the nominal trajectory $x^*(k)$? We have the following result.

Proposition 4.1. Assume that the system (4.1) is subject to a random disturbance on the state $x^*(p)$ at the instant p. Then the α-entropy $H_\alpha(X(p+m))$ of the future random state $X(p+m)$ satisfies the equation

$$H_\alpha(X(p+m+1)) = H_\alpha(X(p+m)) + \ln|\det A(p+m)|, \tag{4.4}$$

and the system is asymptotically insensitive in $\mathcal{R}^n$, w.r.t. the nominal trajectory $x^*(k)$ and the random disturbance, provided that the following conditions are satisfied, for every $m > 0$.

$$E\{X(p)\} = x^*(p) \tag{4.5}$$

$$|\det A(p+m)| < 1 \tag{4.6}$$

Proof. (i) Condition (4.5) is equivalent to condition (3.5) in the present special case.

(ii) This being so, equations (4.2) on the one hand and (2.12) on the other hand yield

$$H_\alpha(X(p+m)) = H_\alpha(\Phi(p+m,p)X(p))$$

where $X(p+m)$ denotes the new random state at $(p+m)$.

(iii) The Jacobian of the transformation $X(p) \to X(p+m)$ is $\det \Phi(p+m, p)$. Let $q(x)$ denote the probability density function of $X(p)$; equation (2.13) yields

$$H_\alpha(X(p+m)) = \frac{1}{1-\alpha} \ln \int_{\mathcal{R}^n} q^\alpha(x)\,|\det \Phi(p+m,p)|\,dx$$
$$= \ln|\det \Phi(p+m,p)| + H_\alpha(X(p)) \tag{4.7}$$

(iv) Next, combining this result with the definition (4.3) for $\Phi(\circ)$, we get

$$H_\alpha(X(p+m+1)) - H_\alpha(X(p+m)) = \ln|\det A(p+m)| \tag{4.8}$$

The system will be A.I. provided that this decrement is negative for every $m > 0$, therefore the result.

4.3 Sensitivity w.r.t Disturbances on the Control.

Problem 4.2. We now assume that the above system (4.1) is disturbed by an additive continuous random noise $W(k) \in \mathcal{R}^n$, in such a way that its equation is

$$X(k+1) = A(k)X(k) + B(k)u(k) + C(k)W(k) \tag{4.9}$$

where $C(k)$ denotes an (n, n) matrix. This disturbance occurs, starting a given instant p; it may be permanent or else volatile, in which case it is then equivalent to assume that $C(k)$ vanishes after a certain range of time. Again we ask the question under which condition the system will reach and follow its reference trajectory $x^*(k)$?

Let $H_\alpha(W(k))$ denote the α-entropy of $W(k)$; we state:

Proposition 4.2. The system (4.9) is asymptotically insensitive, in $\mathcal{R}^n$, w.r.t. the nominal trajectory $x^*(k)$ and the random noise $C(k)W(k)$, $C(k) = 0$ for $k < p$, provided that the following conditions are satisfied; for every $k \geq p$:

$$E\{W(k)\} = 0 \tag{4.10}$$

$$|\det A(k)C(k)| < \exp(-H_\alpha(W(k))). \tag{4.11}$$

Proof. (i) According to condition (4.10), equation (4.9) yields

$$E\{X(k+1\} = A(k)E\{X(k)\} + B(k)u(k)$$

therefore

$$E\{X(k)\}=x^*(k), k\geq p.$$

(ii) Now equations (2.12) and (4.9) provide

$$H_\alpha(X(k+1))=H_\alpha(A(k)X(k)+C(k)W(k))$$

therefore, by using equation (2.10),

$$H_\alpha(X(k+1))\leq H_\alpha(A(k)X(k))+H_\alpha(C(k)W(k)) \quad (4.12)$$

(iii) A calculation similar to Step (iii), equation (4.7) in the proof of proposition 4.1, yields

$$H_\alpha(A(k)X(k))+H_\alpha(C(k)W(k))=\ln|\det A(k)| + H_\alpha(X(k))+\ln|\det C(k)|+H_\alpha(W(k)),$$

and by substituting this result into (4.12) we get

$$H_\alpha(X(k+1))-H_\alpha(X(k))\leq \ln|\det A(k)C(k)| + H_\alpha(W(k)) \quad (4.13)$$

The system will be A.I. provided that

$$\ln|\det A(k)C(k)|+H_\alpha(W(k))<0$$

therefore the result.

Remark. If we desire stability w.r.t. the state and stability w.r.t. the control, then conditions (4.5, 6, 10, 11) should be simultaneously satisfied.

A Direct Generalization. Assume that the system is disturbed in such a way that its equation takes the new form

$$X(k+1)=A(k)X(k)+B(k)u(k)+C_1(k)W_1(k) + C_2(k)W_2(k), \quad (4.14)$$

$C_1(k) = C_2(k) = 0$, $k < p$. Then it is A.I. provided the following conditions are satisfied for $k \geq p$:

$$E\{W_1(k)\}=0 \quad (4.15)$$

$$E\{W_2(k)\}=0 \quad (4.16)$$

$$|\det A(k)C_1(k)C_2(k)|<\exp(-H_\alpha(W_1(k)) - H_\alpha(W_2(k)). \quad (4.17)$$

The proof is quite similar to that for proposition (4.2).

4.4 Analysis of the Conjugate System.

Problem 4.3. We consider the conjugate system associated with (4.1), in the form

$$y(k+1)=A^T(k)y(k)+\tilde{B}(k)v(k) \quad (4.18)$$

$$y(0)=y_0, \text{ given value}$$

We now assume that, at a given instant p, the conjugate system is subject to a random disturbance around the nominal trajectory $y^*(k)$ in such a way that the future state is a random state $Y(k)$, $k \geq p$. Our purpose is to determine the relations which may exist between the sensitivities of the two systems.

We have the following result.

Proposition 4.3. Assume that at a given instant p the original system and the conjugate one are subject to random disturbances around their respective nominal references $x^*(p)$ and $y^*(p)$; and let $H_\alpha(X(p))$ and $H_\alpha(Y(p))$ denote the corresponding α-entropies. Then the future state $X(p+m)$ and the costate $Y(p+m)$ evolve in such a way that the equation

$$H_\alpha(X(p+m))-H_\alpha(Y(p+m))=H_\alpha(X(p)) - H_\alpha(Y(p)). \quad (4.19)$$

is satisfied for every $m \geq 0$.

Proof. We duplicate the proof of proposition 4.1, and we take account of the relation

$$|\det(-A^T)|=|\det A| \quad (4.20)$$

to get the counterpart of equation (4.7) in the form

$$H_\alpha(Y(p+m))=\ln|\det \Phi(p+m,p)|+H_\alpha(Y(p)).$$

Now, comparing this result with equation (4.6) yields

$$H_\alpha(X(p+m)-H_\alpha(X(p))=H_\alpha(Y(p+m)) - H_\alpha(Y(p)). \quad (4.21)$$

therefore equation (4.19)

5 On the Meaning of These Results

In this section, we state a few comments to make clear the practical significance of the above results.

5.1 Asymptotic Insensitiveness and Component-Wise Convergence. The various determinants which appear in the above conditions, are essentially related to the amount of space where the disturbance is distributed. If this region is small, then the system is A.I. in $\mathcal{R}^n$. But this region may be small in $\mathcal{R}^n$, and not necessarily small in a given subspace of $\mathcal{R}^n$; a plane has a zero bulk in $\mathcal{R}^3$ but a nonzero area in $\mathcal{R}^2$; in such a manner that the A.I. conditions in $\mathcal{R}^n$ do not imply the component-wise convergence, that is the convergence for all the components. If we want to be sure of the point-wise convergence, we must require that

$$H_\alpha(X_j(k))\downarrow-\infty \text{ for every component } X_j,\ k\rightarrow\infty$$

or else

$$\sum_j(1/H_\alpha(X_j(k)))\uparrow 0^-,\ k\rightarrow\infty.$$

5.2 Conditional Asymptotic Insensitiveness. We now consider the A.I. condition (4.11) in proposition 4.2. It explicitly refers to the α-entropy $H_\alpha(W(k))$ of the disturbing noise, and in this way, it can be thought of as defining a *conditional asymptotic insensitiveness*. For a given value of det $A(k)$, the larger $H_\alpha(W(k))$ is, the smaller det $C(k)$ is, but here again this ensures A.I. in $\mathcal{R}^n$ only. If we want pointwise A.I. we must consider the effect of the noise on each component taken one at a time.

5.3 The Decomposition Approach. Assume that the disturbance applies on some components of $x(p)$ only. Then we can partition the state vector x into the components ${}^1x := (x_{11}, x_{12}, \ldots, x_{1n_1})^T$, ${}^2x := (x_{21}, x_{22}, \ldots, x_{2n_2})$ where we have renamed the components of the vector x in an obvious way. The dimensions of the two vectors x_1, x_2 are n_1, n_2 so that $n_1 + n_2 = n$. This partition yields the system (4.1) in the form

$$\begin{vmatrix} {}^1x(k+1) \\ {}^2x(k+1) \end{vmatrix} = \begin{vmatrix} A_{11}(k) & A_{12}(k) \\ A_{21}(k) & A_{22}(k) \end{vmatrix} \begin{vmatrix} {}^1x(k) \\ {}^2x(k) \end{vmatrix} + B(k)u(k). \quad (5.1)$$

The matrices A_{11} and A_{22} have the dimensions (n_1, n_1) and (n_2, n_2), respectively, while the matrices A_{12}, A_{21} have the dimensions (n_1, n_2) and (n_2, n_1), respectively. We have a decomposition into two subsystems S_1 and S_2, and proposition 4.1 together with proposition 4.2 apply.

This being so, assume that the state $x(p)$ is disturbed on one component of ${}^1x(p)$ and one component of ${}^2x(p)$ only; then, according to the remark at the end of Section 3, it is sufficient to assure the A.I. of ${}^1x(k)$ and ${}^2x(k)$, respectively, to get the convergence of $X(k)$ to a deterministic variable.

5.4 Asymptotic Insensitiveness and Mean-Square Convergence. If a random variable $X(k) \in \mathcal{R}$ converges to the deterministic one $x^*(k) \in \mathcal{R}$ in the mean square sense, then $H_\alpha(X(k)) \downarrow -\infty$. But the converse is not necessarily true. If $H_\alpha(X(k)) \downarrow -\infty$, this means only that the probability distribution of $X(k)$ tends to be concentrated in $\mathcal{R}$, in other words, the limiting value of $X(k)$ is a discrete random variable. It follows that, on a theoretical standpoint, asymptotic insensitiveness is not quite equivalent mean-square convergence.

Nevertheless, in a practical engineering framework, it is likely that the equivalence will hold since it is rather specific and unusual that continuous random variables converge to discrete random variables (while the converse is trivially common!).

This is the basic reason why we introduced the new term of "asymptotic insensitiveness" rather than to utilize the concept of "asymptotic stochastic stability."

6 Application to Stability Analysis

In this section, we shall examine how can the above results be of use to analyze the stability of sampled-data systems. For the sake of simplicity, we shall explain the main features the approach with an introductory example, and then the generalization will be straightforward.

6.1 An Introductory Example. *Asymptotic insensitiveness and stability.* We consider a two-dimensional system and we are interested in the study of its Lyapunov stability, L.S. in the following, w.r.t. the initial state. We assume that this system is time independent. By using a suitable change of coordinates, its equations are

$$x_1(k+1)=\lambda_1 x_1(k) \quad (6.1)$$

$$x_2(k+1)=\lambda_2 x_2(k) \quad (6.2)$$

and the nominal trajectory to be considered is $x^*(0)=0$. The system is A.I. in $\mathbb{R}^2$ provided that $|\lambda_1\lambda_2|<1$; in other words, one may have $|\lambda_1|<1$, and $|\lambda_2|>1$, and in such a case, the subsystem (6.1) is stable at 0 while the subsystem (6.2) is unstable at 0. The overall system is A.I. in $\mathbb{R}^2$ but it is unstable at 0. It will be stable if and only if we have $|\lambda_1|<1$ and $|\lambda_2|<1$.

Derivation of Stability Conditions. We now consider the system defined by the equation

$$x_1(k+1)=a_{11}(k)x_1(k)+a_{12}(k)x_2(k) \quad (6.3)$$

$$x_2(k+1)=a_{21}(k)x_1(k)+a_{22}(k)x_2(k). \quad (6.4)$$

If we assume that this system is A.I. in $\mathbb{R}^2$, then it may be stable; but it may be unstable as well. If we want to be sure that it is also stable, we must require that both $H_\alpha(X_1(k))$ and $H_\alpha(X_2(k))$ are uniformly decreasing, in other words that the subsystems S_1 and S_2 with the respective states x_1 and x_2, both are A.I.

A.I. conditions for S_1 are easy to obtain: it is sufficient to consider equation (6.3) as subject to a random disturbance on the initial state $x_1(0)$ and to the external disturbance $a_{12}\delta x_2(k)$. Likewise for S_2 via its equation (6.4). So according to propositions (4.1) and (4.2) it is sufficient to have

$$|a_{11}(k)|<1, |a_{22}(k)|<1 \quad (6.5)$$

$$|a_{11}(k)a_{12}(k)|<\exp(-H_\alpha(X_2(k)) \quad (6.6)$$

$$|a_{21}(k)a_{22}(k)|<\exp(-H_\alpha(X_1(k)) \quad (6.7)$$

for every $k\geq 0$. Conditions (6.6) and (6.7) are themselves satisfied provided that

$$|a_{11}(k)a_{12}(k)|<\exp(-H_\alpha(X_2(0)) \quad (6.8)$$

$$|a_{21}(k)a_{22}(k)|<\exp(-H_\alpha(X_1(0)) \quad (6.9)$$

so that conditions (6.5, 8, 9) are sufficient conditions for the stability of the system.

Indeed, if condition (6.8) holds, then one has

$$H_\alpha(X_1(1))<H_\alpha(X_1(0)),$$

therefore

$$\exp(-H_\alpha(X_1(0)))<\exp(-H_\alpha(X_1(1)))$$

so that (6.6) is satisfied for every $k>0$. Likewise (6.9) for (6.7).

Comparison With Other Results. For fixing the thought, and to facilitate the calculation, assume that

$$a_{11}=a_{22}=h, 0<h<1.$$

This being so, as $\text{var}\{X_1(0)\}$ and $\text{var}\{X_2(0)\}$ increase, their α-entropies also increase, so that the admissible ranges for $|a_{12}|$ and $|a_{21}|$ decrease, reducing to zero in the limiting case. In the opposite case, when these variances are zero, the α-entropies are $-\infty$, and a_{12} together with a_{21} may take on any finite values. In such a case the eigenvalues of the system are the solutions of the equation

$$\lambda^2-2h\lambda+(h^2-a_{12}a_{21})=0$$

therefore

$$\lambda_1, \lambda_2=h\pm(a_{12}a_{21})^{1/2}$$

It is clear that one of the λ_i's or both can be larger than unity, in other words, the system can be unstable in Lyapunov deterministic sense.

This somewhat surprising result can be explained as follows. For the one-dimensional system

$$\dot{x}=ax+bw,$$

A.I. is equivalent to L.S. only when $w=0$. For $w\neq 0$, A.I. should then be compared with input-output stability [8]. And it is exactly what happens in the above decomposition. For instance, in equation (6.3), $x_1(k)$ is the state of subsystem S_1 while $x_2(k)$ is considered as an input.

Another apparent discrepancy is the following. The A.I. conditions above, say equations (6.5, 8, 9) does not yield necessarily $|\det A|<1$ as mentioned in proposition 4.1. Again this is due to the fact that the A.I. considered in Section 5.1 is an A.I. conditional to the value of the disturbance on the initial state, while the A.I. in proposition 4.1 is absolute and independent of this disturbance. In this way, the requirement for the latter are obviously stronger than for the former.

6.2 Stochastic Stability. According to the above analysis, we shall introduce the following definition.

Definition 6.1. The discrete system

$$x(k+1)=A(k)x(k) \quad (6.10)$$

has the property of *stochastic conditional asymptotic stability* at zero, given random initial perturbations, whenever each one of its components is A.I.

We have the following result.

Proposition 6.1. The system (6.10) has the property of conditional asymptotic stochastic stabilty at zero provided that the following conditions are satisfied:

$$|a_{ii}(k)|<1, \forall i, \forall k\geq 0 \quad (6.11)$$

$$|a_{ii}(k)|\sum_{j\neq i}|a_{ij}(k)|<\exp\Big(-\sum_{j\neq i}H_\alpha(X_j(0))\Big), \forall k\geq 0. \quad (6.12)$$

Proof. The proof is a direct application of equation (4.17) by considering every subsystem S_i whose state is $x_i(k)$ and inputs are $x_j(k), j\neq i$.

Particular Case of Special Importance. Assume that the disturbances occur on some components of $X(0)$ only, say two components, for fixing the thought. Then we can partition x into two components ${}^1x$ and ${}^2x$ and the corresponding equation is (5.1). Then a condition for stochastic stability is

$$|\det A_{ii}(k)|<1, i=1,2$$

$$|\det A_{11}(k)A_{12}(k)|<\exp(-H_\alpha(X_2(0)))$$

$$|\det A_{21}(k)A_{22}(k)|<\exp(-H_\alpha(X_1(0))).$$

In Section 3, we have shown that in such a case, $X(k)$ converges to the state 0.

The generalization of this model to a larger number of disturbed components is direct.

7 Structural Perturbations

This section is devoted to analyzing the effects of disturbances which affect the structural matrices $A(k)$ and $B(k)$.

7.1 Disturbances on the Matrix *A(k)*.

Problem 7.1. Again we assume that the system (4.1) is designed to run on a trajectory $x^*(k)$ corresponding to the matrices $A^*(k)$ and $B^*(k)$. We assume that at the given instant p, the system is subject to a random disturbance which affects the matrix $A(k)$ in such a way that it becomes

$$\tilde{A}(p) := A^*(p) + \delta\tilde{A}(p) \tag{7.1}$$

with the coefficients

$$\tilde{a}_{ij}(p) := a^*_{ij}(p) + \delta\tilde{a}_{ij}(p); \tag{7.2}$$

and we further assume that this disturbance applies at the instant p only. How do we define the effects of this perturbation?

We have the following result.

Proposition 7.1. Assume that the disturbance applies in such a way that the random coefficients $\tilde{a}_{ij}(p)$ of the matrix $\tilde{A}(p)$ are mutually independent. Then the system (4.1) is asymptotically insensitive in $\mathcal{R}^n$ w.r.t. the nominal trajectory $x^*(k)$ corresponding to $A^*(k)$ and the disturbance, provided that the following conditions are satisfied.

$$E\{\tilde{A}(p)\} = A^*(p) \tag{7.3}$$

$$|x_1^*(p)x_2^*(p)\ldots x_n^*(p)\det A^*(p)| < \exp(-\sum_{ij} H_\alpha(\tilde{a}_{ij}(p))) \tag{7.4}$$

$$|\det A(p+j)| < 1, j=1,2,\ldots \tag{7.5}$$

Proof. (i) Condition (7.3) is the usual requirement on the mathematical expectation to yield $E\{X(p)\} = x^*(p)$.

(ii) At the instant $(p+1)$ one has

$$X(p+1) = A^*(p)x^*(p) + \delta\tilde{A}(p)x^*(p) + B^*(p)u^*(p)$$

where $x^*(p)$ is deterministic. The term $\delta\tilde{A}(p)x^*(p)$ so appears as a disturbance on the control, and according to proposition 4.2, the system is stable provided that

$$|\det A^*(p)| < \exp(-H_\alpha(\delta\tilde{A}(p)x^*(p))) \tag{7.6}$$

and there now remains to determine a lower bound for the right-hand side term of equation (7.6).

(iii) First, with our assumption about the mutual independence of the coefficients, we have

$$H_\alpha(\delta\tilde{A}(p)x^*(p))$$

$$= H_\alpha\Big(\sum_j \delta\tilde{a}_{1j}(p)\cdot x_j^*(p), \ldots, \sum_j \delta\tilde{a}_{nj}/(p)\cdot x_j^*(p)\Big)$$

$$= \sum_{i=1}^{n} H_\alpha\Big(\sum_{j=1}^{n} \delta\tilde{a}_{ij}(p)x_j^*(p)\Big).$$

Next, by using (2.10), one has

$$H_\alpha\Big(\sum_{j=1}^{n} \delta\tilde{a}_{ij}(p)\cdot x_j^*(p)\Big) \le \sum_{j=1}^{n} H_\alpha(\delta\tilde{a}_{ij}(p)\cdot x_j^*(p)\Big)$$

and by using (2.13),

$$H_\alpha(\delta\tilde{a}_{ij}(p)\cdot x_j^*(p)) = H_\alpha(\delta\tilde{a}_{ij}(p)) + \ln|x_j^*(p)|$$

or else, by virtue of (2.12) applied to (7.2),

$$H_\alpha(\delta\tilde{a}_{ij}(p)\cdot x_j^*(p)) = H_\alpha(\tilde{a}_{ij}(p)) + \ln|x_j^*(p)|.$$

After substitution, we finally get,

$$H_\alpha(\delta\tilde{A}(p)x^*(p)) \le \sum_{i,j=1}^{n} (H_\alpha(\tilde{a}_{ij}(p)) + \ln|x_j^*(p)|)$$

and using this upper bound in (7.6) yields condition (7.4).

(iv) Now, at the instant $(p+j+1)$, $j \ge 1$, we have

$$X(p+j+1) = A^*(p+j)X(p+j) + B(p+j)u^*(p)$$

and we are exactly in the framework of proposition 4.1

7.2 Disturbances on the Matrix *B(k)*.

Problem 7.2. We now assume that it is the matrix $B^*(p)$ which, at the instant p, is subject to a random disturbance, in such a way that it becomes

$$\tilde{B}(p) := B^*(p) + \delta\tilde{B}(p)$$

with the coefficients

$$\tilde{b}_{ij}(p) := b_{ij}^*(p) + \delta\tilde{b}_{ij}(p)$$

and we analyze the A.I. of the system subject to such a disturbance.

We have the following result.

Proposition 7.2. Assume that the disturbance applies in such a way that the random coefficients $\tilde{b}_{ij}(p)$ of the matrix $\tilde{B}(p)$ are mutually independent. Then the system (4.1) is asymptotically insensitive in R^n w.r.t. the nominal trajectory $x^*(k)$ corresponding to $B^*(k)$, and the disturbance, provided that the following conditions are satisfied:

$$E\{\tilde{B}(p)\} = B^*(p)$$

$$|u_1^*(p)u_2^*(p)\ldots u_n^*(p)\det B^*(p)| < \exp\Big(-\sum_{ij} H_\alpha(\tilde{b}_{ij})\Big)$$

$$|\det A(p+j)| < 1, j=1,2,\ldots$$

Proof. The proof is similar to the proof of proposition 7.1.

8 Analysis of Linear Continuous Systems

We are now interested in the asymptotic insensitiveness of linear continuous systems, since the related results may be of use in the analysis of sampled-data systems between the sampling instants.

8.1 Definition of the System. We now consider the continuous model associated with the discrete system above, which is

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{8.1}$$

$$x(0) = x_0$$

where $\dot{x}(t)$ holds for the derivative with respect to time, and where the definitions of the different variables are exactly like in (4.1) except they are continuous in time. We assume that the continuity and differentiability assumptions are satisfied for (8.1) to have a solution. We assume further that $A(t)$ and $\int_0^t A(\tau)\,d\tau$ commute for all t.

Definition 8.1. Assume that the random vector X_t is the result of a random disturbance R_t which acts against $x(t)$. We shall say that X_t has the property of A.I. w.r.t. the nominal deterministic value $x(t)$ and the disturbance R_t provided that the following conditions are satisfied:

$$E\{X_t\} = x(t), t \ge 0; \tag{8.2}$$

$$dH_\alpha(X_t)/dt < 0, \text{ for } t \text{ large enough} \tag{8.3}$$

8.2 Sensitivity w.r.t. Disturbances on the Initial State.

Problem 8.1. The system is designed to follow an optimal trajectory $x^*(t)$ defined by a given initial state $x_0 = x^*(0)$. We now assume that x_0 is subject to a random disturbance in such a way it becomes a random variable X_0, and the state X_t is random and defined by the equation

$$\dot{X}_t = A(t) X_t + B(t) u(t), \tag{8.4}$$

$$Pr\{x_0 \le X_0 \le x_0 + dx_0\} = p(x_0). \tag{8.5}$$

Under which condition is X_t A.I.?

Proposition 8.1. Assume that the system (8.1) is subject to a random disturbance on the initial state $x(0)$. Then the α-entropy $H_\alpha(X_t)$ of the future random state X_t is given by the equation

$$H_\alpha(X(t)) = H_\alpha(X_0) + \int_0^t tr\, A(\tau) d\tau \tag{8.6}$$

and the system is asymptotically insensitive in $\mathbb{R}^n$, w.r.t. the nominal trajectory $x^*(t)$ and the random disturbance, provided that the following conditions are satisfied for $t \ge 0$.

$$E\{X_0\} = x_0 \tag{8.7}$$

$$tr\, A(t) < 0, \tag{8.8}$$

where $tr\, A(t)$ denotes the trace of the matrix $A(t)$.

Proof. (i) Only X_0 is random, and for a given value to it, standard result yields

$$X(t) = (\exp\int_0^t A(\tau) d\tau) X_0 + \int_0^t \Phi(t,\tau) B(\tau) u(\tau) d\tau \tag{8.9}$$

with

$$\Phi(t,\tau) := \exp\int_\tau^t A(\Theta) d\Theta \tag{8.10}$$

(ii) This being so, by virtue of condition (8.7), one has

$$E\{X_t\} = x^*(t).$$

(iii) Next, the Jacobian of the transformation $X_0 \to X_t$ is $\exp \int_0^t A(\tau) d\tau$, and we repeat the calculation in the proof of proposition 4.1 to get

$$H_\alpha(X_t) = \ln|\det \exp\int_0^t A(\tau) d\tau| + H_\alpha(X_0).$$

This being so, a standard result in matrix calculus provides

$$\det \exp\int_0^t A(\tau) d\tau = \exp\int_0^t tr\, A(\tau) d\tau \tag{8.11}$$

therefore

$$H_\alpha(X_t) = \int_0^t tr\, A(\tau) d\tau + H_\alpha(X_0), \tag{8.12}$$

and the proof is completed.

8.3 Sensitivity w.r.t. Disturbances on the Control

Problem 8.2. The counterpart of problem 4.2, is the model

$$\dot{X}_t = A(t) X_t + B(t) u(t) + C(t) W_t \tag{8.13}$$

in which $W_t \in \mathbb{R}^n$ denotes a random noise, and $C(t)$ is an (n, n)-matrix. We look for the A.I. of X_t w.r.t. to the nominal reference trajectory $x^*(t)$ in the absence of the noise W_t.

Let $H_\alpha(W_t)$ denote the α-entropy of W_t; we state

Proposition 8.2. The system (8.13) is asymptotically insensitive w.r.t. the nominal trajectory $x^*(t)$ and the random noise $C(t) W_t$ provided that the following conditions are satisfied.

$$E\{W_t\} = 0 \tag{8.14}$$

$$\ln|\det C(t)| + t\, Tr\, A(t) + H_\alpha(W_t) < 0 \tag{8.15}$$

for every $t \ge 0$.

Proof. (i) Condition (8.14) applied to equation (8.13) yields $E\{X_t\} = x^*(t)$.

(ii) The solution for equation (8.13) is (in the engineering framework)

$$X_t = (\exp\int_0^t A(\tau) d\tau) X_0 + \int_0^t \Phi(t,\tau) B(\tau) u(\tau) d\tau + \int_0^t \Phi(t,\tau) C(\tau) W_\tau d\tau; \tag{8.16}$$

and equation (2.12) yields

$$H_\alpha(X_t) = H_\alpha\left(\int_0^t \Phi(t,\tau) C(\tau) W_\tau d\tau\right) \tag{8.17}$$

and by using the property (2.11) for the α-entropy, we get

$$H_\alpha(X_t) \le \int_0^t H_\alpha(\Phi(t,\tau) C(\tau) W_\tau) d\tau. \tag{8.18}$$

(iii) Next, like in the step (iii) in the proof of proposition 4.1, we have

$$H_\alpha(\Phi(t,\tau) C(\tau) W_\tau) = \ln|\det \Phi(t,\tau) C(\tau)| + H_\alpha(W_\tau)$$

and by virtue of the definition (8.10) for $\Phi(t, \tau)$ together with the property (8.11) for matrix functions

$$H_\alpha(\Phi(t,\tau) C(\tau) W_\tau) = \int_\tau^t tr\, A(\Theta) d\Theta + \ln|\det C(\tau)| + H_\alpha(W_\tau) \tag{8.19}$$

Now substituting into equation (8.18), we get

$$H_\alpha(X_t) \le \int_0^t \left(\left(\int_\tau^t tr\, A(\Theta) d\Theta\right) + \ln|\det C(\tau)| + H_\alpha(W_\tau)\right) d\tau \tag{8.20}$$

When the right-hand side term is decreasing as t increases, it is so for the left-hand side one. It is then sufficient to write that the derivative of the right-hand side term is negative to get the result.

In this framework, the problem of analyzing the asymptotic insensitiveness and the stochastic stability of continuous systems can be conducted like for sampled data systems.

8.4 Possible Extensions of the Results. The above results are stated under the condition that the matrices $A(t)$ and $\int_0^t A(\tau)\, d\tau$ commute, which is the case for instance when $A(t)$ is constant. This assumption is required because the exponential matrix $\exp(E(t))$ involves the powers $E^k(t)$, $k = 0, 1, \ldots$ and the derivative of $E^k(t)$ equates $k\dot{E}E^{k-1}$ only when $\dot{E}$ commutes with E.

Of course, this assumption is rather restrictive. Nevertheless one can expect an extension of the result by using the technique of dominant systems as follows.

Consider the vector system

$$\dot{x} = A(x,t) x, x \in \mathbb{R}^n \tag{8.21}$$

Let $E_1, E_2, \ldots, E_k$ denote a partition of R^n, and define x_i as the projection of x in E_i. Let ψ_i denote a scalar norm defined on E_i; then it infers a k-dimensional vector norm $v(x)$ to x such that

$$v_i(x) = \psi_i(x_i) \tag{8.22}$$

Assume now that we can define a matrix $M(A(x, t))$ such that $v(x)$ satisfies the variational inequality

$$\dot{v}(x) \le M(A(x,t) . v(x) \tag{8.23}$$

If further $M(A(.))$ is constant, say M, then one can analyze the dynamical system

$$\dot{y} = M.y$$

to get results on the stability of the original system (8.21).

9 Nonlinear Discrete Systems

9.1 Main Results. The system we now consider is defined by the equation

$$x(k+1) = f(x(k)) + B(k) u(k) \tag{9.1}$$

where x, B and u are those of equation (4.1) and where $f := (f_1, f_2, \ldots f_n)^T$ is a continuous nonlinear $\mathcal{R}^n \to \mathcal{R}^n$ mapping.

We have the following result.

Proposition 9.1. Assume that the system (9.1) is subject to a random disturbance on the state $x(p)$ at the instant p. Then it is asymptotically insensitive in $\mathcal{R}^n$ w.r.t. the nominal trajectory $x^*(k)$ and the random disturbance, provided that the following conditions are satisfied for every $m \geq 0$

$$E\{f(X(p))\} = f(x(p)) \tag{9.2}$$

$$|D(f)/D(x)|_{p+m} \leq M < 1 \tag{9.3}$$

Proof. Condition (9.2) is the usual condition for the statistical average. This being so, after the disturbance occurs, the equation is

$$X(p+m+1) = f(X(p+m)) + B(p+m)u(p+m)$$

therefore

$$H_\alpha(X(p+m+1)) = H_\alpha(f(X(p+m))).$$

By using equation (2.13), we then have

$$H_\alpha(X(p+m+1))$$
$$= \frac{1}{1-\alpha} \ln \int q^\alpha(x(p+m)) \left| \frac{D(f)}{D(x)} \right|_{p+m} dx(p+m). \tag{9.4}$$

Next assume that equation (9.3) is satisfied; then (9.4) provides

$$H_\alpha(X(p+m+1)) = H_\alpha(X(p+)) + \ln M$$

and

$$H_\alpha(X(p+m+1)) - H_\alpha(X(p+m)) < 0$$

therefore the result.

9.2 Practical Consequence. The practical consequence of this result is that we can consider the linearized system associated with (9.1) and apply the preceding result. Indeed, if we define $X(p+m)$ as

$$X(p+m) := x^*(p+m) + \xi(p+m) \tag{9.5}$$

then the equation of the linearized system is

$$\xi(p+m+1) = (D(f)/D(x))_{p+m} \cdot \xi(p+m) \tag{9.6}$$

and condition (9.3) is exactly a condition for its asymptotic insensitiveness.

An Illustrative Example. Consider the second-order system

$$S_1 : x_1(n+1) = x_2(n) \tag{9.7}$$

$$S_2 : x_2(n+1) = f_1(x(n))x_1(n) + f_2(x(n))x_2(n) \tag{9.8}$$

with

$$f_1(0) = f_2(0) = 0$$

and examine the asymptotic stability at zero with regard to perturbations at the instant zero. According to the criteria above, the subsystem S_1 is stable, as one has

$$0 < \exp(-H_\alpha(X_2(0))),$$

and the substem S_2 is stable provided that

$$|Df_2(x)/D(X)|_n < 1$$

and

$$\left| \frac{Df_1(x)}{D(x)} \right|_n \left| \frac{Df_2(x)}{D(x)} \right|_n < \exp(-H_\alpha(X_1(0)))$$

for $n \geq 0$.

10 Conclusions

In the present paper, we have proposed to utilize information theory, and more explicitly the so-called entropy of a random variable to analyze the sensitivity and the stability of multivariable discrete systems in a stochastic framework. The main result we so derived can be summarized as follows. We obtained the concept of asymptotic stability with respect to random disturbances on the ready state, and the concept of conditional asymptotic stability with respect to external disturbances in the form of additional random inputs, and general criteria for this kind of stability are stated. As a result, a multivariable system can be viewed as a set of interconnected subsystems, each one subject to external inputs which are the states of the other systems, and the stability criteria above directly apply. Broadly speaking it looks like if we had identified a somewhat relation or connection between stability in Lyapunov sense and input-output stability, but it is important to understand that, in the stochastic framework, this identification has been possible only because we used the entropy. This approach seems to be of interesting prospects in the study of large-scale systems.

The comments above illustrate by themselves the differences between the results by Siljak (reference [1]) and the present results. The former basically utilizes the technique of Lyapunov functions to analyze global asymptotic stability in the mean, while the latter deals with a somewhat conditional asymptotic stability which is exactly defined by means of the entropy. In this way, the latter is more direct than the former in the applications, and both ones can be thought of as being complementary.

Another interesting problem would be to introduce the effect of quantization in this information theoretic approach. In such a case, the entropy is in the form of a finite summation, it tends to zero with the variance, and the results should be slightly different. We intend to work this question.

References

1 Siljak, D.D., *Large-Scale Dynamic Systems*, North-Holland, New York, 1978.

2 Kushner, H.J., *Stochastic Stability and Control*, Academic Press, New York, 1967.

3 Siljak, D.D., "Stability of Large-Scale Systems Under Structural Perturbations," *IEEE Transactions*, SMC-2, 1972, pp. 657-663.

4 Shannon, C.E., "A Mathematical Theory of Communication," *Bell System Tech. J.*, Vol. 27, 1948, pp. 379-423.

5 Shannon, C.E., "A Mathematical Theory of Communication," *Bell System Tech. J.*, Vol. 27, 1948, pp. 623-656.

6 Aczel, J., and Daroczy, A., *On Measures of Information and Their Characterizations*, Academic Press, New York, 1975.

7 Jumarie, G., "Relativistic Information Theory III. On the Use and the Practical Meaning of the Renyi Entropy in Encoding Problems" (In French), *Annales des Télécommunications*, Vol. 34, Nos 11-12, 1979, pp. 521-530.

8 Zames, G., "On the Input-Output Stabilty of Time Varying Nonlinear Feedback Systems-Part II. Conditions Involving Circles in the Frequency Plane and Sector Nonlinearities," *IEEE Transactions*, AC-11, No. 3, 1966, pp. 465-476.