ON THE ATTAINABILITY FUNCTION
OF STOCHASTIC SYSTEMS

by

Nicolas J. Spyratos

Submitted to the Department of Electrical Engineering, University of Ottawa, in partial fulfillment of the requirements for the degree of 'Master of Applied Sciences'.

Department of Electrical Engineering
Faculty of Science and Engineering
University of Ottawa
Ottawa, Canada
December 1971

ABSTRACT

In this work, the concept of the attainability function F(Roxin, pp. 189-190) is extended in the case of stochastic systems. In particular, the axioms for this function are presented and its basic properties, useful in the theory of optimal control, are investigated. Since these properties are considered with respect to a certain metric on the domain of F, a discussion on the properties of this metric is included. Finally, the time optimal control problem is formulated, using the concept of the attainability function. The possibility of applications of the present theory to practical problems is also discussed.

It should be noted that although one of the two basic assumptions in Roxin's work, namely, that of completeness of the metric on the state space of the system, has been removed, the properties of the function F are proved to be valid. Moreover, the properties described in theorems 2 and 3 are more general.

ACKNOWLEDGMENTS

I am very grateful to my advisor, Professor N.U. Ahmed for his encouragement, understanding and guidance.

I am also grateful to Professor G. Glinski for his initial help and guidance.

Thanks are also due to the members of the staff and the graduate students of the Electrical Engineering Department, for their friendship throughout the period of this work.

The financial assistance of the Government of Ontario, through an Ontario Graduate Fellowship is gratefully acknowledged.

TABLE OF CONTENTS

Page

1. Introduction 1

2. The Axioms of the Attainability Function 4

3. The Properties of the Attainability Function 8

4. The Metric Space $(\mathcal{H},\rho)$ 23

5. Examples 26

6. Conclusion 30

1. INTRODUCTION

Given the initial condition (x_o,t_o) of a system, it is not always possible to determine uniquely the state x of the system at some time $t \geq t_o$. Depending on the factors affecting the system (e.g., control action), we can only talk about the set A(t) of the states which can be attained by the system at time t, otherwise called the "set of attainable states".

Let F be a set-valued function which satisfies certain axioms and which assigns to each triple (x_o,t_o,t) the corresponding set A(t). It is natural to call F the 'attainability function' of the system.

In order to extend the concept of the attainability function to stochastic systems, we note that the state x_t of the system at time $t \geq t_o$ is now a random variable and therefore, each motion of the system is a stochastic process. Then, instead of the initial state of the deterministic system, the initial probability distribution function should be considered (that is, the probability distribution function of the random variable x_{t_o}), while A(t) is now the set of probability distribution functions of the random states x that can be attained by the system at time t.

As an example consider the Itô stochastic-differential equation:

$$dx(t) = F(t,x(t),u(t))dt + G(t,x(t),w(t),u(t))\, dw(t)$$

where u is a control function from a set U of bounded measurable functions. Under certain conditions, the above equation yields a unique probability density function $p^u(x,t/x_o,t_o)$ for each u. Given then the probability density function of the random variable x_o, one can compute the function $p^u(x_t)$ and thus determine the set of attainable probability density functions

$$A(t) = \{ p^u(x_t) \ / \ u \varepsilon U\}$$

This procedure is made clear in an example given in the last section. Observe that as far as we are in a position to determine the set A(t), the attainability function of the system is defined, no matter if the system can be described by a set of differential equations or not. This is exactly one of the advantages of the present approach to problems concerning dynamical systems.

Now, it is clear that one can consider the probability measures μ_t^u instead of the corresponding probability density functions $p^u(x_t)$, in which case A(t) becomes the set of attainable probability measures for the system at time $t \geq t_o$. It is in the last form that A(t) will be used in the definition of the attainability function for a stochastic system. Moreover, certain axioms will be assumed for the function F. These axioms are in fact common properties of large classes of systems which motivate the introduction of the attainability function in an axiomatic way. An example of these properties can be found in the time-optimal control problem (La salle, p. 107, theorem 20.1).

<u>NOTE</u>: In what follows, whenever we write 'measure' we shall understand 'probability measure'.

2. THE AXIOMS OF THE ATTAINABILITY FUNCTION

In this section, the concept of the attainability function F for a stochastic system is defined. This is done in definition 5. Certain concepts, which are needed for the definition of F and later on when discussing its properties, are defined in definitions 1-4.

As it was pointed out in the introduction, the attainability function F is a set-valued function whose range is contained in $P(\mathcal{M})$, the family of all subsets of a space of measures which is defined below.

<u>Definition 1</u> Let M be the state space of the system, A_M a σ-algebra on M and R^+ the set of non-negative real numbers. Define,

$$\mathcal{M} = \{\mu : A_M \rightarrow R^+ \;/\; \mu \text{ is a probability measure on } (M, A_M)\}$$

The concept of continuity will be used for F. This means that topologies should be defined on both the domain and the range of F.

The domain of F is the set

$$Y = \{ (\mu_o, t_o, t) \in \mathcal{M} \otimes R \otimes R \;/\; t \geq t_o \}$$

where R is the set of real numbers.

It will be assumed that Y is equipped with the topology induced by the product topology on $\mathcal{M} \otimes R \otimes R$, with the usual metric considered on R and a metric ρ on $\mathcal{M}$. The nature of the metric ρ will be specified in the next section.

The range of F being the family $K(\mathcal{H})$ of all non-empty closed subsets of the metric space $(\mathcal{H},\rho)$ (this will be required in one of the axioms F should satisfy), is assumed equipped with the Hausdorff metric. In order to define this metric we need the concept of separation (or divergence) of one set from another.

<u>Definition 2</u> Let A and B by any two subsets of a metric space (X,ρ_X). The 'separation' of A from B is defined by

$$s(A,B) = \sup\{d(a,B) \,/\, a \in A\}$$

where $d(a,B)$ denotes the distance of a from B in the metric ρ, that is, $d(a,B) = \inf\{\rho(a,b)/b \in B\}$.

<u>NOTE</u>: The separation of A from B is always considered with respect to the metric on the set whose A and B are subsets. However, whenever confusion might arise, an index is used.

It is easy to see that $s(A,B)$ is not necessarily equal to $s(B,A)$ as the following example shows. Let the metric space (X,ρ_X) be the set of points in R^2 with the usual metric and let its subsets A and B be defined by

$$A = \{ (x,y) \,/\, x = o,\ y > o\} \cup \{(x,y) \,/\, x > o,\ y = o\}$$

$$B = \{ (x,y) \,/\, x > o,\ y = {}^1/x \}$$

Then, $s(A,B) = \sqrt{2}$, while $s(B,A) = 1$. Therefore, s cannot be considered as a metric on P(X). However, using s a metric ρ_H can be defined as follows:

Definition 3 Let K(X) be the family of all non-empty closed subsets of the metric space (X,ρ_X). Then the function

$$\rho_H : K(X) \times K(X) \to R \text{ such that}$$

$$\rho_H (A,B) = \max \{s(A,B), s(B,A)\} ,$$

(which is easily, verified to be a metric on K(X)) is called the 'Hausdorff metric'.

NOTE: Observe that the metric ρ_H restricted to the family of compact elements of K(X) is finite.

Using now the concept of separation the concept of upper quasicontinuity is defined.

Definition 4 Let (W,ρ_W) and (Z,ρ_Z), be two metric spaces and let P(W) and P(Z) denote the families of all subsets of W and Z, respectively. Then, a function $G : W \to P(Z)$ is called 'upper quasicontinuous' at $w_o \varepsilon W$ if for every $\varepsilon > o$ there exists a $\delta > o$ such that $s(\{w\},\{w_o\}) < \delta$ implies that $s(G(w), G(w_o)) < \varepsilon$. Moreover, the function on P(W) defined by $G(A) = \bigcup \{G(w)/w \varepsilon A\}$ is called upper quasicontinuous at $A_o \subset W$, if for every $\varepsilon > o$ there exists a $\delta > o$ such that $s(A,A_o) < \delta$ implies that $s(G(A), G(A_o)) < \varepsilon$.

NOTE: G is called simply upper quasicontinuous if it is upper quasicontinuous at every $w_0 \varepsilon W$.

In defining below the attainability function and studying later on its properties, we shall often be writing $F(\mu_0,t_0,t)$ instead of $F((\mu_0,t_0,t))$.

Definition 5 Let $Y = \{(\mu_0,t_0,t) \varepsilon \mathcal{M} xRxR \,/\, t \geq t_0\}$. Then, a function $F : Y \to P(\mathcal{M})$ is called the 'attainability function' of the system if it satisfies the following axioms.

(i) $F(\mu_0,t_0,t)$, the set of attainable measures for the system at time t, is a closed subset of the metric space $(\mathcal{M},\rho)$, for any $\mu_0 \varepsilon \mathcal{M}$ and $t_0, t \; \varepsilon R$ such that $t_0 \leq t$.

(ii) $F(\mu_0,t_0,t_0) = \{\mu_0\}$

(iii) For every $\mu \varepsilon \mathcal{M}$ and $t_0, t \; \varepsilon \; R$ such that $t_0 \leq t$ there exists a $\mu_0 \varepsilon \mathcal{M}$ such that $\mu \; \varepsilon \; F(\mu_0,t_0,t)$.

(iv) $F(\mu_0,t_0,t_2) = \bigcup \{F(\mu,t_1,t_2) \,/\, \mu \; \varepsilon \; F(\mu_0,t_0,t_1)\}$, for any $t_1, t_2 \; \varepsilon \; R$ such that $t_1 \leq t_2$.

(v) F is continuous in t, in the Hausdorff metric.

(vi) F is upper quasicontinuous in (μ_0,t_0), for fixed t.

REMARK Since $s(F(\mu_0', t_0', t'), F(\mu_0,t_0,t)) \leq s(F(\mu_0',t_0',t'), F(\mu_0,t_0,t') + \rho_H(F(\mu_0,t_0,t'), F(\mu_0,t_0,t))$, it follows from axioms (v) and (vi) that F is upper quasicontinuous.

NOTE: Axiom (iii), which is assumed in Roxin's work, is not used in the proof of the theorems to follow and, therefore, it can be omitted.

3. THE PROPERTIES OF THE ATTAINABILITY FUNCTION

In this section, the properties of the attainability function are investigated. This investigation is based on the assumption that the metric ρ on the set $\mathcal{M}$ is such that the metric space $(\mathcal{M},\rho)$ is locally compact. Certain results concerning this metric space are presented in the next section

In control problems it is very important to know whether the set of attainable measures $F(\mu_o,t_o,t)$ is compact. This question is treated in theorem 1 below. In order to prove this theorem we need the following definition and lemmas 1-4.

Definition 6 Let (X,ρ_X) be a metric space and $A \subset X$. For every $\varepsilon>o$, define

$$S_\varepsilon(A) = \{x\varepsilon X \;/\; d(x,A) <\varepsilon\}$$

NOTE It should be clear that the set $S_\varepsilon(A)$ is always considered with respect to the metric on the set whose A is a subset. Moreover, observe that $S_\varepsilon(A)$ is an open set containing A. It is for this reason that it is often referred to as an ε-neighborhood of A.

Lemma 1 Let (X,ρ_X) be a metric space, $A \subset X$ a compact set and $E \subset X$ an open set containing A. Then, there exists $\varepsilon>o$ such that $S_\varepsilon(A) \subset E$.

Proof Let E^c denote the complement of the set E and consider the distance $d(a,E^c)$ as a function of a on the compact set A. Since the distance is a continuous function, it attains its infimum on A at some point $a_0 \varepsilon A$. Since $a_0 \notin E^c$ and E^c is a closed set, it follows that $d(a_0,E^c) = \varepsilon > o$. Then, for every $x \varepsilon X$ such that $d(x,A) < \varepsilon$, we have that $x \notin E^c$ or, equivalently, that $x \varepsilon E$. Therefore, $S_\varepsilon(A) \subset E$, Q.E.D.

Lemma 2 Let (X,ρ_X) be a locally compact metric space and $A \subset X$ be any compact set. Then, there exists $\varepsilon > o$ such that $\overline{S_\varepsilon(A)}$ is compact.

($\overline{S_\varepsilon(A)}$ denotes the closure of the set $S_\varepsilon(A)$).

Proof Since (X,ρ_X) is a locally compact metric space, for every $x \varepsilon X$ there exists an open neighborhood N_x of x such that $\overline{N_x}$ is compact. On the other hand, since the family of sets $\{N_x / x \varepsilon A\}$ is clearly an open cover of the compact set A, we can find a finite number of points $x_1, ---, x_n \varepsilon A$ such that $A \subset N_{x_1} \cup --- \cup N_{x_n}$. Since the set $N_{x_1} \cup --- \cup N_{x_n}$ is open, it follows from lemma 1, that there exists some $\varepsilon > o$ such that $S_\varepsilon(A) \subset \overline{N}_{x_1} \cup --- \cup \overline{N}_{x_n}$. This last relation implies that $\overline{S_\varepsilon(A)} \subset \overline{N}_{x_1} \cup --- \cup \overline{N}_{x_n}$. Therefore $\overline{S_\varepsilon(A)}$ is compact as a closed subset of the set $\overline{N}_{x_1} \cup --- \cup \overline{N}_{x_n}$ which is compact, as finite union of compact sets.

Lemma 3 Let (X,ρ_X) be a metric space and $A,B \subset X$. Then $s(A,B) < \varepsilon$ implies that $A \subset S_\varepsilon(B)$.

Proof Since $s(A,B) = \sup \{d(a,B) \,/\, a\varepsilon A\}<\varepsilon$, it follows that for every $a\varepsilon A$, we have $d(a,B)<\varepsilon$. Therefore, $a \,\varepsilon\{x\varepsilon X \,/\, d(x,B)<\varepsilon\}= S_\varepsilon(B)$ which implies $A\subset S_\varepsilon(B)$.

Lemma 4 Let (X,ρ_X) be a metric space, $A \subset X$ a closed set and $\{A_n\}$ a sequence of closed non-empty sets converging to the closed non-empty set A_o, in the Hausdorff metric. Then, if A_n is a subset of A, for every $n = 1,2,\text{---}$ so is A_o.

Proof Suppose that A_o is not a subset of A. Then there exists a point $x \,\varepsilon\, A_o$ such that $x \notin A$. Since A is closed, it follows that $d(x,A) = \varepsilon>o$. On the other hand, since $A_n \to A_o$, there exists an integer $n_o = n_o(\varepsilon)$ such that $\rho_H(A_n,A_o)<\varepsilon/2$ for every $n>n_o$. This implies that $s(A_o,A_n)< \varepsilon/2$ for every $n>n_o$. It follows then from lemma 3, that $A_o \subset S_{\varepsilon/2}(A_n)$, for every $n>n_o$. Since each A_n is a subset of A, this last relation implies that $A_o\subset S_{\varepsilon/2}(A)$. This is impossible because $x\varepsilon A_o$ and $d(x,A) = \varepsilon>\varepsilon/2$ and the proof is completed.

Theorem 1 The set $F(\mu_o,t_o,t)$ is a compact subset of the metric space $(\mathcal{M},\rho)$, for every $\mu_o\varepsilon\mathcal{M}$ and $t_o,t\varepsilon R$ such that $t_o\leq t$.

Proof Since the space $(\mathcal{M},\rho)$ is locally compact, for every $\mu_o\varepsilon\mathcal{M}$ there exists an ε_1-neighborhood $S_{\varepsilon_1}(\{\mu_o\})$ such that $\overline{S_{\varepsilon_1}(\{\mu_o\})}$ is compact. By definition 5(ii), $F(\mu_o,t_o,t_o) = \{\mu_o\}$ and therefore for every $\mu_o\varepsilon\mathcal{M}$, there exists $\varepsilon_1>o$ such that

$\overline{S_{\varepsilon_1}(F(\mu_0,t_0,t_0))}$ is compact. On the other hand, upper quasicontinuity of F implies that, for the ε_1 considered above, there exists a $\delta_1 > o$ such that $s(F(\mu_0,t_0,t),$ $F(\mu_0,t_0,t_0)) < \varepsilon_1$ for $|t-t_0| < \delta_1$. It follows then from lemma 3 that $F(\mu_0,t_0,t) \subset S_{\varepsilon_1}(F(\mu_0,t_0,t_0)) \subset \overline{S_{\varepsilon_1}(F(\mu_0,t_0,t_0))}$ for $|t-t_0| < \delta_1$. Therefore, $F(\mu_0,t_0,t)$ being a closed subset of a compact set is compact, for every $t \varepsilon [t_0, t_0+\delta_1)$. We show next that $F(\mu_0,t_0,t_0+\delta_1)$ is also compact. To do this, consider a sequence $\{t_n\}$ in the interval $[t_0,t_0+\delta_1)$ converging to $t_0+\delta_1$. It follows then from continuity of $F(\mu_0,t_0,\cdot)$ that the sequence of non-empty closed sets $F(\mu_0,t_0,t_n)$ converges to the non-empty closed set $F(\mu_0,t_0,t_0+\delta_1)$, while each set in the sequence is a subset of the non-empty closed set $\overline{S_{\varepsilon_1}(F(\mu_0,t_0,t_0))}$. It follows then from lemma 4, that the closed set $F(\mu_0,t_0,t_0+\delta_1)$ is a subset of the compact set $\overline{S_{\varepsilon_1}(F(\mu_0,t_0,t_0))}$ and therefore it is itself compact. Therefore, for any $\mu_0 \varepsilon \ell e$ and $t_0 \varepsilon R$ there exists $\delta_1 > o$ such that the theorem is true for any $t \varepsilon [t_0,t_0+\delta_1]$. It is then clear that if the above procedure is followed again, for $\{\mu_0\} = F(\mu_0,t_0+\delta_1, t_0+\delta_1)$, a $\delta_2 > o$ can be found, such that for any $\mu_0 \varepsilon \ell e$ and $t_0 \varepsilon R$ the theorem is true for any $t \varepsilon [t_0+\delta_1, t_0+\delta_1+\delta_2]$. This implies that given any $t \geq t_0$ there exists an integer n and a corresponding δ_n such that the theorem is true for any $\mu_0 \varepsilon \ell e$, $t_0 \varepsilon R$ and $t \varepsilon [t_0, t_0+\delta_1+\cdots+\delta_n]$ and this completes the proof.

A compactness property is shown now in theorem 2. In order to prove this theorem, we need an equivalent definition of upper quasicontinuity. This definition is given in lemma 5.

Lemma 5 Let (W,ρ_W) and (Z,ρ_Z) be two metric spaces and consider a function $G : W \to P(Z)$. Then the function defined by $G(A) = \bigcup\{G(w) / w\varepsilon A\}$ is upper quasicontinuous at $A_o \subset W$, if and only if, for every $S_\varepsilon(G(A_o))$ there exists $S_\delta(A_o)$ such that $G(S_\delta(A_o)) \subset S_\varepsilon(G(A_o))$.

Proof Suppose first that G is upper quasicontinuous at $A_o \subset W$ and consider any $S_\varepsilon(G(A_o))$. It follows from upper quasicontinuity of G at A_o, that for this particular $\varepsilon > o$ there exists a $\delta' > o$ such that $s(A,A_o) < \delta'$ implies $s(G(A), G(A_o)) < \varepsilon$. Considering now any $\delta > o$ such that $\delta < \delta'$, we have that $s(S_\delta(A_o),A_o) \leq \delta < \delta'$. Then upper quasicontinuity of G implies that $s(G(S_\delta(A_o)), G(A_o)) < \varepsilon$ and it follows from lemma 3 that $G(S_\delta(A_o)) \subset S_\varepsilon(G(A_o))$, Q.E.D.

We show next the "if" part. We have to show that for any $\varepsilon' > o$ there exists $\delta > o$ such that $s(A,A_o) < \delta$ implies $s(G(A),G(A_o)) < \varepsilon'$. For the given ε' consider an ε such that $o < \varepsilon < \varepsilon'$. Then there exists a $\delta > o$ such that $G(S_\delta(A_o)) \subset S_\varepsilon(G(A_o))$. Now, $s(A,A_o) < \delta$ implies that $A \subset S_\delta(A_o)$ and therefore, $G(A) \subset G(S_\delta(A_o))$. It follows then that $G(A) \subset S_\varepsilon(G(A_o))$ and this implies that for every $w\varepsilon G(A)$ we have $d(w, G(A_o)) < \varepsilon$. Therefore, $s(G(A), G(A_o)) \leq \varepsilon < \varepsilon'$, Q.E.D.

Theorem 2 Let L be a non-empty compact subset of the domain of the attainability function F and let $x = (\mu_0, t_0, t)$. Then

$$F(L) = \bigcup\{F(x) \;/\; x \varepsilon L\}$$

is upper quasicontinuous at L and the set F(L) is a compact subset of the metric space $(\mathcal{M}, \rho)$.

NOTE: As an example of such a set L consider the case

$$L = A \times [t_0, t_1] \times [t_1, t_2]$$

where A is a non-empty compact subset of $(\mathcal{M}, \rho)$.

Proof It follows from lemma 5, that in order to show upper quasicontinuity it is enough to show that for every $S_\varepsilon(F(L))$ there exists an $S_\delta(L)$ such that $F(S_\delta(L)) \subset S_\varepsilon(F(L))$. Upper quasicontinuity of F (definition 5, remark) implies that for every $S_\varepsilon(F(x))$ there exists an $S_{\delta'}(x)$ such that $F(S_{\delta'}(x)) \subset S_\varepsilon(F(x))$. Now, the metric space $(\mathcal{M}, \rho)$ is locally compact and since its compact subset L is clearly a subset of the open set $V = \bigcup\{S_{\delta'}(x) \;/\; x \varepsilon L\}$, it follows from lemma 3 that there exists a $\delta > o$ such that $S_\delta(L) \subset V$. This implies that

$$\begin{aligned} F(S_\delta(L)) \subset F(V) &= F(\bigcup\{S_{\delta'}(x) \;/\; x \varepsilon L\}) \\ &= \bigcup\{ F(S_{\delta'}(x)) \;/\; x \varepsilon L\} \\ &\subset \bigcup\{S_\varepsilon(F(x)) \;/\; x \varepsilon L\} \end{aligned}$$

and since each $S_\varepsilon(F(x))$ is clearly a subset of $S_\varepsilon(F(L))$, we obtain that $F(S_\delta(L)) \subset S_\varepsilon(F(L))$, Q.E.D.

Next, we show that the set $F(L)$ is a compact subset of the metric space $(\mathcal{L},\rho)$. Observe first that since, by theorem 1, $F(x)$ is a compact subset of the locally compact metric space $(\mathcal{L},\rho)$ for every $x \varepsilon L$, it follows from lemma 2, that for every $x \varepsilon L$, there exists an $\varepsilon = \varepsilon(x)$ such that $\overline{S_\varepsilon(F(x))}$ is compact. It follows then from upper quasicontinuity of F and lemma 5 that, for this particular $\varepsilon = \varepsilon(x) > o$, there exists a $\delta = \delta(x) > o$ such that $F(S_\delta(x)) \subset S_\varepsilon(F(x))$. Now, since the family of sets $\{S_\delta(x) \,/\, x \varepsilon L\}$ is clearly an open cover of the compact set L, we can find a finite number of points $x_1, ---, x_n \varepsilon L$, such that: $L \subset S_{\delta_1}(x_1) \cup --- \cup S_{\delta_n}(x_n)$. This implies that

$$\begin{aligned} F(L) &\subset F(S_{\delta_1}(x_1) \cup --- \cup S_{\delta_n}(x_n)) \\ &= F(S_{\delta_1}(x_1)) \cup --- \cup F(S_{\delta_n}(x_n)) \\ &\subset S_{\varepsilon_1}(F(x_1)) \cup --- \cup S_{\varepsilon_n}(F(x_n)) \\ &\subset \overline{S_{\varepsilon_1}(F(x_1))} \cup --- \cup \overline{S_{\varepsilon_n}(F(x_n))} \end{aligned}$$

and therefore $F(L)$ is a subset of a compact set (any finite union of compact sets is a compact set). Then, in order to show that $F(L)$ is compact, it is enough to show that $F(L)$ is closed. To do this, we suppose that a sequence $\{\mu_n\}$ in $F(L)$ converges to some point $\mu \varepsilon \mathcal{L}$ and we show that $\mu \varepsilon F(L)$.

Observe that for each $\mu_n \varepsilon F(L)$ there exists $x_n = (\mu_{on}, t_{on}, t_n) \varepsilon L$ such that $\mu_n \varepsilon F(x_n)$ and since L is compact there exists a subsequence $\{x_{n_i}\}$ converging to some point $x \varepsilon L$. It should be then enough to show that $\mu \varepsilon F(x)$. Since F is upper quasicontinuous (definition 5, remark) it follows from lemma 5 that for every $S_\varepsilon(F(x))$ there exists $S_\delta(x)$ such that $F(S_\delta(x)) \subset S_\varepsilon(F(x))$. On the other hand, since $S_\delta(x)$ is a neighborhood of x, there exists an integer $m = m(S_\delta(x))$ such that $x_{n_i} \varepsilon S_\delta(x)$ for every $n_i > m$. This implies that $\mu_{n_i} \varepsilon F(S_\delta(x)) \subset S_\varepsilon(F(x))$ and therefore $d(\mu_{n_i}, F(x)) < \varepsilon$ for every $n_i > m$. Since F(x) is compact, this last relation implies that for every μ_{n_i} there exists a $\nu_{n_i} \varepsilon F(x)$ such that $\rho(\mu_{n_i}, \nu_{n_i}) < \varepsilon$ for every $n_i > m$. Then, the relation $\rho(\nu_{n_i}, \mu) \leq \rho(\nu_{n_i}, \mu_{n_i}) + \rho(\mu_{n_i}, \mu)$ implies immediately that $\nu_{n_i} \to \mu$ and since each ν_{n_i} belongs to the compact set F(x) it follows that $\mu \varepsilon F(x)$, Q.E.D.

Using the results of the previous two theorems a continuity property is shown next.

Theorem 3 Let D be any non-empty compact subset of the space $\mathcal{M} x R$ and define $F(Dx\{t\}) = \bigcup \{F(\mu_o, t_o, t) \;/\; (\mu_o, t_o) \varepsilon D\}$. Then the function F(D,.) is continuous whenever it is defined (that is, for every t which is equal to or greater than t_o, for every $(\mu_o, t_o) \varepsilon D$).

NOTE: As an example of such a set D consider the set $D = A\otimes[t_1, t_2]$, where A is a non-empty compact subset of the metric space $(\mathcal{M},\rho)$.

Proof Since $D\otimes\{t\}$ is a non-empty compact subset of the domain of F, it follows from theorem 2 that the function $F(D,.)$ is upper quasicontinuous. Therefore, we only have to show that for every $\varepsilon>o$ there exists a $\delta>o$ such that $|\tau-t|<\delta$ implies $s(F(D\otimes\{t\}), F(D\otimes\{\tau\}))<\varepsilon$. Suppose this is not true; then for every $\varepsilon>o$ and for every $\delta>o$ the relation $|\tau-t|<\delta$ implies that there exists $\mu\varepsilon F(D\otimes\{t\})$ such that $d(\mu,F(D\otimes\{\tau\}))\geq\varepsilon$. Let now $\{\tau_n\}$ be a sequence converging to t and consider the corresponding, by the previous relation, sequence $\{\mu_n\}$ of points of the set $F(D\otimes\{t\})$. Since, by theorem 2, the set $F(D\otimes\{t\})$ is compact, there exists a subsequence $\{\mu_{n_i}\}$ converging to some point $\mu^o\varepsilon F(D\otimes\{t\})$. Considering then the corresponding sequence $\{\tau_{n_i}\}$ it follows that $d(\mu_{n_i}, F(D\otimes\{\tau_{n_i}\}))\geq\varepsilon$ for every $n_i> n_o = n_o(\delta)$. On the other hand, we have that

$$d(\mu_{n_i}, F(D\otimes\{\tau_{n_i}\})) \leq \rho(\mu_{n_i}, \mu^o) + d(\mu^o,F(D\otimes\{\tau_{n_i}\}))$$

and since $\mu_{n_i}\to\mu^o$ it is enough to show that there exists some integer $m_o = m_o(\delta)$ such that $d(\mu^o,F(D\otimes\{\tau_{n_i}\}))<\varepsilon/_2$ for every $n_i>m_o$. To do this we consider the following three cases.

a. If $\tau_{n_i} > t$, then consider a point $\nu_{n_i} \varepsilon F(\mu^o, t, \tau_{n_i}) \subset F(Dx\{\tau_{n_i}\})$. Then $d(\mu^o, F(Dx\{\tau_{n_i}\}))$ $\leq d(\mu^o, F(\mu^o, t, \tau_{n_i})) = \rho_H(\{\mu^o\}, F(\mu^o, t, \tau_{n_i}))$.

It follows from definition 5(ii) that $\rho_H(\{\mu^o\}, F(\mu^o, t, \tau_{n_i})) =$ $= \rho_H(F(\mu^o, t, t), F(\mu^o, t, \tau_{n_i}))$ and then continuity of F in t yields the desired result.

b. If $\tau_{n_i} = t$, then $d(\mu^o, F(Dx\{\tau_{n_i}\})) = \rho_H(F(\mu^o, t, t), F(Dx\{t\})) = 0$.

c. If $\tau_{n_i} < t$, it follows from definition 5(iv) and the fact that $\mu^o \varepsilon F(Dx\{t\})$, that we can consider a point $\nu_{n_i} \varepsilon F(Dx\{\tau_{n_i}\})$ such that $\mu^o \varepsilon F(\nu_{n_i}, \tau_{n_i}, t)$. Then continuity of F in t implies that, as $\tau_{n_i} \to t$, $\rho_H(F(\nu_{n_i}, \tau_{n_i}, \tau_{n_i}), F(\nu_{n_i}, \tau_{n_i}, t)) < \varepsilon/_2$ for every n_i which is greater than some integer $m_o = m_o(\varepsilon)$ and, since $\mu^o \varepsilon F(\nu_{n_i}, \tau_{n_i}, t)$, it follows that $\rho(\mu_o, \nu_{n_i}) < \varepsilon/_2$, which implies that $d(\mu_o, F(Dx\{\tau_{n_i}\})) < \varepsilon/_2$, for $n_i > m_o$, Q.E.D.

Another property of the attainability function is shown in theorem 4, whose proof requires the following two lemmas.

Lemma 6 Let (X, ρ_X) be a metric space, B a disconnected subset of X and F_1, F_2 two non-empty closed and disjoint subsets of B such that $F_1 \cup F_2 = B$. Then, if E is a connected subset of B we necessarily have either $E \subset F_1$ or $E \subset F_2$.

Proof We prove this lemma by contradiction. Suppose that $E \cap F_1 \neq \phi$ and $E \cap F_2 \neq \phi$. Since $(E \cap F_1) \cap (E \cap F_2)$ $= E \cap F_1 \cap F_2 = E \cap \phi = \phi$, it follows that $E \cap F_1$ and $E \cap F_2$ are two non-empty, closed and disjoint subsets of E such that $(E \cap F_1) \cup (E \cap F_2) = E \cap (F_1 \cup F_2) = E \cap B = E$. This contradicts our assumption that E is connected. Therefore, either $E \cap F_1 = \phi$ or $E \cap F_2 = \phi$ and this proves the lemma.

Lemma 7 Let $C(\mathcal{X})$ be the family of all non-empty compact subsets of the metric space $(\mathcal{X}, \rho)$ and consider a continuous function $f : I = [t_0, t_1] \rightarrow C(\mathcal{X})$ such that $f(t_0)$ is a connected subset of $\mathcal{X}$. Then the set $H = \bigcup \{f(t) \; / \; t \in I\}$ is a compact and connected subset of $\mathcal{X}$.

Proof We show first that H is compact. Since H is a subset of a metric space, it is enough to show that for every sequence $\{\mu_n\}$ of points of the set H there exists a subsequence converging to some point in H. It follows from the definition of H that for every $\mu_n \varepsilon$ H there exists $t_n \varepsilon I$ such that $\mu_n \varepsilon f(t_n)$. Since I is compact the sequence $\{t_n\}$ contains a subsequence $\{t_{n_i}\}$ converging to some point $t^o \varepsilon I$. Consider now, in the sequence $\{\mu_n\}$, the subsequence $\{\mu_{n_i}\}$ corresponding to the sequence $\{t_{n_i}\}$. As $t_{n_i} \to t^o$, it follows from continuity of f at t^o that, for every $\varepsilon > o$ there exists $\delta > o$ such that the relation $\rho_H(f(t_{n_i}), f(t^o)) < \varepsilon/_2$ holds for every $n_i > n_{i_o} = n_{i_o}(\delta)$. Since $\mu_{n_i} \varepsilon f(t_{n_i})$, for every i = 1,2,---, it follows that $d(\mu_{n_i}, f(t^o)) < \varepsilon/_2$, for every $n_i > n_{i_o}$ and since $f(t^o)$ is a compact subset of the metric space $(\mathcal{K},\rho)$ there exists $\nu_{n_i} \varepsilon f(t^o)$ such that $\rho(\mu_{n_i},\nu_{n_i}) < \varepsilon/_2$, for every $n_i > n_{i_o}$. On the other hand, the sequence $\{\nu_{n_i}\}$ contains a subsequence $\{\nu_{n_{i_k}}\}$ converging to some point $\nu_o \varepsilon f(t^o)$. Considering then the corresponding sequence $\{\mu_{n_{i_k}}\}$ it follows from the relation $\rho(\mu_{n_{i_k}},\nu_o) \leq \rho(\mu_{n_{i_k}},\nu_{n_{i_k}}) + \rho(\nu_{n_{i_k}},\nu_o)$ that $\mu_{n_{i_k}} \to \nu_o \varepsilon$ H, Q.E.D.

We show next that H is a connected subset of the metric space $(\mathcal{K},\rho)$. Define first, for each $t \varepsilon I$, the set function $H(t) = \bigcup\{f(\tau) / \tau \varepsilon [t_o,t]\}$ and observe that it is a non-decreasing function of t. Moreover, the function H(t) is continuous. This follows easily from continuity of f and the following relations.

a. For $t' \geq t$ we have:

$$\rho_H(H(t'), H(t)) = \rho_H(H(t') \setminus H(t), H(t))$$

$$= \rho_H(\bigcup\{f(T) / t \leq T \leq t'\}, H(t))$$

$$= \rho_H(f(T_o), f(t))$$

where T_o is an appropriate point in the interval $[t,t']$.

b. For $t' < t$ we have:

$$\rho_H(H(t),H(t')) = \rho_H(H(t) \setminus H(t'), H(t\ ) = \rho_H(\bigcup\{f(T) / t' \leq T \leq t\}, H(t'))$$

$$= \rho_H(f(T_o), f(t'))$$

where T_o is an appropriate point in the interval $[t', t]$. Suppose now that H is not connected. Then we can find two non-empty, closed and disjoint subsets F_1 and F_2 of H such that $F_1 \bigcup F_2 = H$. Since the function H(t) is non-decreasing, the set $f(t_o) = H(t_o)$ is a subset of the set $H = H(t_1)$. Moreover, the set $f(t_o)$ is connected by assumption. It follows then from lemma 6 that $f(t_o)$ is a subset of either F_1 or F_2. Without loss of generality we assume that $f(t_o) \subset F_1$. Define now the set $A = \{t \varepsilon I / H(t) \subset F_1\}$. We show next that the set A is closed. Since A clearly contains t_o, it is enough to show that A is an interval containing its supremum. That A is an interval can be easily seen from the fact that for any $\tau_1, \tau_2 \varepsilon A$ with $\tau_1 < \tau_2$ and any $\tau \varepsilon I$ such that $\tau_1 < \tau < \tau_2$, we have that $H(\tau) \subset H(\tau_2) \subset F_1$, which implies that $\tau \varepsilon A$. On the other hand, let $\tau^* = \sup A$ and consider a sequence $\{\varepsilon_n\}$ of positive numbers such that $\varepsilon_n \to o$ and $\tau^* - \varepsilon_n > t_o$. Since $(\tau^* - \varepsilon_n) \varepsilon A$, it follows that $H(\tau^* - \varepsilon_n) \subset F_1$, for every $n = 1, 2, ---$ and since F_1 is closed, it follows from lemma 4.

$$\lim_{n \to \infty} H(\tau^* - \varepsilon_n) = H(\lim_{n \to \infty} (\tau^* - \varepsilon_n)) = H(\tau^*) \subset F_1$$

which implies that $\tau^* \varepsilon A$. Hence the set A is closed. Similarly one can show that its complement A^c is also closed. Since F_1, F_2 are non-empty sets with $F_1 \cup F_2 = H$, it follows that A and A^c are also non-empty which, along with the fact that A and A^c are closed sets such that $A \cup A^c = I$, implies that the interval I is not connected. This is a contradiction.

<u>Theorem 4</u> If A is a non-empty compact and connected subset of the metric space $(\mathcal{H}, \rho)$ then so is the set

$$F(A \otimes \{t_o\} \otimes [t_o, t_1]) = \bigcup \{F(\mu_o, t_o, t) \,/\, \mu_o \varepsilon A,\ t \varepsilon [t_o, t_1]\}$$

for every $t_1 \geq t_o$.

Proof: Since the set A x $\{t_0\}$ x $[t_0, t_1]$ is a non-empty compact subset of the domain of F as a product of compact sets, it follows from theorem 2 that the set $F(A \times \{t_0\} \times [t_0, t_1])$ is a compact subset of the metric space $(\mathscr{K}, \rho)$. Moreover, since $D = A \times \{t_0\}$ is a non-empty compact subset of the space $\mathscr{K} \times R$, it follows from theorem 3 that the function F(D,.) is a continuous function of t on the interval $[t_0, t_1]$. Finally, since $F(D, t_0) = A$ is a connected subset of the metric space $(\mathscr{K}, \rho)$, the proof of the theorem follows immediately from lemma 7.

4. THE METRIC SPACE ($\mathcal{M}$,ρ)

On the set $\mathcal{M}$ of measures various metrics can be defined so that it becomes a metric space. Two examples of such metrics are given below.

1. <u>The Prohorov Metric</u> Consider the function

$\rho_p : \mathcal{M} \times \mathcal{M} \to R$ such that $\rho_p(\mu_1,\mu_2)$

$= \max \{\varepsilon_{12}, \varepsilon_{21}\}$

where,

$\varepsilon_{12} = \inf \{\varepsilon \,/ \mu_1(A) < \mu_2 (S_\varepsilon(A)) + \varepsilon$, for every closed set $A\}$

$\varepsilon_{21} = \inf \{\varepsilon \,/ \mu_2(A) < \mu_1 (S_\varepsilon(A)) + \varepsilon$, for every closed set $A\}$

The function ρ_p as defined above, is a metric on $\mathcal{M}$, called the 'Prohorov metric' (Billingsley, p. 238)

2. <u>The w* - metric on $\mathcal{M}$</u> Let M be the state space of the system and C (M,R) the set of all continuous and bounded real-valued functions on M. For every point $\mu \varepsilon \mathcal{M}$ consider the family of subsets of $\mathcal{M}$ of the form

$N = N(f_1, ---, f_k, \varepsilon_1, ---, \varepsilon_k)$

$= \{\nu \varepsilon \mathcal{M} / |\int_M f_i d\nu - \int_M f_i d\mu| < \varepsilon i,\ i = 1, ---, k$

where $f_1, ---, f_k$ are elements of C(M,R) and $\varepsilon_1, ---, \varepsilon_k$ are positive numbers. The family of sets N which results if we vary k, $f_1, ---, f_k, \varepsilon_1, ---, \varepsilon_k$, satisfies the axioms of a neighborhood basis for each $\mu \varepsilon \mathscr{M}$. Therefore it defines a topology on $\mathscr{M}$ which is called the 'w*-topology' on $\mathscr{M}$.

<u>NOTE</u>: Observe that a sequence $\{\mu_n\}$ of points of the set $\mathscr{M}$ converges to a point $\mu \varepsilon \mathscr{M}$ in the w*-topology, if and only if, $\int_M f d\mu_n \rightarrow \int_M f d\mu$, for every $f \varepsilon C(M,R)$.

We show next that the w*-topology on $\mathscr{M}$, as defined above, can be metrized, under certain conditions on the state space of the system. The properties of the resulting metric on $\mathscr{M}$ are presented in theorem 5 whose proof requires the following lemmas.

<u>Lemma 8</u> The set $\mathscr{M}$ with the w*-topology can be metrized as a separable metric space, if and only if, M is a separable metric space.

<u>Proof</u> (Parthasarathy, p. 43, theorem 6.2).

<u>Lemma 9</u> The set $\mathscr{M}$ with the w*-topology is a compact metric space, if and only if, M is a compact metric space.

<u>Proof</u> (Parthasarathy, p. 45, theorem 6.4).

<u>Lemma 10</u> If M is a separable metric space, then the set $\mathscr{M}$ with the w*-topology is complete, if and only if, M is a complete metric space.

<u>Proof</u> (Parthasarathy, p. 46, theorem 6.5).

In view of the previous lemmas 8, 9 and 10, the proof of the following theorem is obvious.

Theorem 5 If the state space M of the system is a separable and compact metric space, then the set $\mathcal{M}$ equipped with the w*-topology can be metrized as a complete and compact metric space.

NOTE: Under the conditions of theorem 5, the metric space $(\mathcal{M},\rho)$ becomes a complete and compact metric space. However, theorems 1-4, in previous section, require only that $(\mathcal{M},\rho)$ be locally compact. Therefore, it would be interesting to investigate for weaker conditions on M so that $(\mathcal{M},\rho)$ becomes only locally compact (and not compact and complete).

5. EXAMPLES

The study of questions concerning stochastic systems, using their attainability functions, presents the advantage of viewing the problems in a natural geometric way which, sometimes, suggests a method of solution.

Consider as an example, the case of two stochastic systems described by their attainability functions F_1 and F_2 with initial conditions (μ_{01}, t_{01}) and (μ_{02}, t_{02}), respectively. Moreover, suppose that F_1 is a pursuing system and that we wish to reach the target F_2, in minimum time. If $F_1(\mu_{01}, t_{01}, t) = \{\mu^u(\mu_{01}, t_{01}, t) \,/\, u \in B\}$, where B is a set of bounded measurable functions, then the whole problem can be viewed in the following way. Let,

$$Q(t) = F_1((\mu_{01}, t_{01}, t)) \bigcap F_2((\mu_{01}, t_{01}, t))$$

and consider the following two cases:

(i) Suppose that there exists a time $t \geq \max\{t_{01}, t_{02}\}$ such that $Q(t) \neq \phi$. Then, if $I = \{t / Q(t) \neq \phi\}$ and $t^* = \inf I$, one is interested to know the conditions under which $t^* \in I$, that is, the conditions under which $Q(t^*) \neq \phi$. It is clear then, that these conditions will ensure the existence of an admissible control, yielding an optimal time solution to the problem.

(ii) Suppose next that for every $t \geq \max\{t_{01}, t_{02}\}$, we have $Q(t) = \phi$. Then, letting

$$g(t) = d(F_1((\mu_{01}, t_{01}, t)), F_2((\mu_{02}, t_{02}, t)))$$

(the distance between the sets of attainable measures, at time t) it is natural to ask if there exists an infimum, for the values of g and under what conditions this infimum is one of the values of g. It is then clear that these conditions would ensure the existence of a time and a corresponding admissible control, which would be the approximate solution to the problem. (e.g., if g is a real-valued continuous function on a compact set, then its infimum is finite and belongs to its range).

As it was indicated in the previous example, useful qualitative information can be obtained, from an attainability function formulation of the problem.

Let us now see an example of how the attainability function of a stochastic system is identified.

Consider the system

$$dx = u_1 dt + u_2 dw$$

where, the admissible controls $u = (u_1, u_2)$ are bounded measurable functions such that

$$-1 \leq u_1(s) \leq +1 \quad , \quad -1 \leq u_2(s) \leq +1$$

and w is the Wiener process.

The conditional probability density function $p(x/x_0)$ of the system satisfies the equation,

$$\frac{\partial p(x/x_o)}{\partial t} = u_1 \frac{\partial p(x/x_o)}{\partial x} + u_2^2 \frac{\partial^2 p(x/x_o)}{\partial x^2}$$

whose solution is (Takacs, pp. 103-104)

$$p_\tau(x/x_o) = \frac{1}{\sqrt{2\pi\lambda(\tau)}} e^{-\frac{(x-x_o-a(\tau)^2}{2\lambda(\tau)}}$$

where

$$a(\tau) = \int_o^\tau u_1(s)\, ds \quad , \quad \lambda(\tau) = \int_o^\tau (2u_2(s))^2 ds$$

Supposing now that the initial probability density function is Gaussian with variance σ_o and mean m_o, we obtain the following expression for the total probability density function of the system:

$$p_\tau(x) = \frac{1}{\sqrt{2\pi(\sigma_o^2 + \lambda(\tau))}} e^{-\frac{(x-m_o-a(\tau))^2}{2(\sigma_o^2 + \lambda(\tau))}}$$

Since $\lambda(\tau)$ and $a(\tau)$ are functions of the control u, considering all admissible controls we obtain the set of attainable probability density functions and hence the corresponding set of attainable measures for the system at time τ. The question which now arises is to find an admissible control $u = (u_1, u_2)$ such that $p_\tau(x)$ is a given desired function. It is clear that unless the given function belongs to the attainable set at time τ, such a control does not exist. However, in this case, one is interested to know if a control exists such that the corresponding probability function $p_\tau(x)$ has a minimum distance from the given function. The existence of such a control depends on the attainable set being compact and compactness of the attainable set is ensured by the conditions given in theorem 1.

6. CONCLUSION

In the present work, the concept of the attainability function was extended to stochastic systems. As made clear in the given examples, an approach to the problems through the attainability functions of the systems presents the advantage of viewing the problems in a natural geometric setting.

The properties of the attainability function were studied with the only assumption that the metric space $(\mathcal{H},\rho)$ is locally compact, while the additional assumption of completeness which was made in previous works, was found to be unnecessary. In order to apply these properties in a particular system, after identifying the attainability function of the system, one should verify its axioms and define on $\mathcal{H}$ a proper metric ρ so that the space $(\mathcal{H},\rho)$ becomes locally compact.

In conclusion we should point out that the present theory, providing mostly qualitative information about the system, should be complemented by a computational procedure. We believe that it is in this direction where much effort should be devoted.

REFERENCES

1. AHMED, N. U.

   'Optimal control of a large class of stochastic differential systems with finite memory', Proc. 7th annual Allerton conference on circuit and system theory, October 1969, pp. 44-53.

2. BILLINGSLEY, P.

   'Convergence of probability measures', John Wiley, 1968.

3. LASALLE, J.P. & HERMES, H.

   'Functional analysis and time optimal control', Academic Press, 1969.

4. PARTHASARATHY, K.R.

   'Probability measures on metric spaces', Academic Press, 1967.

5. ROXIN, E.

   'Stability in generalized control systems', J. Diff. eq. 1, 1965, pp. 115-150.

6. TAKACS, L.

   'Stochastic Processes', Methuen, 1960.

CURRICULUM VITAE

Name : Nicolas SPYRATOS

Place of birth : Argostolion , Kefallinia , GREECE.

Date of birth : March 5 , 1942 .

Education :

1) 8th High School Of Athens , 1954-1960 .
2) National Technical University Of Athens ,
   Department Of Mechanical And Electrical Engineering ,
   Dipl. Ing. in Electrical Engineering , 1961-1966 .