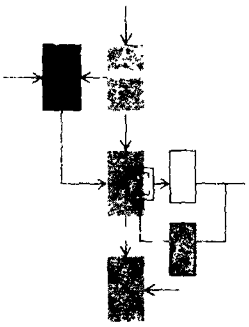
https://ntrs.nasa.gov/search.jsp?R=19660024151 2020-03-16T20:22:14+00:00Z

MAY, 1966

REPORT ESL-R-269
M.I.T. PROJECT DSR 9952
NASA Research Grant NsG-496 (Part)



MINIMAX CONTROLS of UNCERTAIN SYSTEMS

GPO PRICE $ ______

CFSTI PRICE(S) $ ______

Hard copy (HC) 5.00

Microfiche (MF) 1.25

ff 653 July 65

Hans S. Witsenhausen

N66 33441
(ACCESSION NUMBER)
184
(PAGES)
CR-77050
(NASA CR OR TMX OR AD NUMBER)
(THRU)
1
(CODE)
10
(CATEGORY)
FACILITY FORM 602

Electronic Systems Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MASSACHUSETTS 02139

Department of Electrical Engineering

May, 1966 Report ESL-R-269 Copy 59

MINIMAX CONTROLS OF UNCERTAIN SYSTEMS

by

Hans S. Witsenhausen

This study was conducted at the Electronic Systems Laboratory with support extended by the National Aeronautics and Space Administration under Contract NsG-496 with the M.I.T. Center for Space Research (M.I.T. Project DSR 9952). Reproduction in whole or in part is permitted for any purpose of the United States Government.

Electronic Systems Laboratory
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

N66-33441

ABSTRACT

The control of uncertain systems is considered as a decision problem. Some general concepts, such as adaptivity, are analyzed in this light and features common to many design methods are clarified. The special case of worst-case, or minimax design, is considered in more detail. The dynamic programming algorithm is discussed for a class of linear problems with bounded perturbations, bounded control variables, and with sampled output of the state. A dual algorithm using the support functions of reachable sets is proposed. Bounds are obtained, relating the performance of optimal and suboptimal designs, when the criteria have the properties of norms.

Author

ACKNOWLEDGEMENT

The material presented in this report is based on a thesis submitted in May, 1966 in partial fulfillment of the requirements for the Doctor of Philosophy Degree in Electrical Engineering.

The author is happy to have this opportunity to express his gratitude to Professor Michael Athans for his inspiring teaching, his constant encouragement and support, his helpful suggestions and criticisms and his willingness to share an invaluable collection of reprints. The present work also owes much to the comments of Professors Leonard A. Gould, Robert G. Gallager and Roger W. Brockett, who acted as readers.

The friendly advice and encouragement given the author by his Faculty Counselor, Professor Alfred K. Susskind, were highly appreciated.

The author's thinking on the problems of control and on the application thereto of mathematical methods has been greatly influenced by the teaching and the comments of Professors Elijah Polak (visiting from the University of California at Berkeley), Stuart E. Dreyfus (visiting from the Rand Corporation), Arthur E. Bryson Jr. (visiting from Harvard University), Murray Wonham (visiting from Brown University), W. Gilbert Strang (of the Department of Mathematics) and of Dr. Fred C. Schweppe (Lincoln Laboratory).

During the last year of his doctoral program the author enjoyed the generous support of the Fannie and John Hertz Foundation, which is gratefully acknowledged.

CONTENTS

		page
ABSTRACT		ii
ACKNOWLEDGMENTS		iii
INTRODUCTION		viii
CHAPTER I	ON THE DECISION PROBLEM	1
1.1	Definition of the decision problem	1
1.2	Numerical indexing of outcomes	1
1.3	The case of certainty	2
1.4	Partial ordering of decisions	3
1.5	Rules of choice: hedged and unhedged	4
1.6	Utility	5
1.7	Reformulation of the decision problem	6
1.8	Valuations and evaluators	8
1.9	The guaranteed performance evaluator	9
1.10	Expected performance evaluators	10
1.11	Guaranteed expected performance evaluators	11
1.12	The interchange inequality, opvalue and lopvalue	12
1.13	Relation to game theory	14
1.14	Certainty equivalence and Wald's decision theory	17
1.15	Sources	19
CHAPTER II	UNCERTAIN CONTROL PROBLEMS	21
2.1	Introduction	21
2.2	The need for precision	22

CONTENTS (Contd.)

2.3	Plants	page 24
2.4	Control systems and feedback	26
2.5	Valued plants and control problems	30
2.6	Decision by evaluators	31
2.7	Optimization with respect to evaluators and adaptivity	36
2.8	Gap reduction by extension	39
2.9	Constraints	47
2.10	Filtering problems	48
2.11	Identification problems	48
2.12	Information	50
2.13	Sources	50
CHAPTER III	THE TIME FACTOR	53
3.1	Heuristic introduction	53
3.2	Imperfect recall	57
3.3	Clock uncertainty	60
3.4	Timed control problems	62
3.5	Markovian sets	67
3.6	Variable end time	69
3.7	Outline of further developments	70
3.8	A discrete time case	72
3.9	The dynamic programming algorithm	77
3.10	Superobserver description	80

CONTENTS (Contd.)

3.11	Algorithm using internal description	page 84
3.12	Case of independent output uncertainty	88
3.13	Case of internal state output	89
3.14	Introduction of structure	91
3.15	Sources	93
CHAPTER IV	REACHABLE SETS OF LINEAR DIFFERENTIAL SYSTEMS	95
4.1	Introduction	95
4.2	The basic mapping	95
4.3	Superposition of independent constraints	97
4.4	Duality	97
4.5	Hard constraints and the maximum principle	98
4.6	The dual Hamilton-Jacobi equation	101
4.7	Hard constraints and duality	102
4.8	Initial condition constraints	104
4.9	Energy constraints	104
4.10	Sources	108
CHAPTER V	VECTOR ADDITION GAMES	109
5.1	Introduction	109
5.2	Vector addition games	109
5.3	Linear differential systems with end-point criteria	110
5.4	Linear differential systems with general criteria	112
5.5	Single stage case	115

CONTENTS (Contd.)

		page
5.6	The canonic form	116
5.7	Example of reduction to canonic form	117
5.8	The game-theoretic interpretation	119
5.9	The dynamic programming algorithm	120
5.10	The conservation of convexity	122
5.11	The conservation of uniform continuity	123
5.12	Why a minimax principle does not hold	124
5.13	The use of duality	126
5.14	Single-stage duality	130
5.15	Multistage duality	131
5.16	Sources	134
CHAPTER VI	BOUNDS FOR THE PERFORMANCE OF SUBOPTIMAL CONTROLLERS	135
6.1	Introduction	135
6.2	The naive controllers	136
6.3	Main assumptions	138
6.4	The single-stage case	140
6.5	A bound for expected performance	149
6.6	The two-stage case	151
6.7	The multistage case	155
6.8	Sources	156
BIBLIOGRAPHY		157

INTRODUCTION

Frequently, a controller for a plant must be designed in the absence of a full, precise description of this plant; that is, in the face of uncertainty. The selection of a design is then a special case of the general problem of decision under uncertainty.

In the first three chapters of the present work, some of the possible approaches to this problem are considered and their implications in the control context are investigated. An attempt is made to discuss each concept on the basis of the minimum of structure required, uncluttered by irrelevant assumptions.

Only deterministic controllers, as opposed to random controllers, are considered. The performance of controllers is measured by a "supercriterion"; for instance, the expectation of the original criterion for given a-priori probabilities or its supremum for given bounds on the uncertain quantities. The latter case is that of worst-case or minimax design.

Notions to which are given the names "optimal", "feedback" and "adaptive" can then be defined, and this before the introduction of the notion of time. It is shown how optimization over open-loop designs provides a-priori bounds for the improvement possible by the use of adaptive controllers.

In Chapter III, some first steps towards a theory taking the time factor into account are outlined. Rather than pursuing this subject further in full generality, only a specialization is considered more deeply: discrete-time systems with the minimax definition of optimality. (It is to be noted that many continuous-time systems with sampled outputs are reducible to discrete-time form.) The solution of the corresponding optimization problem is immediate in principle, by dynamic programming.

In the case of linear differential systems with sampled output, the minimax optimization algorithm is best described in terms of reachable sets under given constraints. In Chapter IV, the formulas giving the support functions of such reachable sets are derived. In Chapter V, the corresponding dynamic programming algorithm is examined in more detail. A dual algorithm which appears to present some advantages is proposed. It is based on Fenchel's theory of conjugate convex functions.

Since the computational effort for the determination of the minimax design, as above, is considerable, it is tempting to design on the assumption that all uncertain quantities are fixed at nominal values. Such a "naive" design can be found by the less demanding algorithms for optimal control under certainty. Both open and closed loop forms of the naive design can be considered. In Chapter VI the question of the relative merit of the optimal and the various supoptimal controllers is posed. Some inequalities which may begin to throw light on this question are obtained, though further research on this topic is called for.

In summary, the present work is of an exploratory nature. One major advantage of the study of minimax design is the absence of many of the existence problems that are common in the stochastic approach. It is possible to concentrate at once upon the physically or algorithmically relevant difficulties created by the presence of uncertainty. In the end, it should be possible to use any insight gained from minimax studies as a guide in the investigation of alternative approaches. To some degree such a cross-fertilization is exemplified by the bounds on expectations mentioned in the last chapter.

CHAPTER I

ON THE DECISION PROBLEM

1.1 DEFINITION OF THE DECISION PROBLEM

A decision problem consists of three non-empty sets D, N and O, of a function $M:D \times N \to O$, and a transitive relation $\leq$ on O such that between any two elements o_1, o_2 of O at least one of $o_1 \leq o_2$ and $o_2 \leq o_1$ holds:

- D is the set of possible decisions or "action space".
- N indexes the uncertainty of the problem and may be called the set of "states of nature", (no relation to the notion of state of dynamic systems).
- O is the set of outcomes.
- M is the function which determines which outcome will result for a given decision and state of nature.

  The relation $\leq$ defines our preference among the outcomes. We take "less" to mean "better" in conformity with the control theory usage of minimization. Thus $o_1 \leq o_2$ means that o_1 is as good or better than o_2.

1.2 NUMERICAL INDEXING OF OUTCOMES

Assume that we can assign to each element of O a real number in such a way that the order between elements of O agrees with the usual order of the corresponding numbers. That is, there exists a function $\nu:O \to R$ with the property

$$\nu(o_1) \leq \nu(o_2) \Leftrightarrow o_1 \leq o_2$$

Then, if ϕ is any monotone increasing function the composition $\phi \circ \nu$ (defined by $\phi(\nu(o))$) has the same property. Thus the function ν is very far from unique.

For any choice of ν as above we may define a function $W: D \times N \to R$ by $W = \nu \circ M$, that is

$$W(d,n) = \nu(M(d,n))$$

The "payoff function" W suffers from the lack of uniqueness inherited from ν.

1.3 THE CASE OF CERTAINTY

In case the set N has but one element (lacks entropy) the argument n of M and W becomes redundant. In this case of "certainty" consider the set

$$D_{opt} = \{d \epsilon D : W(d) = \inf W(D)\}$$

If it is not empty, its elements are the optimal decisions, because no better outcomes can be obtained than those resulting from such a decision. The set D_{opt} is independent of the particular choice of the function ν. Such is not the case for the set

$$D_\epsilon = \{d \epsilon D : W(d) \leq \inf W(D) + \epsilon\}, \quad \epsilon > 0$$

to which one might turn for help in case D_{opt} is empty. For fixed ϵ any decision d can be brought into D_ϵ by appropriate choice of ν. Therefore, the set D_ϵ is only useful if a particular function ν and a value of ϵ can be agreed upon. Otherwise, we might as well consider the set

$$D_{d*} = \{d \epsilon D : W(d) \leq W(d*)\}$$

$$= \{d \epsilon D : M(d) \leq M(d*)\}$$

and attempt to agree upon the choice of d^*. This is clearly as difficult as the original problem so that nothing has been gained.

1.4. PARTIAL ORDERING OF DECISIONS

Define a relation on D, i.e., among decisions by

$$d_1 \leq d_2 \Leftrightarrow (\forall n \epsilon N) \; M(d_1, n) \leq M(d_2, n)$$

Then this relation is transitive and the relation

$$d_1 \sim d_2 \Leftrightarrow d_1 \leq d_2 \text{ and } d_2 \leq d_1$$

is an equivalence relation.

If we consider the relation $\leq$ among equivalence classes of decisions, then it is a partial order. On the set D itself it is a partial order modulo equivalence. The set $D^* = \{d^* \epsilon D : (\forall d \epsilon D) \; d^* \leq d\}$ is usually empty. When it is not, then it is an equivalence class and its members are optimal since decisions outside D^* cannot yield better outcomes, for any state of nature, than the outcome resulting from a decision in D^* for the same state of nature. Decisions in D^* are called dominant decisions.

In the absence of dominant decisions the partial order enables only the definition of the (possibly empty) set D_m of minimal controllers and of the collection $\mathscr{D}_c$ of complete sets of controllers, by

$$D_m = \{d_m \epsilon D : (\forall d \epsilon D) d \leq d_m \Rightarrow d_m \leq d\}$$

and

$$\mathscr{D}_c = \{D_c \subset D : (\forall d \epsilon D) \; (\exists d_c \epsilon D_c) d_c \leq d\}$$

When D_m is not empty and belongs to $\mathcal{D}_c$ then one may consider the problem to be reduced to the selection of a decision in D_m. Indeed, in that case D_m is the smallest set such that for all decisions in D there exists a better decision in D_m, in the sense of the partial order.

D_m is sometimes called the set of "admissible" or "non-inferior" decisions.

1.5. RULES OF CHOICE: HEDGED AND UNHEDGED

To select a decision on a rational rather than intuitive basis it is necessary, in the absence of dominant decisions, to agree on a rule of choice. Such a rule should be compatible with the given preference relation among the outcomes, hence with the partial order of D defined above.

The rule of choice, unlike the preference relation among outcomes, takes into account the presence of uncertainty of the state of nature and possible a priori knowledge (such as probabilities) about this uncertainty.

In practice there are two approaches: unhedged and hedged rules of choice. Unhedged rules create an order among decisions, expressed by assignment of a real (or extended real) number to each decision in D. The situation becomes, then, the same as in the absence of uncertainty (see Section 1.3).

Hedged rules randomize the decision process by specifying a σ-algebra Σ_D of subsets of D and considering the new problem of selecting an element of the set $\tilde{D}$ of all probability measures of Σ_D. The rule orders the set $\tilde{D}$ by assignment of a real (or extended real) number to each element of $\tilde{D}$. The situation becomes again the same as in the absence of uncertainty but now with $\tilde{D}$ as the action space. If an element of $\tilde{D}$, say $\tilde{d}$, is selected, the corresponding action consists in the activation "at the last minute" of a random device which selects an element d of D according to the probability measure $\tilde{d}$. This decision d is then made in the original problem.

When the set D is countable, the collection of all its subsets is a natural and convenient choice of Σ_D. When the set D is not countable, then Σ_D would be derived from the structure (usually the topological structure) of D. Unfortunately, even the preliminary step of agreeing on a topological structure for D is by no means clear. Just consider the case where D is the set of "all" nonlinear feedback controllers for a given plant.

1.6 UTILITY

The rules of choice used in practice take account of the preference ordering of outcomes by way of the numerical function ν. Some rules have the property that the resulting decisions (if some exist) remain the same when ν is replaced by $\phi \circ \nu$ as in Section 1.2 and the same rule of choice is used. Most rules of choice, however, do not possess this property and require agreement upon a specific choice of ν, the "utility function".

In particular, agreement on a utility function is necessary when expectations under probability measures are involved and also when a concept of ϵ-optimal decision is used.

Axiomatic treatments of the establishment of the utility function may be found in Von Neumann and Morgenstern and in Pratt, Raiffa and Schlaiffer. Essentially, these authors show that if a set of reasonable axioms is accepted, the existence of a unique utility function follows. The axioms in question are essentially reasonable but by no means compelling. The difficulty stems from the following requirement of these axioms. Let a, b, c be three outcomes with b strictly preferable to a and c strictly preferable to b (and hence to a). Consider the mixed outcome f(p) for $0 \leq p \leq 1$ which consists in a probability p of outcome c and probability (1-p) of outcome a, then it is required that there always exists a value p^* of p, with $0 \leq p^* \leq 1$ such that outcomes b and $f(p^*)$ are equivalent.

This axiom fails to hold if a "worst case" point of view is adopted. Thus acceptance of the axioms rules out one of the most simple-minded and logically consistent approaches to the decision problem.

In the sequel we will assume that a specific function ν has been selected. For each rule of choice it will be clear whether the results are or are not invariant under monotone remapping of ν into $\phi \circ \nu$.

1.7 REFORMULATION OF THE DECISION PROBLEM

Once the function ν has been selected, a decision problem takes the form (D, N, W) where D is the action space, N indexes the uncertainty and $W: D \times N \rightarrow R$ (or R_e) defines the utility of the outcome or at least the preference relation among outcomes.

An unhedged rule of choice may now be defined as follows:

Definition: An unhedged rule of choice ρ is a function which associates to each decision problem (D, N, W) in a set P of problems (the domain of the rule) an extended real valued function $J = \rho(D, N, W)$ on the set D of the argument problem and satisfies

(a) $(\forall (D, N, W) \in P)\ (\forall d_1, d_2 \in D)$

$$[(\forall n \in N)\ W(d_1, n) \leq W(d_2, n)] \Rightarrow J(d_1) \leq J(d_2)$$

where $J = \rho(D, N, W)$

(b) Let S_D be the set of all mappings of the set D onto itself which swap two elements of D and leave the others unchanged. Then it is required that

(b1) $(\forall (D, N, W) \in P)\ (\forall \sigma \in S_D)\quad (D, N, W \circ \sigma) \in P$

(b2) $(\forall (D, N, W) \in P)\ (\forall \sigma \in S_D)$

$$\rho(D, N, W \circ \sigma) = \sigma \circ \rho(D, N, W)$$

Requirement (a) expresses the compatibility of ρ with the partial ordering inherited from the outcome preferences. The symmetry requirement (b) expresses the independence of the a-priori knowledge of the state of nature, as embodied in the rule of choice, with respect to the selection of a decision. We conjecture that every meaningful problem can be cast into a form in which this independence is realized. The function J obtained by the rule of choice may be called a supercriterion to distinguish it from the function W which depends on the uncertainty and corresponds to the usual concept of a criterion.

A similar precise definition of the notion of hedged decision rule would depend on whether the σ-algebra on D is considered as part of the problem data to which the rule is applied or is considered to be selected by the rule.

1.8 VALUATIONS AND EVALUATORS

For a given decision problem (D, N, W) there corresponds to each decision d a function on N, defined by W with d fixed.

Definition: The <u>valuation</u> of a decision d is the function $W(d, \cdot): N \to R$ (or R_e). The <u>value mapping</u> V of the decision problem (D, N, W) is the function which associates to every element of D the corresponding valuation.

The range V(D) of V is a subset of the set of all functions from N to R_e. Note that valuations are partially ordered by pointwise inequality and that this partial order is precisely the one which defines the partial order of the corresponding decisions.

An important special class of unhedged rules of choice is the class of evaluators, that is rules assigning a number to a decision solely on the basis of the corresponding valuation. The supremum of the valuation and its expectation under a given probability measure are prime examples of evaluators.

Definition: A function $C: \mathcal{D}(C) \to R_e$ is an <u>evaluator</u> if (a) the domain $\mathcal{D}$ (C) consists of pairs (N, ϕ) where N is a nonempty set and ϕ a function from this set into R (or R_e).

(b) Whenever (N, ϕ_1) and (N, ϕ_2) belong to $\mathcal{D}(C)$ and $(\forall n \in N)\phi_1(n) \leq \phi_2(n)$, that is, $\phi_1 \leq \phi_2$ in the partial order, then $C(N, \phi_1) \leq C(N, \phi_2)$.

When a fixed N is under discussion, the first argument N of C is redundant and need not be written. In that case of fixed N the domain of C is a set $\mathcal{D}(C)$ of functions from N into R or R_e.

An evaluator C is <u>applicable</u> to a decision problem (D, N, W) if all valuations (N, V(d)) are in the domain of C. For fixed N this requirement reads $V(D) \subset \mathcal{D}(C)$. When an evaluator is applicable to a decision problem the function $J: D \to R_e$ defined by $J(d) = C(N, V(d))$ is the corresponding supercriterion. The rule of choice is to select decisions, if some exist, which minimize J.

1.9 THE GUARANTEED PERFORMANCE EVALUATOR

The pessimistic decision maker will consider the worst case resulting from each decision. This amounts to the use of the guaranteed performance evaluator, defined for any pair (N, ϕ) by

$$C(N, \phi) = \sup_{n \in N} \phi(n)$$

One technical advantage of this evaluator is that it is applicable to every decision problem. Note that it is not necessary that the supremum be a maximum, the "worst case" need not exist. One has $J(d) = \sup_{n \in N} W(d, n)$.

The study of this evaluator is motivated by the desire to assess in advance what <u>can</u> happen when a given decision is selected, on any basis whatsoever. One would then compute the guaranteed performance for that decision. It appears worthwhile to have a standard of comparison for the number so obtained, and the best guaranteed performance.

$$\inf_{d \in D} \sup_{n \in N} W(d, n) = \inf_{d \in D} J(d)$$

is eminently suitable as such. The determination of this quantity amounts to considering the decision problem from the point of view of optimizing guaranteed performance.

The guaranteed performance $J(d)$ may very well turn out to be independent of d. This is likely if the outcomes are ordered into just two equivalence classes (success and failure) with the function ν taking only two values. If for each decision there is at least one state of nature leading to failure J will be constant. Another possibility is that J be constant with the value $+\infty$ because $V(d)$ is unbounded for each d. Since extremization commutes with monotone functions, optimality for guaranteed performance is independent of the selection of the utility function ν.

1.10 EXPECTED PERFORMANCE EVALUATORS

For fixed set N, an expected performance evaluator C is defined by a σ-algebra on N and a probability measure μ on this σ-algebra. The domain $\mathcal{D}(C)$ is the set of μ-integrable functions on N, augmented by the functions having $+\infty$ or $-\infty$ as μ-integral. The definition is $C(\phi) = \int_N \phi(n)\, d\mu(n) = \underset{\mu(n)}{E}\, \phi(n)$.

The applicability of such an evaluator to the problem (D, N, W) must be carefully checked in each case. If applicable, the super-criterion is

$$J(d) = \underset{\mu(n)}{E} \quad W(d, n)$$

and the rule of choice is to select decisions minimizing J, if some exist. The optimal performance is

$$\inf_{d \in D} \underset{\mu(n)}{E} \quad W(d, n)$$

1.11 GUARANTEED EXPECTED PERFORMANCE EVALUATORS

There may be a set of probability measures on a common σ-algebra on N, one such measure μ_a for each element a of the index set A. Then one may define an evaluator, for fixed N, by

$$C(\phi) = \sup_{a \in A} \underset{\mu_a(n)}{E} \quad \phi(n)$$

The domain $\mathcal{D}$ (C) consists of the functions ϕ which are μ_a integrable for each $a \in A$, with the values $\pm\infty$ allowed.

In that case the supercriterion is

$$J(d) = \sup_{a \in A} \underset{\mu_a(n)}{E} \quad W(d, n)$$

and the optimal guaranteed expected performance is given by the expression

$$\inf_{d \in D} \sup_{a \in A} \underset{\mu_a(n)}{E} W(d, n)$$

The rule of choice is to select decisions for which this infimum is attained, if some exist.

We note that the two previously considered types of evaluators are special cases of the present one. Expected performance corresponds to the case where A consists of a single element. Guaranteed performance is obtained by letting the σ-algebra be that of all subsets of N and by taking A = N with μ_n the atomic measure with unit weight at point n.

It is only in the latter case of guaranteed performance that the rule of choice is independent of the selection of the utility function ν. This advantage of the guaranteed performance evaluator is lost when it is used as a basis for numerical comparison of decisions and also, of course, when ϵ-optimal guaranteed performance is considered.

1.12 THE INTERCHANGE INEQUALITY, OPVALUE AND LOPVALUE

The set of all extended real-valued functions on N is a complete lattice under the partial order induced by pointwise inequality. (Every subset of such functions has an infimum and supremum under the partial order.)

Given a decision problem (D, N, W) define its minimal valuation ϕ_m by

$$\phi_m = \inf_{d \in D} V(d)$$

that is $\phi_m(n) = \inf_{d \in D} W(d, n)$

An evaluator C is completely applicable to a decision problem if it is applicable and, in addition, the minimal valuation belongs to the domain $\mathcal{D}(C)$.

Note that guaranteed performance is completely applicable to any decision problem.

If an evaluator C is applicable to (D, N, W) then we define the corresponding opvalue (optimum value) as the extended real number.

$$v = \inf_{d \in D} C(V(d))$$

If C is completely applicable then we can also define the lopvalue (lower optimum value) as

$$v' = C(\phi_m)$$

where ϕ_m is the minimal valuation of the problem.

In that case, we have the interchange inequality:

$$v' \leq v$$

Proof: By definition of the minimal valuation

$$(\forall d \in D) \quad \phi_m \leq V(d)$$

under the partial order.

By definition of an evaluator and by the assumption of complete applicability, this implies

$$(\forall d \in D) \quad C(\phi_m) \leq C(V(d))$$

Taking the infinum over all d in D,

$$C(\phi_m) \leq \inf_{d \in D} C(V(d))$$

or

$$v' \leq v$$

as claimed. The case where $v = v'$ is called the zero gap situation.

When $v > v'$ the positive extended real number $v - v'$ is uniquely defined and is called the <u>gap</u>.

For a given decision problem the opvalue , lopvalue and gap all depend on the evaluator used.

The interpretation of the gap is the following. If the decision must be made in the assumed way, the value of the evaluator cannot be reduced below the opvalue v. If, on the other hand, the problem is changed to one in which the actual value of n will be made available shortly before the decision must be selected, so that the decision can be taken under conditions of certainty, then <u>before</u> the value of n becomes known, there is still uncertainty as to the results but it can be asserted that the value of the evaluator can now be reduced no lower than the lopvalue. In a zero-gap situation the effect of "spying" is nil as far as the value of the evaluator is concerned. If the gap is positive it represents the value of "spying" in terms of the evaluator.

1.13 RELATION TO GAME THEORY

The zero gap situation for the guaranteed performance is known in game theory as the case of a pure value. When the gap is positive, game-theory recommends the use of hedged rules of choice. It abandons the guaranteed performance in favor of a merely expected performance whose value is numerically more favorable, despite the fact that the guaranteed performance under the hedged optimal decision procedure may be worse than the guaranteed performance of some unhedged decisions.

EXAMPLE: Consider the game with payoff matrix

	α	β	γ
a	2	2	-1
b	5	-4	-10
c	-4	-1	8

The minimizing player chooses among a, b, c and the maximizing player among α, β, γ.

The pessimistic strategy is to select the strategy for which the worst possible result is as good as possible. This is a for one player α or β (select α) for the other player. The optimistic strategy is the one which would bring the greatest reward if one were able to direct the opponent's move. Here b and γ are optimistic. The most dangerous strategy is the one the opponent wishes for if he is playing his optimistic strategy. Here c and γ are the most dangerous. The equal probability strategy selects at random among abc or $\alpha\beta\gamma$ with probabilities 1/3.

The Von Neumann strategy is the one which gives a saddle-point for the expectation of the payoff. Here its probabilities are $(0, \frac{4}{5}, \frac{5}{9})$ for the minimizing player and $(\frac{2}{3}, 0, \frac{1}{3})$ for his opponent. Finally the maximizing player might be following Murphy's law:* he plays α against a and b, γ against c regardless how the selection of a, b, or c was made.

The values or expectations of the resulting payoff are tabulated below for various combinations of these strategies.

---

* Murphy's law: anything that can go wrong, will.

	(a) pess.	most dg. opt.	= prob	VN	Murphy	guarantee
pessimistic	2	-1	1	1	2	2
optimistic	5	-10	-3	0	5	5
most dangerous	-4	8	1	0	8	8
= prob.	1	-1	-1/3	+1/3	5	8
V.N.	0	0	-7/9	0	6-2/3	8

The calculation of the payoff against Murphy is the expectation of the payoff with respect to the probabilities of a, b, c. The "guarantee" is obtained by applying Murphy's law to the minimizing player's chance device: consider the worst selection among those with positive probability.

It should be clear from this example that no randomized strategy can give a guarantee lower than that obtained by the pessimistic strategy, i.e., the upper value of the game.

In particular, the Von Neumann strategy is not optimal in the sense of guaranteed performance. To clarify this apparent contradiction, distinguish two cases.

Case I: The decision maker (the minimizing player) decides, for reasons which need not concern us, to be able to guarantee the results and therefore uses the guaranteed performance evaluator to assess any type of decision, hedged or unhedged. In this case he can under no circumstances obtain a guarantee better than the opvalue of the guaranteed performance evaluator. If the gap is positive the hedging procedure suggested by game theory, which may yield a better expected performance, is definitely incorrect and he must avoid it.

Case II: The decision maker considers a σ-algebra over N and accepts his ignorance as to the probability measure in force. The set A is now the set of all probability measures over the σ-algebra. He chooses the guaranteed expected performance as evaluator to select unhedged decisions. Then, assuming the technical difficulties of integrability are resolved, he necessarily obtains the same opvalue as for the guaranteed performance evaluator. Indeed, the set A contains in particular the atomic measures with unit weight at a point n. Thus the guaranteed expected performance can not be lower than the guaranteed performance. The opposite inequality is due to the fact that expectation is order-preserving, and we must have equality.

Now the possibility exists that a hedged decision give a better guaranteed expected performance than the opvalue for unhedged decisions. This improvement is precisely what game theory accomplishes and he should avail himself of the possibility. The point is that by choosing guaranteed expected performance he declared himself content with a mere expectation and if such is the case hedging can often yield an improvement.

1.14 CERTAINTY EQUIVALENCE AND WALD'S DECISION THEORY

The zero-gap situation for expected performance is also known as certainty equivalence, and is expressed by

$$\inf_{d \in D} \; \underset{\mu(n)}{E} \; W(d,n) = \underset{\mu(n)}{E} \; \inf_{d \in D} \; W(d,n)$$

Note that in game theory the interchange of extremization with expectation is never considered. This is because the "spying" of game theory is directed against a human opponent. Random devices are considered spy-proof. In the context of control, with nature as "opponent", spying is accomplished by an increase of the measurements taken and this can be done in a situation described by probabilities.

For the case of guaranteed expected performance with a set A of probability measures μ_a over N the same difference of point of view between game theory and the evaluator approach is encountered.

The opvalue v is defined by

$$v = \inf_{d \epsilon D} \sup_{a \epsilon A} \mathop{E}_{\mu_a(n)} W(d, n)$$

The lopvalue v' by

$$v' = \sup_{a \epsilon A} \mathop{E}_{\mu_a(n)} \inf_{d \epsilon D} W(d, n)$$

Define v" the "game-theoretic lower value" by

$$v'' = \sup_{a \epsilon A} \inf_{d \epsilon D} \mathop{E}_{\mu_a(n)} W(d, n)$$

Consider a game in which the minimizing player chooses d in D, his opponent chooses a in A and the payoff is

$$W^*(d, a) = \mathop{E}_{\mu_a(n)} W(d, n)$$

For this game, denoted by (D, A, W*), the game-theoretic upper value is v and the lower value v".

By the interchange inequality we have

$$v \geq v'' \geq v'$$

Therefore the zero-gap situation $v = v'$ implies that the game has a pure value $v = v''$ but the converse is false.

Wald's statistical decision theory amounts to the following:

if $v = v''$ guaranteed expected performance is the rule of choice (whether or not $v'' = v'$)

if $v > v''$ switch to hedged decisions according to the usual game-theoretic procedure applied to (D, A, W^*)

Our point of view is different: v represents the limit on the performance possible in the given problem, as judged by the evaluator; v' represents the limit for the modified problem, in which "spying" allows one to make the decision as a function of n, as judged by the evaluator before the value of n becomes known; v'' has no special significance.

1.15 SOURCES

Game-theory, utility theory and their relation to decision making were first clarified by Von Neumann and Morgenstern [58].*

A broad application of game-theory to statistical decision making was then proposed by Wald [59].

In view of some objections to the pessimism of game theory, when the opponent is nature, a great deal of effort went into the axiomatic study of decision making; see the books by Blackwell and Girshick [10], by Thrall et al. (especially the section by Milnor [55]), by Luce and Raiffa [38].

In more recent times favor has gone to the approach in which a priori probabilities are estimated, however roughly, and the expectation of utility is used as supercriterion. A strong argument

---

*Numbers in [ ] refer to numbered items in the bibliography.

for this procedure can be found in Pratt, Raiffa and Schlaifer [47].

Consider the following statement: "There is no way to avoid having to make a decision as to the rules by which decisions are to be made." It is a shocking statement because of its circularity. One may say that decision theory was developed in an attempt to "disprove" this statement but ended up by "proving" it.

Sometimes the only uncertainty in a decision problem is the choice of the preference relation among the outcomes. Formally this is just a special case: the competing preference relations are indexed by the set N. The problem of vector-valued criteria (see Zadeh [64]) is of this nature.

CHAPTER II

UNCERTAIN CONTROL PROBLEMS

2.1 INTRODUCTION

The distinction between "plant" and "controller" is in many ways an artificial one. For example, the compounding of d.c. generators may be viewed either as an improvement of design of the plant or as a form of feedback control.

The interest in a clear-cut distinction between plant and controller has been reinforced by the advent of the computer. The plant is now taken as some process with actuators accepting inputs in computer signal form (analog or digital) and with sensors providing outputs in the same form. The actuators and sensors are considered part of the plant. The controller is viewed as an on-line computer, with a program, which makes the actuator input signals some function of the sensor output signals.

The uncertainties as to the behavior of the plant, as to the demands that will be made of it and as to the operation of actuators and sensors are all considered as uncertainties of the plant. The controllers under consideration are defined by those operators from sensor outputs to actuator inputs which can be realized with negligible error and uncertainty by the available data processing equipment.

In this view two problems arise:

Problem I: Assuming the plant to be already designed, or given by nature, find a controller such that the system will give satisfactory performance in some sense.

Ideally, the design of a plant, including the choice of actuators and sensors, should take the need for control into account. When faced with the total design task, one needs to consider:

Problem II: How do the solutions of problem I, and especially the attainable performance, depend on design parameters of the plant?

At least abstractly, the total design task reduces to a problem of the first type. One need only consider the controller as setting the values of the design parameters at the beginning of the process. The corresponding actuator is the builder of the plant. Note that what distinguishes design parameters is that they must be fixed at the outset and can not be changed during the process under consideration, they are thus equivalent to adjustable initial conditions.

2.2 THE NEED FOR PRECISION

In this chapter and the one following we state our point of view on uncertain control systems in a precise mathematical fashion. Thus specific definitions are given to terms that have been used with many loose meanings and will certainly continue to be so used. Also a large number of new terms are introduced.

We want to state explicitly that these definitions are not considered to be the only suitable ones. It is clear that loose notions as crucial as "feedback", "adaptive", "optimal", etc., can be made precise in many different ways, according to particular points of view. What is inadmissible, and leads to fallacies, is to use these notions as if they were precise without giving a definition.

Let us point out what the definition of a property must do. It must precisely state the class of objects under consideration and it must provide an unambiguous test to decide whether an object in the class does or does not have the property. In other words it must specify a set and a partition of this set into two complementary subsets.

Many statements about general control notions are not definitions in the above sense or else are vitiated by extraneous assumptions, such as linearity.

The need for extreme precision is best illustrated by listing a few statements which, from our point of view, are fallacious.

Fallacies:

1. Feedback is a property of certain controllers.
2. Feedback is a property that a control system may possess in the absence of any uncertainty.
3. A control system is adaptive if it consists of a plant, a controller which applies inputs to the actuators dependent on sensor outputs and a "supervisor" which changes the structure of the controller in a way dependent on sensor outputs.
4. Filtering, estimation or identification problems can not be considered as control problems.
5. A controller for an uncertain plant should always consist of two independently designable parts: (1) an identifier or estimator which determines the value of the uncertain quantities and (2) a controller designed without taking uncertainty into account.

6. Suppose a plant is operated in the morning for identification and in the afternoon for business. Then, a) the problem as to what to do in the morning may be treated as a separate problem, the results of which are then used to solve the problem of operation in the afternoon.
   b) the entire day's operation is something beyond a control problem.

7. The synthesis of optimal control in feedback form is always at least as good as the open-loop optimal control.

It will turn out that the most basic concepts can be defined independently of the notion of time and therefore of the notion of state. Full advantage is taken of this fact since it is inadvisable to use any unnecessary notions in basic definitions. The crucial role which time will play later on is due to the simplifications that result from the use of state concepts. With the introduction of time new complications will arise. These matters will be examined in Chapter III.

2.3 PLANTS

Definition 2.1 A plant is an ordered quadruple (U, Q, Y, S) in which U, Q and Y are nonempty sets and $S: U \times Q \rightarrow Y$.

U is called the input set, Q the uncertainty set, Y the output set and S the system function. Elements of U are called inputs, elements of Y outputs, elements of Q uncertainties.

<u>Definition 2.2</u> The <u>effective output set</u> of the plant (U, Q, Y, S) is the subset $S(U, Q)$ of Y.

<u>Definition 2.3</u> A <u>determinate plant</u> is a plant whose uncertainty set is a singleton (a set with just one member).

<u>Definition 2.4</u> A <u>filter plant</u> is a plant whose system function is independent of its first argument (the input).

<u>Definition 2.5</u> A <u>mute plant</u> is a plant whose system function is independent of its second argument (the uncertainty).

<u>Definition 2.6</u> An <u>outputless</u> plant is a mute plant which is also a filter plant (i.e., the effective output set is a singleton).

<u>Definition 2.7</u> An <u>inputless plant</u> is a plant whose input set is a singleton.

Filter plants are so named because they arise in filtering problems considered as control problems. Mute plants are so named because their output tells nothing about the uncertainty. Clearly, every determinate plant is mute, since there is nothing to tell.

If A is a nonempty set and $f: U \times Q \times Y \to A$ a function arising in the study of a plant (U, Q, Y, S), then f may always be reduced to a function $g: U \times Q \to A$ by defining $g(u, q) = f(u, q, S(u, q))$.

Since the variables u and y are externally accessible, functions expressible in terms of these variables only have special significance.

<u>Definition 2.8</u> A function $g: U \times Q \to A$ is an <u>external function</u> for the plant (U, Q, Y, S) iff

$$(\forall u \in U)(\forall q_1, q_2 \in Q)\ S(u, q_1) = S(u, q_2) \Rightarrow f(u, q_1) = f(u, q_2)$$

This is the necessary and sufficient condition for the existence of a function $h:U \times Y \rightarrow A$ such that $g(u,q) = h(u, S(u,q))$.

Then the restriction of h to the set $\{(u, S(u,q)) \mid u \in U, \ q \in Q\} \subset U \times Y$ is unique and the values of h on the complement of this set are immaterial.

2.4 CONTROL SYSTEMS AND FEEDBACK

Definition 2.9 The set $\Gamma(P)$ of all controllers for plant P is the set of all functions $\gamma: Y \rightarrow U$ which have the property that for each fixed q in Q the equation $u = \gamma(S(u,q))$ has one and only one solution (dependent on q) for u.

In any control problem a subset $\Gamma \subset \Gamma(P)$ is given: the set of all controllers under consideration. The definition of Γ will take into account causality and any other practical limitations.

If γ is in $\Gamma(P)$ then the simultaneous equations

$$y = S(u,q)$$
$$u = \gamma(y)$$

have exactly one solution for u and y, dependent on q, for each fixed q in Q.

Definition 2.10 A control system is a pair (P, γ) where P is a plant and γ a controller for P.

Definition 2.11 The loop function of control system (P, γ) is the function $L_\gamma: U \times Q \rightarrow U$ defined by $L_\gamma(u,q) = \gamma(S(u,q))$

Definition 2.12 The input mapping of control system (P, γ) is the function $m_\gamma: Q \rightarrow U$ which maps q into the solution

$m_\gamma(q)$ of the equation $u = L_\gamma(u, q)$

Consequently $m_\gamma(q) \equiv L_\gamma(m_\gamma(q), q) \equiv \gamma(S(m_\gamma(q), q))$ holds for all q in Q and γ in $\Gamma(P)$, by definition.

The next step is to make precise the notion of feedback.

Definition 2.13 A controller γ for a plant P is blind if γ is a constant function. It is effectively blind if γ is constant on the effective output set of P.

A blind controller γ is completely defined by the constant value u of $\gamma(y)$. By "the blind controller u" is meant the controller γ with $\gamma(y) = u$ for all y in Y. The input set U may thus also be called the set of all blind controllers for P. When considered as such, U is always a subset of $\Gamma(P)$, which proves that $\Gamma(P)$ is never empty.

Definition 2.14 (P, γ) is an open-loop control system if its input mapping is constant, otherwise it is a feedback control system.

Note that a control system may be open-loop even for a controller γ which is not (effectively) blind. It may, for instance, happen that changes in q produce changes in y insufficient to bring $\gamma(y)$ out of a dead-zone. The system is open loop when γ is constant on the set $\{S(m_\gamma(q), q) \in Y : q \in Q\}$ and then the set $m_\gamma(Q)$ is a singleton.

An immediate consequence of definition 2.14 is

Theorem 2.1 If a control system (P, γ) satisfies one or more of the following 3 conditions:

(1) P is inputless

(2) P is mute (a fortiori: outputless or determinate)

(3) γ is effectively blind (a fortiori: blind)

then it is an open-loop control system.

Thus there is no such thing as a determinate feedback control system. It is only for certain special types of plant that the only way to obtain an open-loop control system is to choose γ effectively blind.

Theorem 2.2 If a plant P has the property that the set $Y_u \equiv \{S(u, q) \epsilon Y : q \epsilon Q\}$ is independent of u and if the control system (P, γ) is open loop, then γ is effectively blind.

Proof: Since (P, γ) is open loop $m_\gamma(q)$ is constant with some value u^*. This implies that $\gamma(y) = u^*$ for all y in the set Y_{u^*}. But, by assumption, Y_u is independent of u, so that $Y_{u^*} = \bigcup_{u \epsilon U} Y_u = S(U, Q)$ the effective output set. Thus γ is constant on this set, hence effectively blind. Q.E.D.

An application of this case is given in

Theorem 2.3 If P is a filter plant and (P, γ) is open-loop then γ is effectively blind.

Proof: Filter plants trivially satisfy the assumption of theorem 2.2.

Q.E.D.

As illustration, consider the case of a tracking servo. Then the set U is the set of all time functions describing possible input signals to the power chain, the set Q is the set of all time functions that the servo may be called upon to track, the set Y consists of pairs: an element of Q is paired with a time function that may be received from the servo output sensors.

Since the first component of an element $y \epsilon Y$ is directly determined by $q \epsilon Q$ and the second component is a known function of $u \epsilon U$ (assuming no uncertainty in the power chain), the mapping S is well defined. Since time plays a role one is restricted to the use of physically realizable (i.e., non-predictive) controllers $\gamma : Y \rightarrow U$ and there may be many other restrictions. But there is no magic reason why the controller γ should depend only on the difference called error.

If the subtraction is carried out in the plant and Y is the set of error time functions one can reconstitute the two terms of the subtraction as long as the power chain is perfectly known. If, on the other hand, the power chain is uncertain also, say because of noise, then Q is a subset of the cartesian product of the set of all time functions to be tracked and the set of all noise time-functions. Now there is a genuine loss if the controller receives only the error signal. Indeed, every realizable γ dependent on the error signal only is realizable if both terms of the subtraction are received but no longer vice-versa, and it can never be harmful to have a choice over a larger set of possible controllers.

2.5 VALUED PLANTS AND CONTROL PROBLEMS

Definition 2.15 A valued plant is a pair (P, K) where $P = (U, Q, Y, S)$ is a plant and $K: U \times Q \to R_e$ is a function, the cost function or criterion of the valued plant.

The value of $K(u, q)$ is interpreted as a cost so that $K(u_1, q_1) < K(u_2, q_2)$ means that (u_1, q_1) leads to a result preferable to (u_2, q_2).

The extended real line R_e is used because

a) real valued criteria are included as a special case.

b) R_e is closed under supremum while R is not.

c) it may sometimes be convenient to introduce constraints by letting $K(u, q)$ be infinite when (u, q) leads to violation of constraints.

Definition 2.16 A control problem (P, K, Γ) is a valued plant (P, K) together with a subset Γ of the set $\Gamma(P)$ of all controllers for plant P.

Definition 2.17 A valued control system (P, K, γ) is a valued plant (P, K) together with a controller γ for P.

Thus a control problem is equivalent to a set of valued control systems with the same valued plant.

<u>Definition 2.18</u> *The <u>payoff function</u> of a control problem (P, K, Γ) is the function $W: \Gamma \times Q \to R_e$ defined by $W(\gamma, q) = K(m_\gamma(q), q)$ where m_γ is the input mapping of (P, γ). W is the restriction to $\Gamma \times Q$ of the payoff function of $(P, K, \Gamma(P))$ which is called the payoff function of the valued plant (P, K).

2.6 DECISION BY EVALUATORS

The control problem (P, K, Γ) can be cast in the form of a decision problem (Γ, Q, W) where Γ is the set (formerly called D) of possible decisions and Q the set (formerly called N) of possible "states of nature".

Accordingly, by Chapter I, one has the following notions:

a) valuations are functions from Q into R_e,

b) the value mapping V associates a valuation to each element of Γ: $V(\gamma) = W(\gamma, \cdot)$,

c) valuations are partially ordered by pointwise inequality and this defines equivalence and partial order among the elements of Γ,

d) the minimal valuation ϕ_m is defined by

$$\phi_m(q) = \inf_{\gamma \in \Gamma} W(\gamma, q)$$

e) evaluators, their applicability and complete applicability to a control problem are defined,

f) the opvalue v and lopvalue v' of control problem (P, K, Γ) with evaluator C (assumed completely applicable)

---

* The cost of implementing controller γ is neglected, otherwise it would have to be included in the definition of W.

are given by

$$v = \inf_{\gamma \in \Gamma} C(V(\gamma))$$

$$v' = C(\phi_m)$$

and satisfy $v' \leq v$ (the interchange inequality).

Definition 2.19 A valuation ϕ on the uncertainty set Q of a plant P is a blind valuation if there exists a blind controller u for P such that $V(u) = \phi$

With this notion the motivation behind the definition of feedback comes to the fore:

Theorem 2.4 If the control system (P,γ) is open-loop and K is any criterion for P then the valuation of (P, K, γ) is blind.

Proof: Definition 2.19 is satisfied by letting u be the constant value of the input mapping of (P, γ). Q.E.D.

A first definition for a notion of adaptivity can now be given.

Definition 2.20 The valued control system (P, K, γ) is strictly adaptive if its valuation is strictly less (in the partial order of valuations) than any blind valuation of (P, K).

This means that for all u in U

a) $K(m_\gamma(q), q) \leq K(u, q) \;\forall\; q \in Q$

b) $(\exists q \in Q)\;\; K(m_\gamma(q), q) < K(u, q)$

this q depending, of course, upon u.

This definition is a first attempt to make precise the following idea: a valued control system is adaptive if it uses feedback to advantage. When, as in definition 2.20, the word "advantage" is defined by the partial order, the requirement so expressed is far too strong to be met by non-trivial systems.

Indeed one has

Theorem 2.5 If the valued control system (P, K, γ^*) with value mapping V is strictly adaptive, then

a) (P, γ^*) is a feedback control system

b) $V(\gamma^*) = \inf V(\Gamma(P))$, that is γ^* is dominant in $\Gamma(P)$ and, a fortiori, in any subset Γ containing γ^*.

Proof: a) By definition 2.20 the valuation of (P, K, γ^*) is not blind and by Theorem 2.4 this implies that (P, γ^*) is not open-loop, hence it is feedback.

b) If γ is any controller for P then, for all q in Q

$$K(m_\gamma(q), q) \geq \inf_{u \in U} K(u, q)$$

because $m_\gamma(q)$ belongs to U. Therefore $V(\gamma) \geq \inf V(U)$ in the partial order. Taking the infimum over γ in $\Gamma(P)$

$$\inf V(\Gamma(P)) \geq \inf V(U)$$

But the opposite inequality must hold because U is a subset of $\Gamma(P)$. Hence

$$\inf V(\Gamma(P)) = \inf V(U) = \phi_m$$

where ϕ_m is the common minimal valuation of the control problems (P, K, U) and $(P, K, \Gamma(P))$. The claim is that $V(\gamma^*) = \phi_m$ and this follows from the fact that $V(\gamma^*) \geq \inf V(\Gamma(P))$ because $\gamma^* \epsilon \Gamma(P)$ while, by definition 2.20, $V(\gamma^*) \leq \inf V(U)$.

QED

An example of a strictly adaptive system is the following: Consider a classical optimal control problem in which the only uncertainty is in the values of the initial conditions. Suppose these initial conditions are available as output at the beginning of the control process. Then the designer will first find the optimum control for each possible initial condition and design a device which applies, as a function of the measured initial condition, the corresponding optimal control. Since no other uncertainty is present, it is immaterial whether this is done by the rapid selection of a time function, which is then blindly applied, or by continuous feedback methods. If at least one pair of possible initial conditions have no common optimal control, then such a design is strictly adaptive.

In the next section, our basic idea of adaptivity is made precise in a weaker but far more practical sense. Instead of defining advantage by the partial order, an evaluator is used. With that type of definition most controllers presently used are adaptive. The realization of adaptivity is the reason for the use of feedback and thus dates back at least to Watt's regulator. The notion that "adaptive controllers" constitute a breakthrough to a new level of the control art appears entirely mistaken. Only a matter of degree is involved. The strength of "modern" control theory is that it takes uncertainty, constraints and nonlinearity into account and uses the power of the computer (on-line and off-line) to achieve, sometimes, an improvement in performance. The idea of "learning" is nothing more than feedback, in the case where U and Y contain functions of time.

It goes without saying that the word "adaptive" could be defined in quite different ways, and a number of such proposals have been made. It seems to us that the definition should

a) be independent of all structural assumptions,

b) take the measure of performance into account, and

c) be closely related to the notion of feedback.

These desiderata are taken as a guideline throughout the present chapter.

2.7 OPTIMIZATION WITH RESPECT TO EVALUATORS AND ADAPTIVITY

When the evaluator C is applicable to the control problem (P, K, Γ) with value mapping V, the corresponding supercriterion $J: \Gamma \rightarrow R_e$ is defined by

$$J(\gamma) = C(V(\gamma))$$

The opvalue v is then given as

$$v = \inf_{\gamma \in \Gamma} J(\gamma)$$

and the optimal controller set Γ^* is defined by

$$\Gamma^* = \{\gamma \in \Gamma : J(\gamma) = v\}$$

Elements of Γ^* are called optimal controllers.

Taking account of the partial order on Γ we have

<u>Theorem 2.6</u> The following five cases are mutually exclusive and exhaustive.

1. Γ^* contains a dominant controller γ^*. That is $V(\gamma^*) \leq V(\gamma)$ for all γ in Γ.
2. Γ^* contains a complete subset Γ^{**} of minimal elements, and Γ^{**} contains at least one pair of non-equivalent controllers. (γ_1 is equivalent to γ_2 if $V(\gamma_1) = V(\gamma_2)$).
3. Γ^* is not empty, its set of minimal elements is not complete.
4. Γ^* is empty and v is finite. Then for $\epsilon > 0$ the set $\Gamma_\epsilon = \{\gamma \in \Gamma : J(\gamma) \leq v + \epsilon\}$ is infinite.

5. Γ^* is empty and $v = -\infty$. Then for all real a the set $\Gamma_a = \{\gamma \in \Gamma : J(\gamma) \leq a\}$ is infinite.

Proof: A consequence of the definitions, just note that $v = +\infty$ implies that Γ^* is non-empty.

Discussion: In case (1) of theorem 2.6 the problem may be considered as solved. In the other cases it is reduced to one of the following:

Case (2): (P, K, Γ^{**}).

Case (3): (P, K, Γ^*) or (P, K, Γ_c) with Γ_c complete in Γ^*.

Case (4): (P, K, Γ_ϵ) for some $\epsilon > 0$.

Case (5): (P, K, ϵ_a) for some real a.

This new problem may be further reduced by the use of a new evaluator C_2 different from C and this process may be repeated. C is the primary evaluator C_2 the secondary evaluator, etc....

Evaluator-based definition of adaptivity:

Definition 2.21 Let C be an evaluator applicable tc the control problem (P, K, Γ) and let γ be an element of Γ. Then the valued control system (P, K, γ) is quasi-adaptive with respect to Γ and C iff for all γ_0 in Γ.

(P, γ_0) is open-loop $\Rightarrow J(\gamma_0) \geq J(\gamma)$. It is adaptive with respect to Γ and C if this holds with strict inequality.

Note that when Γ contains U then (P, K, γ) is (quasi-) adaptive iff $(\forall u \in U)\ J(u) > (\geq)\ J(\gamma)$

The idea expressed by definition 2.21 is the same as before: the adaptive system uses feedback to advantage.

This definition can easily be extended to supercriteria which are not derived from evaluators, for instance to the minimum regret supercriterion

$$J(\gamma) = \sup_{q \in Q} (W(\gamma, q) - \inf_{\gamma' \in \Gamma} W(\gamma', q))$$

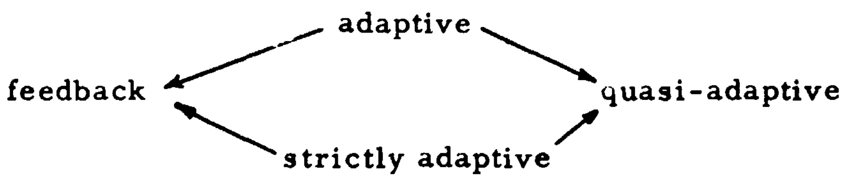
<u>Theorem 2.7</u> Let C be an evaluator applicable to control problem (P, K, Γ) and γ an element of Γ. Then if (P, K, γ) is strictly adaptive, it is quasi-adaptive with respect to Γ and C.

Proof: Let V be the value mapping of (P, K). Let $\gamma_0 \epsilon \Gamma$ be such that (P, γ_0) is open loop. Then by theorem 2.4, the valuation $V(\gamma_0)$ is blind. Since (P, K, γ) is strictly adaptive $V(\gamma) \leq V(\gamma_0)$. By definition of an evaluator $J(\gamma) = C(V(\gamma)) \leq J(\gamma_0) = C(V(\gamma_0))$. Thus definition 2.21 is satisfied. Q.E.D.

<u>Theorem 2.8</u> Let C be an evaluator applicable to (P, K, Γ) and γ an element of Γ. Then, if (P, K, γ) is adaptive with respect to Γ and C, the control system (P, γ) is feedback.

Proof: If (P, γ) were open-loop, definition 2.21 would require $J(\gamma) < J(\gamma)$, a contradiction. Q.E.D.

Four properties: Feedback, strictly adaptive, quasi-adaptive and adaptive have been introduced. Among them there are 12 possible implications. Of these, exactly 4 are true, symbolized by the diagram



2.8 GAP REDUCTION BY EXTENSION

In most cases the opvalue of a control problem is greater than the lopvalue. The gap can be reduced by "spying" and in the control context this means an increase in feedback possibilities due to added sensor instrumentation and more powerful control computers. This process will now be studied in some detail.

Definition 2.22 The control problem (P_2, K_2, Γ_2) with value mapping V_2 is an extension of the problem (P_1, K_1, Γ_1) with value mapping V_1 if the following two conditions are satisfied.

1. P_1 and P_2 have the same uncertainty set
2. The set $V_1(\Gamma_1)$ of all valuations of (P_1, K_1, Γ_1) is a subset of $V_2(\Gamma_2)$.

Thus extension simply means an increase of the set of feasible valuations.

Definition 2.23 Let (P_1, K, Γ_1) and (P_2, K, Γ_2) be two control problem with $P_1 = (U, Q, Y_1, S_1), P_2 = (U, Q, Y_2, S_2)$ such that

$$(\forall \gamma_1 \in \Gamma)(\exists \gamma_2 \in \Gamma_2)(\forall q \in Q)(\forall u \in U)\gamma_1(S_1(u, q)) = \gamma_2(S_2(u, q))$$

then (P_2, K, Γ_2) is called a strict extension of (P_1, K, Γ_1).

The definition of strict extension requires that input set, output set and criterion be the same for the two problems. But the definition is independent of the criterion as long as it is the same for both problems. The condition in the definition is equivalent to the requirement that every loop function feasible in problem (P_1, K, Γ_1) is also feasible in (P_2, K, Γ_2). Thus, instead of an increase of the set of feasible valuations (def. 2.22) an increase in the set of feasible loop functions is considered and this might be called an increase in feedback possibilities.

Theorem 2.9 A strict extension is an extension.

Proof: The first condition of Def. 2.22 is satisfied, in fact U, Q and K are the same in both problems. Let γ_1 be any controller in Γ_1. By Def. 2.23 there exists a controller γ_2 in Γ_2 such that $(\forall q \in Q)(\forall u \in U)\, \gamma_1(S_1(u, q)) = \gamma_2(S_2(u, q))$. Then for all q in Q the equations

$$u = \gamma_1(S_1(u, q))$$

$$u = \gamma_2(S_2(u, q))$$

are identical. Thus their solutions which define the input mappings are the same:

$$(\forall q \in Q)\ m_{\gamma_1}(q) = m_{\gamma_2}(q)$$

Since K is the same in both problems, we have

$$(\forall q \in Q)\ K(m_{\gamma_1}(q), q) = K(m_{\gamma_2}(q), q)$$

In terms of the value mappings V_1 and V_2 this relation becomes

$$V(\gamma_1) = V(\gamma_2)$$

and we have shown that

$$(\forall \gamma_1 \in \Gamma_1)\ (\exists\, \gamma_2 \in \Gamma_2)\ V(\gamma_1) = V(\gamma_2)$$

that is

$$V(\Gamma_1) \subset V(\Gamma_2)$$

which is the second condition of Def. 2.22. Q.E.D.

Theorem 2.10 If $\Gamma_1 \subset \Gamma_2$ then (P, K, Γ_2) is a strict extension of (P, K, Γ_1).

Proof: The conditions of Def. 2.23 are trivially satisfied, by taking $\gamma_2 = \gamma_1$. Q.E.D.

Theorem 2.11 If (P_1, K, Γ_1) and (P_2, K, Γ_2) are uncertain control problems and $P_1 = (U, Q, Y_1, S_1)$, $P_2 = (U, Q, Y_2, S_2)$ with

$$Y_2 = Y_1 \times Y_a \qquad Y_a \neq \emptyset$$

$S_2(u, q) = (S_1(u, q),\ S_a(u, q))$

where $S_a : U \times Q \to Y_a$

and if for all γ_1 in Γ_1 the controller $\gamma_2 : Y_1 \times Y_a \to U$ defined by $\gamma_2(y_1, y_a) = \gamma_1(y_1)$ belongs to Γ_2 then (P_2, K, Γ_2) is a strict extension of (P_1, K, Γ_1)

Proof: The conditions of Def. 2.23 are trivially satisfied by letting γ_2 be the controller associated to γ_1 in the statement of theorem 2.11. Q.E.D.

Theorem 2.10 covers the case of strict extension by addition of controllers. Theorem 2.11 covers strict extension by addition of sensors and controllers using these sensors. Note that the general concept of strict extension is still applicable if one switches to an entirely different set of sensors and controllers, as long as all loop functions remain possible.

<u>Theorem 2.12</u> Let (P_2, K, Γ_2) be an extension of (P_1, K_1, Γ_1). Then the minimal valuations ϕ_{m1} and ϕ_{m2} satisfy

$$\phi_{m2} \leq \phi_{m1}$$

and if an evaluator C is completely applicable to both problems the opvalues v_1 and v_2 and lopvalues v_1', v_2' satisfy the inequalities

$$v_2 \leq v_1 \quad \text{and} \quad v_2' \leq v_1'$$

Proof: Let V_1 and V_2 be the value mappings of the two problems. Then $V_1(\Gamma_1) \subset V_2(\Gamma_2)$ by definition of extension.

Hence $\phi_{m1} = \inf V_1(\Gamma_1) \geq \inf V_2(\Gamma_2) = \phi_{m2}$. By the order-preserving property of evaluators

$$v_1' = C(\phi_{m1}) \leq C(\phi_{m2}) = v_2'$$

and

$$v_1 = \inf_{\phi \in V_1(\Gamma_1)} C(\phi) \geq \inf_{\phi \in V_2(\Gamma_2)} C(\phi) = v_2$$ Q.E.D.

<u>Theorem 2.13</u> Let (P, K, Γ) be a control problem with input set (i.e., set of blind controllers) U and value mapping V. Assume $V(U) \subset V(\Gamma)$, i.e., every blind valuation is feasible. Then

(a) (P, K, Γ) is an extension of (P, K, U)

(b) (P, K, Γ) has the same minimal valuation as (P, K, U)

(c) If evaluator C is completely applicable to (P, K, Γ) with opvalue v and lopvalue v' then C is completely applicable to (P, K, U) with opvalue v_o and lopvalue v'_o which satisfy

$$v'_o = v' \leq v \leq v_o$$

Proof: (a) follows from $V(U) \subset V(\Gamma)$ by definition 2.22.

(b) let ϕ_m and ϕ_{mo} be the minimal valuations of (P, K, Γ) and (P, K, U); then by theorem 2.12

$$\phi_m \leq \phi_{mo}$$

On the other hand, for all q in the uncertainty set Q

$$\phi_m(q) = \inf_{\gamma \in \Gamma} K(m_\gamma(q), q)$$

where m_γ is the input mapping of (P, γ). Define

$$M_\Gamma(q) = \{m_\gamma(q) \in U : \gamma \in \Gamma\}$$

then

$$\phi_m(q) = \inf_{u \in M_\Gamma(q)} K(u, q) \geq \inf_{u \in U} K(u, q) = \phi_{mo}(q)$$

because $M_\Gamma(q)$ is a subset of the input set U. Thus

$$\phi_m \geq \phi_{mo}$$

and since the opposite inequality was previously shown

$$\phi_m = \phi_{mo}$$

(c) Since $V(U) \subset V(\Gamma)$ and $\phi_m = \phi_{mo}$ any C completely applicabie to (P, K, Γ) is completely applicable to (P, K, U).

$v' = v'_o$ follows from $\phi_m = \phi_{mo}$

$v' \leq v$ by the interchange inequality 1.12

$v \leq v_o$ by theorem 2.12.

Q.E.D.

Theorem 2.14 Let (P_i, K_i, Γ_i) for $i = 1, \ldots, n$ be control problems with input sets U_i, value mappings V_i and a common uncertainty set Q. Assume that (P_1, K_1, Γ_1) is an extension of (P_n, K_n, U_n) and that for $i = 1, \ldots, n-1$ $(P_{i+1}, K_{i+1}, \Gamma_{i+1})$ is an extension of (P_i, K_i, Γ_i). Then

(a) The minimal valuations ϕ_{mi} of (P_i, K_i, Γ_i) $i=1,\ldots,n$ and ϕ_{mo} of (P_n, K_n, U_n) are all the same.

(b) If evaluator C is completely applicable to (P_n, K_n, Γ_n) then it is completely applicable to (P_n, K_n, U_n) and to (P_i, K_i, Γ_i) for $i = 1, \ldots, n$.

(c) Let v_i, v'_i and v_o, v'_o be the opvalue and lopvalue of (P_i, K_i, Γ_i) and (P_n, K_n, U_n); respectively, they satisfy

$$v'_o = v'_1 = \ldots = v'_n \leq v_n \leq v_{n-1} \leq \ldots \leq v_1 \leq v_o$$

Proof: (a) It is clear from definition 2.22, that extension is a transitive relation. Consequently (P_n, K_n, Γ_n) is an extension of (P_i, K_i, Γ_i) which in turn is an extension of (P_n, K_n, U_n). Then by theorem 2.12

$$\phi_{mo} \geq \phi_{mi} \geq \phi_{mn}$$

while by theorem 2.13 $\phi_{mo} = \phi_{mn}$ which implies that all minimal valuations are the same; (b) is an immediate consequence of the definition of extension and of (a); (c): the equalities follow from (a); $v'_n \leq v_n$ from the interchange inequality 1.12; the other inequalities from theorem 2.12. Q.E.D.

<u>Theorem 2.15</u> Let (P_i, K, Γ_i) for $i = 1, \ldots, n$ be control problems with common criterion K, input set U and uncertainty set Q. Assume that (P_1, K, Γ_1) is a strict extension of (P_1, K, U) and that for $i = 1, \ldots, n-1$ $(P_{i+1}, K, \Gamma_{i+1})$ is a strict extension of (P_i, K, Γ_i).

Then the claims of theorem 2.14 hold with (P_1, K, U) in the role of (P_n, K_n, U_n).

Proof: By theorem 2.9 a strict extension is an extension. The set of valuations of (P_1, K, U) is the same as that of (P_n, K, U) since K is the same and the difference in output set and system function between P_1 and P_n is immaterial for blind controllers. Therefore, the assumptions of theorem 2.14 are fulfilled and the conclusion follows. Q.E.D.

Theorem 2.15 is the most important in practice. Its implication may be described as follows. We start with an input set, an uncertainty set, a criterion and an evaluator. We determine the opvalue v_o and lopvalue v'_o for the set of all blind controllers. This amounts to solving an uncertain open-loop optimization problem. Then for any strict extension to which the evaluator is completely applicable, i.e., for any increase of feedback possibilities, the new opvalue v will satisfy

$$v'_o \leq v \leq v_o$$

There are two cases:

(a) When $v_o = v'_o$ no improvement of the opvalue is possible regardless of the sensors used. If a blind controller u^* yields v_o it belongs to the set Γ^* of optimal controllers of any strict extension. This does not preclude that Γ^* may contain controllers γ^* preferable to u^* under the partial order, and in fact such is often the case. The control system using γ^* will be quasi-adaptive (as is the one using u^*); it may be strictly adaptive but it can not be adaptive.

(b) When $v_o > v'_o$ then v'_o places a bound on the improvement possible by the use of adaptive systems. To obtain the full improvement (precisely or within ϵ) it is not necessary that every possible loop function $\gamma(S(u,q))$ be realizable (and the corresponding valuation in the evaluators domain) which usually would violate causality. In some cases a realizable strict extension will yield full improvement while in others the bound for realizable strict extensions will lie between v'_o and v_o.

In any event, if we have found a γ^* in some strict extension for which $v = v' = v'_o$ then we can be certain that no improvement of the opvalue can result from further extension, though an improvement with respect to the partial order (or some secondary evaluator) is still possible.

Our final conclusion is that the investigation of open loop systems can provide useful information about the possibilities of adaptive systems, provided a supercriterion based on an evaluator is used.

2.9 CONSTRAINTS

It may be required that some variables related to the control system satisfy given constraints. All system variables are functions of u and q. Consider the set of all pairs (u, q) for which the constraints are satisfied. It is a subset of the cartesian product U x Q and will be called the constraint set.

Optimization of uncertain systems under constraints can be carried out in several ways, such as

(a) The constraints are replaced by a penalty for violation, included in the criterion K. In principle, this can be done by letting $K(u, q) = +\infty$ whenever the constraints are violated. In practice, a smoother real-valued penalty would be used for computational convenience, which implies that the constraints may be violated. If, in fact, violation must be prevented at all costs, the smooth penalty function would be based on an artificial constraint set, a proper subset of the actual constraint set. Once a controller γ has been selected with the help of such a penalty function, a check is made to determine if the use of γ can lead to violation of the actual constraints. If this is the case γ is rejected.

(b) Consider the subset Γ_{cst} of Γ consisting of all controllers γ for which no violation can occur, i.e., for which $(m_\gamma(q), q)$ belongs to the constraint set for all q. If Γ_{cst} is empty the problem has no solution. Otherwise, replace Γ by Γ_{cst} to obtain a new control problem to which the usual methods apply.

(c) If a probability measure on Q is given, let Γ_{cst} be the set of all γ in Γ for which the set of q's for which $(m_\gamma(q), q)$ lies outside the constraint set has zero measure. Then optimize as usual over these controllers. The constraints will be satisfied with probability 1.

The above approaches are mentioned here only to point out the principles involved. In computational algorithms the determination of the constrained optimum would be carried out directly, using, for instance, Lagrange multipliers. The predetermination of Γ_{cst} is too difficult, in general, unless the constraints bear on u alone.

2.10 FILTERING PROBLEMS

Filtering problems may be viewed as uncertain control problems. Let Q_s be the set of signals and Q_n the set of perturbations (noise). Then the uncertainty set Q is the cartesian product $Q_s \times Q_n$ or a subset thereof. Its elements are thus pairs $q = (q_s, q_n)$. The filter γ is to produce an estimate u of some given function of the signal, selected among a set U of possible estimates. The filter receives the corrupted signal $y = S(q_s, q_n)$.

Since the system function S is independent of u the plant is a filter plant. The criterion K will depend on the quality of the estimate, hence on q_s and u, and may also depend on q_n.

2.11 IDENTIFICATION PROBLEMS

For each fixed element q of Q the plant may be considered determinate. The uncertainty of q is then viewed as an uncertainty among a set of determinate plants, indexed by q from Q. Identification has the objective to estimate which plant is in fact realized, that is to estimate q. More generally, one may wish to estimate a function of q; that is, characteristics of the plant. As opposed to filtering problems, identification problems arise in a context where the possibility exists to apply signals to physical actuators. This is done in order to force the plant to reveal some of its characteristics. The controller γ, now called "identifier", receives signal y from the sensors and produces $u = (u_i, u_e)$ where u_i is the input to the physical actuators and u_e the estimate of plant characteristics. The system function is independent of u_e, that is $y = S(u_i, q)$. The criterion K will include the assessment of the quality of the estimate, a function of q and u_e, but, in addition, it will take into account the cost of operating the plant, dependent on u_i and q. Hence K will depend on all variables.

Note that the identification problem is a special case of the uncertain control problem. In the general uncertain control problem there is no separation of γ into an estimator and a controller in some more restricted sense. The problem is simply to select a function γ from a set Γ of functions so as to minimize the supercriterion $J(\gamma)$, and if several minima exist, to select one which is minimal under the partial order and secondary criteria. This procedure, carried out exactly or "within ϵ", subsumes all other considerations.

2.12 INFORMATION

The notion of information can be defined in the framework of control. Consider the strict extension of a control problem resulting from the addition of sensors. The information furnished by the additional sensors can be measured by the decrease in opvalue resulting from the extension, that is by the saving (in terms of utility) which the information makes possible. This type of definition of a measure of information was suggested, in the control context, by Bellman and Dreyfus. It differs from Shannon's definition in that the emphasis is on the use made of the information rather than on its transmission.

2.13 SOURCES

Most authors who have considered the control of uncertain systems do not explicitly distinguish those concepts which are independent of the notion of time. Therefore, the sources mentioned here can not be sharply separated from those listed in Chapter III.

We know of no serious attempt to give a precise definition of the concept of "feedback". This is astonishing, since a great number of definitions of "adaptive" have been suggested, and it is to be expected that the two concepts are strongly related. Among the proposals that have been made, the views of Zadeh [65] should be of interest to those persons whose feelings on the subject differ from ours.

In very recent times the view that controller design is just a special case of decision making has gained acceptance but the progress towards this realization was slow. The link between the statistical decision theorist and the control engineer was provided by the engineers concerned with filtering problems, from Zadeh [63] to Middleton [41]. This had to be so, because, unlike the control engineer, the filter designer could never ignore the presence of uncertainty.

Uncertainty entered control engineering by the stochastic door, with James, Nichols and Philips [30]. This led to a vast development for which we refer to the review paper of Kushner [35].

Feldbaum [17], [18] was among the first to point out clearly that the "dual" problems of identification and actuation of an uncertain plant had to be considered as one single problem. He also stated that the worst case (minimax) approach would be a most interesting alternative to the use of expectations, from an engineering point-of-view. Other workers [51] made the same remark, but also pursued the stochastic approach, mainly on the ground that the mathematical difficulties of minimax design were forbidding. This opinion may have to be revised, now that the full intricacies of stochastic control theory have come to light.

Macko [40] made the point that a design optimal under expectation will usually not be dominant (a fortiori, not strictly adaptive). Thus, when dominant designs do not exist, a decision problem does exist.

Sworder and Aoki [53] discuss the relations between control and decision theory and Sworder [54] applies Wald's procedure to discrete-time problems.

The minimax approach proper (without randomization) has been mostly considered from the point-of-view of game theory and with the assumption (not always verified) that a pure value exists, that is, that the gap is zero. This has been done from the point-of-view of filtering by Johansen [31], of sensitivity by Dorato and Kestenbaum [12] and of genuine conflicts of interest in the theory differential games founded by Isaacs [28].

The preliminary step of computing the worst case for a given design has been considered by Howard and Rekasius [27] and, in a sensitivity context, by Bellanger [4].

The important work of Warga [60] is the only one known to us which considers a minimax control problem without the zero-gap assumption.

The idea of measuring information by the gain in utility it can produce is given in Bellman and Dreyfus [7], Chapter VIII.

CHAPTER III

THE TIME FACTOR

3.1 HEURISTIC INTRODUCTION

Most control systems of interest evolve in time. The inputs u and outputs y are time-functions. An input u may also include initial actions: (e.g., the selection of initial conditions and design parameters) as well as final actions (e.g. the selection of the estimate in an identification problem). Such "once only" actions may be included in a time function description since the sets in which time functions take their values may be different at different times.

The main impact of the presence of the time variable results from causality. For each q in Q the system function S must be a non-predictive mapping of the time functions in U into the time functions in Y. Also the set Γ of controllers can only contain non-predictive mappings γ of Y into U.

Algorithms for optimization may take advantage of causality; this is known as dynamic programming. To apply the ideas of dynamic programming the notion of state is introduced. In the case of uncertain systems different types of states can be defined. All these definitions rely on causality, the basic idea being the following:

The state is a summary of the information available to some observer at some time, adequate for some future purposes.

Various purposes and various observers having access to different kinds of information may be considered, leading in each case to a different definition of a state.

The maximal information that the controller can have at a given time consists of:

1. all the a-priori information that was available to the designer.
2. the current value of the time variable.
3. the restriction of the time functions u and y to past and present times
4. the options that are still open for the choice of control in the future.

A controller may only possess in fact part of the maximal information, because of a limited ability to recall or use past data and because an accurate clock may not be available.

From the point of view of the controller the consideration of the uncertainty q as a time function is unnecessary. At any time t the available information leads to the assertion that q belongs to a subset of Q. Consideration of such subsets is the only meaningful way of describing the time evolution of the uncertainty for an observer located in the controller. As a mathematical device, useful for the development of optimization algorithms, one may consider a superobserver who has access to a great deal more information than the controller. For such an observer the state need not be an external function on the plant (in the sense of def. 2.8 of Chapter II), it can be dependent on the actual value of q. For such an observer it is meaningful to consider q as a function of time. For instance in a differential equation description

$$\dot{x}(t) = f(x(t), u(t), q(t), t)$$
$$y(t) = g(x(t), u(t), q(t), t)$$

$x(t)$ would be the state for such a superobserver, but not in general for an observer located in the controller.

Another basic question for the application of dynamic programming is the following: To what extent does the selection of the earlier values of the control $u(t)$ restrict the freedom in the choice of later values ? The most desirable situation is that there be no such restriction. When there is a restriction it can be eliminated by a reformulation of the problem. As an example consider the case where u is a sequence $u_1, \ldots, u_n$ of real numbers subject to $\sum_{i=1}^{n} u_i^2 \leq E$. Then one can reformulate the problem as one in which the input is a sequence of real numbers $\theta_1, \ldots, \theta_n$ independently restricted by $|\theta_i| \leq 1$, and use the substitution

$$u_i = \theta_i(E - \sum_{j=1}^{i-1} u_j^2)^{1/2}$$

When optimization is carried out by dynamic programming the minimization of the supercriterion $J(\gamma)$ is attempted, where J is defined for the entire control process. In doing this a sequence of subprocesses are optimized, where the subprocesses take place over a subinterval of time and are restricted by the information which has become available to the controller. Control policies which are not optimal for the subprocess may still lead to a minimization of $J(\gamma)$ for the whole process, and this possibility is especially often encountered when using the guaranteed performance evaluator. Indeed if at mid-time the controller has obtained information indicating that q lies in a particularly favorable subset then any control policy may be followed from that point on for which the value of K will not exceed the opvalue of the problem. There may be many such policies, even all possible policies. In dynamic programming, optimization for the subprocess is carried out using a conditional evaluator (for guaranteed performance: the supremum over the subset of Q). While doing so will not reduce the value of J for the whole process, the effect is to pick out of the optimal controller set Γ^* elements which are preferable under the partial order. This is extremely valuable, as has been pointed out before, because an improvement under the partial order is an improvement under <u>any</u> secondary rules of choice one might consider in addition to J.

Many of the questions arising in connection with time have not as yet been resolved in sufficient generality. In the present chapter we only discuss some of them at greater length and then proceed to the simple case of discrete-time problems with the guaranteed performance as evaluator.

3.2 IMPERFECT RECALL

Besides the causality condition, the set Γ of possible controllers may be limited by many practical considerations. One limitation arises when the controllers do not dispose of the maximal information about past events, that is, have imperfect recall.

Perfect recall means that any information on which the selection of $u(t_1)$ can be based can also be used to determine $u(t_2)$ whenever $t_2 > t_1$.

The simplest case is that of a single control station. All sensors and actuators have their signal-level terminals in a single location where the control computer (or operator) will be installed. In that situation imperfect recall would result from the limitations on memory and processing capacity of the computer (operator). Such a limitation is hard to assess because a control program γ can be implemented by retaining only those features of the past data which are known to be sufficient to define the future actions under any circumstances. In this way the memory requirements may sometimes be considerably reduced at the expense of some additional processing (reduction) of the data as it is received. Such is the case, for instance, for the Wiener-Kalman-Bucy controller.

In general it is therefore worthwhile to find a representation of the state of knowledge of the controller which is sufficient for purposes of optimization and provides an economy in data manipulation, both in the calculations to find optimal controllers and in their implementation. For the case of expected performance this question is known as the search for sufficient statistics and has received attention recently in the control context.

In any event, the memory and processing requirements are so difficult to express as manipulatable functions of γ that this type of restriction on implementation must be ignored in practice, at least in a first attack on the problem.

A far more basic limitation on the ability to recall past data arises when control is effected from several separate locations. For instance the control of a space mission may involve sensing and actuation from several stations, widely separated and in relative motion. The situation which thus arises is well known in game theory, where the crucial importance of the information pattern is fully recognized. In game theory, bridge is a two player game in which each "player" consists of two partners (control stations) which do not sense the same information (each senses only his hand, not his partner's). In the control case a computer, or a person able to follow a pre-arranged program, is available at each station. The design task consists in the selection of programs for all stations. Because the stations do not sense equivalent data (if only because of time delays) not every non-predictive function from the combined sensor outputs to the combined actuator inputs is realizable. Indeed such a function would make the actuation at one station depend on sensing at another.

To improve this situation, a communications system between the control stations will have been provided, but noise and time-delays limit its performance.

If there are n stations, station i will apply signal u_{ii} to its actuators and signals $u_{ij} (j = 1, \ldots, i-1, i+1, \ldots, n)$ to the inputs to the channels towards other stations; it will receive signal y_{ii} from its sensors and signals $y_{ij} (j=1, \ldots, i-1, i+1, \ldots, n)$ from the channel outputs.

Now consider the communication channels as part of the plant. One may always do so as a matter of convenience and one may have to, because plant and communications system may not be independent The ability of a spacecraft to communicate depends on its trajectory which may be precisely what is being controlled. Note also that sensing and actuation will themselves involve communication channels. From this point of view $u_i = (u_{i1}, \ldots, u_{in})$ is the input to the plant applied by station i and is given by $\gamma_i(y_i)$ where γ_i is the station's computer program and $y_i = (y_{i1}, \ldots, y_{in})$ is the plant output sensed by the station.

For the plant the input now is $u = (u_1, \ldots, u_n)$ and the output is $y = (y_1, \ldots, y_n)$, related by $y = S(u, q)$ via the system function S (which includes the time delays) and the uncertainty q (which includes the noise).

If Γ_i is the set of non-predictive γ_i at station i then the whole control program $\gamma = (\gamma_1, \ldots, \gamma_n)$ to be designed may be chosen, in the absence of further restrictions, in the cartesian product $\Gamma = \Gamma_1 \times \Gamma_2 \times \ldots \times \Gamma_n$.

When does this constitute a restriction? From a non-relativistic standpoint, one may compare Γ to the set Γ_0 of all non-predictive functions γ mapping Y into U by $u = \gamma(y)$. A restriction exists if for some u, $i \neq j$, $q_1 \neq q_2$ the outputs $S(u, q_1)$ and $S(u, q_2)$ differ in their i^{th} component but not in their j^{th} component.

For relativistic velocities the comparison set Γ_0 does not exist, since only the non-predictive character of the γ_i is physically defined and the times of the different stations are not comparable.

In the sequel only the case of perfect recall will be considered.

3.3 CLOCK UNCERTAINTY

The time variable t, in terms of which the a-priori plant description is given, is usually assumed available to the controller, by means of a perfect clock incorporated in the control computer. Such an assumption is eminently reasonable in the discrete-time case where the clock reduces to a counter. In the continuous-time case the clock must be an analog device and thus necessarily affected by some error. It may happen that this error is not negligible and this places an unusual type of restriction on the set Γ of realizable controllers.

Let the input set U consist of time functions u with range in the set Ω_u and let the output set Y consist of time functions y with range in a set Ω_y. Let t be a time at which a value u(t) is to be applied to the actuators and T^- the set of all past times at which outputs have been sensed.

Assume first that a perfect clock __is__ available. Then a non-predictive controller γ will produce u(t) as a function of the past output time function, i.e., as a function of

$$\{t, \{(\tau, y(\tau)) : \tau \in T^-\}\}$$

In particular, a blind controller will produce u(t) as a function of t. If the clock is not perfect, that is, if its reading is in an uncertain relationship with t, then the clock should be considered part of the plant and its reading part of the output data y. The controller then has no clock at all and this restricts the possible sets Γ of realizable controllers. Indeed the actuator input u(t) must now be produced as a function of the __set__ of all past outputs, i.e., as a function of

$$\{y(\tau) = \tau \in T^-\}$$

a subset of Ω_y.

In other words γ maps subsets of Ω_y (elements of its power set 2^{Ω_y}) into points in Ω_u. A blind controller subject to this restriction has not even access to an imperfect clock and can only apply inputs constant in time.

From the point of view of a superobserver who disposes of a perfect clock, any mapping of 2^{Ω_y} into Ω_u can be reduced to a non-predictive mapping of the set of time-functions Y into the set of time-function U. The source of the restriction is that the converse is false: not every non-predictive mapping of Y into U can be generated by a mapping of 2^{Ω_y} into Ω_u, as the case of blind controllers already shows.

In the sequel it will be assumed that a perfect clock is available.

3.4 TIMED CONTROL PROBLEMS

A satisfactory general theory for timed control problems is yet to be developed and this task appears replete with difficulties. In the sequel, only a few preliminary steps are taken in this direction. First the formal description of control problems, taking time into account, is considered. Such a description is essential to any further developments. Perfect recall and availability of a perfect clock will be assumed throughout.

Definition 3.1 A time set T is a nonempty totally ordered set.

Definition 3.2 A cut θ of the time-set T is a partition of T into two complementary order-convex subsets designated by θ^- and θ^+ with the elements of θ^- preceding those of θ^+. The set Ⓗ of all cuts of T is called the cut-set of T. It is totally ordered ($\theta_1 \leq \theta_2$ iff $\theta_1^- \subset \theta_2^-$) and order-complete.

Definition 3.3 An input range function ω_u is a function which associates a non-empty set $\omega_u(t)$ to every element t of a nonempty subset T_u of a time set T. T_u is called the input time set and the union of the sets $\omega_u(t)$ for all t in T_u is designated by Ω_u. The same definitions are made for output range functions and output time sets with the subscript u replaced by y.

Definition 3.4 The maximal input set for a given input range function ω_u is the set $U_{max}(\omega_u)$ of all functions u from T_u into Ω_u which satisfy.

$$(\forall t \in T_u) \ u(t) \in \omega_u(t)$$

Hence it is the cartesian product of the $\omega_u(t)$. A timed input set U is a nonempty subset of $U_{max}(\omega_u)$; its elements are called inputs. Corresponding definitions are made for a maximal output set $Y_{max}(\omega_y)$ and a timed output set Y.

Notation: For any cut θ the restriction of u to the domain $T_u \cap \theta^-$ is designated by u_θ^- and the restriction of u to the domain $T_u \cap \theta^+$ is designated by u_θ^+. A restriction of u to an empty domain is called an empty function. Since u is completely defined by u_θ^- and u_θ^+ we write

$$u = (u_\theta^-, u_\theta^+)$$

The same notation is used for outputs: $y = (y_\theta^-, y_\theta^+)$.

Definition 3.5 A timed plant $P = (T, U, Q, Y, S)$ consists of a time set T (with cut set Θ), a timed input set U (with input time set T_u and range function ω_u), a non-empty uncertainty set Q, a timed output set Y (with output time set T_y and range function ω_y) and a system function $S: U \times Q \rightarrow Y$ satisfying $(\forall q \in Q)\ (\forall u \in U)$ $(\forall \bar{u} \in U)\ (\forall \theta \in \Theta)\ u_{\theta}^{-} = \bar{u}_{\theta}^{-} \Rightarrow y_{\theta}^{-} = \bar{y}_{\theta}^{-}$ where $y = S(u, q)$ and $\bar{y} = S(\bar{u}, q)$. (Two empty functions are considered equal).

Every timed plant is also a plant in the sense of Chapter II and all definitions and theorems of that chapter are applicable. A valued timed plant is obtained when a criterion $K: U \times Q \rightarrow R_e$ is given. The criterion refers to the whole history of the process, causality does not in any way restrict its choice.

Definition 3.6 A function $\gamma : Y \rightarrow U$ is non-predictive iff $(\forall \theta \in \Theta)$ $(\forall y \in Y)\ (\forall \bar{y} \in Y)$

$y_{\theta}^{-} = \bar{y}_{\theta}^{-} \Rightarrow u_{\theta}^{-} = \bar{u}_{\theta}^{-}$

where $u = \gamma(y)$ and $\bar{u} = \gamma(\bar{y})$

The set of all non-predictive functions from Y into U is designated by Γ_{np}.

Note that a non-predictive function γ is not necessarily a controller and that a controller (as defined in Chapter II) is not necessarily non-predictive. All blind controllers are non-predictive which shows that $\Gamma_{np} \cap \Gamma(P)$ and a fortiori Γ_{np} are nonempty.

<u>Definition 3.7</u> Let P be a timed plant and θ a cut. Then the <u>truncated input sets</u> are defined by

$$U_\theta^- = \{u_\theta^- : u \in U\}$$

$$U_\theta^+ = \{u_\theta^+ : u \in U\}$$

and the <u>truncated output sets</u> by

$$Y_\theta^- = \{y_\theta^- : y \in Y\}$$

$$Y_\theta^+ = \{y_\theta^+ : y \in Y\}$$

<u>Definition 3.8</u> For γ in Γ_{np} define the truncation γ_θ^- of γ by

$$\gamma_\theta^- : Y_\theta^- \rightarrow U_\theta^-$$

$$\gamma_\theta^-(y_\theta^-) = u_\theta^-$$

whenever $\gamma(y) = u$.

The truncation of γ is uniquely defined because γ is non-predictive.

Definition 3.9 For a timed plant $P = (T, U, Q, Y, S)$ and a cut θ the truncated system function S_θ^- is defined by

$$S_\theta^- : U_\theta^- \times Q \rightarrow Y_\theta^-$$

$$S_\theta^- (u_\theta^-, q) = y_\theta^-$$

whenever

$$S(u, q) = y$$

The truncation of S is uniquely defined because S is non-predictive.

Definition 3.10 For a timed plant P, a cut θ, and an element γ of Γ_{np}, the truncated loop function L_θ^- is defined by

$$L_\theta^- : U_\theta^- \times Q \rightarrow U_\theta^-$$

$$L_\theta^- = \gamma_\theta^- \circ S_\theta^-$$

Definition 3.11 A timed controller for a plant P is a function γ from Y into U such that

1. $\gamma \in \Gamma_{np}$
2. $(\forall \theta \in Ⓗ)$ $(\forall q \in Q)$ the equation

$$u_\theta^- = L_\theta^- (u_\theta^-, q)$$

has exactly one solution (fixpoint) in U_θ^-. The dependence of this solution on θ and q is expressed by the timed input mapping $m_{\gamma\theta}: Q \rightarrow U_\theta^-$

$$u_\theta^- = m_{\gamma\theta}(q)$$

The set of all timed controllers for plant P is designated by $\Gamma_T(P)$

Clearly $\Gamma_T(P) \subset \Gamma_{np} \cap \Gamma(P)$, it is a proper subset in general and contains U, i.e., all blind controllers.

Definition 3.12 A timed control problem (P, K, Γ) consists of a timed plant P, a criterion K and a subset Γ of $\Gamma_T(P)$. For γ in Γ the pair (P, γ) is called a timed control system.

3.5 MARKOVIAN SETS

On the way towards definitions of states, with their property of separating past and future, it is necessary to require that the sets involved allow such a separation.

Definition 3.13 A subset Γ of Γ_{np} is markovian iff

$$(\forall \theta \in \textcircled{H})(\forall \gamma \in \Gamma)(\forall \bar{\gamma} \in \Gamma)$$

the function

$$\bar{\bar{\gamma}} : Y \to U$$

defined by

$$(\bar{\bar{\gamma}}(y))(t) = \begin{cases} (\gamma(y))(t) & \text{for } t \in T_u \cap \theta^- \\ (\bar{\gamma}(y))(t) & \text{for } t \in T_u \cap \theta^+ \end{cases}$$

belongs to Γ. An input set U is markovian if it satisfies the above definition when considered as a set of blind controllers.

The assumption of a markovian input set is important even in the derivation of the maximum principle.

Theorem 3.1 The intersection of a nonempty collection of markovian subsets of Γ_{np} is markovian.

Proof: For arbitrary fixed θ, if γ and $\bar{\gamma}$ belong to the intersection they belong to each of the sets, since each set is markovian $\bar{\bar{\gamma}}$ (see Def. 3.13) belongs to each of the sets, hence to their intersection. Q.E.D.

Theorem 3.2 If U is markovian Γ_{np} is markovian.

Proof: Consider γ, $\bar{\gamma}$ in Γ_{np} and a cut θ. Let $\bar{\bar{\gamma}}$ be constructed as in Def. 3.13. It must be shown that $\bar{\bar{\gamma}}$ is non-predictive. Let θ' be an arbitrary cut. For $\theta' \leq \theta$ $u^-_{\theta'} = \gamma^-_{\theta'}(y^-_{\theta'})$ because γ is non-predictive. For $\theta' > \theta$ $u^-_{\theta'}$ is composed of $u^-_\theta = \gamma^-_\theta(y^-_\theta)$, which is independent of y^+_θ and a fortiori of $y^+_{\theta'}$, and of the restriction of u to $\theta^+ \cap \theta'^- \cap T_u$ on which domain it equals $\gamma^-_{\theta'}(y^-_{\theta'})$ which is also independent of $y^+_{\theta'}$. Hence $\bar{\bar{\gamma}}$ is non-predictive. Q.E.D.

Theorem 3.3 $U_{max}(\omega_u)$ is always markovian.

Proof: The condition $(\forall t \in T_u)$ $u(t) \in \omega_u(t)$ places a restriction on u independently at every t in T_u. Since it is the only restriction $U_{max}(\omega_n)$ is markovian. Q.E.D.

3.6 VARIABLE END TIME

In an uncertain control system the time at which a certain event occurs may be dependent on the uncertainty q as well as on the input u applied. If this time is not an external function (Def. 2.8) then the controller is unable to determine its exact value. Even if it is an external function the exact value may only become known to the controller at some other time which can also be uncertain but is necessarily an external function.

The most important uncertain event is the termination of the process. The control process is considered to be terminated when further inputs have no more influence on the value of the criterion K. The controller may be unable to determine whether the process has terminated and γ is to be chosen taking account of this difficulty. While this problem is automatically included in the general description of the control problem as developed so far, it is useful to consider the properties of terminal time more closely.

Given u and q the corresponding <u>final cut</u> $\theta_f(u, q)$ is defined as the infinium, in the complete lattice Ⓗ, of all cuts θ which satisfy

$$(\forall \bar{u} \in U) \; \bar{u}_\theta^- = u_\theta^- \Rightarrow K(u, q) = K(\bar{u}, q)$$

Given γ and q the final cut $\theta_f(\gamma, q)$ is defined as $\theta_f(m_\gamma(q), q)$. Given γ only (including the case of blind controllers u) the latest final cut is defined as

$$\theta_\ell(\gamma) = \sup_{q \in Q} \theta_f(\gamma, q)$$

and the earliest final cut as

$$\theta_e(\gamma) = \inf_{q \in Q} \theta_f(\gamma, q)$$

When, for all γ in Γ, $\theta_e(\gamma) = \theta_\ell(\gamma)$ the terminal time is known, though variable. It is fixed if in addition it is independent of γ.

In any case the designer can determine $\theta_{f\ell} = \sup_{\gamma \in \Gamma} \theta_\ell(\gamma)$.

The problem data is then reduced to the consideration of $U^-_{\theta_{f\ell}}$, $Y^-_{\theta_{f\ell}}$, $S^-_{\theta_{f\ell}}$ and $\Gamma^-_{\theta_{f\ell}} = \{\gamma^-_{\theta_{f\ell}} : \gamma \in \Gamma\}$.

3.7 OUTLINE OF FURTHER DEVELOPMENTS

The development of the general theory of timed control problems will not be pursued further here. Nevertheless an outline of the application of dynamic programming can be given heuristically.

Consider a timed control problem (P, K, Γ) with Γ markovian, and an evaluator C. Let θ be a cut and assume that on θ^- a controller with truncation γ^-_θ was used. As a result some truncated output y^-_θ was observed. Assuming perfect recall and a perfect clock, the state of knowledge of the controller consists of

1. all a priori information, including the evaluator and Γ.
2. the cut θ
3. the pair $\sigma(\theta) = (u^-_\theta, y^-_\theta)$.

Let $Q_\theta = \{q: S_\theta^- (u_\theta^-, q) = y_\theta^-\}$. Because Γ is markovian it is still possible, for fixed γ_θ^-, to switch to any γ in Γ over θ^+. To each such choice of γ corresponds a valuation on Q_θ. Using a conditional evaluator, a number (supercriterion) can be obtained for each γ and the problem of its minimization is a timed control problem over θ^+. An optimal controller and the opvalue have to be determined. This is done for every $\sigma(\theta)$ which can result for any q and any selection of γ_θ^-.

In consequence a payoff function $W(\gamma_\theta^-, q)$ is defined by the corresponding opvalue of the conditional evaluator.

The choice of an optimal γ_θ^- is thereby reduced to another control problem on θ^- for which an evaluator is required. The appropriate evaluator might be called a truncated evaluator.

Since each of the control problems that arise in this fashion can be cut in turn, a recursive (i.e., dynamic programming) type of algorithm is obtained.

The factorization of C into a conditional and a truncated evaluator is trivial for the expected performance and for the guaranteed performance.

An expression for this factorization, valid for arbitrary evaluators, does not appear to be available, nor is it known whether C itself can always serve as truncated evaluator.

Detailed analysis of the case of expected performance has been actively pursued for some time. Far less is known at present for the case of guaranteed performance. The rest of the present work is devoted almost exclusively to this case: the "minimax" control of uncertain systems.

3.8 A DISCRETE TIME CASE

A specific type of timed control problem will now be considered. The optimization procedure for guaranteed performance by dynamic programming will then be described.

Let the succession of events be the following:

Application of input $u_1 \in \omega_u(1)$

Observation of output $y_1 \in \omega_y(1)$

Application of input $u_2 \in \omega_u(2)$

Observation of output $y_{n-1} \in \omega_y(n-1)$

Application of input $u_n \in \omega_u(n)$

The integer n is fixed. According to the well-established but rather unfortunate convention, the same index i is used for the input u_i and the later output y_i. Consequently i by itself is not the time. The time set T is the union of T_u and T_y which interlace. The n elements of T_u and the $n-1$ elements of T_y are labeled by overlapping sets of integers despite the fact that T_u and T_y are disjoint. Accordingly

$$U = \prod_{i=1}^{n} \omega_u(i)$$

$$Y = \prod_{i=1}^{n-1} \omega_y(i)$$

$$u = (u_1, \ldots, u_n)$$

$$y = (y_1, \ldots, y_{n-1})$$

Consider cuts $\theta_0, \theta_1, \ldots, \theta_{n-1}$ where cut θ_i is located before the application of u_{i+1} but after the observation of y_i. Hence θ_0 is the initial cut (θ_0^- is empty).

The notations will be

$$u_i^- = (u_1, \ldots, u_i) \in U_i^- = \prod_{j=1}^{i} \omega_u(j)$$

$$y_i^- = (y_1, \ldots, y_i) \in Y_i^- = \prod_{j=1}^{i} \omega_y(j)$$

Let Q be the uncertainty set.

The system function S is defined by $(S_1, \ldots, S_{n-1})$, that is, for $i = 1, \ldots, n-1$

$$S_i : U_i^- \times Q \rightarrow \omega_y(i)$$

$$y_i = S_i(u_1, \ldots, u_i, q) = S_i(u_i^-, q)$$

The criterion $K: U \times Q \rightarrow R_e$

has values $$K(u, q) = K(u_1, \ldots, u_n, q)$$

$$= K(u_{n-1}^-, u_n, q)$$

The set Γ of controllers will be the set Γ_{np} which is identical to $\Gamma_T(P)$ for this type of plant. Controllers γ are defined by $(\gamma_1, \ldots, \gamma_n)$ such that for $i = 1, \ldots, n$

$$\gamma_i : Y^-_{i-1} \rightarrow \omega_u(i)$$

$$u_i = \gamma_i(y_1, \ldots, y_{i-1}) = \gamma_i(y^-_{i-1})$$

Note that Y^-_0 consists of just one element: an empty function, so that $\gamma_1 = u_1 \epsilon \, \omega_u(1)$. Let Γ_i be the set of all functions γ_i which map Y^-_{i-1} into $\omega_u(i)$. Then $\Gamma = \Gamma_{np} = \prod_{i=1}^{n} \Gamma_i$.

Since the input set $U = U_{max}(\omega_u)$ it is markovian (Theorem 3.3), hence Γ is markovian (Theorem 3.2). The equality of Γ_{np} and $\Gamma_T(P)$ results from the fact that the solution of $u = \gamma(S(u, q))$ for given q is found by composition of a finite number of functions, a procedure which must lead to a unique result.

Indeed this solution proceeds recursively by

$$u_i = \gamma_i(y^-_{i-1})$$

$$u^-_i = (u^-_{i-1}, u_i)$$

$$y_i = S_i(u^-_i, q)$$

$$y^-_i = (y^-_{i-1}, y_i)$$

The truncated system function S^-_i is defined for $i = 1, \ldots, n-1$ by

$$S_i^- : U_i^- \times Q \rightarrow Y_i^-$$

$$y_i^- = S_i^-(u_i^-, q)$$

That is $S_i^- = (S_1, S_2, \ldots, S_i)$. The knowledge of the controller at cut θ_i consists of

a. all a priori information

b. the position of the cut

c. the data $\sigma_i = (u_i^-, y_i^-) \in U_i^- \times Y_i^-$

d. the controller $(\gamma_1, \ldots, \gamma_i)$ used so far.

The information in (a) is fixed by perfect recall. The information in (b) is implicitly contained in (c). The information in (d) is redundant for future purposes because

1. given (c) it provides no additional data about the plant
2. since Γ is markovian, the choice of γ_i in the past has no effect on the freedom in the choice of γ_i in the future.

For these reasons it is the information in (c) which characterizes the state of knowledge of the controller. Note that the assumption of a <u>markovian</u> set of controllers is essential to obtain this reduction.

Note that for the initial cut θ_0 the state of knowledge σ_0 consists of a pair of empty functions, that is, characterizes the absence of data.

Let 2^Q be the power set (set of all subsets) of Q.
Define the function $\hat{Q}_i : U_i^- \times Y_i^- \to 2^Q$.

$$\hat{Q}_i(\sigma_i) = \hat{Q}_i(u_i^-, y_i^-) = \{q \in Q : y_i^- = S_i^-(u_i^-, q)\}$$

Then σ_i must belong to the subset of $U_i^- \times Y_i^-$ defined by

$$\Sigma_i = \{\sigma_i \in U_i^- \times Y_i^- : \hat{Q}_i(\sigma_i) \text{ nonempty}\}$$

Note that $\Sigma_0 = \{\sigma_0\}$ and $\hat{Q}_0(\sigma_0) = Q$.

Let $$\Sigma = \bigcup_{i=0}^{n-1} \Sigma_i$$

and let $\hat{Q}: \Sigma \to 2^Q$

be defined by $\hat{Q}(\sigma) = \hat{Q}_i(\sigma)$

for σ in Σ_i.

Comments: 1. Many systems evolving continuously in time can be modeled by the above formalism. For instance, the input may be a continuous-time function while the output consists of samples delivered at a discrete set of times, fixed in advance. Then y_i is such a sample and T_y the set of sample times. As for u_i it consists of the portion of input function between two consecutive samples. An overlap at one time instant between u_i and y_i is immaterial for most systems. The case of variable, uncertain, terminal time is still within this framework provided the set $\theta_{f\ell} \cap T_y$ is finite; the cardinality of this set determines n.

Similarly, if the inputs are piecewise constant with discontinuities restricted to a discrete set T_u while the outputs are any type of continuous-time functions, the above description can be used.

2. Two states of knowledge of the controller, characterized by different data (u_i^-, y_i^-) may be equivalent as far as the calculation and implementation of optimal control is concerned. If such pairs are considered equivalent it is sufficient to know in which equivalence class the data is located. What this equivalence is depends on the structure of the criterion K and on the evaluator used.

For the case of expected performance, any function of the observed data, such that the value of this function determines the equivalence class in which the data is located serves as a sufficient statistic.

3.9 THE DYNAMIC PROGRAMMING ALGORITHM

To find the opvalue v and an optimal or ϵ-optimal controller γ^* for the problem described in the previous section, for the guaranteed performance evaluator, the method of dynamic programming can be used.

The objective is to replace the calculation of

$$v = \inf_{\gamma \in \Gamma} \sup_{q \in Q} W(\gamma, q)$$

by a large number of simpler problems of the same type.

In the sequel, whenever an infimum is determined, assume that its value is not $-\infty$ (that would be too good to be true). Select $\epsilon > 0$. Then there will be an element, in the set over which the infimum is taken, for which the value does not exceed the infimum by more than ϵ. Any such element will be called a minimizing element. Since the control process has a finite number n of steps, the controller constructed in this manner will have a guaranteed performance not exceeding the opvalue by more than $n\epsilon$. In short, one may act as if all infima were attained.

The algorithm proceeds as follows:

First Step: For every σ_{n-1} in Σ_{n-1} let

$$\sigma_{n-1} = (u^-_{n-1}, y^-_{n-1})$$

Define $$D_{n-1}: \Sigma_{n-1} \to R_e$$

by $$D_{n-1}(\sigma_{n-1}) = \inf_{u_n \epsilon \omega_u(n)} \sup_{q \epsilon \hat{Q}(\sigma_{n-1})} K(u^-_{n-1}, u_n, q)$$

$$g_n: \Sigma_{n-1} \to \omega_u(n)$$

$$g_n(\sigma_{n-1}) = \text{a minimizing } u_n$$

Second Step: For every σ_{n-2} in Σ_{n-2} let

$$\sigma_{n-2} = (u^-_{n-2}, y^-_{n-2})$$

Define $D_{n-2}: \Sigma_{n-2} \to R_e$

by $$D_{n-2}(\sigma_{n-2}) = \inf_{u_{n-1} \epsilon \omega_u(n-1)} \sup_{q \epsilon \hat{Q}(\sigma_{n-2})} D_{n-1}(\sigma_{n-1})$$

where $$\sigma_{n-1} = (u^-_{n-1}, y^-_{n-1})$$

$$u^-_{n-1} = (u^-_{n-2}, u_{n-1})$$

$$y^-_{n-1} = (y^-_{n-2}, S_{n-1}(u^-_{n-1}, q))$$

and let $$g_{n-1}: \Sigma_{n-2} \to \omega_u(n-1)$$

$$g_{n-1}(\sigma_{n-2}) = \text{a minimizing } u_{n-1}$$

Last (n^{th}) Step: $$v = \inf_{u_1 \epsilon \omega_u(1)} \sup_{q \epsilon Q} D_1(\sigma_1)$$

where $$\sigma_1 = (\omega_1, S_1(u_1, q))$$

$$g_1 = \text{a minimizing } u_1$$

The opvalue v has been found and an ($n\epsilon$-) optimal controller is implicitly defined by the g_i ($i = 1, \ldots, n$). The g_i are entirely adequate for implementation. If desired, the corresponding γ_i, dependent solely on the outputs, can be found by recursion:

$$\gamma_1 = g_1$$

$$\gamma_2(y_1) = g_2(\gamma_1, y_1)$$

$$\gamma_3(y_1, y_2) = g_3(\gamma_1, \gamma_2(y_1), y_1, y_2)$$

etc.,

The dynamic programming algorithm is especially advantageous when all sets involved are finite.

Since the sets Σ_i may be difficult to determine, it is worth noting that for the execution of the algorithm they may be replaced by any superset, at the expense of redundant calculations.

3.10 SUPEROBSERVER DESCRIPTION

The problem considered in Sections 3.8 and 3.9 was stated in the basic external description. Frequently an internal description is given from the point of view of a superobserver to whom q appears as an input time sequence $(q_1, \ldots, q_n)$. This superobserver is assumed to be still limited by causality so that his state of knowledge at cut θ_i includes (u_i^-, q_i^-). The uncertainty set Q is then another input set.

Such a description leads to the same dynamic programming algorithm if the structure of Q as a set of time functions is trivial, that is if $\omega_q(1) = Q$ and, for $i > 1$, $\omega_q(i)$ is a singleton. When the time structure is not trivial then simplifications result, provided that Q is a markovian input set. This means that for q and $\bar{q}$ in Q and arbitrary i the sequence $\bar{\bar{q}}$ with $\bar{\bar{q}}_j = q_j$ for $j \leq i$ and $\bar{\bar{q}}_j = \bar{q}_j$ for $j > i$ also belongs to Q.

It may be necessary to reformulate a given superobserver description to insure this property. This reformulation is analogous to the idea of "prewhitening" of stochastic filtering and control problems. More specifically, the concept of an uncertain time function which is only known to belong to a markovian set is analogous to the idea of white noise.

The sequence of events, as viewed by the superobserver, is the following:

system starts in fixed initial "internal state" $x_0 \in \omega_x(0) = \{x_0\}$

application of input $u_1 \epsilon \omega_u(1)$ by the controller

application of input $q_1 \epsilon \omega_q(1)$ by nature

transition to internal state $x_1 \epsilon \omega_x(1)$

production of output $y_1 \epsilon \omega_y(1)$ sent to the controller

application of input $u_2 \epsilon \omega_u(2)$

application of input $q_2 \epsilon \omega_q(2)$

transition to internal state $x_2 \epsilon \omega_x(2)$

production of output $y_2 \epsilon \omega_y(2)$

----

application of input $u_n \epsilon \omega_u(n)$

application of input $q_n \epsilon \omega_q(n)$

transition to internal state $x_n \epsilon \omega_x(n)$.

The internal states x_i represent the state of knowledge of the superobserver, who disposes of an accurate clock and enjoys perfect recall.

The same integers are used for indexing of all variables though the time sets T_x, T_u, T_q and T_y may be considered disjoint.

The input set from the controller $U = \prod_{i=1}^{n} \omega_u(i)$ and the input set from nature $Q = \prod_{i=1}^{n} \omega_q(i)$ are both markovian. The set Γ of possible controllers is the same markovian set Γ_{np} as in Section 3.8 because the controller has access only to u and y (as before) not directly to x.

Let the criterion $K: U \times Q \rightarrow R_e$ be defined as a function solely of x_n, the final internal state, by

$$K(u, q) = k(x_n)$$

$$k: \omega_x(n) \rightarrow R_e$$

This is a criterion because x_n is completely determined by the sequences u and q. The relationships describing the plant consist of an internal state transition equation.

$$x_i = f_i(u_i, q_i, x_{i-1})$$

$$f_i: \omega_u(i) \times \omega_q(i) \times \omega_x(i-1) \rightarrow \omega_x(i)$$

$$i = 1, \ldots, n$$

and an output equation

$$y_i = h_i(u_i, q_i, x_i)$$

$$h_i : \omega_u(i) \times \omega_q(i) \times \omega_x(i) \rightarrow \omega_y(i)$$

$$i = 1, \ldots, n-1$$

The appearance of u_i as argument for h_i corresponds to the possibility of a selection by the controller among different types of output. Since u_i also influences x_i, hence x_n, hence the value of the criterion, the selection of u_i which yields a more useful output may also lead to a higher cost. Thus this description includes the cost of measurement as a factor in the optimization. It may at times be preferable for the controller to select a u_i for which h_i is independent of its last two arguments, that is, decline the possibility of receiving output data because this choice will lead to a saving which is more valuable than the loss of information incurred.

Note that, given an external description as in Section 3.8, an internal description of the above type can always be trivially obtained by letting

$$\omega_q(1) = Q \text{ and } \omega_q(i) = \text{singleton for } i > 1$$

$$q_1 = q$$

$$\omega_x(i) = U_i^- \times Q_i^-$$

$$x_i = (u_i^-, q_i^-)$$

Then x_0 = absence of data

and x_n = (u, q) = arguments of K

so that $k \equiv K$

Also
$$f_i(u_i, q_i, x_{i-1}) = f_i(u_i, q_i, u_{i-1}^-, q_{i-1}^-)$$
$$= (u_i^-, q_i^-)$$

and
$$h_i(u_i, q_i, x_i) = h_i(u_i, q_i, u_i^-, q_i^-)$$
$$= S_i(u_i^-, q_i^-)$$

Of course, this identification produces only a formal change in the dynamic programming algorithm.

3.11 ALGORITHM USING INTERNAL DESCRIPTION

Consider cuts θ_i for $i = 0, \ldots, n-1$ placed just before the application of inputs u_{i+1} by the controller. At cut θ_i a controller with perfect recall disposes of the data $\sigma_i = (u_i^-, y_i^-)$ and in general this will not enable the unique determination of x_i. The case where x_i is always uniquely determined by σ_i will be considered later. Thus there will be a nonempty set $r_i \subset \omega_x(i)$ of all states x_i which can be reached by selecting any q in $\hat{Q}(\sigma_i)$, that is, compatible with the observations. Because U, Γ and Q are all markovian the knowledge of r_i rather than σ_i is already sufficient for the purpose of optimization.

Therefore, one introduces the sets $\omega_r(i)$ where for each $i, \omega_r(i)$ is a collection of subsets of $\omega_x(i)$ sufficiently large to include all r_i that can occur. One may always insure thi by taking $\omega_r(i)$ as the power set (collection of all subsets) of $\omega_x(i)$ but much more restricted collections are usually sufficient when the detailed structure of the problem is taken into account. For example, if $\omega_x(i)$ is a topological space it may be sufficient to consider only compact subsets; if it is also a linear space, compact and convex subsets may be sufficient and then r_i is characterized by its support function which can be more readily manipulated.

The set r_i plays a role analogous to that of the conditional probability distribution of x_i given σ_i in the case of expected performance.

Of course $$r_o = \{x_o\} = \omega_x(0) \text{ and } \omega_r(0) = \{r_o\}$$

To proceed to the dynamic programming algorithm note first that r_o is known and that r_i can be obtained recursively from r_{i-1}, u_i and y_i. Let the compatibility functions c_i be defined for $i=1, \ldots, n-1$ by

$$c_i: \omega_r(i-1) \times \omega_u(i) \times \omega_y(i) \rightarrow \omega_r(i)$$

$$c_i(r_{i-1}, u_i, y_i) = \{f_i(u_i, q_i, x_{i-1}) \in \omega_x(i) : q_i \in \omega_q(i),$$
$$x_{i-1} \in r_{i-1}, \; y_i = h_i(u_i, q_i, f_i(u_i, q_i, x_{i-1}))\}$$

Then $$r_i = c_i(r_{i-1}, u_i, y_i) \quad \text{for } i = 1, \ldots, n-1$$

Note that the compatibility functions are monotone in the sense that $A \subset B$ implies

$$c_i(A, u_i, y_i) \subset c_i(B, u_i, y_i)$$

Now the algorithm can be defined.

First Step: For every r_{n-1} in $\omega_r(n-1)$ define

$$D_{n-1} : \omega_r(n-1) \rightarrow R_e$$

$$D_{n-1}(r_{n-1}) = \inf_{u_n \epsilon \omega_u(n)} \sup_{q_n \epsilon \omega_q(n)} \sup_{x_{n-1} \epsilon r_{n-1}} k(f_n(u_n, q_n, x_{n-1}))$$

and $$g_n(r_{n-1}) = \text{a minimizing } u_n$$

Second Step: For every r_{n-2} in $\omega_r(n-2)$ define

$$D_{n-2} : \omega_r(n-2) \rightarrow R_e$$

$$D_{n-2}(r_{n-2}) = \inf_{u_{n-1} \epsilon \omega_u(n-1)} \sup_{q_{n-1} \epsilon \omega_q(n-1)} \sup_{x_{n-2} \epsilon r_{n-2}} D_{n-1}(c_{n-1}(r_{n-2}, u_{n-1}, y_{n-1}))$$

where $$y_{n-1} = h_{n-1}(u_{n-1}, q_{n-1}, f_{n-1}(u_{n-1}, q_{n-1}, x_{n-2}))$$

and $$g_{n-1} : \omega_r(n-2) \rightarrow \omega_u(n-1)$$

$$g_{n-1}(r_{n-2}) = \text{a minimizing } u_{n-1}$$

Last (n^{th}) Step: Define

$$v = \inf_{u_1 \epsilon \omega_u(1)} \quad \sup_{q_1 \epsilon \omega_q(1)} \quad D_1(c_1(r_0, u_1, y_1))$$

where $y_1 = h_1(u_1, q_1, f_1(u_1, q_1, x_0))$

and let g_1 = a minimizing u_1

Then v is the opvalue of the problem and the g_i define the optimal controller.

If desired, the functions γ_i can be recovered recursively by

$$\gamma_1 = g_1$$

$$\gamma_2(y_1) = g_2(c_1(r_o, \gamma_1, y_1))$$

$$\gamma_3(y_1, y_2) = g_3(c_2(c_1(r_o, \gamma_1, y_1), \gamma_2(y_1), y_2)$$

etc..,

Note that the functions D_i are monotone set functions in the sense that $A \subset B$ implies $D_i(A) \leq D_i(B)$. This algorithm is of a more complex structure than that of Section 3.9 and it requires manipulation of the sets r_i. Its advantage is that the functions f_i, h_i, c_i, k depend only on data at a single time index while in Section 3.9 the functions K, S_i^- and Q_i have truncated sequences as arguments.

The present algorithm becomes especially interesting if simplifying features are present in the problem; these special cases are considered in the next two sections of this chapter.

3.12 CASE OF INDEPENDENT OUTPUT UNCERTAINTY

Assume, for the problem described in Section 3.10, that for each $i = 1, \ldots, n$

$$\omega_q(i) = \omega_q^x(i) \times \omega_q^m(i)$$

where $\omega_q^m(n)$ is a singleton. Then the elements q_i are pairs (q_i^x, q_i^m).

Assume also that the state equation involves only q_i^x and the output equation only q_i^m.

$$x_i = f_i(u_i, q_i^x, x_{i-1})$$

$$y_i = h_i(u_i, q_i^m, x_i)$$

Then we say that the output uncertainty is independent of the state uncertainty. This case arises in particular when the h_i do not depend on q_i at all (though the h_i still need not be one-one with respect to x_i at fixed u_i).

Simplifications of the dynamic programming algorithm are now possible.

Define the reachability function $\rho_i: \omega_r(i-1) \times \omega_u(i) \rightarrow \omega_r(i)$ by

$$\rho_i(r_{i-1}, u_i) = \{f_i(u_i, q_i^x, x_{i-1}) \epsilon \omega_x(i) : q_i^x \epsilon \omega_q^x(i), x_{i-1} \epsilon r_{i-1}\}$$

and the measurement function $m_i : \omega_u(i) \times \omega_y(i) \rightarrow \omega_r(i)$ by

$$m_i(u_i, y_i) = \{x_i \epsilon \omega_x(i) : (\exists q_i^m \epsilon \omega_q^m(i))\ y_i = h_i(u_i, q_i^m, x_i)\}$$

Then the definition of the compatibility function c_i reduces to

$$c_i(r_{i-1}, u_i, y_i) = \rho_i(r_{i-1}, u_i) \cap m_i(u_i, y_i)$$

This expression can be substituted for every appearance of c_i in Section 3.11. Furthermore each appearance of $\sup_{q_i \epsilon \omega_q(i)}$ can be replaced by $\sup_{q_i^x \epsilon \omega_q^x(i)} \sup_{q_i^m \epsilon \omega_q^m(i)}$. Thus one obtains a further subdivision of the total task into subtasks such as the separate determination of ρ_i and m_i followed by intersection and the separate extremization over the sets $\omega_q^x(i)$ and $\omega_q^m(i)$.

Note also that the expression for y_i in the algorithm reduces to

$$y_i = h_i(u_i, q_i^m, f_i(u_i, q_i^x, x_{i-1}))$$

A tremendous simplification occurs when the sets $M_i(u_i, y_i)$ never contain more than one point. This case is the subject of the following section.

3.13 CASE OF INTERNAL STATE OUTPUT

Assume, for the problem described in Section 3.10, that for any i and any arguments the value of the compatibility function c_i is a set of at most one point. Then for all arguments that can actually occur the value of c_i is a set of exactly one point (a singleton).

This case is obtained when, in the problem of Section 3.12, the values taken by the measurement functions m_i are always singletons. Then, without loss of generality, one may assume that the output equation is simply

$$y_i = x_i$$

that is $$h_i(u_i, q_i^m, x_i) \equiv x_i$$

Under these assumptions the dynamic programming algorithm is considerably simplified and requires only operations with the functions k, f_i and ρ_i.

The reachability function ρ_i is now reduced to

$$\rho_i : \omega_x(i-1) \times \omega_u(i) \rightarrow \omega_r(i)$$

$$\rho_i(x_{i-1}, u_i) = \{f_i(u_i, q_i, x_{i-1}) \epsilon \omega_x(i) : q_i \epsilon \omega_q(i)\}$$

The superscript x on q_i^x has become redundant. The algorithm now proceeds as follows.

<u>First Step</u>: Define $D_{n-1} : \omega_x(n-1) \rightarrow R_e$ by

$$D_{n-1}(x_{n-1}) = \inf_{u_n \epsilon \omega_u(n)} \sup_{q_n \epsilon \omega_q(n)} k(f_n(u_n, q_n, x_{n-1}))$$

$$= \inf_{u_n \epsilon \omega_u(n)} \sup_{x_n \epsilon \rho_n(x_{n-1}, u_n)} k(x_n)$$

and $$g_n : \omega_x(n-1) \rightarrow \omega_u(n)$$

by $$g_n(x_{n-1}) = \text{a minimizing } u_n$$

<u>Second Step</u>: Define $D_{n-2} : \omega_x(n-2) \rightarrow R_e$ by

$$D_{n-2}(x_{n-2}) = \inf_{u_{n-1} \epsilon \omega_u(n-1)} \sup_{q_{n-1} \epsilon \omega_q(n-1)} D_{n-1}(f_{n-1}(u_{n-1}, q_{n-1}, x_{n-2}))$$

$$= \inf_{u_{n-1} \epsilon \omega_u(n-1)} \sup_{x_{n-1} \epsilon \rho_{n-1}(x_{n-2}, u_{n-1})} D_{n-1}(x_{n-1})$$

and $$g_{n-1} : \omega_x(n-2) \to \omega_u(n-1)$$

by $$g_{n-1}(x_{n-2}) = \text{a minimizing } u_{n-1}$$

Last (n^{th}) Step: Define

$$v = \inf_{u_1 \epsilon \omega_u(1)} \sup_{q_1 \epsilon \omega_q(1)} D_1(f_1(u_1, q_1, x_0))$$

$$= \inf_{u_1 \epsilon \omega_u(1)} \sup_{x_1 \epsilon \rho_1(x_o, u_1)} D_1(x_1)$$

and $$g_1 \epsilon \omega_u(1)$$

by $$g_1 = \text{a minimizing } u_1$$

Then v is the opvalue of the problem and the g_i define directly the controller functions γ_i by

$$\gamma_i(y_1, \ldots, y_{i-1}) = \gamma_i(x_1, \ldots, x_{i-1}) = g_i(x_{i-1})$$

3.14 INTRODUCTION OF STRUCTURE

Thus far in Chapter 1, 2 and 3 no assumptions on the structure (linear, topological, metric, etc.) of the sets involved has been made. As a consequence no meaningful assumptions on the structure of the functions involved (continuity, linearity, convexity, etc.) could be made either.

Though it is believed that the discussion of the more fundamental ideas should be made without the introduction of irrelevant structure assumptions, it is nonetheless clear that further development of the theory towards application relies entirely on such assumptions.

Hence in the remainder of this work a very narrow class of heavily structured problems will be considered more in detail. This class will be described in Chapter V and is a special case of the class of problems discussed in Section 3.13. In Chapter V an investigation of the reachable sets of linear differential systems and of their support functions is presented. It establishes a bridge between a familiar type of problem and the class considered in Chapter V.

3.15 SOURCES

The application of dynamic programming to the control of uncertain systems evolving in discrete time was proposed in a fundamental paper of Bellman and Kalaba [5], for expected performance, without assumption of perfect recall.

A more detailed application of these ideas is available in Bellman [6], Bellman and Dreyfus [7], and Dreyfus [13].

Among the important recent works on the stochastic discrete-time case are those of Striebel [52], Astrom [1] and, above all, Dynkin [14].

Most recently this subject has also been approached from the point of view of mathematical programming, by Van Slyke and Wets [57] and Wets [61].

For the stochastic continuous-time case the developments begun by Kalman and Bucy [32] and Florentin [20] have led to the results summarized by Kushner [35].

As for the application of dynamic programming to games evolving in time (the "extensive form"), it goes back at least to the classical proof that finite games of perfect information have a saddle-point. For further developments see the treatment of recursive games by Everett [16], of Markov games by Zachrisson [62] and of differential games by Isaacs [28, 29].

For other approaches, see the references discussed at the end of Chapters II and V.

The requirements for sets of admissible inputs stated by Pontryagin et al. [45], Chapter 1, imply a markovian property of these sets.

CHAPTER IV

REACHABLE SETS OF LINEAR DIFFFRENTIAL SYSTEMS

4.1 INTRODUCTION

The minimax control of linear systems, in fixed time, with end point criteria and sampled output of the state will be shown later to reduce to a geometric minimax problem involving reachable sets. Among the reachable sets of interest are the sets of states reachable at a fixed time, under various types of constraints on inputs and initial conditions. These sets are convex and compact for many types of constraints and thus fully described by their support functions. Even when the sets are only compact their convex hulls are sufficient to solve some control problems.

Once a fundamental matrix of the linear system is known, application of the theory of moments or of the maximum principle yields the support functions of the reachable sets with little additional effort.

4.2 THE BASIC MAPPING

Consider the vector system (n components):

$$\dot{x} = A(t)x + f(t) \qquad x(t_0) = \xi \tag{1}$$

and let T be given, $t_o < T < \infty$. A fundamental matrix $\Phi(t)$ is a non-singular solution of

$$\dot{\Phi}(t) = A(t)\Phi(t) \tag{2}$$

Absolutely continuous solutions of Eq. 2 exist for locally integrable A(t) and are non-singular throughout if non-singular for some t. If $\Phi(t)$ is a fundamental matrix, every other fundamental matrix is of the form $\Phi(t)C$ where C is a constant non-singular n by n matrix. The unique transition matrix is

$$\psi(t, \tau) = \Phi(t)\Phi^{-1}(\tau)$$

Then Eq. 1 may be replaced by

$$x(T) = \Phi(T) \left[\Phi^{-1}(t_0)\xi + \int_{t_0}^{T} \Phi^{-1}(\tau)f(\tau)d\tau\right] \quad (3)$$

Note that Eq. 3 makes sense when f is integrable on $[t_0, T]$ and that functions $f(t)$ equal almost everywhere are equivalent as far as Eq. 3 is concerned. Thus we take f to mean an equivalence class of integrable functions. Then Eq. 3 defines a linear mapping of pairs (ξ, f) belonging to $R^n \times [L_1[t_0, T]]^n$ into elements $x(T)$ of R^n. Note that for $1 < p < q$

$$L_1[t_0, T] \supset L_p[t_0, T] \supset L_q[t_0, T]$$

so that Eq. 3 makes sense a foriori for (equivalence classes of) functions f whose components belong to L_p for $p > 1$.

Let $\mathcal{D}$ denote the linear space $R^n \times [L_1[t_0, T]]^n$, the domain of the mapping defined by Eq. 3. Then the problem of reachable sets at time T may be viewed as follows.

Given a set γ in $\mathcal{D}$, a "constraint set", find the image of γ in R^n under Eq. 3. We will concentrate on finding the support function of this image, i.e., the closure of its convex hull.

4.3 SUPERPOSITION OF INDEPENDENT CONSTRAINTS

It may happen that the elements (ξ, f) appear as sums of several such elements, each independently constrained. This means that the constraint set γ is the vector sum of a finite number of sets γ_i each defined independently of the others.

Then, since the mapping is linear, the image of γ is the vector sum of the images of the γ_i and one may determine these separately.

This is particularly convenient in terms of support functions since the support function of a vector sum of sets is the sum of the individual support functions.

Note in this connection that the constraints on f need not always be independent of those on ξ.

4.4 DUALITY

Let us restrict our attention to the subspace $\mathcal{D}_2$ of $\mathcal{D}$

$$\mathcal{D}_2 = R^n \times [L_2[t_0, T]]^n$$

Let the inner product of elements (ξ_1, f_1), (ξ_2, f_2) of $\mathcal{D}_2$ be defined by

$$< (\xi_1, f_1), (\xi_2, f_2) > = < \xi_1, \xi_2 > + \int_{t_0}^{T} < f_1(t), f_2(t) > dt \qquad (4)$$

then $\mathcal{D}_2$ is a Hilbert space and may be identified with its dual.

The basic mapping may be written

$$x = L(\xi, f)$$

and L^* will denote the adjoint of L, i.e., for x in R^n

$$< x, L(\xi, f)> = < L^*x, (\xi, f)>$$

from which

$$L^*x = (\Phi'^{-1}(t_0)\Phi'(T)x, \Phi'^{-1}(t)\Phi'(T)x) \qquad (5)$$

Let γ be a constraint set in $\mathcal{D}_2$. Its support function is defined for $(\hat{\xi}, \hat{f})$ in $\mathcal{D}_2$ by

$$G(\hat{\xi}, \hat{f}) = \sup_{(\xi, f)\epsilon\gamma} < (\hat{\xi}, \hat{f}), (\xi, f) >$$

The support function of the corresponding reachable set is

$$\begin{aligned} H(p) &= \sup_{(\xi, f)\epsilon\gamma} < p, L(\xi, f) > \\ &= \sup_{(\xi, f)\epsilon\gamma} < L^*p, (\xi, f)> \\ &= G(L^*p) \qquad (6) \end{aligned}$$

Thus the support function of the reachable set follows by substitution from the support function of the constraint set.

4.5 HARD CONSTRAINTS AND THE MAXIMUM PRINCIPLE

By "hard constraints" is meant a constraint set γ of the form

$$\gamma = \{(0, \phi(u, t)): u(t)\epsilon\,\Omega \text{ a.e in } [t_0, T]\}$$

where Ω is a compact set in R^r

u are measurable functions from $[t_0, T]$ into R^r

ϕ is a continuous function from $R^r x[t_0, T]$ into R^n

By a theorem of Neustadt the corresponding reachable set is compact and convex.

Thus for any n-vector q the maximum of $\langle q, x(T)\rangle$ subject to the constraint γ is attained for some time function u with corresponding function x and by the maximum principle, a corresponding function p satisfying

$$\begin{cases} \dot{p}(t) = -A'(t)p(t) & p(T) = q \\ \dot{x}(t) = A(t)x(t) + \phi(u(t), t) & x(t_0) = 0 \end{cases} \quad (7)$$

The hamiltonian

$$H = \langle p(t), A(t)x(t)\rangle + \langle p(t), \phi(u(t), t)\rangle$$

has the maximum value

$$\begin{aligned} H_{max} &= \langle p(t), A(t)x(t)\rangle + \max_{\omega \in \Omega} \langle p(t), \phi(\omega, t)\rangle \\ &= \langle p(t), A(t)x(t)\rangle + \sigma(p(t), t) \end{aligned} \quad (8)$$

where $\sigma(p, t)$ is the support function of the set $\phi(\Omega, t)$ in R^n.

then $\frac{d}{dt} \langle p(t), x(t)\rangle = \sigma(p(t), t)$ (9)

and, since $x(t_0) = 0$

$$\langle p(t), x(t)\rangle = \int_{t_0}^{t} \sigma(p(\tau), \tau) d\tau$$

at time T

$$\begin{aligned} \langle p(T), x(T)\rangle = \langle q, x(T)\rangle &= \max_{x \in \gamma} \langle q, x\rangle \\ &= h(q, T) \end{aligned}$$

the support function of the reachable set at time T, evaluated at argument q.

If Φ is a fundamental matrix for A, integration of the adjoint equations gives

$$p(t) = \Phi'^{-1}(t)\Phi'(T)q \tag{10}$$

Thus the final result is

$$h(q,T) = \int_{t_0}^{T} \sigma(\Phi'^{-1}(t)\Phi'(T)q, t)dt \tag{11}$$

which completely describes the reachable set since it is known to be compact and convex.

In the case where $\phi(u,t) = B(t)u$ with an n by r matrix $B(t)$ let $\omega(q)$ be the support function in R^r of the set Ω to which $u(t)$ is constrained. Then the support function $\sigma(q,t)$ of the set $B(t)\Omega$ is, by Eq. 6

$$\sigma(q,t) = \omega(B'(t)q)$$

and Eq. 11 takes the form

$$h(q,T) = \int_{t_0}^{T} \omega(B'(t)\Phi'^{-1}(t)\Phi'(T)q)dt \tag{12}$$

If Ω is defined by $N(u) \leq \rho$ where N is a norm in R^r, let $N^*(\cdot)$ designate the dual norm of norm N. Then

$$\omega(q) = \rho N^*(q)$$

and Eq. 12 becomes

$$h(q,T) = \rho \int_{t_0}^{T} N^*(B'(t)\Phi'^{-1}(t)\Phi'(T)q)dt \tag{13}$$

If A and B are independent of t Eq. 12 becomes

$$h(q, T) = \int_{t_0} \omega(B'e^{A'(T-t)}q)dt \tag{14}$$

Finally if $r = 1$ and Ω is defined by $|u| \leq \rho$

$$h(q, T) = \rho \int_{t_0}^{T} |b'e^{A'(T-t)}q| \, dt \tag{15}$$

4.6 THE DUAL HAMILTON-JACOBI EQUATION

Taking formula 11 as a starting point, note that along a line in (q, T) space, that is, a solution of the adjoint equations

$$\frac{dq}{dT} = -A'(T)q \tag{16}$$

h has a total derivati

$$\frac{Dh(q, T)}{DT} = \sigma(q, T) \tag{17}$$

This suggests that we are faced with the characteristics of a partial differential equation.

Indeed, provided the derivatives exist, differentiation of Eq. 11 yields

$$\frac{\partial h(q, T)}{\partial T} = \sigma(q, T) + q'A(T) \frac{\partial h(q, T)}{\partial q} \tag{18}$$

Since, by Eq. 8,

$$H_{max}(p, x, t) = \sigma(p, t) + p'A(t)x \tag{19}$$

we may write Eq. 18 as

$$\frac{\partial h(p, t)}{\partial t} = H_{max}(p, \frac{\partial h(p, t)}{\partial p}, t) \tag{20}$$

which is the dual of the Hamilton-Jacobi equation

$$\frac{\partial J(x,t)}{\partial t} = -H_{max}\left(\frac{\partial J(x,t)}{\partial x}, x, t\right) \tag{21}$$

which governs the value function J.

While Eq. 20 is integrated forward in time, from the support function of the initial states, Eq. 21 is to be integrated backward in time from the terminal cost function.

Just as Eq. 21 may be obtained as the limiting form of the Bellman equation

$$J(x, t-\Delta t) = \max_{u \in \Omega} J(x+(A(t)x+\phi(u,t))\Delta t, t) \tag{22}$$

equation 20 can be obtained from the dual Bellman equation

$$h(p, t+\Delta t) = h(p+\Delta t A'(t)p, t) + \max_{u \in \Omega} p'\phi(u,t)\Delta t \tag{23}$$

Notice that the duality between Eq. 20 and 21 is analogous to the duality between the backward and forward Kolmogorov equations for stochastic systems.

4.7 HARD CONSTRAINTS AND DUALITY

In the Hilbert space $\mathcal{D}_2$ consider the subspace $\xi = 0$ which is itself a Hilbert space with inner product

$$<\hat{f}, f> = \int_{t_0}^{T} <\hat{f}(t), f(t)> dt$$

Since γ is a set in this subspace, its support function evaluated at an element $\hat{f}$ is given by

$$G(\hat{f}) = \sup_{f \in \gamma} <\hat{f}, f>$$

$$= \sup_{f \in \gamma} \int_{t_0}^{T} <\hat{f}(t), f(t)> dt$$

$$= \sup_{u(\cdot)} \int_{t_0}^{T} <\hat{f}(t), \phi(u(t), t)> dt$$

$$\leq \int_{t_0}^{T} \max_{u \in \Omega} <\hat{f}(t), \phi(u, t)> dt$$

$$= \int_{t_0}^{T} \sigma(\hat{f}(t), t)\, dt \qquad (24)$$

the maximum integrand at any time t is attained for some $u(t)$ since Ω is compact. Either the function $u(t)$ defined in this way is measurable or it can be approximated by a measurable $u(t)$ with range in Ω so that Eq. 24 exceeds $G(\hat{f})$ by less than ϵ. Thus equality holds and

$$G(f) = \int_{t_0}^{T} \sigma(\hat{f}(t), t)\, dt \qquad (25)$$

and by the substitution of Eq. 6 one obtains the support function of the reachable states. The adjoint operator L^* is multiplication by $\Phi'^{-1}(t)\ \Phi'(T)$. Therefore one obtains Eq. 11 in yet another way, by replacement of $f(t)$ by

$$L^* p = \Phi'^{-1}(t)\Phi'(T)p$$

4.8 INITIAL CONDITION CONSTRAINTS

In case $\gamma = \{(\xi, 0): \xi \epsilon \Omega_0\}$ where Ω_0 is a compact, convex set in R^n with support function $h_0(p)$, one may apply directly the duality formula, Eq. 6, to obtain

$$h(p, T) = h_0(\Phi'^{-1}(t_0)\Phi'(T)p) \tag{26}$$

When initial condition constraints are combined with independent hard constraints, i.e.,

$$\gamma = \{(\xi, \phi(u, t)): \xi \epsilon \Omega_0,\ u(t) \epsilon \Omega \text{ a.e.}\}$$

superposition applies and the addition of the support functions for the separate constraints yields

$$h(p, T) = h_0(\Phi'^{-1}(t_0)\Phi'(T)p) + \int_{t_0}^{T} \sigma(\Phi'^{-1}(t)\Phi'(T)\, p, t)\, dt \tag{27}$$

When the derivatives exist, this may be viewed as the solution of the dual Hamilton-Jacobi equation from the initial condition $h(p, t_0) = h_0(p)$.

4.9 ENERGY CONSTRAINTS

By an energy constraint, is meant the subset of $\mathscr{D}_2$

$$\gamma = \{(B_0\mu,\ B(t)u(t)): \mu' Q_0 \mu + \int_{t_0}^{T} u'(t)Q(t)u(t)dt \leq \rho^2\} \tag{28}$$

where Q_0 is a constant symmetric positive definite r' by r'

matrix, B_0 a constant n by r' matrix, B(t) an n by r matrix, Q(t) a symmetric r by r matrix positive definite a.e. for t in $[t_0, T]$, μ has r', u has r components and $\rho > 0$ is a constant. Assume the elements of B and Q are in $L_\infty[t_0, T]$.

Since positive definite matrices have positive definite square roots, one may effect the change of variables

$$\begin{cases} \mu = Q_0^{-1/2}\nu \\ u(t) = Q(t)^{-1/2}v(t) \end{cases} \tag{29}$$

Then γ may be written

$$\gamma = \{(B_0 Q_0^{-1/2}\nu,\ B(t)Q(t)^{-1/2}v(t)) : \nu'\nu + \int_{t_0}^{T} v'(t)v(t)dt \leq \rho^2\} \tag{30}$$

i.e., (ν, v) is constrained to a sphere of radius ρ about the origin of a Hilbert space.

The mapping of this sphere into reachable states is described by

$$\begin{aligned} x(T) &= \Phi(T)\Phi^{-1}(t_0)B_0Q_0^{-1/2}\nu \\ &\quad + \int_{t_0}^{T} \Phi(T)\Phi^{-1}(t)B(t)Q^{-1/2}(t)v(t)dt \\ &= L(\nu, v) \end{aligned} \tag{31}$$

The adjoint L^*, applied to an n-vector p is

$$L^*p = (Q_0^{-1/2}B'_0\Phi'^{-1}(t_0)\Phi'(T)p, Q^{-1/2}(t)B'(t)\Phi'^{-1}(t)\Phi'(T)p) \tag{32}$$

The support function of the unit sphere in Hilbert space is just the norm; for the sphere of radius ρ it is

$$G(\hat{\nu}, \hat{v}) = \rho(\hat{\nu}'\hat{\nu} + \int_{t_0}^{T} \hat{v}'(t)\hat{v}(t)dt)^{1/2} \tag{33}$$

The support function of the reachable set follows from Eqs. 32, 33 by the duality substitution, Eq. 6:

$$\begin{aligned} h(p, T) &= \rho[p'\Phi(T)\Phi^{-1}(t_0)B_0Q_0^{-1}B_0'\Phi'^{-1}(t_0)\Phi'(T)p \\ &\quad + \int_{t_0}^{T} p'\Phi(T)\Phi^{-1}(t)B(t)Q^{-1}(t)B'(t)\Phi'^{-1}(t)\Phi'(T)p\,dt]^{1/2} \\ &= \rho(p'M(T)p)^{1/2} \end{aligned} \tag{34}$$

where

$$\begin{aligned} M(T) &= \Phi(T)\Phi^{-1}(t_0)B_0Q_0^{-1}B_0'\Phi'^{-1}(t_0)\Phi'(T) \\ &\quad + \int_{t_0}^{T} \Phi(T)\Phi^{-1}(t)B(t)Q^{-1}(t)B'(t)\Phi'^{-1}(t)\Phi'(T)dt \end{aligned} \tag{35}$$

Differentiating Eq. 35 one obtains a matrix differential equation for the symmetric, non-negative definite matrix $M(T)$.

$$\frac{dM(T)}{dT} = A(T)M(T) + M(T)A'(T) + B(T)Q^{-1}(T)B'(T)$$

with
$$M(t_0) = B_0Q_0^{-1}B_0' \tag{36}$$

Using this fact, differentiation of Eq. 34 gives a partial differential equation for h(p, T)

$$\frac{\partial h}{\partial T} = \frac{\rho}{2h} p'B(T)Q^{-1}(T)B'(T)p+p'A(T)\frac{\partial h}{\partial p} \qquad (37)$$

or for its square

$$\frac{\partial h^2}{\partial T} = \rho\, p'B(T)Q^{-1}(T)B'(T)p + p'A(T)\frac{\partial h^2}{\partial p} \qquad (38)$$

with initial condition

$$h(p, 0) = \rho(p'B_o Q_0^{-1} B_0' p)^{1/2} \qquad (39)$$

Since the mapping of Eq. 31 is linear and completely continuous the reachable set, image of a Hilbert space sphere, is compact and convex.

When M(T) is positive definite, the reachable set is the ellipsoid

$$\{x: x'M^{-1}(T)\, x \leq \rho\} \qquad (40)$$

When M(T) is singular, the reachable set is confined to the range subspace of M(T). In this subspace M(T) has an inverse and the reachable set is the flattened ellipsoid described by Eq. 40 restricted to this subspace.

Controllability means that the reachable set for unconstrained u, starting with $\xi = 0$, is R^n. Equivalent to this is to say that for $\rho > 0$ $B_0 = 0$ and $Q(t) = I$ the reachable set has an interior, i.e., det $M(T) \neq 0$. Since $\Phi(T)$ is nonsingular this reduces to

$$\det \int_{t_0}^{T} \Phi^{-1}(t)B(t)B'(t)\Phi'^{-1}(t)\, dt \neq 0 \qquad (41)$$

a well-known result.

4.10 SOURCES

The idea of viewing control problems as a propagation of reachable sets may be found in Halkin [25]. A proof of compactness and convexity of such sets is given by Neustadt [44]. For the maximum principle, see Athans and Falb [2] and the references quoted therein. For the usual definition of controllability, see Kalman, Ho and Narendra [33]. For the mathematical background concerning duality, convexity and support functions see the sources given in Chapter V. The application of the theory of moments to such problems is discussed, for instance, by Beckenbach and Bellman [3].

CHAPTER V

VECTOR ADDITION GAMES

5.1 INTRODUCTION

In the present chapter the problem with internal state output of Section 3.13 is considered again and, for the first time in this work, specific assumptions on structure are made. Essentially, a form of linearity is assumed. The problems satisfying these assumptions are called vector addition games. They can arise from a variety of classical control situations and they are reducible to a simple canonic form. Their solution depends on the properties of the criterion such as convexity or homogeneity.

5.2 VECTOR ADDITION GAMES

The problem of Section 3.13 is a vector addition game if it has the following structure.

Let $L_x(i)$ for $i=0,1,\dots,n$

$L_\mu(i)$ for $i=1,\dots,n$

$L_v(i)$ for $i=1,\dots,n$

be real linear spaces.

Let $\omega_x(i)$ for $i=0,\dots,n$ be a subset of $L_x(i)$.

One may always for $i>0$ take $\omega_x(i) = L_x(i)$, while $\omega_x(0)$ is a singleton in $L_x(0)$. The space $L_x(n)$ will also be designated by the abbreviated notation L.

Let μ_i and ν_i for $i=1,\ldots,n$ be functions

$$\mu_i: \omega_u(i) \to L_\mu(i)$$

$$\nu_i: \omega_q(i) \to L_\nu(i)$$

Let A_i, B_i and C_i for $i=1,\ldots,n$ be <u>linear</u> functions

$$A_i: L_x(i-1) \to L_x(i)$$

$$B_i: L_\mu(i) \to L_x(i)$$

$$C_i: L_\nu(i) \to L_x(i)$$

Let the state transition equation be

$$x_i = A_i x_{i-1} + B_i \mu_i(u_i) + C_i \nu_i(q_i)$$

As in Section 3.13 the criterion is given by $K(u,q) = k(x_n)$ where $k:L \to R$. The output equation is $y_i = x_i$ for $i=1,\ldots,n-1$.

The problem is to find a controller in Γ_{np} to optimize guaranteed performance.

5.3 LINEAR DIFFERENTIAL SYSTEMS WITH END-POINT CRITERIA

A first example of a vector addition game is the following.

Let $t_o,\ldots,t_n$ be an increasing sequence of real numbers.

Consider the differential equation for an m-vector valued function ξ on $[t_o, t_n]$

$$\dot{\xi}(t) = H(t)\,\xi(t) + \phi_u(u(t),t) + \phi_q(q(t),t) \quad \text{with} \quad \xi(t_o) = x_o$$

In this differential equation

H is an m by m matrix of integrable functions on $[t_o, t_n]$.

u is an r-vector of bounded measurable functions on $[t_o, t_n]$, selected from a markovian input set U.

q is an r'-vector of bounded measurable functions on $[t_o, t_n]$ selected in a markovian set Q.

ϕ_u and ϕ_q are continuous m-vector valued functions of their arguments.

Assume that outputs $y_i = \xi(t_i)$ are delivered at time t_i for $i-1, \ldots, n-1$ to a controller which selects the restriction u_i of u to $(t_{i-1}, t_i]$ on the basis of $(y_i, \ldots, y_{i-1})$.

The functions q are uncertain and the controller is to be designed to optimize guaranteed performance with the criterion $K(u,q) = k(\xi(t_n))$.

This problem can be reduced to a vector addition game as follows:

Let ψ be the transition matrix corresponding to H.

Let $x_i = \xi(t_i)$ for $i=0, \ldots, n$

u_i = restriction of u to $(t_{i-1}, t_i]$ for $i=1, \ldots, n$.

$\omega_u(i)$ = set of all restrictions u_i for u in U.

q_i = restriction of q to $(t_{i-1}, t_i]$ for $i=1, \ldots, n$.

$\omega_q(i)$ = set of all restrictions q_i for q in Q.

$L_x(i) = R^m$ for $i=0, \ldots, n$.

$L_u(i) = L_q(i) = (L_\infty(t_{i-1}, t_i))^m$ for $i=1, \ldots, n$.

$\mu_i(u_i)$ = restriction of $\phi_u(u(t), t)$ to $(t_{i-1}, t_i]$

$\nu_i(q_i)$ = restriction of $\phi_q(q(t), t)$ to $(t_{i-1}, t_i]$

Then the transition equation and the linear functions A_i, B_i, and C_i are given by

$$x_i = \psi(t_i, t_{i-1})x_{i-1} + \int_{t_{i-1}}^{t_i} \psi(t_i, \tau)\phi_u(u(\tau), \tau)d\tau$$

$$+ \int_{t_{i-1}}^{t_i} \psi(t_i, \tau)\phi_q(q(\tau), \tau)d\tau$$

where the first integral could be written

$$\int_{t_{i-1}}^{t_i} \psi(t_i, \tau)\,[\mu_i(u_i)]\,(\tau)d\tau$$

and similarly for the second integral.

5.4 LINEAR DIFFERENTIAL SYSTEM WITH GENERAL CRITERIA

5.4.1 If the criterion, in the problem of Section 5.3, depends also directly on u, say by way of an integral, then the running value of this integral can be taken as an additional component of ξ and the problem is still of the same type. This reduction, the Bolza to Mayer reduction, is well known in the calculus of variations. Unfortunately it has the effect of hiding in the dynamics of the system many useful properties that the criterion may have, preventing one from taking advantage of such properties. For this reason an apparently much more elaborate alternative way to reduce such a case to a vector addition game is of interest.

Let $\quad x_i = (\xi(t_i), u_1, \ldots, u_i)$

so that $$L_x(i) = R^m \times \prod_{j=1}^{i} (L_\infty(t_{j-1}, t_j])^r$$

$$= R^m \times (L_\infty (t_o, t_i])^r$$

The transitions of x_i are described by the equation of Section 5.3 together with the adjunction of u_i, a linear operation.

Then $x_n = (\xi(t_i), u_1, \ldots, u_n) = (\xi(t_i), u)$ and the criterion k may be any function on $L = L_x(n) = R^m \times (L_\infty(t_o, t_n])^r$.

The output equation $y_i = x_i$ is still valid because of perfect recall.

5.4.2 If the criterion depends also on the time-function ξ, say by way of an integral of a function of $\xi(t)$, $u(t)$ and t, then the Bolza → Mayer reduction would destroy the linearity of the differential equations. It is thus even more advisable to proceed differently. Two cases must be distinguished.

a) k depends only on ξ by way of the values $\xi(t_i)$. Then, because of perfect recall, it suffices to take as internal state

$$x_i = (\xi(t_o), \ldots, \xi(t_i), u_1, \ldots, u_i)$$

that is the state of knowledge of the controller. The linearity is trivially preserved and

$$x_n = (\xi(t_o), \ldots, \xi(t_n),\ u_1, \ldots, u_n)$$

contains all the arguments for the criterion.

b) k depends on the entire function ξ.

Then the reduction to a vector addition game is impossible unless the physical assumptions are slightly modified.

Suppose that at time $t_i (i=1,\ldots,n-1)$ the controller receives, in one fell swoop, the restriction of ξ to $(t_{i-1}, t_i]$. One may imagine a recorder which records on successive sheets, out of view of the controller, and ejects each sheet, corresponding to the time interval (t_{i-1}, t_i) at time t_i that is when the sheet is full. The state of knowledge of the controller at time t_i then includes the restriction of both u and ξ to $[t_o, t_i]$ while the system still has discrete time internal state output.

Such a situation may seem strange but can be considered a good approximation to reality when the sampling times are very closely spaced. The interpretation is then that a continuous output recording of ξ is available but the designer is restricted to the use of controllers which "look" at this output recording only every so often.

Under these conditions a reduction to a vector addition game is again possible.

Let s_i = restriction of ξ to $[t_o, t_i]$

and $x_i = (s_i, u_1, \ldots, u_i)$

Then x_i is the state of knowledge of the controller and may be taken as the output. Linearity is preserved and $x_n = (s_n, u_1, \ldots, u_n) = (\xi, u)$ contains the arguments required for k.

5.5 SINGLE STAGE CASE

For n=1 there are no outputs. This may also be called the open-loop case. The problem is defined just by the criterion K(u, q) and optimization is over a set of blind controllers. Conversely any problem with an outputless plant may be considered as a problem with internal state output (sic) but n=1. It is a vector addition game if one can write

$$K(u,q) = k(B\mu(u) + C\nu(q))$$

indices having become superfluous.

Thus the only linearity requirement is that one have superposition of the effects of u and q in the argument of k. This is a very weak requirement as the following example shows.

Let ξ and η be two vector valued functions on $[t_o, t_1]$ not necessarily with the same number of components, solutions, for given initial conditions, of

$$\dot{\xi}(t) = f_1(\xi(t), u(t), t)$$

$$\dot{\eta}(t) = f_2(\eta(t), q(t), t)$$

where u is a time function selected by the controller in a set U, which need not be markovian and q is an uncertain time function from set Q. Both differential equations may be nonlinear.

Now let

$$K(u,q) = k(\xi, \eta, u, q)$$

that is k may depend on all time functions involved. Then a vector addition problem exists because

a) the linear space L is the product of the function spaces of ξ, η, u and q.

b) $(\xi, \eta, u, \xi) = (q, 0, u, 0) + (0, \eta, 0, q)$ is the space L.

c) $(\xi, 0, u, 0)$ depends only on u while $(0, \eta, 0, q)$ depends only on q.

This general observation leads to innumerable special cases. For instance k might only depend on $\xi(t_i)$ and $\eta(t_i)$. Then

$$(\xi(t_i), \eta(t_i)) = (\xi(t_i), 0) + (0, \eta(t_i))$$

where the first term is determined by u and the second by q, which in effect choose points in the reachable sets at time t_i of the two differential equations. Furthermore one can let $\xi(t_o)$ depend on u and $\eta(t_o)$ on q without destroying the superposition.

5.6 THE CANONIC FORM

A canonic form for vector addition games is obtained by the simple device of extrapolating all effects to time t_n.

The following change of variables is carried out:

$$\bar{x}_i = A_n A_{n-1} \ldots A_{i+1} x_i$$

in particular $\bar{x}_n = x_n$

$$\bar{u}_i = A_n \ldots A_{i+1} B_i \mu_i(u_i)$$

$$\bar{q}_i = -A_n \ldots A_{i+1} C_i \nu_i(q_i)$$

hence $\bar{\omega}_u(i) = A_n \ldots A_{i+1} B_i \mu_i(\omega_u(i))$

and $\bar{\omega}_q(i) = -A_n \ldots A_{i+1} C_i \nu_i(\omega_q(i))$

Then the transition equation becomes

$$\bar{x}_i = \bar{x}_{i-1} + \bar{u}_i - \bar{q}_i$$

where all variables are elements of L for all i. The criterion is given by

$$k(x_n) = k(\overline{x}_n) = k(\overline{x}_o + \sum_{i=1}^{n} (\overline{u}_i - \overline{q}_i))$$

Since the controller has a priori knowledge of the linear functions A_i it can determine $\overline{y_i} = \overline{x_i}$ from the observation of $y_i = x_i$. The knowledge of $\overline{y_i}$ is sufficient for optimization because the transformed problem is still a vector addition game and a fortiori a problem with internal state output and markovian input sets.

The minus sign in the definition of $\overline{q}_i$ is chosen because of the intuitive appeal of the notion of distance in the case where the function k is a norm.

5.7 EXAMPLE OF REDUCTION TO CANONIC FORM

In the case of linear differential systems with end-point criteria, as considered in Section 5.3, the reduction to canonic form is obtained as follows.

The space L is R^m and all new variables are elements of R^m. They are defined by

$$\overline{x}_i = \psi(t_n, t_i)\xi(t_i) = \psi(t_n, t_i)\, x_i \quad \text{for } i=0,1,\ldots n.$$

$$\overline{u}_i = \psi(t_n, t_i) \int_{t_{i-1}}^{t_i} \psi(t_i, \tau)\phi_u(u(\tau), \tau)\, d\tau$$

$$\bar{q}_i = -\psi(t_n, t_i) \int_{t_{i-1}}^{t_i} \psi(t_i, \tau)\phi_q(q(\tau), \tau)d\tau$$

hence $\bar{\omega}_u(i) = \psi(t_n, t_i)\bar{\bar{\omega}}_u(i)$

where $\bar{\bar{\omega}}_u(i)$ is the reachable set at time t_i for

the system $\dot{\xi}(t) = H(t)\xi(t) + \phi_u(u(t), t)$

with initial condition $\xi(t_{i-1}) = 0$

given that the restriction of u to $(t_{i-1}, t_i]$ is constrained

to belong to $\omega_u(i)$.

Similarly $\bar{\omega}_q(i) = -\psi(t_n, t_i)\bar{\bar{\omega}}_q(i)$

where $\bar{\bar{\omega}}_q(i)$ is the reachable set at time t_i for the system

$\dot{\xi}(t) = H(t)\xi(t) + \phi_q(q(t), t)$ with $\xi(t_{i-1}) = 0$

and the restriction of q to $(t_{i-1}, t_i]$ is constrained to $\omega_q(i)$.

A great deal is therefore known about the sets $\bar{\omega}_u(i)$ and $\bar{\omega}_q(i)$. Their compactness, convexity, symmetry, support functions can be determined by the results of classical optimal control theory and duality as in Chapter IV.

In the case of Section 5.4 the space L is infinite dimensional and in order to obtain sets $\bar{\omega}_u(i)$ which are convex it becomes necessary to assume that the set U of input functions is itself convex in $(L_\infty[t_0, t_n])^r$ and that ϕ_u is linear in u.

5.8 THE GAME-THEORETIC INTERPRETATION

From now on, assume that the problem is given directly in the canonic form, so that the overbars become unnecessary.

The determination of the opvalue v is equivalent to the determination of the (pure) value of the following game.

Move 1: minimizing player selects u_1 in $\omega_u(1)$

Move 2: maximizing player selects q_1 in $\omega_q(1)$

Move 3: minimizing player selects u_2 in $\omega_u(2)$

etc., ...

The payoff is $k(x_0 + \sum_{i=1}^{n} (u_i - q_i))$

The game is of perfect information: at every move the player who must act has knowledge of the selections made at all previous moves.

The determination of optimal controllers is equivalent to the determination of an optimal strategy for the minimizing player. Hence the name "vector addition game".

The determination of the lopvalue v' for the control problem is equivalent to the determination of the pure value of another game of perfect information, with only 2 instead of 2n moves, and the same payoff function.

Move 1: maximizing player selects sequence $(q_1, \ldots, q_n)$

with $q_i \in \omega_q(i)$

Move 2: minimizing player selects sequence $(u_1, \ldots, u_n)$

with $u_i \in \omega_u(i)$

Another number of interest is the value v'' of the game of perfect information, with $2n$ moves, played as follows:

Move 1: maximizing player selects q_1 in $\omega_q(1)$

Move 2: minimizing player selects u_1 in $\omega_u(1)$

Move 3: maximizing player selects q_2 in $\omega_q(2)$

etc., ...

By the interchange inequality $v' \leq v'' \leq v$.

If one considers the game of imperfect information, with n moves, played as follows:

Move i: Both players select simultaneously, one u_i in $\omega_u(i)$, the other q_i in $\omega_q(i)$, knowing the selections made at all previous moves

then v is the upper value and v'' the lower value of this game.

5.9 THE DYNAMIC PROGRAMMING ALGORITHM

The solution of the problem by dynamic programming is straightforward in principle. It can be organized as follows:

First Step: Define $E_n: L \rightarrow R_e$

$D_{n-1}: L \rightarrow R_e$

$g_n: L \rightarrow \omega_u(n)$

by

$$E_n(x) = \sup_{q \in \omega_q(n)} k(x-q)$$

$$D_{n-1}(x) = \inf_{u \in \omega_u(n)} E_n(x+u)$$

$$g_n(x) = \text{a minimizing } u$$

Second Step: Define $E_{n-1}: L \to R_e$

$D_{n-2}: L \to R_e$

$g_{n-1}: L \to \omega_u(n-1)$

by

$$E_{n-1}(x) = \sup_{q \in \omega_q(n-1)} D_{n-1}(x-q)$$

$$D_{n-2}(x) = \inf_{u \in \omega_u(n-1)} E_{n-1}(x+u)$$

$$g_{n-1}(x) = \text{a minimizing } u$$

Last (n^{th}) Step: Define $E_1: L \to R_e$

$v \in R_e$

$g_1 \in \omega_u(1)$

by

$$E_1(x) = \sup_{q \in \omega_q(1)} D_1(x-q)$$

$$v = \inf_{u \in \omega_u(1)} E_1(x_0+u)$$

$$g_1 = \text{a minimizing } u$$

Then v is the opvalue and the g_i define the optimal controller, by $u_1 = g_1$ and $u_i = g_i(x_{i-1}) = g_i(y_{i-1})$ for $i = 2, \ldots, n$.

5.10 THE CONSERVATION OF CONVEXITY

It frequently happens that the function k is convex (it might be a norm or a positive definite quadratic form) and that the sets $\omega_u(i)$ are convex (as reachable sets of linear systems).

Under these conditions the functions D_i and E_i will likewise be convex, that is, convexity is conserved in the recursive procedure of Section 5.9.

Indeed, assume D_i is convex, then for $0 \leq \theta \leq 1$

$$E_i(\theta x + (1-\theta)y) = \sup_{q \epsilon \omega_q(i)} D_i(\theta x + (1-\theta)y - q)$$

$$= \sup_{q \epsilon \omega_q(i)} D_i(\theta(x-q) + (1-\theta)(y-q))$$

$$\leq \sup_{q \epsilon \omega_q(i)} (\theta D_i(x-q) + (1-\theta) D_i(y-q))$$

$$\leq \sup_{q \epsilon \omega_q(i)} \theta D_i(x-q) + \sup_{q \epsilon \omega_q(i)} (1-\theta) D_i(y-q)$$

$$= \theta E_i(x) + (1-\theta) E_i(y)$$

Now assume E_i is convex, then for $0 \leq \theta \leq 1$

$$D_{i-1}(\theta x + (1-\theta)y) = \inf_{u \epsilon \omega_u(i)} E_i(\theta x + (1-\theta)y + u)$$

Since $\omega_u(i)$ is convex $\omega_u(i) = \theta\omega_u(i) + (1-\theta)\omega_u(i)$. Hence

$$D_{i-1}(\theta x + (1-\theta)y) = \inf_{u_1 \epsilon \omega_u(i)} \inf_{u_2 \epsilon \omega_u(i)} E_i(\theta x + (1-\theta)y + \theta u_1 + (1-\theta)u_2)$$

$$= \inf_{u_1 \epsilon \omega_u(i)} \inf_{u_2 \epsilon \omega_u(i)} E_i(\theta(x+u_1) + (1-\theta)(y+u_2))$$

$$\leq \inf_{u_1 \epsilon \omega_u(i)} \inf_{u_2 \epsilon \omega_u(i)} (\theta E_i(x+u_1) + (1-\theta) E_i(y+u_2))$$

$$= \theta D_{i-1}(x) + (1-\theta) D_{i-1}(y)$$

Now since k is convex, the convexity of all functions D_i and E_i follows by recursion. QED

Note that the sets $\omega_q(i)$ need not be convex. Similarly if k is concave and the sets $\omega_q(i)$ are convex then all functions E_i and D_i are concave and the sets $\omega_u(i)$ need not be convex for this result.

5.11 THE CONSERVATION OF UNIFORM CONTINUITY

Suppose the function $\phi: L \to R$ has the property that, for all x and y in L

$$|k(x) - k(y)| \leq \phi(x-y)$$

then, for all i,

$$|E_i(x) - E_i(y)| \leq \phi(x-y)$$

and

$$|D_i(x) - D_i(y)| \leq \phi(x-y)$$

Assume

$$|D_i(x) - D_i(y)| \leq \phi(x-y)$$

then

$$D_i(y) - \phi(x-y) \leq D_i(x) \leq D_i(y) + \phi(x-y)$$

Replace x by x-q and y by y-q

$$D_i(y-q) - \phi(x-y) \leq D_i(x-q) \leq D_i(y-q) + \phi(x-y)$$

Take supremum over q in $\omega_q(i)$

$$E_i(y) - \phi(x-y) \leq E_i(x) \leq E_i(y) + \phi(x-y)$$

or

$$|E_i(x) - E_i(y)| \leq \phi(x-y)$$

and this last relation implies $|D_{i-1}(x) - D_{i-1}(y)| \leq \phi(x-y)$ by an entirely similar argument. Hence the claim follows by recursion.

Note that it is not necessary that the sets $\omega_u(i)$ and $\omega_q(i)$ or the function k be convex.

The function ϕ need not be even $(\phi(-x) = \phi(x))$ but, if it is not, it can be replaced by $\min(\phi(-x), \phi(x))$ or by $\frac{1}{2}(\phi(-x) + \phi(x))$, which are.

In case ϕ is a norm, one has the conservation of Lipschitz continuity: whenever k is Lipschitz continuous, with constant λ, with respect to some norm, the functions E_i and D_i are Lipschitz continuous with the same constant with respect to this norm.

The same result holds for Hölder continuity, by letting ϕ be a power of a norm.

More generally, if ϕ can be interpreted as a modulus of uniform continuity with respect to a norm, then the result says that E_i and D_i are uniformly continuous with the same modulus.

Finally if k is a norm, hence convex and Lipschitz continuous with respect to itself with constant 1, then the functions E_i and D_i are Lipschitz continuous with constant 1 with respect to norm k. Furthermore, if the sets $\omega_u(i)$ are convex, E_i and D_i are also convex, by conservation of convexity.

5.12 WHY A MINIMAX PRINCIPLE DOES NOT HOLD

The determination of the opvalue is equivalent to that of the pure value and saddle-point of a game of perfect information, as pointed out in Section 5.8.

In the continuous time case, the theory of differential games provides necessary conditions, which may be called a minimax principle, for saddle-points. These conditions are very similar to the maximum principle for one-sided optimization. On the other hand, discrete time maximum principles are available in the one-sided case, especially with convex constraint sets. Thus one would expect that a discrete minimax principle would apply to the solution of vector addition games, at least with convex constraint sets.

The adjoint equation for the canonic form of a vector addition game is simply $p_{i-1} = p_i$ that is a constant costate, because the state is constant for zero inputs.

The reason for the failure of this approach is the following. In differential games singular surfaces arise, on which the costate undergoes a jump. In the discrete time game, there is the possibility of a jump between any two consecutive moves. Thus the costate equation is really $p_{i-1} = p_i + \lambda_i$ where λ_i is the jump. The occurrence of jumps and the corresponding values of λ_i depend on the behavior of the solution in the large and cannot be determined by local variational techniques. Thus the λ_i are additional unknowns for which there are no simple equations. Hence the costate is completely undetermined and the minimax principle is vacuous.

One case in which a minimax principle can be shown to hold, is the case where the function k is linear. This essentially trivial case can be handled even more easily by duality as in the sequel. To see how the singularities arise, note that the points x for which $E_i(x) = \alpha$ are just those points (if any) for which the set

$x - \omega_q(i)$ is just contained in the set $S = \{x: D_i(x) \leq \alpha\}$ that is the boundaries are in contact. Now the singularity arises for any x for which the boundary of $x - \omega_q(i)$ has two or more contacts, from the inside, with the boundary of S. Even for convex S this is a common occurrence, it leads to a corner in the locus $E_i(x) = \alpha$, and the presence of this corner is impossible to detect by local analysis.

5.13 THE USE OF DUALITY

The sets $\omega_u(i)$ and $\omega_q(i)$ are often closed and convex but they may be known only by their support functions. Let s_i be the support function of $\omega_u(i)$ and σ_i the support function of $\omega_q(i)$.

In case the function k is convex, in particular if it is a norm, the problem can be brought to a dual form.

Note that if k is a monotone increasing function of a convex function it is sufficient to solve the problem for the convex function because extremization commutes with monotone functions.

To simplify the exposition of the duality transformations, the following assumptions and conventions are made.

1. L is assumed finite-dimensional
2. Extended real valued convex functions are used when required, though we do not stop to justify their use. Suffice it to say that the necessary mathematical apparatus has been developed. The occurrence of infinite values is not a complication but a simplification (just as the occurrence

of zero values is often a simplification). When infinite values occur certain technical requirements must be met, they will be tacitly assumed.

The case of infinite dimensional L requires a more sophisticated mathematical apparatus which is still under active development in the current literature. It will not be considered here.

If the sets $\omega_u(i)$ and $\omega_q(i)$ are bounded then the set

$$x_o + \sum_{i=1}^{n} \omega_u(i) - \sum_{i=1}^{n} \omega_q(i)$$

which contains all arguments of k that matter, is also bounded.

Now for k convex and real-valued on L, it follows that k is continuous on L and is Lipschitz continuous on any bounded set. Hence, by conservation of Lipschitz continuity and convexity all the functions E_i and D_i are Lipschitz continuous on any bounded set, convex and real-valued. In particular, every infinum is a minimum and every supremum is a maximum since the constraint sets were assumed closed.

Note also that the support function of a bounded set is convex, real valued, positively homogeneous and Lipschitz continuous with respect to the dual norm with a constant equal to the maximum of the norm on the set.

The practical import of Lipschitz continuity is the following. To compute the maximum of a continuous function on a compact set to finite accuracy in finite time is impossible without additional information. If the function is Lipschitz continuous with a known constant λ then the maximum lies between α and $\alpha + \frac{\epsilon}{\lambda}$ where α is the maximum of the function on a finite ϵ-net. Since compact sets have finite ϵ-nets for all $\epsilon > 0$, that is, are totally bounded, the extremization can at least always be carried out by brute force.

The use of variational techniques is inadequate because there are no local sufficient conditions for a maximum of convex function on a convex set. Many local maxima are the rule.

The duality transformations are based on the following concepts.

Let L^* be the dual of L, the set of all linear real valued functions on L.

For p in L^* and x in L, (x, p) or (p, x) designates the value of p at x.

The support function $s: L^* \to R_e$ of a set A in L is defined by

$$s(p) = \sup_{a \in A} (a, p)$$

The Fenchel transform $k^f: L^* \to R_e$ of a function $k: L \to R_e$ is defined by

$$k^f(p) = \sup_{x \in L} ((x, p) - k(x))$$

k^f is always convex. If k is convex then $(k^f)^f = k$. If k is not convex then $(k^f)^f$ is the "convexification" of k: the supremum of all convex functions which do not exceed k at any point.

Thus if k is convex, it has the duality representation

$$k(x) = \sup_{p \in L^*} \left((x, p) - k^f(p) \right)$$

If k is a norm, its dual norm k^* on L^* is defined by

$$k^*(p) = \sup_{x \in B} k(x)$$

where B is the unit ball $\{x \in L: k(x) \leq 1\}$.

The dual unit ball is $B^* = \{p \in L^*: k^*(p) \leq 1\}$. One has $(k^*)^* = k$ and the representation

$$k(x) = \sup_{p \in B^*} (x, p)$$

If k is a pseudo-norm (because there may be components about which we do not care) then the dual k^* is an extended real valued norm. B is a cylinder while B^* is flattened into a subspace. Thus the sup in the representation of k(x) is on a lower dimensional set, a great simplification.

Note that k^* is the Fenchel transform of the function which equals 0 on B and $+\infty$ elsewhere. More generally, the support function of a set is the Fenchel transform of the function which equals 0 on the set and $+\infty$ elsewhere, the indicator function of the set.

Convention: In the sequel the sets over which the dummy variable of an extremization ranges are not indicated when it is clear which set is meant.

5.14 SINGLE-STAGE DUALITY

In the single-stage (open-loop) case, the opvalue v and lopvalue v' are given by

$$v = \inf_u \sup_q k(x_o + u - q)$$

$$v' = \sup_q \inf_u k(x_o + u - q)$$

This may be written

$$v = \inf_u E(x_o + u)$$

where

$$E(x) = \sup_q k(x - q)$$

But $k(x-q) = \sup_p \left[(x - q, p) - k^f(p)\right]$

$$E(x) = \sup_q \sup_p \left[(x, p) - (q, p) - k^f(p)\right]$$

$$= \sup_p \left[(x, p) + \sigma(-p) - k^f(p)\right]$$

where σ is the support function of set ω_q which need not be convex. The function E is convex by conservation of convexity; this is obvious here because E is expressed as the supremum of a family of linear functions.

Now if ω_u is convex the determination of v reduces to the minimization of a convex function on a convex set, so that local minimality is sufficient.

As for the lopvalue

$$v' = \sup_q \inf_u \sup_p \left[(x_o, p) + (u, p) - (q, p) - k^f(p)\right]$$

since the bracket is concave in p and linear in u, one can, for a convex set ω_u, interchange the extremizations over u and p to obtain

$$v' = \sup_q \sup_p \inf_u \left[(x_o, p) + (u, p) - (q, p) - k^f(p)\right]$$

$$= \sup_p \left[(x_o, p) - s(-p) + \sigma(-p) - k^f(p)\right]$$

where s is the support function of ω_u.

In case k is a norm, one has

$$v = \inf_u E(x_o + u)$$

$$E(x) = \sup_{p \in B^*} \left[(x, p) + \sigma(-p)\right]$$

In case k is linear, $k = (\pi, x)$, one has

$$v = \inf_u \sup_q (x_o + u - q, \pi) = (x_o, \pi) + \sigma(-\pi) - s(-\pi)$$

$$v' = \sup_q \inf_u (x_o + u - q, \pi) = v$$

a zero-gap situation.

5.15 MULTISTAGE DUALITY

Applying the Fenchel transformation to the dynamic programming algorithm

$$E_i(x) = \sup \{D_i(x-q) : q \in \omega_q(i)\}$$

$$D_{i-1}(x) = \inf \{E_i(x+u) : u \in \omega_u(i)\}$$

one obtains the dual algorithm:

$$E_i^f(p) = \sup_x \inf_{p'} \left[(p-p', x) - \sigma_i(-p') + D_i^f(p')\right]$$

$$D_{i-1}^f(p) = E_i^f(p) + s_i(-p)$$

In case the difference $D_i^f - \sigma_i$ is convex, the first equation reduces to

$$E_i^f(p) = D_i^f(p) - \sigma_i(-p)$$

If the difference is not convex then the first equation expresses E_i^f as the convexification of this difference. It can also be written

$$E_i^f(p) = (D_i^f(p) - \sigma_i(-p))^{ff}$$

It is the possible need for convexification which makes the "discrete minimax principle" fail. Note the inequality

$$E_i^f(p) \leq D_i^f(p) - \sigma_i(-p)$$

The determination of the lopvalue is much simpler

$$v' = \sup_{q_1} \sup_{q_2} \dots \sup_{q_n} \inf_{u_1} \inf_{u_2} \dots \inf_{u_n} k(x_o + \sum_{i=1}^{n} (u_i - q_i))$$

for convex k and convex sets $\omega_u(i)$, the dual representation gives, as in the single stage case

$$v' = \sup_p \left((x_o, p) - k^f(p) + \sum_{i=1}^{n} (\sigma_i(-p) - s_i(-p)) \right)$$

If, in the dual dynamic programming algorithm

$$E_i^f(p) = D_i^f(p) - \sigma_i(-p)$$

holds for all p and i (no need for convexification), then

$$D_o^f(p) = k^f(p) + \sum_{i=1}^{n} (s_i(-p) - \sigma_i(-p))$$

so that

$$v = D_o(x_o) = \sup_p [(x_o, p) - D_o^f(p)] = v'$$

and the zero-gap situation prevails.

The convexification apparently requires extremizations on all of L and L* , an impossible procedure for a computer program. In fact, when the primal criterion function is real valued and the sets $\omega_u(i)$ and $\omega_q(i)$ are bounded it is possible to show that extremization over compact sets, determinable by estimation inequalities, is sufficient.

Finally, note that the case of concave k can be treated by taking it as the negative of a convex function. The effect is to replace inf by sup and vice-versa, with corresponding significant changes in the dual algorithm.

One question of interest in uncertain control problems is that of reachability. Given a set, can one find a controller such that the internal state at the final time will belong to this set, regardless of what the uncertain quantities are? For a vector addition game and set to be reached which is closed and convex the reachability problem amounts to considering as criterion k the indicator function of the given set. Then k^f is the support function of this set, so that the dual algorithm deals entirely with positively homogeneous functions.

Much remains to be done to explore the properties and the implementations of the dual algorithm.

5.16 SOURCES

The theory of two-person zero-sum games is due to Von Neumann [58]. The textbook of Karlin [34] contains many of the results of this theory and of the convexity and duality concepts used in this chapter. For the major saddle-point theorems, see Ky Fan [36, 37] and Moreau [43]. The texts on convexity by Eggleston [15] and Valentine [56] were helpful, as well as the beautiful monograph of Lyusternik [39].

The problem of minimizing a convex function on a set known only by its support function has received attention by Goldstein [24] and Gilbert [23].

The theory of conjugate convex functions is due to Fenchel [19], hence the name Fenchel transform. For its developments see Moreau [42] and Rockafellar [48]. An idea of the difficulty of the infinite-dimensional case can be obtained from the paper of Bronsted and Rockafellar [11] and the references quoted therein.

The theory of games as applied to control situations is so far exclusively centered on the differential games introduced by Isaacs, that is, to a continuous-time zero-gap situation [28, 29]. For developments of this approach see Gadzhiev [21], Grishin [22], Berkovitz and Fleming [8, 9], Pontryagin [46], Ho, Bryson and Baron [26].

CHAPTER VI

BOUNDS FOR THE PERFORMANCE OF SUBOPTIMAL CONTROLLERS

6.1 INTRODUCTION

In this chapter vector addition games are considered in which the criterion function k is a pseudo-norm on the real linear space, L, which may be infinite-dimensional.

Thus $$k(\lambda x) = |\lambda| k(x)$$

and $$k(x+y) \leq k(x) + k(y)$$

the notation $k(x) = \|x\|$ will be used.

Since extremization commutes with monotone increasing functions, all results for guaranteed performance can be translated from the pseudo-norm case to the case where k is a monotone increasing function of a pseudonorm, such as a power.

The objective is to obtain bounds for the guaranteed performance of certain suboptimal controllers. These controllers are obtained by optimization under the (incorrect) assumption that the uncertain vectors q_i are fixed at assumed values q_{oi}.

Pseudonorms are considered because their ratios have an interesting dimensionless meaning. Strict norms are a special case but pseudonorms allow the possibility of "don't care" components. The use of pseudonorms creates no difficulties because everything takes place effectively in a quotient space which is strictly normed.

6.2 THE NAIVE CONTROLLERS

In an n-stage vector addition game let $q_{o1}, \ldots, q_{on}$ be assumed values of the uncertain vectors.

A naive blind controller is a controller which applies, blindly, the input sequence $u_{o1}, \ldots, u_{on}$ where $u_{oi} \epsilon \omega_u(i)$ and

$$k(x_o + \sum_{i=1}^{n} (u_{oi} - q_{oi})) = \min_{u_1 \epsilon \omega_u(i)} \ldots \min_{u_n \epsilon \omega_u(n)} k(x_o + \sum_{i=1}^{n} (u_i - q_{oi})) \tag{1}$$

There may be zero, one or many such controllers. We are interested in statements about such controllers on the assumption that some do exist and we want these statements to hold for all those that exist, unless otherwise stated.

Naive feedback controllers are obtained as follows: For each i from 1 to n, consider the truncated problem, just after the observation of x_{i-1}. For each x_{i-1} let $\mu_i^j(x_{i-1})$ $(j = i, \ldots, n)$ be a naive blind controller for the truncated problem, that is $\mu_i^j(x_{i-1}) \epsilon \omega_u(j)$ and

$$k(x_{i-1} + \sum_{j=i}^{n} (\mu_i^j(x_{i-1}) - q_{oj})) = \min_{u_i \epsilon \omega_u(i)} \ldots \min_{u_n \epsilon \omega_u(n)} k(x_{i-1} + \sum_{j=i}^{n} (u_j - q_{oj})) \tag{2}$$

Then the corresponding naive feedback controller is defined by

$$u_i = \gamma_i(x_{i-1}) \equiv \mu_i^i(x_{i-1}) \quad \text{for } i = 1, \ldots, n \tag{3}$$

Again it is assumed that at least one naive feedback controller $(\gamma_1, \ldots, \gamma_n)$ exists, and there may be many.

This construction of the feedback controller is precisely what is called the synthesis of optimal control as a feedback law in the classical case of no uncertainty. If the assumption that $q_i = q_{oi}$ were indeed correct the naive blind controller and the naive feedback controller would both be optimal and give precisely the same value to the criterion, no supercriterion would be needed.

Since the assumption is incorrect the performance of both controllers must be evaluated by a supercriterion, in general these performances will be different and neither will be optimal.

The guaranteed performance J_o of the naive open loop controller $(u_{o1}, \ldots, u_{on})$ is defined by

$$J_o = \sup \left\{ k(x_o - q + \sum_{i=1}^{n} u_{oi}) : q \epsilon \sum_{i=1}^{n} \omega_q(i) \right\} \quad (4)$$

The guaranteed performance J_f of the naive feedback controller is defined by recursion as follows:

$$G_{n-1}(x) = \sup_{q \epsilon \omega_q(n)} k(x + \gamma_n(x) - q)$$

$$G_{n-2}(x) = \sup_{q \epsilon \omega_q(n-1)} G_{n-1}(x + \gamma_{n-1}(x) - q)$$

etc., ...

$$J_f = \sup_{q \epsilon \omega_q(1)} G_1(x_o + \gamma_1(x_o) - q) \quad (5)$$

Finally, designate by J^* the opvalue v of the problem, that is the best guaranteed performance which can be obtained, at least within ϵ, by using truly optimal feedback control, as considered in Chapter V.

Then $J_o \geq J^*$ and $J_f \geq J^*$ by optimality. Note that for the single stage case $J_o \equiv J_f$ because there are no outputs.

We shall say that the naive open loop controller $(u_{o1}, \ldots, u_{on})$ corresponds to the naive feedback controller $(\gamma_1, \ldots, \gamma_n)$ if

$$u_{o1} = \gamma_1(x_o)$$

and for $i = 2, \ldots, n$

$$u_{oi} = \gamma_i\left(x_o + \sum_{j=1}^{i-1} (u_{oj} - q_{oj})\right) \tag{6}$$

6.3 MAIN ASSUMPTIONS

A first assumption is that the sets $\omega_q(i)$ are bounded in pseudonorm. This assures that J_o, J_f, J^* and all other quantities considered will be finite.

Now note that J_o, J_f and J^* are unchanged if one or more of the sets $\omega_u(i)$ and $\omega_q(i)$ is replaced by its closure in the pseudonorm topology. Furthermore if all the suprema over the sets $\omega_q(i)$ involve convex functions, then nothing is changed by replacing one or more of the sets $\omega_q(i)$ by the closure of its convex hull. These replacements will be tacitly assumed.

Clearly the assumed values q_{io} must bear some relation to the sets $\omega_q(i)$. The second basic assumption is that for $i = 1, \ldots, n$ q_{io} belongs to $\omega_q(i)$. By the above it suffices that it belong to the closure of the convex hull of $\omega_q(i)$.

Besides the two basic assumptions, three additional assumptions will be investigated as to their consequences.

Assumption S: For $i = 1, \ldots, n$ the set $\omega_q(i)$ is symmetric about the point q_{oi}.

By the remarks above it suffices that the closure of the convex hull of $\omega_q(i)$ be symmetric about q_{oi}.

Assumption C: The sets $\omega_u(i)$ are convex for all i. (It suffices that their closure be convex).

Assumption P: The pseudonorm is quadratic, that is, it satisfies the parallelogram law

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{7}$$

Assumption P implies that

$$xy \equiv \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) \tag{8}$$

is a pseudo inner product on L. In the present chapter it is denoted by juxtaposition. Hence $x^2 \equiv \|x\|^2$.

One has

$$xy = yx$$

$$x(y+\lambda z) = xy + \lambda xz$$

$$x^2 \geq 0$$

$$|xy| \leq \|x\| \cdot \|y\| \tag{9}$$

The additional properties that

$$x^2 = 0 \Rightarrow x = 0$$

and $|xy| = \|x\| \cdot \|y\| \Rightarrow x, y$ linearly dependent

are only true for strict norms and inner products, they are not required in this chapter.

Note that no assumptions are made on the dimension, completeness or separability of the pseudonormed space L.

6.4 THE SINGLE-STAGE CASE

For the single stage case $(n=1)$ the time index i becomes redundant. By translation of ω_u or ω_q one can take $x_o = 0$ without loss of generality. According to the basic assumptions the assumed value of q is $q_o \in \omega_q$ and $\sup \|\omega_q\| < \infty$.

A naive controller is defined by u_o, with

$$\|u_o - q_o\| = \min_{u \in \omega_u} \|u - q_o\| \tag{10}$$

Define $$J : L \rightarrow R$$

by $$J(x) = \sup_{q \in \omega_q} \|x - q\| \tag{11}$$

then $$J^* = \inf_{u \in \omega_u} J(u) \tag{12}$$

and $$J_o = J(u_o) \tag{13}$$

As a supremum of convex functions J is a convex function. By the triangular inequality

$$\|x - q\| - \|x - y\| \leq \|y - q\| \leq \|x - q\| + \|x - y\|$$

taking the sup over q in ω_q

$$J(x) - \|x-y\| \leq J(y) \leq J(x) + \|x-y\|$$

or
$$|J(x) - J(y)| \leq \|x-y\| \qquad (14)$$

hence J is Lipschitz continuous with the constant 1. Define

$$R_o = \sup_{q \in \omega_q} \|q - q_o\|$$

The most immediate inequalities are summarized in

Theorem 6.1 Under the basic assumptions

$$\max(\|u_o - q_o\|, R_o/2) \leq J^* \leq J_o \leq \|u_o - q_o\| + R_o$$

Proof: $J^* \leq J_o$ because u_o belongs to ω_u

$$J_o = \sup_{q \in \omega_q} \|u_o - q\| = \sup_{q \in \omega_q} \|(u_o - q_o) + (q_o - q)\|$$

$$\leq \|u_o - q_o\| + \sup_{q \in \omega_q} \|q_0 - q\| = \|u_o - q_o\| + R_o$$

Since q_o belongs to ω_q

$$J(u) = \sup_{q \in \omega_q} \|u-q\| \geq \|u-q_o\|$$

By (10) for u in ω_u

$$\|u-q_o\| \geq \|u_o - q_o\|$$

so that
$$J(u) \geq \|u_o - q_o\|$$

taking inf over u in ω_u

$$J^* \geq \|u_o - q_o\|$$

Finally $\|q_o - q\| \leq \|u - q_o\| + \|u - q\|$

or $$R_o = \sup_{q \in \omega_q} \|q_o - q\| \leq \|u - q_o\| + J(u) \leq 2J(u)$$

because $q_o \epsilon \omega_q$

taking inf over u in ω_u

$$R_o \leq 2J^* \qquad \text{QED}$$

A more interesting type of bound is obtained by finding the smallest number α such that $J_o \leq \alpha J^*$ under a given combination of assumptions. These ratio bounds are derived in the theorems that follow.

Theorem 6.2 Under the basic assumptions solely or augmented by C or augmented by P (but not by both) the smallest number α such that

$$J_o \leq \alpha J^* \text{ is } 3$$

Proof: A. First show that the basic assumptions alone imply $J_o \leq 3J^*$. Indeed for all u in ω_u and q in ω_q, by the triangular inequality

$$\|u_o - q\| \leq \|u_o - q_o\| + \|q_o - u\| + \|u - q\|$$

By (10): $$\|u_o - q_o\| \leq \|q_o - u\|$$

Since q_o and q belong to ω_q:

$$\|u_o - q\| \leq 3J(u)$$

Taking inf over u in ω_u:

$$\|u_o - q\| \leq 3J^*$$

Taking sup over q in ω_q:

$$J_o \leq 3J^*$$

B. To show that the bound is the best it suffices to produce an example where $0 < J_0 = 3J^*$ with C satisfied and another such example with P satisfied.

Take $L = R^2$ with the sup norm. Let $\omega_q = \{(0,1), (2,1)\}$, $\omega_u = \{(a,0): a \in R\}$, $q_0 = (0,1)$ and $u_0 = (-1,0)$. Then $J_0 = 3$ and $J^* = 1$ while C is satisfied.

Take $L = R$ with absolute value norm. Let $\omega_q = \{1,3\}$, $\omega_u = \{0,2\}$, $q_0 = 1$, $u_0 = 0$. Then $J_0 = 3$ and $J^* = 1$ while P is satisfied. QED

The proof of Theorem 6.2 requires only the triangular inequality, hence the bound also holds in pseudometric spaces.

<u>Lemma 6.1</u> If assumption P and C hold, then for all u in ω_u

$$(q_0 - u_0)\ (u - u_0) \leq 0$$

Proof: If $\|u - u_0\| = 0$ the inner product vanishes by (9), otherwise by (10)

$$\|u_0 - q_0\| = \min_{u \in \omega_u} \|u - q_0\|$$

Thus for u in ω_u

$$\|q_0 - u_0\|^2 \leq \|q_0 - u\|^2 = \|(q_0 - u_0) - (u - u_0)\|^2$$

or

$$\|u - u_0\|^2 - 2(u - u_0)(q_0 - u_0) \geq 0$$

Since ω_u is convex and contains u_0 and u, for θ in $[0,1]$ $u_0 + \theta(u - u_0)$ belongs to ω_u and

$$\theta^2 \|u - u_0\|^2 - 2\theta(u - u_0)(q_0 - u_0) \geq 0$$

The claim follows by taking

$$\theta = \min \left[1, \frac{|(u-u_0)(q_0-u_0)|}{\|u-u_0\|^2} \right]$$

QED

In other words, this well known property of euclidean spaces actually holds in any real pseudoprehilbert space.

Theorem 6.3 If the basic assumptions are augmented either by S or by S and C or by P and C, then the smallest number α such that $J_0 \leq \alpha J^*$ is 2.

Proof: A. First show that S implies $J_0 \leq 2J^*$.

Since J is convex and S implies $J(x) \equiv J(2q_0 - x)$, q_0 must give J its absolute minimum over L. Thus $J(q_0) \leq J^*$. By Theorem 6.1, $\|u_0 - q_0\| \leq J^*$.

By the triangular inequality

$$\|u_0 - q\| \leq \|u_0 - q_0\| + \|q_0 - q\|$$

take sup over q in ω_q to obtain

$$J_0 \leq \|u_0 - q_0\| + J(q_0) \leq 2J^*$$

B. Now show that P and C imply $J_0 \leq 2J^*$. For u in ω_u

$$\|u - q_0\|^2 = \|u - u_0 + u_0 - q_0\|^2$$

$$= \|u - u_0\|^2 + \|u_0 - q_0\|^2 + 2(u - u_0)(u_0 - q_0 \geq \|u - u_0\|^2$$

(using lemma 6.1.) Thus $\|u - u_0\| \leq \|u - q_0\| \leq \sup_{q \in \omega_q} \|u - q\| = J(u)$.

Also for all q in ω_q and u in ω_u

$$\|q-u_0\| = \|q-u+u-u_0\| \leq \|q-u\| + \|u-u_0\| \leq \|q-u\| + J(u)$$

take sup over q in ω_q

$$J_0 \leq 2J(u)$$

and inf over u in ω_u

$$J_0 \leq 2J^*$$

C. To show that the bound is the best it suffices to produce an example with $0 < J_0 = 2J^*$ for which S and C hold and another such example for which P and C hold.

Take $L = R^2$ with the sup norm, let $\omega_q = \{(-1, 1), (1, 1)\}$. $\omega_u = \{(a, 0): a \in R\}$, $q_0 = (0, 1)$ and $u_0 = (1, 0)$. Then $J_0 = 2$ and $J^* = 1$ while both S and C hold.

Take $L = R$ with the absolute value norm, take $\omega_u = R$, $\omega_q = \{0, 2\}$, $q_0 = 0$, $u_0 = 0$, then $J_0 = 2$ and $J^* = 1$ while both P and C hold. QED

<u>Theorem 6.4</u> If the basic assumptions are augmented by S and P, then the smallest number α such that $J_0 \leq \alpha J^*$ is $\sqrt{2}$.

Proof: To show that the bound holds, note that for all u in ω_u and q in ω_q by assumption S

$$J^2(u) \geq \max\ (\|u-q\|^2, \|u-2q_0+q\|^2)$$

$$= \max(\|(u-q_0)-(q-q_0)\|^2, \|(u-q_0)+(q-q_0)\|^2)$$

$$= \|u-q_0\|^2 + \|q-q_0\|^2 + 2|(u-q_0)(q-q_0)|$$

$$\geq \|u-q_0\|^2 + \|q-q_0\|^2$$

$$\geq \|u_0 - q_0\|^2 + \|q-q_0\|^2$$

by (10). Taking inf over u in ω_u

$$J^{*2} \geq \|u_0-q_0\|^2 + \|q-q_0\|^2$$

$$= \frac{1}{2}\|u_0-q\|^2 + \frac{1}{2}\|u_0-2q_0+q\|^2$$

by P

$$\geq \frac{1}{2}\|u_0-q\|^2$$

Taking sup over q in ω_q

$$J^{*2} \geq \frac{1}{2} J_0^2$$

To show that this bound is the best take the euclidean plane for L, let $\omega_q = \{(-1,0), (1,0)\}$, $\omega_u = \{(x_1,x_2): x_1^2+x_2^2 = 1\}$, $q_0 = (0,0)$, $u_0 = (1,0)$. Then $J_0 = 2$ and $J^* = \sqrt{2}$ while both S and P hold.

QED

Theorem 6.5 If the basic assumptions are augmented by S, P and C, then the smallest number α such that $J_0 \leq \alpha J^*$ is $2/\sqrt{3}$.

Proof: Define $m: L \times L \to R$ by

$$m(u,q) = \max\left(\|q-u\|^2, \ \|2q_0 - q - u\|^2\right) \tag{15}$$

then by S
$$J^2(u) = \sup_{q \in \omega_q} m(u,q) \tag{16}$$

while by (15)

$$4m(u,q) - 3\|u-q\|^2 - \|2q_0 - q - u\|^2 \geq 0 \tag{17}$$

For u in ω_u by lemma 6.1.

$$12(q_0 - u_0)(u_0 - u) \geq 0 \tag{18}$$

and of course

$$\|(2q_0 - q - u) - 3(u_0 - u)\|^2 \geq 0 \tag{19}$$

Adding (17), (18) and (19) and rearranging

$$4m(u,q) \geq 3\|u_0 - q\|^2$$

taking sup over q in ω_q, by (16)

$$4J^2(u) \geq 3J_0^2$$

Take inf over u in ω_u

$$4J^{*2} \geq 3J_0^2$$

To show that the bound is the best take the euclidean plane for L, let $\omega_q = \{(0,0), (4, 2\sqrt{2})\}$, $\omega_u = \{(a,0): a \in R\}$, $q_0 = (2, \sqrt{2})$, $u_0 = (2, 0)$. Then $J_0 = 2\sqrt{3}$ and $J^* = 3$ while S, P and C all hold. QED

A somewhat longer but illuminating alternative proof of Theorem 6.5 proceeds in four steps. By S it is sufficient to consider sets ω_q consisting of 2 points (or their convex hull, a segment). By C it is sufficient to consider sets ω_u which are closed half-spaces. By P it is then sufficient to consider the problem in the plane through the two points, orthogonal to the boundary of the half-space. After this reduction, the bound for 2 points and a half plane can be established by plane euclidean geometry.

In some of the examples showing a bound to be the best in Theorems 6 2 to 6.4, the suboptimal control u_0 is not uniquely determined by (10). This lack of uniqueness can be removed by minor changes (using an additional dimension if necessary) while the ratio changes infinitesimally. Hence a uniqueness requirement could at most lead to the statement of the bound with strict inequality for non-zero J*.

When ω_q has no center of symmetry the next best assumption, to replace S, is that q_0 is an outcenter of ω_q, that is a point at which J attains its minimum over L. In conjunction with the basic assumption that q_0 belongs to the closure of the convex hull of ω_q this leads to a best bound of 2 as for S. But without this basic assumption the bound is 3 unless the unit ball in the quotient space is assumed uniformly rotund. This subject is not pursued because outcenters are hard to determine in practice.

6.5 A BOUND FOR EXPECTED PERFORMANCE

The stochastic version of the single stage case is obtained by assuming that q is a random vector with given distribution on L and letting $J(x)$ be the expectation of $\|q-x\|$.

Equation 10 and assumptions P and C retain their meaning. The basic assumption $q_0 \epsilon \omega_q$ and assumption S must be redefined.

Let A be a σ-algebra on L such that pseudonorm is measurable and $a \epsilon A$, $x \epsilon L$ imply $x-a \epsilon A$.

Let μ be a probability measure on (L, A) such that the expectation of $\|q\|$ is finite and for a in A.

$$\mu(a) = \mu(2q_0 - a)$$

that is, the probability measure is symmetric about q_0.

Define
$$J(x) = \underset{\mu(q)}{E} \|x-q\|$$

$$J_0 = J(u_0)$$

$$J^* = \inf_{u \epsilon \omega_u} J(u)$$

Now we seek the smallest number a such that $J_0 \leq aJ^*$ given that (10), P, C and the stochastic form of the symmetry assumption all hold.

First note that, unlike supremum, expectation does not commute with monotone increasing functions. Hence if $k(x) = \|x\|^2$ the problem is entirely different. When P holds $\|x\|^2$ is a quadratic form and if q_0 is the mean of q

$$\underset{\mu(q)}{E} \|u-q\|^2 - \|u-q_0\|^2 = \underset{\mu(q)}{E} \|q-q_0\|^2$$

Since the right-hand-side is independent of u, optimization against the mean <u>is</u> optimal and $J_0 = J^*$ when J is defined by $J(x) = \underset{\mu(q)}{E} \|x-q\|^2$. But this does not hold for the norm itself. Indeed one has

<u>Theorem 6.6</u> Under the stochastic symmetry assumption with P, C, (10) and $J(x) = \underset{\mu(q)}{E} \|x-q\|$, the smallest number α such that $J_0 \leq \alpha J^*$ is $2/\sqrt{3}$.

Proof (outline): By the stochastic symmetry assumption a bound will hold if it holds for atomic measures assigning equal weights to two points. By C, a bound will hold if it holds for ω_u a closed half-space. By P it then suffices to consider the plane through the two points orthogonal to the boundary of the half-space. In this plane a bound of $2/\sqrt{3}$ can be shows to hold by elementary geometry with a discussion of cases. To prove that this bound is the best let L be the euclidean plane, μ the atomic measure with equal weight at $(0,0)$ and $(2, 2\sqrt{2})$, $\omega_u = \{(a, 0): a \in R\}$, $q_0 = (1, \sqrt{2})$, $u_0 = (1, 0)$. Then $J^* = 2\sqrt{3}$ and $J_0 = 4$ while all assumptions are satisfied. QED

It is remarkable that this bound is the same as in the corresponding case of guaranteed performance (Theorem 6.5.). The underlying reason for this equality is not clear as yet. It may hinge on the fact that R^2 with the ℓ-1 norm and R^2 with the ℓ-∞ norm are not only dual normed spaces but are also isometrically isomorphic.

6.6 THE TWO-STAGE CASE

Consider a naive feedback controller (γ_1, γ_2) for a two-stage vector addition game and let (u_{01}, u_{02}) be the naive blind controller corresponding to it by (6). Then the three numbers J_0, J_f, J^* are in general distinct and satisfy $J^* \leq \min(J_0, J_f)$. Even under the strongest combination of assumptions considered so far (S, P and C) there is then no nontrivial bound involving J^*. Indeed one can have $J^* = J_0 = J_f$ by letting $\omega_q(1)$ and $\omega_q(2)$ be singletons, and one has

Theorem 6.7 For two-stage problems satisfying assumptions S, P and C or any weakening thereof the largest numbers λ_0, λ_1, λ_2 such that

$$\lambda_0 J_0 \leq J^*, \lambda_1 J_f \leq J^*, \lambda_2 J_0 \leq J_f$$

are

$$\lambda_0 = \lambda_1 = \lambda_2 = 0$$

Proof: The inequalities hold because pseudonorms are non-negative. To prove that they are the best it suffices to give an example for which $J^* = 0$ with $J_f > 0$ and $J_0 > 0$ and an example with $J_f = 0$ and $J_0 > 0$.

Let L be the real line with the absolute value norm. Let $\omega_q(1) = \{-1, 1\}$, $\omega_q(2) = \{0\}$, $\omega_u(1) = [-1, 1]$, $\omega_u(2) = [0, 2]$, $x_0 = q_{01} = q_{02} = \gamma_1(x_0) = 0$, $\gamma_2(x) = (1+x)/2$.

Then $J_0 = J_f = 1$ and $J^* = 0$.

Now modify this example by letting $\gamma_1(x_0) = -1$ and $\gamma_2(x) = 1+x$. Then $J_0 = 1$ and $J_f = 0$. QED

Note that the remarks on uniqueness at the end of Section 6.4 apply also to the above examples.

In conclusion, the only non-trivial bound is of the type $J_f \le \alpha J_0$. Intuition suggests that $\alpha = 1$ that is, feedback, even naive, cannot be worse than blind naive control. A lower value than 1 is ruled out by the first example in Theorem 6.7. As we shall see the best bounds under the combination of assumptions considered already for the single stage case, are all greater than 1. The meaning of this result is that a naive feedback controller can be fooled, while a blind controller cannot. The best bound α under given assumptions will be called the fooling factor for these assumptions.

Theorem 6.8 In a two-stage problem, assume that $\omega_u(2)$, $\omega_q(2)$, q_{02} and the pseudonorm satisfy one of the combinations of assumptions required for the single-stage data in any one of theorems 6.2, 3, 4 or 5. Let α be the corresponding bound, that is $3, 2, \sqrt{2}$ or $2/\sqrt{3}$. Then $J_f \le \alpha J_0$ in the two-stage problem.

Proof: by (5)

$$J_f = \sup_{q \in \omega_q(1)} G(x_0 + \gamma_1(x_0) - q)$$

with

$$G(x) = \sup_{q \in \omega_q(2)} \|x + \gamma_2(x) - q\|$$

Also, interpreting (4) with the help of (6)

$$J_0 = \sup_{q\,\epsilon\,\omega_q(1)} H(x_0+\gamma_1(x_0) - q \ , \ x_0 + \gamma_1(x_0) - q_{01})$$

with
$$H(x, \xi) = \sup_{q\,\epsilon\,\omega_q(2)} \| x+\gamma_2(\xi) - q \|$$

since $\gamma_2(\xi)$ belongs to ω_u, we have for all x and ξ

$$H(x, \xi) \geq \inf_{u\,\epsilon\,\omega_u(2)} \ \sup_{q\epsilon\omega_q(2)} \|x+u-q\|$$

$$\geq \frac{1}{a} \sup_{q\epsilon\omega_q(2)} \| x + \gamma_2(x) - q\|$$

by the single-stage result, which applies because (2) and (3) imply that (10) holds when x is considered the initial state of a single-stage problem with $\omega_q(2)$ and $\omega_u(2)$ as constraint sets (and x is taken as origin).

Thus

$$H(x, \xi) \geq \frac{1}{a} \ G(x)$$

in particular.

$$H(x_0 + \gamma_1(x_0) - q, \ x_0 + \gamma_1(x_0) - q_{01}) \geq \frac{1}{a} G(x_0+\gamma_1(x_0)-q)$$

taking sup over q in ω_q in $\omega_q(1)$

$$J_0 \geq \frac{1}{a} J_f$$

QED

Note that this bound is not claimed to be the best, but one has

Theorem 6.9 For two-stage problems satisfying assumptions S, P and C the fooling factor is $2/\sqrt{3}$.

Proof: The bound holds by Theorem 6.5 via Theorem 6.8. To show that it is the best, let L be the euclidean plane.

$$\omega_u(1) = \{(0,0)\},\ \omega_u(2) = \{(a,0): a \in R\},\ \omega_q(1) = \{(1,-\sqrt{2}),(-1,\sqrt{2})\},$$

$$\omega_q(2) = \{(2,\sqrt{2}),\ (-2,-\sqrt{2})\},\ x_0 = q_{01} = q_{02} = \gamma_1(x_0) = (0,0),$$

$\gamma_2(x)$ = orthogonal projection of x on $\omega_u(2)$. Then $J_f = 2\sqrt{3}$, $J_0 = J^* = 3$ while S, P and C hold. QED

Of course the fooling factor for weaker assumptions can only be larger.

In the case of expected performance, Theorem 6.8 is valid with supremum replaced by expectation throughout. For the square of a quadratic norm the fooling factor is 1, this is essentially the Wiener-Kalman-Bucy case with the degeneracy that the output is the exact state vector. But for the norm itself one has

Theorem 6.10 For two-stage problems of expected performance satisfying P, C and the stochastic form of S the fooling factor is $2/\sqrt{3}$.

Proof: The bound holds by Theorem 6.6 via the stochastic parallel of Theorem 6.8. To show that it is the best, let L be the euclidean plane. Let the μ_1 and μ_2 be independent and identical probability measures for q_1 and q_2, namely atomic measures with equal weights at $(1,\sqrt{2})$ and $(-1,-\sqrt{2})$. Let $\omega_u(1) = \{(0,0)\}$, $\omega_u(2) = \{(a,0): a \in R\}$, $x_0 = q_{01} = q_{02} = \gamma_1(x_0) = (0,0)$, $\gamma_2(x)$ = orthogonal projection of x on $\omega_u(2)$. Then $J_f = 2$ and $J_0 = J^* = \sqrt{3}$ while all assumptions are satisfied. QED

Again, under weaker assumptions the stochastic fooling factor can only be larger.

6.7 THE MULTI-STAGE CASE

For the case of an arbitrary number n of stages little is known as yet. An exponential bound a^n on the fooling factor is easily obtained but far too high. Note though that, for otherwise fixed assumptions, the bound must be monotone in n because any n-stage problem is equivalent to an (n+1)-stage problem with trivial first stage.

It is the asymptotic behavior of the fooling factor for large n which is of the greatest interest. It seems that the factor goes to infinity with n. To obtain finite limits one can make the assumption that L has the finite dimension d. The limit for infinite n is then finite for fixed d and bounds on this limit have been found. They go to infinity with d.

Another question of considerable interest is the implication of time invariance of the linear differential system from which the vector addition game is derived. Finally, the continuous-time case merits attention despite its much greater technical difficulty.

6.8 SOURCES

Some discussion of the differences between various suboptimal design methods for uncertain systems may be found in Dreyfus [13] who considers discrete-time problems and expected performance.

The mathematical background of the present chapter is wholly elementary. In fact the proof of the $2/\sqrt{3}$ bound, the least trivial, is nothing more than plane euclidean geometry. For the general mathematical background the introductory texts of Simmons [50] and Royden [49] are far more than sufficient.

BIBLIOGRAPHY

1. Astrom, K. J., "Optimal Control of Markov Processes with Incomplete State Information," Jl. Math. An. Appl., Vol. 10, p. 174, 1965.

2. Athans, M. and Falb, P. L., Optimal Control. An Introduction to the Theory and its Applications, McGraw-Hill, 1965.

3. Beckenbach, E. F. and Bellman, R., Inequalities, Springer, 1961.

4. Bellanger, P. R., "On the Sensitivity of Solutions to Some Optimal Control Problems," Ph.D. Thesis, M.I.T , 1964.

5. Bellman, R., and Kalaba, R., "A Mathematical Theory of Adaptive Control Processes," Proc. Nat. Acad. Sci. USA, Vol. 45, p. 1288-1290, 1959.

6. Bellman, R., Adaptive Control Processes. A guided tour. Princeton, 1961.

7. Bellman, R., and Dreyfus, S.E., Applied Dynamic Programming, Princeton, 1962. (Chapter VIII: "Feedback Control Processes")

8. Berkovitz, L.D. and Fleming, W. H., "On Differential Games With Integral Payoff." In: Contributions to the Theory of Games III, Annals of Math. Study 39, Princeton, 1957.

9. Berkovitz, L. D , "A Variational Approach to Differential Games." In: Advances in Game Theory, Annals of Math. Study 52, p. 127-174, Princeton, 1964.

10. Blackwell, D. and Girshick, M. A., Theory of Games and Statistical Decisions, Wiley, 1954.

11. Bronsted, A., and Rockafellar, R. T., "On the Subdifferentiability of Convex Functions," Proc. Am. Math. Soc., Vol. 16, p. 605-611, 1965.

12. Dorato, P. and Kestenbaum, A., "Application of Game Theory to the Sensitivity Design of Systems with Optimal Controller Structures," Proc. Third Allerton Conference on Circuit and System Theory, University of Illinois, Sept., 1965.

13. Dreyfus, S.E., Dynamic Programming and the Calculus of Variations, New York, Academic Press, 1965.

14. Dynkin, E.B., "Controlled Random Sequences," In: Theory of Probability and Applic. (translation), Vol. 10, p. 1-14, 1965.

15. Eggleston, H.G., Convexity, Cambridge U. Press, 1963.

16. Everett, H. "Recursive Games." In: Contrib. to the Theory of Games, Annals of Math. Study 39, Princeton, 1957.

17. Feldbaum, A.A., "Dual Control Theory," Automation and Remote Control. Part I, Vol. 21, p. 874, 1960, Part II, Vol. 21, p. 1033, 1960, Part III, Vol. 22, p. 1, 1961, Part IV, Vol. 22, p. 109, 1961.

18. Feldbaum, A.A., Optimal Control Systems, Academic Press, New York, 1965.

19. Fenchel, W., "On Conjugate Convex Functions," Canad. J. Math., Vol. 1, p. 73-77, 1949.

20. Florentin, J.J., "Optimal Control of Continuous Time Markov Stochastic Systems," J. Electronics and Control, Vol. 10, p. 473, 1961.

21. Gadzhiev, M.Y., "Application of the Theory of Games to Some Problems of Automatic Control," Automation and Remote Control. Part I, Vol. 23, p. 957, 1963, Part II, Vol. 23, p. 1074, 1963.

22. Grishin, V.P., "A Minimax Problem in the Theory of Analytical Design of Control Systems," Automation and Remote Control, Vol. 25, p. 779, 1964.

23. Gilbert, E. G., "An Iterative Procedure for Computing the Minimum of a Quadratic Form on a Convex Set," Journal SIAM Control, Vol. 4, p. 61, 1966.

24. Goldstein, A.A., "Convex Programming and Optimal Control," Journal SIAM Control, Ser. A, Vol. 3, p. 142-146, 1965.

25. Halkin, H., "The Principle of Optimal Evolution". In" Intern. Symposium on Nonlinear Differential Equations and Nonlinear Mechanics, p.284-301, New York, Academic Press, 1963.

26. Ho, Y.C., Bryson, A.E. Jr., Baron, S., "Differential Games and Optimal Pursuit Evasion Strategies," IEEE Trans. Aut. Control, Vol. AC-10, p. 385-389, 1965.

27. Howard, D.R., Rekasius, Z.V., "Error Analysis with the Maximum Principle," IEEE Trans., Vol. AC-9, p. 223, July, 1964.

28. Isaacs, R., "Differential Games," Rand Corporation Research Memoranda, RM-1391, RM-1399, RM-1411, 1954 and RM-1486, 1955.

29. Isaacs, R., Differential Games, SIAM Series in Appl. Math., Wiley, 1965.

30. James, H.M., Nichols, N.B., Phillips, R.J., Theory of Servomechanisms, New York, McGraw-Hill, 1947.

31. Johansen, D.E., "Solution of a Linear Mean Square Estimation Problem When Process Statistics are Undefined." Proc. JACC, 1965, p. 64-74.

32. Kalman, R.E , Bucy, R.S., "New Results in Linear Filtering and Prediction Theory," ASME J. Basic Eng., Vol. D-83, p. 95, 1961.

33. Kalman, R.E., Ho, Y.C., Narendra, K.S., "Controllability of Linear Dynamical Systems." In: Contributions to Differential Equations, Vol. 1, No. 2, p. 189-213, New York, Wiley, 1962.

34. Karlin, S., Mathematical Methods and Theory in Games, Programming and Economics, (2 Vols.), Addison-Wesley, 1960.

35. Kushner, H.J., "On the Status of Optimal Control and Stability for Stochastic Systems," 1966 IEEE Convention Record, Part 6, p. 143-151, 1966.

36. Ky Fan, "Fixed-point and Minimax Theorems in Locally Convex Topological Linear Spaces," Proc. Nat. Acad. Sci., Vol. 38, p. 121-126, 1952.

37. Ky Fan, "Minimax Theorems," Proc. Nat. Acad. Sci., Vol. 39, p. 42-47, 1953.

38. Luce, R.D., and Raiffa, H., Games and Decisions, New York, Wiley, 1957.

39. Lyusternik, L.A., Convex Figures and Polyhedra, Dover, New York, 1963.

40. Macko, D.S., "Uncertainty and Optimal Control," Report SRC 16-A-62-10, Case Institute of Technology, May, 1962.

41. Middleton, D., "Communication and Statistical Decision Theory," Electro-Technology, Vol. 74, p. 67-82, Nov., 1964.

42. Moreau, J.J., "Fonctions Convexes Duales et Points Proximaux dans un Espace Hilbertien," C.R., Vol. 255, p. 2897-2899, 1962.

43. Moreau, J.J., "Théorèmes 'inf-sup'," C.R., Vol. 258, p. 2720-2722, 1964.

44. Neustadt, L. W., "The Existence of Optimal Controls in the Absence of Convexity Conditions," J. Math. Anal. Appl., Vol. 7, p 110-117, 1963.

45. Pontryagin, L.S., Boltyanskii, V.G., Gamrelidze, R.V , Mishenko, E.F., The Mathematical Theory of Optimal Processes, Translation: Interscience, 1962.

46. Pontryagin, L.S., "On Some Differential Games," Soviet Ac.Sc. Doklady (transl.), Vol. 156, p. 712-716, 1965. Also: Jrl.SIAM on Control, Vol. 3, p. 49-52, 1965.

47. Pratt, J.W., Raiffa, H., Schlaifer, R., Introduction to Statistical Decision Theory, McGraw-Hill, New York, 1965.

48. Rockafellar, R.T., "Duality Theorems for Convex Functions," Bulletin Am. Math. Soc., Vol. 70, p. 189-192, 1964.

49. Royden, H.L., Real Analysis, New York, Macmillan, 1963.

50. Simmons, G.F., Introduction to Topology and Modern Analysis, McGraw-Hill, 1963.

51. Skelton, G.B., "Optimum Fixed Control of an Uncertain Plant," Ph.D. Thesis, M.I.T., 1964.

52. Striebel, C., "Sufficient Statistics in the Optimum Control of Stochastic Systems," Journal of Math. Anal. and Appl., Vol. 12, p. 576-592, 1965.

53. Sworder, D.D., and Aoki, M., "On Control System Equivalents of Some Decision Theoretic Theorems," J.Math. Anal. and Appl., Vol. 10, p. 424-438, 1965.

54. Sworder, D.D., "Minimax Control of Discrete Time Stochastic Systems," J. SIAM Control, Ser. A, Vol. 2, p. 433-449, 1965.

55. Thrall, R.M., Coombs, C.H., Davis, R.L., editors, Decision Processes, New York, Wiley, 1954.

56. Valentine, F.A., Convex Sets, New York, McGraw-Hill, 1964.

57. Van Slyke, R., and Wets, R., "Programming Under Uncertainty and Stochastic Optimum Control," J.SIAM Control, Ser. A, Vol. 4, p. 179-193, 1966.

58. Von Neumann, J. and Morgenstern, D., Theory of Games and Economic Behavior, Princeton, 1944.

59. Wald, A., Statistical Decision Functions, New York, Wiley, 1950.

60. Warga, J., "Minimax Problems and Unilateral Curves in the Calculus of Variations," J. SIAM Control, Ser. A, Vol. 3, p. 91 - 105, 1965.

61. Wets, R., "Programming Under Uncertainty. The Complete Problem," Z. Wahrscheinlichkeitstheorie verw. Gb., Vol. 4, p. 316 - 339, 1966.

62. Zachrisson, L. E., "Markov Games," In: Advances in Game Theory," Annals of Math. Studies, No. 52, p. 211 - 253, Princeton, 1964.

63. Zadeh, L. A., "General Filters for Separation of Signal and Noise," In: Symposium on Information Networks, p. 31 - 49, Polytechnic Institute of Brooklyn, April, 1954.

64. Zadeh, L. A., "Optimality and Non-scalar Valued Performance Criteria," IEEE Trans. Aut. Control, Vol. 8, p. 59 - 60, 1963.

65. Zadeh, L. A., "On the Definition of Adaptivity," Proc. IEEE, Vol. 51, p. 469 - 470, 1964.