PRENTICE HALL

SPACE TECHNOLOGY SERIES

David O. Ellis
and Fred J. Ludwig

SYSTEMS PHILOSOPHY

PRENTICE-HALL INTERNATIONAL SERIES
IN SPACE TECHNOLOGY

HUNT LIBRARY

Carnegie-Mellon University
PITTSBURGH PENNSYLVANIA 1900

DAVID O. ELLIS, PH.D. University of Missouri, is Manager, Data Systems Research and Development, Litton Systems, Inc., of California. Dr. Ellis is widely experienced in systems technology and has contributed numerous research articles to scientific journals both here and abroad.

FRED J. LUDWIG, M.A., University of Arizona, is a research scientist in the Guidance and Control Systems Division of Litton Systems, Inc., of California.

Systems Philosophy

PRENTICE-HALL SPACE TECHNOLOGY SERIES

C. W. Besserer and Floyd E. Nixon, editors

BAZOVSKY *Reliability Theory and Practice*
ELLIS AND LUDWIG *Systems Philosophy*
KOLK *Modern Flight Dynamics*
LLOYD AND LIPOW *Reliability: Management, Methods, and Mathematics*
SAMARAS *Theory of Ion Flow Dynamics*
SCHLESINGER *Principles of Electronic Warfare*
STILTZ, ED. *Aerospace Telemetry*

PRENTICE-HALL, INC.
PRENTICE-HALL INTERNATIONAL, INC., UNITED KINGDOM AND EIRE
PRENTICE-HALL OF CANADA, LTD., CANADA
BERLINER UNION, WEST GERMANY AND AUSTRIA

David O. Ellis

Manager, Research and Analysis Department, Litton Industries
Formerly Associate Professor of Mathematics
University of Florida

Fred J. Ludwig

Research Scientist, Litton Industries
Formerly Research Psychologist
University of Arizona

Systems Philosophy

Prentice-Hall, Inc.
Englewood Cliffs, N. J.
1962

© 1962 by Prentice Hall, Inc., Englewood Cliffs, N. J.

ALL RIGHTS RESERVED. NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM, BY MIMEOGRAPH OR ANY OTHER MEANS, WITHOUT PERMISSION IN WRITING FROM THE PUBLISHER.

LIBRARY OF CONGRESS CATALOG CARD NO. 62–8833

PRINTED IN THE UNITED STATES OF AMERICA

88158–C

Preface

We have witnessed in the past few years the advent of many new branches of technology involving complex constructs both in concept and equipment. The realization of many of these constructs may be conceded as well beyond the states-of-the-arts of previous decades. As a result, developments in these areas have outstripped general understanding; although many special, and a few fairly general, treatments have appeared in technical literature. Running through these various areas of complex constructs is a common thread, the systems concept.

It is the intent of the text portion of this book to provide a discussion of the major key points and probably trends in systems technology; to make them intelligible to both management and the public; and to furnish a general survey of the subject to the scientific generalists and specialists evolving the technology. To this end, technical detail has, for the most part, been relegated to Appendixes.

The Appendixes are completely independent and provide supportive and illustrative material for the text. They constitute the bulk of this book and, it is hoped, will be of interest for the nature of their various special fields. Appendixes 1, 2, 3, and 10 require a modest background in modern mathematics for comprehension. Appendixes 5, 7, 11, 12, 13, 14, and 15 contain relatively new ideas which we feel will soon prove fruitful in practice. The diagrams encountered in the various Appendixes illustrate well the variety of representations employed in systems engineering.

JUL 2 '69
HUNT LIBRARY
CARNEGIE-MELLON UNIVERSITY

The text itself attempts to give a cogent view of "the big picture." It will not, of course, make a "systems man" of a novice overnight. No detailed plans are given either for "design procedure" or for "organization of the system team." No kit of tools for systems engineering is given. If the reader is led to acquaintance with "the nature of the beast"—in this case, systems thinking—we feel success will have been achieved.

Major formulas are numbered consecutively. Since most of the discussion is either original viewpoint or "technical folklore," relatively few references appear. Since much of our recent work has been in airborne control systems, many of the examples are drawn from this area. This should not, however, be construed as limiting the scope of applicability of the concepts involved.

We are particularly indebted to Dr. George Kozmetsky and Mr. Leon Steinman, for many stimulating discussions in this area. Other colleagues who have given us help reflected in this book are: Al Boyajian, Ann Cameron, Dan Cameron, Joe Campeau, Neal Carlson, Bill Cass, Bob Chollar, Fran Dedona, Charles Gordon, Harve Hanish, Vic Hesse, Wayne Irwin, Ron Johnston, Wally Kantor, Bob Levinson, Ann McIntyre, Al Monroe, Art North, Nelson Parker, Frank Schneidermeyer, Ken Smith, Max Sosnow, H D Sprinkle, Jack Thorne, Ken Wilson, Jim Woodbury, and Roger Woods. We are indebted to Professor Sir Ronald A. Fisher, Cambridge, and to Messrs. Oliver and Boyd Ltd., Edinburgh, for permission to reprint an extract from their book *The Design of Experiments*. We are highly appreciative of the aid given us by Nancy Davidson and Alice Rishoff in preparing the manuscript.

THOUSAND OAKS, CALIFORNIA
AND
WOODLAND HILLS, CALIFORNIA

September 15, 1960

Table of Contents

1. BASIC SYSTEM CONCEPTS 1

Section 1.1 NOTION OF SYSTEM 2

1. Introduction, 2 2. System Definition, 3 3. System Description, 4 4. Missions and Systems, 8 5. Summary, 8

Section 1.2 SUBSYSTEMS AND COMPONENTS 9

1. Supersystems and Subsystems, 9 2. Subsystems and Components, 11 3. Synthesis and Projection, 11 4. Summary, 12

Section 1.3 SELECTION AND EFFECTIVENESS 13

1. Criteria, 13 2. Value, 15 3. Optimal Design, 16 4. Human Value and System Design, 17 5. Suboptimization, 17 6. Summary, 18

2. DESIGN AND ANALYSIS 19

Section 2.1 DESIGN OBJECTIVES 20

1. Imposed Constraints, 20 2. Design Freeze, 21 3. Design Objectives, 22 4. Summary, 22

Section 2.2 ANALYSIS 24

1. Analysis, 24 2. Analysis Classification, 26 3. Real Time Systems, 26 4. Requirements and Planning, 28 5. Signal and Noise, 28 6, Feasibility Analysis, 32 7. Prediction and Extrapolation, 33 8. Summary, 37

Section 2.3 DESIGN FLOW 38

1. Mark I and Evolutionary Designs, 38 2. Specification-Design-Analysis, 39 3. System Matrices, 40 4. Growth Potential, 40 5. Summary, 45

3. SPECIAL PROBLEM AREAS 47

Section 3.1 DESCRIPTION 48

1. The Descriptive Model, 48 2. Other Descriptive Problems, 50 3. Simulation, 51 4. Summary, 53

Section 3.2 COMPLEXITY 54

1. Measures of Complexity, 54 2. Difficulties Arising from Complexity, 55 3. Reliability Computations, 55 4. Reliability and Other Criteria and Design Objectives, 56 5. Exponential Decrease in Reliability with Increasing Complexity, 58 6. Exponential Decrease in Reliability, 58 7. Self-Checking, 59 8. Redundant Order Code, 61 9. Parity in Arithmetic Operations, 62 10. Summary, 63

Section 3.3 HUMAN SUBSYSTEMS 64

1. Human Problems, 64 2. Human Input Capabilities, 65 3. Human Sensory Abilities, 65 4. Human Perceptual Abilities, 66 5. Encoding and Decoding, 67 6. Evaluation, 68 7. Judgment Errors, 70 8. Man-Machine Comparisons, 72 9. Man-Machine Interfaces, 77 10. Summary, 81

Section 3.4 SPECIAL PROBLEMS IN CONTROL SYSTEMS 82

1. Feedback, 82 2. More on Thought Processes, 84 3. Comments on Biological Systems, 90 4. Summary, 90

4. OPERATIONAL APPROACH 91

Section 4.1 SYSTEM MANAGEMENT 92

1. The Systems Manager, 92 2. Direction of Procurement, 93 3. Information Flows, 94 4. System Management/Advisory Committee Cooperation, 96 5. Schedules, 97 6. Delegation and Communication, 100 7. Commercial Systems, 101 8. Summary, 102

Section 4.2 SYSTEMS SUPPORT FUNCTIONS 103

1. Support Functions, 103 2. Quality Control, 104 3. Summary, 106

5. SYSTEMS TECHNOLOGY FORECAST 107

Section 5.1 THE NEXT ERA—MAJOR AREAS 108

1. Predictory Assumptions, 108 2. Basic Automation, 110 3. Organizational Automation, 111 4. Compromise Automation, 113 5. Summary, 116

Section 5.2 THE NEXT ERA—SPECIFIC COMMENT 117

1. Basic Automation, 117 2. Organizational Automation, 119 3. Comments on Partial and Compromise Automation, 122 4. Planning, 123 5. Summary, 124

Section 5.3 L'ENVOI 125

APPENDIXES 127

Appendix 1 PRECISE DEFINITION OF SYSTEM 128

1. Terminology, 128 2. The Concept, 129

Appendix 2 NOTION OF GENERALIZED LOGICAL DESIGN 132

1. Terminology, 132 2. The Problem, 134 3. Further Discussion, 136

Appendix 3 NOTION OF DYNAMIC PROGRAMMING 138

1. Summary, 138 2. Terminology, 139 3. Dynamic Programming Processes, 141 4. Principles of Optimality, 142

Appendix 4 PARTIAL DESIGN OF THE HYPOTHETICAL SYSTEM: PORCUPINE 145

1. Introduction, 145 2. System Objectives, 146 3. Hypothetical Test Attack, 146 4. Porcupine Development, 147 5. Channel Capability Requirement, 147 6. Pre-launch Design, 148 7. Strategy Evaluation, 148 8. Necessity for Integrated Systems Apporach, 149 9. Information Flow for Porcupine in Attack Mode, 152 10. FCO Console, 155

Appendix 5 THE SYSTEMS MATRIX APPROACH 165

1. Partial Expansion of Attack Plane Computer Example into Matrices, 165 2. Manipulation of System Matrices, 169 3. Charts Showing Partial Expansion Examples, 175

Appendix 6 A MATHEMATICAL MODEL OF A CONVENTIONAL DIGITAL INTEGRATOR 226

1. Definition, 226 2. Schematic, 227

Appendix 7 GENERALIZED LOGIC—FOUNDATIONS OF A NEW DESIGN PHILOSOPHY 230

1. Summary and Definitions, 230 2. Scope, 231 3. Other Applications, 235

Appendix 8 PRELIMINARY CONSIDERATION OF AN ARMY SURVEILLANCE SYSTEM 236

1. Problem Areas Consideration, 236 2. Diversity of Communication Channel Types, 243 3. Radio and Wire Systems, 244 4. Possible Preference Bases, 247 5. Preliminary Remarks on Gross System Design, 249

Appendix 9 REMARKS ON LOW LEVEL REDUNDACY 253

1. Introduction, 253 2. The Diode Quad, 254 3. The Dual Diode Quad, 258 4. Notation for the Transistor Block, 259 5. The Transistor Block Quad, 260 6. Tentative Conclusions, 263 7. Further Work, 263

Appendix 10 ELEMENTS OF A BEHAVIORAL THEORY OF STATIC DECISIONS 264

1. Introduction, 264 2. Static vs. Dynamic, 265 3. The Static Decision Environment, 265 4. Estimates and Criteria, 266

Appendix 11 POTENTIAL MECHANIZATION OF SAFETY PROCEDURES DURING IN-FLIGHT EMERGENCIES: AN EXAMPLE 268

1. Introduction, 268 2. Presentation of Several Systems, 275 3. Presentation of Information to the Pilot, 276 4. Data Requirements for Automatic Flight Safety, 279 5. Explanation of the Various Instructions and Elements Involved, 281 6. Other Possible Systems Incorporations, 291 7. Concluding Remarks, 291

Appendix 12 RUDIMENTARY STUDY OUTLINE FOR SPACE-ORIENTED INFORMATIONAL SYSTEMS STUDIES 293

1. Preliminaries, 293 2. Objectives, 294 3. Need for Program, 295 4. Plan of Attack, 296

Appendix 13 RELATIVISTIC DOPPLER EFFECT, A NEW APPROACH TO SPACE NAVIGATION 302

1. Abstract, 302 2. Introduction, 304 3. Theory and Techniques, 305 4. System Considerations and Illustrations, 324 5. Interplanetary Guidance, 338

Addendum LIMITATIONS OF CONTEMPORARY TERRESTRIAL DOPPLER NAVIGATION RADARS 342

1. Limitations on Existing Navigation Radars, 342 2. Summary, 344

Appendix 14 APPLICATIONS OF ELECTROMYOGRAPHIC TECHNIQUES IN THE INTEGRATION OF MAN—MACHINE SYSTEMS 345

1. Introduction, 345 2. Preliminary Discussion, 347 3. Physiological Aspects of Neuro Muscular Activity, 348 4. Psychological Considerations of Behavioristic Learning, 358 5. Other Applications, 359

Appendix 15 THE TUNNEL DISPLAY CONCEPT 361

1. Introduction, 361 2. The Contact Analog, 362 3. The Tunnel Display Concept, 364 4. An Approach to Implementation, 374

INDEX 379

1

Basic System Concepts

"Kindred objects kindred thoughts inspire,
As summer clouds flash forth electric fire."

—ROGERS

Section 1.1

Notion of System

"System: an organic or organized whole"

—WEBSTER

1. Introduction

Newspapers are apt to characterize our age as the "Jet Age," "Space Age," or the "Atomic Age." People more familiar with the details of modern technological concepts, however, probably would choose "The Systems Era" as being a more accurate descriptive phrase. The jet engine and nuclear weapon, for example, become parts of a "weapons system," and the nuclear reactor part of a "power distribution system." Management makes use of "systems concept," "systems philosophy," and "systems approach." Engineers and physical scientists speak of "systems analysis," "systems engineering," and "systems theory." Medical, biological, and behavioral scientists discuss "nervous systems," "homeostatic systems," and "social systems." If one asked these people what the word "system" means, many diverse answers utilizing the various professional languages would be obtained. If these answers were translated into simple English and compared to each other, the notion common to them

all might be stated as such: "A system is something which accomplishes an operational process; that is, something is operated on in some way to produce something." That which is operated upon is usually called input; that which is produced is called output, and the operating entity is called the system.

2. System Definition

A *system* is a device, procedure, or scheme which behaves according to some description, its function being to operate on information and/or energy and/or matter in a time reference to yield information and/or energy and/or matter. Schematically, we may indicate this as in Fig. 1.1.

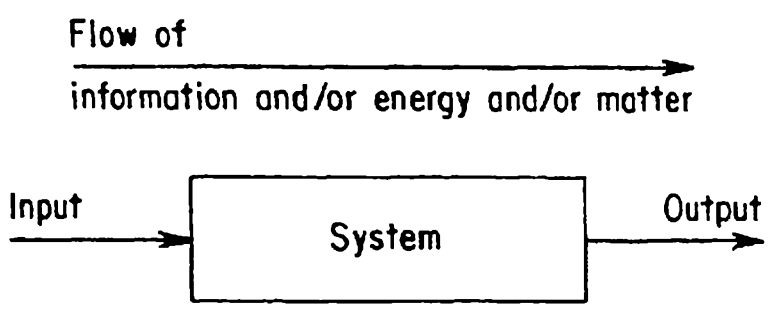


Fig. 1.1. Broad-brush system schematic.

Input and output consist of complexes of information and/or energy and/or matter. The description of behavior (that is, the nature of transformation) may be a deterministic mathematical model, or it may be one that also involves stochastic variables (i.e., randomness). In addition, the description may admit certain variables reflecting the instantaneous condition of the system itself. These possible conditions of the system are called *states*, and their totality is called the *phase space* of the system. A precise mathematical definition of "system," which supports this intuitive definition, is given in Appendix 1.

The term "system" emphasizes that an over-all operational process is under consideration rather than a collection of pieces. No specification has been made as to the complexity of the operational process. Thus, a jet engine is regarded as a system when the process under consideration is conversion of chemicals to thrust, but only a portion of a system when the process is controlled translational motion arising from chemicals. In general, the operations performed by a system depend upon the condition of the system itself, and such conditions (e.g., internal state) may be randomly, historically, or externally determined at various times. Examples of systems are:

(a) A magnetic-core-pulse-circuit which accomplishes the "or" function in logic;

(b) An executive predicting his company's near-future performance on the basis of statistical historical information;
(c) An industry which makes nuclear weapon cores from essentially raw materials.

The choice of terms such as "system," "input," "output," "states," "phase space," and others depends upon standard usage. This choice is certain to be arbitrary if several terms are common.

Note that the notion of system mentioned above applies equally (among other items) to operational processes, physical, or biological devices implementing operational processes, and mathematical models in which some variables are given and others sought. The properties, circumstances, occurrences, and relationships concerning appropriate phenomena may be related by a common vocabulary and a single set of concepts. These yield sufficient generality to describe widely diverse systems without reference to their expression in nature or by artifice.

Systems may ordinarily be classified into three gross categories:

(a) Natural systems are of organic or physical origin. They have antecedents and elements largely beyond the control of man, though not necessarily exclusive of man;
(b) Devised systems are products of design and elaboration by man. They have their physical or organic substances and/or forms modified by the intervention of man;
(c) Hybrid systems are combinations of natural and devised systems.

Natural systems impose the environment containing systems of other classes. Our interest may focus upon them to the extent that their presence affects the design of other systems. Devised and hybrid systems that reflect man's ingenuity, however, are the systems of immediate concern in this book.

3. System Description

The description of a system involves the specification of the following:

(a) The nature of inputs;
(b) The nature of outputs;
(c) The system phase space;
(d) A descriptive model relating inputs, outputs, and system states in time.

Further points to be observed are:

(e) Some notion of time sequencing is required;
(f) A distinction must be drawn between suitable inputs and outputs, and those actually permissible for particular runs of system operation;
(g) During any run, the input and state descriptions must determine (perhaps as a distribution) the output;
(h) Outputs should be essentially identical when inputs and state changes are essentially identical.
(i) The output, at any particular time, in any particular run, should depend only on the history and *not* on the future of that run, up to that time.

These points are accounted for in the precise definition of system in Appendix 1. They actually determine, of course, the rigorous definition together with level of generality assumed.

Our statements in the preceding paragraphs about system description are deceptively simple. This fact will be evident later in the book. However, items (a–d) must be identified, for even the least detailed description usually proceeds by iteration. Each item is described roughly and then, subsequently, in terms of increasing detail.

A. Examples

It may now be helpful to consider a few types of systems.

(a) In classical mechanics, a system of n mass points has a $6n$-dimensional phase space consisting of all possible combinations of positions and momenta. Inputs are forces (function of time), and outputs are sets of trajectories from the initial condition at input application (i.e., system state). A descriptive model for this system is the set of differential equations obtained from Newton's law,

$$F = ma. \tag{1}$$

This system is, of course, a deterministic one.

(b) A classical machine that exemplifies the deterministic, mechanical class of systems is the pendulum governor of a steam engine. The engine rotates a shaft from which two iron balls are suspended. As the velocity of shaft rotation increases, the iron balls move outward (owing to the

effect of the centrifugal force exerted on them). The plane of motion of the balls to the shaft, the vertical axis, is then a function of the velocity of the shaft. Increasing or decreasing velocities move this plane up or down the vertical axis. When the plane approaches the horizontal, an array of levers is actuated to close a valve and to depress a throttle arm, thus decreasing the velocity of the shaft rotation. When the velocity of rotation is too slow, the levers attached to the iron balls keep a throttle open.

The pendulum governor arrangement illustrates several features of interest. During operation, input (information indicative of the angular velocity of the driven shaft) is given by the relationships momentarily existing among the components of the pendulum-lever array. The momentary condition of the pendulum-lever array constitutes one of the states within the phase space of the system. Successive states themselves, by means of mechanical linkages within the system, provide outputs (actuation means) for valve closure and throttle control. This is but one (partial) description of the system. We might wish to regard the range of variations in steam pressure as the phase space of the system. Or, we may prefer to consider the angular velocities of the shaft, or the linear travels of a piston per unit time as states of the system.

(c) Another simple classical example is the customary concept of a fixed-product factory whose inputs consist of personnel, product orders, raw or semi-processed materials, and commercially available forms of energy. The states are truly factory conditions in the usual sense, and outputs are product batches. The descriptive model is the industrial engineering layout and process description. In this system we find stochastic variables entering both in inputs (personnel condition variance, for example) and in states. This example indicates that a system may be described in many ways owing to trade-offs between definitions of states, inputs, and outputs (in this connection the personnel conditions might have been included in states rather than as input properties) and also owing to different levels of detail as are suitable to the purpose of description.

The purpose for which the description serves conditions the level of detail sought in preliminary analysis. If an inclusive operational description of the system were the objective, it would be necessary to isolate all elements involved in the system, and to determine precisely the nature of all relationships through time which all elements would bear toward all others. Ordinarily, in the case of operational systems it is not necessary to undertake, even theoretically, such a massive venture. A more reasonable

approach would be incremental and would proceed in a step-wise fashion upward in the complexity scale.

(d) An even simpler example of a deterministic system than (a) is that of a counting register in a digital computer. In practice, digital computers are presently limited to the use of bistable devices characterized by high reliability. Having only two possible electrical states, the storage elements can represent only two digits. Thus, it develops that digital computers are dependent upon a system of numbers composed of only two digits. Whereas on paper we are accustomed to arranging ten different digits into all possible numbers, the computing device can arrange only two digits to represent those same numbers. Since a flip-flop or other storage element can store only a single binary digit, the presentation of a complete number requires a registering device with as many storage units as there are digits in the binary expansion of the number. Here, the state at any given time is the number the register holds at the time in binary expansion; the inputs are "bits" (that is a "0" or a "1"); and the outputs are the resulting states with, perhaps, overflow registry. Note that here the same object may play different roles of state and output at different times. The descriptive process noted is Peano's successor function modified by modularity. This function may be defined by:

$$f(1) = 2; f(n + 1) = n + 2; n = 1, 2, \cdots. \tag{2}$$

(e) As a more complex example we may consider a strategic bomber system in which the inputs include (among hundreds of items) mission definition, estimate of present position at any time, condition of environment with respect to electronic countermeasures, etc.; the descriptive model is partially unknown and is of such complexity as to necessitate its replacement by gross approximations; output is a choice between mission completion or abort. The design requirements for the system include recognition of the vast geographical distance between one nation and its potential enemies. Considerable emphasis is given provisions for self-preservation of the craft and its members during flight in order that the aircraft system may be virtually assured of successfully performing its dual function as a weapon of psychological chastisement and physical destruction. The strategic bomber system is a contingently effective system unless we are able to define its value in terms of comfort rendered the supporting populace.

4. Missions and Systems

To avoid confusion, it is well to point out that in the usage customary to military technology the operational processes considered are usually called *missions*, and the physical complexes implementing these are called *systems*, even though both definitions may satisfy our present notion of system.

This distinction between mission and system is a convenient one and will be used whenever we wish to consider systems intentionally designed to carry out specified processes. It might be thought that systems engineering deals exclusively with events after mission specification. In fact, many common techniques apply to both mission design and system design and, as we shall see later in the book, these generally modify each other.

5. Summary

In this section we have defined the concept of system and considered the associated concepts of input, output, and state. The examples we have given illustrate the wide class of objects which it may be convenient to regard as systems.

Section 1.2

Subsystems and Components

"Here and elsewhere we shall not obtain the best insight into things until we actually see them growing from the beginning . . . ."

—ARISTOTLE

1. Supersystems and Subsystems

It is clear that if the outputs of a system A are of the same nature as the inputs of a system B and are otherwise restricted as necessary, we may combine A and B as indicated in Figure 1.2 into a system C having the same nature of inputs as A, the same nature of outputs as B, and a phase space lying in the product (in the usual Cartesian sense) of the phase spaces of A and B (the entire product would be utilized unless some combinations of A-states and B-states are outlawed). Similarly, regardless of the nature of systems A and B, a system D may be formed as shown in Fig. 1.3 whose inputs and outputs are time-indexed complexes of those of A and B taken together, and whose phase space again lies in the product

of those of A and B. Both C and D are *supersystems* of A and B; A and B conversely are *subsystems* of either the system C or D.

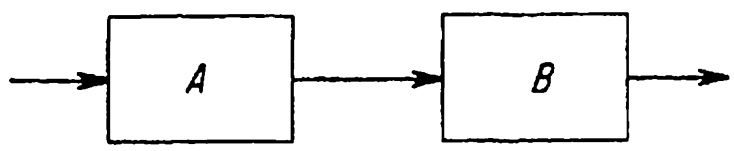


Fig. 1.2. C: Series-combination of systems.

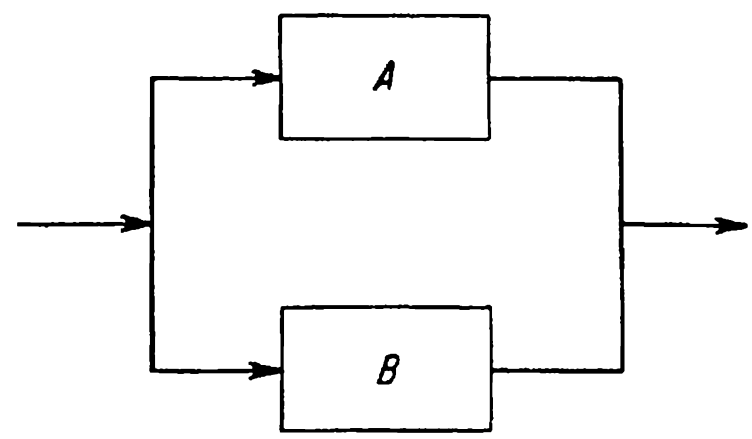


Fig. 1.3. D: Parallel-combination of systems.

A communications network system affords an illustrative example of series-combination. Certain properties, i.e., inputs, phase space, and output attributes entering into the communications structure, are independent of the environment in which the system functions (submarine or surface media, for example) and of the immediate purposes which the system serves. Thus, though extent of dispersion and displacement patterns of the subsystems comprising the supersystem may vary radically and message traffic rates undergo extreme variation from segment to segment, the same functions are performed by all segments of the system-information transfer. Information transfer is accomplished by equipments that are functionally similar and permit compatible linkages, making possible the extension and maintenance of communication services.

Troop deployment practices provide examples of subsystem additiveness without loss of fundamental characteristics through dispersion of the unit element. In military philosophy, to the extent that men are similarly motivated, similarly disciplined, and physiologically and psychologically maintained, they may be lumped more or less into a variety of units. Alternatively, they may be individually displaced in the field without altering the essential characteristics of the soldier. Although it is recognized that individual differences do exist and, further, that the state of the individual soldier may vary, these considerations rarely enter explicitly into military engagement planning.

2. Subsystems and Components

By additive processes, series-combinations of suitable systems, interlaced with parallel-combinations as desired, may be constructed into larger and larger systems. Thus, hierarchies of subsystems may be developed: subsystems of subsystems of subsystems, etc. In any given discussion of larger systems, the lowest level of a system is traditionally distinguished by being called a *component*. It should be emphasized that the terms *subsystem* and *component* are purely relative to some established context. This relativity of viewpoint may cause much confusion if not clearly understood. The designer of an aircraft-type of vehicular weapons system speaks in his shop of radars, computers, and airframes as components. A subcontractor may speak in his shop of the computer† as "the system," of a magnetic drum memory as a subsystem, and of read heads for the drum memory as components. When the prime and subcontractor hold technical information exchanges, each tends to think his own language is that being used by the other. This hierarchy effect is illustrated in the table of Fig. 1.4.

SYSTEM HIERARCHY EXAMPLE

Agency	*System*	*Component*
CDR in Chief	National Defense System	Strategic Air Command
SAC Chief	SAC	Bomber wing
Wing CDR	Bomber wing	Bomber
Prime contractor	Bomber	Computer
Subcontractor	Computer	Logic circuit
Circuit designer	Logic circuit	Diode

Fig. 1.4

3. Synthesis and Projection

The process of constructing supersystems is called *synthesis*. The inverse process, that of reducing a given system to subsystems, is called

† A more intelligent approach by a prime would generally be to place the so-called "informational subsystem" (computers, displays, and computer and display support equipment) with one subcontractor.

projection. There are many ways in which a given system may be projected into subsystems. Generally, projection is restricted by placing constraints on the class of allowable subsystems into which the projection is intended. Although it is not included in this book, work has been carried out to establish partially a formalism of projections based on the theory of oriented linear graphs.

Rarely are synthesis or projection used in simple forms. In system design, the systems engineer generally tries alternate syntheses from subsystems known to be available or under development. Similarly, projection of a given system into alternate classes of subsystems is quite common since it enables selection of techniques subject to criteria, as we shall consider in the next section. While synthesis for a specified mission is a creative art, projection can, in some cases, be accomplished methodically. Some such cases are discussed in the second part of Appendix 2 under the heading of "Generalized Logical Design." Synthesis and projection usually intermingle in system design even as do invention and deduction in all engineering activities. The highlights of an example illustrating the interplay of projection and synthesis, as well as system hierarchy effects, are given in Appendix 4 entitled "Partial Design of Hypothetical System: Porcupine."

4. Summary

In this section we have discussed making "little ones from big ones" and "big ones from little ones." Although synthesis is usually regarded as the initial step of systems engineering, in complex designs, projection and synthesis function as correlates of the requirements and capabilities relations that will be discussed in Section 2.2. (Analysis). Appendix 4 illustrates this point.

Section 1.3

Selection and Effectiveness

"I've measured it from side to side,
'Tis three feet long and two feet wide."
—W. WORDSWORTH

1. Criteria

We must consider briefly now the basic questions of *value* associated with systems concepts. These may arise at several distinct levels in systems considerations. Suppose, for example, it is desired to select among several alternatively proposed systems for implementing a given mission. Before an intelligent choice is possible, we must answer as a minimum the questions:

(a) What constitutes adequate performance of the mission in terms of mission parameters?

(b) Should two or more systems appear capable of adequately performing the mission, what is a further measure of performance in mission parameters that will indicate which system will better perform the mission?

(c) For each proposed system, what functions of system parameters do the mission performance criteria involve?

HUNT LIBRARY
CARNEGIE-MELLON UNIVERSITY

(d) For each proposed system, what is the sensitivity of the system-parametrized mission performance measures to design detail, availability of subsystems in the state-of-the-arts, and quality of systems support?

(e) How shall the cost or penalty factors associated with procurement of each proposed system be computed in order to obtain a basis for comparison?

The answers to these questions are necessary and not merely window-dressing. Unless a system is chosen which is at least adequate for the mission, no system may well be preferable. If system performance is too sensitive to developmental and operational factors, fallibility is likely to intervene. If several systems could carry out the mission, a choice should be made on a rational basis. Some possible considerations are:

(f) Select that system maximizing effectiveness which satisfies some cost constraint. A homely example is provided by the case of the householder who is constrained to operate within a minimal budget. The householder is disgusted to find a leaky roof. He asks for estimates of the cost of adequate repairs. Finding none within his limited means and lacking credit, he may "do-it-yourself." He accumulates the necessary materials, borrows a ladder and a few tools, and applies his energies.

(g) Select that adequate system of minimum cost. During the westward expansion of this country in the later half of the nineteenth century, there was general need among newly established settlements for the services of a miller. Prospective local millers were familiar with the existence of metal shafts, pulleys, and associated milling gear used in developed communities, and they were also keenly aware of the cost and time involved in acquiring such components for their mill-site. Thus, they devised adequate milling systems from local materials. Hardwoods, suitably seasoned, were fashioned into shafts, pulleys, etc., that were required for the operational success of the venture. Such scraps of metal as were available were used to strengthen stressed areas. Leather was frequently substituted for metal on bearing surfaces. Such primitive systems, though of slight dollar cost, were adequate for satisfying the milling needs of many early American communities.

(h) Select that adequate and not-too-sensitive system within some cost constraint maximizing some specified criterion function monotone increasing in effectiveness and monotone decreasing in cost. Present-day double entry bookkeeping procedures afford common examples of this type of system. The financial success (or performance) of the thoroughly

documented enterprise may be measured and compared historically on the basis (criterion function) of net profit per unit period.

What type of selection basis is used will depend, of course, on weights assigned value of mission capabilities, scarceness of resources, and many other factors in the politico-economic environment. This is a matter which is quite difficult and should be of major concern, since selection of improper criteria for optimization means beginning work solving the wrong problem. By *optimization* we mean merely maximization or minimization as appropriate of the selected criterion function(s).

2. Value

One of the major difficulties in this area is the lack of adequate theories of value. Thus, rough and not always meaningful attempts are usually employed to reduce performance to some simplified parameter and to convert penalties to, perhaps, equivalent dollar costs.

All theories of value are, of course, products of the mind of man. One's own directions of interest enter the arena of thought immediately upon the introduction of abstractions such as value. The social scientist, for example, is apt to consider value in a social context, possibly humanitarian. The military man is likely to think of this concept in terms of destruction potential remaining. The cyberneticist may feel that value is measured indirectly as a function of energy conserved or of approximate maintenance of entropy level. It is clear that *criteria* are formed from some sense of *value*. Total value is not a simple measure, such as the parameters of energy conserved, kill potential, or sale price. Value concepts actually include many things much more difficult to grasp numerically. These simple parameters are only substitutes with which we work compromises. Thus, we speak of (but do not measure) "sentimental values" of objects inherited or received as gifts. An excellent example of the difficulty involved in measuring a problem with connotative value associations is given in the preliminary material of Appendix 4, in which we grope to measure "maximum protection of the task force."

The "value space," defined in Appendix 10 (entitled, "Elements of a Behavioral Theory of Static Decision") as an upper conditionally complete vector lattice, is probably the closest approximation that can be achieved in developing a concept that includes all the elements we speak of as "value."

3. Optimal Design

The concept of optimal design is one which is not customarily well-defined. In forming an information processing system by the interconnection of units (subsystems), for example, some of the interpretations possible are:

(a) That configuration in which the least requirements (complexity, bandwidth, accuracy, etc.) are placed on the units;

(b) That configuration in which the information is most reliably passed from one unit to another;

(c) That configuration in which the inter-unit equipment (wires, cables, amplifiers, etc.) is kept to a minimum;

(d) That configuration in which the inter-unit translations (transformations: AC to DC, analog to digital, etc.) are kept to a minimum.

As we recognize, these are criteria reduced to simplified parameters as mentioned previously under *value*.

Actually, it is only in rare cases that optimization in a single one of these criteria is apropos. A much more appropriate concept is that of an allowable region in the multi-dimensional space defined by the system parameters and constraints on them in which a weighted value function accounts for suitable mixtures of the optimizations criteria. This, too, is a compromise type of criterion, but superior to those commonly in use. The weighting need not be linear. For purposes of example, this concept is presented in Appendixes 2 and 3 in two simple systems contexts. The concept should be expanded to the level of system design and exploited rather than merely expounded. It is within such a context that well-defined regions may be delineated giving meaning to the concepts of adequacy, "goodness," and finally, optimality. Within such a context, system design techniques that permit simultaneous preparation of computer programs may be developed. These could be used to select immediately not only optimal cases but other cases sometimes of interest, such as "minimal (in cost, or weight, etc.) but adequate," which actually represent saddle-point and/or constrained variational problems.

4. Human Value and System Design

Within our culture it is undesirable to equate the value of a human life against the value of inanimate equipment. In the absence of adequate value concepts the systems designer generally makes the assumption that an individual in a system must suffer no deprivation or minor injury, even though compensable or reparable.

5. Suboptimization

One of the major dangers in optimization of choice among several alternative systems or of choice among alternative detailed designs of a given system is that of reduction to suboptimization. It may be that in criterion selection one relevant factor is overweighted at the expense of others equally important, or that some subsystem is optimized rather than the over-all system. Some examples are:

(a) In the choice of a central control computer for a manned aircraft, physical weight and volume (i.e., compactness) are certainly of significance. To base a selection, however, on a difference of weight or volume which is small relative to the normal variations in weights or volumes of the crew members, is surely foolish unless all other factors are truly equal.

(b) In a business or military organization it is all too frequently found that some organizational subsystem is more concerned with maximizing performance and growth in its own functions than in furthering over-all organization objectives.

(c) In systems utilizing both human and machine subsystems, it is a common fault to suboptimize around the humans. While traditional human engineering strives to permit humans to perform most easily and reliably their functions within the systems performance context, it sometimes does so to the detriment of over-all system performance.

Clearly, the entire question of optimization versus suboptimization is merely a restatement of the question of balanced judgment versus overemphasized consideration of certain criteria (frequently the result of simple failure to understand all criteria). Unfortunately, experience indi-

cates that in real-world situations suboptimization almost invariably precludes optimization. On the other hand, simple theoretical examples can easily be found in which optimization is consistent with certain suboptimizations. This consistency merely means that the suboptimizer's criterion is also selected by the optimizer as his criterion. An example would be a corporation in which both the board of directors and the comptroller's unit elect to maximize,

$$\int_0^\infty n(t)\ dt, \tag{3}$$

where $n(t)$ is net profit. This is not, of course, a practical election since such factors as cash on hand, commitment of resources, survival of the firm, pressures from stockholders, etc., must be accounted for in an intelligently run business.

6. Summary

In this section we have considered the concepts of value, criteria, optimality, and suboptimality. Things to remember are:

(a) "Value" is a very difficult concept, and compromise substitutes are often used instead to formulate criteria;
(b) In the real world, suboptimization generally precludes optimization. Therefore, one good reason why systems are "system-engineered" is that, supposedly, all aspects are examined.

2

Design and Analysis

"When any great design thou dost intend,
Think on the means, the manner and the end."

—DENHAM

Section 2.1

Design Objectives

"Accurst ambition, how dearly I have bought you"
—DRYDEN

1. Imposed Constraints

As has been previously indicated, the primary consideration in system design is that of obtaining a system adequate in performance which fulfills the system requirements arising from mission specification and satisfies whatever additional constraints are specified. These additional constraints frequently have the nature of generalized commodity limitations such as:

(a) Limits on time to delivery;
(b) Limits on development and/or production costs;
(c) Performance limits of a vehicular system utilizing an informational subsystem under design.

An outstanding example of sacrifice of operational effectiveness to traditional constraints in the design of airborne weapon systems is provided by the record of U.S. fighter planes' combat performance during

the Korean action. The principal competitor for air supremacy against these jets was the Soviet MIG. Both aircraft were of the latest state-of-the-art in contemporary fighter systems. The pilots of both types of aircraft were probably equally motivated and were certainly expertly trained. Soviet design, however, sacrificed liberally the traditional constraints bearing upon the pilot's personal safety through the reduction of armor and the virtual elimination of various secondary protective measures. The MIG's were considerably lightened and, therefore, more maneuverable. The frequent tactical consequences, as a result of the additional loading of personnel safety features (suboptimization about the pilot), were that the U.S. planes were outperformed by the MIG's. This relative maneuverability was tested subsequently with a MIG delivered to the U.S.A. by a Soviet defector. The tests showed the MIG definitely superior. The U.S.A. nevertheless had frequent combat success over Korea owing to the Soviet action of rotating pilot units from all parts of Russia and its satellites in order to spread combat training. A temporary measure of pilot safety during the early phases of the individual engagement, dearly bought in terms of decreased maneuverability, resulted in the enhanced probability of double loss of pilot and aircraft.

These additional constraints are artificial as compared with systems requirements arising from mission specifications. They are not, however, necessarily artificial within the systems context and, hence, may best be referred to as *imposed constraints*. Such constraints directly or indirectly influence preliminary system conceptualization and design in a significant fashion.

2. Design Freeze

There appears to be no valid procedure by which we may predict the exact appearance of technological breakthroughs nor, for that matter, the effects, immediate or otherwise, which such breakthroughs may have upon companion equipments. The practical solution to these problems of incertitude usually assumes the form of "freezing" system design significantly in advance of prototypic developmental models.

The conceptual design of the system is stabilized not as the result of knowledge or appreciation of disparate rates of advance in componentry or subsystems arts, but as the result of conjectured limits on permissible elapsed time to delivery.

3. Design Objectives

In addition to system requirements as such, propositions called *design objectives* are frequently present in a design problem. These vary from a basic desire to utilize certain classes of phenomenological devices to the attainment of additional systems capabilities (beyond those required for the specified mission) at essentially no additional cost in time or money which may or may not be feasible in the design context. Examples of each end of this spectrum are readily found in airborne digital control systems. Specifications frequently call for the use of the latest developmental semiconductors rather than those tested by long experience, or for extensive use of solid-state components for a system in which other components have been proved effective. Specifications may also call for silicon rather than germanium semiconductors because of their higher thermal operating range. This is certainly a luxurious capability in a machine destined to occupy, with the pilot, an air conditioned cockpit.

In an oversimplified sense, design objectives specify system characteristics which are regarded as desirable but not essential for mission accomplishment. One of the most frequently recurring examples is that of imposing research-type targets on a developmental system. Thus, in specifying a system for development in an area where state-of-the-art life to first failure is between 200 and 500 hours, it is sometimes stated: "mean time to failure of 1000 hours shall be a design objective."

Design objectives are not necessarily a good or a bad thing. They have been with us in commercial areas for many years under the concepts of packaging and other marketability considerations. A product designed to accomplish a function, if it is to sell in the presence of competition, must be attractive to prospective buyers. It must not be so large or ill-shapen as to appear clumsy, nor so small as to fall psychologically beneath its price. Thus, shape, format, and size become objectives of design *whether or not they have anything to do with accomplishment of function.*

4. Summary

We have classified various constraints upon systems into three classes:

(a) *Systems requirements* or *primary constraints* which arise from mission specification;

(b) Imposed constraints which are usually generalized commodity limitations;
(c) *Design objectives* which are reflections of goals other than mission or function accomplishment.

Section 2.2

Analysis

"Consider well what your strength is equal to, and what exceeds your ability."
—HORACE

1. Analysis

Most devised and hybrid systems defy easy comprehension by virtue of inherent detail. In consequence, not one form of analysis, but many are required to describe the internal relations of the system as well as interactions with its environment.

In the conceptual approach to system formulation it is often of value to reduce the complex system problem to those elements with which one expects to exhibit gross causative or responsive effects within the system. In some instances, the major features of the system may be planarized, so to speak, within a conceptual structure which retains the significant features of the system of interest. Where sufficient time and imaginative talents are available, this approach may be of considerable heuristic value.

In the techniques common to the experimental sciences, the approach

tacitly conveys the assumption that it is possible to attain valid inferences from experimental results. Fisher's† words are relevant in this respect,

> "I have assumed . . . that it is possible to draw valid inferences from the results of experimentation; that it is possible to argue from consequences to causes, from observations to hypotheses; as a statistician would say, from a sample to the population from which it was drawn, or, as a logician might put it, from the particular to the general. It is, however, certain that many mathematicians, if pressed on the point, would say it is not possible rigorously to argue from the particular to the general; that all such arguments must involve some sort of guesswork, which they might admit to be plausible guesswork, but the rationale of which, they would be unwilling, as mathematicians, to discuss. We may at once admit that any inference from the particular to the general must be attended with some degree of uncertainty, but this is not the same as to admit that such inference can not be absolutely rigorous, for the nature and degree of the uncertainty may itself be capable of rigorous expression. In the theory of probability, as developed in its application to games of chance, we have the classic example proving this possibility. If the gamblers' apparatus are really *true* or unbiased, the probabilities of the different possible events, or combinations of events, can be inferred by a rigorous deductive argument, although the outcome of any particular game is recognized to be uncertain. The mere fact that inductive inferences are uncertain cannot, therefore, be accepted as precluding perfectly rigorous and unequivocal inference."

There are those individuals who, as a result of specialized training if not through personal inclination, prefer to approach problem solving on the basis of generalization from special cases (induction). There are also those of opposite natures who are most comfortable and, possibly, most efficient when they seek out directly the generalized expression descriptive of the phenomena of interest. As in most life activities, intellectual or otherwise, moderation is probably a "good" thing. The scientific generalist will reflect his type by selecting and combining from both deductive and inductive techniques those approaches that appear promising. The more conventional scientist or engineer will probably favor a step-wise approach based upon an examination of a series of demonstrable events. Clearly, at the level of preliminary conceptual development toward the system problem solution there is no question of which approach is "right" or

† R. A. Fisher, *The Design of Experiments* (New York: Hafner Publishing Co., 1953), pp. 3–4.

"wrong." Both types of approaches are applicable and of merit; both have demonstrated ample justification for their inclusion within problem solving philosophy.

The variety of analytical approaches applicable to a given system is conditioned by the developmental stage in which the system exists at the time of analysis. At this point let us restrict our area of interest to complex hybrid systems analyses and consider in further detail those approaches which are roughly classifiable in terms of objectives.

2. Analysis Classification

In addition to frequent wholesale quantities of mathematical and numerical analysis of a miscellaneous nature in systems problems, it is possible to distinguish by goal orientation three primary types of analysis. These are:

(a) *Requirements analysis* attempts to answer questions of the form: What capabilities must a system possess in order to implement this given mission or to meet this given specification?
(b) *Capabilities analysis* attempts to answer questions of the form: What missions can this given system implement or what specifications can it meet?
(c) *Feasibility analysis* attempts to answer questions of the form: Can this given system implement this given mission or meet this given specification?

Generally, all three of these types of analysis are required in a reasonably complex systems problem. We shall note in the next section how they are related to the over-all problem. Here we shall merely consider specific examples of the questions each type attempts to answer.

3. Real Time Systems

Associated with any system context is a time scale usually called real time. Suppose that at some initial time α, inputs adequate to carry out the required processing are available and the results (i.e., outputs) of processing these particular inputs are required by the systems context at time β, where

$$\alpha < \beta. \tag{4}$$

Suppose further that the time required for processing is γ and the time required to distribute appropriately the results of the processing is δ (δ is frequently called transportation lag). In order for the system to do useful processing, we must have

$$\alpha + \gamma + \delta \leq \beta. \tag{5}$$

If

$$\gamma = \beta - \alpha - \delta, \tag{6}$$

the system is said to be operating in real time. If

$$\gamma < \beta - \alpha - \delta, \tag{7}$$

the system is said to be faster than real time. Should a system be faster than real time and also possess capabilities to carry out secondary missions with respect to which it is also faster than real time, it is sometimes possible to interlace processing so that more than one mission can be accomplished in real time. Such a possibility may well be reflected in a design objective calling for faster than real time operating with respect to the primary mission.

One of the basic questions, then, to be answered by requirements analysis on a system is to determine $\beta - \alpha - \delta$ and thus establish minimum allowable rate of processing. Such determinations may be propagated through an extended hierarchy of subsystems. If, for example, the system under consideration is a control digital computer which repetitively utilizes continuously available sensory data to solve a set of equations yielding position of a vehicle, accuracies of the sensory data and performance characteristics of the vehicle together with specification of the mission of the vehicular supersystem must establish the basic rate (called iteration rate or major cycle) at which the computer must furnish outputs. This rate, together with computer system design, must in turn establish limits on memory access times, logical switching delays, etc. These limits together with consideration of such component operating characteristics as duty cycles in turn establish limits on component operating times and, hence, at least in a synchronous computer, ultimately establish a lower bound on clock rate. It should be noted, however, that many trade-offs exist when a variety of computer designs are considered such as paralleling of logical operations to reduce clock rate requirements while maintaining over-all iteration rate.

4. Requirements and Planning

Requirements analysis applies, of course, not only to specific problems, but is, in fact, the basis of long range planning. In order to plan intelligently for participation in the technological business opportunities of the future or for discharging a responsibility such as national defense, it is essential to extrapolate conditions and trends so that one may anticipate the nature of systems which will be needed in the future. Generally, of course, the sharpness of requirements detail falls off rapidly with extent of extrapolation. The planning aspects are considered at greater length in Chapter 5 of this book.

5. Signal and Noise

As in the case of requirements analysis, capabilities analysis operates at both "broad-brush" and detail levels. It is necessary for those engaged in work in any particular systems area to stay abreast of the states-of-the-arts impinging on that area if they are to remain competitive either in the sense of business competition or in the sense of military potential. This in turn implies continuing capabilities analysis of new developments in the impinging areas either of a "technological breakthrough" nature or merely of an evolutionary improvement nature.

In *real time systems* above, we considered the question of processing rate of a system. Of almost equal importance in some missions involving information subsystems are the accuracy, precision, and signal/noise ratio of the sensory subsystems since, clearly, the nature of inputs to an information processing system limits information characteristics of the outputs. The basic questions bearing on signal/noise ratio may be conceptualized as follows:

Measurement is plagued by statistical difficulties generally called collectively "noise." In any discussion of the study of the nature of signals in the presence of noise, there must be a clear understanding of the definition of the basic terms, signal and noise. To illustrate best the meanings that are generally imparted to these words, a simplified model will be used. If there exist two objects, one of which is termed the observer and the second, that which is being observed, then there exist various relationships between the two objects. One such relationship is the dis-

tance between the two. A second is the velocity of the observed object relative to that of the observer. In any event, the fact that these relationships exist independently of the measuring equipment that will later be used to measure these relationships is intuitively plausible. In this sense the relationship of distance between two objects cannot contain noise, nor does it constitute what is termed signal. It is only when measuring equipment is applied that both signal and noise are identified. Using for the sake of discussion the illustration of the relationship of distance, it can at once be realized that if a steel tape were used and were stretched between two points, measurement could be obtained. The relationship of distance, although a constant, could easily produce a large number of different readings of the tape, owing to effects such as temperature variations causing expansions or contractions of the tape, or stress being applied to the tape which causes it to stretch or contract. In any event, any given reading from the tape represents the combined effects of both signal, which is that part of the reading that exactly corresponds to the relationship, and noise, which is the additional or modified portion of the reading owing to other extraneous effects. The definition then of signal is as follows: signal is that portion of an instrument reading which has direct correspondence to the relationship being measured. The definition of noise is simply the actual instrument reading minus the reading on the instrument if it were reading only signal. *It is to be remembered that we are only considering instruments to illustrate the concept of noise. Any spuriously generated portion of any signal or perception is "noise."* Noise poses a tremendous number of problems, because it is created by such a wide variety of causes. In the case of a radar, for example, that is being used to determine the range from an observer to an observed target, noise is introduced into the reading of range by virtue of angular and amplitude scintillation. The amount of scintillation and, hence, the magnitude of noise depend upon the general shape of the target, its rates of motion through its environment, atmospheric conditions, the value of the range itself, and a large number of other causes that cannot be easily controlled or measured in themselves.

Of interest in this regard is the type of noise introduced within the internal systems environment by the inclusion of a human being. Rarely is the design of a man-machine system of such a simplified arrangement as to permit the human component to act as a single element having but a simple iterative act to perform that may be linearly predictable. In many modern high performance systems the human operator may be called upon to perform a variety of acts or to intersperse a variety of

observations within a sequence of acts which are contingent on observations. To reduce system instabilities which may be precipitated by the inherently fallible operator acting as an integrator of multiple information formats and action patterns, many arrangements have been proposed and implemented which have the effect of "unburdening" the operator of much of the complexity of an unaided control situation. The operator may be provided also with immediate knowledge of the effects of his own responses through a variety of control techniques generally classified as "quickening." (Both unburdening and quickening are discussed at greater length in Chapter 3.) Unburdening and quickening are of value in attaining optimal man-machine control system performance. Both techniques, however, constitute noise sources within the system in that modifications of the direct correspondence between sensed state and sensory indicator are necessary to compensate for the inclusion of the nonlinear component (man) within the system.

It is not possible to separate completely that portion of a reading from an instrument corresponding to the relation that is being measured and that portion contributed by noise. Obviously, a great deal must be learned about the noise and the signal. To begin with, the simplest concepts that have been employed for this attempted separation process have emphasized the fact that, whereas the signal information is constrained to lie within a known pre-established bandwidth, the noise may lie well outside of this bandwidth. This in itself, however, does not guarantee that there will not be a certain amount of noise lying within the bandwidth of the signal. But a first attempt for eliminating noise would be to filter all information about a certain frequency, which would then leave the signal plus a smaller amount of noise. To refine this process still further, additional information must be had concerning the signal and the noise. The additional information that noise appears to be uniformly distributed over a very large spectrum, whereas a signal is located almost along a spike in the frequency distribution diagram has been utilized in the designing of a radar PPI (Plot Position Indicator) display. By using the PPI scope to perform integration, it can be seen that the random noise portion of the signal will build up at a very slow rate and reach a much lower saturation level than will the repeated signal spike. Even though the signal-to-noise ratio (that is, ratio of power contained in each) at any moment is less than 1, as the integration process proceeds, a bright spot will appear on the scope indicating that the noise has been partially separated from the signal, building up the signal well above this signal-

to-noise ratio of unity. The foregoing material indicates that, before any improvement in signal-to-noise ratio can be obtained, more information must be had about the nature of the noise involved. In fact, if the noise spectrum can be defined, techniques exist, complicated as they may be, for designing what has been termed the optimum filter. This method can be applied and used as a guide in determining the type of filter operation required. The solution does not necessarily represent the type of filter that would be mechanized since it can conceivably be extended to cover all modes and methods of nonlinear filtering that might be desired. A second possibility is to define more closely the spectrum of signal and to adjust reading of this spectrum very closely by such techniques as Microlock.† If, for example, a number of discrete readings of a measurement are obtained as a function of time, and it is known that the quantity being measured is changing as a linear function of time, then a problem of fitting the best straight line to a large number of sample points is the manner in which the signal is partially extracted from the noise. To date, there appears to be no single sequence of arithmetic operations to optimize the partial extraction of signal from a combination of signal and noise.

It should now be apparent that in the design context of the computing portion of an informational system (for example) it is necessary to establish the capabilities of the input sensors at least with respect to:

(a) Signal/noise at the sensory output,
(b) Types of exterior noise passed by the sensor,
(c) Types of noise introduced by the sensor,

inasmuch as these will directly affect the computing requirements if filtering or smoothing should be necessary.

Concerning capabilities of sensors themselves, full information on their transfer characteristics is frequently not known even to the manufacturer, and even less is generally known regarding their interior signal characteristics. Surprising as this may at first thought appear, it has become unfortunately evident in recent studies on navigational instrumentations, for example. Only the crudest empirical approximations are available, for instance, to indicate how the error-noise statistics of navigational Doppler radars reflect varying terrain characteristics. Having set out to optimize central computer system information flows and interconnections with multiple navigational sensory subsystems (for example, inertial platform

† Microlock is a Hallamore proprietary technique for obtaining weak signals in the presence of noise by phase-locking signal and receiver.

+ Doppler radar + TACAN equipment + stellar tracker + central digital computer), one finds that first one must undertake the deep analyses of the subsystems one would expect to find in the technical literature or accompanying commercially available subsystems. This means, of course, that manufacturing has outstripped engineering and that, consequently, applications analysis must suffer.

6. Feasibility Analysis

In distinction to requirements analysis and capabilities analysis which operate over a spectrum of detail levels from specific detailed questions to "broad-brush" considerations, feasibility analysis, as customarily employed, is directed to specific questions. This is, undoubtedly, due to the fact that feasibility analysis generally follows a series of tentative decisions and might well be likened to testing of the hypothesis in the scientific method. Feasibility analysis may also be thought of as combining basic requirements and capabilities techniques for purposes of re-evaluation (in some cases this re-evaluation turns out to be the now famous "agonizing reappraisal"). Thus, there is a strong resemblance conceptually to the more complex problems of classical mechanics in which the basic equation $F = ma$ is utilized by sequentially assuming certain variables fixed and operating now upon one side and now upon the other.

One of the most obvious types of questions arising in feasibility analysis is that of possible conflict between a systems requirement for certain operating speeds or a certain degree of reliability and a design objective calling for limitation to a class of phenomenological devices. Thus, for example, only a small number of years ago it was common, in the case of airborne computing systems, to find a design objective specification calling for maximum use of semiconductors and other solid state devices which at that time were relatively unknown with respect to their effect on system reliability. Basic data were needed as well as parallel theoretical study to ascertain whether substitution of semiconductor circuitry for vacuum tube circuitry could sustain the reliability performance then prevalent. Subsequently, as we know, semiconductors rendered vacuum tubes virtually extinct in this area although the problem must now be once again reappraised in view of the potential radiation environments for equipment to be utilized in nuclear powered vehicles, and in vehicles penetrating space radiation areas such as the Van Allen Belt.

7. Prediction and Extrapolation

A simple example demonstrating a question of feasibility is whether or not one may, in a given problem, utilize prediction and extrapolation to enable a digital computer to overcome certain systems difficulties introduced by itself. We depart from standard practice in this book at this point by introducing technical details.

Sample data theory shows that a digital control system may be represented by an ideal computer in series with a sampling element, a delay element, and a holding circuit. This is illustrated in Fig. 2.1.

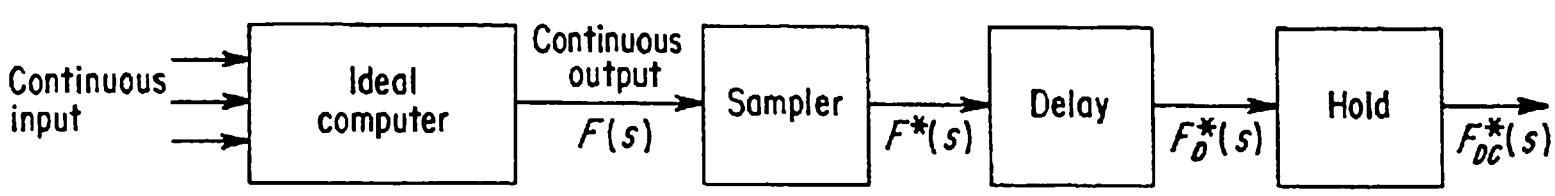


Fig. 2.1. A digital control system.

The first element in the model, the ideal computer, is defined as a computing element with the property that it accepts continuous inputs and produces continuous answers as required by the particular problem the system is to solve. It introduces no delays and has no sampling processes associated with it. Assume it is an analog computer with an infinite pass band.

Because of the discrete nature of the operation of a digital computer, however, a sampling process is present in the forward loop. It can be shown that if the Laplace transform of one of the outputs of the ideal computer is $F(s)$, then the Laplace transform of the sampled output $F^*(s)$ will be

$$F^*(s) = \frac{1}{\tau} \sum_{n=-\infty}^{\infty} F(s + jn\omega); \qquad j^2 = -1, \tag{8}$$

where $\omega = 2\pi f = 2\pi/\tau$, and τ is the sampling period.

In series with the sampling element, there is an element introducing a delay equal to the sampling period τ. This element is present because from the time the digital computer receives its inputs to the time a result is obtained, a certain amount of computation must take place and this consumes time.

The Laplace transform of the sampled and delayed output signal can be shown to be:

$$F^*(s)_D = e^{-sT}F^*(s)$$

$$= e^{-sT} \sum_{n=-\infty}^{\infty} F(s + jn\omega). \tag{9}$$

Finally, there is the "boxcar" element in the model of the system. This element is used to hold the output of the digital computer between samples. The Laplace transform of the sampled, delayed, and clamped output of the system is:

$$F^*(s)_{DC} = e^{-sT}(1 - e^{-sT}) \sum_{n=-\infty}^{\infty} F(s + jn\omega). \tag{10}$$

What would be most desirable is the output of the ideal computer, $F(s)$, since it would not contain the distortion introduced by the sampling, clamping, and holding processes inherent in the digital computer. Fig. 2.2

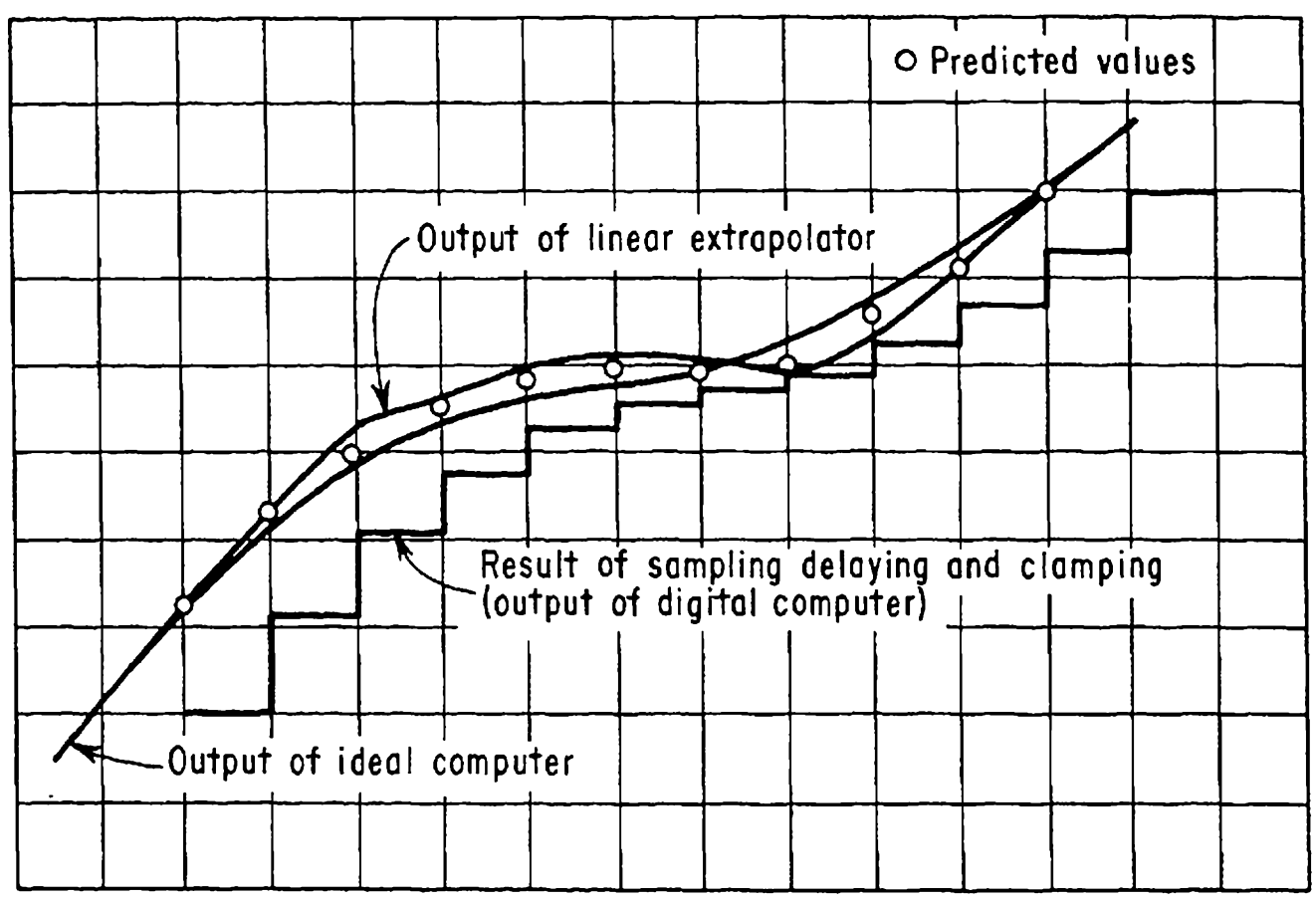


Fig. 2.2. Effect of an extrapolator-predictor on the output of a digital computer.

is a graph showing a typical output of an ideal computer and also the signal after it has been sampled, delayed, and clamped. There are two techniques which can sometimes be used to counteract the effects mentioned above; these are prediction and extrapolation. Prediction is gen-

erally used in order to counter the effect of the digital delay while extrapolation is used mainly to reduce the effects of the sampling and clamping.

The prediction can be designed in different forms, the simplest of which is linear. In this type of prediction only the present and the immediately previous output of the system are examined, and on the basis of these values the next value is determined. Higher order prediction can also be used; i.e., the last three points can be examined in order to predict the next value.

The extrapolator may be described as a device which produces additional outputs of the digital computer between sampling periods. It is possible to introduce many extrapolation points between samples. The extrapolation process can be broken into two parts: the first part consists of examining a certain number of previous values of the output of the computer in order to determine best what the output of the computer should be between samples; the second part consists of dividing, in a linear fashion, the time interval between iterations into several parts.

The first part of the extrapolation process may be grouped with the prediction mentioned above, so that the prediction not only will counter the effect of the digital delay, but will also determine the best form for the initial conditions to be placed in the extrapolator.

An illustration of the effects of prediction and extrapolation is considered next.

If the present and the previous outputs of the digital computer are f_i and f_{i-1}, respectively, then the predicted value f_{i+1}, on the basis of a straight line, is:

$$\begin{aligned} f_{i+1} &= f_i + f_i - f_{i-1}, \\ &= 2f_i - f_{i-1}. \end{aligned} \tag{11}$$

The predicted output of the ideal computer at the next step will be:

$$\begin{aligned} f_{i+2} &= f_i + 2(f_i - f_{i-1}), \\ &= 3f_i - 2f_{i-1}. \end{aligned} \tag{12}$$

The numbers f_1 and f_2 constitute the predicted output of the ideal computer for the present and for the next sampling point. Note that these numbers are computed quite simply as linear functions of f_0 and f_{-1}, which are respectively the present and previous outputs of the digital computer. The values f_1 and f_2 will now be used to set the linear extrapolator. Assume that the value in the linear extrapolator at the present time is f_0; the predicted output of the ideal computer is f_1. Rather than replace f_0 by f_1, which would produce a discontinuity in the output of the

extrapolator, it would be more desirable to provide a smooth transition to f_2. The equation that encompasses this is:

$$f_1 = f_{K-1} + \frac{f_2 - f_0}{n}$$
$$K = 1, \cdots; n, \tag{13}$$

where the linear extrapolator is to provide n values between each iteration. Using this equation, it is only necessary to compute f_2.

A block diagram of the digital system would now have the predictor-extrapolator incorporated as in Fig. 2.3.

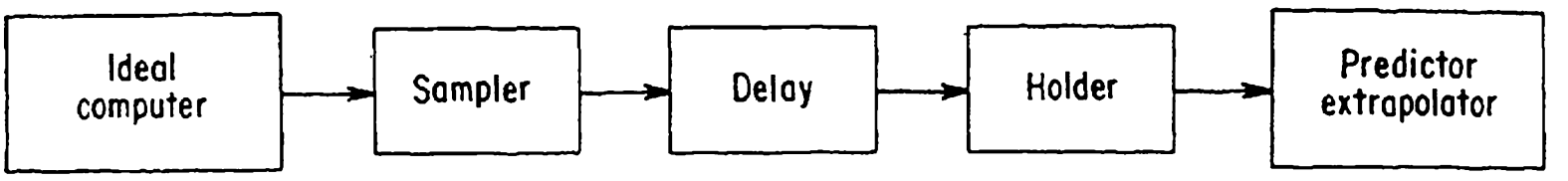


Fig. 2.3. Predictor-extrapolator incorporation.

Figure 2.2 shows the improvement obtained by the use of the prediction and extrapolation. The predictor-extrapolator has produced an output which is smoother and more closely approximates the output of the ideal computer than does the output of the digital system without such a mechanism.

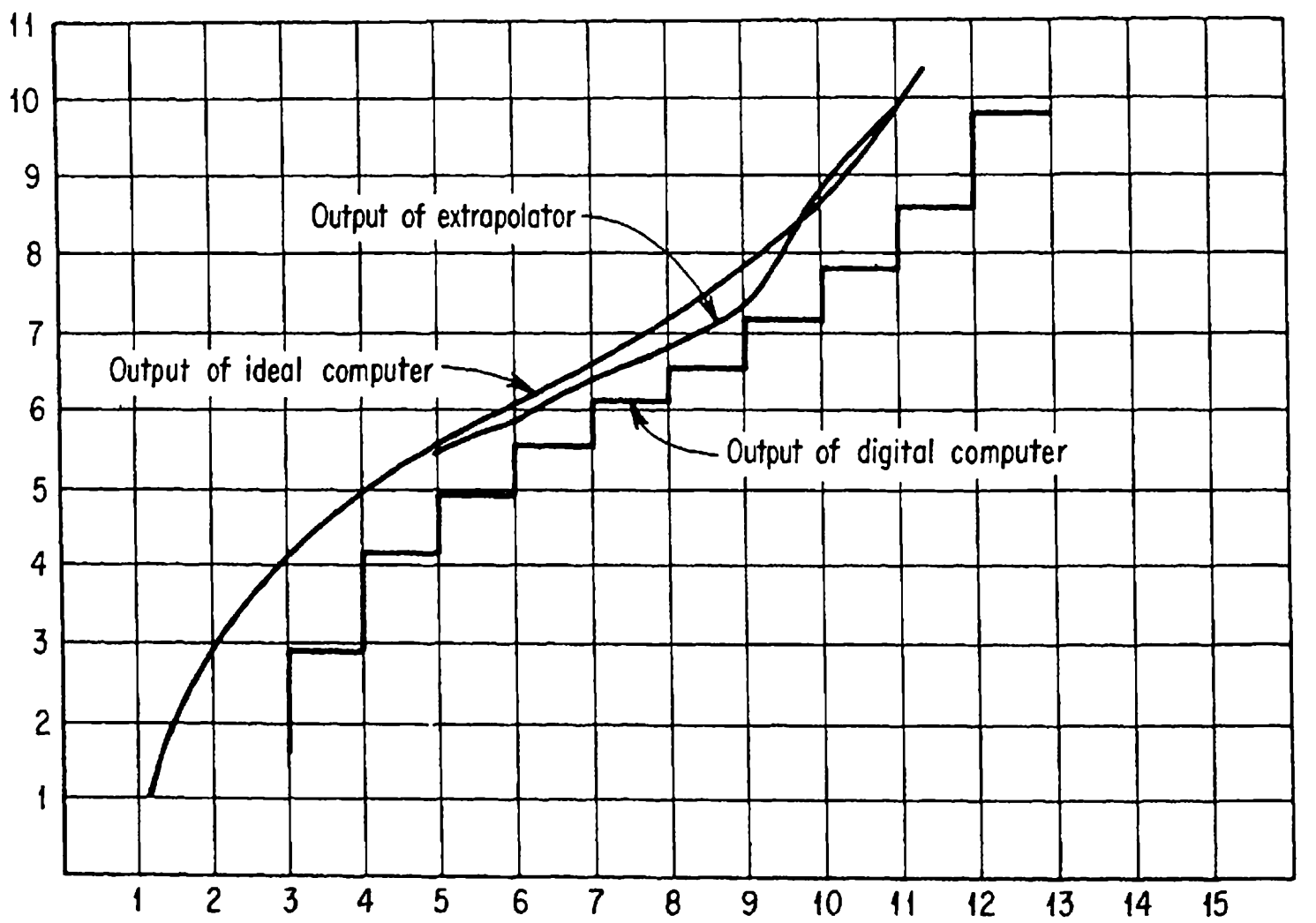


Fig. 2.4. Effect of circular prediction and linear extrapolation.

A higher order type prediction may be used which fits a circular arc through the previous three points available to predict the next value.

The equation of a circle in the (f_k, k) plane at the kth iteration is:

$$(f_k - a)^2 + (k - b)^2 = c^2. \tag{14}$$

Taking the first forward difference:

$$2(f_k - a)\Delta f_k + (\Delta f_k)^2 + 2(K - b) + 1 = 0. \tag{15}$$

Taking the second forward difference:

$$2(f_k - a)\Delta^2 f_k + 2(\Delta f_k)^2 + 2\Delta f_k \Delta^2 f_k + 2\Delta f_k \Delta^2 f_k + (\Delta^2 f_k)^2 + 2 = 0. \tag{16}$$

Solving for a:

$$a = f + \frac{2(\Delta f_k + \Delta^2 f_k)^2 - (\Delta^2 f_k)^2 + 2}{2\Delta^2 f_k}. \tag{17}$$

Obtaining b at $k = -2$:

$$b = \frac{2\Delta f_{-2}(f_{-2} - a + \Delta f_{-2}) - 3}{2}. \tag{18}$$

Evaluating at $k = 0$:

$$(f_0 - a)^2 + b^2 = c^2. \tag{19}$$

Substituting and solving for f_2:

$$\begin{aligned} f_2 &= s + (f_0 - a)^2 + b^2 - (2 - b)^2 \\ &= a + (f_0 - a)^2 + fb - 4, \end{aligned} \tag{20}$$

where a and b are as evaluated at $k = -2$.

Examination of the equations above shows that the computation is much more involved than is the case for linear prediction; however, a better fit is obtained. Fig. 2.4 shows the results obtained using circular prediction combined with linear extrapolation.

8. Summary

In this section we classified the most common types of system analyses into requirements analysis, capability analysis, and feasibility analysis types; the concepts of real time, noise, and prediction-extrapolation were introduced.

Section 2.3

Design Flow

"All wonder is the effect of novelty upon ignorance."
—JOHNSON

1. Mark I and Evolutionary Designs

Although there occur upon occasion so-called one-of-a-kind systems, the much more common case is that of a series of systems (to carry out a mission) evolving either through routine improvements in subsystems and their organization into the system, or through necessity to meet expanded requirements resulting from elaboration of the basic mission. It is convenient in the more common case to think of an evolving spectrum of systems from the first, Mark I, to those approaching but never quite attaining the ultimate (the ultimate is that system best satisfying selection criteria with only cosmological restrictions on impinging states-of-the arts). Thus, there are two basic design flows to be considered:

(a) Mark I design,
(b) Evolutionary design.

Clearly, Mark I design provides, so to speak, the initial condition for evolutionary design.

The system designer, in working from the reference of the Mark I system having limited preplanned evolutionary capability, is subject to severe approach constraints which are of an extra-systemic nature. He must, for instance, undertake an exhaustive examination and classification of the Mark I system's characteristics within all included areas of operation. He must discern similarities potentially useful in perpetuating the major systems characteristics through the steps necessary for an evolved system. In effect, the major portion of his activities relate to the preservation of the essential characteristics of the Mark I system, despite enlargement, subject, of course, to such minor modifications as may be compatible with the imposed demands upon the system which stimulated development beyond the Mark I phase. In this situation the designer is virtually committed to the preservation of system performance capabilities characteristic of the Mark I system. In order to discharge this responsibility he must locate and identify, within the extant Mark I system, areas of operational similarity and incorporate these within the successor system. The system thus becomes larger without becoming better.

The desirability for careful consideration of evolutionary design planning in Mark I design is thus clear.

2. Specification—Design—Analysis

The methodology associated with system specification usually becomes intimately mingled with the functions of system design and systems analysis, although regarded in the abstract as being prior to and independent from the latter functions. The intricacies and trade-offs, arising from hardware constraints and other real-world difficulties associated with systems work, distort the theoretical information flow coarsely depicted in Fig. 2.5 into that represented by Fig. 2.6.

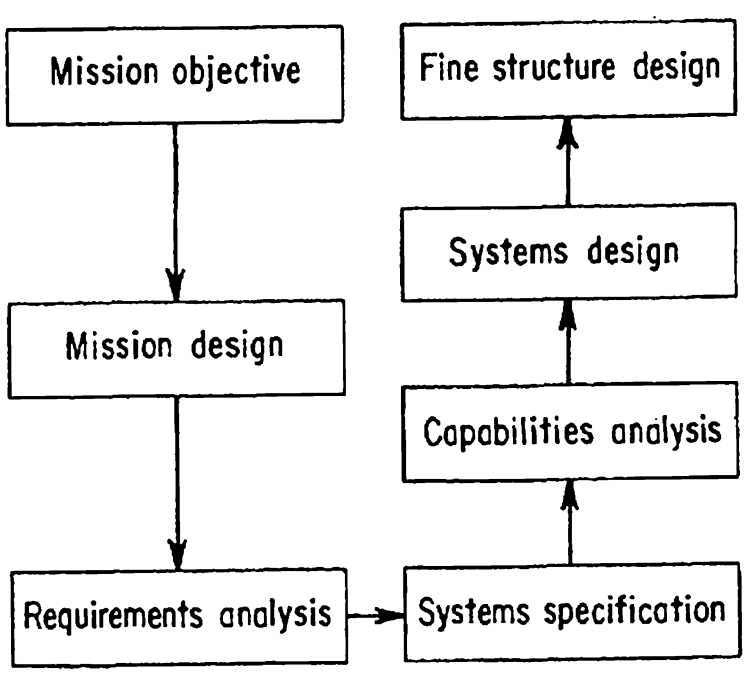


Fig. 2.5. Theoretical information flow for Mark I design.

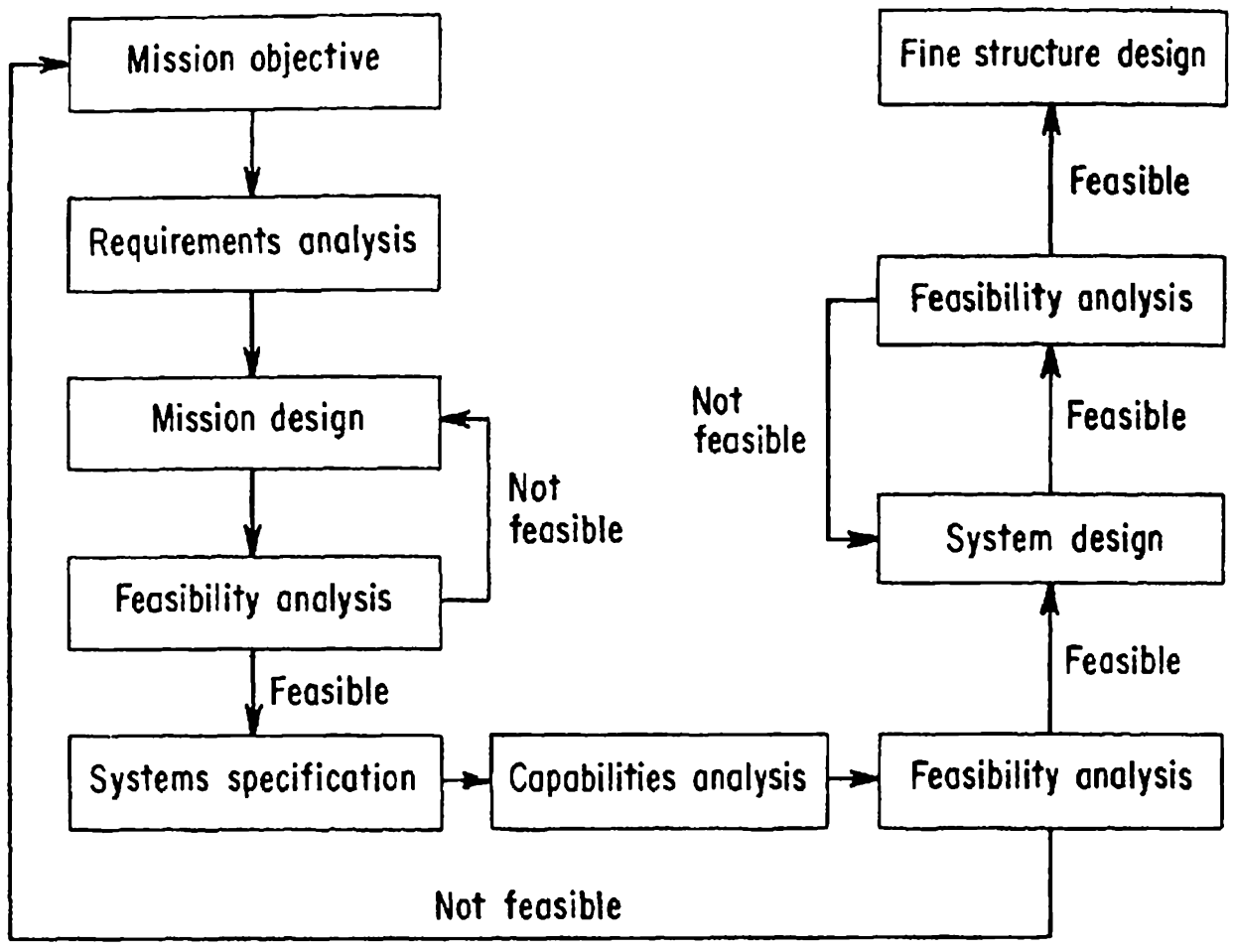


Fig. 2.6. Real-world information flow for Mark I design.

3. System Matrices

A method of organizing system information for design flow and control is desperately needed until a general systems theory is available in full form. Such a method has been developed utilizing the concepts of economic analysis. The charts and discussion in Appendix 5 give the first-pass matrices evolved for defining the central-control-computer oriented portion of, first, an advanced conceptual high-performance aircraft; secondly, a modern submarine command system; and thirdly, a manned space vehicle. The method is, however, of general applicability, and suitable coding of matrix entries followed by algebraic manipulation may be made to yield not only interdependence of subsystems and components, but the degree of such interdependence and measure of effects of marginal alterations. Such matrices may also be used in the automation of control for weapons system management. This subject is discussed in Chapter 4.

4. Growth Potential

As indicated previously, a methodology for selecting evolving systems during Mark I design is needed.

This statement will seem more cogent if examples of systems designed without sufficient attention given to evolutionary or growth potential implications are considered. There is no dearth of such examples within recent technological history. Note, for instance, the following:

(a) *Integration of cockpit instruments displays within the configuration of a modern high-performance fighter aircraft.* Early design efforts have accomplished a degree of integration of the display/control interface equipments which reasonably satisfy most operational requirements upon the aircraft system. However, additional demands for further integration of the display/control complex will arise in response to increased tactical versatility expectations. Satisfactions of these demands will be reflected in successor systems design. The design team will discover that, to accommodate the increased demands upon the man-machine system which arise by virtue of altered modes necessitated by redefinition of the mission, it will be necessary to modify the central computer in a radical fashion to supply computation capabilities that meet these new demands. Thus, redefinition of the mission, calling forth envisioned extensions of tactical capabilities, impressed new demands upon the man-machine system, which are not met adequately without substantial modifications of interim display arrangements. These in turn require extensive alterations in the structure, size, form factor and other attributes of a central control computer that was originally conceived as possessing sufficient flexibility to cope with such growth contingencies for some time.

(b) *Attack aircraft control.* Increased diversity and complexity in the types of armaments associated with modern attack-type aircraft have resulted in substantial design changes with accompanying modifications in form factors, speed, memory capabilities, and power requirements related to a central control computer. The Mark I attack plane of an arbitrary class may have supported a modest arsenal of fire capabilities, i.e., a single type of missile having one type of warhead, one type of guidance system, and one launch mode with a mechanical release. As diversity has increased the number of useful, available weapons of this type of mission, new requirements have been imposed upon the central control computer. There may be several types of mid-course missile guidance computations required which were not necessary with an earlier system. Several attack modes may be available, each of which is of great utility in meeting its anticipated class of threats. The complexity of the armament may be so great that checkout procedures demanding special

sensory-computer-display linkages not previously required are prescribed. New demands are imposed upon the computation means, which may not be adequately satisfied by modularity in design or by the addition of new special purpose computational blocks. Spatial limitations may be such that a complete redesign of the computer package is indicated. In the worst case, state-of-the-art is exceeded so that it is necessary to compromise the potential effectiveness of the system because of contemporary computer size, weight, packaging, or form factors, or some combination of these.

(c) *Submarine controls.* Let us consider an example from another domain—that of the submarine. In consequence of vastly extended endurance capabilities following provision of nuclear power sources, the submarine is receiving attention from many quarters. On the basis of extended capabilities it has come to be considered as having great potential as a strategic instrument in addition to its conventional functions as a tactical device. Thus, alterations in the performance characteristics of the submarine which have ensued in consequence of the incorporation of nuclear power resources have resulted in re-examination of the operational characteristics of this type of vessel. The enhanced responsiveness of the system which emerged in consequence of the reduction of latency factors associated with an earlier generation of submarines has led to advanced concepts of control. These will require for their satisfactory expression new families of sensors, specialized displays, and computation capabilities far in excess of those required in earlier submarines. The advances in the power domain have thus led to a class of requirements for new display, control, linkage, and computer techniques which will more or less "bootstrap" the submarines into a thoroughly new design epoch. With the increased responsiveness and endurance capabilities have come increased demands upon existing internal and external communications. Therefore, another dimension must be examined if the submarine is to emerge as a consistent, thoroughly cohesive representation of the existing technologies in many realms.

Growth potential integration affords exceptional promise as a methodology that answers the need for selecting evolving systems during Mark I design. It has been much praised, seldom elaborated, and little practiced. The evolutionary design approach may be indicated in gross fashion as follows:

(1) *Operational analysis.* Examine the operational process (usually rather loosely defined) as a whole and in relation to its environment or

other operational processes to clarify and make precise the objectives to be accomplished. These objectives will become what may be called primary requirements (or constraints) for the system.

(2) *Constraint analysis.* Ascertain the additional requirements, called imposed requirements (or constraints) on the system. These characteristics may appear in the form of commodity constraints (i.e., cost limitations, deadlines, etc.).

(3) *Feasibility analysis.* Determine the class of systems attainable in current states-of-the-arts which satisfy the primary and imposed requirements. These systems are said to be feasible.

(4) *Extension analysis.* Determine, insofar as possible, the necessary characteristics of an ultimate system meeting the primary requirements and those imposed requirements regarded as irrevocable. Ultimate system is defined as one best satisfying selection criteria with only cosmological restrictions on the states-of-the-arts.

(5) *Growth potential analysis.* Each feasible system permits evolution toward an ultimate system in a fashion that may be roughly classified as:

(a) Requirement of technical breakthrough followed by total redesign;
(b) Discrete steps of technical advance, each followed by major redesign;
(c) "Smooth" modification taking advantage of advances in states-of-the-arts without major redesign and resulting downtime.

If (*a*), (*b*), and (*c*) are indicated on the same time scale, they may be compared very roughly as in Fig. 2.7.

The system cost associated with a mode L of evolution may, apparently, be roughly approximated by

$$K \int_0^T L(t)\, dt.$$

From Fig. 2.7

$$\int_0^T a(t)\, dt > \int_0^T b(t) > \int_0^T c(t)\, dt. \tag{21}$$

Feasible systems having an evolvability characteristic of the c-type are said to possess growth potential.

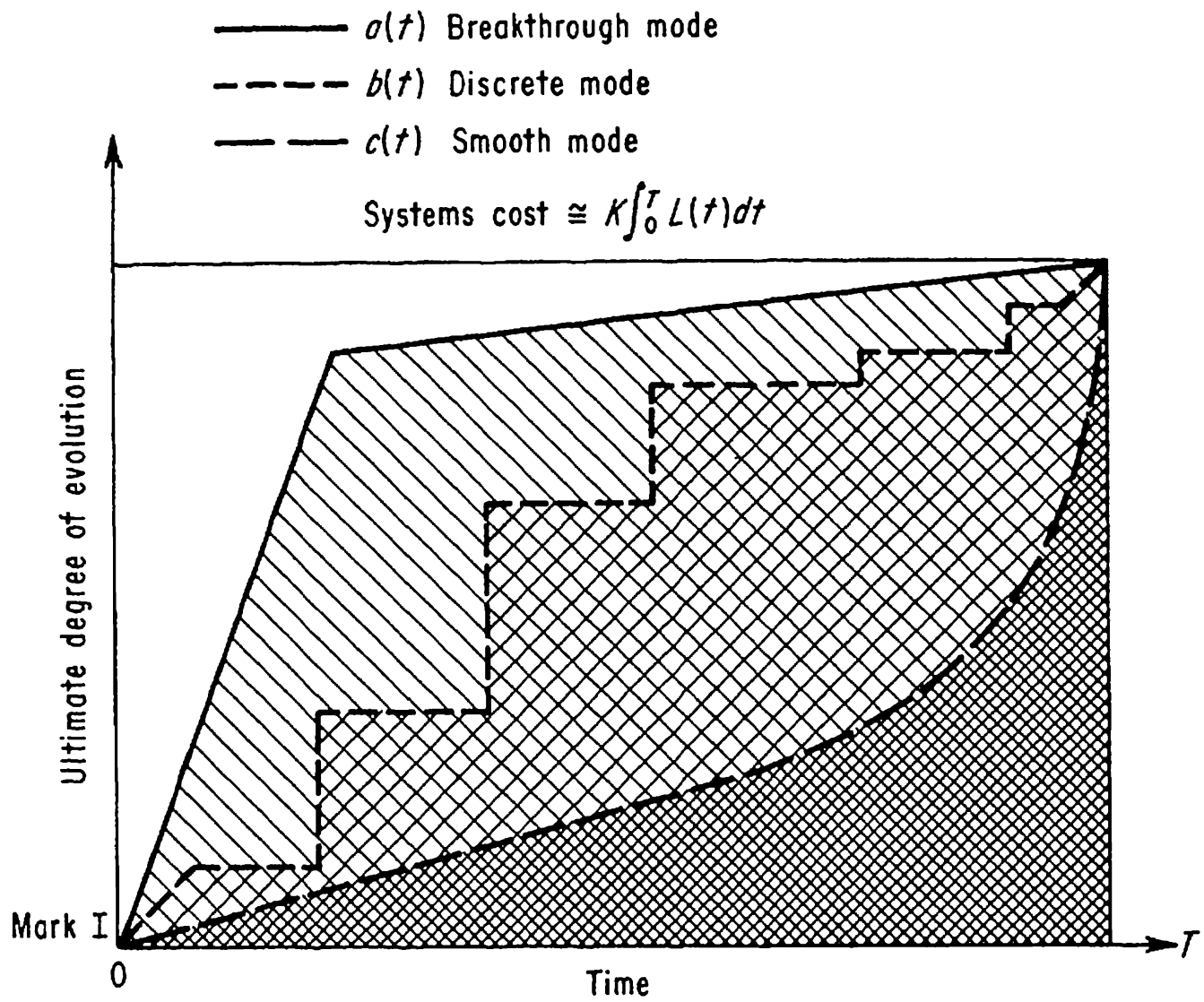


Fig. 2.7. Basic modes of systems evolution.

(6) *Evaluation analysis.* Ascertain a suitable set of criteria for relative evaluation of feasible systems.

(7) *Initial optimization.* Select from those feasible systems possessing growth potential one which is optimal with respect to the evaluation criteria. This becomes the initial or Mark I system.

(8) *Evolution analysis.* Indicate, to the extent permitted by planning under uncertainty, an expected evolutionary path from Mark I toward the ultimate as new technology and technique become available.

The design of contemporary systems without the application of growth potential integration is an approach known colloquially as "kluging." Such inadequate planning may prove very costly and also cause the loss of capability to attain the objectives of the implemented process whenever redesign is necessary. Mark I systems designed without extension analysis or with characteristics *a* or *b* of Fig. 2.7 represent decreasing degrees of kluging.

The above-mentioned evolutionary approach is recapitulated schematically in Fig. 2.8.

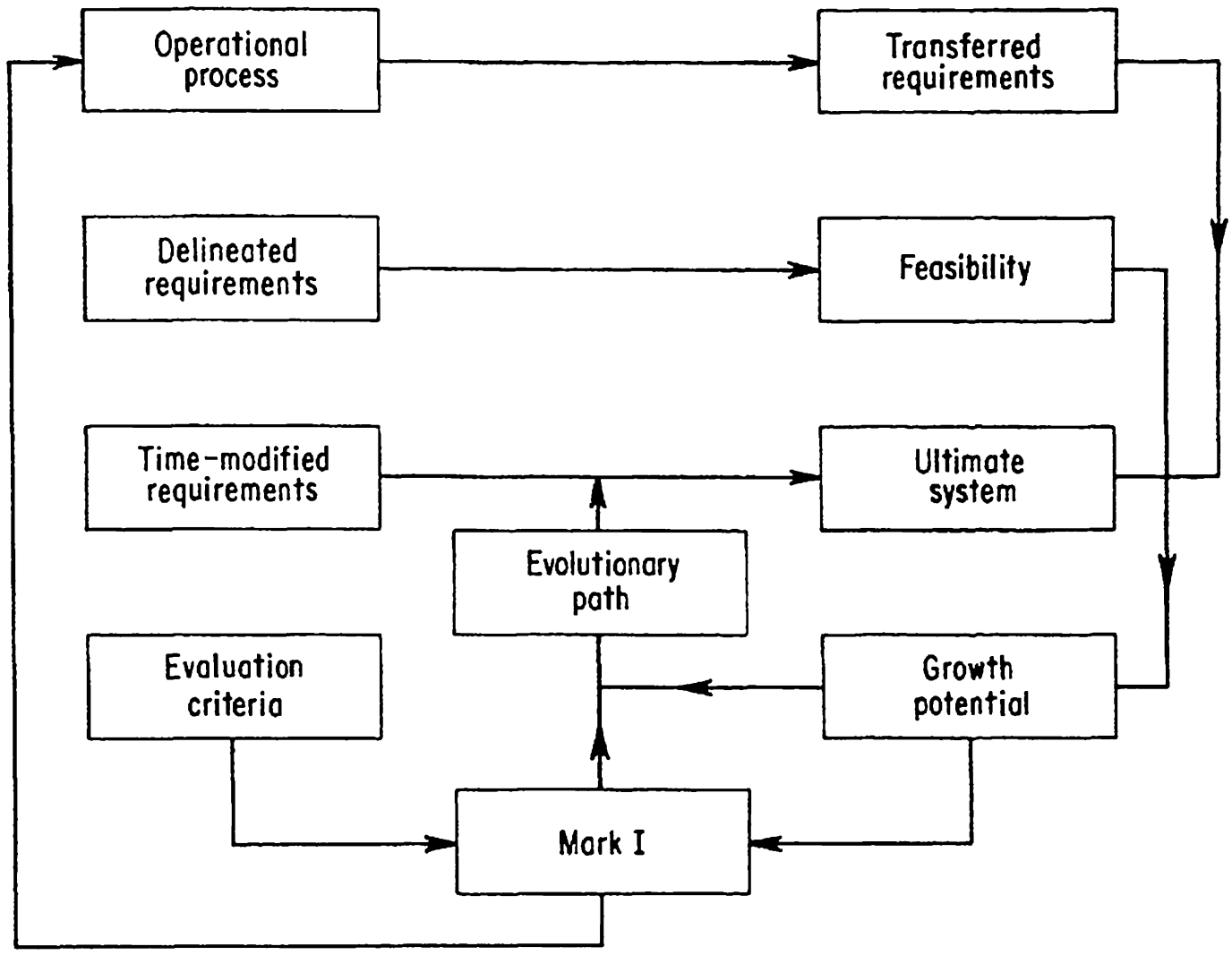


Fig. 2.8. Growth potential integration.

5. Summary

In this section we have considered evolution of systems, the results of "kluging" (design without adequate evolutionary planning), and methods to avoid "kluging." We have also indicated how real-world factors distort theoretical design flow.

3

Special Problem Areas

"And thus the whirligig of time brings in his revenges."
—SHAKESPEARE

Section 3.1

Description

"Reason is a very light rider,
and easily shook off."
—SWIFT

1. The Descriptive Model

In Chapter 1 we considered an aircraft system as an example and indicated that its description was replaced by gross approximations because of complexity and unknown factors. In Chapter 2 we pointed out the lack of available information on noise characteristics of some sensory equipments. Let us now consider as a systems problem a central control computer in a specified aircraft system. One of the missions to be interlaced in the computer (either by time-sharing or by provision of parallel processing capabilities) is navigation. The navigation concept employed might rely on an approach as simple as dead reckoning or on one as complex as optimal-mixing-utilization of information from an inertial platform, a Doppler radar, a stellar tracker, and a radio-fix subsystem such as TACAN or DECCA. There will generally be a requirement associated with autocontrol which may be as simple as commands for mode changes in an autopilot or as complex as the actual stabilization and director loop

computations. There may be a requirement for ballistic computations in an armament mode and a requirement for programming flight in various fuel-management/cruise-control modes such as maximum endurance, maximum range, minimum time to target, or others. As we move into the era of integrated cockpits, in which traditional indicators and controls are replaced by integrated displays and control consoles, the computer will be required to synthesize the display information, drive the display generating equipment, and interpret to actuators the pilot's control commands in a quickened (that is, anticipatory) fashion. Probably, there will be additional requirements for computing functions of air data measurements and such housekeeping functions as checks for flap position and wheel position. Other requirements on this general level of detail that may be present are automatic operation of IFF (Identification, Friend or Foe) and communications equipment and lock-out of certain maneuvers the pilot might attempt which would result in dynamic instability or other potentially fatal conditions (such as low altitude pitch-down beneath the pilot's observance threshold).

If the computer is to be truly designed into the supersystem (as it should be to achieve the reliability and economy of power, weight, and space desirable for airborne equipment), its mission should be completely specified prior to design. Such a specification will consist of a total descriptive model of the system of equations the computer is to simulate, and it must include:

(a) Every equation associated with any required mode;
(b) Ranges, accuracies, and maximum possible rates of change of all input variables;
(c) Times of availability and forms (voltages, synchropositions, etc.) of all input variables;
(d) Types and, at worst, statistical distributions of all errors anticipated in input variables;
(e) Accuracies, ranges, anticipated maximum rates of change and availability time required of all outputs to meet supersystem specifications;
(f) Other necessary equations and/or decision rules supplied from basic theory or empirical data fitting that relate variables in the over-all set of equations so that it becomes theoretically possible to solve the set uniquely for all output variables as functions of input variables and constants appearing in equations;

(g) Specifications of all constants appearing;
(h) All additional computational constraints such as, for example, preference between iterative and algorithmic methods of solving certain subsets of the set of equations.

Now the obtainance and preparation of all this material (a–h) is a formidable task indeed and is certainly not made easier by the fact that generally sensory and actuation equipment as well as the airframe itself are in, at best, developmental phases when work must be started on computer design in order to meet delivery schedules. Thus, much of the descriptive model will be based on educated guesses rather than factual information. Therefore, it will be subject to, hopefully, small (but probably major) perturbations as these development phases progress. The necessity for drastically revising the model with respect to an entire functional mode is not uncommon because of such factors as:

(i) Some sensory component, such as a radar under development, may prove incapable of performing as hoped;
(j) Some supersystem component may be behind development schedule and require substitution in order to meet supersystem schedules;
(k) Inferior analysis or erroneous information may have caused basic inadequacy of some equations;
(l) Realistic or whimsical change in policy or objectives of the customer for the supersystem may revise the supersystem mission;
(m) Computer implementation of the model and imposed constraints taken together may be beyond the state-of-the-art.

It is to be emphasized that we have not considered here any of the difficulties associated with design and fabrication of the computer itself, but merely the problem of describing what the computer is to do. Division of responsibilities for such models will be discussed in Chapter 4.

2. Other Descriptive Problems

In the preceding subsection we discussed via an example some of the descriptive problems arising in system requirements. There are also, of course, serious descriptive problems associated with the capabilities of subsystems and components. The methods available at the present for describing subsystem and component behavior are almost entirely con-

fined to linear system behavior.† Unfortunately, no known physically existing system behaves linearly. Although some systems approximate linear behavior over a severely limited operating range, inf act, (as is highly desirable in some instances, namely, adaptive control‡), it is essential to admit nonlinear behavior. Inadequacy of the available descriptive tools may not, however, be the most serious deficiency. Even as wider ranges of phenomenological devices are being conceived, we find inadequacies in the basic physical theories upon which their operational explanation is hypothesized. To see this it is only necessary to recall in the near past the refutation of both Maxwellian and Lifshitz-Landau electromagnetic theory by the advent of cryogenic phenomena, the choice in physics to give up the parity principle in order to continue maintenance of the older tenets of the conservation religion, and the failure of the standard gravitational theories and astronomic data in conjunction to predict the perturbations later observed in certain artificial satellitic orbits. We even observe upon occasion "tests" of theories in which measurements are made with instruments whose operational explanation is hypothesized upon subtheories of the theory being tested.

We have discussed in this section some of the types of descriptive problems in systems context. Section 3.3 includes a discussion of descriptive problem orders of magnitude more forbidding than these, namely, the description of human components in man/machine systems.

Appendix 6 gives as an example a precise logical description of a digital integrator. Despite the simplicity of this familiar computing element, its analysis has generally been based on primarily heuristic considerations.

3. Simulation

The conclusion of this subsection on description appears to be the proper place to discuss briefly simulation as the term is used in systems analysis and system engineering. As we observe from Section 1.1, a system simulates its descriptive model. The term *simulation* customarily refers to actual operation of some second system believed capable of simulating the descriptive model of another system. Thus, for example, an analog

† A system is linear if its behavior can be described by linear differentio-difference equations with time as the independent variable.

‡ Mode or state changes with changing environment to maintain optimal performance.

computer is frequently employed to study the dynamics of an airframe. If the actual descriptive model of the airframe were known (which it never is) and if a precise and accurate, and inexpensive analog (say electrical) satisfying this model could be achieved (which it never is), it would clearly be more advantageous to operate the analog than to construct and fly the airframe to obtain response characteristics.

Legitimate simulation generally arises from one of two conditions:

(a) The descriptive model involved is intractable to analysis so that response characteristics for general inputs cannot be predicted;

(b) It is desired to utilize the simulating system as a training device for human systems to be associated with it in an eventual supersystem, or as a testing device for other subsystems either prior to the physical existence of the simulated system or for economic reasons.

While condition (b) constitutes inherently a good case for simulation, condition (a) has been greatly overworked historically. Frequently, simulation has been employed as a substitute rather than as a follow-up to analytical studies. This has an effect similar to that of obtaining special solutions to a differential equation rather than the general solution. Thus, if specifications are altered prior to the final design freeze (as they generally are), the work supposedly accomplished by simulation may well have been merely a waste of time and, frequently, of much larger amounts of money than would have been required for serious analytical effort. Simulation in this sense is a legitimate *last-resort tool* which too often is regarded as the entire tool kit. Unfortunately, the masses of data produced by simulation studies have had a certain attraction for the design customer regardless of actual value, and this has undoubtedly encouraged the trend.

To illustrate our point more concretely, we may point out the glamour accorded Monte Carlo methods. These methods are basically nothing more than techniques for estimating (subject to certain confidence limits) certain output distributions (statistical distributions) by appropriately sampling (according to sampling theory) given or hypothesized input and state distributions and simulating system behavior. This constitutes a legitimate approach to truly intractable stochastic description models, but is evidently nonsense when the output distributions are deducible with 100 per cent confidence.

Finally, a moot point is brought up in the concept of using simulation on a digital computer to check supposedly successful analysis. This would certainly be of value if programmers and coders were available at a level

of perfection beyond analytical personnel. Actually, it is of doubtful value when compared economically with checkout against a second independent analytical effort.

4. Summary

For optimal design a complete descriptive model of the system is essential. This is very difficult to obtain and is frequently subject to change.

There are other problems of description besides the system model itself. Many such difficulties reflect poor basic theory.

Simulation is a much overused tool.

Section 3.2

Complexity

"The number is certainly the cause.
The apparent disorder augments the grandeur."
—EDMUND BURKE

1. Measures of Complexity

The most obvious measure of complexity of a system is

$$\log_2 \aleph \tag{22}$$

where $\aleph$ is the cardinal number (i.e., number of elements) of the system phase space. Special account is required for countably infinite phase spaces, and the definition inherently assumes meaningfulness for other infinite phase spaces which essentially implies assumption of the generalized continuum hypothesis.† This measure is quite appealing for such classical systems as relay networks or flip-flop registers. Where there are significant trade-offs in system description between inputs and states or

† *Cf.* Hans Hahn, *The World of Mathematics*, (New York: Simon and Schuster, 1956), III, 1593–1611.

between outputs and states as in the factory sample of Chapter 1, it is clearly inadequate. A more satisfactory measure would be

$$\sup_{d \in D} \log_2 \aleph_d, \tag{23}$$

where D is the class of all possible descriptions of the system in question suitably restricted as to the levels of generality of interest. This, however, is generally quite intractable to analysis. In any case, the measure should reflect intuitively the number of realizeable simple components involved (together with their states) which is established by the level of generality of interest.

Currently, reference is made to *complex systems*. This concept covers a spectrum of complexity, and an idea of the ends of the current interpretation of such a spectrum may be gained from perusal of Appendixes 7 and 8. Appendix 7 discusses a shipboard fire control system representative of the lower end of the spectrum. Appendix 8 considers certain aspects of a hypothetical surveillance system representative of the upper end of the spectrum.

2. Difficulties Arising From Complexity

Aside from descriptive difficulties and, currently, fabrication difficulties arising from complexity, the real haunt associated with complexity is compounded of reliability and operability. Clearly, unless components are perfect (and they are not), larger aggregates in more or less random assembly have lower probability of continued simultaneous operation within specifications than do smaller aggregates.

3. Reliability Computations

It is theoretically possible to arrive at estimates, including confidence limits, for the time distribution of probabilities of transient and total failures of any system provided sufficient information is given regarding the restraints on, and fine structure of, the system. Fine structure refers here to all components and their interconnections, interaction with environment, and the totality of interrelationships established by these. Unfortunately, the information available for a given system is invariably insufficient for these purposes. Furthermore, insufficient study has been

devoted thus far to the elaboration of a systematic methodology which would utilize available information to permit arrival at realistic estimates. For these reasons, and perhaps others, rather strong compromising assumptions are customarily admitted regarding such matters as uniformity of componental characteristics, statistical independence of component failures, and impossibility of transient failure of a component followed by failure in another mode. These assumptions characteristically lead to compromised estimates which are in disagreement with empirical observations. Thus, certain classes of systems which, according to compromised estimates, "should have died long ago" are found to be still operating satisfactorily while other classes which, according to compromised estimates, "should work like a charm" exhibit behavior described by "there is always some little thing wrong." Some initial attempts have been made to replace the customary compromising assumptions with more realistic assumptions, but, to date, the surface has barely been scratched in this area. To attain a satisfactory degree of understanding of system reliability and a useful methodology for predicting it, the approach indicated schematically in Fig. 3.1 is apparently a necessity.

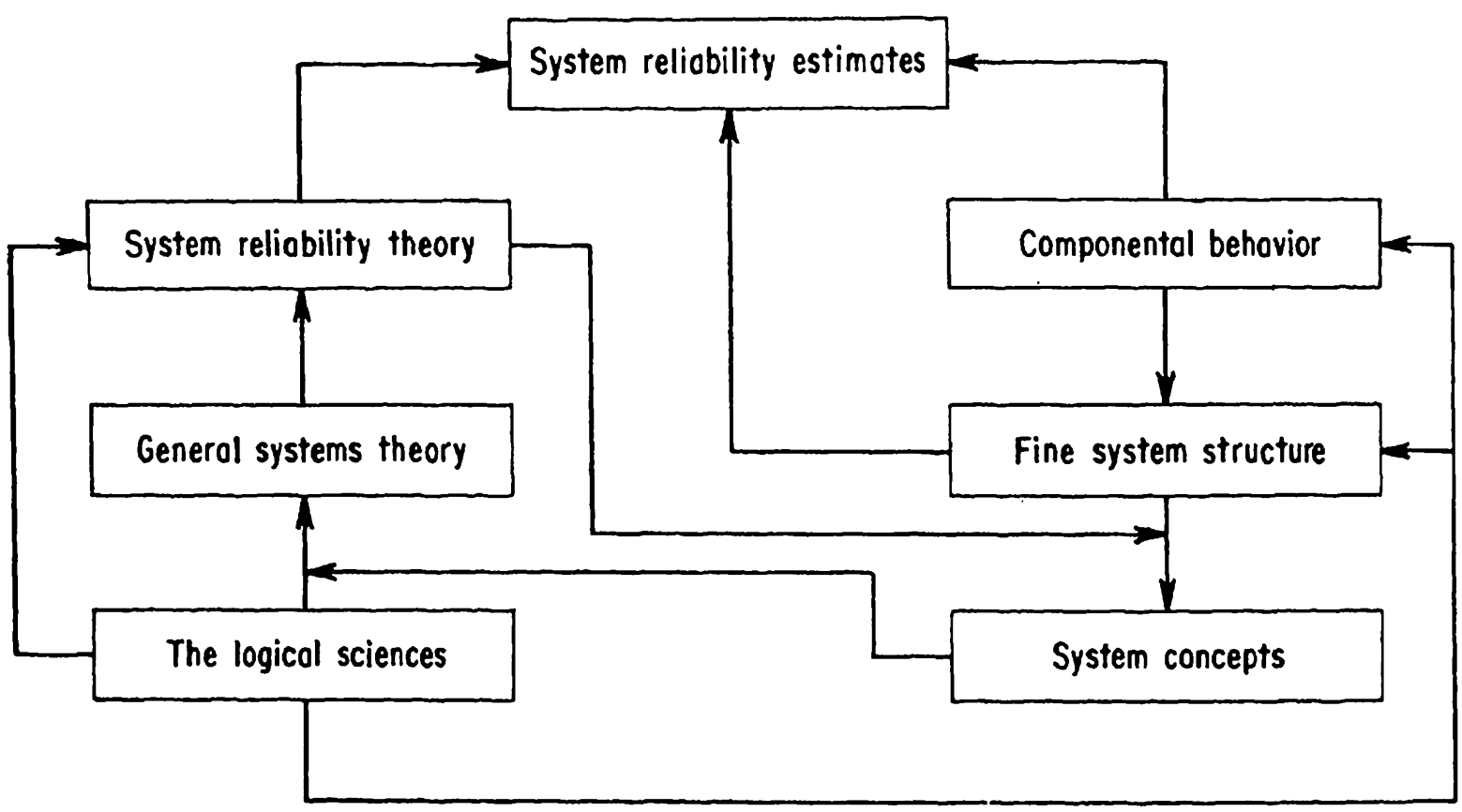


Fig. 3.1. Approach to system reliability concepts.

4. Reliability and Other Criteria and Design Objectives

As an example of the present emphasis on reliability relative to other system criteria, let us consider again the airborne digital com-

puter. Within the present state-of-the-art, relatively complex airborne digital computers can be constructed having volumes of under two cubic feet and weights of under forty pounds. It is a reasonable expectation that a strong miniaturization and packaging program could wring out these figures by a factor of $\frac{2}{3}$ to $\frac{3}{4}$. Recognizing that, in the distant future, airborne computers may be required of much vaster complexity, it is evident that effort in microminiaturization is necessary. It is also evident, however, that for most purposes the size and weight of airborne computers will be so negligible in comparison to the vehicle utilizing them that this is no longer a primary design objective. What is of prime significance at this state is reliability. This is a distinct change from the early days of digital computers when it was far from obvious how an airborne vehicle could accommodate such a machine. In graphic form, the situations may be compared as illustrated in Figs. 3.2 and 3.3

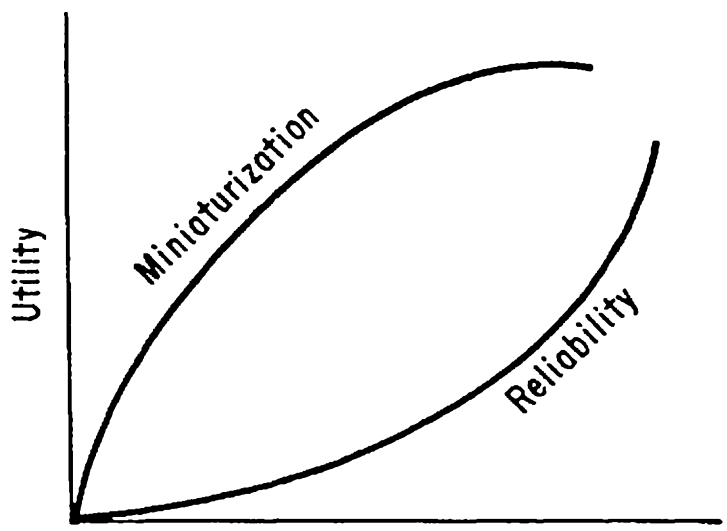


Fig. 3.2. Early stage of marginal utility in airborne digital computers.

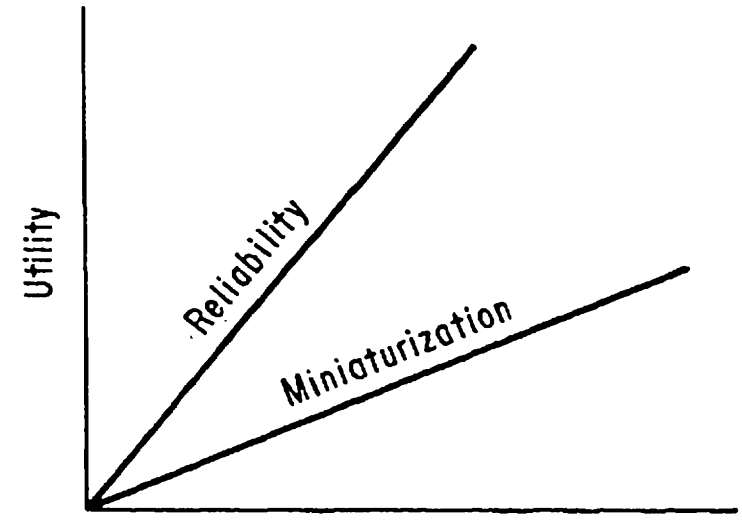


Fig. 3.3. Late stage of marginal utility in airborne digital computers.

It is clear that the first criterion must at present be reliability, as even a computer one cubic micron in volume and one micromilligram in weight is only of academic interest unless its faithful operation can be reasonably assured throughout a critical mission. As compactness (size and weight) has been, so have the problems of speed and capacity been temporarily conquered. Latter day computers offer speeds and capacities adequate for the computational requirements of latter day aircraft. Thus, by elimination, our second and third criteria for the present time become flexibility and operability, respectively, in order that a machine may be most fully utilized. Producibility is our fourth present criterion, for machines are needed in quantity and without long lags induced by inflated

testing and checkout requirements. To summarize, we should weigh the state-of-the-art against the primary criteria of

(a) Reliability,
(b) Flexibility,
(c) Operability,
(d) Producibility,

while retaining in mind those of

(e) Capability (speed and capacity),
(f) Compactness,
(g) Power requirement,

when looking to the future.

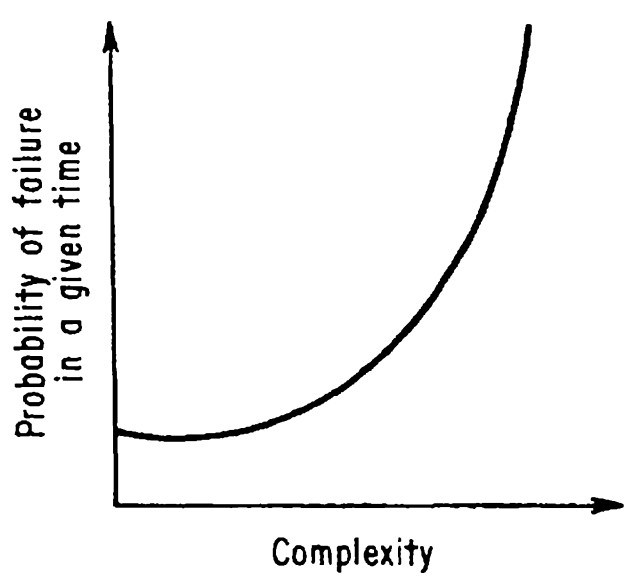


Fig. 3.4. Exponential decrease in reliability.

5. Exponential Decrease in Reliability with Increasing Complexity

As one may observe empirically as well as predict intuitively, the growth of complexity within a relatively fixed system design approach at a given state-of-the-arts causes roughly the exponential decrease in reliability graphically shown in Fig. 3.4.

6. Exponential Decrease in Reliability

This is clearly a restatement of the "law of diminishing returns" and indicates saturation in a very rapid fashion. Many present day systems have duty cycles (100 × uptime/total time) of from one to ten per cent mainly because of the saturation in complexity. Of course these figures are not completely accurate, owing to lack of complete statistical independence of components and instrumentation.

There is a certain paradoxical aspect to the reliability *vs.* complexity question in that certain designs may, in many circumstances, enhance reliability while simultaneously increasing complexity. This is accomplished on a component or higher subsystem level by "sharing-the-load"

or "stand-by" capabilities. An analysis of a typical approach of this kind is shown in Appendix 9. Unfortunately, the same difficulties pointed out previously in this section make difficult an exact evaluation of the utility of low-level redundancy techniques.

7. Self-Checking

Self-checking of a system for errors is a further method offering relief for lack of reliability. Although no general approaches to this concept currently exist, techniques are well known in special classes of systems. As an example, consider the use of redundant codes in digital computers: the principle of redundant codes is undoubtedly the oldest and most frequently used method of detecting computer errors. If not all possible coded words are utilized, the remainder can be referred to as inadmissible conditions. Certain errors which may be made in computation typically result in inadmissible conditions. Thus, the apparent redundancy may be used as a partial error checking system.

In the normal 8—4—2—1 code, there are sixteen possible combinations of four binary bits that identically are the binary numbers zero to fifteen. In the binary coded decimal machine, only ten of these combinations are used; namely, the binary numbers zero to nine. The remaining six combinations of binary ten to fifteen are not used and are called redundant states because under normal operation these combinations will never exist. Knowing this permits the computer controller to utilize a detection circuit that will recognize one of these disallowed combinations when it comes up. Thus, any error that results in one of these combinations will be detected immediately by the detection circuit.

This procedure will not catch all of the possible errors, for an error (which would not be detected) might result in an allowed combination. Consider that the probability of detecting an error in a randomly defined code is directly dependent on the ratio of allowed to disallowed states. Therefore, a five-bit code representing our decimal digit would have only ten allowed states as before, but would now have twenty-two disallowed states. Obviously, this increases the probability of a random error producing a disallowed state.

If one chooses allowed states intelligently, however, one is able to manipulate the system so that a single random error will always produce

a disallowed state. Note, for example, the following codes that could be used for the ten allowed states:

00011	01010
00101	10010
01001	10100
10001	01100
00110	11000

It will be noticed immediately that all of these codes have exactly two out of the five bits. Changing any single bit in any code will always produce a disallowed code. A very large number of such codes can be developed, and the choice of the one to use is usually based upon the ease with which the code can be used in the machine.

A very commonly used coding arrangement is built upon a principle known as parity. In this method, any normal coding technique convenient in the normal machine operations can be used. Then an extra bit, called the parity bit, is added to each code. This bit is always assigned a value which will make the total number of ones in the code even (even parity) or odd (odd parity). Odd parity is most commonly used, since it rules out the all-zeros state as an allowed state. With odd parity, a single change of a bit in a code will automatically make the number of ones even, signifying a disallowed code.

Parity is most commonly used as a check on memory devices and information transfer operations. Each time a number is stored in or read from the memory, its parity bit is checked. This will inform the machine when a bit has been dropped or picked up in the transfer process. Since the memory tends to be the least reliable part of digital circuitry, the parity check has distinct value. Parity also has the advantage of being comparatively inexpensive to utilize.

The principle of parity can be extended in several directions to improve reliability further. As a simple example of what might be done, a complex parity arrangement will be developed for a binary machine using a sixteen-bit number length. Adding a single parity bit to the number increases its length to twenty-one bits, and it enables the machine to detect single errors in the number. In this form the number 33,123 would appear as:

1, 1000, 0001, 0110, 0011.

The number has been divided into four groups of four bits each for con-

venience, with the parity bit appearing as the first bit on the left. Odd parity is being used.

If the parity application now is extended by dividing the number into four groups of four bits and adding a parity bit to each group, the following representation of 33,123 is obtained:

0, 1000
0, 0001
1, 0110
1, 0011

The parity bits again are set off to the left of each group of four bits. This added parity adds very little to the over-all error detection capability of the machine. It merely locates the error more closely. However, it now allows the addition of still more parity bits that will accurately locate the error. The new parity checks will be performed on each column of the groups of bits as arranged above. In this particular case, the column parities are shown below:

0, 1000
0, 0001
1, 0110
1, 0011
———————
1, 0011

The parity bits on the left of the table are called "row parities," and the parity bits at the bottom of the table are called "column parities." The bit at the lower left corner is a parity bit checking the other parity bits. It may be considered either a "row parity" checking the four "column parities" or a "column parity" checking the four "row parities."

With the arrangement as described, it becomes apparent that any single error can be located exactly. Also, any combination of two errors can be detected.

The foregoing example is given only as an illustration of how the parity principle can be expanded.

8. Redundant Order Code

The discussion of redundant codes so far has concerned only the numbers actually being operated upon by the computer. The information

applies equally well, however, to such things as addresses and order codes. Nearly all general-purpose machines have built-in detection networks that sense the presence of an order code that has no meaning in the machine. This form of checking is comparatively cheap and can prevent some very peculiar operations that might occur if the machine tried to follow one of the disallowed order codes.

9. Parity in Arithmetic Operations

The use of parity as described above is adequate to finding errors in the processes of storing and transmitting numbers, but it is restricted to operations that do not normally change the number in any way. This means that arithmetic operations are immediately ruled out. To make the parity usable in the rest of the machine, it is necessary to provide a means of generating the parity of the new numbers as they come from the arithmetic elements. This procedure restores the parity but provides no check on the arithmetic operations themselves.

In order to check arithmetic operations by parity methods, it is necessary to consider the basis of parity and why it gives information. The parity bit (or bits) is merely a tag that is attached to a number which yields some information about the number. If the number ever changes, the hope is that the parity will indicate this fact. Thus, the whole parity principle involves making up a tag at one time and then using it, subsequently, from time to time to see if it is still correct.

If the parity bit is to be useful in checking an operation that changes a number, then the tag for the new number should be made up in advance and then used to check the new number as it is produced. To illustrate how this might be done, consider a simple adder. In the add process it is found that the parity of the sum of two digits is always the same as the sum of the two individual parities except when there is a carry from the new lower significant place. From this it can be shown easily that the parity of the sum of two numbers is the same as the sum of the parities of the numbers being added and the parity of the carries produced during the addition process. A predicted parity based on these operations is formed during the addition operation itself, and then compared with a parity generated from the sum as it comes out of the adder. However, there are some pitfalls in this method that should be avoided. For instance, a failure in the carry generation will produce compensating errors in both the predicted parity and the parity of the sum.

10. Summary

In this subsection we found that complexity in a system becomes a problem mainly because of reliability considerations. Present theoretical bases are insufficient for realistic reliability estimates. Redundancy and self-checking are two methods sometimes employed to combat the loss in reliability caused by increasing complexity.

Section 3.3

Human Subsystems

> *"Biology is truly a land of unlimited possibilities. We may expect it to give us the most surprising information, and we cannot guess what answers it will return in a few dozen years to the questions we have put to it. They may be of a kind which will blow away the whole of our artificial structure of hypothesis."*
>
> —S. FREUD

1. Human Problems

The problems arising in connection with human subsystems are, for the most part, heightened versions of those issuing from other subsystems. The human subsystem, however, is associated with quantities of myth, morality, and superstition, which in some measure are also intrinsic to the analyst. Further, this component also has an individuality which seemingly defies quantizable description except under rigorously controlled experimental conditions and then only in very restricted activities. Hence, the problem of describing human responses to a variety of environments and selected inputs is of primary concern and timeliness.

2. Human Input Capabilities

It is readily apparent from the technical literature in the areas of human engineering, industrial engineering, operations research, and management science that many studies have been carried out whose orientation is toward the determination of the characteristics of human response to and utilization of certain types of information flow. It is, unfortunately, equally apparent that almost all of these studies have been of an empirical or heuristic nature so that their end products consist of prescriptions for solution of very narrow problems or of statistical inferences with generally weak support. A few significant empirical rules-of-thumb have been formulated, such as:

(a) In general, human beings are incapable of simultaneously utilizing more than six relatively uncorrelated channels of information;

(b) The human being, as a sampled-data monitor, can utilize many more than six channels provided complex abnormalities among the channels are locked out.

(c) The complexity of abnormalities among channels to which a human being can effectively respond may be increased with training but apparently asymptotically approaches a bound far below that predictable by simple theories of the central nervous system.

Let us consider some topical areas and illustrative examples having content bearing upon certain aspects of human response characteristics.

3. Human Sensory Abilities

As indicated above, there are certain restrictions upon the responses of human beings to environmental stimuli in general, and upon specific inputs in particular. Despite these limitations, there are several domains wherein human capacities frequently exceed those of devised sensory equipments, i.e., in terms of interdependent sensory thresholds, or in the amount of energy necessary for the excitation and response elicitation from certain exclusively natural (in this particular context, human) receptors. Two such receptors, about which a large amount of

basic information has been accumulated, are the eye and the ear. The discriminative efficiency of these primary human sensors, within their ranges of operability, has rarely been surpassed by artificial sensors. Of course, certain analogous mechanical means have been devised which are sensitive to energy changes in realms beyond the wavelength bands to which human eyes and ears are sensitive.

4. Human Perceptual Abilities

Perception includes those processes of immediate conscious cognition or awareness mediated through the sensory channels. Sorting activities provide examples of a perceptual process of limited scope involving the comparison of objects with each other or with given physical standards. In the broadest sense, one may consider such operations as requiring only elementary or simplex decision-making activities in that secondary process or intermediate steps are not usually involved. Sorting is routine, mechanical, and iterative in those circumstances requiring the manipulation of marked material objects in accordance with some protocol involving only observations of the marks and positioning of the objects following observation. [Intimately related to the fundamental sensory capabilities are the unique (as yet) human capabilities in the generalization of stimuli, (that is, gestalt perception). Energy patterns, if minutely observed by an adequate objective process associated with a given object or object class, differ with the conditions of ambient light, noise level, etc. Nonetheless, essential qualities are perceptible and a familiar object is identifiable.]

The sorting process is indeed simple when gross differences only are of interest. The sorting task becomes more complex as subtlety is introduced. Thus crude sorting is well within the range in which human beings may operate at high efficiency if they are suitably motivated and possess the sensory capabilities which permit discernment of differences. However, other factors must be considered if the sorting activity is not an isolated one. A continuing activity of monotonous sorting at a crude level would soon induce fatigue, boredom and, likely, a general deterioration of whatever motivation previously existed so that defects in judgments would appear with significant frequency. Fortunately, machines accomplish such onerous tasks exceedingly well and, although requiring attention, do not yield to psychological or physiological need or frailty.

5. Encoding and Decoding

Recognition of the similarities between the activities of encoding and decoding, insofar as the human organism is involved in transformation processes, suggests a unified approach to these topics. Significantly, either process requires change of a given set of arbitrary symbols or representations into another, and each is but a variant of the general process of coding. Ashby† cites an example of the ubiquity of coding which is encountered in an ordinary reaction between the human organism and its environment. He asks that we

> consider the comparatively simple sequence of events that occurs when a "Gale Warning" is broadcast. It starts as some patterned process in the nerve cells of the meteorologist, and then becomes a pattern of muscle-movements as he writes or types it, thereby making it a pattern of ink marks on paper. From here it becomes a pattern of light and dark on the announcer's retina, then a pattern of retinal excitation, then a pattern of nerve impulses in the optic nerve, and so on through his nervous system. It emerges, while he is reading the warning, as a pattern of lip and tongue movements, and then travels as a pattern of waves in the air. Reaching the microphone it becomes a pattern of variations of electrical potential, and then goes through further changes as it is amplified, modulated, and broadcast. Now it is a pattern of waves in the ether, and next a pattern in the receiving set. Back again to the pattern of waves in the air, it then becomes a pattern of vibrations traversing the listener's eardrums, ossicles, cochlea, and then becomes a pattern of nerve-impulses moving up the auditory nerve.

Apart from the inanimate types of codings involving manipulations within the nonhuman environment cited in Ashby's example, there are no less than eight major transformations throughout which something has been preserved, though the superficial appearances and psycho-physiological processes may have changed. Undoubtedly, more complex examples could be cited, but in these complexity and numbers of transformations required by the human organism do not appear to be the main sources of response difficulty. Rather, the very variety of disturbances discernible by the human being is inimical to optimal coding performance. Let us

† W. Ross Ashby, *An Introduction to Cybernetics* (New York: John Wiley and Sons, 1956).

next consider a simplified example: the activities involved in solution of the ordinary crossword puzzle. We may disregard certain elementary knowledge required or implied by the physical format of the puzzle itself and concern ourselves solely with the transformations involved. For each word of a puzzle, a definition or an ordered assemblage of clues is given. The would-be solver knows at the outset only that the word sought consists of so many letters in a certain sequence and that these letters are to be aligned vertically and horizontally. If the solver is a "good speller," his pathway may be eased somewhat. He is further aided by the possession of an extensive vocabulary. A good memory, both of long and short term, will also help him, because devisers of crossword puzzles frequently display in their work an extreme affinity for archaic and obsolete word forms. On the other hand, if the experience of the puzzle-solver does not include some previous exposure to such word-forms, his task will prove more difficult. The puzzler begins his task by taking each word in order, if systematic in this sense, and by checking and eliminating, as he proceeds, incompatible letter or word linkages. He is able by means of his previous linguistic experience to anticipate many incompatibilities prior to setting his choice of the word required into the proper boxes. The point is that there exists a common body of linguistic experience which enables individuals of widely divergent backgrounds to communicate more or less effectively, even via the cumbersome device of a crossword puzzle. Judgment enters, to be sure, in choice of words spoken or written, but common words of nearly universal use may be employed should any doubt exist in the expressor's mind concerning the interpretation by the listener or reader.

6. Evaluation

Evaluating, or summation of the judgment processes basic to complex response, consists of several processes forming a mental activity complex. First, perception defines the object, message, or circumstances of immediate concern. Experience facilitates perception in that it enriches perceptual scope and possibly short-cuts the perceptual process. Habitual patterns of thought also affect the process, as well as categorization within the limits of individual experience. Similarities are sought between the item of momentary interest and those previously encountered within the individual's experience, either through direct exposure or via the recorded experiences of others.

The distinctive quality of human performance in the solution of complex problems is as yet unparalleled (with the exception of certain iterative problem-classes) by machines. We are presently unable to provide accurate descriptions of human thought processes at levels much beyond the simplest stimulus-response pattern. However, we can recognize the unique, human-associated character of these processes and attach to them such labels as "reasoning," "judgment," and "selective recall." Conventionally, we reserve the use of such labels for the generalized description of inadequately understood human thought processes. Despite this lack of precision in definition, it is clear that the human organism enjoys certain capabilities not as yet embodied in other machines. A suggested, though inexhaustive, list of these capabilities follows:

(a) Extracting something recognizable from a dynamic environment because it has been previously experienced. In psychological idiom this capability is referred to as *stimulus generalization*;
(b) Discerning poorly defined, unquantized, and frequently irrelevant items of information, and subsequently having such items "on call" selectively;
(c) Perceiving analogous behavior in dissimilar systems without change in the fundamental receptor structure;
(d) Self-assessment or, in the system idiom, determination of a gross approximation to the state occupied;
(e) Utilizing information which is qualitative in nature, and further assigning such qualitative information relational positions (upon a continuum of uncertain dimensions);
(f) Using a nonrigorous language structure, i.e., nonrigorous in the sense that there is no isomorphic relationship between symbol and designatum;
(g) *Synaesthesia*, i.e., concomitant sensation (in certain instances of an altogether subjective nature) of experiencing data in terms of another sense than the one actually being stimulated, as in color hearing, in which sounds seem to have characteristic colors. These phenomena are not clearly understood, but to the extent that they are operable they serve as stimulus reinforcement means (they function as built-in sensory redundancy);
(h) Operating conceptually in "universes" not apparent as a direct result of experience; i.e., conceptualization of geometries not locally Euclidean, spaces of power greater than that of the continuum, etc.

7. Judgment Errors

We have now considered a parial list of the complex processes invoked by the activities fundamental to decision-making. There are many more, to be sure, and the underlying principles at the present state of our knowledge of these others are probably as poorly known. Let us now explore the types of errors commonly encountered and perpetuated in human communication and response, either directly or by instrumental means. Systematically speaking, these are the types of errors caused by the presence of subsystems composed of series-combinations of human beings within a supersystem context. A partial list would include errors stemming from:

(a) *Overloading*, i.e., unrealistic output demands upon human perception and information handling capacity. An example in point is provided by the situation of the pilot in modern military aircraft who must rely for attitude and director information upon a battery of discrete displays. These are inadequately integrated so that the pilot must synthesize the information displayed within some permissible interval for corrective control as well as implement the effector action required for control adjustment.

(b) *Fatigue*. As duty cycles are extended and the action density of an operator within a given system increased, the frequent result is that the probability of human fallibility is enhanced with consequent unreliability penalties placed upon the system.

(c) *Multiple functions performed simultaneously*. Linearity in the output of human performance occurs only in simple repetitive processes done singly. There may be certain simple sequences of action where predictable linearity within reasonable limits might be anticipated. However, as a general rule there are severe limitations upon the number of discrete types of action of which the human being is capable without nonlinear performance. Thus, as variety in the task inputs increases, linearity of performance decreases and so occasions the interposition of various compensatory electromechanical functions in, for example, a hybrid control system.

(d) *Monitoring without informational feedback*. Here, the human being's performance quality suffers not as the result of incapability but from the necessity for variety in activity, albeit only of an intellectual nature. Task variety, if not overdone to the extent of inducing operator confusion,

sustains alertness and decreases the effects of fatigue and hence lessens system instabilities that are due to human causation.

(e) *Uncontrolled ambient environment*, i.e., the necessity for recording or transforming under difficult environmental conditions. Conditions of ambient noise or light, ventilation, and temperature, if not properly controlled, may induce degraded operator performance.

(f) *Intellectual disparities*. Although of late there has been a surfeit of tests purporting to measure individual predisposition to and adequacy for certain types of operator performance requirements, none has been so presumptuous as to claim certainty in the prediction of on-the-job adequacy of the tested operator.

(g) *Dynamic anthropometry*. Frequently the human operator works with devices and in surroundings not compatible with individual dynamic anthropometry. Considerable study has been devoted to the problems occasioned by restrictive work environments and by having instruments form a portion of the environment. As man enters new environmental domains at even greater variance with his normal environment than are those associated with advanced vehicular systems common to the domains of earth's atmosphere, it is likely that more restraints will be imposed upon the operator. The problems associated with error-generation by the human component in such systems have already become critical and are in large measure accountable for the retardation of man's entry into extraterrestrial environments.

(h) *Precision demands*. These may exceed human sensory capabilities. Such conditions commonly exist and frequently result from inadequate prior human engineering studies. Psychologists and physiologists have acquired large amounts of fundamental information about the limitations and extent of human sensory capacities in a great variety of situational environments. The relevant literature abounds in experimental and collative studies directed toward the determination and cataloging of such information.

(i) *Motivational, emotional, psychophysiological operator status*. Intractable as such psychological states are to analysis, it is clear that they account for a significant proportion of operator-induced system errors. This may only happen through the distractive effects which they impose upon the attention-requiring functions of the operator.

(j) *Response incompatibility*. The "natural" habit or response pattern may be incompatible with required visual-motor functions. Conflict may exist between status conditions as reflected by indicator readings and

proprioceptive sensations, or in other instances between the types and sequences of visual-motor functions required for human response within the complex display/control array characteristic of the hybrid system. As in the case of Item (h) above, this situation is apt to develop if insufficient attention is given to preliminary human engineering studies prior to the inclusion of the human component within the operational system context.

8. Man-Machine Comparisons

It will be of interest to compare man with machine subsystems for certain functions. A summary of the more or less traditional conclusions in this area follows. These conclusions are of general interest, but clearly do not, in fact, have the totally universal applicability frequently claimed for them.

(1) SENSORY ABILITY. It is often stated that the human sensing capacities of vision and audition generally surpass those of machines. The sensitivities of these two important sensory modalities of the human organism are frequently cited as challenges to the ingenuity of the instrument maker and the engineer. In the distance receptors of the ear and eye there are artificial considerations which have not been applied in the design of counterpart instruments. We are speaking here of phenomena which are truly unique to the human organism. Examples of such features are, in the eye, the effects of after-images which result from the lag in refocussing, or the perception of the quality "whiteness" as appearing whiter when observed against a black background. Similar phenomena are noted for the ear. These features are consequences of built-in limitations of anatomical and evolutionary origins.

Man is a multi-purpose assemblage of sensory and motor capabilities. These capabilities interact in a way fundamentally similar to the workings of machines which incorporate various feedback schemes, but in a manner much more complex—not more complex in terms of the basic action relationships between error sensing and corrective adjustment, but in terms of concomitancy and sequencing of various acts of *both* sensory and motor nature. Thus, we are probably justified in saying that man senses not just one stimulus in even the most rigidly controlled experimental situation; he senses many. He will therefore always select, from a variety of responses available to him, one which is more or less appropriate to the situation.

Hence, we may say that, quantitatively, the sensing capabilities of human receptors are but approximated, not known. It follows that the interrelationships between sense modalities are even less capable, at the present, of quantification.

(2) PERCEPTUAL ABILITY. "Man excels at comprehending complex data presented pictorially or symbolically." This statement is progressively weakened by almost daily reports in the technical journals concerning new achievements in electromechanical pattern recognition. In the mechanization of these processes, ordering schemes are brought to bear in the form of rapid discriminative subroutines. Given an experimental library (mechanical memory) of vast dimensions such as that envisioned by present computer technology, and access speeds such as those within our grasp and presently employed, we are quite able to match or exceed human competence in these matters. Mechanized sensing and information transmission now depend upon the presence or absence of atomisitic elements in the stimulating source.

However, it should be noted that while both man and machine are currently capable of recognizing patterns *from prescribed universes*, machines (unlike man) can not yet discern relational patterns among items from non-prescribed universes.

(3) PREDICTION. "By being alert for changes, human beings can frequently anticipate undesirable conditions." The statement is no doubt true. It no longer, however, applies uniquely to human beings, for adaptive systems, or systems exhibiting "intelligent" behavior, have usurped another niche in provinces formerly enjoyed only by organisms. The designers of so-called adaptive systems refer to "goal-oriented" behavior—toward the achievement of an objective by the machine system.

(4) FLEXIBILITY. Humans possess a degree of flexibility of action which may protect them against total system failure. Programming of a given machine for corrective or alternate action in all conceivable emergencies is not generally feasible. However, man is frequently capable of inventing expedient alternatives or repairs in such situations.

(5) MEMORY. Man is reasonably efficient in tasks requiring long-term memory.

(6) JUDGMENT AND REASONING. Men are needed to make judgments when it becomes impossible to reduce operations to logical, preset procedures. Here again the flexibility mentioned in (4) must be noted.

(7) SPEED AND POWER. Clearly, machines can be devised to make specific movements more smoothly, quickly, and powerfully than man.

(8) REPETITIVE OPERATIONS. Machines excel in repetitive and routine tasks; unlike humans, machines do not become bored or inattentive.

(9) COMPUTATIONS. Machines designed to perform specific computing operations are more efficient than man.

(10) SIMULTANEOUS ACTIONS. Machines can be devised to carry on a greater number of simultaneous activities than man.

(11) COMPUTER CHARACTERISTICS. Three papers are combined into a section entitled "Human Beings as Computers" in *IRE Transactions*, Vol. EC-6(1957), pp. 190-202. These articles probably provide the best estimates currently available of certain computer-oriented characteristics of man.

Their most salient propositions can be summarized as follows:

(a) The human central nervous system is composed of about 10^{10} neurons, each of which acts, relating a neuron to a component, as a distributed repeater.

(b) The principal function of the central nervous system is similar to that of computer for an enormously complex, multiple closed-loop (between 10^3 and 10^6 loops) servosystem.

(c) The output side of a neuron is called an axon. The principal axon connections number about 2×10^6 from each eye to the brain, about 5×10^4 from the ears to the brain, and about 10^6 from the remainder of the body of the brain.

(d) To date, there is no known transfer function covering all effective connections between neurons (these connections are called synapses). It is known, however, that all synapses behave as nonlinear filters and that, in general, beat-frequency coding is present.

(e) On the level of production of biochemical tools to sustain life (subconscious nervous control, chemical control, and metabolic coordination), man processes an estimated 3×10^{24} bits of information per day. The general pattern for these computations is contained in a "blueprint" of between 10^5 and 10^9 bits of non-redundant information. "Blueprintwise," man is only one or two orders of magnitude more complex than a bacterium.

(f) Current estimates place the quantity of information simultaneously storeable in the conscious human memory between 1.5×10^6 and 43×10^9 bits. The smaller number represents about 1000 times the complexity of the multiplication table, and the larger number represents the storage over 80 years of a person with perfect memory who receives from his environment and actively processes about 25 bits per second for 16 hours per day. A "reasonable" capacity is usually estimated to be about 10^8 bits.

(12) Decision Input Capability Comparison. Of particular importance is the ability of data acceptance for decision processes. If we observe Fig. 3.5, we see that increasing sophistication of mechanization leads to cojoined (branching and/or parallel) operations. Now, patterns of cojoined operations generally involve branch or decision points. Decisions possess two characteristics of particular interest in any automation consideration. These are the following:

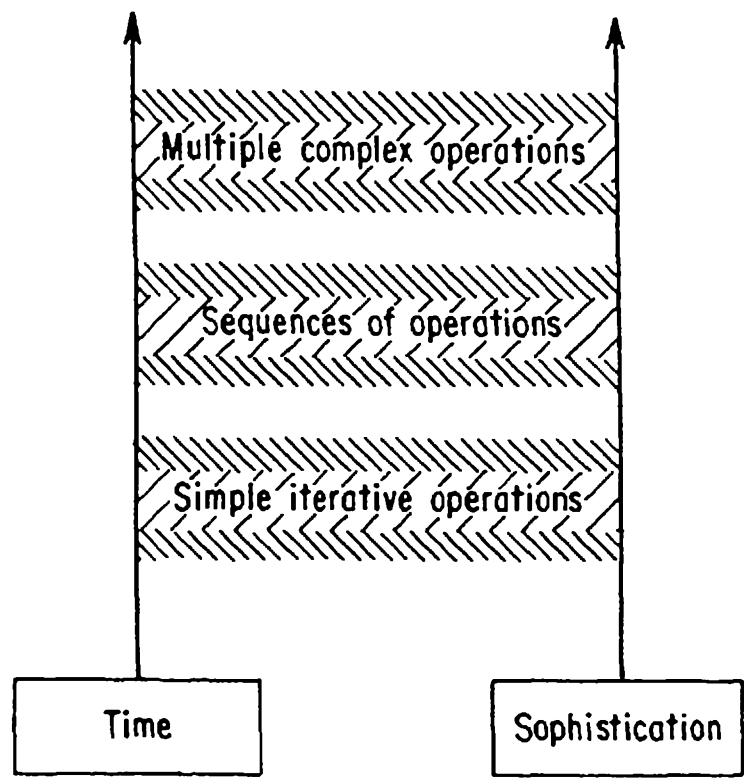


Fig. 3.5. Evolution of complexity of mechanization.

(a) *Necessary complexity*, the variety in amount and kind of information necessary to reach a "logical" decision.

(b) *Resolution*, the degree of specification of the necessary information.

The capabilities of humans and machines to reach "logical" decisions vary with respect to these characteristics and may be compared symbolically as in Figs. 3.6 and 3.7. Thus, for example, humans possess some effective capability in problems of very low resolutions, for which the machine is at a loss, but humans will quickly saturate under growing complexity while the machine, theoretically at least, retains relatively high capacity. We now encounter a paradox. Assuming that sufficient resolution is given, we note that decision processes may be symbolically represented as in Fig. 3.8.

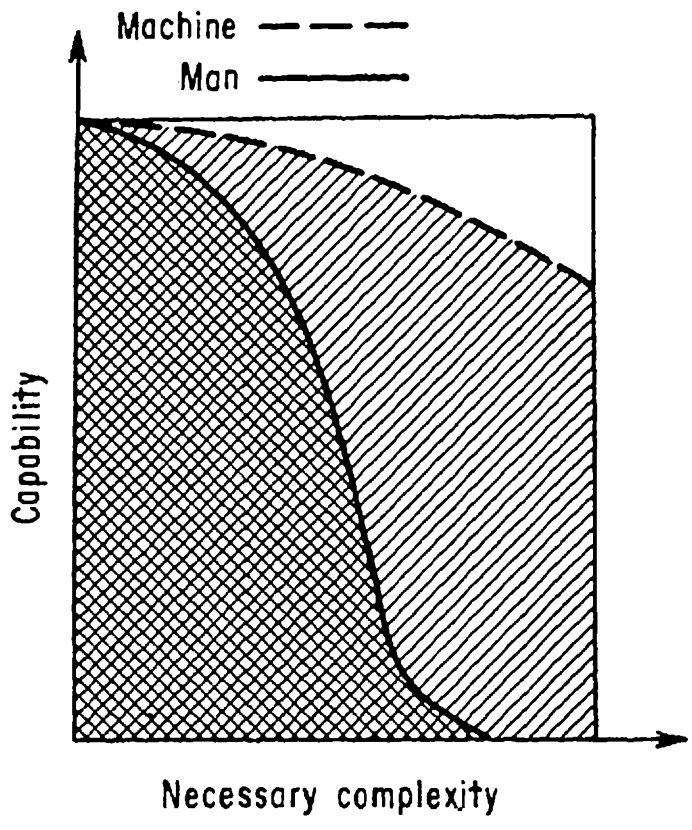


Fig. 3.6. Comparative theoretical response profiles: necessary complexity.

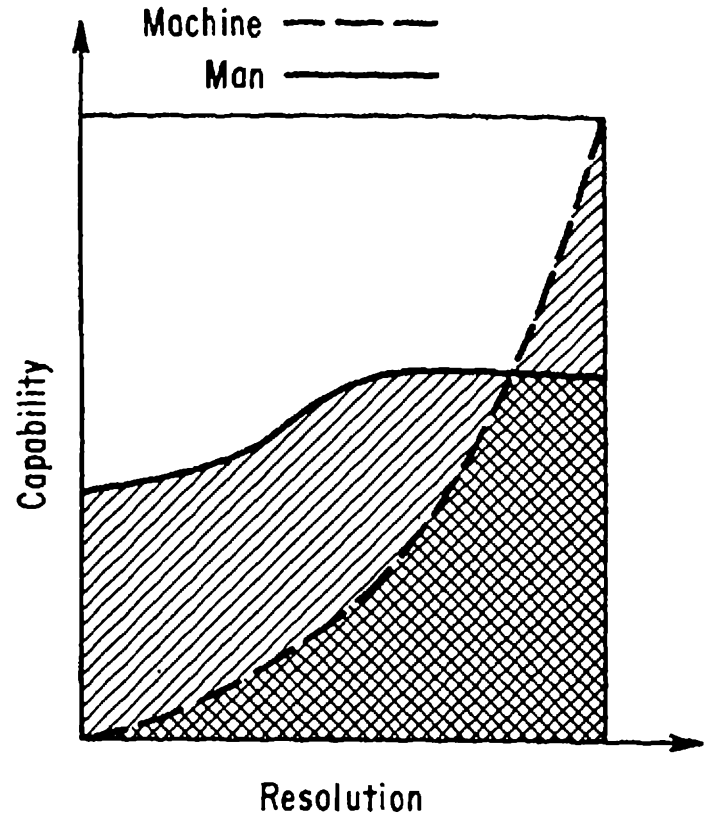


Fig. 3.7. Comparative theoretical response profiles: resolution.

By an atomistic decision we mean selection of one from a small number of (frequently two) alternatives. Machines make atomistic decisions easily and can, in some cases, handle simplex decisions. There are, however, very few machines that make complex decisions. This job is usually reserved for humans! We are here caught between the rock of theoretical technique and the hard place of technology. There will have to be significant advances permitting reduction in cost, size, and failure probabilities

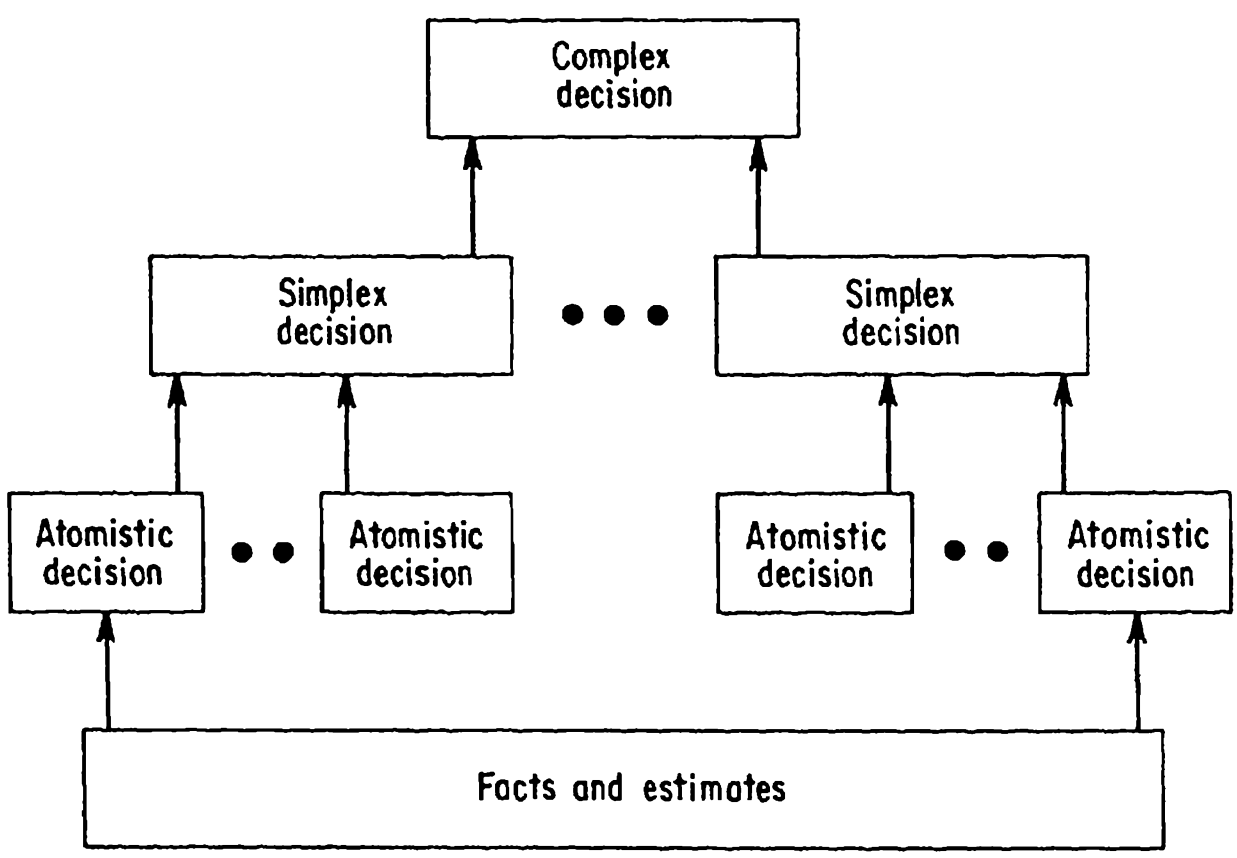


Fig. 3.8. Conceptual structure of decisions.

of hardware before we can actually construct the complex arrays envisioned by our theoretical potential and expect them to operate. For example, consider that our largest present-day computing machines, already so complex that twenty-hours operation without downtime is considered excellent performance, are unable to cope in full detail in the time required to make a solution of value with any of the problems:

(a) Meteorological forecasting,
(b) Simulation of a full scale strategic air war,
(c) Social Security System accounting.

9. Man-Machine Interfaces

The man/machine interface problem is one of the most difficult compatibility problems in system design. We discuss it now in certain of its simpler aspects.

General Consideration of the Display Art. A time framework is of value in organizing the discussion of this topic. The expression of current or imminently attainable technological knowledge in a display distinguishes the display as realized and therefore potentially conventional. Those displays which embody in their expression the characteristics of both conventional displays and highly integrative displays may be considered as advanced or interim displays since they represent progress toward highly integrative displays. An ultimate integrative display exists only as an intellectual construct.

The types of displays in military use with which we are here concerned are means of presenting information which may serve as the basis for decision and operator action. We exclude those displays or elements which play a passive role and are, in effect, ends in themselves. Such displays may be contemplated at leisure "after the ball is over" and frequently are but records of the passage or sequence of arbitrary units. Our immediate interests are in those displays or elements with which an operator interacts and which provide the bases for selection or modification of courses of action. We will now consider some of the characteristics of conventional, and interim and advanced military displays.

Conventional Displays. In general, conventional displays as of now are made up of discrete display elements, each of which presents distinctive information relative to changing values in physical areas of opera-

tional interest. Increased complexity of operation and, therefore, information requirements for control are thereby expanded until in some instances the array of dials, gauges, and indicators on the display looks like the end-of-the-line in pin-ball machines.

Now, engineering psychology has tried to reduce chaos to mere confusion in these matters, and has, in fact, (through experimental effort and the application of psychological principles) made certain advances in the treatment of conventional displays. The engineering psychologist may now speak intelligently about the characteristics of a good indicator, the design of numerals, letters and indices, and appropriate numerical progressions. He may manipulate the placement of these elements, for example, to achieve a rational operational sequence between display and control, and display and display, in terms of importance, and frequency of use. Such procedures make the best of a bad situation.

Let us consider those types of military displays that focus upon fresh approaches, for these represent efforts to come to grips with fundamental problems associated with display presentation.

INTERIM AND ADVANCED DISPLAYS: *General Discussion.* Interim displays reflect efforts toward the gradual evolution of simplified and integrative presentations. Conventional indicators are used in conjunction with more advanced instruments which combine several elements of information required for a given function. The basic improvement is reduction in the number of instruments. Primary instruments are apt to be unaffected by these modifications. Secondary instruments are more amenable to integration in the early stages of improved display development, and it is here that the noticeable improvements are gained with the least expenditure of effort.

There are certain principles which should govern the acceptable combination of interim displays. These principles are:

(a) Combine only those forms of information which bear a common relationship;
(b) Maintain the common factor of interpretation (fixed and moving, scale values, etc.) the same;
(c) Avoid confusing the operator by superfluous information.

Advances in the design of interim displays are reported with increasing frequency in military and trade journals. These advances are frequently motivated by the kluging-type of "systems approaches." The increasing density of information to be handled in the absence of evolutionary plan-

ning requires a continuing redesign of the control systems from the standpoint of the human operator. Essentially, the man and the machine components have to be integrated into a reliable network.

Another interesting point in interim-displays is the tendency to present anticipatory information at a rate consistent with the actual changes occurring as results of control manipulations. Check-reading is thus facilitated, and the operator is able to answer the self-question, "Am I about where I want to be, or not?" In other circumstances the question may be, "Is the operation running about the way I think it should be, or not?"

The preceding comments may convey the impression that display developments are quite independent of technical constraints. Unfortunately, such is not the case. While it is easy to speculate about the ideal attributes of a given display, compromise must be effected between what we would like and what may be done.

(4) MAN-OPERATED CONTROL SYSTEMS.†,‡ Empirical evidence suggests that the simpler the tasks imposed upon the human operator of a control system, the more precise and less variable become his responses. The

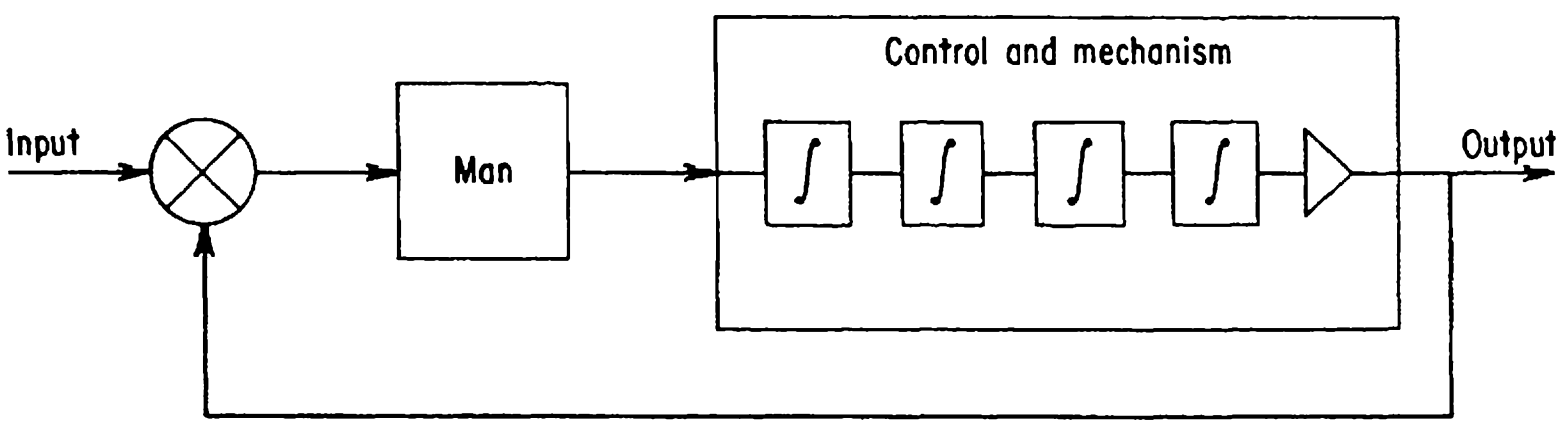


Fig. 3.9. A control system requiring quickening.

optimal performance of the element "man" in man-machine control system performance can be obtained only when the mechanical components of the system are designed so that the human being need only act as a simple amplifier. In systems embodying such a constraint upon design, such performance is achieved through "unburdening" (relieving the operator of the task of acting as an integrator) and "quickening" (provid-

† H. P. Birmingham, A. Kahn, F. V. Taylor, "A Human Engineering Approach to the Design of Man-operated Continuous Control Systems," *NRL Report 4333*, (Wash., D. C.: NRL), April 7, 1954.

‡ H. P. Birmingham, A. Kahn, F. V. Taylor, "A Demonstration of the Effects of Quickening in Multiple-Coordinate Control Tasks," *NRL Report 4380*, (Wash., D. C.: NRL), June 23, 1954.

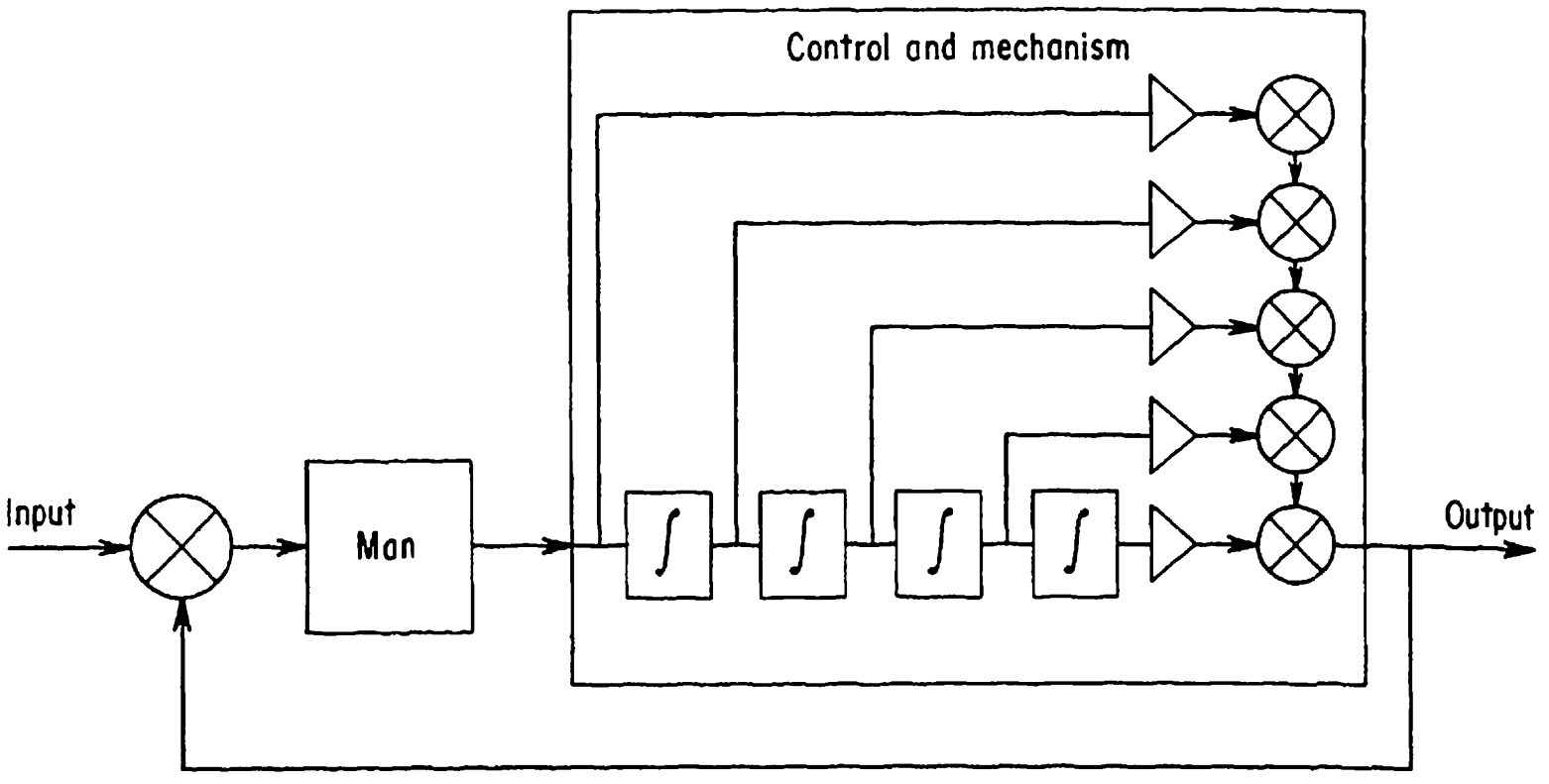


Fig. 3.10. Potential quickening inducing instability.

ing the operator with immediate knowledge of the effects of his own responses).

The control system block diagrammed in Fig.3.9 illustrates an application that requires a display quickening. This device is intended to operate on an input which consists only of step function position changes, so spaced that the full correction of any one step may be accomplished before the next requires action. Time constants of the four integrators are long, and this, coupled with the fact that the integrators shift the phase of the input through 360 degrees, causes the system to be quite unstable.

An obvious solution for the quickening of this device as diagrammed in

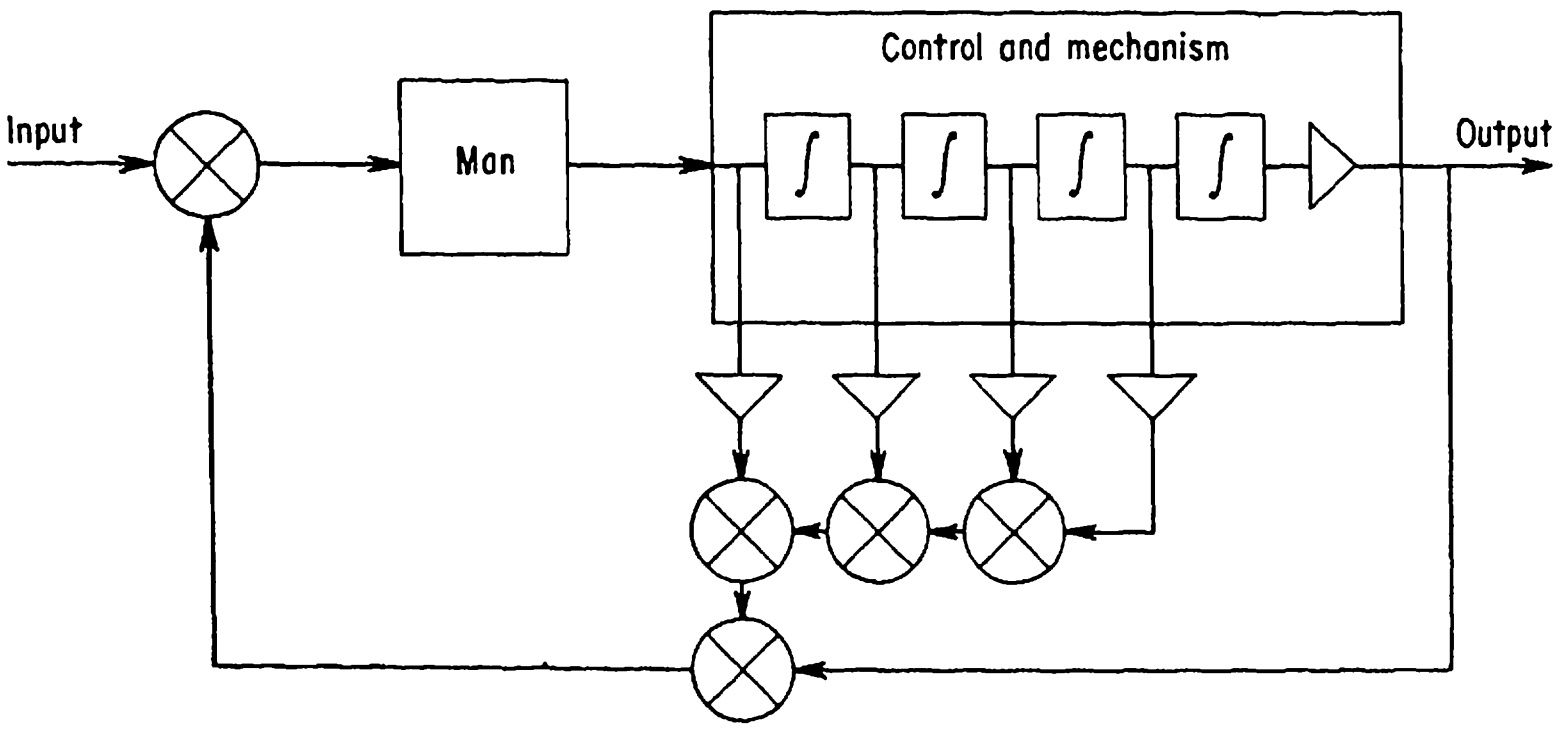


Fig. 3.11. Preferred quickening.

Fig 3.10 is ruled out in circumstances that prevent making changes that directly affect the output of the control system. The problem may be solved by picking off the components of (in the case of a mobile system) position, rate, acceleration, and the first derivative of acceleration, amplifying them properly and adding them algebraically in the feedback loop going to the display (Fig. 3.11). The display may then be given the form of a simple arrangement of double-pointer dials, with one pointer responding to ordered input and with the other controlled by the quickened feedback. With this arrangement, the man has only to operate his control so that the follow-up pointer matches the input-pointer at all times. Tests run with such devices have indicated that instability may be nearly (if not completely) eliminated by this means of quickening. The preceding paragraphs should not be interpreted to mean that quickening cannot be applied to the output of a system under any circumstances. Quickening, as a concept, is essentially the same as "aiding" (such as used for many years in aided tracking) and may be just as effective in the output as in the input.

10. Summary

Having indicated some of the problems associated with human subsystems (primarily those of inadequate description and informational interfaces) and having compared certain characteristics of men and machines, we may well raise the question of why human subsystems should be utilized at all in system design. As will be pointed out in Chapter 5 in connection with compromise automation, there are certain theoretical reasons why this must be. More practically, as has been humorously pointed out by many, man represents a fairly high capability control element, already inexpensive in mass production and producible by inexperienced labor.

Section 3.4

Special Problems in Control Systems

"How slow the time to the warm soul, that in the very instant it forms, would execute a great design."

—THOMSON

1. Feedback

Control systems depend upon utilization of system output for what is termed *feedback*. The output is compared against desired performance, and the discrepancy information is used as a portion of input or to control the states of the system (i.e., "mode-switching"). The two concepts are indicated schematically in Figs. 3.12 and 3.13.

The comparison element may be so simple as to accomplish an addition or subtraction of signals or so complex as an evaluation by a human being of displayed performance information versus desired performance. As will be noted from Figs. 3.12 and 3.13, these types of control systems characteristically contain a closed-loop about the basic functional system. Conceptually, open-loop control, assuming very faithful system response, is

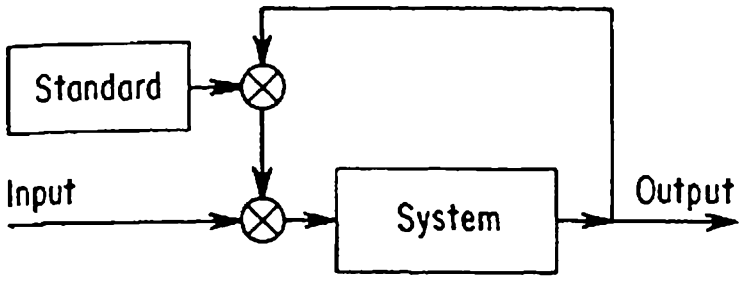


Fig. 3.12. Input feedback.

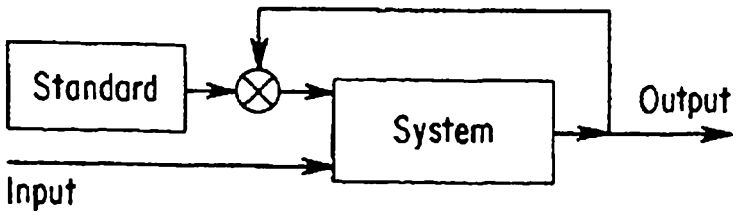


Fig. 3.13. Mode-switching feedback.

also possible where input-shaping or mode-switching is preprogrammed. Typically, we will have both types present in a complex system. If, for example, the airborne guidance system of Figs. 3.14 and 3.15 is utilized, and if the loop is closed through the pilot, his entering of checkpoint information would tend to null the errors of the inertial subsystem.

One of the major problems in feedback systems is that of stability. Unless the comparison element and the mode of injection of the error (discrepancy between actual and desired performance) information are

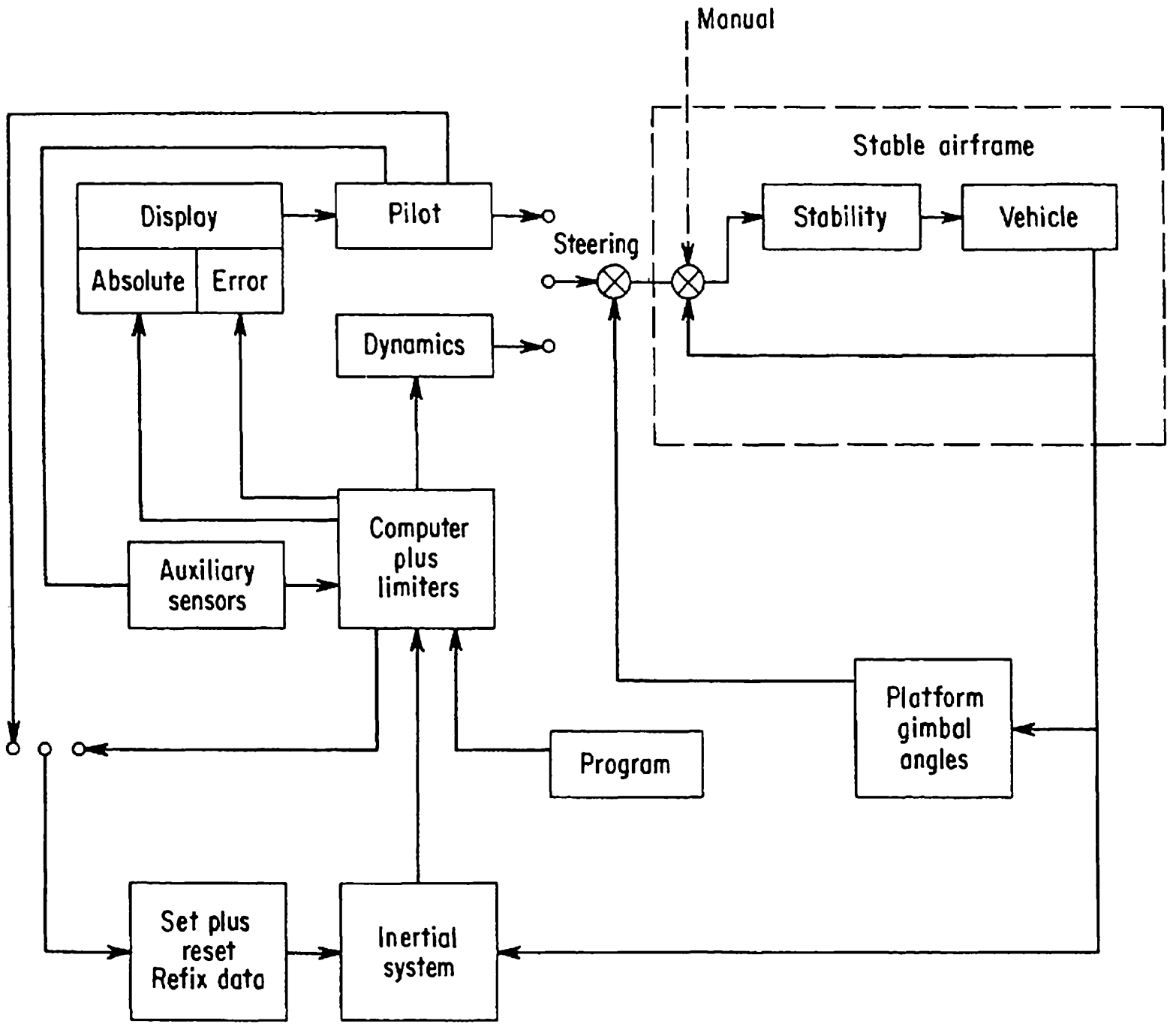


Fig. 3.14. Rudimentary guidance configuration.

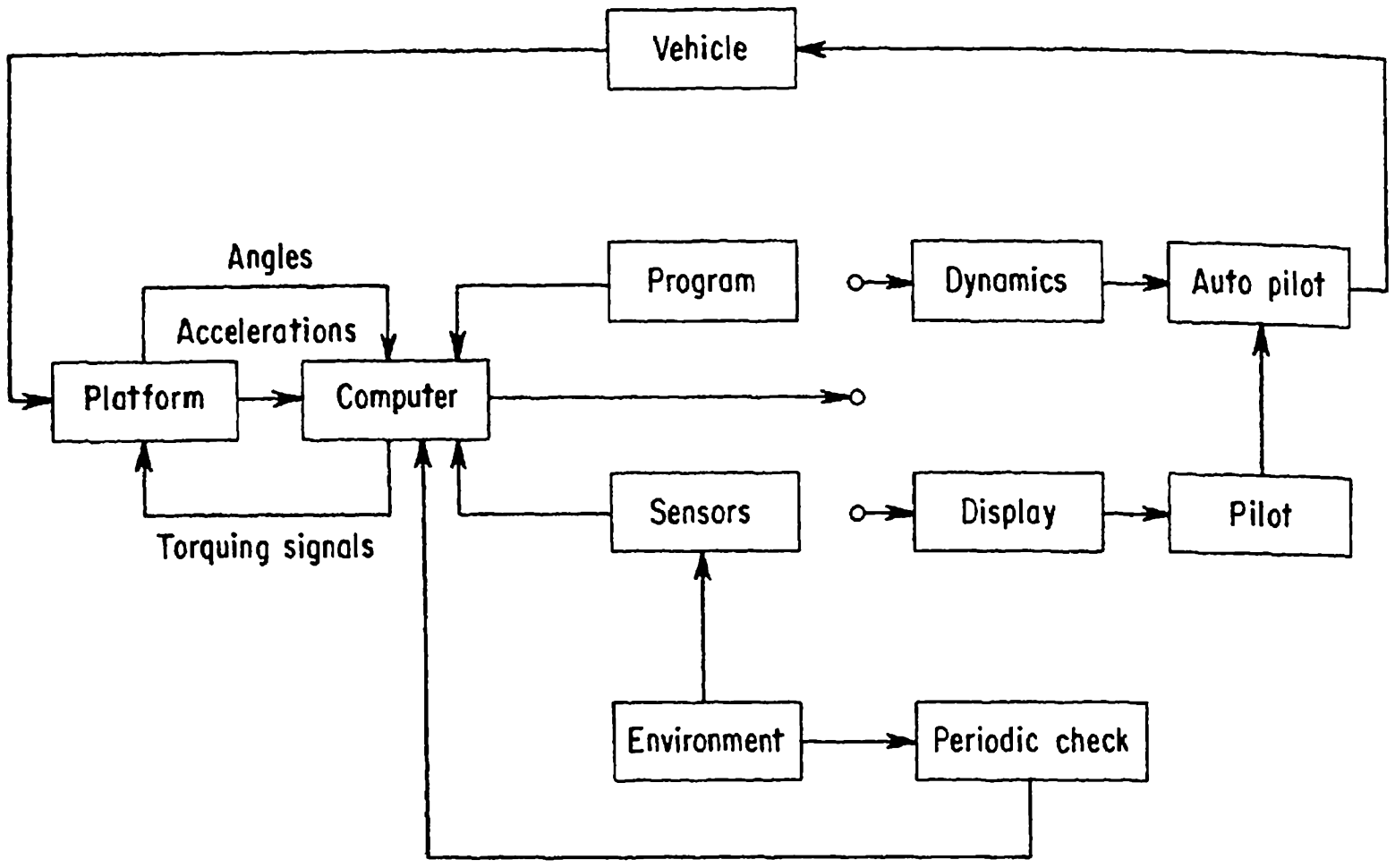


Fig. 3.15. Guidance-oriented view of control configuration.

properly selected, the system may "hunt" (oscillate about desired performance without convergence), or "overshoot," or "undershoot" without nulling. It is this very point to which much of traditional systems analysis is addressed.

2. More on Thought Processes

A second problem area is devised versus natural cybernetic systems. The term cybernetics is, for our purposes, synonymous with control and is usually preferred to control in discussions having a biological orientation. A source of frustration is our failure thus far to capitalize upon many of the desirable features of natural cybernetic systems in our designs. This general area of application to equipment design of the principle of organismic operation may be designated *Bionics*. This does not necessarily mean that we should attempt to synthesize a complete human nervous system but rather that we wish to design into our control systems certain of the desirable features of the nervous system.

It is of interest to note here a singular distinction of the human or protohuman mentality (shared, though in lesser degree, by the anthropoid apes) in the capability for improvisation. Improvisation is the logical consequence of those attributes of human thought which are frequently

referred to as versatility, flexibility, ingenuity and, for that matter, a host of other characteristics. In the instance of Kohler's apes, elementary forms of problem-solving were evidenced; apes stacked boxes to reach a desired object at a level beyond the jumping reach of the ape, and they made use of a long pole to secure a food object considerably beyond reach. It is indeed paradoxical and, to some extent, premature that such complex mental processes, which are incompletely understood, have provided the stimulus for new theoretical approaches to adaptive control, heuristic programming, and related ideas.

The distinctiveness of human performance in the solution of complex problems (as yet unapproachable via the conventional instruments of current computer technology) appears to rest upon several specific characteristics.

Of importance among the nonstructural features that set man apart from the thousands of other contemporary species of creatures are his verbalizing capabilities. It must be realized that these capabilities are not the result of distinctive anatomico-physiological endowments of speech. Man possesses an associative function enabling him to relate a verbalized sign and an object referenced. This trait appears to be lacking among infrahuman forms. Although we speak of "languages" existing among lower forms, as for instance, the drumming of the beaver, or the elaborate premating exhibitions of various forms, e.g., that of the peacock, these are not true languages any more than are those gurglings that human infants use to express themselves. Language, though certainly arbitrary in form, is exclusively associated with human culture. Exceptionally, it appears possible that porpoises may communicate in true language.

There has been much argument about the symbolism employed in human thought processes. Some authors state that all thought is verbalized, i.e., we think in terms of words and phrases. Others believe that we employ some means of mental shorthand which transcends the ordinary definitions of language. Such arguments do not matter if we consider the functional use of language as a medium for communication of ideas or the transfer of information.

It seems reasonable to assume (as the Whorf-Sapir hypothesis does) that our knowledge of the "real" world about us is conditioned by, if not restricted to, the constructions of language which convention, within a language group, attaches to given phenomena observable through the sensory modes of man. As yet no means, experimental or otherwise, have been devised to refute or verify the hypothesis that our interpretation of

"real" world events depends upon language structure as well as the pragmatic relationships existing therein.

Linguistic considerations are clearly important when giving machines a "thinking" capability is considered. If success in mastering the technical complexities can be achieved, in what language or in what system of notation shall our "thinking" machine think? "Adaptive control," "heuristic programming," "self-organizing systems," all have as their conceptual basis the assumption that it is possible to duplicate the processes distinguishing the human receptor-brain-effector linkage by devised analogous structure and circuitry without comprehension of the basic psychophysiological phenomena associated with human mental activity. Generally, the ultimate purpose of artificial "learning" or adaptive systems is twofold: the system must sense changes inherent to the controlling (computation)—controlled system as well as extrasystemic changes, i.e., those occasioned by alterations of the environment wherein the "controlled" system operates. We are again forcibly reminded of the fact that there is as yet no decision theory of sufficient adequacy to describe satisfactorily in terms of sufficient generality the conditions of even a static decision environment. The problems associated with decision processes in a dynamic environment are certainly several orders of magnitude greater in complexity, and we have forged only the most primitive theoretical tools to approach these.

Theoretical oversimplification of the effects of temporal and spatial summation with regard to information flow in neural networks has resulted in inadequate cybernetic models. In a very restricted sense, models of this class produce output information for an assemblage of input stimuli equivalent to the output of the biological system represented. The entire question of the means and avenues by which the organism utilizes and processes this information is generally avoided in cybernetic model development.

The cybernetic machine is an amazing example of redundant machine fabrication. A system upon which the requirements of automatic self-examination and self-repair are imposed must be capable of operation (albeit, degraded operation) despite the presence of components exhibiting transiently refractory behavior within its structure. Redundancy is mandatory in such a system, since the failed components may fall within diagnostic and repair portions of the system.

We shall now consider what some of the advantages of natural cyber-

netic systems are and how people have tried to devise analogous advantages.

A. Primary Advantageous Systems Characteristics and Previous Attempts to Implement Them

(a) Fabrication-in-the-Large. Here we refer to those substances of natural or synthetic origin which retain their distinctive physicochemical properties independently of quantity. This should enable one, at least conceptually, to "grow" equipment. Certain chemical memory devices approach satisfaction of this criterion but are inadequate to date because of slow response characteristics. They are further restricted by unsatisfactory environmental operating ranges. Perforated ferrite plates require auxiliary winding. Thin-film processes present the unresolved technical problem of controlling uniformity of deposition and, as a result, identity characteristics. The technique of beam-spraying, though not yet fully evaluated, appears to be a distinct possibility for fabrication-in-the-large.

(b) Statistically Enhanced Response. The motive here is to avoid dependence upon a particular component. Current approaches employ low-level redundancy techniques at the cost of size and weight. Major operational difficulties stem from uncertainty as to trouble source. Characteristically, elements adjacent to primarily affected components continue to operate (albeit, marginally). Hence, autodiagnostic processes are precluded during operation. Distributed systems employing the concept of superposition of response involve another kind of redundancy, which may be described as a smoothing process removing or nullifying noise.

(c) Self-test, or Self-observation. In the simplest sense, the operational problem here is that of determining which machine element(s) has immediate and/or delayed effects upon which other element(s). Machine design fixes the number of states wherein such effects are noticeable. The major problems associated with rapid and precise definition of the temporal and spatial origin and extent of unacceptable discontinuities and their dispersion among these states remain inadequately solved.

To date, efforts in the area of mechanized self-test include supplementary marginal programming. This entails the continuous use of subordinate and discrete machine elements and assemblages at the cost of increased programming effort. Other techniques rely upon built-in reiterative computational routines.

(d) SELF-REPAIR. Theoretically, self-repair functions are realizable in large inanimate systems. Unfortunately, technical difficulties obstruct the way to efficient, practical solutions of problems in this area. The use of "avoidance" or redundant paths by circuitry or analogous means compounds machine complexity and, therefore, within present design limitations, decreases reliability factors. It must also be realized, of course, that self-repair and self-diagnosis are not independent. Furthermore, actual self-repair implies some type of redundancy in the self-diagnostic section, which is in itself subject to component failures.

B. REQUIREMENTS AND COMPLEXITY

As automatic information processing and control systems become more complex by virtue of ever-increasing needs for capability (speed, capacity, and the simultaneous and concurrent control of many inputs and outputs) and compactness (size, weight, and geometry), reliability and operability must of necessity decrease if present standard philosophies of mechanization are utilized with standard componentry. Historically, Complexity has been unable to escape its unwanted companion, Unreliability. Technical demands for increased complexity are unrelenting and, within present mechanization philosophies, appear to admit of no compromise. Therefore, unconventional approaches are needed to develop complex but reliable equipment.

C. SATURATION OF THE STATE-OF-ART

For the most part, studies which have been performed to date in cybernetic systems have dealt with systems which were modeled to some extent after biological systems, but which generally were postulated to have far simpler and more uniform properties; that is, many complexities of operation and response have not been considered or, if considered, have remained implicit. Most such models have followed rules inspired not by the cybernetic system itself, but rather by the past development of mathematical logics and by present state-of-the-art in digital techniques.

If success is to be possible, the approach employed must begin with a study of the biological neural networks themselves. No properties of these networks revealed by former and current neuro-physiological investigations should be omitted from a comprehensive analysis. Both the microscopic properties of individual fibers and the macroscopic properties of

networks should be treated, the latter being derivable from the former by relations analogous to those relating quantum statistics to electromagnetic theory or statistical mechanics to thermodynamics.

Items to be specially considered include:

(a) COMPLEXITY OF FABRICATION. Present methods of fabrication rely upon the use of assemblages of dissimilar materials having distinct physical properties in an effort to duplicate the working of principles inherent in biological systems or envisioned in the abstract. Materials having appropriate characteristics for a given function are assembled in accordance with design specifications to produce machines of greater or less applicability to given classes of problems. The process of manufacture of such compounds requires special treatment. Present processes concerned with the production of a wide variety of discrete parts, their subassembly and inclusion as functioning portions of computing and control devices require specialized time-consuming fabricatory processes and an amount of subsequent treatment and manipulation which might stem from a technically primitive theoretical organization.

Redundancy principles provide an operational means to bridge the gap between materials presently available and ultimately desirable cybernetic machine design. However, redundant machines generally require larger numbers of subcomponents than do their non-redundant counterparts. These remarks indicate that redundant systems should be capable of condensation by utilizing subcomponentry, which is more efficient in energy conversion, is quickly and economically fabricated from materials present or producible in the system, and possesses the capability of being automatically incorporated within the system.

(b) RELIABILITY AND OPERABILITY. Significant advances must be achieved in reduction in cost, size, and failure probabilities of hardware, before we can achieve the physical expression of intricate arrays presently implied by theory.

(c) FLEXIBILITY AND ADAPTATION. The design constraints imposed by the factors of speed, accuracy, and power requirements vary with machine function and purpose. Therefore, a tendency has developed within the computer systems culture to produce a large variety of special-purpose machines in response to rapidly expanding operational needs. Effort devoted toward the development of such specialized devices customarily includes the adaptation of generally accepted techniques to a special purpose and hence does not advance the creative level of the art. Modifica-

tion and synthesis fulfill the conditions of operational expediency—the needs of the moment.

3. Comments on Biological Systems

The constellation of physiological qualities which characterizes living matter is partially known. Gross behavioral attributes of automaticity, irritability, metabolic and reproductive activity are displayed by living materials at all levels of cellular organization. The principles underlying these dynamic processes and their quantitative measurement, however, are poorly understood at present. Hence, the development of adequate models and consequent conceptual progress are substantially retarded.

An important objective in this connection is the clarification of the principles and techniques underlying, among other life activities, those of fabrication-in-the-large—a property held in common by all self-contained or "information-tight" biological systems. This quality is exemplified in processes associated with reproduction, as well as those of physiological repair and partial or complete regeneration. These processes of repair and regeneration are thought to be initiated in response to homeostatic organismic (or cellular) demands as expressed in terms of metabolic gradients, i.e., arising from differences in physiological function (a field concept) rather than from differences in anatomical structure. Thus, the organism exhibits the processes of self-test, self-tending, self-repair, and (dependent upon the complexity of cellular organization) partial or complete regeneration of segments as a fact-in-being.

A cybernetic approach in depth to the understanding and application of activity-maintaining principles in biological systems appears potentially capable of suggesting parallelisms between such systems and their artifactual analogs. This approach may in turn lead to advances in logics and organization philosophy.

4. Summary

We have considered in this subsection two special problem areas in control systems. These are:

(a) Proper use of feedback,
(b) Carry-over to designed systems of the desirable properties of biological control systems.

4

Operational Approach

"For they can conquer who believe they can."

—VIRGIL

Section 4.1

System Management

"He is unfit to manage public matters, who knows not how to rule at home his household."

—HENRY FORD

1. The Systems Manager

With the advent of more and more complex weapons systems, the fact has become clear that very few companies contain all the talents and facilities necessary to design, fabricate, and support an entire weapons system. The fact that the military agency responsible for procurement is attempting to obtain such a system from industry proves that it feels itself even less suited to do the job. Such considerations, together with our society's tendency to be haunted by the specter of "big business" monopoly while accommodating itself to work with oligopoly, have been responsible for the concept formation of the "System Management Contract." The developed concept is certainly a logical one and should operate admirably if the common pitfalls are avoided. In essence, one company is envisioned as taking responsibility for the entire weapons system and

subcontracting to other companies, as members of the system team, subsystems falling within their primary areas of competency.

2. Direction of Procurement

Direction of a system procurement will generally be shared by the eventual user (Operations), the buyer (Advisory Committee), and the contractual coordinator and developer (System Management). This relationship is indicated in Fig. 4.1. The various elements of this triumvirate fulfill their roles (see Fig. 4.2), these being to a degree dependent on both the system in question and their own organizational philosophies. A smooth and expedited flow of information, cooperation and agreement on procedures must form the basis of this relationship. In particular, to obtain maximum team performance the

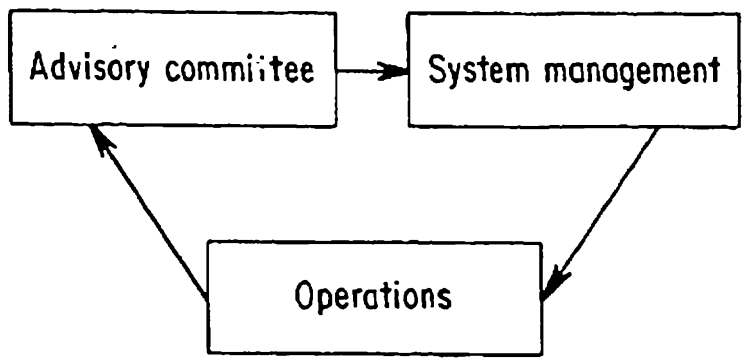


Fig. 4.1. Directive triumvirate.

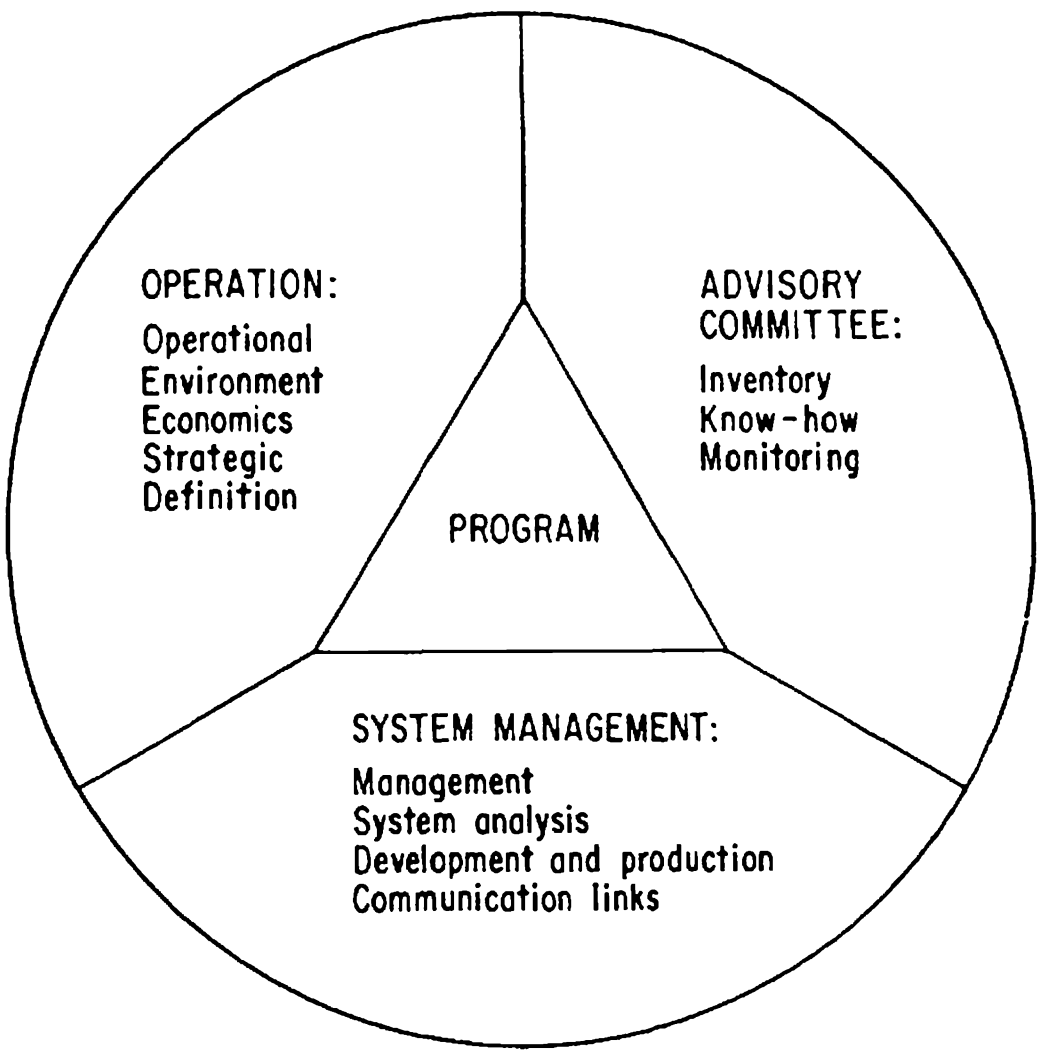


Fig. 4.2. Control of program effort.

Advisory Committee must behave in a permissive manner, allowing certain responsibilities, not classically so assigned, to devolve in part onto System Management. Examples of such responsibilities are:

(a) Initiation of procedural proposals to eliminate certain delays in guidance and approvals in standard from Advisory Committee
(b) Partial assumption of review, testing, and standard specification responsibilities
(c) Initiation of flows of relevant information to Advisory Committee

Thus, it is clear that a high level of organizational capability as well as technical prowess should be expected from System Management. The practice of management working in full cooperation with many others (Systems Management) is one that must be better learned. Successful past performance, and a willingness to learn from experience and incorporate necessary change into schemes are of great help, but designing the right organization is the requisite for achieving technical competence. Primary goals of the "right" organization are:

(d) Creation of operationally effective weapon systems
(e) Provision of such systems at minimum cost within a delivery schedule
(f) Facilitation of long range development of competent technical and administrative personnel

These goals must actually dominate the traditional short-term profit motif if successful System Management is to be established.

Proper execution of a System Management program requires the integration of all aspects of design, development, production, operational support, and liaison. In short, the problem cannot be attacked piecemeal; it must be attacked as a whole.

The talents and knowledge brought to the program by the Directive Triumvirate may be partially schematized as in Fig. 4.2.

3. Information Flows

In the preceding item (c) we emphasized initiation of information flows by System Management. These are really of two kinds. One kind, that which maintains state-of-arts information for the program effort, may be handled as in Fig. 4.3. The second kind of flow consists of com-

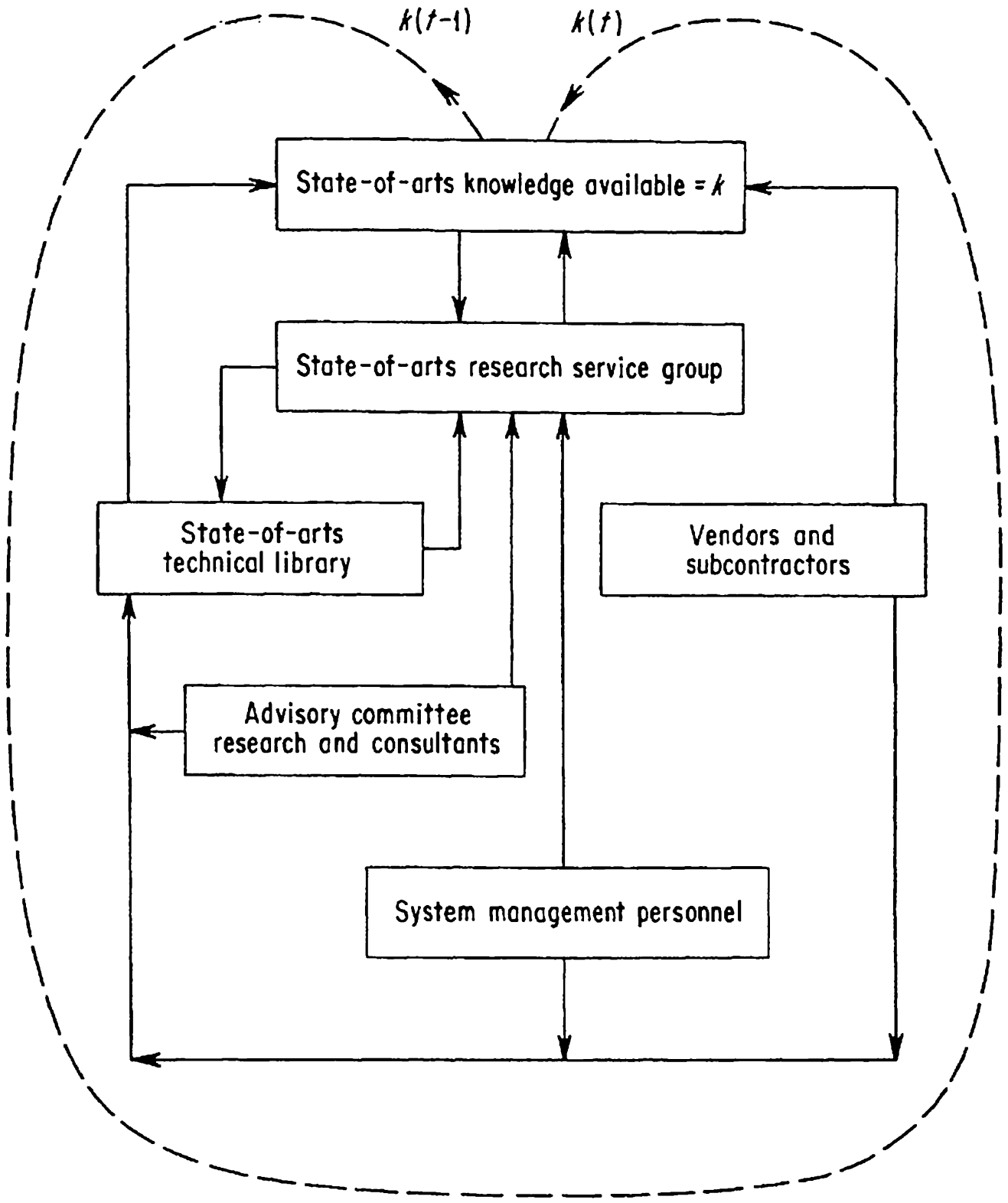


Fig. 4.3. Up-dating state-of-the-arts knowledge.

munication (status reports, progress reports, and engineering changes) necessary to carry out the program. Blockages in the communication system, a most serious problem in System Management, may occur in any one of three steps in the communication process:

(a) *Initiation.* Those who have information may fail to tell others who need it.
(b) *Transmission.* Those who should transmit the information may not do so.
(c) *Reception.* Those who receive the information may not be able to assimilate it.

The dissemination and clear understanding of pertinent data, objectives, and reports of problems encountered and progress made by all

groups taking part in a weapons system program is critically important to its success. Establishment of informal, as well as formal, rapport between System Management and the subcontractors is as much a management function of such a program as is the organization of technical and support manpower, and of physical and financial resources. There is a direct correlation between the technical quality of a weapons system and the degree of understanding and liaison established among its members. Such understanding, however, certainly does not imply the approval of extreme wastage, as for example, the matching of man-for-man by the sub-contractor and the System Management, or by the System Management and the Advisory Committee. Cooperation among members should be based on the true understanding of the problem involved. The correlation noted implies the need for communication on levels other than the formal and legal requirements of customary contractual agreements. Where informal relationships are "good," technical quality is usually high, because mutual understanding of the problems involved tends to prevent misdirected effort, attempted assignment of blame, and, as is typically the case, the resultant obscuration of malperformance behind legalistic language and voluminous but non-pertinent data. The same principles which determine good structure and control within any organization apply with equal force here. These principles must be maintained in order to keep the program within control limits and prevent destructive fragmentation of the effort into unrelated investigations and technical *cul-de-sacs*.

4. System Management/Advisory Committee Cooperation

The problem of "teaming" the capabilities of System Management with those resident in the Advisory Committee is a matter of minimizing the extent of review imposed by each successively higher level of organization. Great advantages can be realized in this area, especially where "performance" specifications are adapted, specifically:

(a) Collateral functions (i.e., groups of counterpart personnel), working together, should be permitted to establish realistic specifications that are specifically appropriate for each subsystem or component without detail review by higher authority. Such a course of action implies that the executive personnel have confidence in the technical competence of their operating-level per-

sonnel, and of course, such competence is a requisite for organizational success. Review, then, should be made on the basis of performance rather than conformance to form.

(b) Agreement of collateral functions should be formal on test parameters and methodology, but informal on conduct of tests, and form and quantity of test data. This puts observation of tests on a permissive rather than a compulsory basis and should raise the confidence level of test data since methodology will have received maximum technical consideration. No amount of data will remove doubt in the absence of confidence in test methodology. These two principles form the basis for a "permissive atmosphere," reflecting the modern philosophy of "management by exception." They maximize the utilization of specialists and elicit sound judgment and perspective. They decrease the probability of designing and testing to non-operational conditions and thereby avoid heavy economic loss.

5. Schedules

Schedules provide a means for planning, controlling, and reviewing progress of a program. In a broader sense they are a prime element in the organizational and managerial philosophy of any effort designed to accomplish tangible goals. They reveal the interdependencies of all elements of the total effort, and the relationships of the parts to the whole. They establish the magnitude of the task to be accomplished by the organization. Detail schedules indicate the time loci and time ranges of critical decision points affecting development of equipment, engineering changes, logistics, and economics. These points must be known with some precision by System Management in order to assure meaningful communication, upward to the Advisory Committee and outward to sub-contractors, so that intelligent and positive action can be taken in each area of responsibility. There is a fundamental limitation on the degree of precision that can be expected, because scheduling is always accomplished under conditions of uncertainty. It would be a mistake to assume that the degree of uncertainty is a function only of the state of particular technologies encompassed by the program. A great many other influences are involved. The fact that a program does or does not depend upon major technological breakthroughs does not assure accomplishment of planned schedules. (The fact may, however, reduce the dollar cost.) For any program there

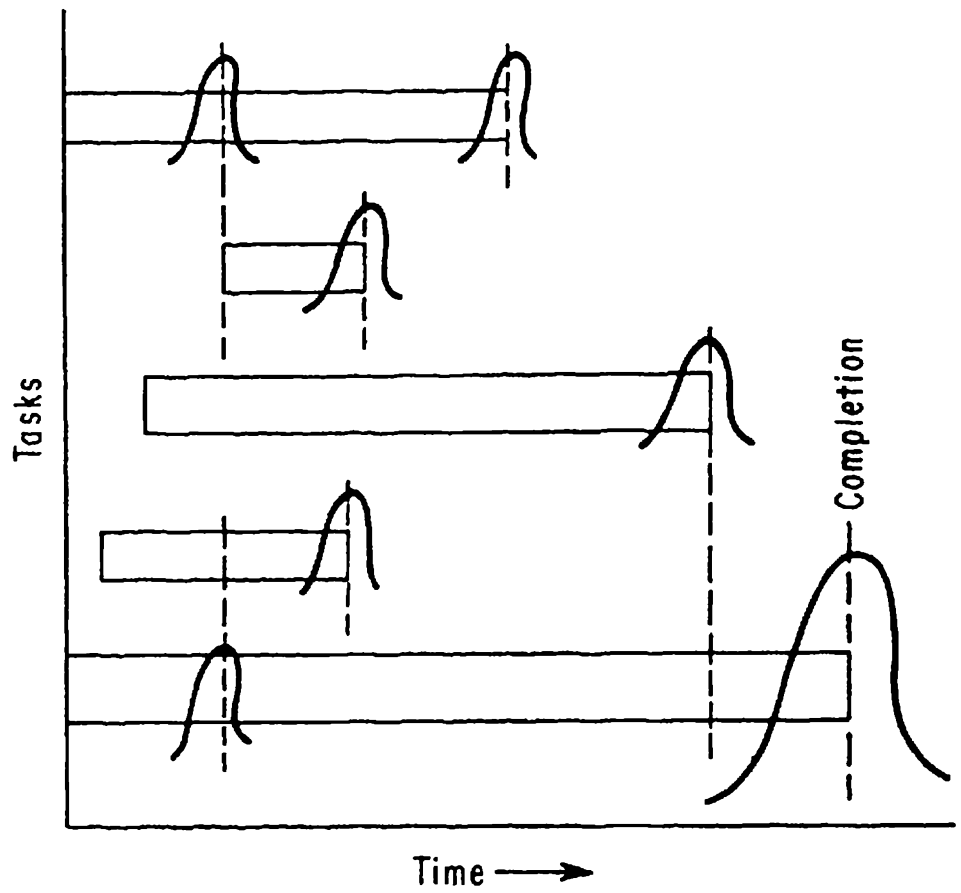


Fig. 4.4. Generalization of scheduling probabilities.

is a family of probabilities concerning schedules. These probabilities are not additive; some are dependent and some are independent. Certain principles derive from this concept of scheduling. Chief among them are:

(a) The level of reliability of detail estimates and forecasts is relatively low.
(b) Because of (a), detail forecasts of operations should be made to cover relatively short time periods.
(c) Over-all scheduling should be based on non-sensitivity of detail schedules.
(d) "Slack" is not only inevitable, but necessary, if any solution is feasible at all. It is otherwise impossible to purge unfeasible elements from the control system.
(e) "Slack" can be used to advantage in adjusting short term schedules, and in controlling the application of resources by diversion in accordance with the magnitude of the "problem."
(f) Such "slack" as exists within the System Management total organization can be applied to advantage in attenuating oscillations in the program.

It is essential to assure that important relationships are not overlooked, for such an oversight would induce series of chain-reaction crises. The magnitude of a weapons system program and the available time will generally require use of tabulating and computing equipment to maintain

timely surveillance. However, careful control of this data as generated is essential, or the surveillance will introduce, instead of prevent, crisis and confusion. Systems responsibility will permit System Management to take advantage of integrated, scientific data processing, and thus materially reduce the time lost as a result of conventional communication and decision methods.

The skilled use of modern computational techniques to accomplish the coordination of information, lead times, schedules of requirements, and procurement of requirements makes it possible for the System Management to initiate corrective action at the rate required to maintain optimum coordinated progress.

By proper application of incremental adjustments to a master schedule, the consequence of the adjustments on total costs can be quickly available. Furthermore, the interrelationship schedules permit rapid interrogation for determining consequences of conjectural changes on the total technical and organizational system.

The danger of producing over-control when attempting to adjust minor anomalies should be recognized. The necessary damping without loss of flexibility is obtained by maintaining available alternative adjustments at critical points in the schedule and taking full advantage of the "slack" that is necessarily present in all feasible schedules.

In any event, the overriding criterion is maintenance, by continuous re-estimation of requirements, of "always feasible schedules" for meeting delivery dates as specified by the Advisory Committee. The important consideration is assurance that reasonable departures from forecasts cannot seriously disrupt the major schedule.

Until sufficient detail specifications have been derived to permit responsible estimation of the magnitude of efforts to produce, test, and sub-contract, it is more economical, as well as more certain, merely to assure existence of feasible alternatives to accomplish programming for cost estimation.

The following are basic remarks on scheduling and also on costs:

(a) "Minimum cost" is equivalent to "maximum efficiency" within the schedule.†

(b) Maximum efficiency is achieved if the schedule minimizes the maximum magnitude (rate) or effort (expenditure—dollars, manpower, materials), in *all* categories.†

† A. Charnes, W. W. Cooper, B. Mellon, "A Model for Optimizing Production by Reference to Cost Surrogates," *Econometrica*, XXIII No. 3 (July, 1955), 307–23.

(c) Two major factors define the over-all constraints on any program:
   1. Time available
   2. Economic capability (including availability of appropriate manpower)

Almost any program can be "accomplished" if limits are not set on these two factors. All practical schedules are planned on the basis of rather definite limits on *both*. The System Management always makes the fundamental assumption that time is a prime system parameter. The question of whether any schedule proposed by the procuring agency is realistic or not or whether it may be "too tight" then becomes a question of economics. The question to be asked is, "How much work can be accomplished in t time?"; not, "How much work can be accomplished with n dollars?" The quantitative answer to the first question will determine what n dollars are. The procuring agency must then determine whether n dollars are within its range of capability for the particular procurement. This range of capability is primarily an economic consideration for which no real formula exists. It is beyond the realm of System Management's decision-making ability to assess the strategic situation in which the weapons system finds its operational use, for it cannot make the appropriate politico-economic evaluations necessary.

6. Delegation and Communication

Throughout the text of this subsection great stress has been placed on the significance of communications. The basis of all proposed organizations, functions, and relationships is the need for positive communication, for gross deficiencies in this respect can negate the best of technical approaches. The design and interposition of the most appropriate communication links within each segment of System Management's internal and external organization, and between System Management and the Advisory Committee, are basic System Management responsibilities. It is interesting to note that, according to reports available, the military's experience with the systems manager approach has been "uneven." It is not surprising that the method has worked well when the major aspects of a program (i.e., operational characteristics, conceptual grasp of the problem, and interjection of current military point of view) have been involved; the method, however, has proved less than satisfactory when it has been used as a device to control more detailed considerations (i.e.,

logistics and reliability). The problem of coordinating phases of a systems program is one that all managements face.

There is a considerable divergence to be noted between the conception of policy, and its detailed communication, interpretation, and implementation throughout the organization. Top management cannot reasonably dictate the specific manner in which "everything" is to be done. One of the most important aspects of management skill, therefore, is the manner by which the management creates an atmosphere in which communication flows progressively upward and downward so that there is an awareness of problems at the appropriate levels of organization. Without this atmosphere there can be no real perspective about the relative importance of each task, and about the relationships of tasks to each other and to the major objective. To the extent that communication is ineffectual, there will be confusion and imbalance of efforts at progressively lower levels of the total program. As a corollary principle it can be stated with confidence that the number and intensity of miscommunication items increase as the extent of detail in the task level increases. This principle is one possible explanation of noted difficulties where details of specifications, qualifications, testing and evaluation, logistics, and interchangeability are concerned.

At the "management" level, on the other hand, agreements can usually be more easily reached, because there are fewer people and fewer details involved, and only the major aspects of the problem are dealt with. The emphasis at this level should be placed on resolution of conflicts of interest and on setting direction. At successive working levels the emphasis is obviously on execution—on the skill and ingenuity of responsible parties to find the best ways of accomplishing stated objectives, and to transmit knowledge of the existence of unfeasible aspects of the effort upward. One requirement of a System Management philosophy is that detail problems can usually be resolved at the level of effectivity of the particular organizational unit. The attention of responsible parties at successively higher levels is therefore seldom diverted to problems which are properly functions of lower levels.

7. Commercial Systems

We have now considered at some length the System Management concept as it involves weapons systems. This limitation (i.e., weapons systems context) is purely an artificial one induced by the fact that com-

mercial systems have lagged behind weapons systems in the degree of complexity procurable and the urgency of requirements. In the commercial case, it is reasonable to anticipate that the System Management concept will be adopted as complexity increases. The only major forseeable alterations are evolution of portions of the definitive and informational responsibilities of operations upon the System Management and absorption of the essential remainder of operations into the Advisory Committee with the result that the Directive Triumvirate would be reduced to two parties.

8. Summary

We have discussed the way weapons systems procurements are (or at least should be) handled. The general techniques can be directly applied to commercial systems. Primary points to remember are that scheduling "slack" is to some degree inevitable and that the optimal use of "slack" is of the utmost importance.

Section 4.2

Systems Support Functions

"The actions of men have two effects—primary and secondary. Often the secondary effect is of more importance than that of the primary."

—ELBERT HUBBARD

1. Support Functions

Under the heading of support functions are generally gathered (among other items):

(a) Field service
(b) Training and training aids
(c) Specification preparation
(d) Handbook and manual preparation
(e) Checkout equipment
(f) Spares and provisioning

(g) Drawing maintenance
(h) Auxiliary equipment
(i) Engineering service to customer
(j) Reliability programs

2. Quality Control

Quality control of one's own production and *acceptance testing* of subcontractor's deliveries are, of course, merely opposite sides of the same coin. Much has been written concerning applicable procedures for these functions, and we need only remark that control of the entire *reliability program* and selection of procedures should lie in the hands of one well-versed in fundamental statistics and its operational applications, as well as in the context of the subsystems and components involved. The reliability program must, of course, be also linked from the beginning of the program to the design effort so that reliability is "designed in" to the maximum feasible.

World War II gave substantial impetus to research in techniques for the selection of personnel for certain classes of tasks; it helped initiate large scale efforts directed toward the development of means for the assessment of *training*. However, during the World War II period and since, a relatively small amount of investigation was directed toward the more fundamental area of training techniques as such. Clearly, with the increasing complexity in hybrid systems, the problem of adequate training provision are accentuated.

Psycho-physiological information concerning man continues to accumulate rapidly. In systems areas the need for synthesis of this information by other than simple incremental process assumes frightful proportions when viewed in the light of information available against what must be known. Thus, as an example, although we believe our species to be poised upon the threshold of space-entry, we have not successfully achieved to date, within a terrestrial setting, a space-equivalent environment. To be sure, certain specialized aspects have received much attention, as for example, brief periods of weightlessness, human occupancy of near-vacuums, and "g" effects through centrifugation and flight. Many of the training-associated problems caused by our entry into new technological areas will be those evoked by the design and guided evolution of synthetic devices constructed to simulate extra-terrestrial environments. Other less spectacular classes of simulation equipments will be those substituting for

more expensive or scarce equipments, or those which will provide more efficient means for the presentation of principles and operational procedures than would otherwise be available. Whatever the purpose served by the training device, or in an extended context, the training program in full, within the advancing technology, the ultimate criterion for the measurement of success will be individual operational proficiency. Unfortunately, the considerable amount of work done on the development of techniques for measuring and predicting the transfer value of the effects of training has been restricted to classes of hybrid systems having long-established developmental histories. Such is not the case with respect to the non-terrestrial human occupancy systems presently envisioned. There is no "space-folklore" except for that manufactured before the fact by imaginative writers. Thus, it is unlikely that in time to come any space-trip navigator will be able to refer to his performance as "flying by the seat of his pants." Experience is the precursor of idiom.

OPERATIONAL CHECK-OUT. Whereas the System Manager would naturally regard production check-out as a part of the production process and whereas production check-out is customarily accomplished with one or a few sets of special bench equipment, operational check-out is likely to prove one of his major areas of concern. Unfortunately, this fact is frequently not realized in the earlier stages of a systems program where conceptual matching of the check-out system to the primary system or primary subsystems may easily be as large a program item as the primary item(s). Under varying programs the required number of check-out systems may be less than, equal to, or greater than the number of primary systems produced. Check-out systems may be partially or wholly integrated physically with primary systems or may be physically separate.

The basic extant concepts of check-out procedure encompass:

(a) Total check-out to try every possible failure mode of the primary system
(b) Carefully designed (in the sense of experimental design) sequences of checks to permit deduction with pre-assigned confidence level of which failure mode, if any, subsists
(c) Marginal checks to ascertain whether any of the failure modes regarded as highly probably subsist or whether the system is capable of performing in at least some "acceptable" degraded mode

One of the major design objectives in operational systems is provision of capability for performance in degraded modes rather than so-called

total failure. This really amounts to "building-in" back-up systems to enhance "fail-safe" operation of a supersystem.

Any of the check-out concepts can be rendered a multi-stage check-out by redefining failure modes. Thus, accompanying present day modular fabrication, a frequently prevailing philosophy is that of check-out in operational situations only to the module level, followed by check-out at leisure of faulty modules. "Throw-away" modules, of course, are not further checked.

We shall not discuss at this point the questions of preventive maintenance, and "continuous" *vs.* "discrete" checkout except to say that the trend in complex systems will no doubt be toward inclusion of more and more self-diagnostic features centered in the above-mentioned concept (b). Although the customer will determine the check-out concept and subsystem operational level, the check-out system generally will have some design advantage over the primary system in that its functional specification will be determined by the primary system design.

The *logistic concepts* (field service, spares, and provisioning) associated with a complex system may well warrant a large-scale effort in themselves. The primary danger is application in a hackneyed way of tradition, standard operating procedure, and protocol. There is, for example, little future in designing clever control procedures that release nuclear cores to bombers if their application implies an interval longer than that required for bombing-up between red alert and strike on the airfield concerned; Mobility of the missile system aimed at lessening vulnerability is of little value if it significantly lessens operational capability.

Finally, emphasis should be placed on the orderly maintenance of all *documentary material* associated with a system program, for there are frequently literally thousands of drawings and blueprints alone associated with a given subsystem. Confusion in handling such material is sure to produce a great waste of time, effort, and money.

3. Summary

We have spoken briefly of the primary support functions:

(a) Quality control and reliability program
(b) Training
(c) Operational checkout
(d) Logistic concepts
(e) Documentary material

5

Systems Technology Forecast

"If a few little efforts are made, here and there, to begin thinking about the range of possibilities, *there will be material to sketch out, as in a great chess game, some preliminary questions, so that better players than ourselves can ultimately develop a strategy."*

—GARDNER MURPHY

Section 5.1

The Next Era—Major Areas

"If you want to get somewhere else, you must run at least twice as fast as that."
—LEWIS CARROLL

1. Predictory Assumptions

In the 1960–1980 era we may anticipate three major politico-economic trends which will directly affect requirements in both types and numbers for devised and hybrid systems. These are:

(a) Reinforcement of socialistic tendencies will result in increased widespread demands for greater organization-member interaction; further uniformization of a populace that possesses more conveniences and leisure time; greater stabilization of the politico-economic environment in order that a status quo might be firmly established; and increasing dominance over organizational control functions by centralization agencies.

(b) By general agreement, established commitment to space exploration and the extreme costs of uniform or, more probably, accelerated progress will soon lead to demands for concomitant space exploitation.

(c) The now existing demands for scientific improvements to be incorporated into our terrestrial environment may be expected to become greater. Demands will be made for direct benefits, such as weather control and agricultural applications of nuclear techniques, and also for regional technological advances to lessen the threat posed to the "haves" by the "have-nots." These types of demands are, of course, closely linked to the heightened socialistic impetus mentioned in item (a).

Although it would be most fascinating to attempt a detailed prediction of the socio-politico-economic environment of 1980 based on these trends, we shall leave this to more Orwellian sources.

Extrapolative speculation is quite unsatisfactory, particularly in a global context where the players, though occupying the same playing-field, subsidize their own referees and play by rules which are mutually unintelligible. It does appear reasonable to believe, however, that the severe socio-economic derangements resultant from long-continued maintenance of an alert national military posture would enhance significantly the conditions described in (a) and (c) above. In these circumstances space exploration efforts, item (b), might be ultimately curtailed drastically, or, conversely, expanded militarily. The immediate sequelae of a total conflict would, no doubt, constitute psychological and economic hardships that would derange cultural and technological structures and institutions to the point where radical social and economic adjustments would absorb all national energies and wealth for some time. If a global conflict were to happen, it is likely that considerably more primitive hypotheses than those implied by items (a–c) would become of primary importance.

The politico-economic structure we envision during the period of 1965–1980 will be one in which trends now present will, by extension, be technologically embellished. These trends are becoming increasingly explicit to the critical observer. Attention will be focused in this section only on those future eventualities that will affect the development of systems requirements.

2. Basic Automation

The first step in automation in almost any context is mechanization of simple iterative procedures. These include operation of machine tools, monitoring and adjustment of parameters in well-defined material/energy flow processes, and completely specified accounting procedures. Although some mechanization of this nature may utilize directly the capabilities of a centralized information system in more sophisticated installations, the more frequent case will be that of specific decentralized mechanization. The level of procurement of such mechanization may be decided upon classical economic justification grounds by the procuring entity, and, in the U.S., these entities will generally be private companies. It is more probable that the relative impossibility of adopting any other course of action—the only other remotely possible course being that of conceding to the financial and procedural demands of the labor force to find an alternative to automation—will soon force such procurement, provided that reasonably successful equipment be available, whose cost will not preclude the economic survival of the procuring entity.

Reasonably successful equipment, in this context, denotes those equipments which have a serviceability life contingent upon the automated processes they support and are satisfactory temporarily, in providing continuity or physical linkages between discrete processes necessary for the control of multiple operations. Such equipments are in a sense exterior to the controlled/controlling processes involved and as such require special treatment for their (transiently) successful employment in achieving the degree of automation compatible with the product flow or fabricative processes involved as of the technological moment. It is quite probable that at the transient level eliciting basic automations, as examples of specific decentralized mechanization increase and become more generally established throughout affected industries, competitive positions will depend upon management sensitiveness to basic automation requirements, as well as upon variable rates of accomplishment in rendering commercially available automating equipments compatible within a given manufacturing-system complex.

On the other hand, there will be those companies who will participate in highly competitive markets in which little slack remains in profit margins. Here, economically justifiable procurement of mechanization

by one competitor will serve as a motivation for all other competitors to precipitate similar action to make economic survival practicable for them.

It is only those firms in a monopolistic or quite stable oligopolistic situation who can offer resistance to both of the motivational pressures so far mentioned. These in turn are subject to pressure to mechanize if only through anticipation of threat to their admirable positions by more progressive newcomers, particularly those newcomers arising as spin-offs from firms seeking expansion in related areas.

Another problem paralleling the general establishment of decentralized mechanization within classes of industry is that of enlarged central control responsibilities. When manufacturing operations consist of truly iterative operations of one type throughout (a situation becoming increasingly rare), enlarged central control responsibilities would merely entail increased capacity in the controlling mechanisms and an enlarged monitoring staff. The more commonly encountered industrial situation, however, is that of large numbers of discrete sequential operations of great diversity. In this situation, problems associated with smooth transitions to central control become most difficult. These matters are more fully discussed in the succeeding section, "Organizational Automation."

It appears a reasonable assumption that by 1980 essentially all functions at the basic level will have undergone at least experimental mechanization with various degrees of success. Since there are also definite indications that development costs will rise more rapidly than others in the near future, the systems producer will probably find the key to success in the basic area if he concentrates upon modular approaches to solve large classes of problems with small equipment variations and makes the necessary parallel applications studies. Although there is reason to believe that for several years demand will outstrip production and, thus, saturate all competent producers, the more progressive firm will nevertheless not only offer sound equipment but also precede its sales efforts with appropriate applications studies. The firm that adopts such a policy will no doubt be able to pick and choose among the more profitable areas of concentration.

3. Organizational Automation

The second level of mechanization to be considered is that of sequences of one-shot or iterative procedures. This level of mechanization is technologically contemporary with that of basic automation and generally constitutes a more or less compatible assemblage of its operations and

equipments characteristic. It is to be expressly noted that such applications do not necessarily connote technical improvements upon the methodologies and equipments common to basic automation procedures. Rather, the evolutionary change, if such it may be called, between basic automation and organizational automation is one of compoundment, one in which the basic processes and equipments remain essentially unchanged but are selectively multiplied to accomplish (at the level of efficiency permitted by the sensor-actuation mechanization state-of-the-art) the control processes required. It is to be emphasized, however, that initial over-all system design is prerequisite to successful organizational automation and that "kluging-up" of pieces as complexity increases is doomed to failure. Organizational automation generally includes such examples as operation of production lines, mechanization of a firm's entire accounting system, and scientific computations. It is primarily at this level of sophistication that the general purpose digital computer currently excels and holds further promise whether it works alone in certain areas or manages decentralized devices at the basic level in other areas. Although certain factors such as compactness (size, weight, and geometry), resistance to nuclear environmental conditions, and low power consumption will be of importance for systems implementing restricted classes of specialized missions, by far the greatest emphasis will be placed on componentry and fabrication techniques so as to permit the construction of reasonably reliable systems possessing progressively higher basic capabilities (speed, memory capacity, flexibility).

In this area there will be lessening continuation of present day system types but many direct logical extensions of these. Certain classes of space vehicles, for example, will require central control computers quite similar to those in present aircraft with the exception that functional modes will be expanded by generalization. Thus, for example, inertial/radio-fix navigation may become stellar supervised inertial, and fuel-management/cruise-control may become energy-management/trajectory-control. Further probable changes, primarily in external and internal environmental sensory techniques, may be exemplified by such transitions as ordinary Doppler to relativistic Doppler and magnetic deviation matching to gravitational anomaly matching.

For a manned space vehicle an entirely new class of internal environmental sensory problems is anticipated because of its occupancy by an essentially mammalian organism. Problems of environment-stabilization for the operator/occupants, as well as special protective means (elicited

subliminally) necessitated by the need for protection against hazards known to exist in extraterrestrial environments, differ from those encountered in present aircraft. Previously unencountered reference-realms will no doubt require drastic transformations of display-control complexes using enhancement and extension of present unburdening and quickening philosophies. In this area there will also be heightened pressure to develop general purpose computers with capabilities considered ultra-high today. The design objectives of very high speeds and very large memories reflect primarily a class of major problems mentioned previously in Chapter 3 such as social security accounting, realistic weather prediction, and near-real-time stimulation of large scale unagglomerated military-economic complexes quite beyond present machines, but still of a level of conceptual sophistication consistent with present general purpose applications. One further area deserving special consideration at this level is automatic checkout of systems. The extensive variety and mounting complexity of componentry and intracommunicative-directive circuitry, particularly in high performance vehicular and strategic systems, underscores the need for maximum reliability in operability status checkout in minimum time. While electromechanical checkout devices are not infallible, human fallibility is notorious in such iterative procedures. Ample literature attests that human-error frequency in stressed circumstances rises sharply (statistically). In effect this constitutes an implicit requirement for the monitoring of quality in accomplished human checkout functions by sensoric computational means—for an example of automating checking procedures to avoid human error, see Appendix II, "Potential Mechanization of Safety Procedures During In-Flight Emergencies."

4. Compromise Automation

Componentry perfect in regard to predicted behavior is not physically realizable even from a theoretical viewpoint chiefly because of the operation of quantum-mechanical uncertainty principles. At any level of the states-of-the-arts there is some upper bound on permissible system complexity, beyond which reliability/operability is not attainable. A second fundamental limitation upon automation is that the finiteness of entropy at any time implies a lack of complete information concerning the universe so that there will necessarily exist missions which are inherently incapable of *a priori* logical description and, hence, of total mechanization. (We are assuming tentatively that these basic physical "laws"

HUNT LIBRARY
CARNEGIE-MELLON UNIVERSITY

are perpetuable.) As an example illustrative of both classes of limitations upon prescience in operational circumstances, consider the construct of a stylized military aircraft mission during the in-flight phase. There are several major classes of uncertainties that typically influence the quality of mission performance. Significant among these are unpredictabilities concerning the exact geometry of a projected flight path inclusive of alterations necessitated by special attack modes, the dynamics of an inclusive tactical environment, and the successive reliability states of the vehicular system inclusive of operator/occupants as well as time dependencies of such states. The systemic effects of uncertainties of these classes (as well as more subtle effects) may occur independently or concomitantly; they may or may not interact; they may either singly or in concert be trivial or massive. The direct and indirect operational consequences of successive uncertainties arbitrarily ordered in time and other dependency domains define a flight management problem of considerable magnitude (though in different chronologies) for both designer and controller of the airborne system. In a manned-type system we might speculate upon the altered characteristics of the flight management problem at two levels of information availability. In the (presently unrealizable) situation of "completely reliable" predictive information *in extenso*, the aircraft operation could be pre-programmed throughout. In this sense "pre-programmed" refers to the state of information available to the aircraft controller as being "complete" in that command information would be available throughout the flight. A pre-programmed or command course might therefore be "read out" by the controller from display means using sophisticated techniques for the alteration of visual and other sensory stimuli necessary to elicit appropriate control responses. Since there is no *theoretical* difficulty in anticipating approximate aircraft attitudes throughout successive portions of the flight, the performance characteristics of the aircraft may be likewise (still in the conceptual domain) anticipated and displayed. Natural obstacles or (man-made) hostile objects may be made known, and avoidance information presented. Note that in this kind of mission the sole operator requirement for information is deviation from the plan; that is, it is necessary only to present those data which enable him to correct for deviations from command performance.

The pre-planned mission corresponds to well-defined and complete strategy. Unfortunately, such strategy is generally not available. This may be due primarily to the inadequacy of surveillance intelligence, meteorological and other types of environmental information, or it may

be due to the fact that the mission itself is of a tactical nature and is perhaps directed against targets of opportunity. Furthermore, the tactical environment in a hostile situation is apt to be dynamic, and it is often not clear to the launching source just what counter-measures may be directed toward that vehicle during the intermediate course to target or upon arrival.

The presently typical flight management problem is one in which certain segments of the flight path are specified as necessary for avoidance of known countermeasures, refueling rendezvous, and known obstacles. Thus, the system operator must of necessity elaborate dynamic planning when he is operating the aircraft so that he may join together, in a reasonable fashion, the pre-planned segments.

The demand upon the operator for dynamic planning in transit implies requirements for the presentation of information concerning aircraft attitude, geographico-temporal locations, and system-capabilities-remaining, as well as information concerning previously unsuspected enemy tactical countermeasures. The point of great significance to be noted here is that the requirements for system operational management, other than that automatically provided, always exists when the course of action cannot be specified in advance. The system designer's objective, then, in simplifying sensory/pilot system interfaces, and thus de-stressing the human subsystem, is that of supplying, in optimal fashion, comparative monitoring information capabilities under pre-planned conditions to the extent feasible, as well as absolute performance information under dynamically planned conditions.

It appears that as the state-of-the-art in sensory-computational areas advances we may anticipate large improvements in instrumented sensitivity and enhanced actuator means. Despite these advantages we shall, for a considerable time to come, no doubt be submissive in system planning to considerations of natural forces and processes.

Thus, basic limitations imply a non-terminating requirement for man-machine systems to implement some missions regardless of level of technological achievement. The obvious fact that we are nowhere near theoretical limits in either of the two specified directions merely underscores this requirement. In problems involving high complexity or inadequate *a priori* information, we will continue to utilize directly the capabilities of one or more humans.

Another area which springs immediately to mind in this regard is detailed exploration of extraterrestrial bodies either for scientific purposes or with a view to establishing the feasibility of economic or political

exploitation. While it will be possible for some purposes to instrument such exploration in a way that involves men only remotely, it is much more probable that we shall wish to place men directly into this environment with their protection and physical and sensory capabilities suitably enhanced by machine aids. Another area involving men directly, one which appeals to humane sentiments and one which will prove to be a sound business venture for qualified firms, is that in which our advanced mechanization techniques would be used to aid the branch of surgery that treats of prosthesis. The U.S.S.R. has allegedly demonstrated in this regard an advanced form of artificial limb which functions on much the same basis as a natural limb. While 1965–1970 developments in the prosthetic area will undoubtedly be directed toward improving the condition of those suffering natural or accidentally induced defects, there is another later potential direction which should not be ignored. Should space colonization be initiated in environments inimical to inherent human adaptive capabilities, it may be desirable from the viewpoint of the colonizers themselves to obtain a certain degree of physical adaptation by prosthetic means. Many consider space colonization the means to counter the pessimistic assumptions formulated by Malthus; for it must be admitted that, as of now, the Malthusian view is definitely supported by contemporary trends.

Generally speaking then, there will be definite requirements for man-machine systems in the era we are considering, but these must be of a higher order of integration of capabilities than those currently extant or in present planning.

The notion of further man-machine integration, perhaps even to the extent of physical integration, may be distasteful to some. Yet, consider Muller's comment:† "Man must eventually take his own fate into his own hands, biologically as well as otherwise· · · ."

5. Summary

In this section we have indicated in what way near-future automation will develop and why such development seems probable to us. We have noted how increased socialization, increased centralization, and greater demands for return on major centralized expenditures will be instrumental in hastening this change. Automation trends have also been partially classified according to the logical sophistication involved.

† H. J. Muller, *Scientific Monthly*, XXIX (1929), p. 481.

Section 5.2

The Next Era—Specific Comment

> *"The whole concert of industrial operations is to be taken as a machine process, made up of interlocking detail processes, rather than as a multiplicity of mechanical appliances each doing its particular work in severalty."*
>
> —THORSTEIN VEBLEN

1. Basic Automation

A standard illustration in this area is the high concentration of effort that has been placed in the field of automatic machine control. The early mechanization of six-degree-of-freedom automated milling machines is an established fact.

In order to succeed in this field, one must not start by duplicating the efforts that have been made, but must choose instead a research program

enabling one to be, at a predetermined date in the future, in a position to build more advanced systems than will be available through a more fixed route to the end objective. Machine tools in general involve a great deal more than the simple automatic operation of these devices. The following list is representative of the many details that could be automated in the context of basic automation, so preparing the way for a smoothed transition to organizational automation.

(a) Machine set-up
(b) Raw stock loading
(c) Machine tool operation
(d) Machined part cleaning and de-burring
(e) Machined part inspection
(f) Feedback from inspection to machine set-up correction
(g) Computation-sensitive adaptive machine tool which adjusts speeds and tools by determination of raw material characteristics
(h) Machine input sensors directly to utilize standardized drawings (making parts from drawings)

Each of these items mentioned are rich in spin-off potential results suited to all phases of computing and controls. For example, item (g) should result in a new philosophy and subsequent generation of adaptive processes suitable in the ultimate sense for devising computers that can universally modify their internal logical structure to suit operation within a tremendous variety of different machines. The effect of the inclusion of such operations would be to eliminate the need for modifications and extensions of existing designs by logicians and engineers by enabling the machine to encompass other classes of problems by self-adaptation. Input devices having sufficient flexibility so as to permit using them in conjunction with a great variety of simple pictorial representations would require that work be performed in areas such as pattern recognition and information theory. Development of these input devices would stimulate progress in the development of computing techniques similar to those employed by the perceptron.

Examination of the concept of basic automation reveals that, contrary to the belief of many individuals in the automation field, organizational automation is not a whole composed of basically automated devices as its parts. Not only is this concept misleading, but any company which pursues the doctrine of "do a piece now and add on later" is most assuredly going to bring into existence a system so subject to failure that practical economic use of the system will rapidly become an impossibility. What,

then, is the true purpose of basic automation? The answer to this question lies in the economic factors primarily associated with commercial manufacturing. It is felt that the artificiality of the pervasive military (controlled demand) market would not allow a fully developed concept of automation to prove its greatest competitive merit. A second major consideration is one of time. Basic automation is, by far, the simplest problem and will yield to solution at the earliest date. The interpretation of this aspect of basic automation must not rule out the specialized cases of basic automation in military equipment that are not primarily oriented toward the avoidance of fundamental economic penalties. Here, in viewing the over-all problem of military basic automation, an entirely different approach must be taken. However, there are undoubtedly some areas of overlapping study. Certainly reliability, even though the environments are different, must be a common factor in both economic basic automation as well as military basic automation. Another such region of common concern is found in areas of flexibility. For example, it is highly desirable that an automated machine in a factory be capable of performing many different types of operations resulting in different end products, and that the change from one product to another be made as readily as possible. In the case of military basic automation, a desire has been expressed for systems capable of taking on different functions as the mission for the military machine changes. An airplane, for example, may be requested to fly reconnaissance missions for some period of time and then might be called upon to do strategic bombing for a second period of time. The logistic advantage of having a flexible system able to accommodate all likely missions to be encountered by a particular aircraft is obviously one of great importance.

To sum up then, we are presently living in the era of military basic automation. Present designs and concepts are lacking in the essential areas of reliability, flexibility, and adaptability. The era of economic basic automation is just beginning. It should not be said that the large amount of work that will of necessity be devoted to developing a high degree of capability in the areas of basic automation will go to waste. Many of the techniques and tools for designing these equipments will be applicable in part to the much larger problem of organizational automation.

2. Organizational Automation

Military aspects of organizational automation present a set of problems, the full significance of which is beginning to make its effects

felt on the military. Should, for example, some Command reach a decision to automatize completely its warehouse and inventory system, there would be introduced operational and situational complexities that cannot be handled by any existent planning or scheduling techniques; i.e., organizational automation is deemed desirable, but no satisfactory method of determining whether a proposed solution is adequate is offered. Moreover, combinations of events may occur which could effectively tie the system into "knots," make the system become unstable and erratic and exhibit "neurotic tendencies" similar to those found in human beings; that is to say, the ability to perform ordering or ascertaining relative weighting in advance of the design and application of a system is impossible. Because of the many facets of control requiring feedback, modes of instability can most definitely appear. A hypothetical illustration of what could happen, assuming a very simple case, is presented next.

Suppose that a machine is being used to fill orders from a warehouse, maintain an inventory of stock on hand, and initiate requests for stock transfer from depots to the warehouse. Let us further constrain the system to have a finite capacity above which the system becomes saturated, that is, just so much stock can be kept on hand; any surplus must be returned to the depot. Let us further suppose that a finite transfer delay is encountered between requesting a shipment from the depot to the warehouse. Consider stock item "A." For a given period of time a constant rate of withdrawal is initiated from the warehouse by demand of the stock item from inventory. In anticipation of running out of stock, orders based upon the withdrawal rates are placed with the depot for shipments to be made to keep up with the rate of withdrawal. Just as the first of these shipments begin to arrive at the warehouse, the demand for stock part "A" temporarily ceases to exist. In a short period of time the warehouse is completely restocked. Over-shipments begin to arrive. These necessitate return shipments to the depot. Subsequent future demand for stock part "A" continually creates either the problem of being overstocked, which necessitates return shipments, or that of being understocked and running out.

The problem is obviously introduced by the transport delays between the depot and the warehouse, and the inability to be able to make exact predictions of the amount of any particular item that will be demanded at any particular time. Thus, a very undesirable condition, which is certainly costly, may result. In addition, if this depot supplying a large number of houses is automated, then its operations can also become highly unstable

because of the same causes. Solutions to these problems have been attempted by providing overflow buffering so that once a shipment has been made from the depot a part can be kept in temporary storage close to the warehouse. Studies to determine maximum rates of demand anticipated, in order to determine the size of the inventory, have been made. There is danger, however, in maintaining large inventories to meet high rates of demand; if the particular item kept in stock is suddenly discontinued, removal of this item from stock is required. This very obvious example can be compounded again and again to produce so many combinations of problems that a proper or adequate over-all design cannot currently be made.

If one looks for those industries which will undergo organizational automation first, one will conclude that they will be those industries which are already to a great extent operating in an automatic fashion. Among this class are the petroleum and refining industry, the chemical industry, the iron and steel industries, electric power generation and, of course, the telephone system, which to a very great extent is becoming organizationally automated. A study of automation procedures in the telephone system serves to point out some of the problems that must be avoided in future organizational automation programs. The telephone industry has grown over a period of many years by a process of addition and modification. This industry is operating and making a profit perhaps because of its economic structure rather than the efficiency of its automated system. A tremendous amount of money and human effort is required to keep that system in a state of reasonable repair. Some portions of the installation are over fifty years old and may give out at any time. If this were to happen telephone service would be rendered inoperative to some areas until major revision of the system be made. The major result of such a failure would be to inconvenience a number of people. On a competitive basis, failures of this type would have to be kept to a minimum, or company profit would be decreased. A second aspect of the telephone system is the large number of sequential duplicative operations that must be performed owing to the manner in which the system grew, namely, piece by piece. Each new piece requires sets of relays and amplifiers and, hence, is less reliable than if it had been possible to lay out the entire telephone system with anticipated demands on the system in mind. Of course this possibility did not exist for a system such as the telephone system, because techniques for planning and anticipating demands on the system for years into the future were not available at the time the tele-

phone networks were first constructed. The economic factors associated with removal of the system and replacing it by one which is designed in an integrated manner are of such moment that "good business practices" appear to be endangered. Thus, the network may continue to grow much in the same manner as it has in the past by a process of additions to and gradual modernization of existing equipments. This must not be the case in, say, the automated automobile factory. Failures of the system will directly affect the operating costs and unit price, giving the competitor with the more reliable installation a decided long-term advantage in unit price. Industrial survival must then be based upon the initial over-all system design for organizational automation.

3. Comments on Partial and Compromise Automation

From the foregoing discussion it might appear that the road to success in the field of organizational automation is an easy one to travel on. Nothing could be farther from the truth. The main road block to be bypassed or avoided is associated with the problems which will arise if a direct application of basic automation techniques are employed in a sequential manner in order to build up to organizational automation. Even after the goal of true organizational automation is sighted and the path charted, many troublesome areas will remain. Organizational automation may be viewed as a sort of ultimate goal, a lofty objective that in reality may never be completely attained. Automated systems will not suddenly spring into full bloom in one fell swoop. A build-up will occur from the very simple to the very complex. This does not mean that the build-up will proceed along the lines of sequential basic automation, however, but that each additional area encompassed by automatic systems will have to be evolved as a system rather than as a complex of individual parts. In addition, original planning will have to include a very realistic growth potential factor for the system. Partial automation will answer many of the more complex realistic problems that will face companies in the near future. Among the largest of these problems will be the obtaining of satisfactory electronic components. Whereas it might be thought that the environment is easily controlled and hence would pose no problem in terms of decreased component life, it becomes apparent that when the electronic equipment is to be associated with active mechanical/thermodynamical/hydrodynamical elements and systems, it

is not always possible to maintain all of the electronics in the ideal control environment. Measurements that must be made on active machinery to determine the performance of this machinery may indicate an environment of high shock and vibration. One may argue that electrical information can be communicated from the machine to a remote site where the electronic equipment is located, but the solution of incorporating brushes and sliprings into the machine might well be a less reliable one than the placement of transistors and other electronic components directly on or in the machine.

Another major problem which must be faced is that whereas, in the beginning, phases of automation modifications and adaptations of existing equipments and machines may suffice, in later stages studies will surely indicate the need for development of new machine complexes. Our problems here will certainly center around the various moving parts and linkages, the problems of bushings and bearings which in the past have been the shortcomings of many mechanical devices. Studies then should be made to attempt to eliminate, wherever possible, moving parts, to combine functions so as to reduce to a bare minimum the number of moving parts needed in the system, and possibly to look for techniques other than those that we today think of as machine processes to perform many of the operations needed.

A third major problem encompasses the investigations and studies of fail-safe operation. If we are to be realistic, then we must admit the fact that the equipment will have a finite life expectancy. When one of the elements in the system fails, it is mandatory that protective devices be set in motion so as to arrest the activity of the system in those areas where self-destruction would surely follow. An even more subtle problem is the protection of equipment which may undergo temporary overload since the effects of these overloads make themselves felt in life expectancy of the overloaded equipment. What, then, is the proper balance to be achieved between an investment in protective devices and their reliability as compared to shut-down for maintenance of the automated system? This is a question which only study can answer.

4. Planning

If compromise automation is the realistic path to be traveled, then one must be prepared in the early phases of automation to sell what the market will demand, namely machines and devices which fall under the

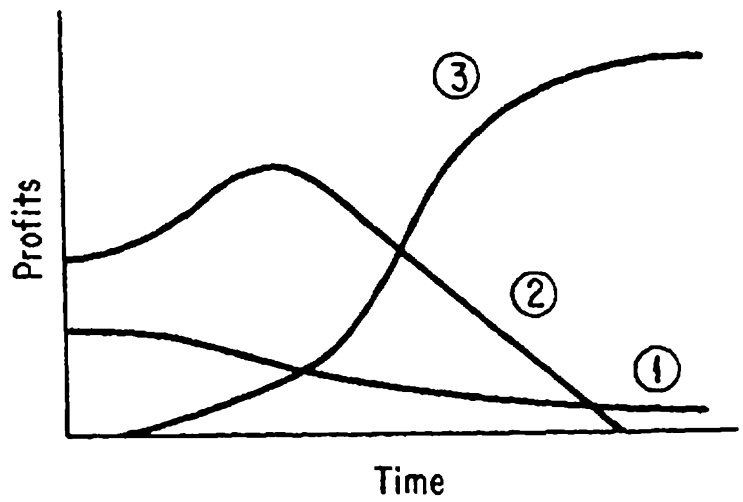


Fig. 5.1. Automation-type comparisons.

heading of basic automation. These will surely be in demand until such time as the complexity of the aggregate structures of these machines cause owners' profits to begin to decrease. When this happens one must be prepared to evolve rapidly a series of systematized automatic installations. Although basic automation is not the path leading to the ultimate, a good market will exist for a number of years in this field. The longer range research problems that must be solved before we can approach the ultimate should be done during this period.

Fig. 5.1 develops schematically the probable break-point philosophy that we have discussed. In this figure, Curve ① represents non-automated methods; Curve ② basic automation methods; and Curve ③ compromise automation methods.

5. Summary

We have in this section considered a few specific applications in automation. The distribution in time of methods has been conjectured, although the exact abscissae of the crossover points of Curves ①, ②, and ③ depend on many factors not yet clear. It should be emphasized once again that "kluging" does not provide an acceptable path from basic automation to organizational or compromise automation.

An additional and somewhat controversial point is that of general purpose versus special purpose systems. As we noted in Section 3.2, complexity is a significant problem in systems engineering. Obviously, generality of purpose will be reflected directly in complexity. As indicated in the discussion of compromise automation in Section 5.1, there is an upper bound on complexity compatible with reliability at any given level of the states-of-the-arts. On the other hand, completely special purpose design does not provide for slight alterations in missions nor for growth potential. Thus, the system designer must "walk a narrow line" between overcomplication and oversimplification.

Section 5.3

L'Envoi

"Man is a device by means of which the system brings itself to perfection."

—ALMEGA

System technology holds great promise for all of us in both materialistic advantages and intellectual aspirations. The main tools for a new industrial revolution are now in our hands. They are systems engineering and advanced hardware techniques (particularly microelectronics and micromechanics). Certain requisites are needed, however, if these tools are to be implemented. Among these are:

(a) An appropriate attitude context unhampered by tradition.

(b) Serious planning at all organizational levels, including national and international, *by those competent to do so. Most important, specific goals must be set and specific plans evolved to meet these goals.*

(c) A sound theory underlying phenomenological explanation, rather than a mutilated patchwork of special purpose physical theories.

(d) A behavioral theory describing individual and organizational action as it is and will be, rather than as some think it should be. Of

prominent importance will be an adequate value theory. For an example of how such a theory might begin, consult Appendix 10.

(e) For the particular area of manned space flight, greater sophistication will be increasingly required not only in studies as outlined in Appendix 12, but also in the basic elements of communication from man to actuators, more accurate guidance techniques, and communication from sensors to man. Preliminary thoughts in these last three areas are presented in Appendixes 13, 14, and 15.

Appendixes

Appendix 1

Precise Definition of System

"The scope of the process is larger than the machine."

—THORSTEIN VEBLEN

1. Terminology

A familiarity with set theory is assumed. We write capitals for sets, and small letters for elements. "p is an element of S" is written "$p \in S$." "$S \subset T$" means, as usual, "$p \in S$ implies $p \in T$." If S and T are sets,

$$S \otimes T = \{(s, t) \mid s \in S \quad \text{and} \quad t \in T\}$$

where $\{x \mid \pi\}$ means the set of all x satisfying requirement π.† In particular, if $R \subset S \otimes S$, R is called a relation in S.

If $R \subset S \otimes S$, one frequently prefers to write xRy for $(x, y) \in R$. If $R \subset S \otimes S$, R is called a partial ordering of S, provided R is:

[1] *Antisymmetric*: xRy and yRx imply $x = y$.
[2] *Reflexive*: For all x in S (abbreviated $\forall x \in S$), xRx.
[3] *Transitive*: If xRy and yRz, then xRz.

† Provided the existence of such a set does not involve an antinomy.

A relation in S satisfying [2] and [3] but not necessarily [1] is called a quasi-ordering. If R is a partial ordering in S, one usually writes $x \leq y$ for xRy and says that S is a partially ordered set. A partially ordered set S is a directed S if, given $x \in S$ and $y \in S$, there is $z \in S$ so that $x \leq z$ and $y \leq z$. A directed subset of a partially ordered set is a subset which is directed by the original partial ordering.

If $R \subset S \otimes T$, and if $(x, y) \in R$ together with $(x, z) \in R$ imply $y = z$, and if for each $x \in S$ there is $y \in T$ with $(x, y) \in R$, then R is called a mapping of S into T and we write $f: S \to T$ meaning $xf = y$ if, and only if, $(x, y) \in R$. In college mathematics, one usually writes $f(x) = y$, rather than $xf = y$. If $f: S \to T$ and $A \subset S$, then Af is understood to be $\{y \in T \mid$ there is $x \in S$ with $xf = y\}$. If $f: S \to T$ and $Sf = T$, the mapping is said to be onto T. If $f: S \to T$ and $x \in S$, $y \in S$, $x \neq y$ imply $xf \neq yf$, the mapping is said to be 1–1 or bi-uniform. If a mapping $f: S \to T$ is both onto and bi-uniform, there is a mapping $f^{-1}: T \to S$ so that $xff^{-1} = x$; $\forall x \in S$ [note the collegiate for xff^{-1} is $f^{-1}(f(x))$]. The set of all mappings of S into T is written T^S. In particular 2^S is usually interpreted as the set of all subsets of S since it can be mapped in a bi-uniform way onto this set, as follows: given $f: S \to 2$ (i.e., xf is either 0 or 1; $\forall x \in S$), make correspond to f the set $\{p \in S \mid pf = 1\}$. If $f: S \to T$ and $A \subset S$, then $f/A: A \to T$ is defined by $x(f/A) = xf$ for $x \in A$.

A bi-uniform mapping $f: S \to T$ of one partially ordered set onto another is called an order isomorphism provided $x \leq y$ if, and only if, $xf \leq yf$. A mapping $f: S \to T$ where S is a directed set is called a net.

If S is a partially ordered set; and if $f: S \to T$; and if for $t_1, t_2 \in T$, we define $t_1 \leq t_2$ if and only if there are $s_1, s_2, \in s$ with $s_1 f = t_1$, $s_2 f = t_2$, and $s_1 \leq s_2$ in S; then $\leq$ defined in T will be a quasi-ordering and is said to be induced by f.

If S is a directed set, and there is a fixed set of nets $\{\varphi\}$ on S under consideration, then we write $\bar{S}$ for

$$\{\varphi/R_t \mid \varphi \in \{\varphi\};$$

R is a directed subset of S, and $t \in R\}$, where $R_t = \{x \in R \mid x \leq t\}$.

2. The Concept

Definition. A system $\mathfrak{S}$ is an object

$$\{T, \tau, \Gamma, \Sigma, \Omega, \{\Gamma_R\}, \{\Sigma_R\}, \{\omega_{\gamma\sigma}\}, \{\tilde{\omega}_{\gamma\sigma}\}\}$$

subject to postulates 1 to 5 below:

POSTULATE 1. Γ, Σ, and Ω are sets.

POSTULATE 2. T is a directed set, $(T, \leq)$, and τ is a set of directed subsets of T.

CONVENTION: Names are assigned as follows:

Symbol	*Name of Set*	*Name of Element*
T	Chronology	Time
τ	Staging space	Run
Γ	Input space	Input argument
Σ	Phase space	State
Ω	Output space	Output argument

CONVENTION: If $R \in \tau$, certain nets over R are named as follows:

Net	*Name*
$\gamma : R \to \Gamma$	Input
$\sigma : R \to \Sigma$	Staging
$\omega : R \to \Omega$	Output

POSTULATE 3. For each $R \in \tau$,

$$\Gamma_R \subset \Gamma^R \quad \text{and} \quad \Sigma_R \subset \Sigma^R.$$

CONVENTION: The sets Γ_R and Σ_R are called the spaces of R-admissible inputs and R-admissible stagings, respectively.

CONVENTION: If $R \in \tau$ and $t \in R$, we denote by R_t the set $\{s \in R \mid s \leq t\}$.

CONVENTION: if $R \in \tau$, $\gamma \in \Gamma_R$, the sets $R\gamma$ and $R\sigma$ are understood to bear the quasi-orderings induced from R by γ and σ, respectively. Subsets of $R\gamma$, for example, are understood to inherit this quasi-ordering.

POSTULATE 4. If $R \in \tau$, and $\gamma \in \Gamma_R$, and $\sigma \in \Sigma_R$, there is a mapping

$$\tilde{\omega}_{\gamma\sigma} : \Gamma \otimes \Sigma \to \Omega$$

defined and called the $\gamma\sigma$ correlatant.

CONVENTION: If $R \in \tau$, and $\gamma \in \Gamma_R$, and $\sigma \in \Sigma_R$, the mapping

$$\omega_{\gamma\sigma}: R \to \Omega,$$

defined by

$$\omega_{\gamma\sigma} = \left(\frac{\gamma}{R_t}, \frac{\sigma}{R_t}\right) \tilde{\omega}_{\gamma\sigma}, \tag{24}$$

is called the $\gamma\sigma$ resultant.

POSTULATE 5. If $R \in \tau$, $S \in \tau$, $f: R \to S$ is an order isomorphism, $\gamma \in \Gamma_R$, $\sigma \in \Sigma_R$, $\hat{\gamma} \in \Gamma_s$, $\hat{\sigma} \in \Sigma_s$, $\hat{\gamma} = f^{-1}\gamma$ and $\hat{\sigma} = f^{-1}\sigma$, then

$$\omega_{\hat{\gamma}\hat{\sigma}} = f^{-1}\omega_{\gamma\sigma}. \tag{25}$$

Appendix 2

Notion of Generalized Logical Design

"Round numbers are always false."—JOHNSON

1. Terminology

A familiarity on the part of the reader with elementary set theory is assumed. We write capitals for sets and small letters for elements. The proposition "p is an element of S" is written "$p \in S$." The proposition "$S \subset T$" means, as usual, "$p \in S$ implies "$p \in T$." If π is a property relevant to certain elements, $\{X \mid \pi\}$ denotes the set of all X with property π, provided that the existence of such a set does not invoke an antinomy. If S and T are sets, $S \otimes T$ is defined as $\{(s, t) \mid s \in S \text{ and } t \in T\}$. In particular,

$$R^n = R \otimes R^{n-1}$$

is defined by recursion.

If $A \subset S \otimes T$, and if $(x, y) \in A$ together with $(x, z) \in A$ imply $y = z$, and if for each $x \in S$ there is $y \in T$ with $(x, y) \in A$, then A is called a mapping of S into T and we write "$a: S \to T$" meaning "$xa = y$ if and

only if $(x, y) \in A$." In collegiate mathematics, one usually writes $a(x) = y$ rather than $xa = y$. If $a: S \to T$ and $B \subset S$, Ba is understood to be $\{y \in T \mid$, there is $x \in B$ with $xa = y\}$. If $a: S \to T$ and $Sa = T$, the mapping is said to be *onto* T. If $a: S \to T$ and $y \in S$ together with $y \neq x$ imply $xa \neq ya$, the mapping is said to be *bi-uniform* or *one-to-one*. If $a: S \to T$ is both bi-uniform and onto, there is clearly a mapping $a^{-1}: T \to S$, also biuniform and onto, which is called the *inverse* of $a: S \to T$, so that for all $x \in S$ and all $y \in T$, $xaa^{-1} = x$ and $ya^{-1}a = y$ [note that the collegiate for xaa^{-1} is $a^{-1}(a(x))$ and not $a(a^{-1}(x))$].

Let A be a set. An *n-ary* operation, $*$, in A is a mapping $*: A^n \to A$. If $n = 1$, we have a unary operation; if $n = 2$, we have a binary operation, etc. A *finitary operation* is an n-ary operation, with n not necessarily specified but regarded as some positive integer. In the case of a binary operation, one usually writes $x*y = z$ rather than $(x, y)* = z$. An algebra, $A(*_1, *_2 \cdots)$, is a set A together with a number (finite or infinite) of finitary operations in A.† Let $A(*_1, *_2, \cdots)$ be a fixed algebra which we shall abbreviate $A(\Theta)$, where Θ denotes the set of operations. A finitary operation defined in A (not necessarily a member of Θ) we shall call a *function* in A. A function in A is said to be *algebraic* if it can be expressed in terms of variables over A and constants from A by some finite sequence (including parallel steps) of members of Θ. A (Θ) is said to be *functionally complete* if all functions in A are algebraic. If $\Phi \subset \Theta$, Φ is said to be *algebraically complete* if every algebraic function in $A(\Theta)$ is an algebraic function in $A(\Phi)$. Trivially, Θ is always algebraically complete in the algebra $A(\Theta)$.

A relation in S is a set $R \subset S \otimes S$. A relation R in S is said to be a *partial ordering* of S provided it has the properties:

[1] *Reflexivity*: $(a, a) \in R$, for all $a \in S$.

[2] *Antisymmetry*: $(a, b) \in R$ and $(b, a) \in R$ imply $a = b$.

[3] *Transitivity*: $(a, b) \in R$ and $(b, c) \in R$ imply $(a, c) \in R$.

If R is a partial ordering in S, we usually write $a \leq b$ for $(a, b) \in R$. The proposition "$a < b$" is defined as "$a \leq b$ and $a \neq b$." A partial ordering, R, in S is said to be *linear* (or *simple*) if $a \in S$ and $b \in S$ imply either $(a, b) \in R$ or $(b, a) \in R$. S is said to be a *lattice* under a partial ordering

† Sometimes mappings involving auxiliary variables from other sets are also admitted, but we shall not include such since they are not needed for the purposes of this operation.

R, provided the following conditions subsist:

[4] If $x \in S$ and $y \in S$, there is $u \in S$ so that $u \leq x$, $u \leq y$; and if $z \leq x$ and $z \leq y$, then $z \leq u$.

[5] If $x \in S$ and $y \in S$ there is $v \in S$ so that $x \leq v$, $y \leq v$; and if $x \leq z$ and $y \leq z$, then $v \leq z$.

One usually writes $x \wedge y$ for the element u of Condition [4] (whose uniqueness is assured by Condition [2]), which is called the *meet* of x and y. One usually writes $x \vee y$ for the element v of Condition [5], which is called the *join* of x and y. It is easily shown that in the lattice the following conditions subsist:

[6] *Idempotency*: $x \wedge x = x \vee x = x$.

[7] *Commutativity*: $x \vee y = y \vee x$ and $x \wedge y = y \wedge x$.

[8] *Associativity*: $x \wedge (y \wedge z) = (x \wedge y) \wedge z$ and $x \vee (y \vee z) = (x \vee y) \vee z$.

[9] *Alternation*: $x \wedge y = x$ if and only if $x \vee y = y$.

One may also show that if $L(\wedge, \vee)$ is an algebra with two binary operations satisfying Conditions [6, 7, 8, and 9], and if one defines "$x \leq y$" by "$x \wedge y = x$," then L is a lattice in which meet and join coincide with the original operations of L.

If $L(\wedge, \vee)$ is a lattice, and if R^+ represents the non-negative real numbers, a mapping $\nu: L \rightarrow R^+$ is called a norm in L provided the following conditions subsist:

[10] *Strict Isotonicity*: $x < y$ implies $x\nu < y\nu$, where the ordering relation for numbers is the customary one.

[11] *Modularity*:

$$x\nu + y\nu = (x \wedge y)\nu + (x \vee y)\nu.$$

2. The Problem

An algebraic design problem may be described as follows:

[1] There is given an algebra, $A(\Theta)$.

[2] There is given an algebraic function, f, in $A(\Theta)$.

[3] There is given an algebraically complete $\Phi \subset \Theta$. Φ is called the *logical philosophy*.

[4] There are given devices which accomplish the implementation of members of Φ. The class of all such given devices is called the *hardware philosophy*.

[5] By means of whatever identities may be extant relating members of Θ, f is made to range over all its equivalent forms as an algebraic function in $A(\Phi)$. This set of forms is written F_f.

[6] Associated with each variable in f and with the output of each device in the hardware philosophy is an *availability criterion* describing in what ways the output may be distributed as input to sets of members of the hardware philosophy.

[7] Associated with each $\varphi \in F_f$ are all flow diagrams exhibiting how φ may be implemented within the hardware philosophy. Those diagrams, which do not violate any of the availability criteria, constitute a set $D\varphi$ (which may, of course, be null).

[8] It is convenient to define

$$D_f = \bigcup_{\varphi \in F_f} D\varphi.\dagger \tag{26}$$

[9] There is given a set of normed lattices,

$$\{P_i(\wedge, \vee, \nu_i)\}_{i=1}^{n},$$

called *penalty factors*.

[10] There are given positive integers

$$\{v_i\}_{i=1}^{n}$$

called *relevancies*.

[11] There are given mappings,

$$p: D_f \to P_i; \qquad i = 1, 2, \cdots, n,$$

called *subpenalties*.

[12] The direct product,‡

$$P = \prod_{i=1}^{n} P_i,$$

is called *penalty space*, and the mapping $\pi: D_f \to R^+$ defined by

$$x\pi = \sum_{i=1}^{n} v_i \cdot x\nu_i, \tag{27}$$

is called the *penalty*.

† Where $\bigcup$ indicates join of sets.
‡ $P_1 \otimes P_2 \cdots \otimes \cdots \otimes P_n$ with operations taken by coordinates.

[13] The mapping

$$xp = x\pi - \inf_{y \in D_f} y\pi \tag{28}$$

is called the *regret.*

[14] The problem of logical design is that of selecting a $\varphi \in F_f$ with the property that there is $x \in D_\varphi$ with $xp = 0$, or at least, xp being very small.

[15] The *problem of hardware design* is that of selecting $x \in D_\varphi$ with $xp = 0$, or at least, xp very near $\inf_{y \in D\varphi} yp$.

3. Further Discussion

[1] It is, in general, nonsense to speak of optimal or suboptimal logical design in the absence of a definite hardware philosophy and penalty evaluation.

[2] The algebra A is generally discrete, and in most current cases, only unary and binary operations are considered. It is no more difficult, however, to state the problem for more general algebras enabling subsumption, for example, of continuous variables. In most cases, the algebra A is also functionally complete, i.e., all finitary functions are algebraic. Functional completeness is required primarily in general purpose designs.

[3] The availability criteria are essential since sets of devices generally have compatibility requirements, and there are generally limitations as to how an output of one device may be distributed as inputs to other devices.

[4] We have restricted our attention to simple-output devices. For other types a modification of the algebraic formalism is required. In general, one would anticipate, subject to input availability criteria, that multi-output devices would permit consideration as sets of single-output devices (i.e., vector valued).

[5] There may be one or many flow-diagrams for a given φ depending on parallel *vs.* sequential responses of devices and upon parallel *vs.* sequential ambiguity in φ. Alternatively, the case may subsist where all flow-diagrams for φ violate availability criteria.

[6] We have yielded some generality in assuming numerical penalties (i.e., negative utilities), but there is currently extant no optimization theory in lattices. We have not, however, yielded to the customary assumption of linear ordering other than in valuation. The penalty factors themselves are generally inherent in objectives of and specifications for the design and may or may not have fixed bounds as a necessity. The norms them-

selves may be objectively or subjectively assigned depending on the problem. The relevancies may likewise be subjective or objective. It is felt that rational relevancies suffice for all "realistic" problems. Changes of norming scales then permit restriction of relevancies to integral values.

[7] If it is difficult for one to think in terms of general normed lattices, one may think of either of the following special cases:

[a] P_i is a subset of the non-negative real numbers and ν_i is a non-negative, monotone strictly increasing, real-valued function on P_i. Thus, P is a subset of the non-negative cone in Euclidean n-space.

[b] P_i is a subset of the non-negative cone of Euclidean n_i-space, and $\nu_i(\lambda)$ for $\lambda \in P_i$ is a linear combination with positive coefficients of the coordinates of λ. Thus, P is a subset of the positive cone of Euclidean

$$\left(\sum_{i=1}^{n} n_i\right) \text{space.}$$

[8] For many logical design problems, there are formalistic methods of associating penalties with flow diagrams so that an algebraic method can be devised for this purpose. We have, however, been unable to convince ourselves that the existence of such methods is assured in all general cases.

[9] Although it may become an insurmountable difficulty to apply strictly the description given to a particular problem, the description should at least place various facts in correct relative perspective.

[10] In many problems, attainment of minimum regret should yield to the already well-known computational devices of linear, quadratic, functional, and dynamic programming.

Appendix 3

Notion of Dynamic Programming

"For her own breakfast she'll project a scheme, nor take her tea without a stratagem."

—EDWARD YOUNG

1. Summary

We now investigate the theory of one-person decision processes, otherwise called dynamic programming. In accordance with our personal philosophy relating to applications of mathematics, we give a definition of dynamic programming processes which is obviously too general for use per se, but one which, we hope, may be suitably compromised to a useful level of generality at a later date by restriction of its components. First, we shall define the notion of "optimal policy" but shall not enquire as to existence of such, as this clearly requires, in general, additional structure of a nature probably topological. The notion of "good policy" is not even defined since this requires a satisfactory theory of approximation in lattices which is not currently extant. However, an example is given which illustrates the notion of "good policy" in a case where monotoneity

and measurability yield satisfactory approximation. As to actual results, we content ourselves for the present with verification of a version of what Bellman calls the "Principle of Optimality."†

2. Terminology

A binary relation R on a set S (i.e., R is subset of $S \otimes S$) is called a *partial ordering* of S if and only if R has the properties:

[1] *Reflexivity*. For all $a \in S$, $(a, a) \in R$.
[2] *Antisymmetry*. If $(a, b) \in R$ and $(b, a) \in R$, then $a = b$.
[3] *Transitivity*. If $(a, b) \in R$ and $(b, c) \in R$, then $(a, c) \in R$.

If R is a partial ordering of S, one frequently writes either $a \leqslant b$ or $a \leq b$ for $(a, b) \in R$ and speaks of the *poset* [abbreviation for partially ordered set] $(S, \leqslant)$ or $(S, \leq)$, respectively. If R is fixed so that $\leqslant$ or $\leq$ is understood, one speaks merely of the poset S. The following conventions are understood in any poset $(S, \leq)$:

[4] $a < b$ if and only if $a \leq b$ and $a \neq b$,
[5] $a \geq b$ if and only if $b \leq a$,
[6] $a > b$ if and only if $b \leq a$ and $a \neq b$.

If $(S, \leq)$ is a poset and $k \in S$, one writes

[7] S^k for $\{x \in S \mid k < x\}$,
[8] S_k for $\{y \in S \mid y \leq k\}$.

If $(S, \leq)$ is a poset and $T \subset S$ one writes

[9] T^+ for $\{x \in S \mid y \in T \text{ implies } y \leq x\}$,
[10] T_+ for $\{u \in S \mid v \in T \text{ implies } u \leq v\}$.

It is obvious that $T \cap T^+$ and $T \cap T_+$ each have at most one element. The element, if any, in $T \cap T^+$ is called the last *element* of T. The element, if any, in $T \cap T_+$ is called the *first element* of T.

If T^+ has a first element, the element is called the *join* of T and written $\bigvee_{x \in T} x$, $\sup_{x \in T} x$, or $\text{lub}_{x \in T} x$. If T_+ has a last element, the element is called the *meet* of T and is written $\bigwedge_{x \in T} x$, $\inf_{x \in T} x$, or $glb_{x \in T} x$.

If for each element x in a poset S one has $\{x\}^+ \cup \{x\}_+ = S$, S is a *loset* (abbreviation for linearly ordered set).

† Richard Bellman, "The Theory of Dynamic Programming," *Bull. Amer. Soc.*, LX (1945), 503–515.

A poset in which each non-null subset has a first element is a *woset* (abbreviation for well-ordered set). Clearly, every woset is necessarily a loset.

A poset $(\Delta, \leqslant)$ is a *diset* (abbreviation for directed set) if and only if $\{\alpha, \beta\}^+ \neq \square$† for each $(\alpha, \beta) \in \Delta \otimes \Delta$. If $(\Delta, \leqslant)$ is a diset, and B is any set, a mapping $\mathfrak{n}:\Delta \to B$ is called a *net* in B based on Δ. If $\Gamma \subset \Delta$ and $(\Gamma, \leqslant/\Gamma \otimes \Gamma)$‡ is a diset, Γ is called a disubset of Δ. Clearly, if Γ is a disubset of Δ and $\gamma \in \Gamma$, then Γ^γ and Γ_γ are also disubsets of Δ. If Γ is a disubset of Δ, $\mathfrak{n}$ is a net based on Γ, and $\gamma \in \Gamma$, the nets $\mathfrak{n}/\Gamma^\gamma$ and $\mathfrak{n}/\Gamma_\gamma$ are denoted by $\mathfrak{n}^\gamma$ and $\mathfrak{n}_\gamma$, respectively. One sees by induction that if Γ is a finite subset of a non-null diset, then $\Gamma^+ \neq \square$.

A non-null poset $(\Lambda, \leq)$ is a *lattice* if and only if

$$\bigvee_{\gamma \in \{\alpha,\beta\}} \gamma \quad \text{and} \quad \bigwedge_{\gamma \in \{\kappa,\beta\}} \gamma$$

exist for each (α, β) in $\Lambda \otimes \Lambda$. One usually writes $\alpha \vee \beta$ and $\alpha \wedge \beta$ for

$$\bigvee_{\gamma \in \{\alpha,\beta\}} \gamma \quad \text{and} \quad \bigwedge_{\gamma \in \{\alpha,\beta\}} \gamma,$$

respectively. One sees by induction that if Γ is a non-null finite subset of a lattice, then $\bigvee_{\gamma \in \Gamma} \gamma$ and $\bigwedge_{\gamma \in \Gamma} \gamma$ exist. For any lattice Λ, the mappings

$$\vee : \Lambda \otimes \Lambda \to \Lambda$$

and

$$\wedge : \Lambda \otimes \Lambda \to \Lambda$$

may be thought of as algebraic operations. They have the properties:

[11] *Commutativity*. For all (α, β) in $\Lambda \otimes \Lambda$,

$$\alpha \vee \beta = \beta \vee \alpha \quad \text{and} \quad \alpha \wedge \beta = \beta \wedge \alpha.$$

[12] *Associativity*. For all (α, β, γ) in $\Lambda \otimes \Lambda \otimes \Lambda$,

$$\alpha \vee (\beta \vee \gamma) = (\alpha \vee \beta) \vee \gamma$$

and

$$\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma.$$

[13] *Idempotency*. For all α in Λ,

$$\alpha \vee \alpha = \alpha \wedge \alpha = \alpha.$$

[14] *Alternation*. $\alpha \vee \beta = \alpha$ if and only if $\alpha \wedge \beta = \beta$.

† $\square$ designates the null set.

‡ $\leqslant/\Gamma \otimes \Gamma$ indicates the relation $\leqslant$ restricted to the subset $\Gamma \otimes \Gamma$ of $\Delta \otimes \Delta$.

Conversely, if $(\Lambda, \vee, \wedge)$ is a double groupoid (i.e., $\vee$ and $\wedge$ are binary operations under which Λ is closed) and if [11], [12], [13], and [14] subsist,† and if one defines $\alpha \leq \beta$, if and only if $\alpha \wedge \beta = \alpha$, then $(\Lambda, \leq)$ is a lattice in which join and meet are the original operations $\vee$ and $\wedge$, respectively. Thus, a lattice may be regarded either from the viewpoint of poset theory as a poset with extrema for finite subsets or from the viewpoint of universal algebra as a double semilattice with alternation.‡

A lattice Λ is called *upper conditionally complete* if and only if $\Gamma \subset \Lambda$ and $\square \neq \Gamma^+ \neq \Lambda$ imply the existence of $\vee_{\gamma \epsilon \Gamma} \gamma$; that is, every non-null subset which is bounded above has a least upper bound. Lower conditional completeness is dually defined, and the conjunction of the two properties is referred to as merely conditional completeness.

3. Dynamic Programming Processes

DEFINITION. A dynamic programming process, which may be indicated as DPP $\langle (\Delta, \leqslant), B, \mathfrak{N}, A, t, (\Lambda, \leq), \nu \rangle$, is a mathematical system consisting of the objects listed as [1] to [7] below, subject to the postulates [8] to [12] below.

[1] $(\Delta, \leqslant)$ is a diset called "stage space."

[2] B is a set called "decision space."

[3] $\mathfrak{N}$ is a set of nets in B, each of which is based on some di-subset of Δ. $\mathfrak{N}$ is called "policy space."

[4] A is a set called "phase space" whose elements are called "states."

[5] t is a mapping

$$t: A \otimes \mathfrak{N} \rightarrow A.$$

t is called the "transition operator."

[6] $(\Lambda, \leq)$ is an upper conditionally complete lattice called "value space."

[7] ν is a mapping

$$\nu: A \rightarrow \Lambda.$$

ν is called the "payoff."

† As a postulate list this is somewhat redundant. *Cf.* David Ellis, "Notes on the Foundations of Lattice Theory," *Publicationes Mathematicea* (Dubrecen), I (1950), 205–208.

‡ A groupoid which is associative is called a semigroup. A commutative semigroup in which every element is idempotent is called a semilattice. *Cf.* David Ellis, "An Algebraic Characterization of Lattices Among Semilattices, *Portugalise Mathematica,* VIII (1949), 103–106.

[8] If $\mathfrak{n} \in \mathfrak{N}$, and $\mathfrak{n}$ is based on Γ and $\gamma \in \Gamma$, then $\mathfrak{n}_\gamma \in \mathfrak{N}$ and $\mathfrak{n}^\gamma \in \mathfrak{N}$.

[9] If Γ is a disubset of Δ and if $\gamma \in \Gamma$ and if $\mathfrak{n}_1 \in \mathfrak{N}$, $\mathfrak{n}_2 \in \mathfrak{N}$, $\mathfrak{n}_1$ is based on Γ_γ and $\mathfrak{n}_2$ is based on Γ^γ, then there is $\mathfrak{n} \in \mathfrak{N}$ based on Γ with

$$\mathfrak{n}_\gamma = \mathfrak{n}_1 \quad \text{and} \quad \mathfrak{n}^\gamma = \mathfrak{n}_2. \tag{29}$$

[10] If $\square \in \mathfrak{N}$, then, for all $a \in A$,

$$(a, \square)t = a. \tag{30}$$

[11] If $\mathfrak{n} \in \mathfrak{N}$ is based on Γ and $\gamma \in \Gamma$, then, for all $a \in A$,

$$(a, \mathfrak{n})t = \langle (a, \mathfrak{n}_\gamma)t, \mathfrak{n}^\gamma \rangle t. \tag{31}$$

[12] For any $a \in A$ and any disubset Γ of Δ,

$$\langle (a, \mathfrak{n})t\nu \mid \mathfrak{n} \in \mathfrak{N} \text{ and } \mathfrak{n} \text{ based on } \Gamma^+ \rangle = \square. \tag{32}$$

CONVENTIONS.

[13] $\mathfrak{N}(\Gamma) = \{\mathfrak{n} \in \mathfrak{N} \mid \mathfrak{n} \text{ is based on } \Gamma\}$. (33)

[14] One writes $a\,\mathfrak{n}$ for $(a, \mathfrak{n})t$.

[15] If $\mathfrak{n} \in \mathfrak{N}$, $\Gamma(\mathfrak{n})$ denotes the base of $\mathfrak{n}$.

[16] If $\mathfrak{N}(\Gamma) \neq \square$, a mapping

$$\nu\Gamma : A \to \Lambda \tag{34}$$

is defined as follows:

$$a\nu\Gamma = \bigvee_{\mathfrak{n} \in \mathfrak{N}(\Gamma)} a\mathfrak{n}\nu.$$

This indicated join exists by Postulate [12]. The mapping $\nu\Gamma$ is called the *upper value* over Γ of the DPP.

REMARK. If $\mathfrak{N}(\Gamma) \neq \square$ and $\gamma \in \Gamma$, then $\mathfrak{N}(\Gamma^\gamma) \neq \square$ and $\mathfrak{N}(\Gamma_\gamma) \neq \square$, by Postulate [8].

DEFINITION. A policy $\mathfrak{n} \in \mathfrak{N}(\Gamma)$ is called an optimal Γ policy in $D \subset A$ if $\mathfrak{n}\nu/D = \nu\Gamma/D$. An optimal Γ policy in A is called merely an optimal Γ policy.

4. Principles of Optimality

We now contemplate a fixed DPP and assume $\mathfrak{N}(\Gamma) \neq \square$ for the Γ under consideration.

LEMMA 1. If $\gamma \in \Gamma$ and $\mathfrak{n} \in \mathfrak{N}(\Gamma\gamma)$,

$$a\mathfrak{n}\nu\Gamma_\gamma \leq \nu\Gamma.$$

Proof.

$$a\mathfrak{n}\nu\Gamma_\gamma = \bigvee_{\mathfrak{n}' \in \mathfrak{N}(\Gamma\gamma)} a\mathfrak{n}\mathfrak{n}'\nu.$$

By Postulate [9], there is $\mathfrak{n}'' \in \mathfrak{N}(\Gamma)$ with $\mathfrak{n}''/\Gamma_\gamma = \mathfrak{n}$ and $\mathfrak{n}''/\Gamma^\gamma = \mathfrak{n}'$. By Postulate [11], $a\mathfrak{n}\mathfrak{n}' = a\mathfrak{n}''$ so

$$a\mathfrak{n}\mathfrak{n}'\nu = a\mathfrak{n}''\nu \leq \bigvee_{\mathfrak{n}'' \in \mathfrak{N}(\Gamma)} a\mathfrak{n}'''\nu = a^\nu\Gamma.$$

LEMMA 2.

$$\bigvee_{\mathfrak{n} \in \mathfrak{N}(\Gamma\gamma)} a\mathfrak{n}\nu\Gamma_\gamma \leq a\nu\Gamma.$$

Proof. By Lemma 1,

$$a\nu\Gamma \in \{a\mathfrak{n}\nu\Gamma\gamma \mid \mathfrak{n} \in \mathfrak{N}(\Gamma_\gamma)\}^+.$$

LEMMA 3.

$$a\nu\Gamma \leq \bigvee_{\mathfrak{n} \in \mathfrak{N}(\Gamma\gamma)} a\nu\Gamma\gamma.$$

Proof. Let $\mathfrak{n} \in \mathfrak{N}(\Gamma)$. By Postulate [8] and Postulate [11],

$$a\mathfrak{n}\nu = a\mathfrak{n}_\gamma\mathfrak{n}^\gamma\nu \leq a\mathfrak{n}_\gamma\nu_\gamma \leq \bigvee_{\mathfrak{m} \in \mathfrak{N}(\Gamma\gamma)} a\mathfrak{m}\nu_{\Gamma\gamma}.$$

Thus,

$$\bigvee_{\mathfrak{m} \in \mathfrak{N}(\Gamma\gamma)} a\mathfrak{m}\nu\Gamma\gamma \in \{a\mathfrak{n}\nu \mid \mathfrak{n} \in \mathfrak{N}(\Gamma)\}^+.$$

THEOREM. If $\mathfrak{N}(\Gamma) \neq \square$ and if $\gamma \in \Gamma$, then

$$a\nu\Gamma = \bigvee_{\mathfrak{n} \in \mathfrak{N}(\Gamma\gamma)} a\mathfrak{n}\nu\Gamma\gamma.$$

Proof. Proposition is conjunction of Lemmas 2 and 3.

COROLLARY. If $\mathfrak{m}$ is an optimal Γ policy, and if $\gamma \in \Gamma$, then $\mathfrak{m}^\gamma$ is an optimal Γ^γ policy in $A\mathfrak{m}_\gamma$.

Proof. Let $c \in A\mathfrak{m}_\gamma$. Then, for some $a \in A$, $c = a\mathfrak{m}_\gamma$. Thus, applying the above-mentioned theorem,

$$c\mathfrak{m}^\gamma\nu = a\mathfrak{m}_\gamma\mathfrak{m}^\gamma\nu = a\mathfrak{m}\nu = a\nu\Gamma = \bigvee_{\mathfrak{n} \in \mathfrak{N}(\Gamma\gamma)} a\mathfrak{n}\nu\gamma.$$

A. An Example

The purpose of the example is to illustrate the components of a DPP, to show that optimal policies rarely exist, and to demonstrate what one means by using the term "good policy" in appropriate circumstances where value space possesses a pseudonorm or norm, and there is monotone convergence in this pseudonorm toward upper values.

To obtain the example we define components as follows:

[1] $(\Delta, \leqslant)$ is the set of positive integers in their usual ordering.

[2] B is the interval $[0, 1)$.

[3] $\mathfrak{N}$ is the set of all nets in B whose bases are either intervals $[n, n + k]$ or rays $[n, \infty)$ of integers. The null-net based on the null-set is also included in $\mathfrak{N}$.

[4] A is the interval $[0, 1]$.

[5] If $a \in A$, and $\mathfrak{n}$ is based on $[n, n + k]$, let $a\mathfrak{n}$ be $ax_1x_2 \cdots x_{k+1}$, where x_i is the value of $\mathfrak{n}$ at $n + i - 1$ and juxtaposition indicates ordinary multiplication. If $\mathfrak{n}$ is based on $[n, \infty)$, let $a\mathfrak{n}$ be

$$\lim_{k\to\infty} ax_1x_2 \cdots x_k,$$

where the x_i have meaning as before.

[6] Let $(\Lambda, \leq)$ be the lattice of Borel sets† in $[0, 1]$ ordered by inclusion.

[7] Let $a\nu = [0, a]$.

Then, the null-net is an optimal $\square$ policy. For Γ any interval or ray,

$$a^\nu\Gamma = [0, a).$$

There are in this case no optimal Γ policies, but there are "good" Γ policies in the sense that for $\epsilon > 0$ there is $\mathfrak{n}_\epsilon\mathfrak{N}(\Gamma)$ with measure $(a\nu_\Gamma \oplus a\mathfrak{n}_\epsilon\nu) < \epsilon$.‡

† We use Borel set in the sense of Hausdorff rather than Halmos.

‡ $\oplus$ indicates symmetric difference.

Appendix 4

Partial Design of the Hypothetical System: Porcupine

"Tonight yon pilot shall not sleep,
Who trims his narrow'd sail."
—HOLMES

1. Introduction

We consider here a hypothetical land (or perhaps, carrier) based weapons system to be used for defense of a task force on land (or at sea) against bomber attack. The rationale for bomber attack consideration is given in the second paragraph below. We shall examine only the design planning highlights at various system hierarchy levels terminating in control and display console as a subsystem of the information subsystem. Conceptually, let us assume that the system is a two-man (pilot-navigator and fire-control officer) aircraft carrying multiple missiles (probably having nuclear warheads) of about 100 mile range. Furthermore, we make the following assumptions. The development of a multi-channel tracking radar is feasible. The Porcupine (aircraft) has the capability of communicating with Airborne Early Warning aircraft and/or SAGE type

ground systems for aid in acquisition of targets and aid in tracking in an electronic countermeasures environment. The Porcupine's missiles are equipped with combinations of seekers permitting appropriate lock-on by countermeasures homing, active radar detection or infra-red detection. Of course, both the multi-channel Doppler (hereafter called the Porcupine radar) and the missiles' seekers have limitations on look-angle and range. The Porcupine radar has a range of about 125 miles, and the seeker a range of approximately ten miles; countermeasures homing, however, has a range of 100 miles. The Porcupines are on station (through rotation) in a ring whose circumference is about 150 miles from task force center. The missiles travel (once launched) at Mach 3+.

2. System Objectives

The objective of the Porcupine System is provision of a high degree of protection for task forces in the 1960–1965 era. The primary mission to be implemented is attrition of a large portion of a more or less massive enemy bomber strike against the task force or, perhaps, total attrition of a sneak attack. The natural performance criterion would be maximum reduction of threat to task force, but this criterion would be difficult to achieve. A reasonable measure of success would be 80 per cent attrition.

3. Hypothetical Test Attack

Studies of aircraft orders-of-battle and probable SUSAC and SUTAC operations in the 1960–1965 era have been carried out. An apparently reasonable working hypothesis on massive attacks in this era is as follows: From one direction about three wings of craft at Mach 2+ and of Badger weight class are distributed vertically in three bands: high altitude, normal altitude, and "on-the-deck." Each band is distributed horizontally in an inverted wedge, the front of which has about 45° angle subtended at the task force from 550 miles. One additional wing attacks in a similar vertical stratification and horizontal deployment at the 180° point with slight time lag or, perhaps, one wing at the 120° point and one at 240°. Such an attack might well be mounted against a task force participating directly or providing logistic support in a brush-

fire war. In the case of all-out war, such an attack is also good enemy planning since comparison against SAC and TAC facilities, missile bases, and industrial and population centers indicates task forces as having higher mobility and lower priority as strategic missile targets. AEW or SAGE or similar ground radar coverage makes the probability of detection in a sneak attack rise very rapidly with the number of craft involved. Five craft is about maximum for a feasible mission.

4. Porcupine Development

Preliminary mission design indicates deployment on patrol of 16 to 36 Porcupines together with a suitable number of AEW's or overground radar on a perimeter about 150 miles from task force center. Patrol endurance requirements of at least four hours (to attain operational feasibility) more or less imply subsonic capability for the Porcupine. Since the AEW craft (for example) are capable of "looking out" 400 miles, the closing rates involved imply the potential launching from task force of a second wave of Porcupines at time of AEW detection of targets, which would meet the targets just across the original patrol perimeter. Note that planning of the illustrative attack precludes concentration of Porcupines from the opposite semicircle even with adequate warning times.

5. Channel Capability Requirement

In the best case, then, perhaps 36 Porcupines could be brought to bear upon the primary three-wing attack. As previously remarked, a reasonable system objective appears to be about 80 per cent attrition outside task force defenses per se (which might include fighters and ground-to-air missiles). Thus, each Porcupine must necessarily have multi-missile capability since closing time from Porcupine acquisition to target passage is roughly twice missile-target closing time, and since completely sequential operation (i.e., fire—wait for kill; fire—wait for kill; etc.) is inadequate even with 100 per cent kill probability. Using a safety factor of 2, a minimum capability is approximately simultaneous handling of six missile-target combinations per Porcupine; a less marginal requirement would be twelve to sixteen.

6. Pre-launch Design

The most drastic implication of the minimum capability requirement is that of a necessity, once missile-target assignments have been made, to maintain simultaneously all selected targets within the Porcupine Radar look cone as well as feasible missile trajectory cones from first firing to last close (since closes may not have ordering in time identical with firings). This requirement is clearly a four-dimensional limitation upon admissible flight envelopes for the Porcupine during combat. Depending upon the versatility of the Porcupine's missile itself (i.e., "*G* load limit" and response to steering signals), trajectory cone constraints may or may not be stringent.

7. Strategy Evaluation

The ultimate (i.e., completely automatic) system would examine all targets available to it, evaluate each group of six (or twelve or sixteen) targets for threat to task force, evaluate admissible envelopes for each group, and complete its firings in an optimal (i.e., maximum reduction of threat) feasible fashion. Unfortunately, it is not yet clear how threat to task force can be evaluated in a Porcupine on patrol, nor how the solution of the ensuing staggering problem of combinatorial computations for optimization can be made airborne. (The number of possible combinatorial choices rises into the billions with only a few targets.) These questions should be studied as growth potential.

There are, however, reasonable substitutes in semiautomatic mode with dynamic (i.e., sequential) rather than combinatorial planning or with partial combinatorial planning by the FCO. Here, target selection is done by the FCO on the basis of his judgment and kill probabilities furnished him. The FCO assigns missile-target combinations sequentially. The Porcupine is constrained (by director information to the pilot) to fly in admissible envelopes for maintenance of look and other feasibility cones. In the sequential case, each combination assigned by the FCO after the first represents a new constraint (and, therefore, reduction of admissible envelope) on the Porcupine. Those that would result in inconsistency are rejected. This type of control is, of course, developed by application of the principle of dynamic programming, which replaces one

high-dimensional decision with successive low-dimensional decision problems.

Between the dynamic semiautomatic mode and the optimal automatic mode lies an entire spectrum of pre-launch designs that should be studied for growth potential as one main source of system evolution. It should be noted that this will be a central problem not only in the Porcupine System per se but generally in multi-channel vehicular weapons systems (and, perhaps, even fixed-locus weapons systems) as long as there are temporogeometric constraints upon the sensory and/or actuation equipment utilized. These design considerations offer an example of the logical flow from mission requirements to preliminary equipment specifications. We have here shown it arising as a portion of the informational subsystem from the assumed mission requirements.

Another necessary element of the pre-launch situation is an indication of the operational readiness of each missile. While in a fully automatic mode maximum readiness information might be necessary for optimization, it is probable that ternary information is adequate in the semiautomatic mode, i.e., combat ready, degraded operation probable but reasonable confidence of non-abort, probable abort.

8. Necessity for Integrated Systems Approach

As indicated, the pre-launch situation required simultaneous consideration of craft, missiles, targets, and interior electronics. This is only one of many factors dictating the necessity for an integrated systems approach to the Porcupine System. Attempts to give isolated consideration to portions of the system can easily result in fallacy. It might be concluded, for example, that the Porcupine radar scan rates imply a requirement for very fast computation over-all, but such a conclusion would be erroneous on several grounds. A filter bank read-out from the Porcupine radar is only a very small portion of the data processing problem. Moreover, it in itself does not involve so much computational as sorting and correlative processing. The Porcupine processing problem involves a great many aspects, and undue emphasis on any one may so distort the system that apparently very difficult but actually totally irrelevant problems may appear. Approaches to semiautomatic pre-launch provide a good example of how directed study can reduce an apparently insurmountable problem (i.e., the combinatorial target selection and tracking) to a much more feasibly attacked problem without significant impair-

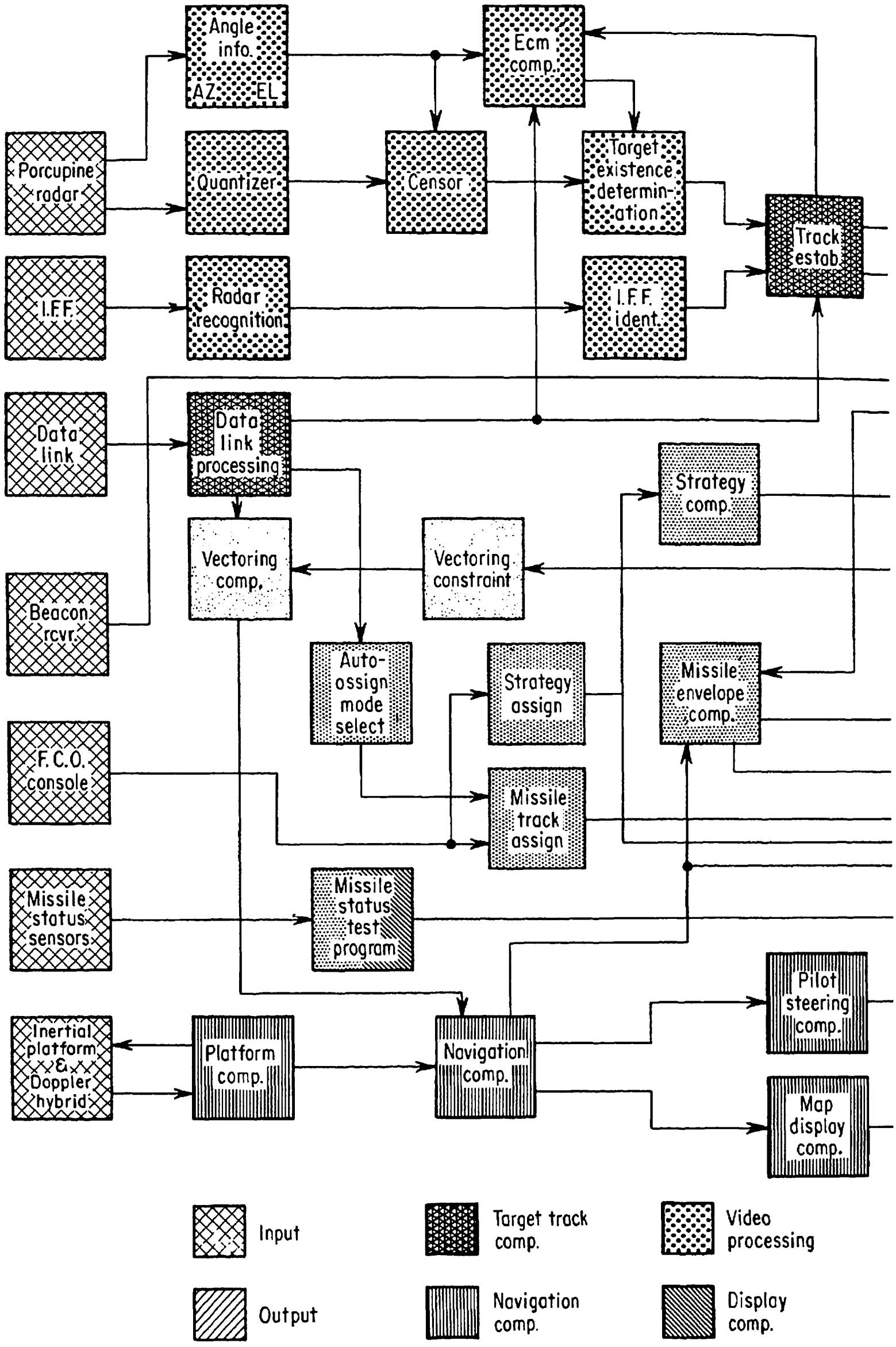
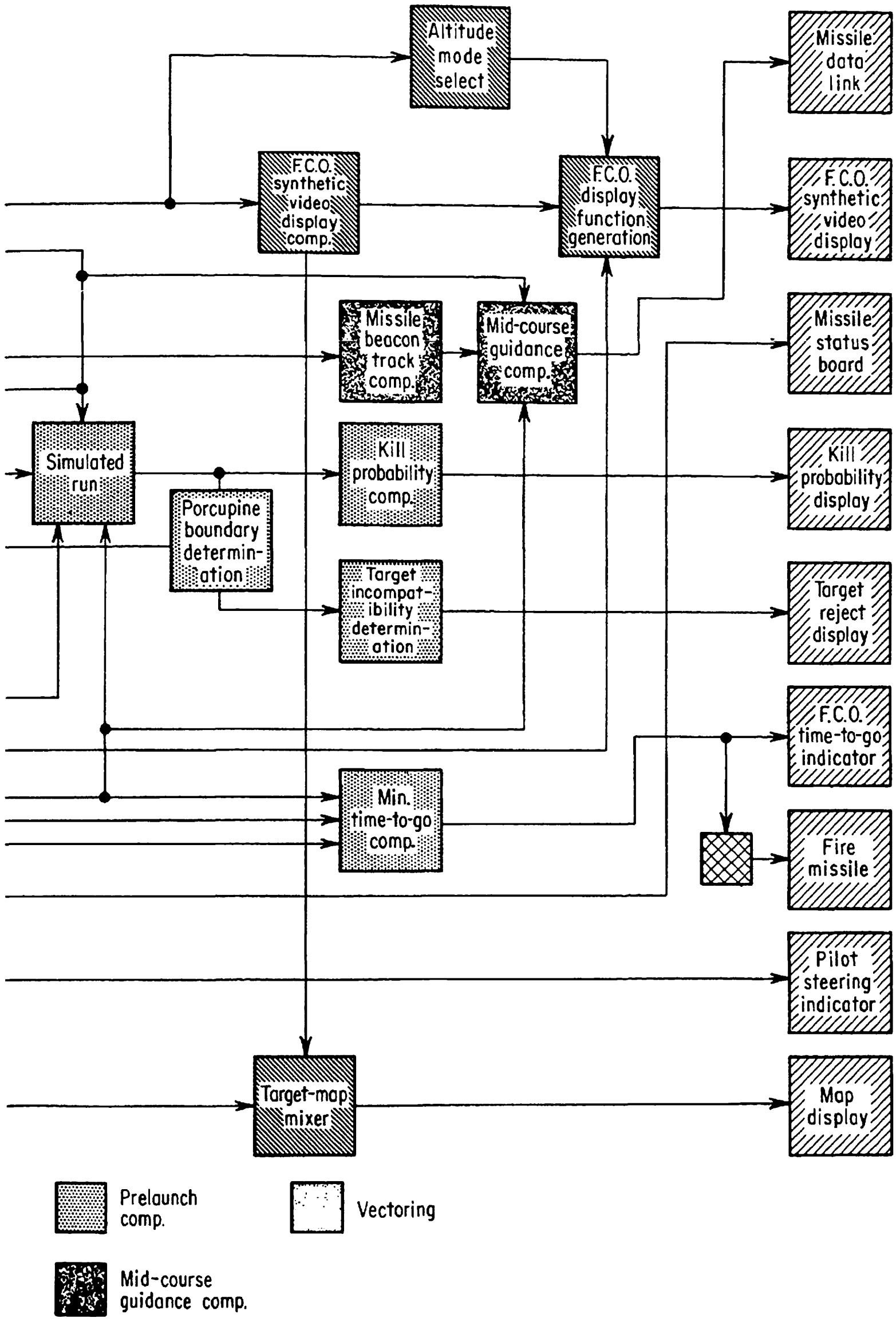


Fig. A4.1. Information flow diagram.


Altitude mode select
Missile data link
F.C.O. synthetic video display comp.
F.C.O. display function generation
F.C.O. synthetic video display
Missile beacon track comp.
Mid-course guidance comp.
Missile status board
Simulated run
Kill probability comp.
Kill probability display
Porcupine boundary determination
Target incompatibility determination
Target reject display
F.C.O. time-to-go indicator
Min. time-to-go comp.
Fire missile
Pilot steering indicator
Target-map mixer
Map display
Prelaunch comp.
Vectoring
Mid-course guidance comp.

ment of system capabilities. In attempting to consider seriously the craft and its missiles as independent entities, for example, one might well ignore the fact that missile assignment and firing are necessarily a function of not only present but future behavior of both targets and the Porcupine itself as well as procedure evaluation in the Porcupine electronics.

9. Information Flow for Porcupine in Attack Mode

The following information is related to Fig. A4.1. The inputs to the system are shown on the left hand side of this diagram. A list of these inputs are:

(a) Porcupine radar
(b) IFF
(c) Data link (RF, secure)
(d) Beacon receiver
(e) FCO console
(f) Missile status sensors
(g) Platform and Doppler

On the right hand side of the chart are nine output systems which are driven from the computer. These outputs are:

FCO console

(a) FCO synthetic video display
(b) Missile status board
(c) Kill probability displayed on main display tube
(d) Target reject indicator
(e) Time-to-go indicator

FCO and pilot

(f) Horizontal situation display

Pilot

(g) Pilot steering indicator

Missile

(h) Fire missile
(i) Missile data link

The small box on the arrow entering the fire missile box is the fire control switch which is held closed by the FCO, enabling the missile to be fired automatically by the computer when the proper initial conditions have been achieved.

The major functional portions of the processing have been separated into six categories by shade-coding the information flow diagram. A description of these areas is presented next:

(1) VIDEO PROCESSING. This portion of the system performs the necessary statistical operations on the video coming from the Porcupine radar and establishes the existence of targets. Another portion of this operation is the evaluation of the IFF data to indicate which of the targets are friendly. When in the ECM, or jammed mode, only angle information is available for target existence determination. Two modes will be available. These are:

(a) Triangulation from AEW and/or Porcupine data.
(b) Self-contained synthetic range determination obtainable from angle information only by suitable filtering in the ECM computation.

(2) TRACK ESTABLISHMENT. The information obtained from video processing together with data link information is used to establish target tracks which serve as identification of the targets and provide for each target a code under which all pertinent target information is filed. When a missile is assigned to a given target, the track number will serve to identify the particular missile assigned to that target.

(3) DISPLAY COMPUTATIONS. The many coordinate transformations, switching operations, and logical decisions needed to generate the output displays are performed in the portion of the diagram referred to under this heading. Information generated goes to three displays on which are presented pilot steering information, navigation information, and fire control information. In this last display, the FCO can elect three modes of display. In all modes, the synthetic video is available. These are:

(a) Altitude mode
(b) Kill probability mode
(c) Missile envelope mode

(4) PRE-LAUNCH COMPUTATIONS. Perhaps the most complex of the six functions of the Porcupine processing is the pre-launch computations. It is in this area alone that the objectives of the system will be met or not.

PRELIMINARY ESTIMATES FOR A PORCUPINE SYSTEM

Function	*Type of process*	*Computation rate (probable)*	*Remarks*
Radar data processor	Non-arithmetic processing Censoring Target Establish IFF Processing angle computations	Need buffer or sync. drum. Could use mag. cores or delay line +16 FF diode tree	Sp. purpose computation
Target track computation	USC-2 processing and target track estimate.		0 to 254 targets can be handled 29×332 (24 bit) words are used
Pre-launch	Kill probability Intercept prediction Boundary condition Missile checkout	35,760 bits/sec/target × 3 strategies × 16 targets	Might get very complicated. Effectiveness of system depends on these computations.
Midcourse guidance	Steer missile to lock on and point seeker, missile tracking	92,000 bits/sec/missile × 8 missiles	Should be flexible for tactics change
Navigation	Doppler inertial	2% integrators at 120 integrations per sec	DDA programmable
Vectoring intercept	Vector missile into firing condition	31,600 bits/sec	Data come from USC 2 and own navigation system
Displays	FCO synthetic video with altitude and kill prob., missile status., nav. disp. and pilot steering	Not applicable	Reducible to about 4 cu. ft. installation

Fig. A4.2

If simultaneous multiple target capability is to be obtained, the following computations are mandatory.

(a) Missile firing envelope computations. These computations must be carried out for all possible strategies available to the fire control officer such as:
   a. Boost-glide-boost
   b. Boost-to-impact
   c. Boost-dive

(b) Simulated missile runs for each strategy to determine highest kill probability for saturated case.

(c) Porcupine boundary determination. If multiple targets are to be fired against, it must be determined that *all* targets against which missiles are to be fired will remain in the radar look angle for the duration of the flight of the missiles. Another question which must be answered by the computer is what bounds are placed on the flight path for the Porcupine so as to keep track of the targets against which it is to launch missiles.

(d) Kill probability calculations.

(e) Target incompatibility computation.

(f) Minimum time-to-go computations.

(5) NAVIGATION COMPUTATIONS.

(6) VECTORING COMPUTATIONS PRIOR TO AND AFTER LAUNCH.

(7) MIDCOURSE GUIDANCE COMPUTATIONS.

Thus, missile status determination is viewed as part of the display computations as well as a portion of the pre-launch computations.

Each of the boxes shown in Fig. A4.1 could be further analyzed into a full page detailed block diagram of information flow.

Figure A4.2 shows preliminary estimates and rates associated with Fig. A4.1 for the hypothetical Porcupine System. Throughout this section, six-channel capability has been assumed.

10. FCO Console

Consideration of the foregoing information flow makes it possible now to rough out the tasks and equipment associated with the FCO (fire control officer). Fig. A4.3 shows a probably near-optimal sequence of

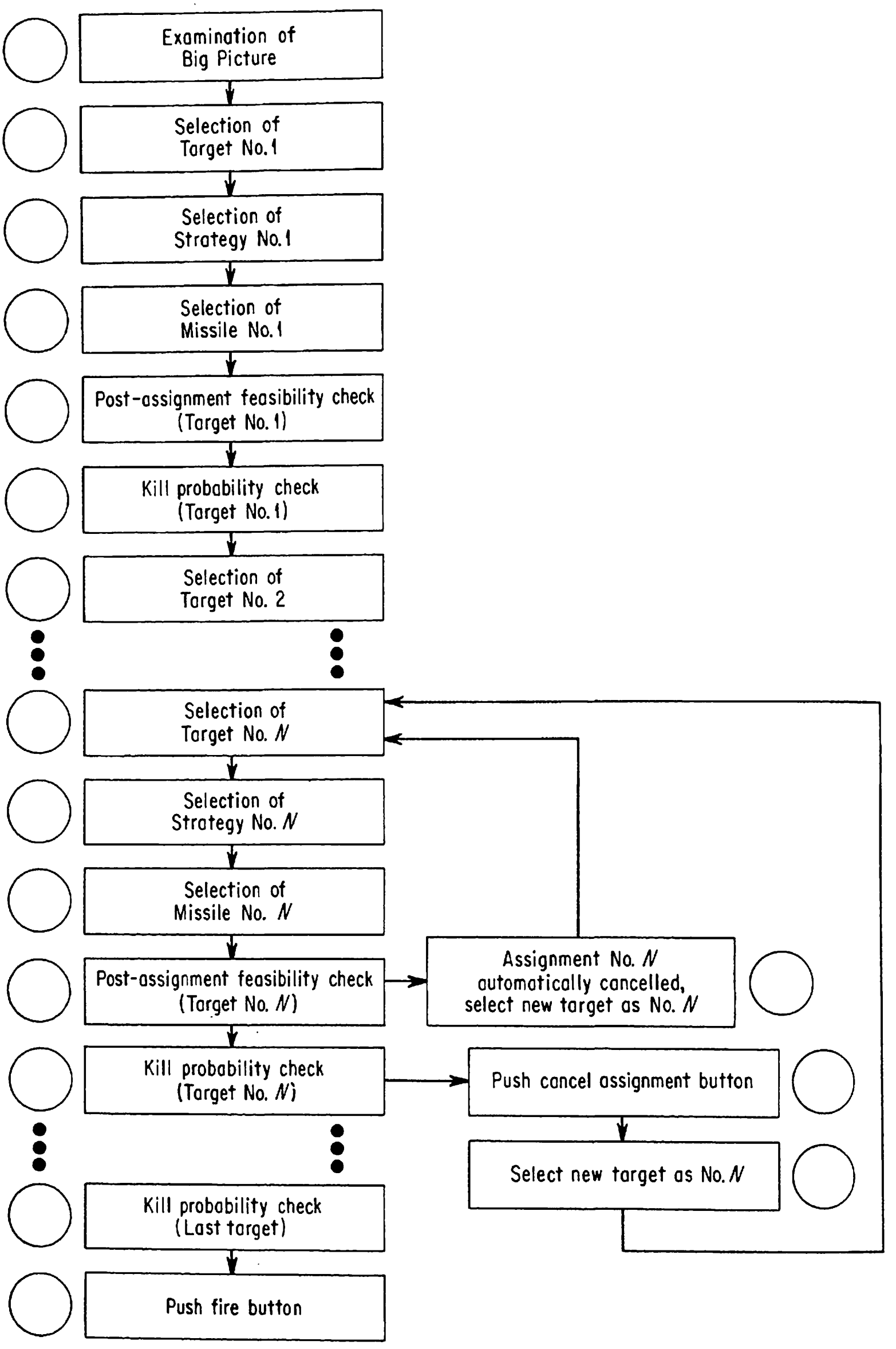


Fig. A4.3. FCO attack mode sequence.

FCO actions in the attack mode assuming a semi-automatic system. Fig. A4.4 details the display and control items potentially needed to implement the sequence of Fig. A4.3. Finally, Fig. A4.5 shows a possible design for the FCO Display/Control Console, one which incorporates good human engineering practice.

Fig. A4.4 DISPLAY/CONTROL ELEMENTS POTENTIALLY EMPLOYED ON FCO CONSOLE

Grouping	*Item description*	*Function*
Central Scope	Item A: Missile selection and assignment mode	Growth potential permitting the FCO to have optional selection between manual and automatic modes.
	Item B: Pencil controls	
	(1) Reposition	A symbol which has drifted away from the target, which it identifies, is moved back into the correct position.
	(2) Cancel assignment	Cancels the assignment of the missile which was chosen for the designated target.
	(3) Incoming accept	Acknowledges acceptance of an incoming target handover.
	(4) Outgoing	Accomplishes transmission of target in formation for handover to other CAPF.
	(5) Target data entry	Enters into the computer target data that has been punched into the target category entry keyboard.
	(6) Erase symbols	Selectively removes target identification symbols from the main display scope.
	(7) Other	For growth potential with respect to use of the pencil probe.

Fig. A4.4 DISPLAY/CONTROL ELEMENTS POTENTIALLY EMPLOYED ON FCO CONSOLE—continued

Grouping	*Item description*	*Function*
Central Scope —Continued	Item C: Pencil probe	To designate to the computer which target is the one for which action is to be taken.
	Item D: Fire	When this button is depressed and left untouched, all missiles which have been assigned are fired in the required order when the time-to-go for each missile becomes zero. A red light under the switch remains lit until all such firings have been completed, and the button stays in its depressed condition. If, at any time, the button is again touched, the firing is aborted, the red light goes out, the button is returned to its original state. This button is not used in the automatic mode.
	Item E: Target altitude	Permits selection of absolute target altitude (ambiguity which might exist when line passes through more than one target blip may be resolved through the use of the selective digital target altitude read-out as described in the proposal. Solid lines are to be employed).
	Item F: Kill probability	Provides a kill probability representation using dashed rather than solid lines.
	12″ CRT	Visual display.

Fig. A4.4 DISPLAY/CONTROL ELEMENTS POTENTIALLY EMPLOYED ON FCO CONSOLE—continued

Grouping	*Item description*	*Function*
Main Scope Controls	Item A: Brightness	To adjust to brightness (intensity of the scope presentation).
	Item B: Focus	To make adjustments in focus of the scope display. Focus is defined in terms of minimum spot size.
	Item C: Video gain	To make adjustments in the video gain (contrast) of the scope display.
	Item D: North centering	To position the center of the scope in vertical direction.
	East centering	To position the center of the scope in horizontal direction.
	Item E: Display radius	Allows the operator to select the scope display range as one of four values: 50, 150, 250, or 500 nautical miles.
	Item F: Stabilization	Permits the operator to select the ground stabilization mode of the scope display.
	Item G: Display mode	Select the video displayed as AEW information, AI information, or both.
	Item H: Target symbols displayed	Provides means for designating the class of target symbols which the fire control officer wishes displayed.

Fig. A4.4 DISPLAY/CONTROL ELEMENTS POTENTIALLY EMPLOYED ON FCO CONSOLE—continued

Grouping	*Item description*	*Function*
Auxiliary Scope Controls	Item A: Brightness	Permits the operator to adjust the brightness (intensity) of the scope presentation.
	Item B: Focus	Permits the operator to make adjustments in focus of the auxiliary scope.
	Item C: Persistence	Permits the operator to adjust (within limits) the persistence of the auxiliary scope display.
	Item D: North centering	To position the center of the scope in vertical direction.
	East centering	To position the center of the scope in horizontal direction.
	Item E: Display mode	Permits the operator to select one of three types of displays to be provided on the auxiliary scope. The three positions are labeled SIF (Security Identification Feature), EXP (expanded display), and PI (position indicator).
	5″ CRT	Visual display.
Target Information Complex	Item A: Digital read-out	Presents height, range, velocity, and threat number selectively for designated targets.
	Item B: Time-to-go	Informs FCO of the length of time until the missile should be fired as computed continuously by the tactical computer.

Fig. A4.4 DISPLAY/CONTROL ELEMENTS POTENTIALLY EMPLOYED ON FCO CONSOLE—continued

Grouping	*Item description*	*Function*
Target Information Complex—continued	Item C: Category entry keyboard	Permits the FCO to indicate various classifications of tracked targets.
	(1) Friend (2) Hostile (3) Unknown	
	Relative to target number	
	(1) Single (2) Few (3) Many	
	Relative to target type	
	(1) Air (2) Missile (3) Surface (4) Subsurface (5) Decoy (6) Chaff	
Missile Status	Armament	Specifies the type of warhead in the missile.
	Seeker	Specifies the type of seeker in the missile.
	Malfunction	The missile is malfunctioning, but the computer has determined that the malfunction(s) will not completely destroy the probability of missile effectiveness.
	Assigned	Indicates that the missile has been assigned.

Fig. A4.4 DISPLAY/CONTROL ELEMENTS POTENTIALLY EMPLOYED ON FCO CONSOLE—continued

Grouping	*Item description*	*Function*
Missile Status —continued	Ready	Indicates that a checkout program has been completed and the missile is ready to fire.
	In-flight	Indicates that the missile has already been fired and that it is at present in flight.
Missile Command Panel	Assign	Assigns the missile to the target designated by the probe and initiates the procedures necessary to prepare the missile for firing.
	Destroy	To destroy or detonate a missile which is in flight.
	Disarm	To disarm the warhead of the missile while the missile is in flight.
	Jettison	The craft jettisons the designated missile when this button is depressed.
Numerical Entry Key		Permits the FCO to enter the data-link addresses of other craft for communication of target data and associated operations, and to enter coded information for transmission to NTDS.
		(Button designations self-explanatory, with the exception of "Other" which are included for growth potential.)

Fig. A4.4 DISPLAY/CONTROL ELEMENTS POTENTIALLY EMPLOYED ON FCO CONSOLE—continued

Grouping	*Item description*	*Function*
Radar Controls	Item A: Radar mode	Selects the track or search mode of radar.
	Item B: Antenna slew	Moves the center of the radar sweep relative to the movement of the control in the vertical and horizontal planes.
	Item C: Horizontal Item D: Vertical	Control the angular limits of the radar sweep in the vertical and horizontal planes when the radar is in the search mode.

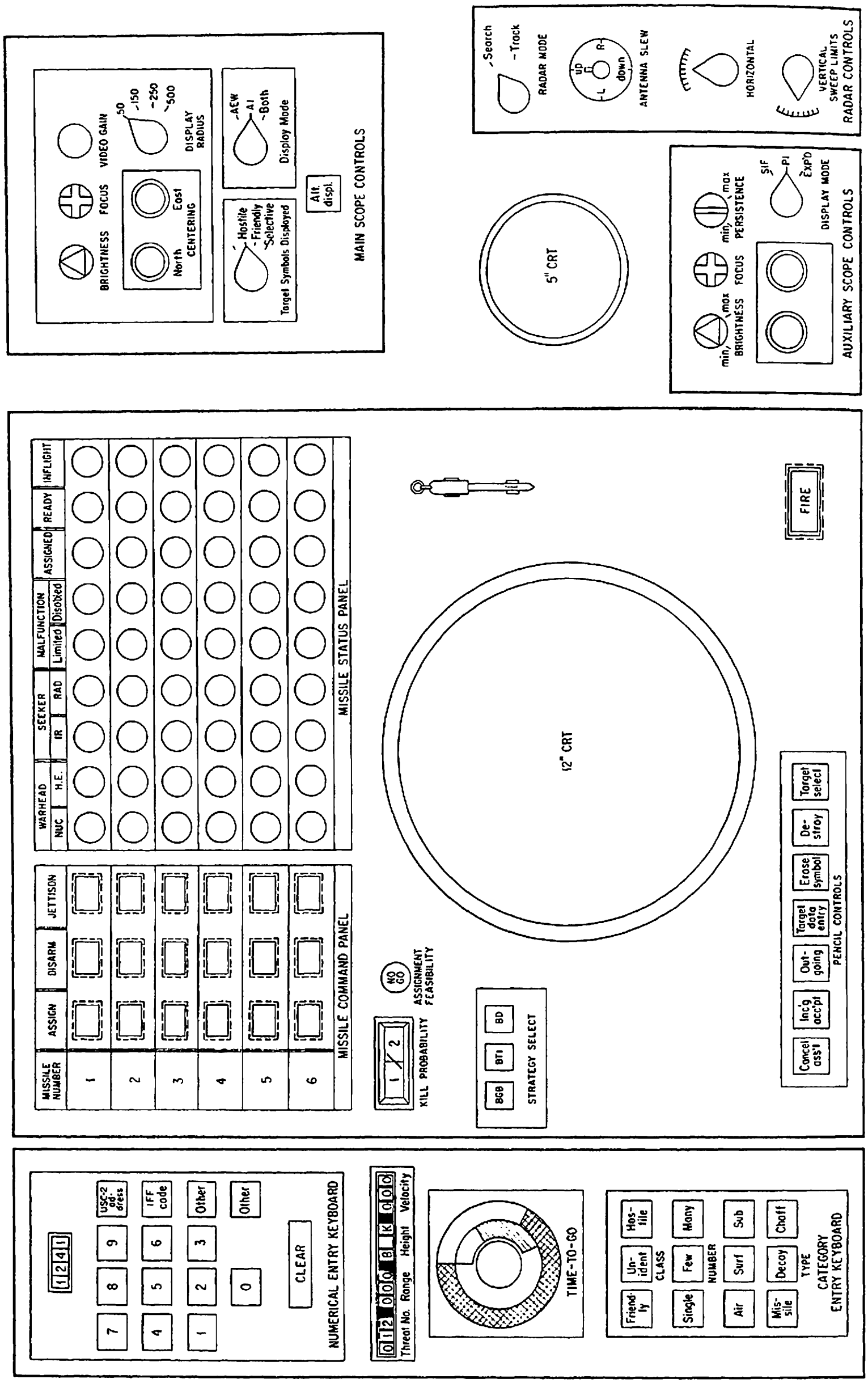


Fig. A4.5. Preliminary design of FCO console.

Appendix 5

The Systems Matrix Approach

"...A good notation has a subtlety and suggestiveness which at times make it seem almost like a live teacher."

—BERTRAND RUSSELL

1. Partial Expansion of Attack Plane Computer Example into System Matrices

A. Approach To Requirements

Any method for specifying a central control digital computer must include a coordinated approach to the requirements of the utilizing system and imply feasibility in terms of basic sciences. The elements which must be coordinated are information and its control, human engineering, display, implementation, and test. Once the approach to the computer's requirements has been established, the elements of feasibility can be investigated in terms of operational, environmental, configurational, and material requirements.

It is essential that all physical requirements be fully defined by the

over-all systems contractor as early as possible, for only then can functional and systems requirements be established and computer performance requirements settled. Design and mechanization can then logically proceed toward development of the optimum computer.

B. General Framework For The Establishment Of Requirements For Central Control Computers For Airborne Man-Machine Systems

To indicate clearly the requirements for a central control computer for airborne man-machine systems, the terms "Over-all Mission Requirements," "Systems Requirements," "Computer Performance Requirements," and "Computer Design Requirements" are introduced. Chart 1 (Fig. A5.1), a flow diagram, shows how each of the requirements above is interrelated, and how each is derived from the others. In addition, it utilizes current thinking regarding the fundamental steps for each of these basic requirements. The detailed breakdown concerns three major areas of computer design: logic, circuit design, and mechanical design. From this phase of the requirements study come the functional specifications for the computer.

Computer Physical Requirements is an approach to the establishment of the remaining portions of the central control computer requirements to provide a set of physical specifications for the computer. These requirements are derived from following a logical sequential process.

The first step is to establish the operational requirements for the computer; the second, to define the environment in which it is to operate; the third, to select the basic configuration for the organization of the computer; and the fourth and final step, to select the basic materials from which the computer can be constructed in order to meet the design requirements. Actually, this fourth item might well be placed near the top level of a development requirements structure, since the material should be selected in accordance with the functional and physical specifications generated by the systems and research requirement study presented herein.

Analysis of this framework to date has shown that it is possible to establish a set of compatible charts for the determination of requirements which meet the following objectives:

(a) The charts present the total man-machine system requirements in an easily comprehensible form.

(b) The charts are flexible enough to present the details of the system.
(c) The charts can be used to establish areas for required research.
(d) The charts show the effects of changes in the various requirements.
(e) The charts indicate accomplishments to date as the program progresses.

In structure, the charts exhibit a great many of the characteristics of matrices. Many of the mathematically-derived matrix operations will be found applicable in manipulating the information contained in these charts.

The entire assemblage of 18 charts, grouped together at the end of this appendix, is horizontally stratified in four levels. Connecting each level, there is a single chart that relates levels in a vertical fashion. Chart 1 (Fig. A5.1) indicates the relationships between Charts A through D–4 (Fig. A5.19).

Chart A (not included in the set) is called "The Over-all Mission Matrix." It should relate the set of specified mission requirements, called "Man-Machine System Source Data," to the general system, called "Man-Machine System." The structure and details of this chart should be determined by the combined efforts of the military and the over-all system contractor. The first-level charts, A–1 (Fig. A5.2) through A–4 (Fig. A5.5), further analyze the man-machine system in terms of:

A–1 Equations for basic modes of flight control
A–2 Sensors
A–3 Man-machine relationship
A–4 Actuators

Chart A/B (Fig. A5.6) correlates the first and second levels by setting up a matrix between them. Here, the "Man-Machine System" is related to the factors comprising "Computer Performance Requirements."

Over-all Mission/Missions Requirements, in the man-machine system context, is a statement of the purpose that the system is to accomplish. Specific policies and plans for the man-machine system in its tactical sense are developed. The statement lists the relevant factors for its successful accomplishment in terms of the machine and man. As noted in Chart 1: High-Performance Aircraft (Fig. A5.1), the system required by the mission requirements is further analyzed into (1) the equations for the basic modes of flight and control, (2) basic sensors, (3) basic man-machine relationships, (4) and basic actuators. Each of the requirements above can and must be broken down into the various phases from their

fundamental characteristics to their basically detailed characteristic. The noted four basic modes of requirements make it possible to establish the central control computer requirements.

The term *Central Control Computer Performance Requirements* denotes the detailed requirements for carrying out the over-all mission requirements. As such, it involves establishing the compatibility, in detail, for and between the equations, the sensors, the man-machine, and the actuators.

Computer performance requirements, therefore, is a concept that determines in a compatible way which information is relevant, how it is to be obtained, and how it is to be presented in the man-machine system in order to accomplish the requirements of the over-all mission/missions. The result of these compatibility studies is, of course, the establishment of the functional and physical design requirements for the central control computer.

Computer Functional Design Requirements relates to the manner in which information is processed by the central digital computing system.

The second level (B) breaks down the computer performance requirements in terms of:

B–1 Computer equations
B–2 Computer-sensor compatibility
B–3 Man-machine compatibility
B–4 Computer actuator compatibility

The B/C matrix relates computer performance requirements to computer parametric requirements covering the three major design parameters: language structure, communication rate, and conversion rates.

The third level (C) relates computer parametric requirements to:

C–1 Logical design
C–2 Circuit design requirements
C–3 Mechanical design requirements

Chart C/D relates the computer parametric requirements to the physical aspects of computer design (computer design requirements).

The fourth level (D) related computer parametric requirements to:

D–1 Operational requirements
D–2 Environment
D–3 Configuration
D–4 Materials

The necessary detailed specifications for development of the computer are obtained from Levels 3 and 4. Level 3 supplies the functional specifications, and Level 4 the physical specifications.

A brief discussion of the types of data that can be entered into these matrices in a coded form is presented next to indicate the extent of applicability of the model embodied in the structure of these charts.

RELATIONSHIPS. It is important to establish whether a relationship exists between any row-column intersection of a particular matrix. Where a relation exists, a "one" can be entered. Where no relation exists, a "zero" can be entered. One immediate use of this coding form would be to identify rapidly all items affected by any changes made at any level in the structure.

DATA SOURCES. In order that documented material be made available to all groups, companies, and agencies with regard to their areas of activity, a multivalued code could be used to provide the following information:

(a) No information available.
(b) Information process of preparation by Company X.
(c) Information documented in Report No. Y. (Referenced to available index.)

TARGET DATES. Coding in terms of target dates would allow compatible timing schedules to be evolved, and it would either dictate the operational target for the entire system or indicate required speed-up areas and system changes needed to meet earlier operational dates.

SOURCE OF SUPPLY. In the development of any large system, literally thousands of companies and military agencies supplying components, subassemblies, subsystems, and systems operate in an intertwined complex of contractual agreements. Coding in terms of participating companies and agencies would allow for (a) rapid informal information exchange, (b) clarification of areas of responsibility, (c) and reduction of duplicated effort.

2. Manipulation of System Matrices

Manipulation of the matrices which constitute the format for the requirements study charts can be best explained by a few simple examples.

To minimize the complexity of the examples, a fictitious pair of three-by-three matrices were created. These are:

Matrix 1 Basic transducers *vs.* major computation

Matrix 2 Major computation *vs.* display

MATRIX 1

	Major Computation		
Basic Transducers	*Navigation*	*Attitude control*	*Air data computation*
Platform	1	1	0
Doppler	1	0	0
Airspeed	1	0	1

MATRIX 2

	Display		
Major Computation	*Horizontal display*	*Vertical display*	*Auxiliary displays*
Navigation	1	1	0
Attitude control	0	1	1
Air data computation	1	1	1

Although the detail coding contained in these matrices is hypothetical, the matrices nevertheless indicate what might be a typical set of relationships existing among the various aspects of the system. For example, the over-all performance of the navigation portion of the system could certainly be related directly to the platform, Doppler, and airspeed accuracies as indicated by the three 1's under Navigation in Matrix 1.

It is desired to make a logical transformation of variables. A type of matrix multiplication is used. A brief review of matrix multiplication is presented first to indicate the essential differences between what is termed "logical" transformations as compared to arithmetic transformations.

Given two matrices, A and B, where

$$A = \begin{vmatrix} a_{11}a_{12}a_{13} \\ a_{21}a_{22}a_{23} \\ a_{31}a_{32}a_{33} \end{vmatrix} \quad \text{and} \quad B = \begin{vmatrix} b_{11}b_{12}b_{13} \\ b_{21}b_{22}b_{23} \\ b_{31}b_{32}b_{33} \end{vmatrix} \tag{38}$$

and multiplication is defined by

$$A \times B =$$

$$\begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}, & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32}, & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}, & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32}, & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31}, & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32}, & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{vmatrix}. \tag{39}$$

Consider the term $a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$. An *arithmetic* operation is indicated. If the logical operations of "and" and "or" are used instead, the "and" operation replacing multiplication and the "or" operation replacing addition, then a logical matrix multiplication by definition results. (This is, in fact, the well-known Boolean matrix multiplication.) It should be recalled that

AND

$$\begin{aligned} a \cdot b &= 1 \text{ for } a = 1, b = 1 \\ a \cdot b &= 0 \text{ for } a = 0, b = 0 \\ a &= 1, b = 0 \\ a &= 0, b = 1; \end{aligned} \tag{40}$$

OR

$$\begin{aligned} a + b &= 1 \text{ for } a = 1, b = 1 \\ a &= 1, b = 0 \\ a &= 0, b = 1 \\ a + b &= 0 \text{ for } a = 0, b = 0. \end{aligned} \tag{41}$$

Applying these rules to matrices 1 and 2, the result is | Matrix 1 | × | Matrix 2 | =

MATRIX 3

Basic Transducer	Display *Horizontal display*	*Vertical display*	*Auxiliary displays*
Platform	1	1	1
Doppler	1	1	0
Airspeed	1	1	1

The resulting matrix indicates the incidence relations between the three instruments and the three displays. It is also interesting to take the arithmetic product.

| Matrix 1 | × | Matrix 2 | =

MATRIX 4

Basic Transducer	Display *Horizontal display*	*Vertical display*	*Auxiliary displays*
Platform	1	2	1
Doppler	1	1	0
Airspeed	2	2	1

Some indication of the number of ways in which an instrument affects a particular display is afforded. For example, the platform affects the vertical display in two ways:

(a) Through the navigation computation,
(b) Through the attitude control computation.

The airspeed measurement affects the vertical display in two ways:

(a) Through the navigation computation,
(b) Through the attitude control computation.

Next, a slight change will be made in Matrix 1, and its effect on the product matrix examined. If the navigation computation is made independently of the airspeed measurement, Matrix 1 becomes

MATRIX 1′

Basic Computation	Major Computation *Navigation*	*Attitude control*	*Computation*
Platform	1	1	0
Doppler	1	0	0
Airspeed	0	0	1

The new logical product matrix becomes

MATRIX 3′

	Navigation	*Vertical*	*Auxiliary*
Platform	0	1	1
Doppler	1	1	0
Airspeed	1	1	1

It should be observed that no logical change has been encountered. The arithmetic matrix indicates the extent of the modification.

MATRIX 4′

	Navigation	*Vertical*	*Auxiliary*
Platform	1	2	1
Doppler	1	1	0
Airspeed	1	1	1

The net effect of the change in Matrix 1 is now seen to be a reduction in the number of ways in which the airspeed affects the horizontal and vertical display. This is a desirable condition for an instrument that does not have high precision.

Assuming that in Matrix 2 the displays are the independent variables and the major computations are dependent variables, and that in Matrix 1 the major computations are independent with the instruments as dependent variables, it is seen that the arithmetic product matrix describes the set of equations relating the displays as independent variables to the instruments as dependent variables.

If the 1, 0 coefficients used in Matrices 1 and 2 are accepted for the moment as being the actual weighting coefficients, then a matrix inversion can be performed on the arithmetic product matrix to "solve" for instruments in terms of the displays. The results are:

(a) Platform $\cong$ 2 Vertical − Horizontal†

(b) Doppler $\cong$ Horizontal − Auxiliary (42)

(c) Airspeed $\cong$ Horizontal − Vertical + Auxiliary

Some interesting conclusions can be drawn from the three equations. With the system as described, increased requirements on the horizontal display, holding the requirements on the auxiliary and vertical displays constant, will be reflected as increased requirements on the Doppler and airspeed but will decrease the requirements on the platform. Increased vertical display requirements will cause a large increase in requirements on the platform, will not change the Doppler requirements, and will decrease the requirements on the airspeed. Finally, increased requirements on the auxiliary display will produce increased airspeed requirements, decreased Doppler requirements, and will leave the platform unaffected.

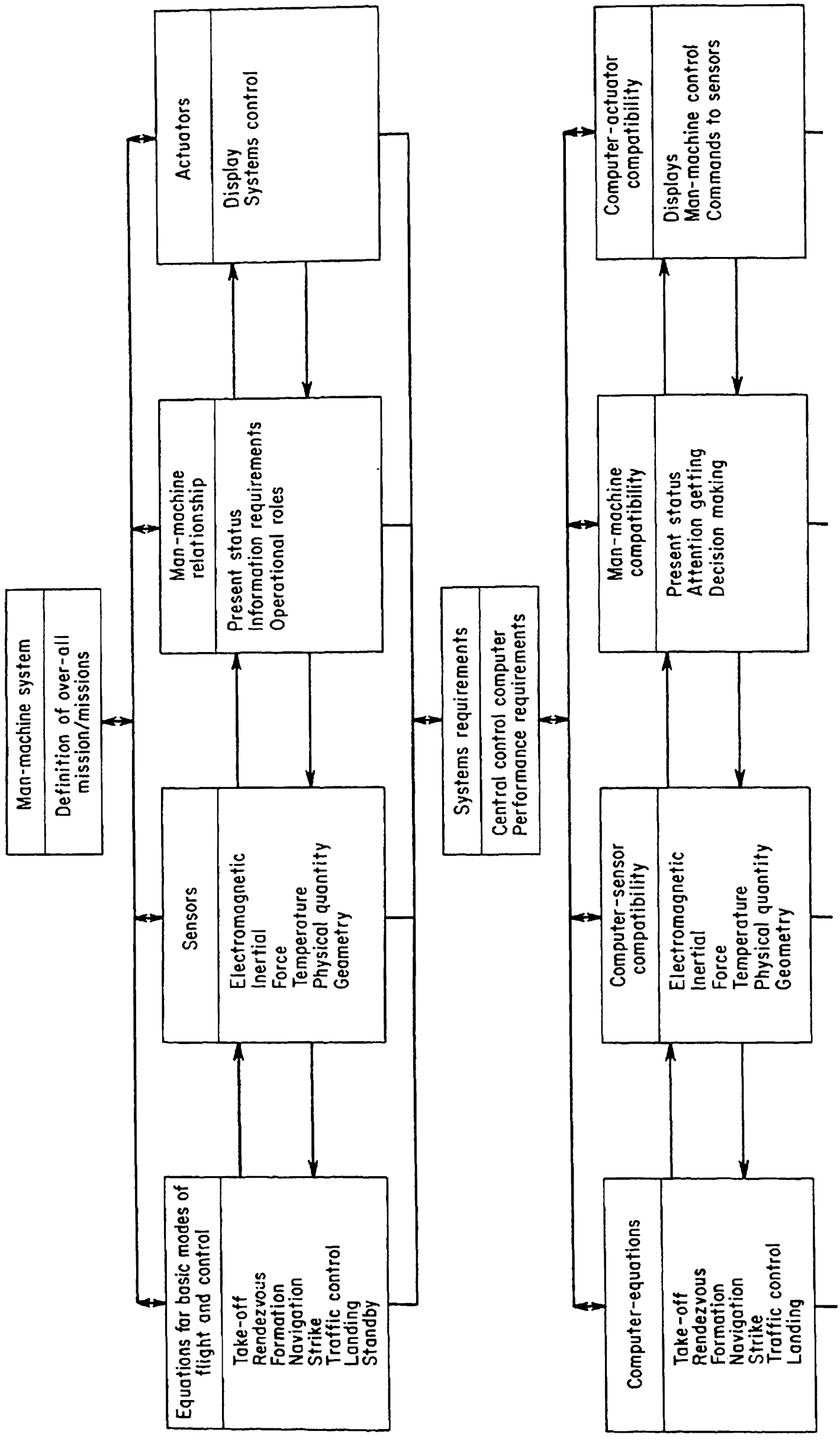
The results may appear rather odd until they are examined in light of the constraints imposed on the equation in demanding that no changes in requirements be allowed two of the variables while the third is being varied. If Matrices 1 and 2 are multiplied by a weighting matrix, which would indicate the relative effects that the various factors had upon each other, a still more useful product matrix and inverted product matrix would result. It is the weighting matrix that requires intense study to develop the concept fully into a working tool.

† The symbol $\cong$ indicates a qualitative relationship.

3. Charts Showing Partial Expansion Examples

The following charts show partial expansion of computer examples into matrices. (The first example noted has been described in the preceding text):

(a) Attack plane computer into matrices, (Figs. A5.1–A5.19).
(b) Submarine command computer into system matrices (upper levels only), (Figs. A5.20–A5.24).
(c) Manned lunar command computer into matrices (upper level only), (Figs. A5.25–A5.29).


Man-machine system
Definition of over-all mission/missions
Equations for basic modes of flight and control
Take-off
Rendezvous
Formation
Navigation
Strike
Traffic control
Landing
Standby
Sensors
Electromagnetic
Inertial
Force
Temperature
Physical quantity
Geometry
Man-machine relationship
Present status
Information requirements
Operational roles
Actuators
Display
Systems control
Systems requirements
Central control computer
Performance requirements
Computer-equations
Take-off
Rendezvous
Formation
Navigation
Strike
Traffic control
Landing
Computer-sensor compatibility
Electromagnetic
Inertial
Force
Temperature
Physical quantity
Geometry
Man-machine compatibility
Present status
Attention getting
Decision making
Computer-actuator compatibility
Displays
Man-machine control
Commands to sensors

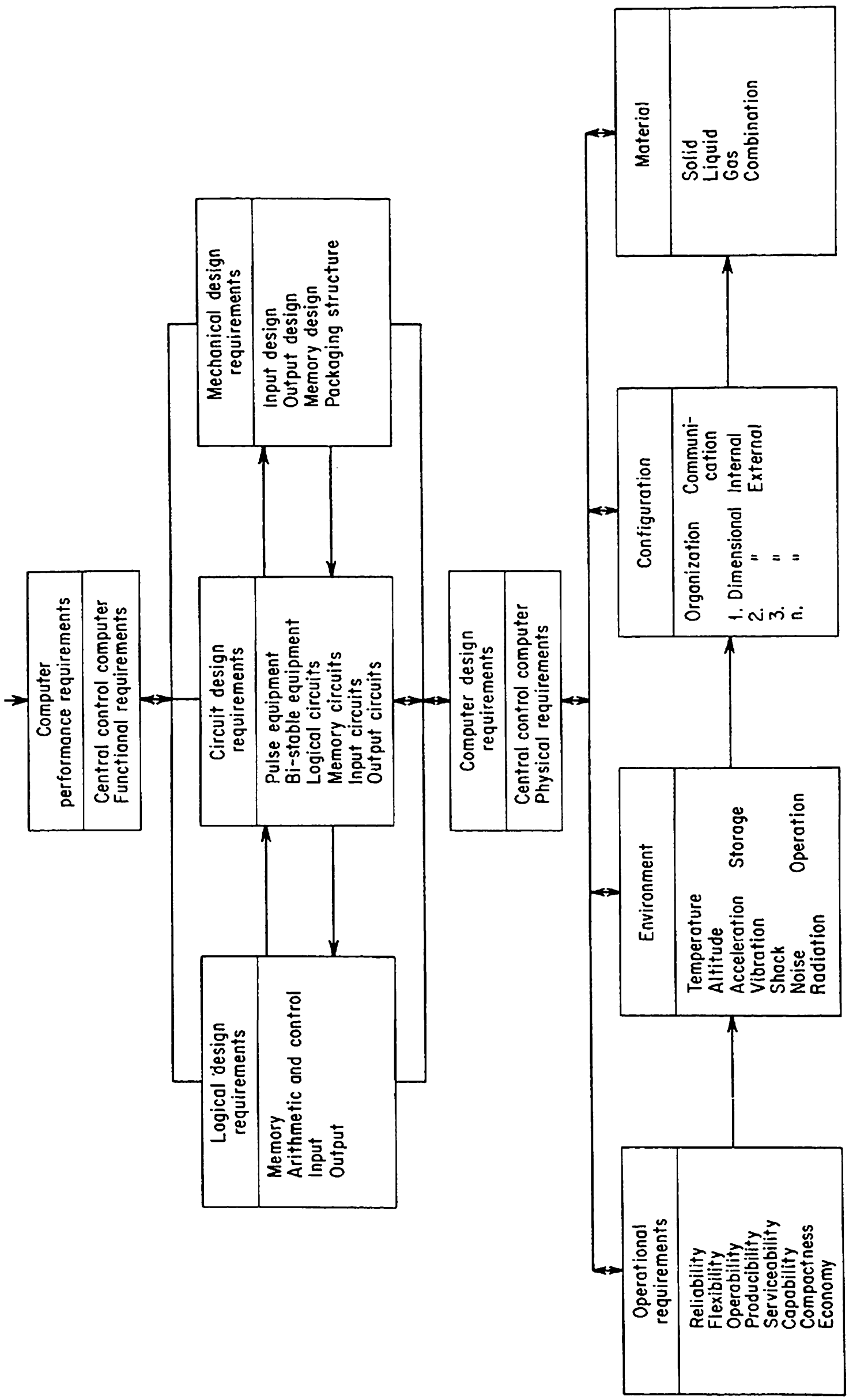


Fig. A5.1. Chart 1 (high-performance aircraft).

Chart A-1:

Advanced Program Systems / *Equations for Phases of Flight and Control*			*Auto control*	*Power plant control*	*Arm. control*	*Auxiliary functions*
			Flight path stabilization with pilot maneuvering Transonic pitch trim change compensation Pitch damping Yaw damping Fire-control tracking Landing tracking Formation flight tracking Cruise and rendezvous tracking Traffic control and external command tracking Altitude or rate of climb hold or programming Mach number hold or programming Cruise cont. tracking for opt. spec. range or end Pilot manual control as a standby	Maximum specific range or specific endurance Fire control (combat conditions) Take-off Mach number hold Altitude hold Formation hold Traffic control—holding pattern	Weapon activation Weapon slaving Weapon release	Landing gear Inertial navigation system control Arresting hook Speed brakes Cockpit and equipment environment control
Take-off	Carrier					
	Shore					
Landing	Carrier					
	Shore					
Traffic control	Noncombat					
	Combat					
Strike	Air-Ground					
	Air-Air					
Rendezvous						
Formation						
Navigation	Assigned alt.	Max E				
		Max R				
	Optimum	Max E				
		Max R				
	Assigned alt. and air-speed					

Fig. A5.2

Vertical display function generator	*Horizontal display function generator*	*Other display commands*	*Communication and identification control*	*Air data*	*Navigation*
Basic contact analog display Impact point Command or predicted flight path Runway symbol Take-off decision point Air target symbols Ground target Firing point symbol Rendezvous and formation symbol Traffic control Landing flight path Runway indication	Present position Aircraft heading Angle between A/C heading and true ground track Position of other aircraft and ships Ground track angles of each target Coding selection for other A/C (alt., type, & ident.) Range prediction Flight plan marker Point of return signals	Output to engine condition indicator Warning temperature indicators Estimated time of arrival indicator Endurance indicator Mode of activation switch	Voice of communication TACAN ADF IFF & SIF Radar	Temperature Pressure Density Relative wind Absolute wind	Inertial Doppler Decca

Chart A-2:

Advanced Program Systems		*Auto control*	*Power plant control*	*Arm. control*	*Auxiliary functions*
Sensors		Flight path stabilization with pilot maneuvering Transonic pitch trim change compensation Pitch damping Yaw damping Fire-control tracking Landing tracking Formation flight tracking Cruise and rendezvous tracking Traffic control and external command tracking Altitude or rate of climb hold or programming Mach number hold or programming Cruise cont. tracking for opt. spec. range or end Pilot manual control as a standby	Maximum specific range or specific endurance Fire control (combat conditions) Take-off Mach number hold Altitude hold Formation hold Traffic control—holding pattern	Weapon activation Weapon slaving Weapon release	Landing gear Inertial navigation system control Arresting hook Speed brakes Cockpit and equipment environment control
Electro-magnetic	Audio transceiver				
	Data link				
	IFF				
	Tracking radar				
Inertial	Stable platforms				
	Accelerometers				
	Rate gyros				
Force	Pressure detectors				
	Airstream direction detectors				
	Elastically restrained synchros				
Temperature	Total temp. transducer				
	Thermocouples & temp. bulbs				
Geometry	Synchros				
	Mechanical angle detectors				
	Switches				
Physical quantity	Buoyant probes				
	Capacitance probes				
	Acoustic gages				
	Digital gages				

Fig. A5.3

Vertical display function generator	Basic contact analog display Impact point Command or predicted flight path Runway symbol Take-off decision point Air target symbols Ground target Firing point symbol Rendezvous and formation symbol Traffic control Landing flight path Runway indication
Horizontal display function generator	Present position Aircraft heading Angle between A/C heading and true ground track Position of other aircraft and ships Ground track angles of each target Coding selection for other A/C (alt., type, & ident.) Range prediction Flight plan marker Point of return signals
Other display commands	Output to engine condition indicator Warning temperature indicators Estimated time of arrival indicator Endurance indicator Mode of activation switch
Communication and identification control	Voice of communication TACAN ADF IFF & SIF Radar
Air data	Temperature Pressure Density Relative wind Absolute wind
Navigation	Inertial Doppler Decca

Chart A-3:

Advanced Program Systems		*Auto control*	*Power plant control*	*Arm. control*	*Auxiliary functions*
Man-Machine Relationship		Flight path stabilization with pilot maneuvering Transonic pitch trim change compensation Pitch damping Yaw damping Fire-control tracking Landing tracking Formation flight tracking Cruise and rendezvous tracking Traffic control and external command tracking Altitude or rate of climb hold or programming Mach number hold or programming Cruise cont. tracking for opt. spec. range or end Pilot manual control as a standby	Maximum specific range or specific endurance Fire control (combat conditions) Take-off Mach number hold Altitude hold Formation hold Traffic control—holding pattern	Weapon activation Weapon slaving Weapon release	Landing gear Inertial navigation system control Arresting hook Speed brakes Cockpit and equipment environment control
Info for present status or modes of operation	Normal				
	Stand-by				
	Emergency				
Attention-getting information requirements	Normal				
	Standby				
	Damage control info.				
Operational decision roles	Prediction				
	Reprogrammer				
	Appraiser				
	Controller				

Fig. A5.4

HIGH-PERFORMANCE AIRCRAFT.

Vertical display function generator	Basic contact analog display Impact point Command or predicted flight path Runway symbol Take-off decision point Air target symbols Ground target Firing point symbol Rendezvous and formation symbol Traffic control Landing flight path Runway indication										
Horizontal display function generator	Present position Aircraft heading Angle between A/C heading and true ground track Position of other aircraft and ships Ground track angles of each target Coding selection for other A/C (alt., type, & ident.) Range prediction Flight plan marker Point of return signals										
Other display commands	Output to engine condition indicator Warning temperature indicators Estimated time of arrival indicator Endurance indicator Mode of activation switch										
Communication and identification control	Voice of communication TACAN ADF IFF & SIF Radar										
Air data	Temperature Pressure Density Relative wind Absolute wind										
Navigation	Inertial Doppler Decca										

Chart A-4:

Advanced Program Systems		*Auto control*	*Power plant control*	*Arm. control*	*Auxiliary functions*
Actuators		Flight path stabilization with pilot maneuvering Transonic pitch trim change compensation Pitch damping Yaw damping Fire-control tracking Landing tracking Formation flight tracking Cruise and rendezvous tracking Traffic control and external command Altitude or rate of climb hold or programming Mach number hold or programming Cruise cont. tracking for opt. spec. range or end Pilot manual control as a standby	Maximum specific range or specific endurance Fire control (combat conditions) Take-off Mach number hold Altitude hold Formation hold Traffic control—holding pattern	Weapon activation Weapon slaving Weapon release	Landing gear Inertial navigation system control Arresting hook Speed brakes Cockpit and equipment environment control
Display amplifiers	Vertical				
	Horizontal				
	Eta & endurance				
	Power condition				
	Auxiliary: Warning signals				
	Auxiliary: Stand-by instrumentation				
System control	Man				
	Machine				
	External				
Mechanisms amplifiers	Power control				
	Armament control				
	Flight path control				
	Mech. adv. control				
	Aux. function				

Fig. A5.5

Vertical display function generator	Basic contact analog display Impact point Command or predicted flight path Runway symbol Take-off decision point Air target symbols Ground target Firing point symbol Rendezvous and formation symbol Traffic control Landing flight path Runway indication														
Horizontal display function generator	Present position Aircraft heading Angle between A/C heading and true ground track Position of other aircraft and ships Ground track angles of each target Coding selection for other A/C (alt., type, & ident.) Range prediction Flight plan marker Point of return signals														
Other display commands	Output to engine condition indicator Warning temperature indicators Estimated time of arrival indicator Endurance indicator Mode of activation switch														
Communication and identification control	Voice of communication TACAN ADF IFF & SIF Radar														
Air data	Temperature Pressure Density Relative wind Absolute wind														
Navigation	Inertial Doppler Decca														

Chart A/B:

Advanced Program Systems	*Auto control*	*Power plant control*	*Arm. control*	*Auxiliary functions*
Computer requirements	Flight path stabilization with pilot maneuvering Transonic pitch trim change compensation Pitch damping Yaw damping Fire-control tracking Landing tracking Formation flight tracking Cruise and rendezvous tracking Traffic control and external command tracking Altitude or rate of climb hold or programming Mach number hold or programming Cruise cont. tracking for opt. spec. range or end Pilot manual control as a standby	Maximum specific range or specific endurance Fire control (combat conditions) Take-off Mach number hold Altitude hold Formation hold Traffic control—holding pattern	Weapon activation Weapon slaving Weapon release	Landing gear Inertial navigation system control Arresting hook Speed brakes Cockpit and equipment environment control
Range of variation				
Accuracy				
Maximum derivation				
Sampling period				
Sampling rate				

Fig. A5.6

HIGH-PERFORMANCE AIRCRAFT.

					Basic contact analog display Impact point Command or predicted flight path Runway symbol Take-off decision point Air target symbols Ground target Firing point symbol Rendezvous and formation symbol Traffic control Landing flight path Runway indication	*Vertical display function generator*
					Present position Aircraft heading Angle between A/C heading and true ground track Position of other aircraft and ships Ground track angles of each target Coding selection for other A/C (alt., type, & ident.) Range prediction Flight plan marker Point of return signals	*Horizontal display function generator*
					Output to engine condition indicator Warning temperature indicators Estimated time of arrival indicator Endurance indicator Mode of activation switch	*Other display commands*
					Voice of communication TACAN ADF IFF & SIF Radar	*Communication and identification control*
					Temperature Pressure Density Relative wind Absolute wind	*Air data*
					Inertial Doppler Decca	*Navigation*

Chart B-1: HIGH-PERFORMANCE AIRCRAFT.

Computer Equations	*Computer Requirements*		*Range of variables*	*Accuracy*	*Maximum derivatives*	*Sampling period*	*Sampling rate*
Take-off	Carrier						
	Shore						
Landing	Carrier						
	Shore						
Traffic control	Noncombat						
	Combat						
Strike	Air-Ground						
	Air-Air						
Rendezvous							
Formation							
Navigation	Assigned altitude	Max E					
		Max R					
	Optimum	Max E					
		Max R					
	Assigned altitude and air speed						

Fig. A5.7

Chart B-2: HIGH-PERFORMANCE AIRCRAFT.

Computer-Sensor Compatibility \ *Computer Requirements*		*Range of variables*	*Accuracy*	*Maximum derivatives*	*Sampling period*	*Sampling rate*
Electro-magnetic	Audio transceiver					
	Data link					
	IFF					
	Tracking radar					
Inertial	Stable platforms					
	Accelerometers					
	Rate gyros					
Force	Pressure detectors					
	Airstream direction detectors					
	Elastically restrained synchros					
Temperature	Total temperature transducer					
	Thermocouples & temperature bulbs					
Geometry	Synchros					
	Mechanical angle detector					
	Switches					
Physical quantity	Buoyance probes					
	Capacitance probes					
	Acoustic gages					
	Digital gages					

Fig. A5.8

Chart B-3: HIGH-PERFORMANCE AIRCRAFT.

Man-Machine compatibility	*Computer Requirements*	*Range of variables*	*Accuracy*	*Maximum derivatives*	*Sampling period*	*Sampling rate*
Present status	Normal					
	Normal					
	Emergency					
Information requirements	Orientation information					
	Quantitative information					
	Director information					
Operational rule	Decision maker					
	Reprogrammer					
	Appraiser					
	Controller					

Fig. A5.9

Chart B-4: HIGH-PERFORMANCE AIRCRAFT.

Computer-Actuator Compatibility \ *Computer Requirements*		*Range of variables*	*Accuracy*	*Maximum derivatives*	*Sampling period*	*Sampling rate*
Display amplifiers	Vertical					
	Horizontal					
	Eta & endurance					
	Power condition					
	Warning signals					
	Stand-by instrumentation					
System control	Man					
	Machine					
	External					
Mechanical amplifiers	Power control					
	Armament control					
	Flight path control					
	Mech. adv. control					
	Aux. function control					

Fig. A5.10

Chart B/C: High-Performance Aircraft.

Computer Performance Requirements / *Computer requirements*	*Range of variables*	*Accuracy*	*Maximum derivatives*	*Sampling period*	*Sampling rate*
Language structure Word length					
Coding					
Memory structure					
Sequence of operations					
Scaling					
Communication rate Timing of operations					
Basic bit rate					
Basic slewing rates					
Conversion rates *Read-In* From man					
From sensors					
From data link					
Read-Out To man					
To actuator					
To communications					

Fig. A5.11

Chart C-1: HIGH-PERFORMANCE AIRCRAFT.

Computer Requirements / *Logical Design Requirements*	*Language structure*					*Communication rate*			*Conversion rate*					
	Word length	Coding	Memory structure	Sequence of operations	Scaling	Timing of operations	Basic bit rate	Basic slewing rate	*Read-In* From man	From sensors	From data link	*Read-Out* To man	To actuator	To communications
Memory														
Volatile memory														
Recirculating														
Destructive														
Non-volatile memory														
Fixed														
Alterable														
Arithmetic & control														
Program structure														
Operational functions														
Input														
Buffering requirements														
Input mode sequences														
Address inputs														
Outputs														
Buffering requirements														
Output mode sequences														

Fig. A5.12

Chart C-2: HIGH-PERFORMANCE AIRCRAFT.

Computer Requirements / *Circuit Requirements*	*Language structure*					*Communication rate*			*Conversion rate*					
	Word length	Coding	Memory structure	Sequence of operations	Scaling	Timing of operations	Basic bit rate	Basic slewing rate	*Read-In* From man	From sensors	From data link	*Read-Out* To man	To actuator	To communications
Pulse equipment														
Bi-stable equipment														
Logical circuits														
Memory circuits														
Input circuits														
Output circuits														

Fig. A5.13

Chart C-3: High-Performance Aircraft.

Computer Requirements / *Mechanical Design Requirements*	*Language structure*					*Communication rate*			*Conversion rate*					
	Word length	Coding	Memory structure	Sequence of operations	Scaling	Timing of operations	Basic bit rate	Basic slewing rate	*Read-In* From man	From sensors	From data link	*Read-Out* To man	To actuator	To communications
Input design														
Output design														
Memory design														
Packaging structure														

Fig. A5.14

Chart C/D: HIGH-PERFORMANCE AIRCRAFT.

Computer Requirements / *Computer Design Requirements*	*Language structure*					*Communication rate*			*Conversion rate*					
	Word length	Coding	Memory structure	Sequence of operations	Scaling	Timing of operations	Basic bit rates	Basic slewing rates	*Read-In* From man	From sensors	From data link	*Read-Out* To man	To actuator	To communications
Logical design requirements														
Memory														
Volatile memory														
Recirculating														
Destructive														
Non-volatile memory														
Fixed														
Alterable														
Arithmetic & control														
Program structure														
Operational functions														
Inputs														
Buffering requirements														
Input mode sequences														
Address inputs														
Outputs														
Buffering requirements														
Output mode sequences														

Circuit design requirements														
Pulse equipment														
Bi-stable equipment														
Logical circuits														
Memory circuits														
Input circuits														
Output circuits														
Mechanical design requirements														
Input design														
Output design														
Memory design														
Packaging structure														

Fig. A5.15

Chart D-1:

Computer Design Requirements / *Operational Requirements*	*Logical*							
	Memory Volatile memory	Recirculating	Destructive	Non-volatile memory	Fixed	Alterable	*Arithmetic & Control* Program structure	Operational functions
Reliability								
Flexibility								
Operability								
Producibility								
Serviceability								
Capability								
Compactness								
Economy								

Fig. A5.16

HIGH-PERFORMANCE AIRCRAFT.

Design				*Circuit design*						*Mechanical design*			
Inputs Buffering requirements	Input mode sequences	*Outputs* Buffering requirements	Output mode sequences	Pulse equipment	Bi-stable equipment	Logical circuits	Memory circuits	Input circuits	Output circuits	Input design	Output design	Memory design	Packaging structure

Chart D-2:

Computer Design Requirements	*Logical*							
Environment	*Memory* Volatile memory	Recirculating	Destructive	Non-volatile memory	Fixed	Alterable	*Arithmetic & Control* Program structure	Operational functions
Temperature								
Altitude								
Acceleration								
Vibration								
Shock								
Noise								
Radiation								

Fig. A5.17

HIGH-PERFORMANCE AIRCRAFT.

Design				*Circuit design*						*Mechanical design*			
Inputs Buffering requirements	Input mode sequences Address inputs	*Outputs* Buffering requirements	Output mode sequences	Pulse equipment	Bi-stable equipment	Logical circuits	Memory circuits	Input circuits	Output circuits	Input design	Output design	Memory design	Packaging structure

Chart D-3:

Computer Design Requirements		*Logical*							
Configuration		*Memory* Volatile memory	Recirculating	Destructive	Non-volatile memory	Fixed	Alterable	*Arithmetic & Control* Program structure	Operational functions
Communication	External								
	Internal								
Organization	*n*-dimensional								
	three-dimensional								
	two-dimensional								
	one-dimensional								

Fig. A5.18

HIGH-PERFORMANCE AIRCRAFT.

Design				*Circuit design*						*Mechanical design*			
Inputs Buffering requirements	Input mode sequences Address inputs	*Outputs* Buffering requirements	Output mode sequences	Pulse equipment	Bi-stable equipment	Logical circuits	Memory circuits	Input circuits	Output circuits	Input design	Output design	Memory design	Packaging structure

Chart D-4:

Computer Design Requirements	*Logical*							
Materials	*Memory* Volatile memory	Recirculating	Destructive	Non-volatile memory	Fixed	Alterable	*Arithmetic & Control* Program structure	Operational functions
Solid								
Liquid								
Gas								
Combination								

Fig. A5.19

HIGH-PERFORMANCE AIRCRAFT.

Design				*Circuit design*						*Mechanical design*			
Inputs Buffering requirements	Input mode sequences Address inputs	*Outputs* Buffering requirements	Output mode sequences	Pulse equipment	Bi-stable equipment	Logical circuits	Memory circuits	Input circuits	Output circuits	Input design	Output design	Memory design	Packaging structure

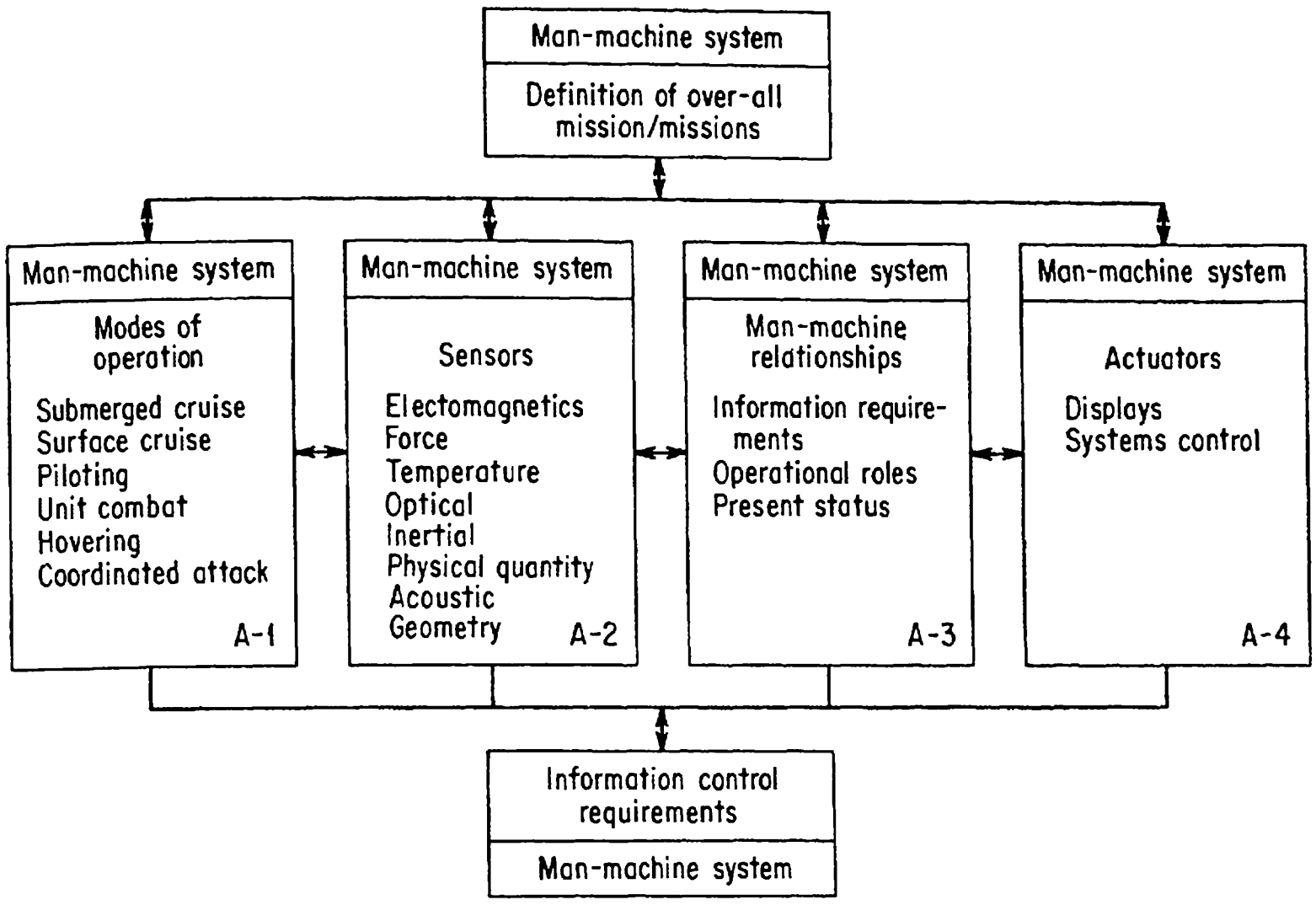


Fig. A5.20. Chart 1 (submarine).

Chart A-1:

Man-machine system		*Environmental data*																								*Engineering*						
		External											Internal																		
Modes of operation		Temperature	Temperature gradient	Salinity	Pressure	Density	Currents	Ship/Ocean noise level	Ice cover	Sea state	Climate and weather	Humidity	Oxygen	Temperature	Radiation	Casualties	Pressure	Reactor	Heat exchanger	Turbine(s)	Generator(s)	Propulsion motors	Hydraulic pressure	Pneumatic pressure	Auxiliaries
Submerged cruise																															
Surface cruise																															
Piloting	Entry																														
	Sortie																														
	Maneuvering																														
Diving																															
Surfacing																															
Unit combat																															
Hovering																															
Coordinated attack																															

Fig. A5.21

										Direct voice	*Communications*
										Radio	
										Underwater telephones	
										PA systems	
										Interstation telephones	
										Annunciator	
										Weapon status	*Armament*
										Weapon slaving	
										Weapon activation	
										Weapon launch	
										Weapon tracking	
										Target tracking and quickening	
										Target identification	
										Hit probability	
										Heading	*Submarine: local environment relationships*
										Yaw	
										Turn rate	
										Roll and roll rate	
										Pitch and pitch rate	
										Depth and depth rate	
										Bottom separation	
										Bouyancy	
										Trim	
										Present position	*Navigation data*
										Probable position error	
										Command course	
										Speed	
										ETA	
										Obstacles	
										Other ships	

Chart A-2:

Man-machine system		*Environmental data*																*Engineering*							
		External										Internal													
Sensors		Temperature	Temperature gradient	Salinity	Pressure	Density	Currents	Ship/Ocean noise level	Ice cover	Sea state	Climate and weather	Humidity	Oxygen	Temperature	Radiation	Casualties	Pressure	Reactor	Heat exchanger	Turbine(s)	Generator(s)	Propulsion motors	Hydraulic pressure	Pneumatic pressure	Auxiliaries
Electro-magnetic	Radar																								
	TV																								
	IFF																								
	Radiation monitor																								
	Compass																								
	Magnetometer																								
	Pyrheliometer																								
Force	Pressure detector																								
	Ocean current direction detector																								
	Cavitation indicator																								
	Pitometer																								
	Anemometer																								
Tempera-ture	Bathythermograph																								
	Thermocouples																								
	Temperature bulbs																								
Optical	Astro-camera																								
	Star-tracker																								
	Periscope																								
	Interferometer																								
Inertial	Stable platforms																								
	Accelerometers																								
	Rate gyros																								
	Gravimeter																								
Physical quantity	Buoyant probes																								
	Digital gages																								
	Hydrograph																								
	Conductrometric instrument																								
Acoustic	Sonar/Transducer																								
	Fathometer																								
	Acoustic gages																								
Geometry	Gas indicators																								
	Synchros																								
	Stadiometer																								
	Mechanical angle detectors																								

Fig. A5.22

Communications	Direct voice
	Radio
	Underwater telephones
	PA systems
	Interstation telephones
	Annunciator
Armament	Weapon status
	Weapon slaving
	Weapon activation
	Weapon launch
	Weapon tracking
	Target tracking and quickening
	Target identification
	Hit probability
Submarine: local environment relationships	Heading
	Yaw
	Turn rate
	Roll and roll rate
	Pitch and pitch rate
	Depth and depth rate
	Bottom separation
	Bouyancy
	Trim
Navigation data	Present position
	Probable position error
	Command course
	Speed
	ETA
	Obstacles
	Other ships

Chart A-3:

Man-machine system		*Environmental data*															*Engineering*								
		External										Internal													
Man-machine relationships		Temperature	Temperature gradient	Salinity	Pressure	Density	Currents	Ship/Ocean noise level	Ice cover	Sea state	Climate and weather	Humidity	Oxygen	Temperature	Radiation	Casualties	Pressure	Reactor	Heat exchanger	Turbine(s)	Generator(s)	Propulsion motors	Hydraulic pressure	Pneumatic pressure	Auxiliaries
Information for present status or modes of operation	Normal																								
	Standby																								
	Emergency																								
Attention-getting information requirements	Normal																								
	Standby																								
	Damage control information																								
Operational decision roles	Prediction																								
	Repro-grammer																								
	Appraiser																								
	Controller																								

Fig. A5.23

SUBMARINE.

Communications						*Armament*						*Submarine: local environment relationships*									*Navigation data*					
Direct voice	Radio	Underwater telephones	PA systems	Interstation telephones	Annunciator	Weapon status	Weapon slaving	Weapon activation	Weapon launch	Weapon tracking	Target tracking and quickening	Target identification	Hit probability	Heading	Yaw	Turn rate	Roll and roll rate	Pitch and pitch rate	Depth and depth rate	Bottom separation	Bouyancy	Trim	Present position	Probable position error	Command course	Speed	ETA	Obstacles	Other ships

Chart A-4:

Man-machine system		*Environmental data*																*Engineering*							
		External										Internal													
Actuators		Temperature	Temperature gradient	Salinity	Pressure	Density	Currents	Ship/Ocean noise level	Ice cover	Sea state	Climate and weather	Humidity	Oxygen	Temperature	Radiation	Casualties	Pressure	Reactor	Heat exchanger	Turbine(s)	Generator(s)	Propulsion motors	Hydraulic pressure	Pneumatic pressure	Auxiliaries
Display amplifiers	Vertical																								
	Horizontal																								
	ETA & endurance																								
	Power condition																								
	Warning signals																								
	Standby instrumentation																								
System control	Man																								
	Machine																								
	External																								
Mechanisms amplifiers	Power control																								
	Armament control																								
	Cruise path control																								
	Mechanical advantage control																								
	Auxiliary functions control																								

Fig. A5.24

														Direct voice	*Communications*
														Radio	
														Underwater telephones	
														PA systems	
														Interstation telephones	
														Annunciator	
														Weapon status	*Armament*
														Weapon slaving	
														Weapon activation	
														Weapon launch	
														Weapon tracking	
														Target tracking and quickening	
														Target identification	
														Hit probability	
														Heading	*Submarine: local environment relationships*
														Yaw	
														Turn rate	
														Roll and roll rate	
														Pitch and pitch rate	
														Depth and depth rate	
														Bottom separation	
														Bouyancy	
														Trim	
														Present position	*Navigation data*
														Probable position error	
														Command course	
														Speed	
														ETA	
														Obstacles	
														Other ships	

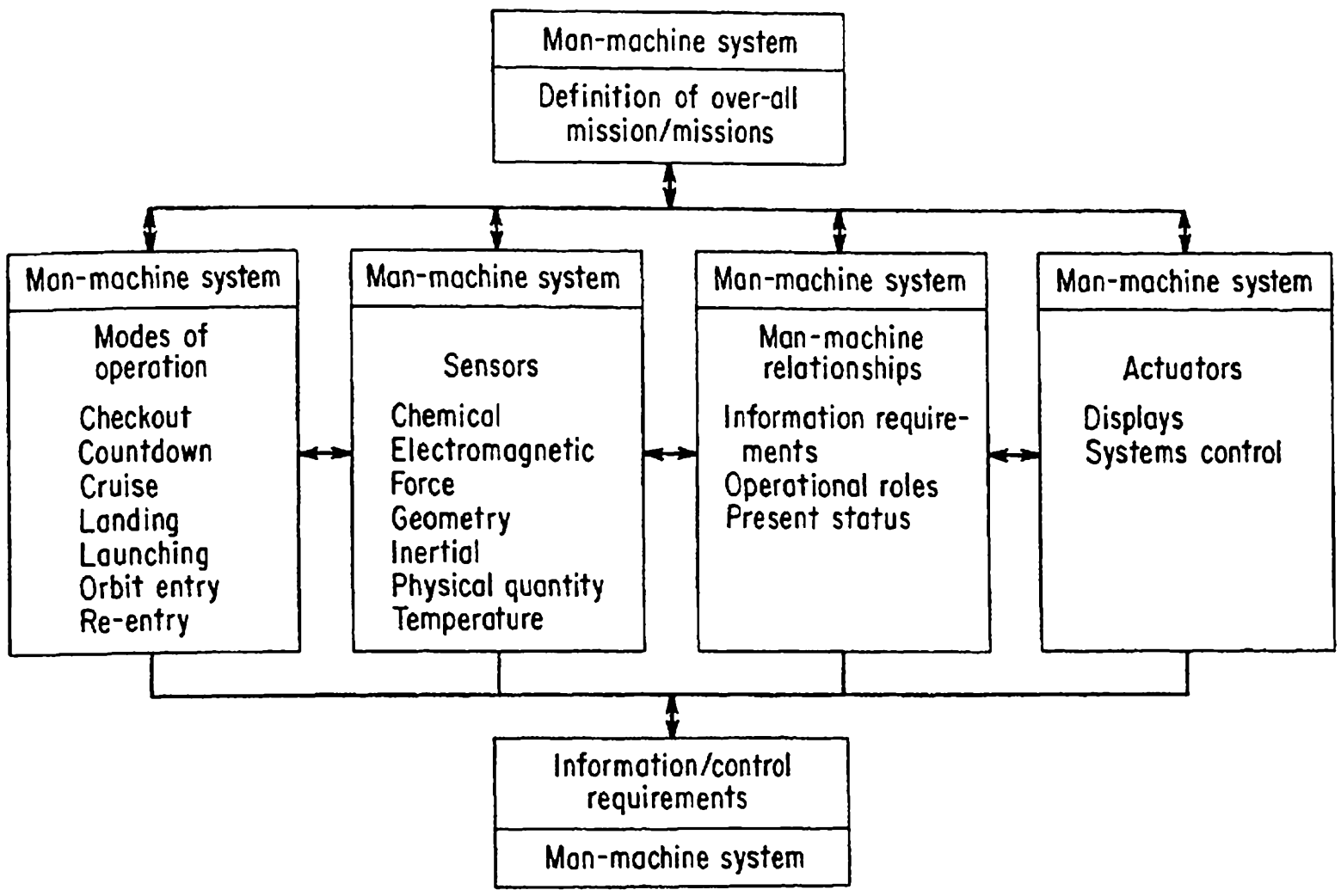


Fig. A5.25. Chart 1 (manned lunar vehicle).

Chart A-1:

Man-machine system	*Communication*			*Engineering*							
Modes of operation	Direct voice	Radar	Radio	Auxiliaries	Generators	Heat exchanger	Hydraulic pressure	Pneumatic pressure	Propulsion motors	Reactor	Turbines
Checkout											
Countdown											
Cruise											
Landing											
Launching											
Orbit-entry											
Re-entry											

Fig. A5.26

Manned Lunar Vehicle.

Environmental data																		*Local environment data*						*Navigation data*					
Internal									*External*																				
Acceleration	Air pressure	Carbon dioxide	Humidity	Noise	Oxygen	Radiation	Temperature	Vibration	Absolute wind direction and speed	Air density	Air pressure	Corpuscular radiation	Electromagnetic radiation	Electron density	Meteor size & density	Relative wind	Temperature	Absolute attitude	Heading	Pitch & pitch rate	Roll & roll rate	Turn rate	Yaw	Command course	Estimated time of arrival	Obstacles	Present position	Probable pos. error	Speed

Chart A-2:

Sensors	*Man-machine system*	*Communication*			*Engineering*							
		Direct voice	Radar	Radio	Auxiliaries	Generators	Heat exchanger	Hydraulic pressure	Pneumatic pressure	Propulsion motors	Reactor	Turbines
Chemical	Carbon dioxide indicator											
	Oxygen indicator											
Electromagnetic	Audio transceiver											
	Data-link											
	Geiger counter											
	IFF											
	Magnetometer											
	Photon counter											
	Pyrheliometer											
	Radar tracking											
	Radiation monitor											
	Space-speed indicator											
	TV											
	Vehicle charge indicator											
Force	Airspeed indicator											
	Airstream direction detectors											
	Anemometer											
	Elas. restrained synchros											
	Pressure detector											
	Vibration detector											
Geometry	Mech. angle detector											
	Stadiometer											
	Synchros											
Inertial	Accelerometers											
	Gravimeter											
	Rate gyros											
	Stable platform											
Optical	Albedo measurement											
	Astro-camera											
	Interferometer											
	Star-tracker											
Physical quantity	Acoustic gages											
	Buoyant probes											
	Capacitance probes											
	Digital gages											
	Hydrograph											
Temperature	Infra-red detector											
	Temperature bulbs											
	Thermocouples											

Fig. A5.27

MANNED LUNAR VEHICLE.

Environmental data																			*Local environment data*						*Navigation data*					
Internal									External																					
Acceleration	Air pressure	Carbon dioxide	Humidity	Noise	Oxygen	Radiation	Temperature	Vibration	Absolute wind direction and speed	Air density	Air pressure	Corpuscular radiation	Electromagnetic radiation	Electron density	Meteor size & density	Relative wind	Temperature	Absolute attitude	Heading	Pitch & pitch rate	Roll & roll rate	Turn rate	Yaw	Command course	Estimated time of arrival	Obstacles	Present position	Probable pos. error	Speed

Chart A-3:

Man-machine system		*Communication*			*Engineering*							
Man-machine relationships		Direct voice	Radar	Radio	Auxiliaries	Generators	Heat exchanger	Hydraulic pressure	Pneumatic pressure	Propulsion motors	Reactor	Turbines
Information for present status or modes of operation	Normal											
	Standby											
	Emergency											
Attention-getting information requirements	Normal											
	Standby											
	Damage control information											
Operational decision roles	Prediction											
	Reprogrammer											
	Appraiser											
	Controller											

Fig. A5.28

MANNED LUNAR VEHICLE.

Environmental data																		*Local environment data*						*Navigation data*					
Internal									External																				
Acceleration	Air pressure	Carbon dioxide	Humidity	Noise	Oxygen	Radiation	Temperature	Vibration	Absolute wind direction and speed	Air density	Air pressure	Corpuscular radiation	Electromagnetic radiation	Electron density	Meteor size & density	Relative wind	Temperature	Absolute attitude	Heading	Pitch & pitch rate	Roll & roll rate	Turn rate	Yaw	Command course	Estimated time of arrival	Obstacles	Present position	Probable pos. error	Speed

Chart A-4:

Man-machine system		*Communication*			*Engineering*							
Actuators		Direct voice	Radar	Radio	Auxiliaries	Generators	Heat exchanger	Hydraulic pressure	Pneumatic pressure	Propulsion motors	Reactor	Turbines
Display Amp	Eta and endurance											
	Horizontal											
	Power condition											
	Standby instrumentation											
	Vertical											
	Warning signals											
Mechanisms amp	Auxiliary functions control											
	Flight path control											
	Mechanical advantage control											
	Power control											
System control	External											
	Machine											
	Man											

Fig. A5.29

Environmental data	Internal	Acceleration													
		Air pressure													
		Carbon dioxide													
		Humidity													
		Noise													
		Oxygen													
		Radiation													
		Temperature													
		Vibration													
	External	Absolute wind direction and speed													
		Air density													
		Air pressure													
		Corpuscular radiation													
		Electromagnetic radiation													
		Electron density													
		Meteor size & density													
		Relative wind													
		Temperature													
Local environment data		Absolute attitude													
		Heading													
		Pitch & pitch rate													
		Roll & roll rate													
		Turn rate													
		Yaw													
Navigation data		Command course													
		Estimated time of arrival													
		Obstacles													
		Present position													
		Probable pos. error													
		Speed													

Appendix 6

Mathematical Model of a Conventional Digital Integrator

"Every man's credit and consequence are proportioned to the sums of which he holds in his chest."

—JUVENAL

1. Definition

A conventional digital integrator is a device having two inputs, α and β; one output, γ; four auxiliary variables ρ, η, ξ, and ζ; and it satisfies the following postulates in which quantized time is referenced by non-negative integral subscripts:

[1] There are fixed positive integers r and q so that

$$\alpha_i = \pm 2^{-q} \quad \text{and} \quad \beta_i = \pm 2^{-r}; \quad \forall_i \tag{43}$$

[2] $$\eta_0 = 2^{-1} \quad \text{and} \quad \rho_0 = 2^{-(q+1)} \tag{44}$$

[3] $$0 \leq \eta_i < 1; \quad \forall i\dagger \tag{45}$$

[4] $$\eta_{i+1} = \eta_i + \beta_i/2; \quad \forall i. \tag{46}$$

† i is read "for all i."

[5] $\rho_{i+1} = \rho_i + \xi_i - \zeta_i$ and $\zeta_i = (\gamma_i + 2^{-q})/2; \quad \forall i$ (47)

[6] $\xi_i = \eta_i \max(a_i, 0) - (1 - \eta_i) \min(a_i, 0); \quad \forall i$ (48)

[7] $\zeta_i = 2^{-q}(1 + \operatorname{sgn} \min(\rho_i + \xi_i - 2^{-q}, 0)); \quad \forall i$ (49)

2. Schematic

There are two conventional representations employed as shown in Figs. A6.1 and A6.2.

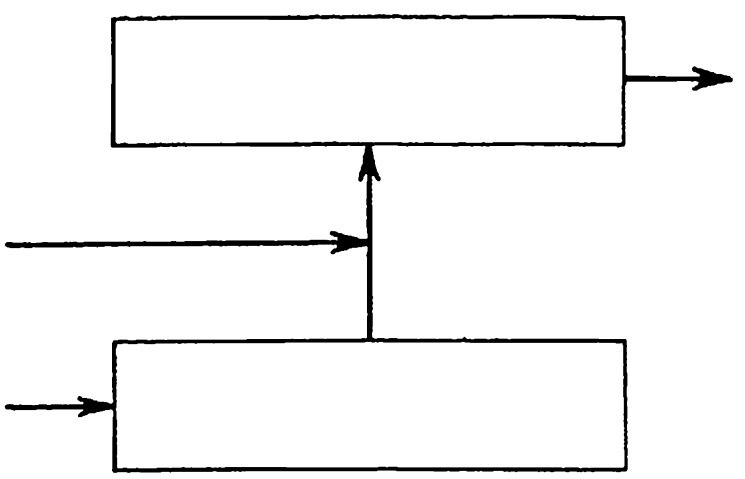


Fig. A6.1. Digital integrator (block schematic).

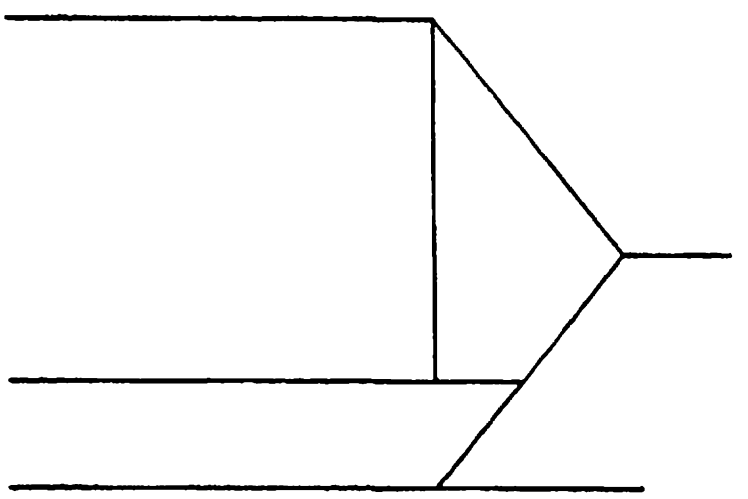


Fig. A6.2. Digital integrator (line schematic).

THEOREM 1.

$$\eta_i = \left(1 + \sum_{j=0}^{i-1} \beta_j\right)/2; \quad \forall i. \tag{50}$$

Proof. By virtue of [2] and [4], the proposition follows immediately by induction anchored at 1 = 0.

THEOREM 2.

$$0 \leq \xi_i < 2^{-q}; \quad \forall i.$$

Proof. The proposition follows immediately from [1], [3], and [6].

THEOREM 3.

$$\xi_i = |\alpha_i|/2 + \{[\max(\alpha_i, 0) + \min(\alpha_i, 0)]\sum_{k=0}^{i-1} \beta_k\}/2; \quad \forall_i. \tag{51}$$

Proof. Substitute from Theorem 1 into [6] and simplify.

COROLLARY 1.

$$\xi_i = 2^{-q-1} + \alpha_i \sum_{k=0}^{i-1} \beta_k/2; \quad \forall i. \tag{52}$$

Proof. Rewrite Theorem 3 substituting from [1] and recalling max (x, y) + min $(x, y) = x + y$ for real numbers.

LEMMA 1.

$$0 \leq \rho_i; \quad \forall i.$$

Proof. By [2], $\rho_0 > 0$. Employing this relation as anchor, make the inductive hypothesis $0 \leq \rho_i$. From [5],

$$\rho_{i+1} = \rho_i + \xi_i - \zeta_i.$$

From [7],

$$\zeta_i = 2^{-q} \quad \text{or} \quad 0$$

according as $\rho_i + \xi_i \leq 2^{-q}$ or $\rho_i + \xi_i < 2^{-q}$, respectively. From Theorem 2,

$$\xi_i \geq 0.$$

Thus

$$\rho_{i+1} = \rho_i + \xi_i - 2^{-q}$$

if $\rho_i + \xi_i \geq 2 - q$, and

$$\rho_{i+1} = \rho_i + \xi_i$$

if $\rho_i + \xi_i < 2^{-q}$ and each of ρ_i, ξ_i is non-negative.

LEMMA 2.

$$\rho_i < 2^{-q}; \forall i.$$

Proof. Again anchor with $\rho_0 = 2^{-(q+1)}$ and suppose $\rho_i < 2^{-q}$. From Theorem 2 and Lemma 1, we have $0 \leq \rho_i < 2^{-q}$ and $0 \leq \xi_i < 2^{-q}$ so that

$$\rho_i + \xi_i < 2^{-q} + 2^{-q}.$$

If $\rho_i + \xi_i \geq 2 - q$, then

$$\zeta_i = 2^{-q}, \quad \text{by [7]},$$

and

$$\rho_{i+1} = \rho_i + \xi_i - 2 - q < 2 - q.$$

If $\rho_i + \xi_i < 2^{-q}$,

$$\zeta_i = 0, \quad \text{by [7]},$$

and

$$\rho_{i+1} = \rho_i + \xi_i < 2 - q.$$

The induction is now complete.

THEOREM 4.

$$0 \leq \rho_i 2^{-q}; \quad \forall i.$$

Proof. Lemmas 1 and 2.

THEOREM 5.

$$\sum_{1=0}^{n} \gamma_i = 2 \sum_{1=0}^{n} \alpha i(\eta_i - 1/2) - 2\rho_{n+1}. \tag{53}$$

Proof. Combine both parts of [5] to eliminate ξ_i, solve for γ_i, substitute for ξ_i from Corollary 1, apply Theorem 1, and sum from $i = 0$ to n.

APPLICATION 1. The device can be used to accomplish numerical integration to within an error of 2^{-q+1} by taking $\gamma = \Delta(\int y\,dx)$, $\alpha = \Delta x$, $\beta = \Delta y$, $\eta = (1 + y)/2$. Since $\eta_0 = 1/2$, $y_0 = 0$.

Proof. By Theorem 5,

$$\int y\,dx \sim \sum_{l=0}^{n} \Delta i \left(\int y\,dx\right) \sim \sum_{l=0}^{n} (\Delta_i x) y_i - 2\rho_{n+1} \tag{54}$$

and, by Theorem 4,

$$0 \leq \rho_{n+1} < 2^{-q}.$$

Thus,

$$0 \leq 2\rho_{n+1} 2^{-q+1}$$

and the approximation is under but by less than 2^{-q+1}.

REMARK 1. We may, if we wish, regard α, β, y, ν; r (correspondent of ρ) as being in exterior code and η, ρ, ξ as being in machine code where the relation is

$$\frac{1 + e}{2} = m. \tag{55}$$

Then $r_0 = -1$, $r_n = 2\rho_{n+1} - 1$, so the error $-2\rho_{n+1}$ becomes $r_0 - r_n$.

Appendix 7

Generalized Logic-Foundations of a New Design Philosophy

"It was a saying of the ancients, 'Truth lies in a well'; and to carry on this metaphor, we may justly say that logic does supply us with steps, whereby we may go down to reach the water."

—DR. I. WATTS

1. Summary and Definitions

We here consider by example a systematic approach to problems which admit of description in terms of the lattice-theoretic structure of a suitable space of functions of time. The terms used in this generalized logic approach are defined as below:

(a) *Space of functions of time:* a set of functions of time which include with any two functions their point-wise (i.e., continuously in time) maximum and minimum.

(b) *Lattice theoretic structure of such a space:* composed of all conditions expressible in terms of a finite number of lattice polynomials, inequality test, and equality tests. (It is evident that equality tests may be composed from inequality tests.)

(c) *A lattice polynomial of functions of time:* a function of time by any finite combination of the taking of point-wise minima and/or maxima of a finite number of the given functions.

(d) *Equality test between two functions of time:* either a function which is not zero, only when two given functions are equal; or, a function which is zero, only when the two given functions are equal.

(e) Inequality test between $f(t)$ and $g(t)$: construed to be either a function which is zero when $f(t) \leq g(t)$ and is non-zero otherwise, or a function which is non-zero when $f(t) \leq g(t)$ and is zero otherwise.

2. Scope

A large number of decision problems are of sequential or continuous time extension nature consisting of comparison of the changing magnitudes of certain input variables. Such problems are describable in terms of the lattice theoretic structure of a suitable function space. Currently, the solution of such problems is being implemented by the brute force application of large and complex computing systems. A timely example of such an approach is the SAGE (semi-automatic ground environment) system intended for utilization by ADC. The generalized logic philosophy renders superfluous the use of a large computer in such a system. A hardware philosophy exists for accomplishing these very same objectives and other more general objectives, together with a logical design philosophy for simultaneously avoiding loss of accuracy and minimizing required numbers of decisions. The method indicated utilizes, in a certain new and novel fashion, certain applications of signal bypass, dynamic programming, and variable permuting techniques, in such a fashion as to retain precision and suboptimize circuitry in certain ways. The method also furnishes solutions of a particularly simple and effective nature for the particular problems of general control assignments (e.g., fire control with many targets and many guns with or without interlocks) and uniform multivariable control.

A. Example

A simple example of a familiar type of decision problem is that in which it is desired to order a number of inputs pointwise in time with respect to magnitude so that they may be picked off in increasing or decreasing order. To accomplish the implementation of the solution of this problem, consider the elementary symmetric functions. Let these functions be translated into lattice polynomials by replacing arithmetic multiplication with minimum, and arithmetic addition with maximum. We then have, for n variables, n lattice polynomials of these variables which will, in fact, order them with respect to magnitude.

A problem that may be used to illustrate this method may be found in the field of general control assignment. The situation is one which confronts the shipboard fire control officer, who is called upon during time of attack to deploy properly his ship's guns to provide a maximum protection against—let us say—enemy aircraft. To design a system that will provide a high confidence level in the protection of the ship, a verbal statement of the strategy to be employed by the machine must first be given.

Given n guns, n tracking radars and one search display system. With m targets, $m > n$, select n closest targets, lock one gun on each, subject to the condition that no gun be forced through a certain region, called its forbidden zone, which is defined in terms of elevation and azimuth boundaries; or with m targets, $m \leq n$, lock at least one gun on each target. Upon acquiring a target a gun will track and fire until such time as the target is shot down (operator indicates hit), is out of range (automatic from track radar), or causes the gun to enter a forbidden zone (decision elements), at which time the gun will then acquire the nearest of those targets not being tracked which is compatible with its forbidden zone. If no target is available, it may join in sharing a target with some already tracking gun. This is, of course, only one possible strategy of many, but it will serve for illustrative purposes.

The theoretical system to be described can be broken into a number of basic blocks, as shown in Fig. A7.1. These blocks, when properly interconnected, would perform the aforementioned strategy. The system includes a search and display block. The search and display block contains a large search radar, and a series of displays of azimuth, elevation, and range for all targets that can be picked up by the large search radar.

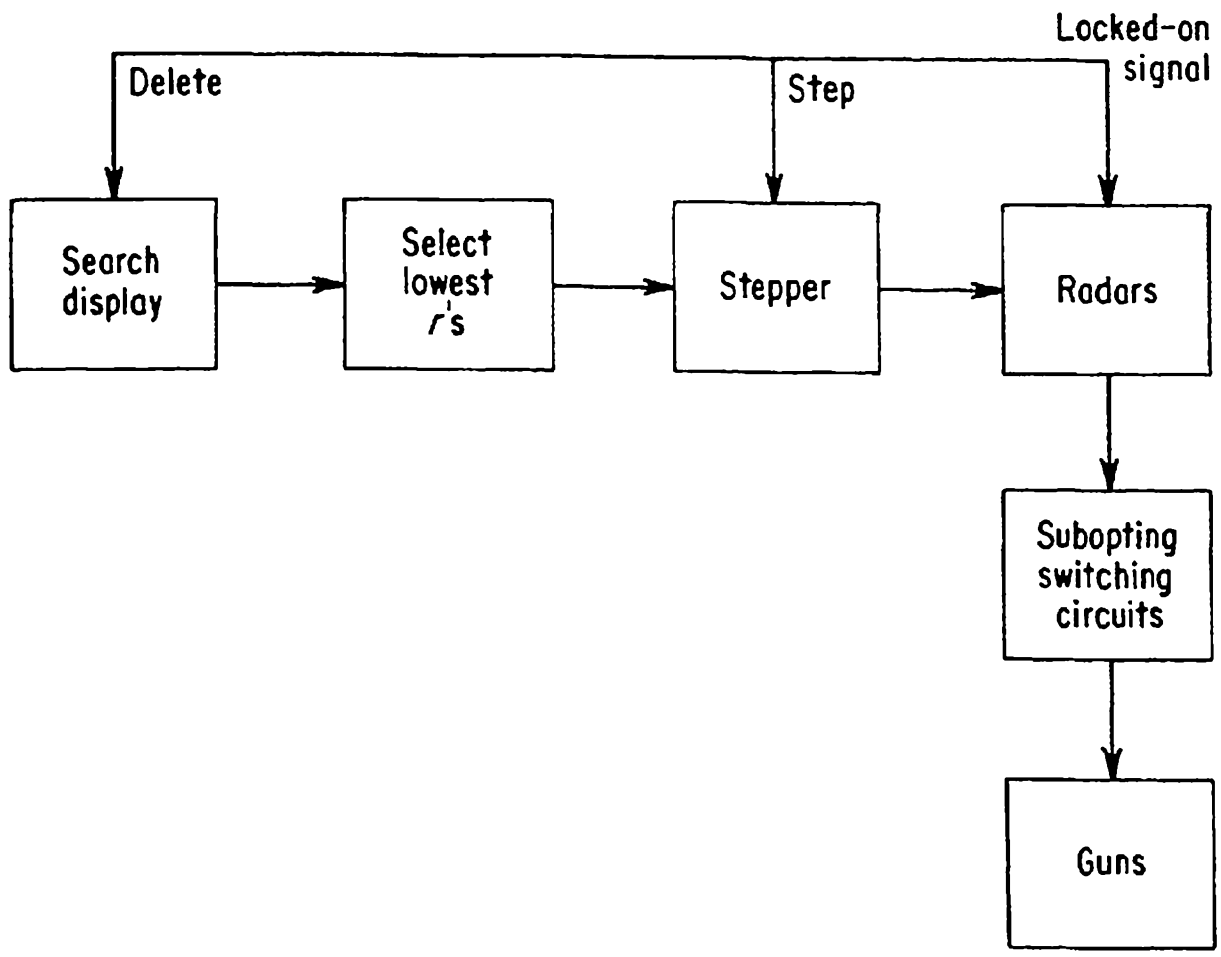


Fig. A7.1. Rudimentary shipboard fire-control system design (block diagram).

The second block is labeled "select lowest r's." The function of this block is to select, from the values of range presented to it by the search display, the closest n values of r and order them in terms of their magnitudes. The outputs of the "select lowest r's" box enter what has been termed the step-up-box. This unit performs the target-radar combination assignments by cyclically permuting those radars which are not tracking targets and assigning closest targets not being tracked to these. This process is accomplished by using decision elements in conjunction with stepping switches. When Radar 1 has locked onto a target, this target is deleted by a feedback to the search display. All other radars are stepped up such that Radar No. 2 is renamed No. 1, and the unit accepts the new closest target. Information from all radars is fed to a block called the subopting switching circuits. In this block, targets and guns are paired by a decision process that investigates the relationship between a target's position and the gun's forbidden area. Initially, at the start of an attack, the various guns are given number names by a senior fire control officer. The numbering scheme for the guns is based on the officer's estimate of the attack situation. These initial names of guns are presented to the subopting unit as positions or contact numbers on a bank of rotary stepping switches, there being one gun per each contact. That gun which has been named 1 is then examined with respect to the

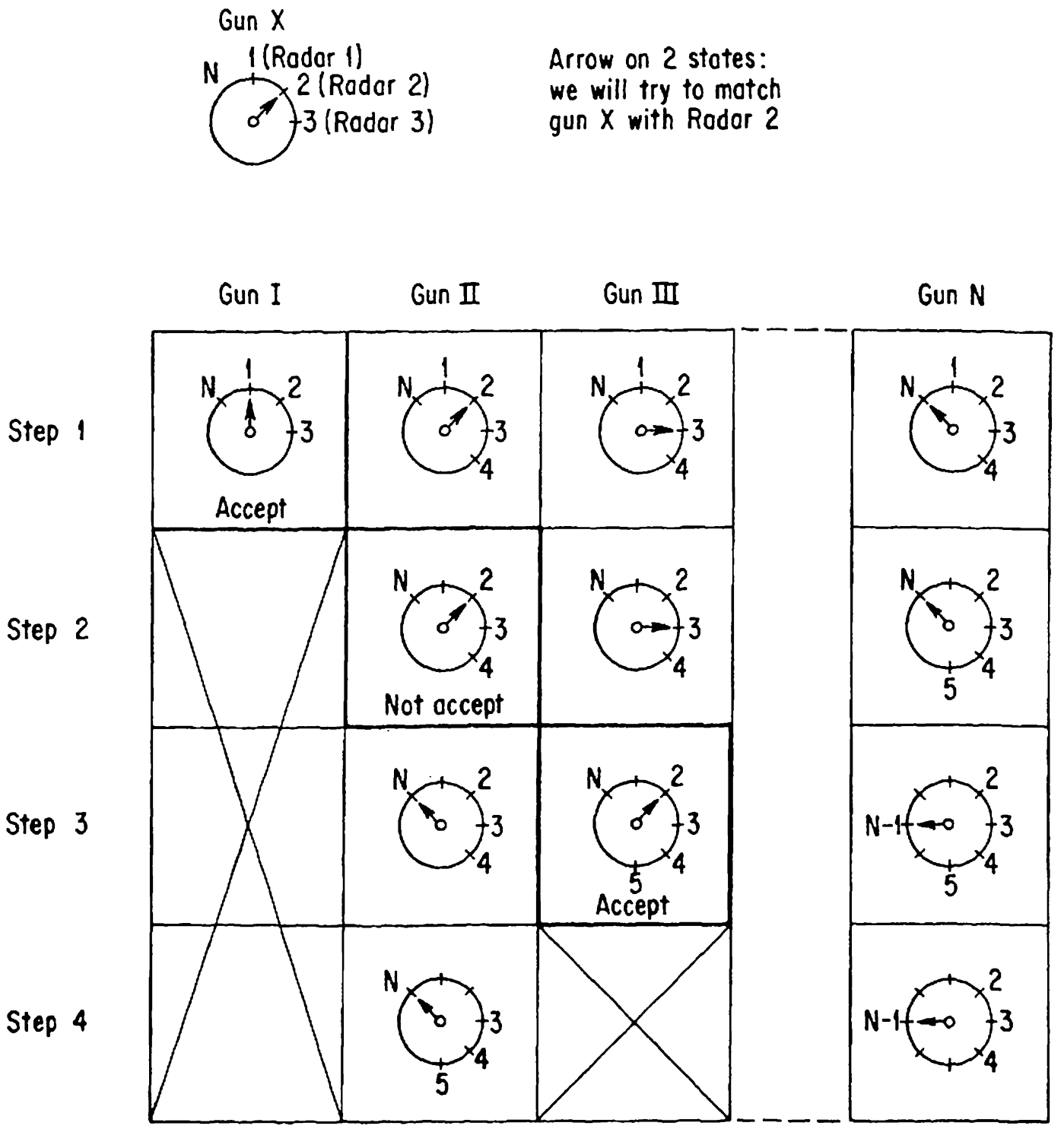


Fig. A7.2. Gun acquisition of radar (flow diagram).

closest target in an attempt to determine whether or not it can fire at this target (*see* Fig. A7.2). If it can, then Gun 1 via Radar 1 acquires Target 1, and Target 2 is now presented for gun target pairing. Gun 2 is investigated and if, for example, it is found unsuitable for Target 2, Gun 2 is demoted and becomes the last gun name assigned. All other guns with the exception of Gun 1 are promoted. Such procedure permits the new Gun 2, previously named Gun 3, to investigate Target No. 2. It is apparent then that for each target all guns not firing at targets can be cyclically permuted until a target gun pair is formed. Moreover, the priority assignment of targets is maintained, the most dangerous target (Target No. 1 is the closest) being assigned a gun first. It is, of course, apparent that *a priori* criteria other than proximity may be utilized by suitable system modification.

Upon the shooting down of a target, a gun operator indicates a hit, and

the freed radar is given a name in the step-up box, which is always assigning nearest untracked ranges to radars that do not have targets to track. Whenever a target is acquired by a gun, it no longer has a name in the numerical coding on the stepping switches. The pin, or contact, associated with this gun essentially goes dead or has a recognizable signal upon it which indicates that this gun is shooting at a given target. To clarify the theory of the arrangement of signals on these stepping switches, we cite a typical example. Let us assume 20 gun positions and 20 tracking radar, there being one pin or contact on the stepping switch for each tracking radar, and, of course, one such switch for each gun. The mechanical position of the contact arm designates the gun name. Utilizing the techniques of dynamic programming and permuting variables within the general framework of the established philosophy result in an overwhelming reduction in the hardware and hardware complexity that would be required in a large digital computer and solve the strategy problem with almost as high a confidence level as could be achieved in the digital computer.

3. Other Applications

It is possible within the context of this philosophy to design systems yielding multivariable uniform control in which a number of variables are modified such that they are all made equal by a process that corrects the variable that departs from equality by the largest amount. A system might be designed, for example, in which several motors of different types and sizes are to be driven at equal speeds with only one motor being energized at any given instant of time, thereby reducing an over-all power load on a generator. The motor which is energized is that motor which requires the greatest correction. A similar approach can be used in telemetering systems for bandwidth compression by transmitting only those signals which have changed by more than some predetermined amount since the last transmission, or by transmitting those signals having the greatest fluctuations with a higher priority assignment than those functions which are slowly varying in time. The fundamentals underlying the strategy discussed enjoy the capability of application to many system problems currently being faced in commercial and military development. These would include missile firing problems, advanced early warning situations, and, from the commercial and military aspect, the generalized problem of handling air traffic in the vicinity of airports.

Appendix 8

Preliminary Consideration of an Army Surveillance System

"Get a thorough insight into the index by which the whole book is governed and turned like fishes, by the tail."
—SWIFT

1. Problem Areas Consideration

A. Introduction

We here indicate certain questions and difficulties arising in a preliminary study of requirements for an army surveillance system. The material is purely illustrative and does not refer to any particular existing system.

B. General Remarks

An army surveillance system (considered herein as at level of field army or above) is of a vast nature not only because of its physical size

and spatial distribution, but also because of its advanced standing in the hierarchy of complexity. Its equipment components themselves frequently possess complexity comparable to that of a full-blown modern-weapons system. These equipment components interact with an extensive array of more or less specialized human components. In addition to data acquired by inherent sensory equipment, the system accepts inputs from a broad spectrum of foreign and domestic sources, including sources dealing with economic, political, meteorological, geographical, and technical news, as well as intelligence sources such as, in the United States, NSA, CIA, Naval Intelligence, SAC Reconnaissance, ATIC, and USAF Security Service. Contemplation of the system in block diagram form at a gross level of detail is trying, and serious understanding of its flow charts is little short of impossible while the necessary complexity is compounded by a wide variety of data and information formats and media.

C. Temporal Data Requirements

All of the Army planning and evaluating missions depend to a greater or lesser extent upon possession of accurate current information of a surveillance nature. Such information is essential to individuals evolving tactics based upon enemy order of battle estimates, those extrapolating known factors for strategic planning purposes, and those evaluating the results of field exercises. The needs for faster and more uniform acquisition, transmission, and processing of information from surveillance activities are radically apparent in the era of pentomic divisions and missile commands that have adopted the philosophy of extreme mobility and flexibility.

D. Equipment Diversity

Any currently contemplated integrated operational surveillance system will involve an encyclopedic *in extenso* catalog of equipment characteristics. There will be communication channels, which will vary from (physically) simple voice links to multiplexed pulse-modulated encrypted secure data links. There must be considered the complete range of commercially available data processors and general purpose computers, as well as special military oriented machines for literally "on the spot" tactical data processing. There will be cryptological devices varying from

Hagelin-type machines at the front to exotic forms of wired rotor machines in communication nodes. There will be reconnaissance equipments varying from photographic devices with telefax follow-up to entire air-ground data-acquisition systems. It will be necessary to utilize large complexes of these equipments toward a common end without introduction of unnecessary time lags, loop instabilities, or system failure modes.

E. Information Format Diversity

Let us first consider the tactical information requirements of an organizational level well below that of a field army.

F. Tactical (Surveillance) Information Requirements for Company Forward Action

Surveillance information would, in a maximum context, subsume all data relating to the capacity of a given political entity to assume and implement offensive or defensive postures at an arbitrary time. Within such an extended context, as war strategy, a portion of national energies would be devoted to the processes of collection, collation, evaluation, synthesis, and interpretation of information from all available sources. The objective of these activities would be a description in depth, modifiable in time, of all elements composing the technological, economic, and political fabric in terms of potential for hostile activities of any magnitude.

One may approach the problem of discerning types of surveillance data by considering the features of a conflict in microcosm; viz., a company-sized forward action. Employment of this approach does not appear to compromise completeness of the listing, although depth is sacrificed, e.g., full enemy potential for troop and material replenishment. The major categories of surveillance information required for objective offensive or defensive behavior could be unified under such broad aspects as those designated as follows:

(a) Task organization.
(b) The tactical situation.
(c) The mission.
(d) Manner of execution of the mission.
(e) Administration and logistics.
(f) Command and communication.

Let us now examine each of these aspects and derive in this manner a list of information that can be characterized as having planning or decision-making value.

Task organization

(a) Units involved.
(b) Types of weapons organic to these units and to those elements which may be temporarily attached.
(c) Unit(s) being held in reserve.
(d) Strength of units and weapons providing support; e.g., artillery, tank companies, aircraft.
(e) Firing priorities of the supporting units.

The tactical situation

(a) Location and strength of enemy forces.
(b) Points of greatest resistance.
(c) Potential breakthrough points.
(d) Disposition and engagement of enemy in surrounding areas and evaluation of conflict outcomes.
(e) Estimate of indirect support time as contingent upon outcome of engagements nearby.
(f) Direction of movement of enemy troops and approximations of time of contact with friendly forces.

The mission

Certain units reinforced by others or elements thereof, within a contact zone of predetermined dimensions defined geographically, may attack (or retrogress). These may also combine with other units to seize (or retire to) another objective without further command or on order.

Manner of execution of the mission

(a) Composition of the elements which are to lead the attack to a level of sufficient details.
(b) Disposition and strength of accompanying units.
(c) Attack paths of all participating elements.
(d) Points of assault of all elements.
(e) Location of reserve elements.
(f) Location of firing sites.
(g) Location of outposts.
(h) Location of all supporting elements.

(i) Manner of lifting and placing supporting fire; i.e., following SOP or specified procedures.
(j) Anticipated consolidation procedures.
(k) Anticipated reorganization.

Administration and logistics

(a) Location of designated assembly areas and commands for initial distribution of supplies and ammunition.
(b) Location of all forward sources having accessibility potential.
(c) Locations of primary and contingent supply paths.
(d) Command personnel identification.
(e) Other.

Command and communications

(a) Location of CO and party, initially and successively.
(b) Strength of command party and composition.
(c) Communications axes.
(d) Communication means.
(e) Primary and conditional communication connectivity patterns.

Other types of data, though implied above, have not been listed discretely. A partial listing of these follows:

Maps

(a) General: (local and extended) of an appropriate scale.
(b) Special situation maps indicating movements and objectives, contour intervals, roads, trials, wooded areas, artifactual and material obstacles.

Meteorological information

(a) Diurnal temperature changes.
(b) Wind direction and velocity.
(c) Humidity.
(d) Forecasting summaries.
(e) Sunrise-sunset times.

Enemy reports or logs

(a) Field message center logs.
(b) Field crypto center logs.
(c) Field teletype operators logs.
(d) POW interrogation reports.

Weapons characteristics

(a) Type (bore).
(b) Range.
(c) Warhead (conventional, HE or atomic).
(d) Impact area size.
(e) Minimum manpower levels required.
(f) Rates of fire.
(g) Standard operating procedures.
(h) Mobility and mobility means.

Terrain characteristics and ecology

(a) Normal.
(b) Abnormal (as under storm conditions).
(c) Texture (rocky, sandy, loamy).
(d) Flora and fauna.
(e) Airfield possibilities.
(f) Natural water supply.

Biographical information

(a) Personal idiosyncrasies of officers.
(b) Tactical weaknesses or strengths.
(c) Personal tactical histories.
(d) Special accomplishments.

Vehicles available

(a) Types and quantity.
(b) Operating characteristics.
(c) Speed as estimated in light of terrain characteristics.
(d) Plans for functional use.

Enemy preparations (special)

(a) Booby traps.
(b) Identity and extent of mine fields.
(c) Mines in use—operation, manner of deactivation.
(d) Quaint practices or penchants.

Supporting equipment (general)

(a) Characteristics.
(b) Special maintenance or testing procedures.

G. Customary Formats

A cursory examination of the indicated requirements will disclose formats as follows:

File documents

(a) Printed text.
(b) Maps and terrain diagrams.
(c) Photographs.
(d) Equipment diagrams and schematics.

Intelligence and captured documents

(a) Printed text (frequently foreign).
(b) Maps and terrain diagrams.
(c) Photographs.
(d) Equipment diagrams and schematics.
(e) Encrypted text or code groups.

Immediate sensory formats

(a) Meteorological charts.
(b) Photographs.
(c) Sketches.
(d) Radar maps.
(e) Raw data on enemy order of battle.
(f) Varying computer print-out formats.
(g) Operator displays of arbitrary format employing visual and/or auditory integration of information.
(h) Text encrypted in varying forms according to local generating equipment.
(i) Signal forms resulting from particular communications equipments such as teletype and AM radio.

H. Additional Difficulties Implied by Information Diversity

In addition to the format diversity problem per se, certain other problems encountered may be:

(a) The problem of associating confidence levels with information sources varying in physical remoteness.

(b) The problem of interpretation of textual résumés obtained without reduction by means of preassigned rules from direct observation; that is, the problem of subjectivity.

Facilitation of secure links. Diversity of signal forms may well require highly specialized cryptological equipment with its implied system costs in training and logistics.

Personnel and training. The utilization of multiple operating languages (signal forms, codes, etc.) implies strict requirements for operating personnel and extensive training for those selected.

Lack of time-sharing. Growth of equipment requirements and the consequent reduction of system reliability is one direct reflection of failure to utilize equipment characteristics in compatible means.

Standby capabilities. In order to achieve standby capabilities and modes of degraded operation as opposed to system failure, it is necessary to attain a reasonable degree of interchangeability of equipment function. This in turn requires a high degree of compatibility.

Checking and adaptation. The ability of a complex information system to accomplish error checking and correction is very sensitive to the information formats utilized, and distinct formats generally require quite distinct system features to enable this ability.

2. Diversity of Communication Channel Types

The multitude of communication channels may be summarized under the broad headings listed below. Composites of these systems also occur:

Radio
Wire
Invisible light and heat
Visual
Auditory 20–20,000 cps
Spatial translocation
Miscellaneous

A. DEFINITIONS

Radio systems include channels that transfer intelligence by means of ethereal electromagnetic waves or pulses within a frequency range between 2 kilocycles and 3×10^9 kilocycles. Radio systems are particularly useful, since no physical material need be provided between antennae. Radio and wire carriers are discussed in the latter part of this section.

Invisible light and heat systems include systems utilizing the frequency spectrum above that of 3×10^9 kc, but not including the visual light range. Security and modulation considerations are related to those of line of sight radio.

Wire systems include channels that transfer intelligence by means of wires or cables. These systems necessitate the provision of a material conductor. Security is enhanced by the fact that under normal operating procedures the enemy could receive transmissions only by direct connection or very close juxtaposition to the wires or cables.

Visual systems include systems that utilize modulation of light sources, flag signals, semaphores, light arrays, etc. These systems operate on a line of sight basis. Security of visual systems is difficult to maintain except through the use of codes. Coding and reading of visual signals depend largely on the skill of human operators.

Mechanical systems include signals carried primarily by mechanical sound waves generated in a conducting medium (e.g., air, water). These systems are, in greater part, limited to short distances. Such systems include, among others, sound and sonar.

Spatial transformations include forms of communication that do not come under the classifications above. Unformalized and ill-defined parapsychological phenomena and gesture languages are considered as belonging to this category.

3. Radio and Wire Systems

This subdivision will deal primarily with a more detailed account of radio and wire systems. The methods of modulation that can be and are used in radio and to some extent in wire systems are described below. Some of these techniques are also applicable in other means of communications. The term "modulation" refers to the method by which the

intelligence is impressed upon the channel. Processes of modulation may be divided into two broad categories; continuous wave (cw) and non-continuous (pulse).

Continuous wave modulation, as the name implies, utilizes those forms of modulation that do not interrupt the carrier.

(1) Modification of carrier amplitude is called amplitude modulation (AM). The signal intelligence is made to vary or control the amplitude parameter of the carrier. Phase relations remain constant.
(2) Modification of the carrier phase or angular modulation is used in two ways.
    (a) Frequency modulation (FM) is that method that uses the intelligence signal to vary the instantaneous frequency parameter of the carrier. The amplitude remains constant.
    (b) Phase modulation is the method that uses the signal intelligence to vary the instantaneous phase excursions of the carrier. Amplitude remains constant.

Pulse or non-continuous techniques of modulation utilize methods by which either the amplitude or time occurrence of some characteristic of a pulse carrier are controlled by instantaneous samples of the modulating (intelligence) wave.

Pulse modulation types

(1) Pulse-time modulation (PTM) refers to those processes in which instantaneous samples of the modulating (intelligence) wave control the time occurrence of some characteristic of a pulse carrier, with carrier amplitude remaining constant.
    (a) Pulse-position modulation: Each instantaneous sampling of the intelligence wave controls the time position of a pulse in relation to some timing pulse.
    (b) Pulse-duration modulation (PDM): Each instantaneous sampling of the intelligence wave controls the time duration or pulse width. This type of modulation is also called pulse-width modulation (PWM).
    (c) Pulse-frequency modulation. The modulating wave is used to frequency-modulate a carrier composed of a time series of pulses.
    (d) Additional methods include modified time reference and pulse-shape modulation.

(2) Pulse amplitude modulation (PAM). Instantaneous samplings of the modulating wave control the amplitude of the various carrier pulses with the time of occurrence of each pulse and the pulse width remaining constant. Forms of this modulation include single polarity pam and double polarity pam.

(3) Pulse-code modulation (PCM). The modulating wave is sampled, quantized, and coded. Types are:

(a) Binary pulse-code modulation in which the code for each element of information consists of one of two distinct values (e.g., pulses and spaces). In wire telegraphy, to cite an example, a direct current pulse and space are used.

(b) Ternary pulse-code modulation in which the code for each element of information consists of any one of three distinct values (e.g., positive, zero, negative).

(c) N-ary pulse code modulation in which the code for each element of information consists of any one of N possible distinct values.

Pulse modulation is commonly used in a time-division-multiplex system in which other pulses may be sent over the same channel provided they are synchronized to appear within the time space of another string of pulses. A separate channel or a portion of a channel is used for synchronization.

A. Sidebands

In all basic modulation types there normally appear two sidebands, one on each side of the nominal center frequency of the channel. Three sideband techniques worthy of note are summarized below:

(1) *Suppressed carrier modulation.* When using this type of modulation, it is necessary to transmit only enough carrier that may be used at the receiver to control a locally generated carrier. The local carrier may then reduce the effective percentage of modulation.

(2) *Single-sideband modulation.* The one sideband that is transmitted can carry the intelligence of a double sideband signal at a substantial saving of space within the frequency spectrum. Single-sideband (AM) techniques suffer a disadvantage when the modulating signal contains low frequencies. As the frequency approaches zero, it becomes very difficult to suppress the adjacent portion of the unwanted sideband.

(3) *Vestigial sideband modulation.* This type of modulation, under low frequency conditions, is able partially to eliminate, but not completely to suppress, the unwanted sideband. Under certain conditions this method may be used to transmit intelligence accurately in a portion of the frequency spectrum smaller than that of double sideband.

B. Transmission Classification

Types of intelligence to be transmitted over a channel may be summarized under five headings. These headings are telegraphy, telephony, facsimile, television, and special.

Telegraphy is that method in which keying takes the form of long and short pulses in the ratio of 3:1. This may be accomplished by:

(1) On-off keying.
(2) Keying of a modulating audio frequency or frequencies.
(3) Keying the modulated emission.
(4) Frequency shift keying.
(5) Keying modulated pulse PAM, PDM, PFM.
(6) Use of teletypewriter or other semiautomatic and automatic telegraph equipment.

Telephone is that method by which voice is transmitted by wire or radio. Carrier frequencies may be used in both. Occasionally, no carrier frequency is used in wire transmission. When using a carrier, the modulation may be any or a combination of the above-mentioned methods.

Facsimile and *television* transmit pictures, maps, etc. These methods usually use amplitude and/or frequency methods of modulation (AM or FM).

Special types of information include, among others, crosstalk between computers, and remote control programming of computers.

4. Possible Preference Bases

A. Radio *vs.* Wire Transmission

Radio facilities are, in general, more portable and can be installed more rapidly than corresponding wire facilities. Radio is employed in initial

phases of combat and for communication with highly mobile elements such as airplanes and tanks.

Wire facilities are more suitable for handling the bulk of ground communication between permanent, semipermanent, or temporarily established points in a combat zone.

Main characteristics

(a) Radio systems are smaller and lighter than corresponding wire systems.
(b) Wire systems require less power.
(c) Radio channels are limited by frequency assignment, whereas wire circuits are limited only by material and manpower available.
(d) Radio systems are readily intercepted or jammed, whereas wire and cables are more subject to interruption by physical damage.
(e) Radio communication ranges are less predictable than wire.
(f) Radio stations are more subject to position-finding by enemy forces than are wire stations.
(g) Radio and wire systems are both affected by weather conditions, but wire systems are, in general, more stable than radio.

B. Telephony *vs.* Telegraphy

Telephone affords closer personal contact between stations, and this is especially important when ideas need to be developed or special circumstances explained quickly. Characteristic telephony troubles arise from too much talk, lack of permanent records (unless special recorders are used), and the possibility of security breaches.

Telegraphy is especially suited to one-way messages that require little explanation. In most cases both telegraph and telephone are on the same channel.

Traffic capacity

(a) Telephony is faster than hand key and manually operated teletypewriters.
(b) Previously punched tapes read on tape reading devices raise telegraphy rates.
(c) Multi-channel voice-frequency carrier telegraph systems can expand total intelligence per unit time.
(d) Little training is needed for telephone operation as compared to telegraphy operation.

C. Modulation Types: AM *vs.* FM

(a) AM systems are less susceptible to carrier drift.
(b) FM systems are less susceptible to noise and level variations.

At the present time there is much interest in single-sideband frequency modulation systems. These systems provide less noise and a greater use of frequency spectra.

5. Preliminary Remarks on Gross System Design

A. General Remarks

In considering a surveillance system, it must be kept in mind that the system is a dynamic entity and functions not *in vacuo* but as a part of the army system that utilizes it for control purposes. Conceptually, it is the sensory (or, colloquially, front-end) subsystem of the entire army system and is used in control of the entire system, both for sensitivity to initial conditions and feedback. The surveillance system will consist typically of the conceptual subsystems depicted in Fig. A8.1. The criteria governing design of all or part of the surveillance system are of necessity of complexity beyond that which we are presently capable of handling by formal techniques. Approximations must therefore be resorted to in selection of the system characteristics, including that of common language. Even if we were capable of solving in full generality the formidable mathematical problems anticipated, we would need to make use of some basic ingredients in these problems that are themselves approximations (for example, man-machine transfer functions) owing to present lack of knowledge. One must, on the other hand, avoid selection of criteria which lead to suboptimization rather than entire system optimization, as these are generally mutually exclusive.

Basic Sensors. Guidelines in this area are:

(1) Requirements for judgment in conversion.
(2) Allocation of duties among men and machines for conversion.
(3) Complexity of conversion equipment.

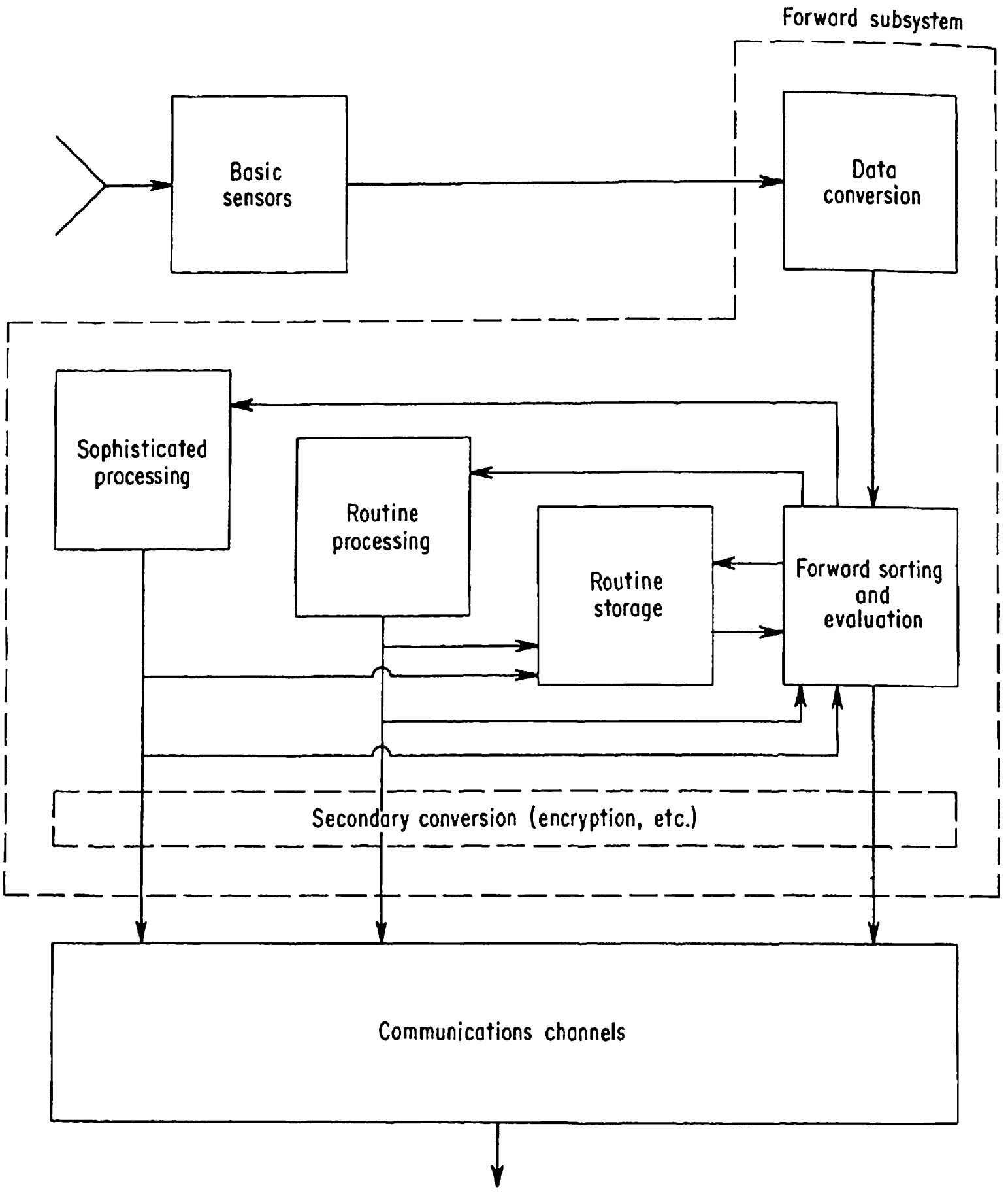


Fig. A8.1. Conceptual flow diagram of the surveillance system.

FORWARD SORTING AND EVALUATION. Although certain data will be converted and committed automatically to communication channels, portions of the information will not be susceptible at least at present to deterministic procedures. Contingencies must be allowed. Standard operating practice in this area must be established and possible modification enhancing performance considered. The primary problem will again be that of allocation of duties between men and machines.

SOPHISTICATED PROCESSING. This subsystem is composed of technical experts and large-scale information machinery. Consideration must be given to input-output characteristics of commercial machinery, as well as to appropriate portions of special military oriented equipments and to the actual problem categories to which these machines will be addressed by the technical experts. Consideration of inter-machine operation should not be limited to compatible codes but should be extended to consideration of universal coding (i.e., common procedural technique for programming).

ROUTINE PROCESSING. The previous remarks on sophisticated processing apply with equal weight to the machine portions of this sub-system. The manual portion should probably be governed by established standard operating practice, modified as proven desirable to enhance performance consistent with constraints.

ROUTINE STORAGE. The primary point to be considered here is that this subsystem does not through inadvertency or ineptness in sorting and evaluation or through lack of system planning subsume functions of the other subsystems, thus becoming a convenient "sink."

SECONDARY CONVERSION. Ease of encryption and subsequent decryption, facility of recording data decoded for human utilization, and ease of switching into alternate channels will be the primary matters of interest in this rather nebulous subsystem.

COMMUNICATION CHANNELS. The channels available for transmission of data and conclusions to higher organizational levels are relatively fixed, and the primary consideration of them will appear as constraints on code selection.

PARTIAL CRITERIA. Although suboptimizations at the expense of system optimization should be avoided, reasonable weight must be assigned to the following points:

(1) Elimination of unnecessary handling of data between origination and utilization.
(2) Various priority categories of information such as those requiring immediate reaction.
(3) Degree of complexity of conversion equipment.
(4) Realistic constraints regarding utilization of existing equipment and typical personnel.

(5) Number of channels required; over-all and degree of utilization of channel capacity (relative efficiency).
(6) Facility of entering standby or alternate modes of operation.
(7) Operability of system (not only regarding maintenance and test, but also vulnerability and degraded modes of operation).
(8) Degree of immediate capability of automation of checking and correction.

Appendix 9

Remarks on Low-Level Redundancy

"Defendit numerus: There is safety in numbers."

—ANON

1. Introduction

Redundant blocks of components (parallel-serial cascades in general) intended to allow construction of "reliable" circuitry from inadequate components have aroused much interest recently. However, the use of parallel-serial cascades is not to be conceived of as a panacea. "Noise bursts" and extreme environmental conditions generally knock out cascades as readily as single components.† We shall consider in this appendix some relatively simple cascades with various somewhat unrealistic assumptions concerning their behaviour, and we shall exclude failures of the primarily perverse nature which necessarily involve statistically dependent component behavior. Furthermore, we shall limit attention to "simple failures" (excluding, for example, a transistor having an open emitter lead which later develops a base-to-collector short).

† *See* J. R. Duffett, "Some Mathematical Considerations of Redundancy," Radioplane Company, 1956.

2. The Diode Quad

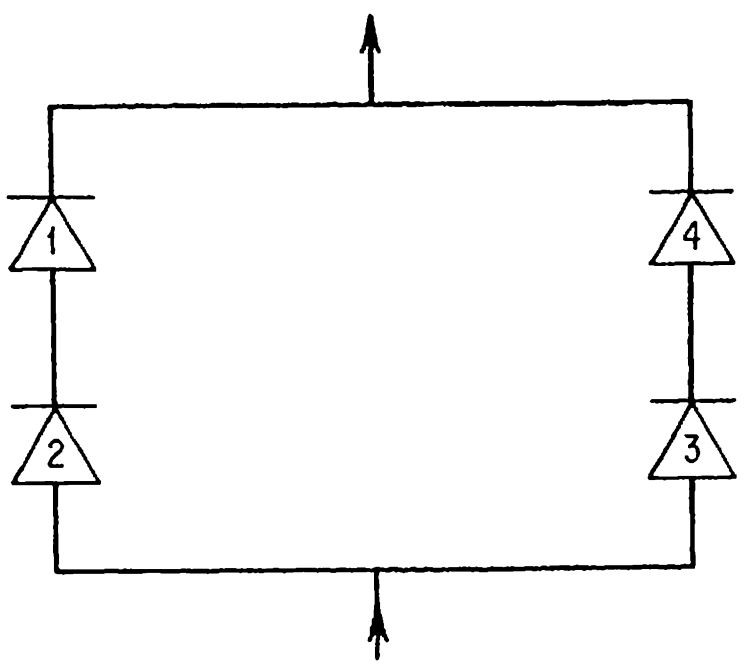


Fig. A9.1. Diode quad.

Consider a device as shown in Fig. A9.1.

HYPOTHESIS 2.1. Each diode has three possible states as follows:

[1] G, denoting normal operability.
[2] S, denoting short circuiting
[3] N, denoting open circuiting.

CONVENTION 2.1. We write X_i for state X of diode i, and ξ_i (i.e., Greek correspondent) for a suitable probability measure of event X_i.‡

HYPOTHESIS 2.2. The failure behavior of a collection of diodes shall be of a nature implying statistical independence.

THEOREM 2.1.

$$G_i \vee S_i \vee N_i = 1. \tag{56}$$

Proof: Hypothesis 2.1.

COROLLARY 2.1.

$$\gamma_i + \sigma_i + \nu_i = 1. \tag{57}$$

Proof: From Theorem 2.1, $G_i \vee S_i \vee N_i = 1$, and, by customary interpretation of the conditions in Hypothesis 2.1,

$$G_i \wedge S_i = S_i \wedge N_i = N_i \wedge G_i = 0.$$

Apply twice the probability axioms: If $D = A \vee B$, and $E = A \wedge B$, then $\delta = \alpha + \beta - \epsilon$. If $X = 0$, $\xi = 0$.

CONVENTION 2.2. We write λ_i for

$$\sigma_i + \nu_i = 1 - \gamma_i. \tag{58}$$

‡ It is to be understood that probabilities considered simultaneously are over coherently referenced sample spaces. We may, for example, think of the states of diodes in a quad after some fixed tour of duty.

Thus, λ_i is the failure probability of the ith diode. We write η and ζ for the probabilities of normal operability and failure, respectively, of the quad (device of Fig. A9.1).

HYPOTHESIS 2.3. The quad fails if and only if at least one side is shorted or both sides are open.

THEOREM 2.2. The quad failure condition is

$$(S_1 \wedge S_2) \vee (S_3 \wedge S_4) \vee (N_1 \wedge N_3) \vee (N_2 \wedge N_3) \vee (N_1 \wedge N_4) \vee (N_2 \wedge N_4). \quad (59)$$

Proof: Hypothesis 2.3 and direct logical computation and simplification.

COROLLARY 2.2. The quad failure probability is

$$(\dagger)\zeta = \sigma_1\sigma_2 + \sigma_3\sigma_4 - \sigma_1\sigma_2\sigma_3\sigma_4 + \nu_1\nu_3 + \nu_2\nu_3 + \nu_1\nu_4 + \nu_2\nu_4 - \nu_1\nu_2\nu_3 - \nu_1\nu_2\nu_4 - \nu_1\nu_3\nu_4 - \nu_2\nu_3\nu_4 + \nu_1\nu_2\nu_3\nu_4. \quad (60)$$

Proof: Theorem 2.2 and arguments similar to those in the proof of Corollary 2.1.

COROLLARY 2.3.

$$\begin{aligned}(\dagger\dagger)\zeta = {} & \nu_1\nu_2(1 - \lambda_3\lambda_4) + \nu_1\nu_3(1 - \lambda_2\lambda_4) + \nu_1\nu_4(1 - \lambda_2\lambda_3) \\ & + \nu_2\nu_3(1 - \lambda_1\lambda_4) + \nu_2\nu_4(1 - \lambda_1\lambda_3) + \nu_3\nu_4(1 - \lambda_1\lambda_2) \\ & - \lambda_2(1 - \lambda_3\lambda_4)\nu_1 - \lambda_1(1 - \lambda_3\lambda_4)\nu_2 - \lambda_4(1 - \lambda_1\lambda_2)\nu_3 \\ & - \lambda_3(1 - \lambda_1\lambda_2)\nu_4 - (1 - \lambda_4)\nu_1\nu_2\nu_3 - (1 - \lambda_3)\nu_1\nu_2\nu_4 \\ & - (1 - \lambda_2)\nu_1\nu_3\nu_4 - (1 - \lambda_1)\nu_2\nu_3\nu_4 + \lambda_1\lambda_2 + \lambda_3\lambda_4 - \lambda_1\lambda_2\lambda_3\lambda_4. \quad (61)\end{aligned}$$

Proof: Convention 2.2, Corollary 2.2., and direct computation.

(1). *Special Case*

HYPOTHESIS 2.1.1. The diodes are identical with respect to probability characteristics; that that is,

$$\lambda_i = \lambda_j \quad \text{and} \quad \nu_i = \nu_j; \quad i, j = 1, 2, 3, 4. \quad (62)$$

CONVENTION 2.1.1. In the special case, we write

$$\lambda_i = \lambda, \nu_i = \nu, \sigma_i = \sigma; \quad i = 1, 2, 3, 4. \quad (63)$$

THEOREM 2.1.1. In the special case,

$$(*)\zeta = 6(1-\lambda^2)\nu^2 - 4\lambda(1-\lambda^2)\nu - 4(1-\lambda)\nu^3 + 2\lambda^2\left(1-\frac{\lambda^2}{2}\right). \quad (64)$$

Proof: Corollary 2.3 and Convention 2.1.1.

CONVENTION 2.3. We assume λ_i constant.

$$\bar{\zeta} = \frac{1}{\lambda}\left[2(1-\lambda^2)\nu^2 - 2\lambda(1-\lambda^2)\nu^2 - (1-\lambda)\nu^4 + 2\lambda^2\left(1-\frac{\lambda^2}{2}\right)\nu\right]_0^\lambda. \quad (65)$$

Proof: Apply Theorem 2.1.1 to

$$\bar{\zeta} = \frac{1}{\lambda}\int_0^\lambda \zeta d\nu.$$

COROLLARY 2.1.1.

$$\bar{\zeta} = 2\lambda^2\left(1-\frac{\lambda^2}{2}\right) - \lambda^3(1-\lambda) = \lambda^2(2-\lambda). \quad (66)$$

Proof: Evaluate in Theorem 2.1.2.

THEOREM 2.1.3. The minimum and maximum values of ζ are at

$$\nu = \frac{3(1+\lambda) - \sqrt{3(1+\lambda)(3-\lambda)}}{6}$$

and

$$\nu = \frac{3(1+\lambda) + \sqrt{3(1+\lambda)(3-\lambda)}}{6}, \quad (67)$$

respectively.

Proof: Set $d\zeta/d\nu = 0$ and test roots in $d^2\zeta/d\nu^2$.

COROLLARY 2.1.2. The maximum and minimum values of ζ are

$$1 \pm \frac{(3-\lambda)(1-\lambda^2)}{9}\sqrt{3(3-\lambda)(1+\lambda)}. \quad (68)$$

Proof: Substitute into (*) from Theorem 2.1.3.

REMARK 2.1.1. From (*),

$$\zeta\,|_{\nu=0} = 2\lambda^2\left(1-\frac{\lambda^2}{2}\right) \quad \text{and} \quad \zeta\,|_{\nu=\lambda} = 4\lambda^2(1-\lambda) + \lambda^4. \quad (69)$$

REMARK 2.1.2. Since the restrictions, $0 \leq \nu \leq \lambda \leq 1$, are required, the maximum-yielding value of ν in Theorem 2.1.3 is excluded and only the minimum is attainable. We have as examples:

EXAMPLE VALUES OF ζ

λ and ν		Values of ζ			
λ	ν opt.	$\zeta \mid \nu = 0$	$\bar{\zeta}$	$\zeta \mid \nu = \tau$	ζ min.
.1	.0333	.0197	.019	.036	.01749
.01	.00416	.0001998	.000199	.000396	.000078
.001	.00036	.00000199999	.000001999	.000003996	.0000002

Fig. A9.2

(2). *General Case*

THEOREM 2.2.1. Necessary conditions for attainment of a minimum (††) are:

$$(**)\ \nu_2(1 - \lambda_3\lambda_4) + \nu_3(1 - \lambda_2\lambda_4)a + \nu_4(1 - \lambda_2\lambda_3) - \lambda_2(1 - \lambda_3\lambda_4)$$
$$- (1 + \lambda_4)\nu_2\nu_3 - (1 - \lambda_3)\nu_2\nu_4 - (1 - \lambda_2)\nu_3\nu_4 = 0. \quad (70)$$

$$_1(1 - \lambda_3\lambda_4) + \nu_3(1 - \lambda_1\lambda_4) + \nu_4(1 - \lambda_1\lambda_3) - \lambda_1(1 - \lambda_3\lambda_4)$$
$$- (1 - \lambda_4)\cdot(\nu_1\nu_3) - (1 - \lambda_3)\nu_1\nu_4 - (1 - \lambda_1)\nu_3\nu_4 = 0. \quad (71)$$

$$\nu_4(1 - \lambda_1\lambda_2) + \nu_1(1 - \lambda_2\lambda_4) + \nu_2(1 - \lambda_1\lambda_4) - \lambda_4(1 - \lambda_1\lambda_2)$$
$$- (1 - \lambda_2)\nu_1\nu_4 - (1 - \lambda_1)\nu_2\nu_4 - (1 - \lambda_4)\nu_1\nu_2 = 0. \quad (72)$$

$$\nu_3(1 - \lambda_1\lambda_2) + \nu_1(1 - \lambda_2\lambda_3) + \nu_2(1 - \lambda_1\lambda_3) - \lambda_3(1 - \lambda_1\lambda_2)$$
$$- (1 - \lambda_2)\nu_1\nu_3 - (1 - \lambda_1)\nu_2\nu_3 - (1 - \lambda_3)\nu_1\nu_2 = 0. \quad (73)$$

REMARKS 2.2.1. Shown on Fig. A9.3 is the Wronskian matrix of (††). This matrix clearly fails to be positive definite and, thus, the customary sufficient condition for a minimum fails; this indicates necessity for extensive analysis of the behavior of ζ in the region of the roots of (**). We therefore abandon this general case for the present study.

WRONSKIAN OF

0	$(1-\lambda_3\lambda_4)-(1-\lambda_4)\nu_3-(1-\lambda_3)\nu_4$
$(1-\lambda_3\lambda_4)-(1-\lambda_4)\nu_3-(1-\lambda_3)\nu_4$	0
$(1-\lambda_2\lambda_4)-(1-\lambda_4)\nu_2-(1-\lambda_2)\nu_4$	$(1-\lambda_1\lambda_4)-(1-\lambda_4)\nu_1-(1-\lambda_1)\nu_4$
$(1-\lambda_2\lambda_3)-(1-\lambda_3)\nu_2-(1-\lambda_2)\nu_3$	$(1-\lambda_1\lambda_3)-(1-\lambda_3)\nu_1-(1-\lambda_1)\nu_3$

Fig. A9.3

3. The Dual Diode Quad

Consider a device as in Fig. A9.4. Using notation and methods similar to those of Section 2 (The Diode Quad), we find

THEOREM 3.1. The dual quad failure condition is

$$(N_1 \wedge N_4) \vee (N_2 \wedge N_3) \vee (S_1 \wedge S_2) \vee (S_1 \wedge S_3) \vee (S_2 \wedge S_4) \vee (S_3 \wedge S_4). \quad (74)$$

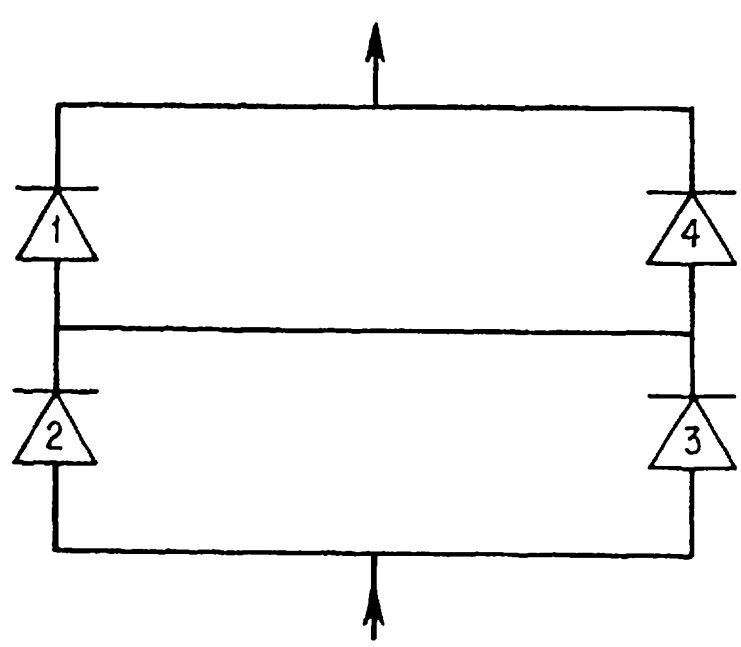


Fig. A9.4. Dual diode quad.

S FROM(††)

$(1-\lambda_2\lambda_4)-(1-\lambda_4)\nu_2-(1-\lambda_2)\nu_4$	$(1-\lambda_2\lambda_3)-(1-\lambda_3)\nu_2-(1-\lambda_2)\nu_3$
$(1-\lambda_1\lambda_4)-(1-\lambda_4)\nu_1-(1-\lambda_1)\nu_4$	$(1-\lambda_1\lambda_3)\nu_1-(1-\lambda_3)\nu_1-(1-\lambda_1)\nu_3$
0	$(1-\lambda_1\lambda_2)-(1-\lambda_2)\nu_1-(1-\lambda_1)\lambda_2$
$(1-\lambda_1\lambda_2)-(1-\lambda_2)\nu_1-(1-\lambda_1)\nu_2$	0

THEOREM 3.2.

$$\bar{\xi} = 6(1-\lambda^2)\sigma^2 - 4\lambda(1-\lambda^2)\sigma - 4(1-\lambda)\sigma^3 + 2\lambda^2\left(1-\frac{\lambda^2}{2}\right). \quad (75)$$

4. Notation for the Transistor Block

Consider a device as in Fig. A9.5.

CONVENTION 4.1. We shall write:

P for the event of a short between Y and Z.
R for the event of an XY or XZ short.
A for the event of opening at X, Y, or Z.

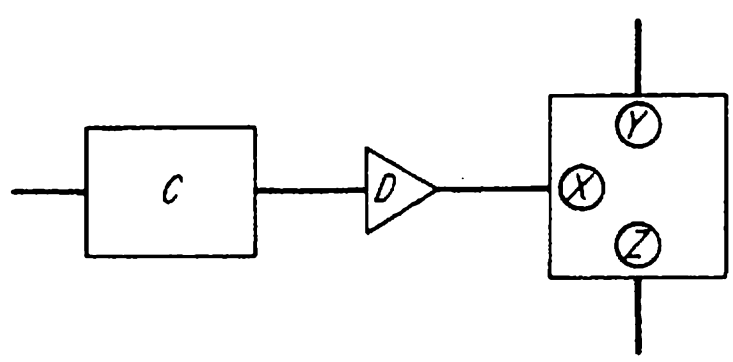


Fig. A9.5. Balance block.

N and S for opening and shorting, respectively of D.
B for shorting on balance circuit C.
M for opening on balance circuit C.

Corresponding Greek letters will indicate the probability of these various events.†

5. The Transistor Block Quad

Consider a device as in Fig. A.9.6.

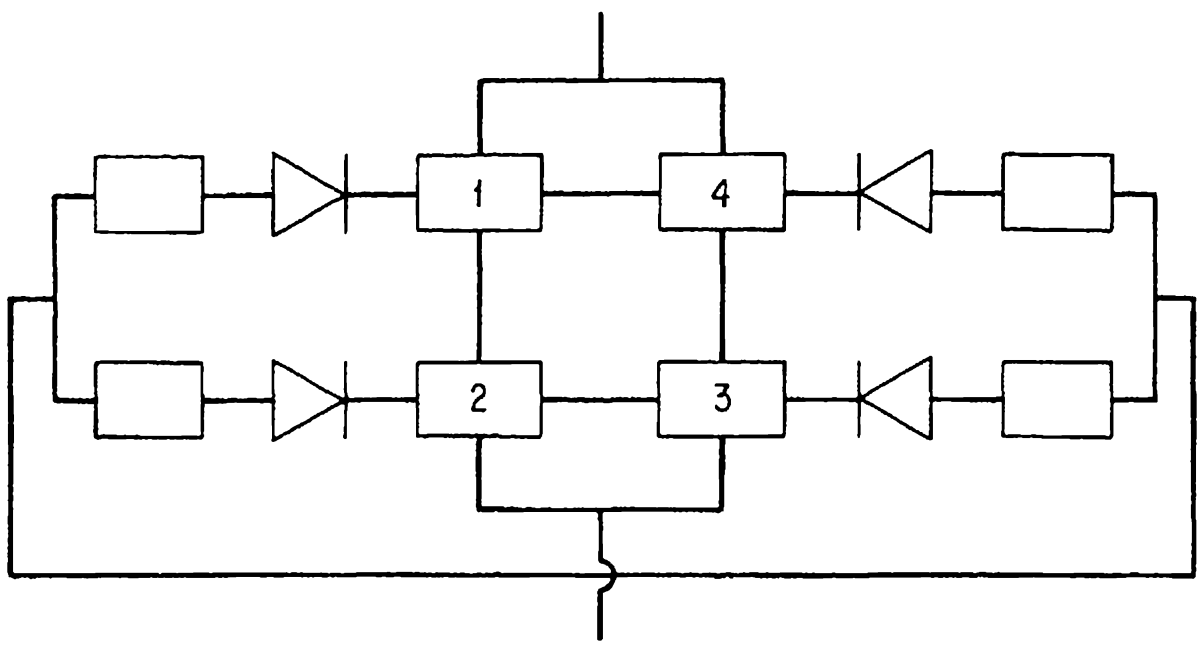


Fig. A9.6. Transistor quad.

HYPOTHESIS 5.1. The failure condition for the device of Fig. A9.6 is

$$
\begin{aligned}
&[(A_1 \vee M_1 \vee N_1 \wedge (A_3 \vee M_3 \vee N_3)] \\
&\qquad \vee (A_1 \vee M_1 \vee N_1) \wedge (A_4 \vee M_4 \vee N_4)] \\
&\qquad \vee (A_2 \vee M_2 \vee N_2 \wedge (A_3 \vee M_3 \vee N_3)] \\
&\qquad \wedge [(A_2 \vee M_2 \vee N_2) \wedge (A_4 \vee M_4 \vee N_4)] \\
&\vee [(P_1 \vee R_1) \wedge (P_2 \vee R_2] \vee [(P_3 \vee R_3) \wedge (P_4 \vee R_4] \\
&\vee [B_1 \wedge S_1 \wedge R_1] \vee [B_2 \wedge S_2 \wedge R_2] \vee [P_3 \wedge S_3 \wedge R_3] \\
&\qquad \vee [B_4 \wedge S_4 \wedge R_4]. \qquad (76)
\end{aligned}
$$

† Recall the admonition stated in footnote ‡ on page 252.

HYPOTHESIS 5.2.

(1) $A_i \wedge R_i = P_i \wedge A_i = 0$. (slightly unrealistic) (77)

(2) $B_i \wedge M_i = 0$. (78)

(3) All elements fail in a nature of statistical independency.

CONVENTION 5.1. We denote:

Probability of device failure by ζ.
Probability of transistor failure by τ.
Probability of diode failure by λ.
Probability of balance circuit failure by ψ.

THEOREM 5.1. We have

$$\pi_1 + \rho_i + \alpha_i = \tau_i, \tag{79}$$

$$\nu_i + \sigma_i = \lambda_i, \tag{80}$$

and

$$\rho_i + \mu_i = \psi_i \tag{81}$$

Proof: Direction from Convention 5.1, Hypothesis 5.2, and previous sections.

HYPOTHESIS 5.3. If ξ is any of π, ρ, α, τ, ν, σ, λ, β, μ, or ψ,

$$\text{then } \xi_i = \xi_j \text{ for } i, j = 1, 2, 3, 4. \tag{82}$$

HYPOTHESIS 5.4. The probability of simultaneous occurrence of more than one of the events listed below is negligible, and we may regard them as pairwise mutually exclusive:

(1) Open circuit (first four equal brackets in Hypothesis 5.1).
(2) Block short (center two square brackets in Hypothesis 5.1).
(3) Quad short (center two square brackets in Hypothesis 5.1).

THEOREM 5.2. We have

$$\begin{aligned}\xi = {} & 4(\alpha + \mu + \nu - \alpha\mu - \alpha\nu - \mu\nu + \alpha\mu\nu)^2 \\ & -2(\alpha + \mu + \nu - \alpha\mu - \alpha\nu - \mu\nu + \alpha\mu\nu)^3 \\ & - (\alpha + \mu + \nu - \alpha\nu - \alpha\mu - \mu\nu + \alpha\mu\nu)^4 \\ & + 2(\pi + \rho - \pi\rho)^2 - (\pi + \rho - \pi\rho)^4 \\ & + 4\beta\sigma\rho - 4(\beta\sigma\rho)^2 - 4(\beta\sigma\rho)^3 + (\beta\sigma\rho)^4.\end{aligned} \tag{83}$$

Proof: Direct computation using previous hypothesis.

COROLLARY 5.1. Ignoring higher degree terms in each type of failure we have an approximation:

$$\begin{aligned}\zeta &= 4(\alpha + \mu + \nu)^2 + 2(\pi + \rho)^2 + 4\beta\sigma\rho \\ &= 4(\alpha + \mu + \nu)^2 + 2(\pi - \alpha)^2 + 4(\phi - \mu)(\eta - \nu)\rho. \qquad (84)\end{aligned}$$

REMARKS 5.1. Upon assuming $\alpha = \pi = \rho = \tau/3$,

$$\nu = \sigma = \lambda/2 \quad \text{and} \quad \beta = \mu = \phi/2, \qquad (85)$$

one finds from Corollary 5.1,

$$\zeta = 1/3(4\tau^2 + 3\lambda^2 + 3\phi^2 + 4\lambda\tau + 4\tau\phi + 6\lambda\phi + \tau\lambda). \qquad (86)$$

REMARKS 5.2. Upon assuming $\alpha = \pi = \rho = \tau/3$, $\sigma = 1$, and $\beta = \mu = \phi/2$, one finds that for diodes deleted we have:

$$\xi = 1/3(4\tau^2 + 6\tau\phi + 3\phi^2). \qquad (87)$$

COROLLARY 5.2. (Local Summary). Upon making various simplifying assumptions, the majority of which tend to overestimation, one finds, upon applying the principle of insufficient reason to failure types, that an approximation to failure probability is:

$$\zeta = 1/3(4\tau^2 + 3\lambda^2 + 3\phi^2 + 4\lambda\tau + 4\tau\phi + 6\lambda\phi + \tau\lambda\phi) \qquad (88)$$

for the device of Fig. A9.6, and is

$$\zeta_\alpha = 1/3(4\tau^2 + 6\tau\phi + 3\phi^2) \qquad (89)$$

with the diodes deleted from the same device and the possible additional signal loss probability caused by this ignored.

REMARKS 5.3. Sample approximate values of ζ from Corollary 5.2 with "reasonable" failure probabilities are:

SAMPLE VALUES OF ζ

τ	λ	ϕ	ζ	ζ_α
10^{-2}	10^{-2}	10^{-3}	4.01×10^{-4}	1.54×10^{-4}
10^{-3}	10^{-3}	10^{-3}	8×10^{-6}	4×10^{-6}
10^{-2}	10^{-3}	10^{-3}	1.64×10^{-4}	1.54×10^{-4}
10^{-3}	10^{-2}	10^{-3}	1.37×10^{-4}	4×10^{-6}

Fig. A9.7

6. Tentative Conclusions

Subject to the restrictions pointed out in Section 1 and to the various simplifying assumptions made in the various applicable sections we have:

(1) A rule of the thumb, long suspected, is that (quad failure probability) $= k$ (max probability of component failure)^2, where $1 \leq k \leq 10$, for cases considered.

(2) In the case of a diode quad, one should employ the quad per se with the understanding that most failures are due to faulty semiconductor material, but one should switch to the dual quad noting that most failures are of a mechanical nature. The swing factor between optimum failure modes and worst failure modes of uniform diodes with "reasonable" failure probability in a quad or dual quad may be as high as 20.

(3) Other things being equal, if the probability of signal loss owing to back emf through a base short in a transistor is fairly negligible, the D-diode may well be omitted in the transistor block quad.

(4) Putting aside questions of such subtlety as degrading, derating, and overloading, a good case can be made for quads *vs.* single semiconductors as a reliability gain in *normal operational environments.*

7. Further Work

Additional studies have been carried out as refinements of the work discussed here. One of these has been a full scale study of the sensitivity of ξ for the diode quad to the special case hypothesis 2.1.1. This study, utilizing experimental design principles to approximate the ξ density function, used several hundred samples based on probability distributions constructed from several years' data.

Appendix 10

Elements of a Behavioral Theory of Static Decisions

"As the ancients wisely say
Have a care o'th' main chance,
And look before you ere you leap;
For as you sow y'ere like to reap."
—BUTLER

1. Introduction

Several major behaviorally inclined attempts at description of the decision process have been made.† Unfortunately, none of these gives an empirically satisfactory performance. In this appendix, we attempt to initiate a new approach under a philosophy not of immediate simplification but rather of the utmost meaningful generality.

† *Cf.* H. A. Simon, *A Behavioral Model of Rational Choice,* (Santa Monica: The RAND Corporation, 1953), p. 365. *See also* L. J. Savage, *The Foundations of Statistics,* (New York: Wiley, 1954).

2. Static vs. Dynamic

It is felt that a satisfactory static theory would permit an intelligent expansion into dynamics. The subject is so rich conceptually that dynamic considerations from the beginning tend to have a "last straw" effect. Here, therefore, only the static case will be considered.

3. The Static Decision Environment

We shall now elaborate the static decision environment (SDE) for a fixed utilizing entity (i.e., a decision-making element). Unless otherwise explicitly noted, all items are referenced both to the fixed utilizing entity and to a fixed time "τ." R^+ will denote the non-negative real numbers.

(1) There is a well-defined system (whose definition may be unknown to us) called the *universe*. The phase space of the universe is written S^{**} and called *universe space*.

(2) There is an equivalence relation $R \subset S^{**} \otimes S^{**}$ called *resolution*.

$$S^* = S^{**}/R \tag{90}$$

is called resolvable space.

(3) There is an equivalence relation $I \subset S^* \otimes S^*$ called *inherent indifference*.

$$S = S^*/I \tag{91}$$

is called *cognizant space*.

(4) There is an upper conditionally complete vector lattice $\vee$ called *value space*.

(5) There is a set A called *action space*.

(6) There is a mapping $c: A \to \vee$ called *cost*.

(7) There is a mapping $u: S \otimes R^+ \to \vee$ called utility. $u(s, \varphi)$ is the "utility" to the fixed utilizing entity of having state "s" obtain at time "$\tau + \varphi$."

(8) There is a mapping $\delta: R^+ \to R^+$ called *discount*. $\delta(\varphi)$ is the weight assigned time "$\tau + \varphi$" relative to other times by the utilizing entity. Depending upon motivation in the decision under consideration, δ might be smoothly decreasing for a typical operational decision,

exponentially decaying for a crisis situation, or, among many other forms, almost a delta-function for a timing problem.

(9) There is a mapping $\pi: S \otimes A \otimes R^+ \otimes S \rightarrow [0, 1]$ called *transition* probability. $\pi(s_1, a, \varphi, s_2)$ is to be regarded as the utilizing entity's estimate of the probability so that if state "s_1" obtains at time "τ" and if act "a" is enacted at time "τ," then state "s_2" will obtain at time "$\tau + \varphi$." By $\hat{\pi}$ we shall indicate the appearance of mapping π in a process where all variables but "s_2" are temporarily held constant.

(10) There is a mapping $\rho: S \rightarrow [0, 1]$ called *present estimate.* $\rho(s)$ is the probability as estimated by the fixed utilizing entity that state "s" obtains at time "τ."

(11) We shall assume that, by joint virtue of a generality of integrals utilized and by the conceptual limitations of the utilizing entity, all integrations indicated below are valid.

(12) The *expected value* of enacting "a" at time "τ" is

$$e(a) = \int_S \int_0^\infty \int_S u(s_2, \tau + \varphi)\delta(\tau + \varphi)\pi(s_1, a, \varphi, s_2)\rho(s_1)\, d\hat{\pi}\, d\varphi\, d\rho - c(a) \tag{92}$$

(13) The expected penalty associated with enacting "a" at time "τ" is (assuming "a" bounded above):

$$v(a) = \sup_{k \epsilon a} e(k) - e(a). \tag{93}$$

(14) The *expected value* of enacting "a" at time "τ" given "s_1" is

$$e(a \mid s_1) = \int_0^\infty \int_S u(s_2, \tau + \varphi)\delta(\tau + \varphi)\pi(s_1, a, \varphi, s_2)\, d\hat{\pi}\, d\varphi - c(a). \tag{94}$$

(15) The regret associated with enacting "a" at time "τ" given "S_1" is

$$v(a \mid s_1) - \sup_{k \epsilon A} e(k \mid s_1) - e(a \mid s_1). \tag{95}$$

4. Estimates and Criteria

As will be immediately recognized, we have not specified:

(a) How utilizing entity is to estimate ρ. (For example, the Laplace principle may be invoked.)

(b) How utilizing entity is to estimate π.

(c) What criteria utilizing entity will involve to select "a." Expectation may be maximized, regret minimized, or any other of the many criteria that have been suggested may be used. A subset of A on the interpretation of other constraints may be discounted.

These methods of estimation and the selection criteria utilized are themselves the parameters of the behavioral theory, which operates, as a result, at a level of logical sophistication above cutomary parametric statistical theories. It is to be emphasized that these parameters will generally be different for different static decision environments even for the same utilizing entity (for example, there will generally be a difference of criteria employed in high confidence situations and in low confidence decisions). We thus specifically deny, for example, Savage's form of the "sure thing principle," at least in the context of his other postulates. Inference of the behavioral parameters in various standardized situations will permit classification of a utilizing entity, and subsequent prediction *vs.* secondary experimentation will permit inference of his degree of rationality *in the sense of consistency*. This is, then, an empirically based and operationally conceived static decision theory whose fruitfulness will depend upon good experimental design, adequate interpretation of results, and, in some instances, rapport with the subject of investigation with respect to *interpretation of the theory elements*.

Appendix 11

Potential Mechanizations of Safety Procedures During In-Flight Emergencies; An Example

". . . to insure that aviation safety is stressed in connection with mission accomplishment . . ."
NAVPERS 10822-A

1. Introduction

This discussion will be directed toward adaptation of a computer system to relieve pilots of both required decision making and physical tasks when emergencies arise. Only the specific emergency of power loss during take-off is presented here. The data, detector, and actuator requirements, and the system difficulties to be expected, arising from both technical and subjective situations are presented. The appendix describes the procedure for utilizing the T-33. Slightly modified computers would be required for different types of planes. More highly modified computers would be required if multi-engined aircraft were considered.

In this appendix, several charts are used to describe functionally the several procedures. Fig. A11.1 shows the procedures undertaken *by the*

pilot in the event of a power loss during take-off and the decisions that he must make to implement a specific procedure. It should be noted that the J-33-A20 engine has no emergency fuel system. Therefore, when there is less than 80 per cent power, flight should not be attempted. See Fig. A11.1. Most of the pilot's decisions are "trained" into him (automatic reactions) and stem from the two basic questions:

(1) Can the plane be stopped in the room allowed?—and if not—

(2) Can the plane fly?

Pilot decisions, valid though they may seem, are in reality estimates. A computer system is needed to provide mathematically sound, evaluated *decisions* that will reduce pilot error resulting from inaccurately estimated decisions. Furthermore, it will be expected that this system be expanded so as to control automatically those *actions* that the pilot normally performs today. Computer action is to be based on the *computed* choice for best chance of survival of both pilot and aircraft. Override capabilities, or the amount of override capabilities, may be limited by pilot acceptance or other constraints.

Figure A11.2 shows the fifteen (only fifteen are possible in this case) separate sub-procedures that would be followed. These sub-procedures would result from the pilot's decisions indicated in Fig. A11.1. Figure A11.2 does not show the decisions the pilot made, only the conclusions that were reached and the actions that had to be performed as a result of his conclusions. This same type of chart could be used to illustrate what the computer's decisions based on factual data, not estimates, would be and what procedures would be adapted as a result of these decisions. The steps that the computer would initiate would not necessarily follow the same sequence as those shown in Fig. A11.2.

If the man is essentially "short circuited" from the system (no override), and the proposed actuators and detectors are available, or can be available, the situation will be that represented by Fig. A11.3. This chart shows the procedures to be decided by the computer and the remaining steps to be performed by the pilot. The steps that can be done automatically are not shown; neither are the decisions. *It should be noted that there are essentially only four steps carried out by the pilot; all other operations are done automatically. These four steps are concerned with either control of the aircraft or the act of leaving the aircraft.*

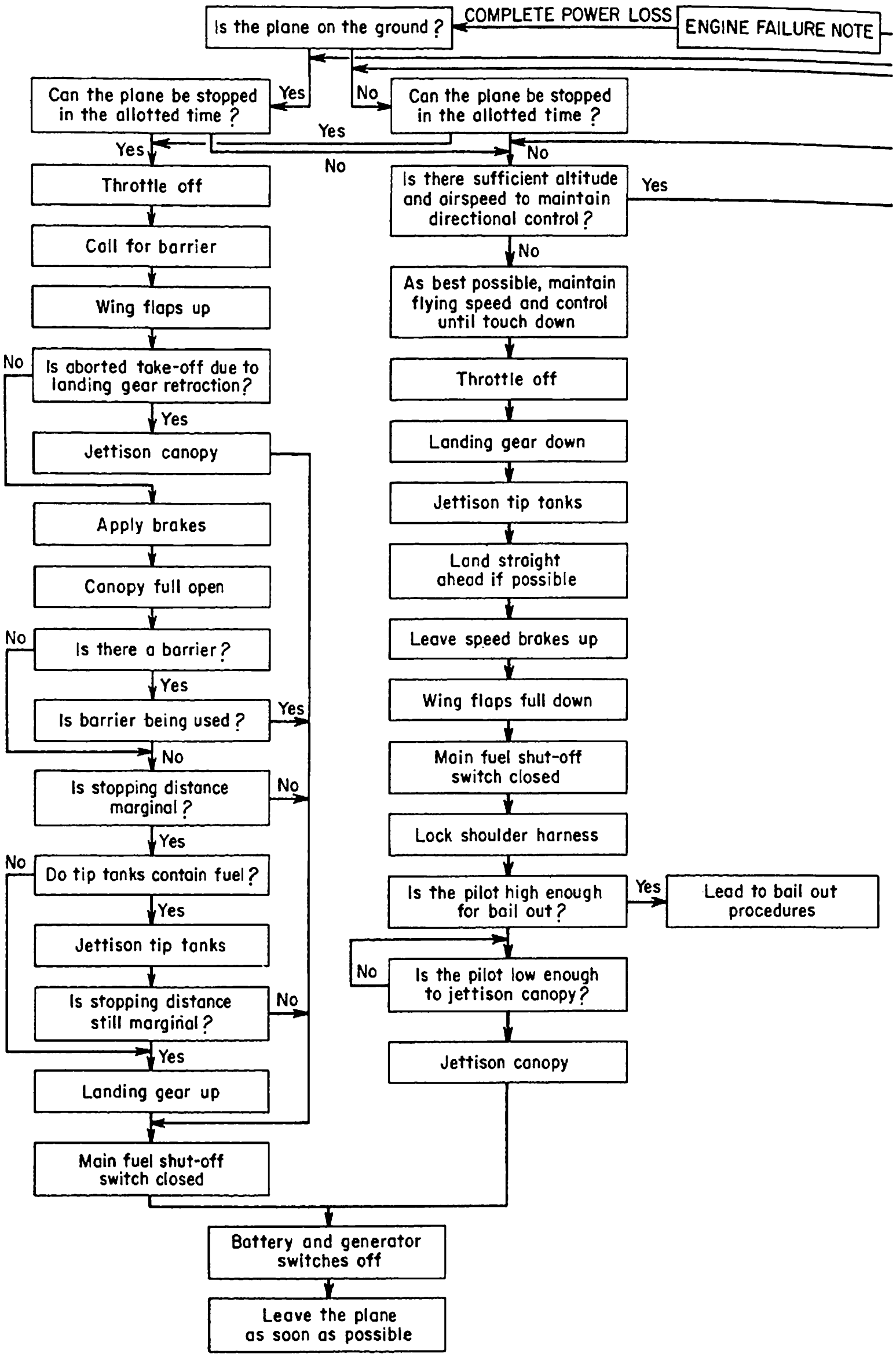
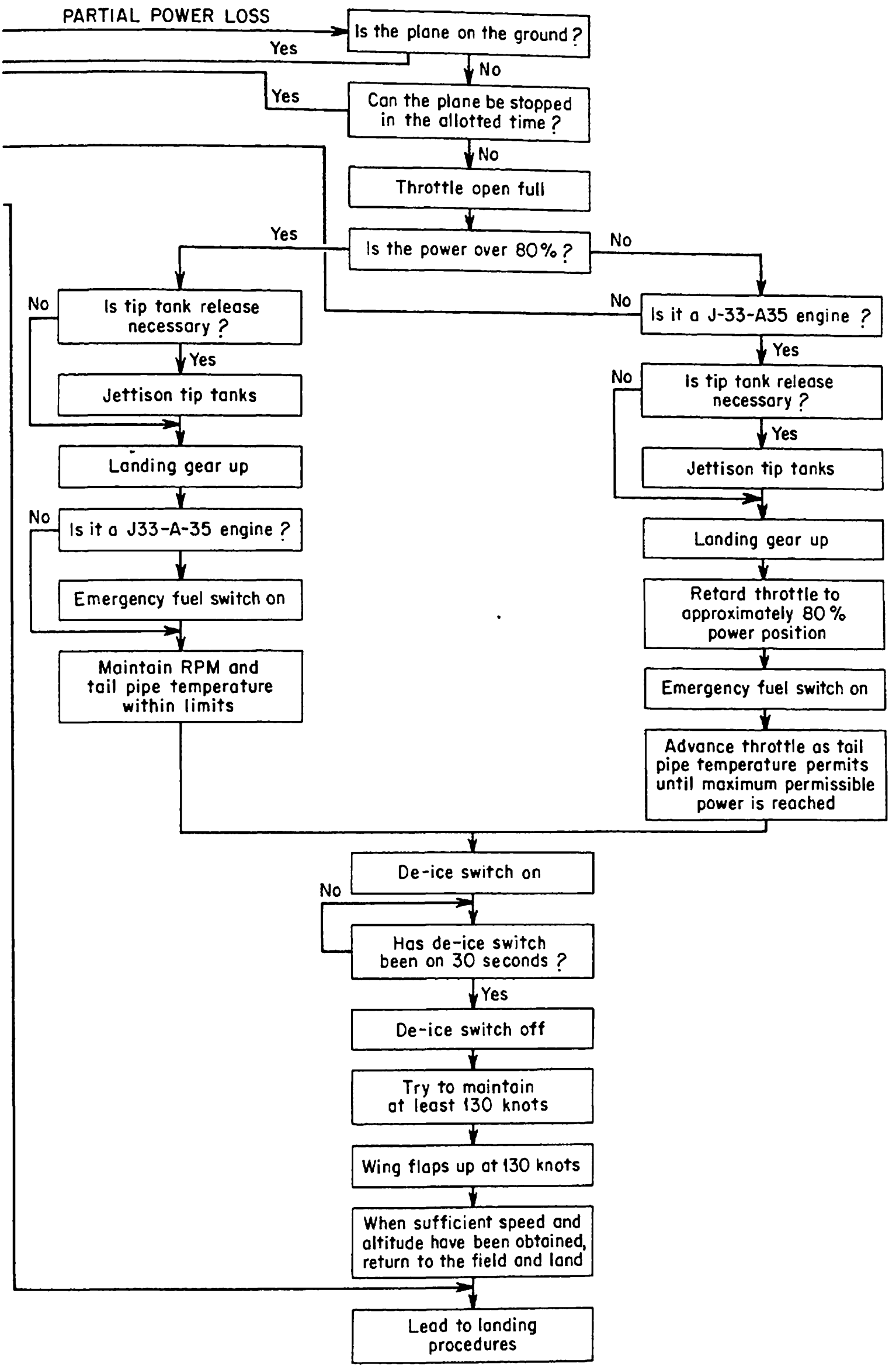


Fig. A11.1. Decisions and procedures (pilot decisions made and procedures followed when power failure occurs during take-off).


PARTIAL POWER LOSS
Is the plane on the ground ?
Yes
No
Can the plane be stopped in the allotted time ?
Yes
No
Throttle open full
Is the power over 80% ?
Yes
No
Is tip tank release necessary ?
No
Yes
Jettison tip tanks
Landing gear up
Is it a J33-A-35 engine ?
No
Emergency fuel switch on
Maintain RPM and tail pipe temperature within limits
Is it a J-33-A35 engine ?
No
Yes
Is tip tank release necessary ?
No
Yes
Jettison tip tanks
Landing gear up
Retard throttle to approximately 80% power position
Emergency fuel switch on
Advance throttle as tail pipe temperature permits until maximum permissible power is reached
De-ice switch on
No
Has de-ice switch been on 30 seconds ?
Yes
De-ice switch off
Try to maintain at least 130 knots
Wing flaps up at 130 knots
When sufficient speed and altitude have been obtained, return to the field and land
Lead to landing procedures

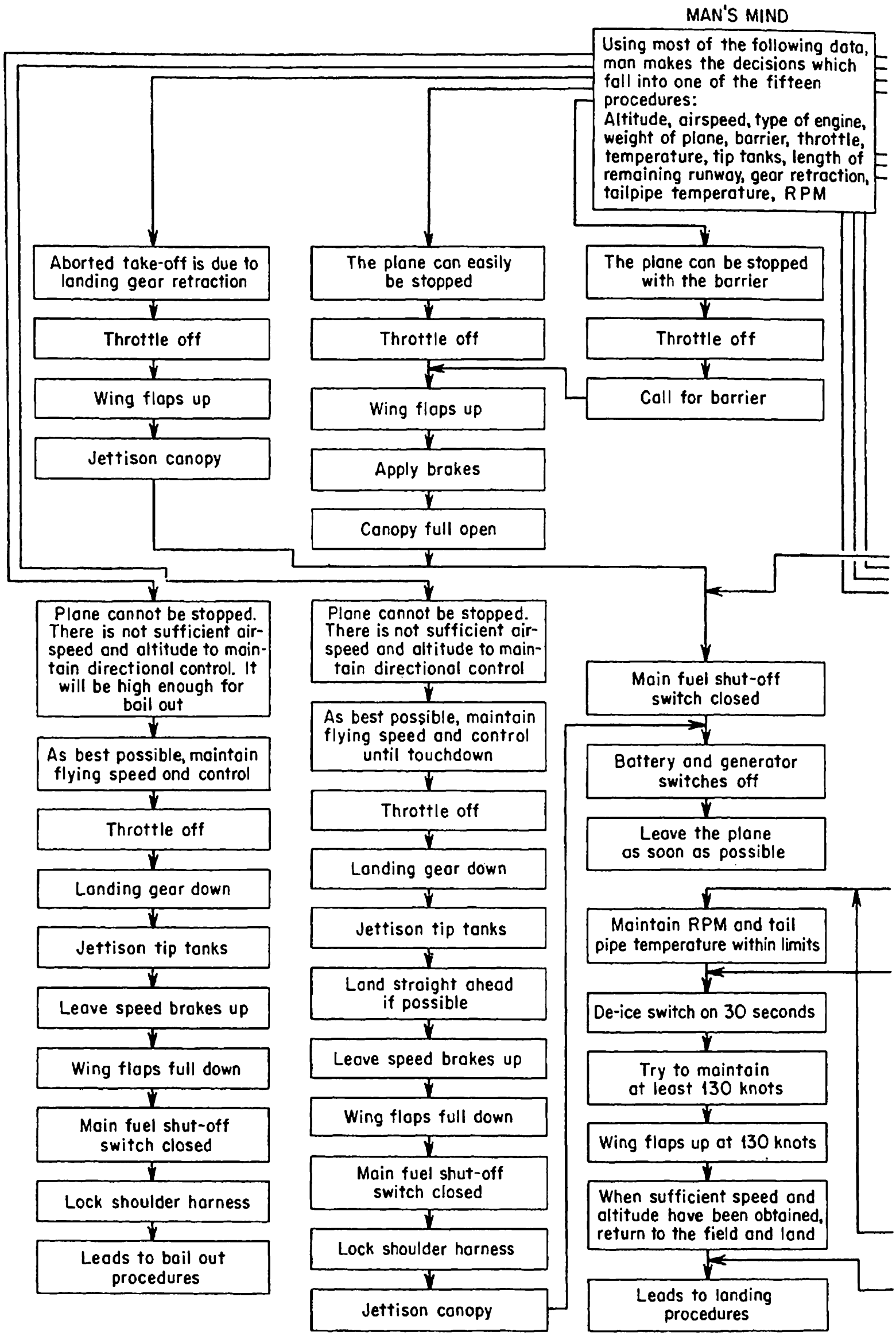
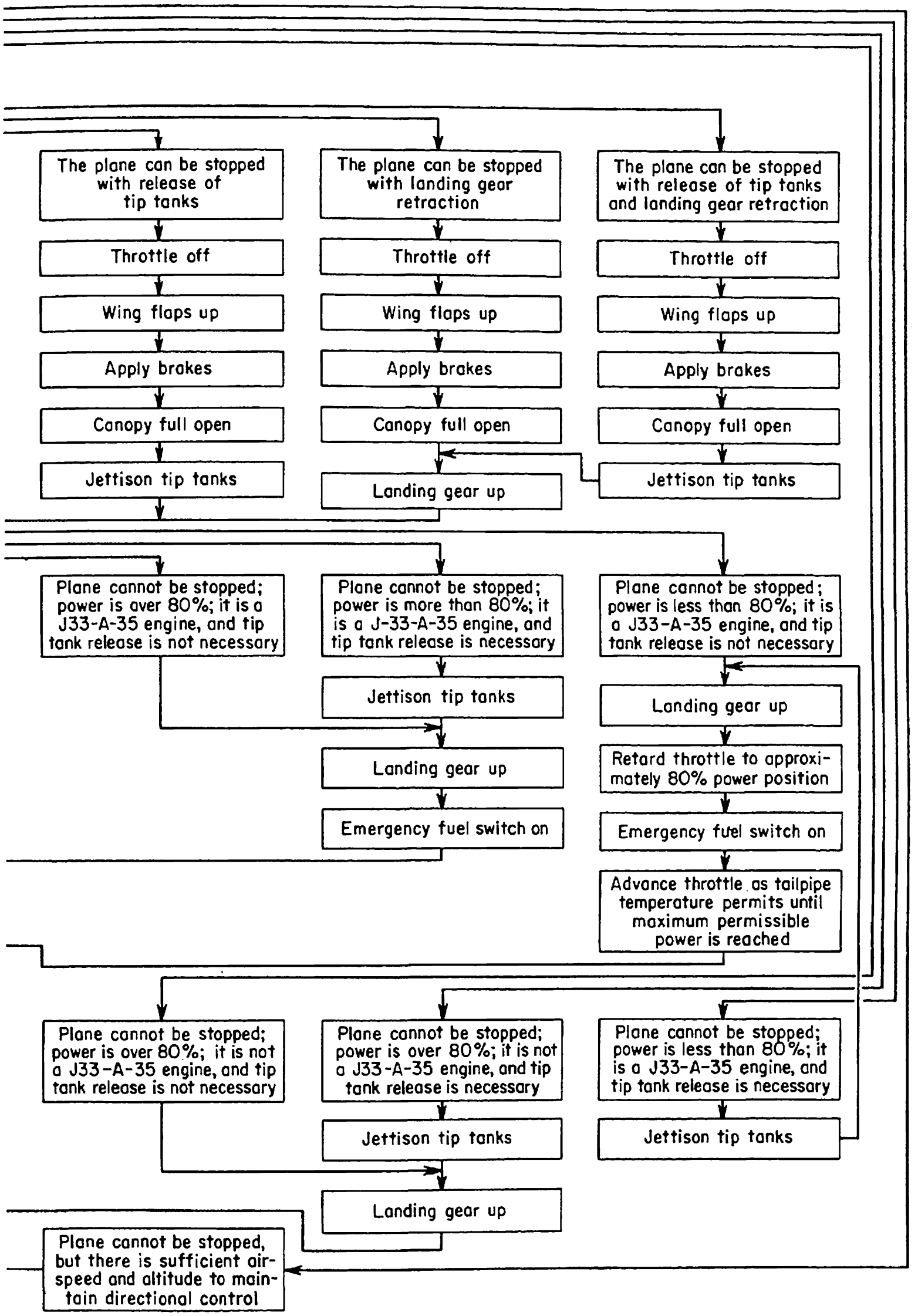


Fig. A11.2. Fifteen separate procedures.


The plane can be stopped with release of tip tanks
Throttle off
Wing flaps up
Apply brakes
Canopy full open
Jettison tip tanks
The plane can be stopped with landing gear retraction
Throttle off
Wing flaps up
Apply brakes
Canopy full open
Landing gear up
The plane can be stopped with release of tip tanks and landing gear retraction
Throttle off
Wing flaps up
Apply brakes
Canopy full open
Jettison tip tanks
Plane cannot be stopped; power is over 80%; it is a J33-A-35 engine, and tip tank release is not necessary
Plane cannot be stopped; power is more than 80%; it is a J-33-A-35 engine, and tip tank release is necessary
Jettison tip tanks
Landing gear up
Emergency fuel switch on
Plane cannot be stopped; power is less than 80%; it is a J33-A-35 engine, and tip tank release is not necessary
Landing gear up
Retard throttle to approximately 80% power position
Emergency fuel switch on
Advance throttle as tailpipe temperature permits until maximum permissible power is reached
Plane cannot be stopped; power is over 80%; it is not a J33-A-35 engine, and tip tank release is not necessary
Plane cannot be stopped; power is over 80%; it is not a J33-A-35 engine, and tip tank release is necessary
Jettison tip tanks
Landing gear up
Plane cannot be stopped; power is less than 80%; it is a J33-A-35 engine, and tip tank release is necessary
Jettison tip tanks
Plane cannot be stopped, but there is sufficient airspeed and altitude to maintain directional control

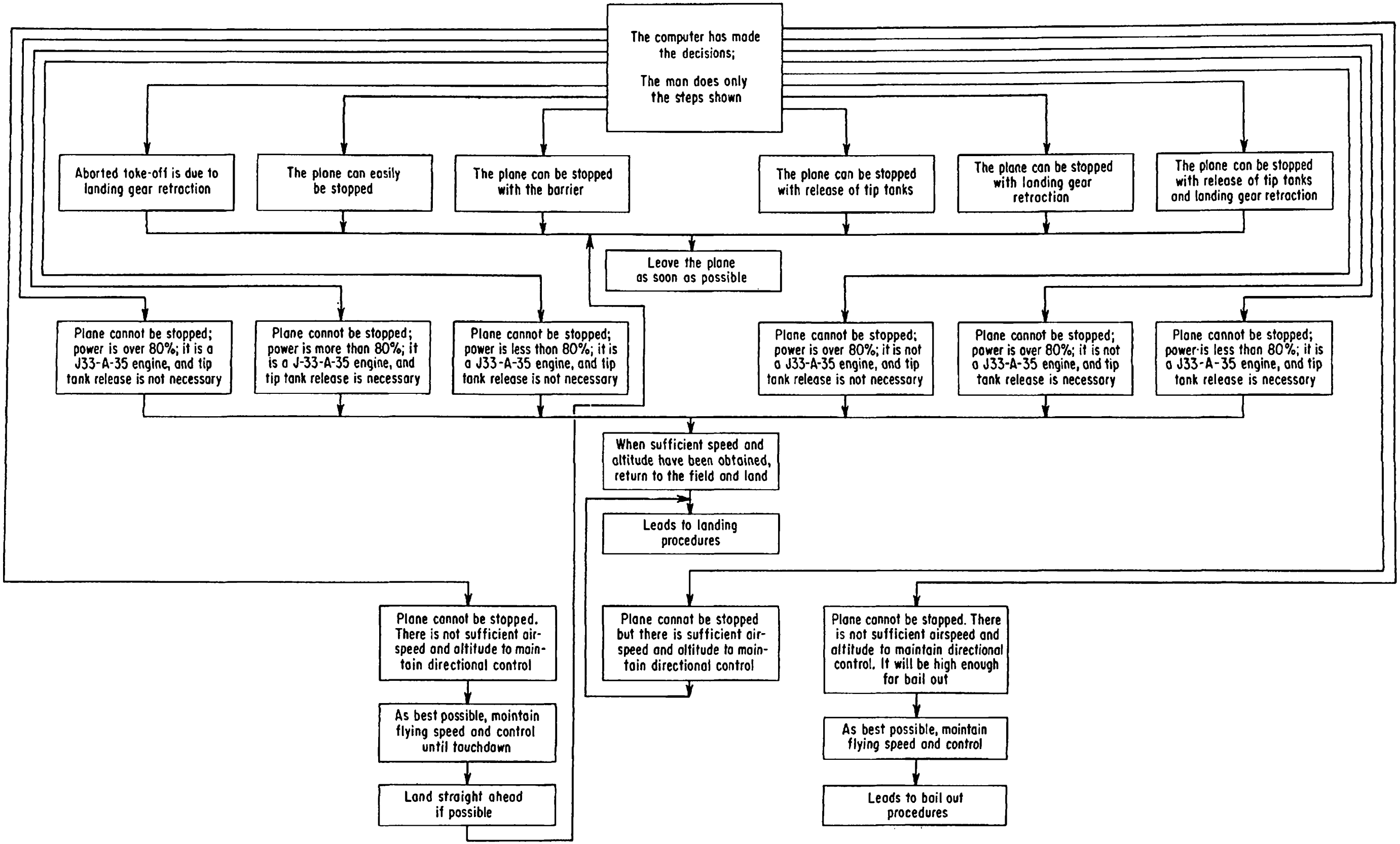


Fig. A11.3. Computer decisions pilot action required.

2. Presentation of Several Systems

Some of the actuators that will be needed might not presently be available, or desirable. This leads to the possibility of several systems, each having various advantages and disadvantages, the choice of which will depend on judgment and imposed constraints. The several systems are:

(1) *Only aircraft control left to the pilot*

advantages — The crisis is completely a mathematical problem. It will render the best mathematical solution for survival, limiting human error to a minimum.

disadvantages — Pilots may object to the sacrifice of personal decision making.

(2) *Aircraft control and other specific controls left to the pilot*

advantages — This system will put more actions into the pilot's hands, which should tend to make the pilot trust the computer's decisions. Such a system could alleviate some of the troublesome areas, such as calling for information relative to the barrier.

disadvantages — The system is now subject to human error, owing to "forgetting," ignoring, confusion, etc.

(3) *Aircraft control with complete override left to the pilot*

advantages — This system is advantageous only if the computer is out of order.

disadvantages — The computer is actually of little use, as the pilot would probably tend to distrust the computer's decision and not use it at all, relying on his "estimates."

(4) *Aircraft control with partial override left to the pilot*

advantages — This system can be advantageous if only items that are not absolutely essential are left for override; the computer can recompute if it is overridden.

disadvantages — The system is again subject to human error and suffers a time loss, since it must make the necessary recomputations when it is overridden.

(5) *Aircraft control, other specific controls, and complete override left to the pilot*

advantages — This system will be advantageous only when the computer is out of order.

disadvantages — The disadvantage is the same as number three above.

(6) *Aircraft control, other specific controls, and partial override left to the pilot*

advantages — This system is advantageous if certain actuators are not presently available, and override capability is not in critical areas.

disadvantages — Again, the system is subject to human error, both in specific controls and override.

In this discussion only the first of the six possible systems will be presented in full. The other five would consist of permuted sections of the first and/or override capabilities.

3. Presentation of Information to the Pilot

Regardless of which system is accepted, the pilot must be informed of the computer's decision so that he will not act in direct contradiction to it, especially if override is available.

If only control is left to the pilot, the pilot can be adequately informed by a four-lighted signal: one color meaning the plane will be stopped; another that the plane cannot be stopped and crash landing is imminent; another indicating that bailout is the best chance for survival; and the other color meaning that the plane is able to fly.

A more elaborate system that gives the pilot more information about what to expect can consist of the items that follow; these are represented in Fig. A11.4.

(1) A gross warning indication is presented so that a quick glance at the panel indicates *if anything*, from landing gear that is not locked to a loss in power, is not as it should be. This is done by merely wiring all the indicators together so that if any one of the present indicators lights up, the *one main indicator* will light up also. A

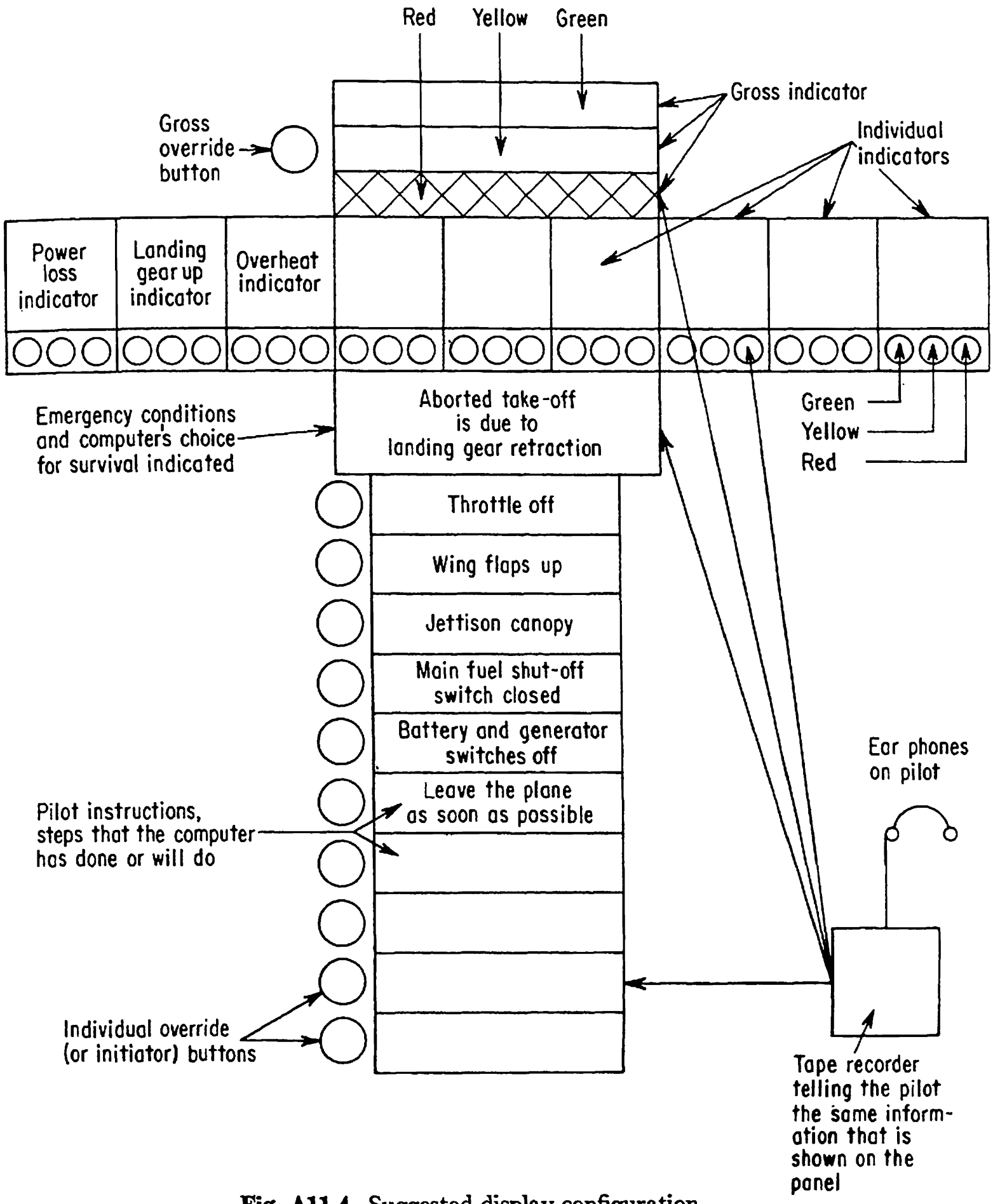


Fig. A11.4. Suggested display configuration.

green light (or no light at all) might be used to indicate all is well, and a red light to indicate something is wrong. The lighting system will, of course, depend upon human engineering considerations. The arrangement can be modified so that a yellow light indicates that something is wrong, and a red light indicates that the computer has started to take action to correct the situation.

(2) Directly under the gross warning indicator, a row of similar lights (both color and number depending upon human engineering con-

siderations) is used to indicate the exact trouble areas. This means that if the landing gear is not down when it should be, the pilot sees the warning on the gross panel and, as a result, immediately scans the lower panel to see what is wrong so as to initiate the desired action. Essentially, all warning indicators are combined into one panel for easy viewing. These first two steps involve only wiring and rearrangement of the panel indicators. It is recognized that the incorporation of these features will influence cockpit design.

(3) Directly under both the gross and individual warning lights, *instructions* could appear to the pilot as to what he is to do, the steps that the computer has done or will do, what is happening, and what the computer's decision for survival is. Which, if any, of these instructions will appear will depend on both human engineering and the computer system selected. A way of presenting these instructions or steps will have to be devised, but evolving such a sequence does not seem too difficult a task—possibly, a cycling tape could be used for each instruction panel. Instructions should appear in the sequence desired.

(4) In conjunction with the item above, or as a substitute for it (dependent upon human engineering considerations), a tape recording device may be attached to the ear phone of the pilot which, again, tells the pilot what is happening, what he should do, and the steps that the computer has done or will do. In this manner both visual and auditory information and directions will be provided the pilot.

(5) Depending also on the computer system desired, override capabilities may be incorporated. A *gross override button* and *individual step override buttons*, or some such button system, can be considered as initiators to produce the actions desired (again depending upon system design).

Of the items listed above, all are needed in some form, but probably the most important single item, and the most conveniently realizable, is tape recordings. This one item alone could be used to tell the pilot what is happening, what will be or is being done, and what he should do, thus eliminating the other items considered above (except override).

Irrespective of the form of the warning system decided upon, some indication of the events that are to happen should be directed to the pilot.

4. Data Requirements for Automatic Flight Safety

The computer's mathematical decision depends on the data that it receives, and the majority of these data will be in the form of lift or drag information. The data that will be required, possible reasons for their requirements, and difficulties of obtaining the data are listed below. Figure A11.5 shows the data required for the various decisions shown on Fig. A11.2.

(1) The *amount of power loss*, or, more specifically, loss of thrust at which the computer will begin its operations must be indicated. This information is given to the pilot from the tachometer, which is a self-generating instrument that indicates engine speed in percentage of the maximum allowable rpm and can be obtained for computer use at this point.

(2) *Altitude* information is given to the pilot through the altimeter, but cannot be taken directly from this instrument. There is a possibility of getting this information from the autopilot system. If not, added instrumentation is required. Altitude information will be required in several phases.

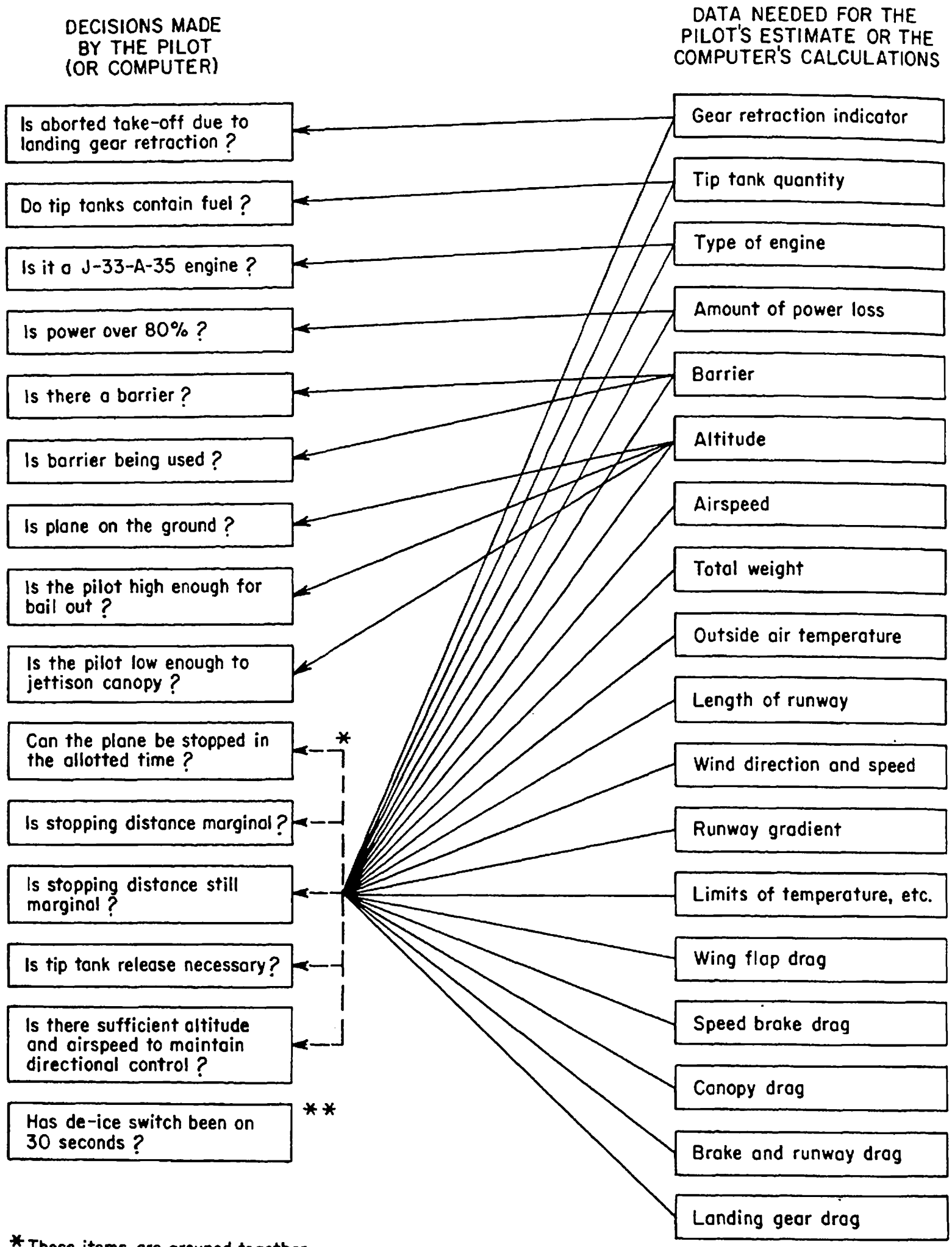
(3) *Airspeed* information is required to determine lifting forces. The same constraints and possible method of obtaining the information as indicated in the paragraph above applies here also.

(4) The *total weight* of the plane can be taken from the addition of the gross weight and the fuel quantity indicators. It will be required for drag and lifting forces.

(5) The *type of engine* will probably have to be built in (as a constant) for each computer built. This information is needed because the J33-A-20 engine has no emergency fuel system.

(6) *Barrier* information is most essential in stopping requirements, but the method of supplying it is a problem. It will be discussed in greater detail in a section to follow.

(7) The *outside air temperature* must be supplied so that the air density can be computed to determine lifting forces. This information is present and is obtainable as an analog voltage signal. However, in order that the accuracy of the indicating instrument be not greatly affected, a high impedance pick-off will be required.



* These items are grouped together because most are essentially duplications of the same question or depend on all of the aerodynamic data presented. When thought of as questions that the computer must decide, there will not be as many questions asked.

** The data for this question are neglected as it will be controlled by a timer, and the data are not required for further calculations. For clarity the time element was eliminated.

Fig. A11.5. Data required for decisions and calculations.

(8) *Tip tank quantity* (gross and fuel weight) information is needed to compute both total weight and braking power when released or not released.

(9) The *length of the runway* will present a problem. It must be known (so that length of remaining runway can be calculated or measured), but there is no indicator for it. Therefore, it must be supplied at each take-off, probably by the pilot during his before-flight check-list.

(10) An indicator for *gear failure* (other than the ordinary one now present) must be devised to indicate that the plane cannot leave the ground. This should not be a difficult task.

(11) *Wind-direction and speed* must be supplied to determine lifting forces. This information is probably already available and should present no problem. If it is not present, the pilot may have to supply it (from base operations) himself during his before-flight check.

(12) *Runway gradient* will also be needed and the section above applies here also.

(13) *Limits of temperature ranges*, etc., will be needed.

(14) *Wing flaps in the DOWN position* will supply braking power. This data will be needed as drag forces and may not be presently available directly.

(15) *Speed brakes in the DOWN position* will supply additional braking power. This information, too, may not be available.

(16) *A full open canopy* will supply braking power that must be calculated.

(17) *Brake application*, along with *type of runway*, will be needed.

(18) *Landing gear retraction* forces will be required.

Items (8) and (14–18) are essential knowledge in determining whether or not the plane can be safely stopped in the required space. This knowledge, however, may not be available, and if so, accurate tests must be made to determine it.

Items (6, 9, 11 and 12) will probably have to be supplied by the pilot during before-flight check-out.

5. Explanation of the Various Instructions and Elements Involved

Because of the duplication of many of the steps in the fifteen procedures, there are only twenty-two separate instructions. Assuming that

the pilot will be left with only directional control of the aircraft and that proper actuators can be made, *only four of the twenty-two steps are left for the pilot.* If override or specific controls are left to the pilot, this number will increase.

Descriptions of the twenty-two different instructions that the pilot carries out—the units affected and how these units operate—will now be presented to indicate the nature and scope of work required to bring an automatic system into being. Additional features of the proposed system will be discussed in some detail.

The first nine of the twenty-two instructions are concerned with fuel control or related subjects that seem amenable to computer control. The nine instructions are:

(1) *Throttle OFF*
(2) *Maintain rpm and tailpipe temperature within limits*
(3) *Try to maintain at least 130 knots*
(4) *Retard throttle to approximately 80 per cent power position*
(5) *Advance throttle as tailpipe temperature permits until maximum permissible power is reached (J33-A-35 engines only)*
(6) *Main fuel shut-off switch OFF*
(7) *Battery and generator switches OFF*
(8) *Emergency fuel switch EMERG (J33-A-35 engines only)*
(9) *De-ice switch ON (30 sec)*

Items (1–5) of these nine instructions related to fuel control are directly associated with use of the throttle. The throttle is a lever that sets the main fuel control unit, which, in turn, regulates the fuel pressure to the engine combustion chambers. The resulting fuel flow determines the engine speed. When the throttle is in the OFF position, all fuel to the engine is shut off (the main fuel shut-off switch could also accomplish this), unless the starting fuel control is energized during automatic starting. The throttle slot in the front cockpit is staggered to provide straight fore and aft motion in the operating range between IDLE and FULL, while outboard and aft motion is required when moving from the IDLE to the OFF position. The slot in the rear cockpit runs straight fore and aft with no IDLE stop. The throttle cannot be closed from the rear cockpit, because the front throttle must be pushed outboard to pass the IDLE detent. This throttle control arrangement, in some emergency cases, is certainly not a desirable one.

The main fuel control unit is possibly a "shorter, easier, or more direct route" to fuel, rpm, or thrust control. The main fuel control unit is a wide-range governor, which is compensated to limit maximum engine speed, to maintain approximately constant rpm regardless of airspeed or altitude, to limit acceleration temperature, and to limit engine deceleration to that required to maintain combustion. To do this, it must control the amount of fuel delivered to the combustion chambers, bypassing any fuel in excess of that required of the throttle setting, engine speed, altitude, and airspeed. Starting fuel control is included for automatic fuel metering during starting, so providing simplified engine starting and controlled acceleration to idling speed. When first energized, the starting fuel control, which directs fuel to numbers 7 and 14 burners for initial ignition and then to all other burners as fuel pressure builds up, provides a more accurately controlled fuel flow during the start, thus reducing the possibility of a "hot start."

The main fuel control unit seems to be the best point of attack for control of instructions (2–5). Step 1, as mentioned above, can probably be controlled by the main fuel shut-off switch—an arrangement which might lead to other problems. Steps 2–5 will require computer functions to control the valve at the desired opening so that the properly imposed limits of temperature and rpm are not exceeded. These relationships appear to be provided for in the main fue control unit, so it may be possible to use it directly as the control. Certainly, it can be used as the source of data to determine if the aircraft can fly to begin with. A thorough understanding of throttle control will have to be undertaken to realize a complete solution to this problem.

Steps 6, 7, and 8 are also fuel controls, but they pose no great problem as they are all switch-operated.

(6) The main fuel shut-off switch is a d-c powered switch located in both cockpits. It has only two positions: ON and OFF. It should be shut off only when the computer says that the plane can either be stopped or it cannot fly. In the latter case, it should be shut off before contact with the ground, even though it appears toward the end of the pilot's procedures. Therefore, some sort of tie-in with the altimeter will be desired, but this should pose no problem. As mentioned above, it could also be used, in essence, to put the throttle into the OFF position (no fuel). The consequences of such an operation will have to be investigated, but at first glance, no damage to the throttle is anticipated.

(7) The battery and generator switches have only two positions: ON and OFF. The front cockpit has a separate switch for both the battery and the generator, while the aft cockpit has one switch that controls both. The two switches in the front cockpit are in series with the one switch in the aft cockpit; either circuit may be de-energized by the switches in the front cockpit, but the switches in *both* cockpits must be ON to energize either circuit.

The battery and generator switches are presented in this section because they are connected indirectly with the fuel system. They should be shut off at all times before contact with the ground. This, however, may lead to problems if taken as a function of altitude as the plane may be on the ground. Once the battery is shut off most of the remaining steps cannot be carried out. Possibly a change in altitude rather than the altitude above the terrain could be used.

(8) Emergency fuel switch is d-c powered and is a three-position toggle switch labeled TAKE OFF AND LAND, EMERG., and OFF. It controls the spring loaded open solenoid-operated bypass valve in the emergency system, the spring loaded closed solenoid-operated bypass valve in the main system, and a pressure switch which is energized in the TAKE OFF AND LAND position. *The TAKE OFF AND LAND position is ordinarily used at all times below 5,000 feet above terrain, because it gives automatic protection against flameout caused by main system failure.* This feature can be added to the system easily.

In the EMERG position the main system bypass valve is opened, the emergency system bypass valve is closed, and the engine operates on the emergency fuel system. (The starting fuel switch must be OFF before the main fuel system bypass valve can be energized by the emergency fuel switch.)

When the emergency fuel switch is in the OFF position, the switch circuit is de-energized. Now, the main system bypass valve is closed, and the emergency system bypass valve is open. The *OFF position is selected over 5,000 feet or more above the terrain because automatic switch-over is not desirable at altitude.* If the main system fails and switch-over occurs, it is possible that a sudden burst of unmetered fuel in the combustion chamber could cause overheating and turbine overspeeding. In the event of a complete electrical failure, the engine can be operated only on the main

fuel control, and there is no emergency fuel system in the J33-A-20 engine.

The emergency fuel control unit is mechanically linked to the throttle and main fuel control. This unit provides manual alternate control of engine power for use in the event of failure of the main side of the engine-driven pump or of the main fuel control system. It consists of a throttle valve, an altitude-compensated relief valve, and a solenoid-operated by-pass valve (spring loaded open). The relief valve is adjusted to provide 100 per cent engine rpm on the ground on a 100°F day. Available full throttle rpm will vary with free air temperature and altitude. When the free air temperature is above 100°F, care must be taken to prevent engine rpm from exceeding 100 per cent. The altitude compensation attempts to maintain constant engine rpm for a given throttle setting regardless of airspeed and altitude. Engine overspeeding is possible when operating on the emergency fuel control system.

(9) The fuel filter de-icer system is provided to remove ice accumulation from the low pressure fuel filter. It consists of a five-gallon (5 minutes) alcohol supply carried in the right water-alcohol injection tank, an electrically driven pump, a differential pressure switch, and an amber warning light. The tank is connected through the fluid pump and a solenoid shut-off valve to the low pressure fuel filter.

The fuel filter de-icer switch is in both cockpits. It is a toggle switch with two positions: ON and OFF. It is spring loaded to OFF. The ON position turns on the pump and opens the solenoid valve so that alcohol is pumped into the filter. If the filter is iced, the alcohol will dissolve the ice accumulation, which reduces the pressure drop.

The fuel filter de-icer system is somewhat useless, however. The switch should be left ON for 30 seconds (this, to be done automatically, will entail the use of a timer), which means that approximately 30 seconds will be lost before proper fuel rate is restored; 30 seconds can be disastrous. Also, the only direct indicator for fuel icing is at the low pressure fuel filter, but it is possible for icing to occur further down the fuel system and not be directly indicated. Indirect indications of icing are:

(a) Fluctuating rpm,
(b) Unable to obtain a higher rpm than has been used,
(c) Apparent loss of throttle control.

Several conclusions can be made. The ultimate solution will be to discover a method of extracting the water from the fuel, thus completely eliminating the problem. Such a fuel is reported to exist, but as of now is not being used. However, since the fuel is not in use, a system could be developed to take care of the problem. When any one or a combination of the indirect indications of icing occur, or the warning system indicates icing, automatic alcohol injection could be accomplished. Both indicators and actuators would be involved. In addition, as a further precaution against flameout caused by fuel starvation from icing, automatic alcohol injection could be supplied in discrete quantities when the outside air temperature or fuel temperature is below freezing (altitude could also be used to indicate proper injection times). It has been recommended by some Air Force personnel that alcohol injection be done by pilots with a 15-second de-ice switch time for each 30 minutes of flight. This system could be established and is to be desired as many flameouts are the result of icing; it is objectionable in that alcohol injection brings about a loss in power. Inasmuch as water-free fuel is not used, some such system in any event can decrease chances of a flameout or partial power loss caused by fuel starvation.

The next seven instructions to be discussed pertain to the braking power required to stop the aircraft, or to the reduction of its weight. These steps are listed below:

(10) *Canopy full open*
(11) *Tip tank jettison*
(12) *Landing gear UP (or DOWN)*
(13) *Wing flaps UP*
(14) *Apply brakes*
(15) *Call for barrier*
(16) *Speed brakes DOWN*

(10) The canopy has several controls, but since we are considering emergency procedures, only the quickest method to produce the desired results will be discussed.

The canopy can be opened by either of the two d-c powered switches, one in each cockpit, which are wired in parallel. They are spring loaded to the OFF position and are wired directly to the battery. They have only two positions: OPEN and OFF.

The canopy itself is a single unit covering both cockpits. It is hinged at the aft end and is raised or lowered by an electrically or mechanically

operated chain and sprocket actuator at the base of the remover, which is attached to the canopy cross member. When closed the canopy seal is inflated by engine air pressure in order to seal the canopy frame against the cockpit sill and windshield. The canopy must be manually locked when closed by an internal locking handle. A system to unlock and open the canopy should pose little difficulty. It also seems desirable that the system (or some other device) include an automatic locking mechanism so that the pilot could not forget this item. Canopy jettison will be discussed below.

(11) Tip tank jettision is accomplished by the bomb salvo button. This releases the entire external load and is present in both cockpits. It is wired directly to the battery. Airspeed might enter into this operation but would not be essential. The recommended speed for jettison is 260 knots, but it has been accomplished at 130 to 435 knots. The pilot, considering the safety for persons below, might desire this operation to remain in his hands.

(12) The landing gear is controlled by a lever in either cockpit. Several systems are available to the pilot in case the main one does not work properly. Landing gear retraction time at 100 per cent rpm is approximately three seconds; at 50-60 rpm it is approximately five seconds.

The landing gear lever has two positions: UP and DOWN; and it is held in either position by a spring loaded lock which must be released by pressing the button on the end of the lever before it can be moved and the gear repositioned. Actuators will be needed for this operation, as it is done manually at the present time. This should not be a large problem area, especially if the before-mentioned "gear down" problem is worked out.

(13) The wing flaps are controlled by interconnected levers, or on earlier aircraft by independent switches connected in parallel in each cockpit. They have three positions—UP, OFF and DOWN—but apparently operate in such a way that control among the three positions is possible. The wing flaps are actuated by two electric motors mechanically interconnected by a flexible shaft so that if one motor fails, the remaining motor will actuate both flaps, though retraction and extension time will be slightly increased. There is a wing flap position indicator but no alternate system. This should pose no great problems, but the introduction of incremental control might.

(14) The brake system is independent of the before-mentioned main hydraulic system and is operated manually. There is no emergency system. Actuators will have to be devised to apply braking power (if it is to be incorporated into the computer). The nose should be down when brakes are applied. This condition might present various problems.

(15) Five out of the fifteen sub-procedures depend upon the data concerning the barrier. The actual barrier data can be supplied by base operations to the pilot during before-flight checkout; the pilot, in turn, can supply these data to the computer. Since this information is essential, it seems desirable, as a check against forgetting by the pilot, to install some mechanism that will not let the engine start unless the information is put into the computer. This could also be done for other information requirements that the pilot might have to supply. However, the act of calling for the barrier's use will present a problem. Some actuator, possibly using a specific frequency (which would entail universal use at all runways) must be built to set the barrier up automatically when it will be needed. If no barrier data is present, the computer must function on the assumption that there is no barrier. Such an assumption could lead to needless harm to the aircraft (as when landing gear will be retracted). The problem of supplying the data and the actuators to set them up appears to indeed be a large one, and the act of calling for barrier use might have to remain in the pilot's hands. If this is the case, a way of notifying him that he should call for it could be accomplished (i.e., by the tape recorder discussed previously).

(16) The speed brake switch is actually a lever control in later type aircraft, having three positions—UP, OFF and DOWN—appearing in both cockpits on the throttle and having incremental control (probably by not holding the switch in one position until fully up or down). Switches, wired so that the front cockpit has control only when the aft switch is UP, are installed in trainer aircraft types; this feature is not necessarily to be desired for other aircraft. When the aft switch is in the UP position, d-c power is directed to the front switch, and the speed brakes travel in the direction selected by the switch. When the aft switch is in the OFF position, power to the valve actuator solenoid is cut off, the spring loaded hydraulic valve returns to the neutral position, and speed

brake travel is stopped. When the aft switch is in the DOWN position, power is directed to the valve actuator (down) solenoid, and the brakes go out regardless of the position of the front switch. Speed brake switches may be carried in the UP position when not in use.

Earlier aircraft have only two selections—UP and DOWN—and are wired in much the same manner. There is no incremental control. There is no speed brake position indicator in either type of aircraft, and they operate only when there is sufficient hydraulic pressure.

It should be noted that the use of speed brakes in any of these procedures is not used except in the UP position. In the procedure where the barrier is used, they are not desired, because they might interfere with the barrier. However, they might be used as braking power in the other procedures where the plane can be stopped, although this remains to be seen. The braking power might not be enough to be advantageous at low speeds, especially since their extension will give a nose-up tendency which might lead to an undesired stall or to less control by the pilot. In some cases, however, they might be effective enough to be desired.

Two instructions are given only for crash landings or bailouts (canopy jettison does not appear in the bailout procedures, but it would if the bailout procedure were continued). These are:

(17) *Lock shoulder harness*
(18) *Jettison canopy*

(17) Locking the shoulder harness automatically may be possible if proper actuators are made. This is actually a desired operation, for the pilot consumes valuable time in performing this step. It is necessary for him to look away from the panel, which means additional time lost owing to eye readjustment back to panel. Presently the shoulder harness locking function is not controlled electrically.

(18) The canopy is jettisoned in only a few of the procedures. There are definite reasons for this. It is jettisoned when one of two situations is apparent. If a fuel leak is anticipated (as in aborted takeoff caused by landing gear retraction, or a crash landing) it is jettisoned; if bailout is anticipated, it is jettisoned. But here, too, it is jettisoned as a function of altitude because of the possibility of air damage to the eyes of the pilot. If bailout is apparent, it should not be jettisoned until bailout altitude is obtained; if a

crash is forthcoming, it should not be jettisoned until the aircraft is only one foot above the terrain. In any case, it should be accomplished before contact with the ground so that escape from the aircraft is not endangered. To treat canopy jettison as a function of altitude may engender problems, as it is not desired in all cases (as on the ground). In aborted take-off caused by landing gear failure, the canopy should be jettisoned as soon as possible.

Presently, to jettison the canopy in later versions of the T-33, either the right hand seat arming lever or the T-handle on the floor adjacent to the arming lever must be raised. These methods, however, also arm the ejection seat in the same operation; there is no switch system to jettison the canopy.

It seems no large problem to redesign the whole system and allow an electrical circuit rather than a mechanical system for canopy jettison, taking both time and accident prevention into account.

The whole area of bailout seems to be a large problem area. Elaborate ejection seats and easily released belts are all features that are making great strides for safety, and these features will probably be added soon to existing planes so that the whole procedure of bailout—from locking or unlocking belts to opening the chutes—can be accomplished automatically. However, these future developments do not apply to the set of procedures under immediate discussion. The treatment of bailout procedures must await discussion at a later date.

Of the 22 instructions the pilot may have to follow, 18 have been presented. Most of the remaining four functions cannot be handled by computers unless an elaborate radar system is established or the plane is operated by remote control. Three of these instructions lie in the field of pilot control, which should not be taken away from the pilot. The fourth is the act of the pilot leaving the plane.

(19) *When sufficient speed and altitude have been obtained, return to the field and land*
(20) *As best possible, maintain flying speed and control*
(21) *Land straight ahead if possible*
(22) *Leave the plane as soon as possible*

Under analysis, Instruction 19 might be partly a computer function, as some indication could be made when enough speed and altitude are obtained for safe flying. However, the pilot will know the actual speed and altitude needed, and the step, computer-wise, is not needed.

Instruction 20 will have to be carried out by the pilot, even though the safe flying speed will be known by the computer.

Instruction 21 must be carried out by the pilot, as only he can select the safest place, for both himself and others, to crash-land.

The man must remove himself from the plane after it has stopped. Therefore, Instruction 22 must be carried out by the pilot unless some type of automated ejection is desired.

6. Other Possible Systems Incorporations

The added features considered in the foregoing discussion will make flying conditions safer. Some of these appear rather futuristic, but all are ideas which will eventually become working apparatus. Some (speed brakes, automatic alcohol injection, ejection seats) have already been discussed; others are presented at this time.

Since there are already heat detectors sensitive enough to detect flying aircraft (as in heat seeking missiles), a combination of radar and heat detectors could be used to warn the pilot of (or to avoid, if actuators were built) approaching aircraft. Development of such a warning would be very helpful in traffic control and instrumental, of course, in helping to find the ultimate answer to mapping displays—that of detecting heat changes on land and water.

A similar method using radar, altitude, and/or heat detectors, with proper actuators (or just a warning system), could be used for avoidance of mountains, etc., when visibility is low (or at all times).

Other "safety features of the future" will undoubtedly be incorporated with physiological data, such as fatigue, for example, that will be taken from the pilot to warn the ground crews of various dangers to him. Possibly, remote control might be established. Inasmuch as these safety features of the future are a complete field in themselves, they will not be discussed at this time.

7. Concluding Remarks

The problems concerning data, detectors, and actuators that a computer system will need to relieve the pilot during power loss on take-off have been presented. These appear quite feasible of solution in most areas. The example procedure presented is the longest and most detailed emergency

procedure for the T-33, involving by far the most decisions, but it must be noted that if the actuators and detectors can be built, only four steps involve the pilot's actions. However, the procedure described applies only to power loss during take-off. There are various other procedures to be considered, and they might occur in combination with this or other emergencies. To illustrate some of the complications that might take place, this interesting story is told:

During take-off roll an aircraft lost power. The pilot, committed to the take-off as sufficient runway in which to stop the aircraft was not available, jettisoned his external tip tanks and managed to get airborne.

His troubles, however, had just begun. His normal fuel system failed; (the failure, actually, should have been anticipated by him, as a high percentage of power losses result from fuel system failure). He switched to the emergency fuel system and requested an immediate landing on the nearest runway. At this time, his emergency system failed, and an overtemperature condition forced him to reduce power continually rather than risk a fire or explosion. At 3200 feet the engine flamed out. As the runway was approached in a flameout pattern, his landing gear wouldn't come down. By making quick use of the emergency system, however, he was able to land.

These problems are representative of ones that will arise. The computer system described in this appendix could handle the situations discussed except for the emergency landing gear extension and the flameout landing pattern, both of which could eventually be incorporated into the system.

Appendix 12

Rudimentary Study Outline for Space-oriented Informational Systems Studies

"In life everything lies in the mass—materials are a mob—a man's measure is his ability to select, reject, organize."
—ELBERT HUBBARD

1. Preliminaries

This suggested program studies astronautic systems, constructs designed to permit accomplishment of certain types or classes of astronautical missions. These, in turn, are operational processes which possess characteristics leading to the realization of certain objectives and which are performable partially or wholly in space. As used herein, the concept of astronautical system will be understood to imply utilization of a vehicle and propulsion.

It will be of convenience to classify roughly astronautical missions according to the distances from earth involved:

Magnitude 1. Pseudoterrestrial missions as exemplified by missions involving orbiters and skip-orbiters.

Magnitude 2. Quasiterrestrial missions which might be regarded as pseudoterrestrial except for probable need of extra-terrestrial referencing, as in missions involving lunar exploratory vehicles.

Magnitude 3. Missions which are definitely extraterrestrial but intrasolar.

Magnitude 4. Missions which are extrasolar but intragalactic.

Magnitude 5. Extragalactic missions.

For purposes of current realism and well-definedness, emphasis on these varying magnitudes will be in inverse order to their indices.

2. Objectives

The objectives of the suggested study program are:

(1) Elaboration of descriptions of representative astronautical missions in accord with planned and projected national policy regarding objectives and mission techniques.

(2) Establishment of the information and control requirements for systems implementing these missions subject to the assumption of suitable vehicle and propulsion.

(3) Elaboration conceptually of certain sensory equipments based on new applications of physical principles and analysis of combinations of these and known sensory equipments to resolve questions of feasibility, adequacy, minimality, and sub-optimality with respect to both information requirements and anticipated restraints concerning compactness, operability, and reliability.

(4) Elaboration conceptually of certain actuating equipments based on new applications of physical principles and analysis of combinations of these and known actuating equipments to resolve questions of feasibility, adequacy, minimality, and suboptimality with respect to both control requirements and anticipated restraints concerning compactness, operability, and reliability.

(5) Establishment of the general requirements for essentially real time processing of the data flows implied by sensory combinations to satisfy the input requirements of actuating combinations, the combinations involved, and the real time concept, which is to be a function of mission.

(6) Resolution of questions of feasibility, adequacy, minimality, and sub-optimality for state-of-the-art computers regarding processing requirements.

(7) Delineation of areas in sensory equipments, actuating equipments, and computers requiring further research and development so that feasible systems can become available for varying missions.

(8) Establishment of methodology leading to the optimal integration of man and machine in those systems implementing missions in which man-machine systems are required or deemed desirable.

(9) Elaboration conceptually of certain display equipments, either known or projected, to permit its utilization by men either aboard system vehicle or elsewhere, or for the purpose of gaining relevant system and mission information.

(10) Analysis of over-all system configurations to resolve questions of feasibility, adequacy, minimality, and optimality concerning varying missions.

3. Need for Program

Although the major subsystems have been elaborated in some detail for contemporary airborne systems, little is known of the effects of their substantially unaltered inclusions within astronautic systems. Certain claims have been made regarding the ease of extrapolating ballistic missile subsystems to astronautic subsystems, but these must necessarily be only partly "educated guesses." Not only is it evident that ballistic trajectories are not universally desirable, but it is also foolhardly to assert that these techniques will apply in systems which are to implement missions that, as yet, have not even been defined.

The present era offers a previously unequalled opportunity to begin systems engineering efforts with truly fundamental analytical efforts rather than with elaboration of systems configurations from combinations of ill-defined points of equipment availability and compromised missions. Trite as such truisms may appear, it must be realized that we must begin

at the beginning and utilize all the planning possibilities at our command to cope with the extrasystem unknowns ahead. Intrasystem unknowns will simply add the last back-breaking straw.

4. Plan of Attack

Phase 1. Representative realistic and futuristic missions will be precisely described, with emphasis decreasing with increasing magnitude. There will, however, be no placement of extreme emphasis on Magnitude 1. Especial attention will be directed to establishment of appropriate referencing and coordinate systems and to energy requirements as a function of trajectory.

Phase 2. For each selected mission, the information and control requirements will be deduced by analytical and, if necessary, introspective and simulatory methods.

Phase 3. Known and projected sensory equipments will be elaborated conceptually. These will include, but not be necessarily limited to, the following:

3.1 Navigation equipment
- (a) Inertial equipment with star-track smoothing and/or intermittent correction.
- (b) Equipment based on the Lorentz theory of time contraction in relativistic mechanics implemented by matching of atomic clocks which are aboard and at known exterior data-linked locations.
- (c) Equipment based on sensing of anticipated gravitational field resultants with or without coupled accelerometers.
- (d) Equipment based on premises of sensing Doppler effects in the infra-red and/or optical portions of the spectrum.

3.2 Exterior surveillance equipment
- (a) Infra-red detectors.
- (b) Radar systems.
- (c) Closed TV systems.

3.3 Vehicle surveillance equipment
- (a) Thermodynamic detectors (temperature, pressure, etc.)
- (b) Analysis equipment for atmospheric content and rates of change thereof.

(c) Direct thrust sensing equipment.
(d) Fuel-remaining sensors.

Special attention should be directed to the question of thermodynamic relations between the system and its environment. It may well be that a whole new range of physical variables will become significant, e.g., the various free energy functions.

3.4 Attitude sensory equipments
(a) Angular accelerometers
(b) Navigation coupled inertial equipment
(c) Star-track or parallax equipment

3.5 Data-link receiving equipment.

Phase 4. Known and projected actuating equipments will be elaborated conceptually. These will include, but not necessarily be limited to:

4.1 Actuators for modifying engine and control surface attitudes.
4.2 Equipments for initiation and cut-off of engine runs.
4.3 Equipments for regulation of power plants providing programmable and controllable thrust (as opposed to on-off thrust).
4.4 Switching for various instrument and control equipments combinations.
4.5 Data-link transmission equipment.

As previously noted following 3.3 above, elaboration of the information and control requirements may well call-out as necessary or desirable entirely new sensory and/or control equipments. The equipments requirements must also be related closely to the allocation of duties in man-machine systems.

4.6 Auxiliary power for instruments
(a) Optical energy conversion equipment
(b) Equipment to utilize kinetic energy associated with space particles (micrometeoritic, etc.)
(c) Equipment to utilize energy of non-optical cosmic radiation.

Phase 5. Combinations of sensory equipments and combinations of actuating equipments will be analyzed and, if necessary, simulated for missions (both with and without human augmentation) in order that questions of feasibility, adequacy, minimality, and suboptimality be resolved. Areas requiring remedial research and development will be delineated.

Phase 6. The data flows implied by sensory and actuating combinations selected for various missions will be elaborated, and corresponding processing requirements deduced. These processing requirements will be categorized to include, but not necessarily be limited to:

6.1 Navigation
6.2 The interrelated control processing functions of autopilot, powerplant control, cruise control, and fuel and/or energy management.
6.3 "Armament" control in the broad sense of transmission of matter and/or energy from system to environment for purposes other than obtainance of thrust or sensory information.
6.4 Surveillance monitoring both in the sense of collision avoidance and in the sense of retention of system parameters "in control," including diagnosis and initiation of remedial action for "out of control" situations.
6.5 Provision for data-link communication.
6.6 Driving of displays for utilization by humans either aboard or, with data-link, elsewhere.
6.7 Miscellaneous housekeeping and auxiliary functions (for example, safety interlocks).

Phase 7. Elaboration conceptually of known and projected computer configurations. These will include, but not necessarily be limited to, hybrid and pure combinations of the following:

7.1 Incremental-type control computers with either incremental or integral internal communication.
7.2 Serial and parallel general and special purpose digital computers.
7.3 Analog function generators.
7.4 High capacity random and serial access memories.
7.5 Special purpose arithmetic units and function generators operating in modular logics.
7.6 Generalized logical units (which accomplish certain types of monitoring and decision-making over function space).

Phase 8. Computer configurations will be analyzed and, if necessary, simulated from the viewpoints of missions, implementations in existing and projected components and organizational techniques, input-output compatibility with sensory and actuator combinations, capability, compactness, operability, reliability (with and without high and/or low-level redundancy), and potential of providing degraded performance rather than total failure to resolve questions concerning feasibility, adequacy,

minimality, and sub-optimality. Areas requiring remedial research and development will be delineated. The components and structural innovations to be considered will include, but not necessarily be limited to, the following:

8.1 Semi-conductors of both germanium and silicon fabrication (including tunnelling devices) and of all types of junction combinations; also, suboxides.
8.2 Magnetic devices (cores, twistors, transfluxors, frequency wheels, etc.) and thin films.
8.3 Ferro-electric and ferro-resonant devices.
8.4 Magnetic recording devices (drums, core matrices, tapes, etc.).
8.5 Cryogenic devices (cryotrons, persistent current devices, etc.).
8.6 Chemical devices (solion, metachromic devices, etc.).
8.7 Microwave devices (traveling wave tubes, shaped waveguide, tuned cavities, etc.).
8.8 Optical devices (phosphor memory devices, glo-gates, etc.).
8.9 Waveform processing (internal communication achieving relaxed requirements for amplitude precision), also called frequency transform).
8.10 Modular approximation of function.
8.11 Utilization of generalized logical structures capable of utilizing both digital and analog forms of information.
8.12 Utilization of bundle-type logic to achieve statistically enhanced responses as opposed to redundancy per se.
8.13 Programming flexibility achieved through micro-programming techniques.

It should be noted that quite distinct programs and actual computer capability may be necessitated by varying portions (boost-out, free-fall, trajectory management, re-entry, etc.) and modes within a mission.

Phase 9. Determination insofar as possible will be made of optimal allocation of duties in man-machine systems implementing missions in which such systems are required or deemed desirable. Primary initial emphasis in this area shall be on mission accomplishment rather than human ease of operation.

Phase 10. Optimal forms of display for human utilization will be studied in view of duties allocations. Here again, for example, optimality refers to enhancement of mission accomplishment rather than allowance of utilization of degraded quality operators.

Phase 11. Throughout the program, careful attention must be focused on mission accomplishment (that is, system aspects). Methodology elaborated, wherever possible, should be such as can be carried further to analyze new and extended missions rather than be a one-shot technique. Definite efforts should be made to prefer techniques incorporating classes of problems rather than individual problems. Approaches should be continually screened to provide new methodology for dealing with systems problems in general; that is, not only should the objectives previously delineated be achieved, but conceptual and communications groundwork should be laid for future systems efforts.

Phase 12. Systems configurations will be analyzed *in toto* (insofar as possible) for mission requirements to resolve questions of feasibility, adequacy, minimality, and optimality.

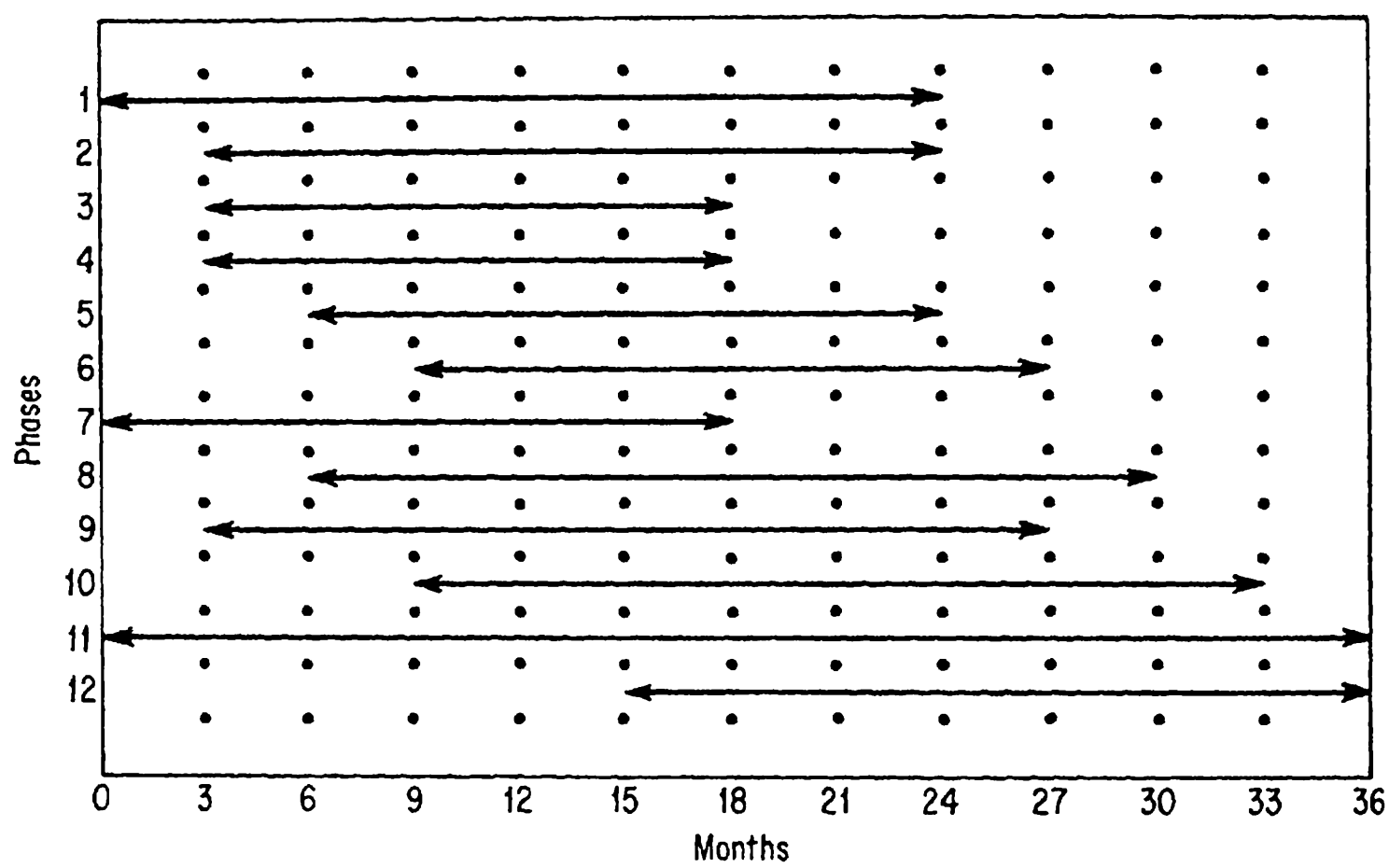


Fig. A12.1. Estimated scheduling, phases 1–12.

Phase 13. Research and development studies will be made to remedy deficiencies in sensory and actuator equipment areas.

Phase 14. Research and development studies will be made to remedy deficiencies in computer component, organization, and structure areas.

Phase 15. There will be development, construction, and testing of prototype computers for varying astronautical systems.

Phase 16. There will be development, construction, and testing of prototype astronautical systems.

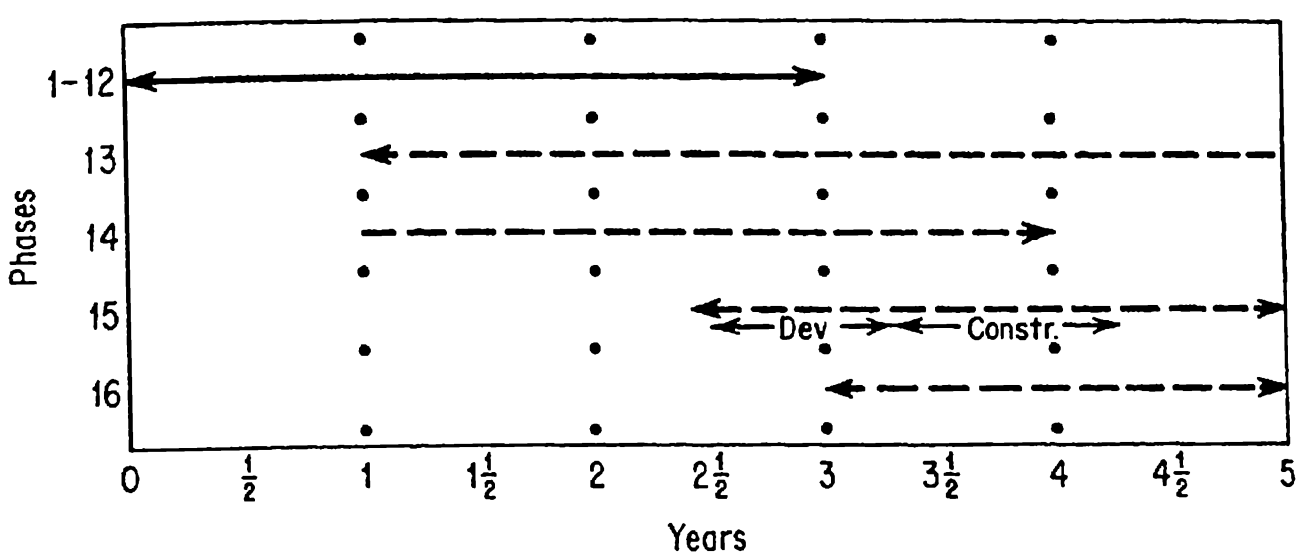


Fig. A12.2. Estimated scheduling, phases 1–16.

Appendix 13

Relativistic Doppler Effect, A New Approach to Space Navigation

"The best way is always the simplest way after it is learned."

FRANK B. GILBRETH

1. Abstract

As a result of recent efforts in evaluation of the unconventional application of known principles and the generation of new concepts for the relatively primitive art of astronavigation, some ideas which may be significant have been developed.

Results to date have indicated that a rather minimal (spaceborne) midcourse and terminal precision navigation and guidance system may be achieved by the utilization of relativistic Doppler effects in conjunction with the application of what is perhaps a new concept in terminal guidance. This concept utilizes either an explicit or an implicit target bearing and

range deduction by an accelerometer-gravity field vector discrimination scheme.

Apparently, it is possible and feasible to effect precision terminal guidance *without* direct target observations of any kind, and *without* self-contained accelerometers, if the masses of the bodies having appreciable influence are known fairly precisely beforehand, and if the gravitational anomalies of these bodies are crudely known. This scheme is made possible by the use of Doppler effects to obtain acceleration information, for which the theory and some techniques are subsequently elaborated.

Furthermore, it appears that guidance and the deduction of planet location are possible without operating directly upon velocity information but merely by utilizing acceleration information.

A significant conclusion that arose from this study regarding investigations of astronavigation instrumentation concerns precision spectral frequency analyzers, which have constituted one of the serious limitations of contemporary analog Doppler systems. It appears that it is feasible to effect an improvement in this area by utilizing simple digital techniques.

Investigations, which show promise, of visible star spectral shift by electronic detection are being explored by application of Kerr cells.

A preliminary evaluation of the concept of velocity determination by the matching of atomic clocks indicates that precision velocities with accuracies of the order of centimeters per second can be achieved with degraded present state-of-the-art performance of masers. In interplanetary operations, the availability of precision range to the earth reduces by at least one or two the set of optical celestial observations required.

Possibly highly significant is the conclusion, deduced from a relativistic analysis, that there exists a vehicle path in which the Doppler shift relating to a single source is zero. More generally, there exist families of paths in which the Doppler shift relating to a single source is a constant. These curves are logarithmic spirals. Since the path of a continuous thrust vehicle is also a logarithmic spiral, the application of this phenomenon for a steering criterion is suggested.

Furthermore, it can be shown that for a path of conic section (elliptical, parabolic, hyperbolic, etc.), steering criterion can be effected by continually changing the expected frequency shift according to the conic desired. Theoretically, one consequence of this idea is that not only velocity and range but also bearing could be deduced from a single source without the establishment of a base line.

2. Introduction

Studies in the area of astronavigation problems have indicated that astronavigational techniques, analysis, and development are lagging. They are not commensurate with the state-of-the-art in trajectory analysis and vehicle designs. Only a small effort has been given to the delineation of implementation, the design of instruments, and to the examination of the real necessity of certain instrument-derived information. The existence of vastly better approaches than a mere extrapolation of terrestrial navigation techniques to space is suspected. There is a need for study effort (in the areas of contemporary perspectives, utility, and efficiency) of what presently is felt to be required information. Furthermore, new concepts and novel implementation techniques need to be generated. Spaceborne equipments are complex and of dubious reliability. They carry heavy physical and energy penalities, and make other demands upon vehicle designs concerning configuration (location, protrusions, vehicle attitudes, etc.).

Motivated by these deficiencies in current design, preliminary studies have been made in several areas. The first of these was in the utilization of relativistic phenomena and associated Doppler effects, together with an examination of their implications. The aim always tended toward optimizing the utilization of information derivatives and the investigation of novel instrumentation of radar and visible spectrum Doppler analysis.

The potential application of the Doppler effect to space navigation is vastly superior to conventional radar techniques for obtaining range and bearing. The following summary indicates the nature of this superiority:

(a) Considering degraded laboratory performance of MASERS with frequency accuracies of 10^{-11} and crystal-controlled oscillators with accuracies of 10^{-8}, and "reflected" (re-transmitted or transponded) vehicle-generated frequency signals, it is theoretically possible to achieve vehicle velocity determination in the order of 3 centimeters per second. Conceivable improvements of this accuracy are possible.

With a base line of 6000 miles, and a spaceborne oscillator of 10^{-8} accuracy, at a lunar unit distance, a bearing accuracy of one microradian can be obtained. Using effectively "coherence" between oscillators, this accuracy might be improved by two orders of magnitude.

(b) No line-of-sight tracking is required, except for reasons of energy propagation efficiency, and even this requirement is slight.

Accurate line-of-sight directing may be accomplished by a conventional data link, if needed.

(c) Using a novel digital differential analyzer-type Doppler spectral frequency tracker, precise acceleration information can be derived directly from the mixer.

(d) In interplanetary operations, the availability of precision radial velocity and range from the geocenter reduces the required sets of optical observations by at least one or two, thereby reducing the number of instruments required.

3. Theory and Techniques

A. Gravitational Field Vector and Range Deduction by Acceleration-Gravity Discrimination

In an effort to formulate gravity effects on accelerometers during powered phases in the target region, it was revealed that the gravity influence could be used as a guidance *aid* rather than a contaminant (as is the situation in terrestrial navigation). Furthermore, in the limiting case, platform-stabilized accelerometers or inertial platforms are theoretically not required for generation of prime navigation data.

They are needed for vehicle attitude control and other systems considerations, however. It apparently is possible and feasible to effect *precision* terminal guidance without target referencing sensors of any kind, if the masses of the earth and target (for lunar flight)—or earth, sun, and target (for interplanetary flights)—are known with a reasonable degree of accuracy. This technique becomes possible through the use of precise externally-derived acceleration and bearing information, such as is possible with an earth-based Doppler radio (or stellar radiation spectrum Doppler shift).

The basic nótion may be explained in the following simple manner; first, consider the general case of a vehicle having on board three accelerometers with an inertially fixed orientation. In a gravitational region, these accelerometers will indicate the vector sum of actual vehicle acceleration in inertial space plus a component resulting from the total gravitational fields of all

of the celestial bodies influencing the accelerometers.

$$\vec{a}_{\text{ind.}} = \vec{a}_{\text{true}} \oplus \vec{g}_{\text{bodies}}. \tag{96}$$

Assume a set of three mutually orthogonal accelerometers. Then,

$$a_{x_i} + a_{y_i} + a_{z_i} = a_{x_t} + a_{y_t} + a_{z_t} + g_{x_b} + g_{y_b} + g_{z_b}. \tag{97}$$

Now, if the true components of linear acceleration a_{x_t}, a_{y_t}, a_{z_t} can be obtained from another source, then the direction and magnitude of the gravitational field $\bar{g}_b$ can be *deduced*. From Eqs. (96) and (97),

$$\begin{aligned} g_{x_b} &= a_{x_i} - a_{x_t} \\ g_{y_b} &= a_{y_i} - a_{y_t} \\ g_{x_b} &= a_{z_i} - a_{z_t}. \end{aligned} \tag{98}$$

There remains the problem of discriminating between the gravity fields of the various influencing bodies. In the general case, if the mass bearing and range to all bodies save one are known or are determined in flight, the range and bearing of the unknown body can be deduced.

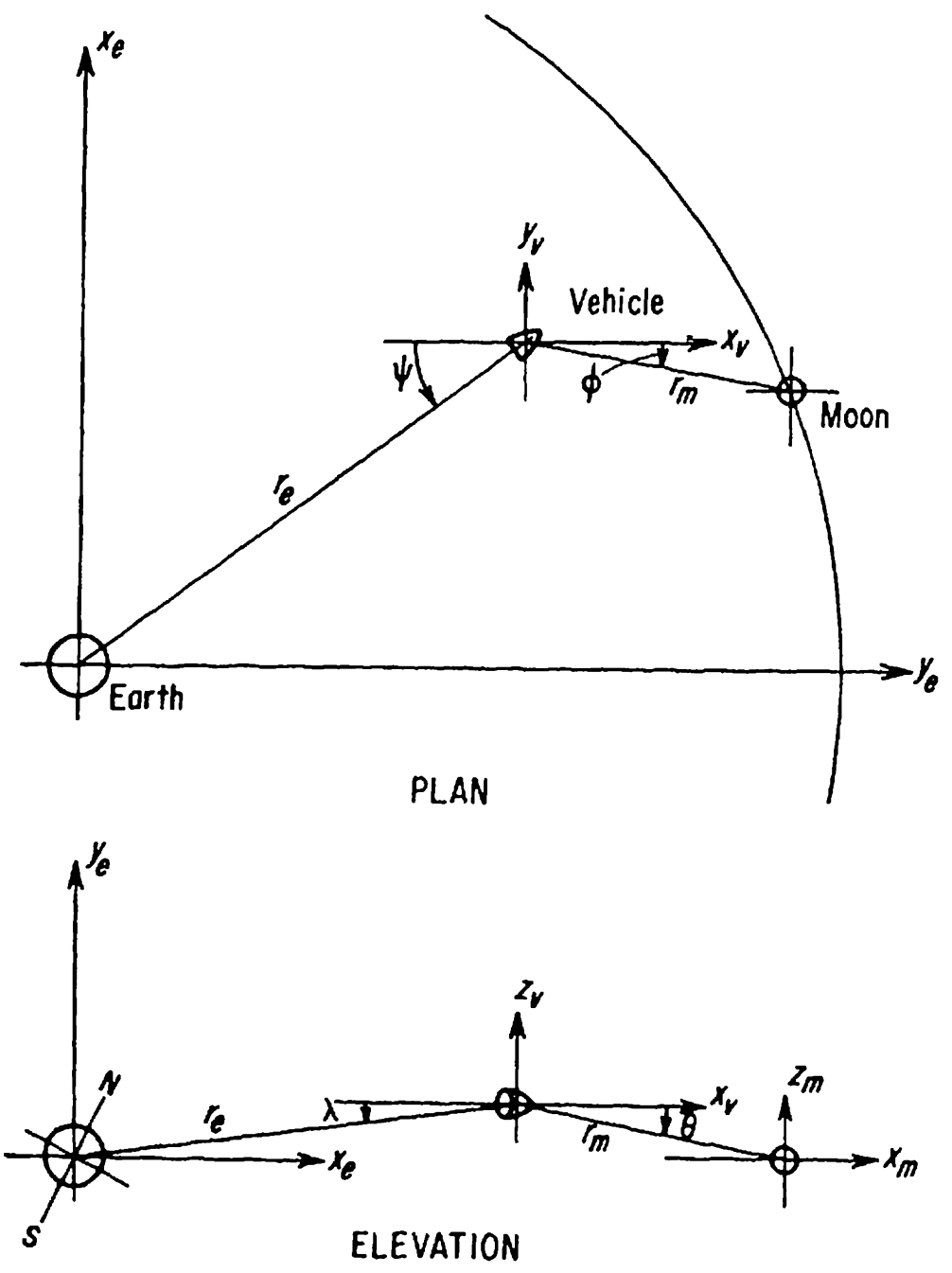


Fig. A13.1 Geometrics of gravity field vector deduction scheme.

Consider the case of the earth-moon and sun system, with the moon as a target, and a hypothetical x, y, z system with the xy plane parallel to the lunar ecliptic plane. (See Fig. A13.1.) The total acceleration indication at the vehicle point V will be

$$\begin{aligned} a_{i_x} &= a_{x_t} + g_{x_e} + g_{x_m} + g_{x_s} \\ a_{i_y} &= a_{y_t} + g_{y_e} + g_{y_m} + g_{y_s} \\ a_{i_z} &= a_{z_t} + g_{z_e} + g_{z_m} + g_{z_s}. \end{aligned} \tag{99}$$

Assuming at first a spherical† earth and moon, at a certain range the gravitational field may be assumed to be independent of declination, δ. The sun's gravitational field may also be considered small and relatively constant.‡ Then, resolving the components of the earth's field upon the hypothetical vehicle coordinate system, one obtains

$$\begin{aligned} a_{x_i} &= a_{x_t} - g_e\left(\frac{a_e}{r_e}\right)^2 \cos\lambda \cos\psi + g_m\left(\frac{a_m}{r_m}\right)^2 \cos\theta \cos\phi + g_{s_x} \\ a_{y_i} &= a_{y_t} - g_e\left(\frac{a_e}{r_e}\right)^2 \cos\lambda \sin\psi + g_m\left(\frac{a_m}{r_m}\right)^2 \cos\theta \sin\phi + g_{s_y} \\ a_{z_i} &= a_{z_t} - g_e\left(\frac{a_e}{r_e}\right)^2 \sin\lambda + g_m\left(\frac{a_m}{r_m}\right)^2 \sin\phi + g_{s_z}. \end{aligned} \tag{100}$$

From these three equations, the three unknowns, r_m, ϕ, and θ, can be obtained by solving simultaneously. Thus, the range and bearing to the target are obtained implicity.

An additional interesting result for the condition of the vehicle being between powered epochs is that the vehicle may be considered to be in a state of "free fall" between the earth-moon-sun system. Then the accelerometers would indicate zero, and target position could be deduced without recourse to spaceborne accelerometers and associated platform and equipment merely by registering of the vehicle hypothetical coordinate system in a memory, noting the vehicle's instantaneous acceleration and making implicit deductions. Based upon the relationships between the kinetic and potential energy and present course of the vehicle, miss distances could be calculated, short corrective bursts may be applied, and the target position deduction resumed.

† The effects of oblateness are discussed in a later section.

‡ If an error analysis should indicate that the variation in g's is appreciable, compensations easily can be effected.

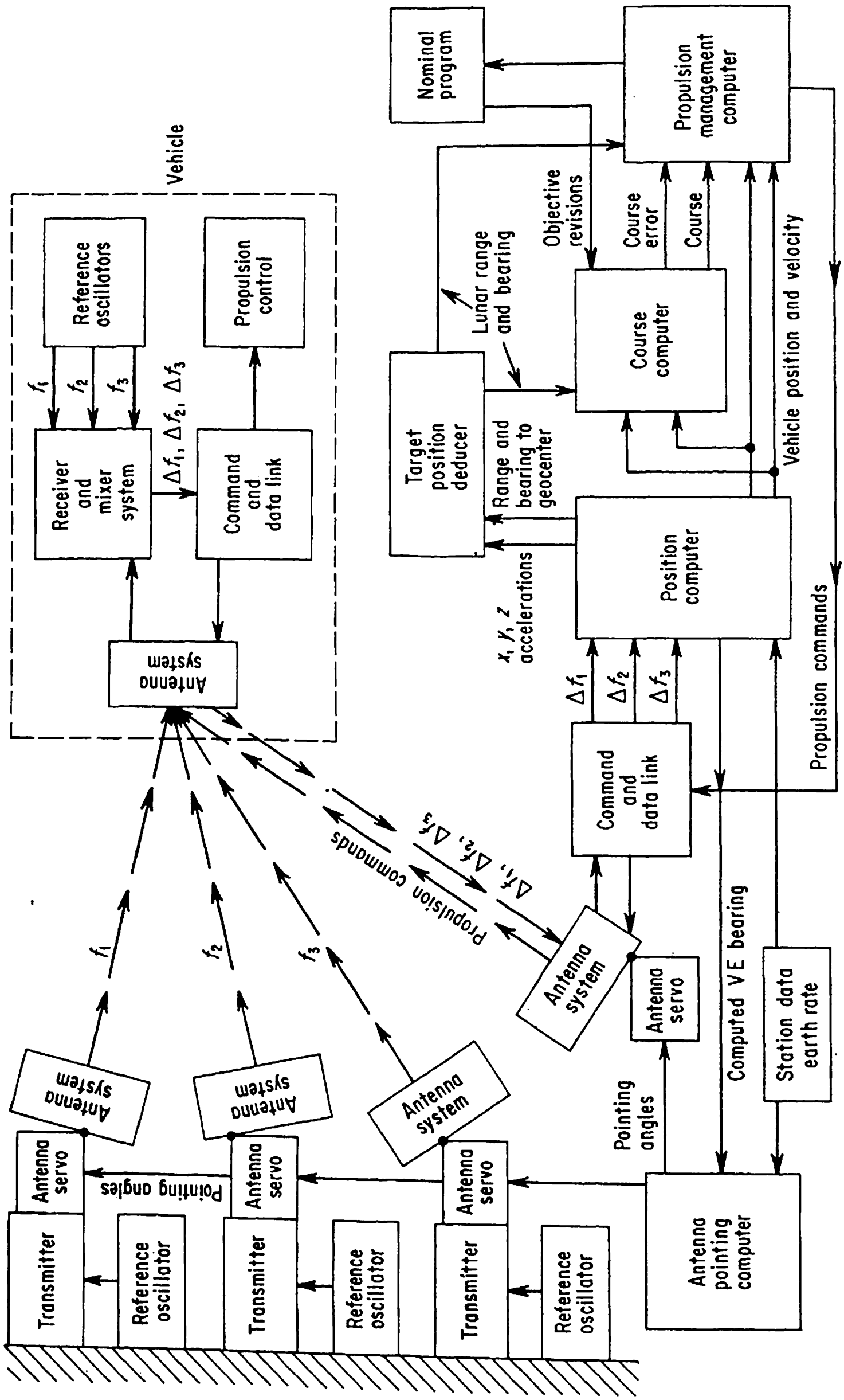


Fig. A13.2 Command system.

For the purpose of discussion, consider a lunar probe of hyperextreme minimality. If satisfactory attitude control could be achieved by spin stabilization,† if feasible axial-only corrective and retro impulses‡ could be achieved, and if the power requirement of a non-directed antenna§ could be tolerated (after trade-off considerations), a command system in the first order delineation would take the form of Fig. A13.2. In the case of the temporal correlation scheme there would result further reduction of spaceborne equipment because of the absence of a transmitter.

For a more realistic case, energy minimization and management (engine orientation), reconnaissance requirements, antenna orientation, and boost-out guidance considerations will require the incorporation of a stellar orientation supervised inertial platform. System considerations are elaborated later.

The above-mentioned scheme illustrated in Fig. A13.2 implies accurate information derived from the Doppler effect, obtained accelerations, and a knowledge of the mass (and to a lesser degree, gravitational anomalies) of the bodies concerned. Data examined¶ suggest that mass is known and gravitational anomalies of the earth and moon are known to an accuracy of at least 10^{-5}. This is approximately the accuracy required of the navigation bearing and range information when a range differential trajectory is flown.

B. Summary of the General Notion of a Spaceborne System

The basic notion of the scheme may be stated thus: if the vehicle radiated three distinct frequencies to three separate base-line stations, and if each of these stations amplified and retransmitted these frequencies back to the vehicle, the vehicle could operate upon the resultant frequency shift, or it could transmit the shift information to the command system.

Another modification, for accuracy reasons discussed below, essentially would effect "coherence" of the three earth signals by using a common oscillator and transponding to the three stations on a time-sharing basis.

† R. W. Buchheim, "Lunar Instrument Carrier-Attitude Stabilization," RAND #*RM-1730*. Requirements can be satisfied by an 80 rpm spin about roll axis.

‡ E. C. Heffern, "Lunar Instrument Carrier-Vehicle Design Considerations," RAND #*RM-1732* (secret).

§ R. T. Gabler, "Lunar Instrument Carrier-Tracking and Communication," RAND #*RM-1731* (confidential).

¶ Smithsonian tables.

However, in this case, information theory considerations and equipment feasibility questions need to be resolved.

Another scheme would be a passive system, using the matching of atomic clocks (MASERS), having a spaceborne frequency reference which is compared with an identical transmitted frequency at the base. However, the theoretically extractable accuracies of information appear to be below those obtainable with the active system discussed above, for the same order of oscillator accuracy.

C. Relativistic Doppler Phenomenon

(1) Theory: General. This discussion is a mathematical inquiry into the feasibility of using the Doppler effect of electromagnetic signals to determine the position of a space vehicle at any point of its path. As a preliminary consideration, we should hold in abeyance the effect of the acceleration on the Doppler shift and limit the development to special relativistic effects only.

It is well known that the Maxwell field equations of electromagnetic theory are Lorentz invariant. It can easily be shown that the wave equation (obtained from Maxwell's equations) and its general solutions are also Lorentz invariant. It will suit our purpose then to present briefly the relativistic transformation of wave motion.

In some medium whose index of refraction shall be assumed to be unity, let there be a source L of electromagnetic signal, the emission to be represented as

$$\psi = \exp 2\pi i(kr - ft + \epsilon), \tag{101}$$

a solution of the wave equation in a reference system S,

where f = frequency,
k = reciprocal wavelength,
ϵ = phase.

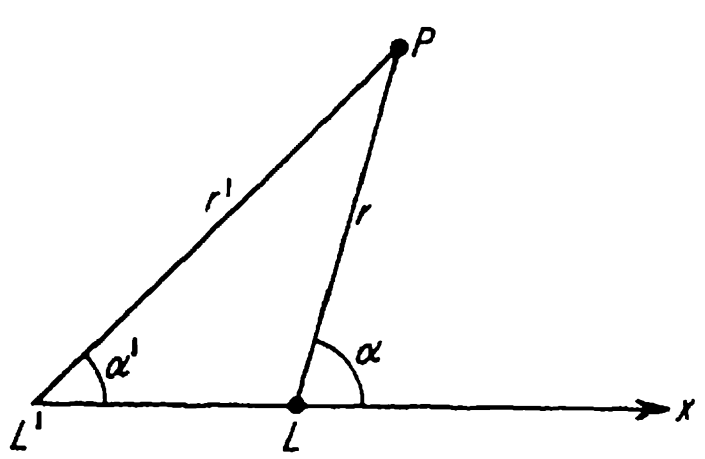


Fig. A13.3 Coordinate angles.

Let α be the angle of a ray from L to P, an observer at rest in the S system, with respect to which α also is measured relative to the X-axis as in Fig. A13.3.

Then $r = x \cos \alpha + y \sin \alpha$ so that equation (101) can be written as

$$\psi = \exp 2\pi i(kx \cos \alpha + ky \sin \alpha - ft + \epsilon). \tag{102}$$

For an observer in a system S', in simple relative translation with respect to S, the same signal beam is represented by

$$\psi' = \exp 2\pi i(k'x' \cos \alpha' + k'y' \sin \alpha' - f't' + \epsilon'). \tag{103}$$

In view of the Lorentz invariance of the solutions of the wave equation, it follows that the application of the Lorentz transformation to Eq. (101) must yield Eq. (104) at all times and places. The Lorentz transformation for the coordinates and time is given by:

$$\begin{aligned} x' &= \gamma(x - vt) & \gamma &= (1 - \beta^2)^{-1/2} & x &= \gamma(x' + vt') \\ t' &= \gamma(t - vx/c^2) & \beta &= v/c & t &= \gamma(t' + vx'/c^2). \end{aligned} \tag{104}$$

Introducing the transformations

$$y = y'; \qquad z = z'; \qquad x' = x'(x, t) \quad \text{and} \quad t' = t'(x, t), \tag{105}$$

we find the condition

$$\begin{aligned} kc \cos \alpha + ky \sin \alpha - ft + \epsilon &= k'\gamma(x - vt) \cos \alpha' + k'y \sin \alpha' \\ &\quad - f'\gamma(t - vx/c^2) + \epsilon'. \end{aligned} \tag{106}$$

Equating coefficients of x and t and recalling that $f/k = f'/k' = c$, we find the relations

$$f = \gamma(vk' \cos \alpha' + f') = \gamma f'\left(1 + \frac{v}{c} \cos \alpha'\right) \tag{107}$$

or

$$f = f'\gamma(1 + \beta \cos \alpha'), \tag{108}$$

$$k \cos \alpha = \gamma(k' \cos \alpha' + f'v/c^2) = k'\gamma\left(\cos \alpha' + \frac{v}{c}\right) \tag{109}$$

or

$$k \cos \alpha = k'\gamma\ (\cos \alpha' + \beta). \tag{110}$$

Dividing Eq. (110) by Eq. (108) we have

$$\cos \alpha = \frac{\cos \alpha' + \beta}{1 + \beta \cos \alpha'}. \tag{111}$$

Also, since $k \sin \alpha = k' \sin \alpha'$, we have from Eq. (108)

$$\tan \alpha = \frac{\sin \alpha'}{\gamma(\cos \alpha' + \beta)}. \tag{112}$$

From Eq. (111) and Eq. (112), we find

$$\sin \alpha = \frac{\sin \alpha'}{\gamma(1 + \beta \cos \alpha')}. \tag{113}$$

We note that Eq. (111) shows that when $\alpha' = 0$, then $\alpha = 0$. Thus, we have the condition that the relative motion of the source and observer lies along the line of sight. Under this condition, Eq. (107) becomes

$$f' = f\sqrt{\frac{1 - \beta}{1 + \beta}}. \tag{114}$$

This is known as the longitudinal or radial Doppler shift. The corresponding wavelength relationship is

$$\lambda' = \lambda\sqrt{\frac{1 + \beta}{1 - \beta}}. \tag{115}$$

Again, Eq. (111) shows that if $\alpha = \pi/2$,. then $\cos \alpha' = -\beta$ so that Eq. (107) becomes

$$f = f'\sqrt{1 - \beta^2}. \tag{116}$$

Correspondingly,

$$\lambda' = \lambda\sqrt{1 - \beta^2}. \tag{117}$$

These last two equations express the transverse Doppler effect and would be observed from a source, the relative motion of which is perpendicular to the line of sight.

We note further that Eq. (116) expresses the time dilation,

$$t' = t\sqrt{1 - \beta^2}, \tag{118}$$

since $f' = 1/t'$ and $f = 1/t$. Similarly, Eq. (117) is an expression of the length contraction directly

$$l' = l\sqrt{1 - \beta^2}. \tag{119}$$

We can see that a discussion of the Doppler effect is tantamount to a discussion of time dilation.

We shall suppose now that at the start and end of some period of acceleration of a space vehicle, its speeds are v_1 and v_2, respectively. Let $v_2 \geq v_1$ so that $\beta_2 \geq \beta_1$. If now an observer on the vehicle knows the rest frequency f_0 of the source, he will decide, on the basis of Eq. (107), that

$$f_1 = f_0\gamma_1(1 - \beta_1 \cos \alpha_1) \quad \text{and} \quad f_2 = f_0\gamma_2(1 - \beta_2 \cos \alpha_2), \tag{120}$$

where we have assumed that the motion is that of recession so that $f_2 < f_1$. We now establish the relation

$$\frac{f_1 - f_2}{f_0} = \gamma_1(1 - \beta_1 \cos \alpha_1) - \gamma_2(1 - \beta_2 \cos \alpha_2), \tag{121}$$

which when expanded to terms of second order in β yields

$$\frac{f_1 - f_2}{f_0} = (1 - \beta_1 \cos \alpha_1)\left(1 + \frac{1}{2}\beta_1\right) - (1 - \beta_2 \cos \alpha_2)\left(1 + \frac{1}{2}\beta_2\right), \tag{122}$$

so that, to first order in β, we have

$$\frac{f_1 - f_2}{f_0} = \beta_2 \cos \alpha_2 - \beta_1 \cos \alpha_1. \tag{123}$$

Correspondingly,

$$\lambda_1 = \lambda_0/\gamma_1(1 - \beta_1 \cos \alpha_1) \tag{124}$$

$$\lambda_2 = \lambda_0/\gamma_2(1 - \beta_2 \cos \alpha_2), \tag{125}$$

and since $\lambda_2 > \lambda_1$

$$\frac{\lambda_2 - \lambda_1}{\lambda_0} = [\gamma_2(1 - \beta_2 \cos \alpha_2)]^{-1} - [\gamma_1(1 - \beta_1 \cos \alpha_1)]^{-1}, \tag{126}$$

which to second order in β is

$$\frac{\lambda_2 - \lambda_1}{\lambda_0} = \left(1 - \frac{1}{2}\beta_2\right)(1 + \beta_2 \cos \alpha_2) - \left(1 - \frac{1}{2}\beta_1\right)(1 + \beta_1 \cos \alpha_1), \tag{127}$$

so that to first order in β we have

$$\frac{\lambda_2 - \lambda_1}{\lambda_0} = \beta_2 \cos \alpha_2 - \beta_1 \cos \alpha_1. \tag{128}$$

We note from Eqs. (123) and (122) that

$$\frac{\lambda_2 - \lambda_1}{\lambda_0} = \frac{f_1 - f_2}{f_0}. \tag{129}$$

(2) Accuracies. The smallest frequency change of electromagnetic signals which can presently be measured is interestingly enough in the microwave region, though one might have suspected it to be in the optical region. This is due to the techniques available by using MASERS.

We see from (123) or (128) that when the motion is along the line of sight, then $\cos \alpha_2 = \cos \alpha_1 = 1$, so that the smallest change of speed which can be detected would be given by Eq. (128) as

$$\frac{\lambda_2 - \lambda_1}{\lambda_0} = \frac{v_2 - v_1}{c}. \tag{130}$$

A conservative estimate for $(\lambda_2 - \lambda_1)/\lambda_0$ by the use of MASERS is 10^{-10} so that

$$v_2 - v_1 = c \cdot 10^{-10} = 3 \text{ cm/sec.} \tag{131}$$

In the event that $\beta_1 = \beta_2 =$ constant, then

$$\cos \alpha_2 - \cos \alpha_1 = 1/10^{10}\beta \tag{132}$$

or

$$-2 \sin \tfrac{1}{2}(\alpha_2 + \alpha_1) \sin \tfrac{1}{2}(\alpha_2 - \alpha_1) = 1/10^{10}\beta. \tag{133}$$

Over a short time interval during the motion, $\alpha_1 \simeq \alpha_2 + \Delta\alpha$, so that we can write the approximation

$$-\Delta\alpha \sin (\alpha_2 + \Delta\alpha/2) = 1/10^{10}\beta. \tag{134}$$

For a given β, $\Delta\alpha$ will be smallest when $\sin (\alpha_2 + \Delta\alpha/2) = 1$; i.e., when the motion is perpendicular to the line of sight. Thus for a speed of 33, 500 mph, $\beta = 5 \cdot 10^{-5}$ so that

$$\Delta\alpha = 2 \cdot 10^{-6} \text{ radians.} \tag{135}$$

For $\Delta\alpha = 1$ microradian $= 4.84 \cdot 10^{-6}$ radians we would have to have

$$\beta = 2.06 \cdot 10^{-5}. \tag{136}$$

Hence, for a speed $v = 13{,}830$ mph at right angles to the line of sight, it would be possible to ascertain an angular change no smaller than one microradian.

(3) Curves of Constant Frequency Shift. It would be desirable to know the minimum changes of v and α at conditions other than along and perpendicular to the line of sight. To do this we shall examine first a special case of Eq. (107), namely the condition that $f = f'$. It would appear at first sight that such a restriction is trivial, since it is satisfied when there is no relative motion. However, interestingly enough, if the relative motion is that of a logarithmic spiral, the condition $f = f'$ will prevail as a special case of the general condition $f = kf'$, where k is some constant other than unity. From Eq. (107) we obtain, with the restriction $f = f'$, the relation

$$\sqrt{1 - \beta^2} = 1 - \beta \cos \alpha, \tag{137}$$

and on squaring both sides we obtain

$$\beta^2(1 + \cos^2\alpha) - 2\beta \cos \alpha = 0, \tag{138}$$

for which we obtain the trivial solution $\beta = 0$. The other solution is

$$\beta = \frac{2 \cos \alpha}{1 + \cos^2 \alpha}. \tag{139}$$

We see that the denominator is always positive. Solving Eq. (139) for $\cos\alpha$ we find

$$\cos\alpha = \frac{1}{\beta}(1 - \sqrt{1 - \beta^2}), \tag{140}$$

where the positive value of the square root has been discarded since $0 \leq \cos\alpha \leq 1$. Thus, with the restriction $f = f'$, there is an unambiguous correspondence between β and $\cos\alpha$; and for a given β, there is only one value of $\cos\alpha$, such that $f = f'$.

We can now demonstrate that $f = f'$ is uniquely satisfied if the motion takes place on a logarithmic spiral. Let the motion be described by a radial distance ρ from the origin in a plane containing the x and y axes (see Fig. A13.4). The equation of the radius (corresponding to the line of sight) to any point (x_0, y_0) of the path is $y = m_1x$. The equation of the tangent to the curve (corresponding to the velocity of the point) at the point (x_0, y_0) is $y = m_2x + b$. At the point (x_0, y_0), the tangent is given by $y' = f'(x_0)$ so that $b = y_0 - y_0'x_0$. Hence, we have $y = y_0'(x - x_0) + y_0$. Clearly, $m_1 = y_0/x_0$ and $m_2 = y_0'$. Now let $\tan\alpha = m$ and we will have

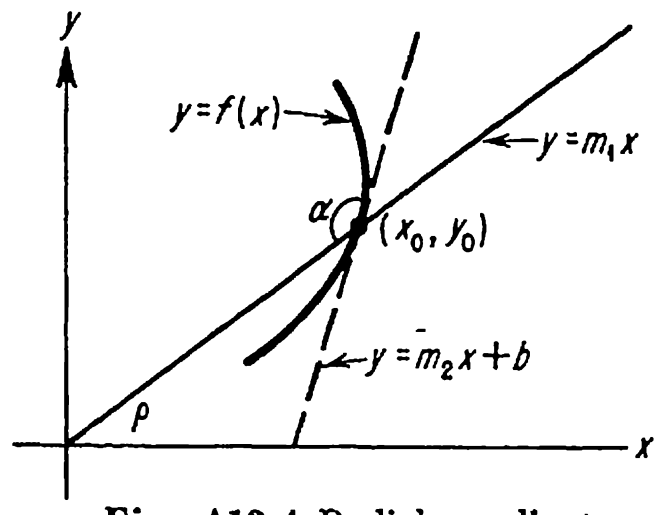


Fig. A13.4 Radial coordinate.

$$m = \frac{m_2 - m_1}{1 + m_1m_2}. \tag{141}$$

The restriction $f = f'$ now is noted by the fact that m equals a constant. Introducing m_1 and m_2, we obtain

$$m = \frac{y'x - y}{x + y'y}, \tag{142}$$

where the subscripts have been dropped since the point (x_0, y_0) is arbitrary. Solving for y', we find

$$y' = \frac{y + mx}{x - my}. \tag{143}$$

Solving this differential equation, the result is

$$(x^2 + y^2)^{1/2} = C \exp\left[\frac{\tan^{-1}(y/x)}{m}\right]. \tag{144}$$

Let $\tan^{-1}(y/x) = \theta$, $C = \rho_0$. And since $(x^2 + y^2)^{1/2} = \rho$, we have the equation of the logarithmic spiral in polar coordinates

$$\rho = \rho_0 e^{\theta/m}. \tag{145}$$

(4) Expressions for the Radial Component of Velocity. The expression (107)

$$f = f'\gamma(1 + \beta \cos \alpha')$$

and a similar expression

$$f' = f\gamma(1 - \beta \cos \alpha), \tag{146}$$

which arise when the transformation is $x = x(x't')$ and $t = t(x't)$, can be used to relate a frequency shift to the radial component of the velocity of the observer relative to the source, or of the source relative to the observer.

If a signal from $B(f_B)$ is received by $A(f_A')$ the second expression gives

$$f_A' = f_B\gamma(1 - \beta_r), \tag{147}$$

where $c\beta_r = c\beta \cos \alpha_B$ = radial component velocity of the observer as measured from the frame of reference of the source. Solving for β_r

$$\beta_r = \frac{f_B - f_A'}{\gamma f_B} + 1 - \frac{1}{\gamma} \tag{148}$$

or

$$\gamma\beta_r = \frac{f_B - f_A'}{f_B} + \gamma - 1. \tag{149}$$

Thus, to find the radial component of velocity, a relativistic correction must be applied to frequency shift measurement which depends upon the ratio of the speed to the speed of light. This is small but may not be negligible. At speeds of 10^6 cm/sec (25,000 mph), it changes velocity measurement by about 30 cm/sec.

If, though, a signal from $A(f_A)$ is received by $B(f_B)$, re-radiated and received by $A(f_A')$, the expressions (107) and (108) must be combined

$$f_A = f_B\gamma(1 - \beta \cos \alpha_B) \quad \text{for the first signal,} \tag{150}$$

$$f_A' = f_B\gamma(1 + \beta \cos \alpha_B') \quad \text{for the second signal.} \tag{151}$$

Here, it has been assumed that the relative speed has not changed during the time taken for the signal to travel from A to B and back again, but that the motion of A with respect to B has caused a change in the angle α_B to

change to α_B'. These two expressions can be combined to give

$$\frac{f_A}{f_A'} = \frac{1 - \beta \cos \alpha_B}{1 + \beta \cos \alpha_B'} = \frac{1 - \beta_r}{1 + \beta_r'} \tag{152}$$

or

$$f_A\beta_r' + f_A'\beta_r = f_A' - f_A, \tag{153}$$

where $c\beta_r$ is the radial component of relative velocity when the signal leaves A and $c\beta_r'$ when it returns to A. The non-relativistic Doppler shift expression holds for the mean velocity $\bar{\beta}_r$ defined by

$$\bar{\beta}_r = \frac{f_A\beta_r' + f_A'\beta_r}{f_A + f_A'}. \tag{154}$$

(This becomes the usual average of β_r' and β_r if $f_A = f_A'$, and if it lies between β_r and β_r'.) Then with expression (153) this becomes

$$\bar{\beta}_r = \frac{f_A' - f_A}{f_A - f_A'}, \tag{155}$$

the non-relativistic Doppler shift expression. If, further, the speed of motion is much less than the speed of light, the frequency shift

$$f_d = f_A' - f_A \ll f_A' + f_A = 2f_B, \tag{156}$$

where f_B = frequency of transmitter. The usual radar Doppler shift expression

$$f_d = 2\bar{\beta}_r f_T = 2\frac{\bar{v}_r f_T}{c} \tag{157}$$

results.

Equation (157) indicates that a non-relativistic expression can be employed if it is used carefully.

A single measurement of a Doppler shift can lead to the knowledge of one component of velocity, and a time integration of this, to the radial distance between A and B, provided an initial value of the distance is known.

D. Relativistic Doppler Techniques

The Doppler technique can be applied to an astronautic vehicle as shown in the block diagram of Fig. A13.5. Briefly, the technique would be applied as follows. One or more gated *c-w* signals are generated in the vehicle

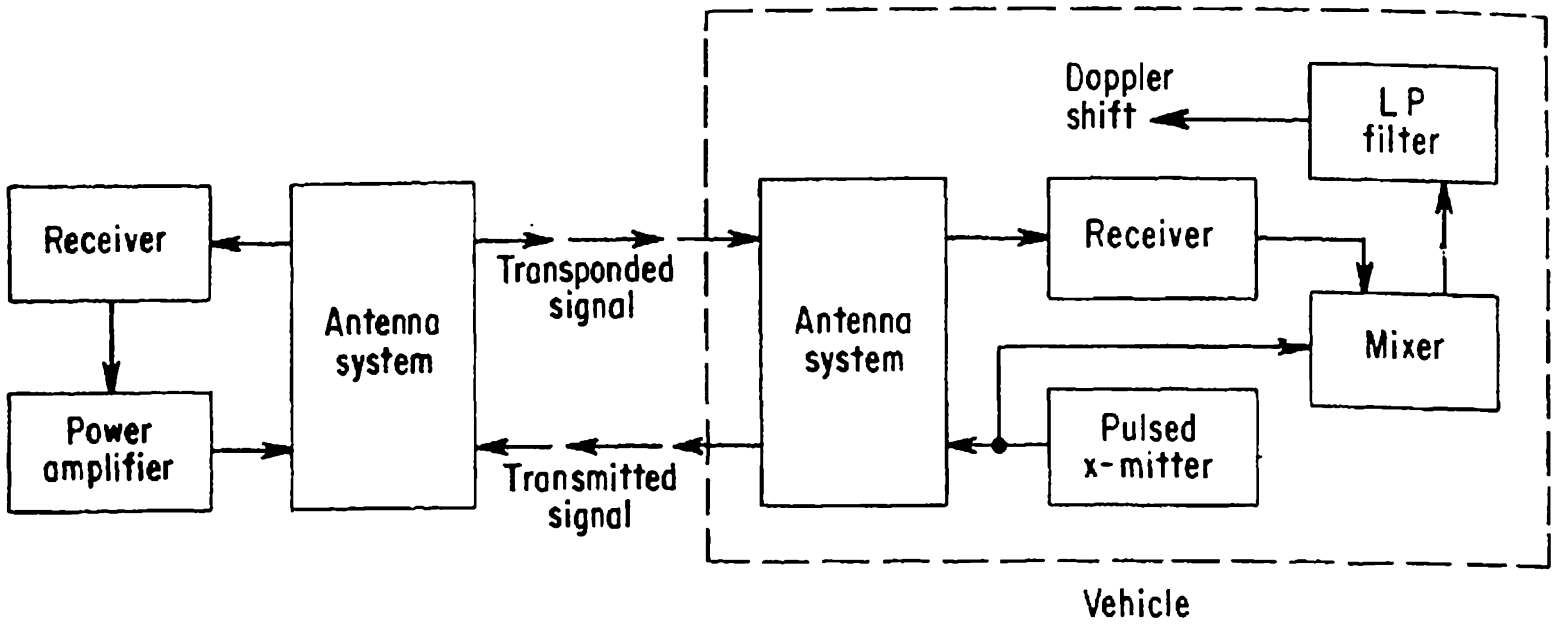


Fig. A13.5 "Reflected" Doppler scheme.

(A) and are transmitted to one or more stations (B, etc.) on the earth where they are reamplified and retransmitted to the vehicle. Then, return signals are mixed with the output of the transmitter oscillator in the vehicle, and the combined signal then is filtered. The output of the filters is the Doppler shift. Hereafter, this technique will be referred to as the transponded (or reflected) Doppler scheme.

The comparatively recent advent of "atomic amplifiers" provides a means of attaining the stabilities required for utilization of a second technique. The block diagram of Fig. A13.6 indicates one possible form of a

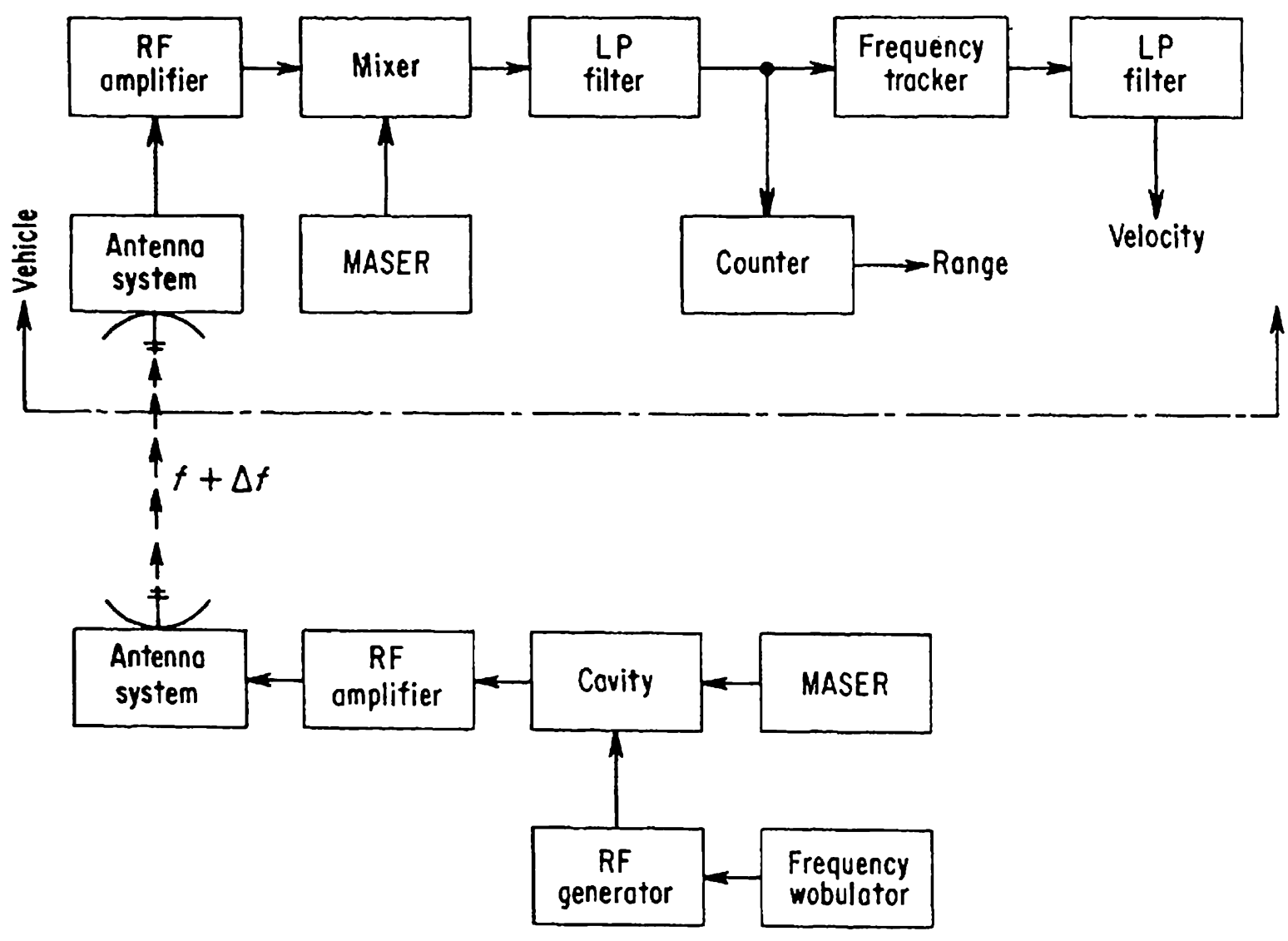


Fig. A13.6 MASER clock matching system.

Doppler system instrumentation employing the MASER (i.e., microwave amplification by stimulated emission of radiation). The frequency of the microwave transmission from the source (presumably earth located) is stabilized, and frequency is controlled by the maser operating in conjunction with a radio-frequency generator. The radio-frequency generator is modulated to provide a broadening of its frequency spectrum. This insures that the MASER cavity will still be excited at its resonant frequency in the presence of drift in the *rf* generator. The cavity is coupled to an *rf* amplifier, such as a traveling wave tube, the output of which is transmitted. The vehicle system contains a mixer which combines the received energy and the reference-frequency energy as obtained from its maser. The unwanted sidebands produced in the mixing process are again eliminated by means of a low-pass filter, resulting in the desired Doppler frequency shift.

The Doppler frequency in either case, when appropriately corrected, gives a component of the velocity which may be integrated by a counting device to give the change in distance, or to read velocity directly after suitable rectification and filtering.

E. Simplified Error Analysis

Of the several deleterious effects associated with a terrestrial-based Doppler system, which are discussed in the addendum to this appendix, only two are potential sources of difficulty in the astronautic application; i.e., transmitter frequency stability and Doppler spectrum tracker linearity. Design of the latter is discussed elsewhere in this appendix. The velocity error which may be expected to arise from transmitter drift may be determined to first order from the variation of one of the Doppler equations:

$$f_d = \frac{2}{c} f_T v_r \quad \text{round trip} \tag{158}$$

or

$$f_d = \frac{f_T v_r}{c} \quad \text{one way,} \tag{159}$$

where again f_d—is the measured difference frequency,

v_r—is the radial component of velocity of the vehicle with respect to the transponder or transmitter on the earth,

f_T—is the nominal frequency of the transmitter.

For the transponded scheme the first of Eqs. (158) applies and gives

$$\delta f_d = \frac{2}{c} f_T \delta v_r + \frac{2v_r}{c} \delta f_T - \frac{2v_r}{c^2} f_T \delta c. \tag{160}$$

Hence,

$$\frac{\delta v_r}{v_r} = \frac{\delta f_d}{f_d} - \frac{\delta f_T}{f_T} + \frac{\delta c}{c}. \tag{161}$$

Therefore, the relative error in determining the vehicle velocity is equal to the sum of the relative errors made in determining the Doppler shift, the transmitter frequency, and the speed of light.

The relative error in determining the Doppler shift will be a function of the accuracy of mechanization of the Doppler spectrum tracker. This error very likely will be the dominating system error. No literal accuracy figure can be quoted for the tracker mechanization. As a measure of the design problem involved, however, consider the fact that present *analog* trackers are good to 10^{-4}, and various theoretical arguments require that the presently proposed *digital* mechanization have a relative accuracy of at least 3×10^{-6}. Consequently, it is necessary to obtain an improvement factor of 100 over existing trackers. This should be possible, although abundant study will be needed.

Transmitter stability is at least 10^{-8} at present. A minimum requirement here is of the order of 10^{-7}; hence, little if any trouble is expected from this error source.

Since the velocity of light is presumed to be known to a relative accuracy of 10^{-6}, it follows from Eq. (161) that the relative error in velocity cannot be less than 10^{-6}. This is just within the desired accuracy bounds. (Calibration flights could reduce this error considerably.)

The second technique (Fig. A13.6) places a much more stringent requirement on the frequency standard. Consider the equation

$$\delta f_d = \delta f_s - \delta f_r, \tag{162}$$

where f_s is the frequency standard that is mixed with the received frequency f_r. Because of the independence of f_s and f_r, δf_d is of the order of δf_s or δf_r, whichever is larger. Equation (159) can be written

$$v_r = c\frac{f_d}{f_T} = c\frac{f_s - f_r}{f_T}. \tag{163}$$

The present situation is that MASERS (atomic clocks) operating in the microwave region have an estimated stability as high as 10^{-13} with a sta-

bility of 10^{-11} realized in established laboratory procedure, while optical interferometry permits a determination, in practice, of wave length to an accuracy of 10^{-8}. The optical limitation is due to the accuracy obtained in metrology on good ruled lines under the best viewing conditions.† Thus we can say that the smallest value of $f_s - f_r/f_T$ can be estimated conservatively as 10^{-10} to allow for degradation in a non-laboratory-type of MASER.

Thus, the smallest change of speed which can be detected would be

$$\delta V_r = C \times 10^{-10} = 3 \text{ cm/sec.} \tag{164}$$

Using optical interferometry this would be

$$\delta v_r = 300 \text{ cm/sec.} \tag{165}$$

The microwave result is comparable (for speeds of about 10^6 cm/sec or 25,000 mph) to the expected uncertainties resulting from previous considerations involving frequency spectrum trackers.

Because this technique is applicable to natural stellar sources, as well as MASER controlled terrestrial ones, it possibly can be utilized to advantage when the small-angle triangulation consideration to follow renders a terrestrial base line inadequate.

Radial distance will be determined by a direct count of the zero crossings of the Doppler shift. Hence,

$$\begin{aligned} N &= f_d \cdot t \\ &= \frac{2f_T}{c}(v_r) \cdot t \\ &= \frac{2f_T}{c} s. \end{aligned} \tag{166}$$

Therefore:

$$\delta N = -\frac{2f_T}{c^2} s\delta c + \frac{2s}{c}\delta f_T + \frac{2f_T}{c}\delta s \tag{167}$$

$$\frac{\delta N}{N} = \frac{\delta f_T}{f_T} - \frac{\delta c}{c} + \frac{\delta s}{s} \tag{168}$$

or

$$\frac{\delta s}{s} = \frac{\delta N}{N} - \frac{\delta f_T}{f_T}\frac{\delta c}{c}. \tag{169}$$

† Baird and Smith, *J. Op. Soc. A.*, Vol. XLVIII, No. 5, May, 1958, p. 300.

Now, $\delta N/N$ is simply the relative error in the count of the zero crossings of the Doppler frequency. For a signal-to-noise ratio of sufficient magnitude, this error rapidly approaches zero. Consequently, $\delta s/s$ can be made to the order of 10^{-6}.

If there are several base stations, B, C, D, etc., separated by known displacements, then the components of velocity in several (usually non-orthogonal) directions can be determined. The Doppler shifts can all be counted to give the distances AB, AC, AD, etc. With distances BC, BD, CD, etc. also known, it is possible to determine the position and velocity vectors of A relative to any stations. To illustrate, consider only A, B, and C (Fig. A13.7).

The angles a and α can be obtained from the law of cosines. The accuracy to which these angles will be determined can be estimated in the following manner for the case where s_1 and s_2 are large compared to d:

$$\alpha \simeq \frac{d \sin \theta}{s_1} \tag{170}$$

$$\delta\alpha = \frac{e \sin \theta}{s_1} - \frac{d \sin \theta \, \delta s_1}{s_1^2} \tag{171}$$

$$\frac{\delta\alpha}{\alpha} = \frac{e}{d} - \frac{\delta s_1}{s_1}, \tag{172}$$

$\dfrac{e}{d}$ should be no greater than $\dfrac{\delta s_1}{s_1}$, and

$$\frac{\delta\alpha}{\alpha} - 2\frac{\delta s}{s} \simeq 10^{-6}. \tag{173}$$

As long as θ is large, $30° \leq \theta \leq 90°$, and

$$\cos \theta \simeq \frac{s_1}{s_2}, \tag{174}$$

$$\sin \theta \, \delta \theta \simeq \frac{\delta s_1}{s_1} - \frac{s_1}{s_2^2} \delta s_2 = \frac{s_1}{s_2}\left(\frac{\delta s_1}{s_1} - \frac{\delta s_2}{s_2}\right). \tag{175}$$

If $\sin \theta \simeq \frac{1}{2}$, and $s_1 \simeq s_2$

$$\delta\theta \simeq 2\left(\frac{\delta s_1}{s_1} - \frac{\delta s_2}{s_2}\right). \tag{176}$$

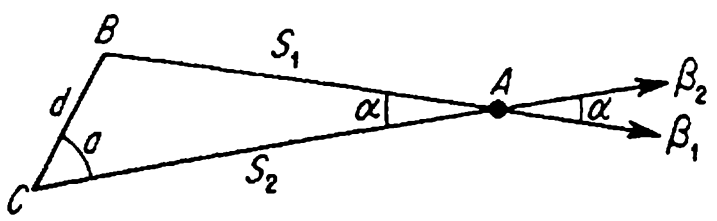


Fig. A13.7 Angle determination.

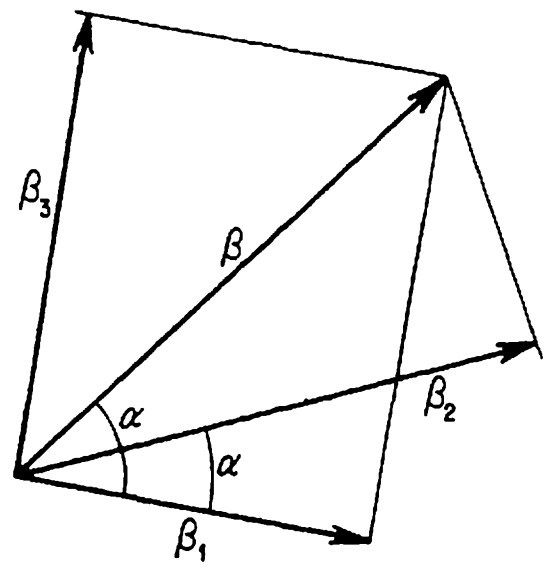


Fig. A13.8 Vector resolution.

At this point it is possible to conceive of two situations:

Case 1. In which δs_1 and δs_2 are uncorrelated, as would be the case if the radiation sources for f_1 and f_2 measurements were uncorrelated.

Case 2. In which there is a correlation in δs_1 and δs_2 so that they are of the same sign and nearly the same magnitude, as would be the case if the same frequency generator were used for the two radiation sources.

The error in $\delta\theta$ would be much less in Case 2. However, it is easier to estimate $\delta\theta$ for Case 1, since

$$\delta\theta \simeq \frac{\delta s}{s} \simeq 10^{-6} \text{ radians.} \tag{177}$$

An improvement factor of 10 or 100 might be realized in Case 2.

The components of velocity β_1 and β_2, together with the angle γ, determine a vector in this plane such that (see Fig. A13.8)

$$\beta_1 = \beta \cos \gamma \tag{178}$$

$$\beta_2 = \beta \cos (\gamma - \alpha). \tag{179}$$

It is useful to know the component of β perpendicular to either β_1 or β_2. If β_3 is the component of β perpendicular to β_1, then

$$\beta_3 = \beta \sin \gamma. \tag{180}$$

Eqs. (178) (179) can be used to determine β_3:

$$\beta_2 = \beta \cos \delta \cos \alpha - \beta \sin \gamma \sin \alpha \tag{181}$$

$$= \beta_1 \cos \alpha - \beta_3 \sin \alpha$$

$$\beta_3 = \beta_1 \operatorname{ctn} \alpha - \beta_2 \csc\alpha. \tag{182}$$

The accuracy to which β_3 may be determined here, as with the determination of θ, may be significantly improved if there is correlation in the errors in β_1 and β_2. However, considering random errors:

$$\delta\beta_3 \simeq \delta\beta_1 \operatorname{ctn} \alpha - \delta\beta_2 \csc \alpha - \beta_1 \csc^2 \alpha\delta\alpha + \beta_1 \operatorname{ctn} \alpha \csc \alpha\delta\alpha. \tag{183}$$

For small α, $\operatorname{ctn} \alpha \simeq \csc \alpha \simeq \frac{1}{\alpha}$, and

$$\delta\beta_3 \simeq \frac{\delta\beta_1}{\alpha} - \frac{\delta\beta_2}{\alpha} - \frac{\beta_1\delta\alpha}{\alpha^2} + \frac{\beta_2\delta\alpha}{\alpha^2} \tag{184}$$

$$\simeq 2\frac{\delta\beta}{\alpha} - \left(\frac{\beta_1 - \beta_2}{\alpha}\right)\frac{\delta\alpha}{\alpha}:$$

Where $\beta_1 - \beta_2$ is of the order of α,

$$\delta\beta_3 \simeq 2\frac{\delta\beta}{\alpha} - \frac{\delta\alpha}{\alpha} \simeq 2\delta\beta\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) \simeq 2\frac{\delta\beta}{\alpha}. \tag{185}$$

For a distance of the order of the moon and a reasonable separation, d, of the station of the earth,

$$\alpha \simeq \frac{1}{40}, \quad \text{and} \quad \frac{\delta\beta_1}{\beta} \simeq 10^{-6}; \tag{186}$$

$$\frac{\delta\beta_3}{\beta} \simeq 4 \times 10^{-5}. \tag{187}$$

If there were a correlation between $\delta\beta_1$ and $\delta\beta_2$, this figure probably would be improved.

4. System Considerations and Illustrations

A. General

Since the feasibility and theoretical accuracies of the relativistic Doppler and gravity field vector discrimination schemes have been examined, it is the purpose of this discussion to present a rudimentary delineation and evaluation of several possible illustrative implementation systems, to state the known or suspected major problem areas, and, where possible, to indicate a method of attack.

The techniques and configurations delineate elementary functional units

which must be more extensively delineated in the future. They should not, in any way, be interpreted as firm conclusions but as a preliminary basis which is subject to slight or extensive revision from the outcome of further system and error analysis.

Also, since the geometrics, dynamics, and transit times of lunar and interplanetary flights are considerably different in character and orders of magnitude, the configurations and performance requirements of the two operations necessarily will be considerably different, and possibly vastly so. The projected operational period will greatly influence the methodology; i.e., the existence of a geo-lunar command base line, the capability of precision electronic measurement of the visible spectral Doppler shift, etc., conceivably could obviate the need for spaceborne optical transits.

Lastly, the systems discussed do not delineate the boost-out guidance techniques, except that boost-out considerations are kept in mind regarding compatibility and utilization and sharing of functional elements. It might be presumed that a modified form of contemporary "Q" guidance scheme (or alternative ICBM schemes) will be employed.

Two main philosophies are presented with various possibilities or combinations: self-contained *vs.* command guidance. Since the degree of complexity involved in position and course computation, energy management, and antenna (ground and spaceborne) direction is quite high in the lunar mission case, it is suggested that the command system philosophy be pursued to the maximum. However, the over-all system and operations aspects will be considered; i.e., the technical problems of efficient data communication, closed-loop stability (computer and propagation delays) versus trade-offs with spaceborne equipment weight, power penalties and reliability, vehicle design considerations, etc.

The atomic clock comparison scheme and implementation will be elaborated for application in interplanetary operations.

B. Lunar Guidance

The excellent bearing, and velocity accuracies obtainable in the relativistic Doppler scheme combined with the gravity field vector discrimination scheme make them possible as a primary mode of midcourse and terminal guidance.

Several possibilities are presented in varying degrees of sophistication and/or redundancy, and with command versus self-contained guidance philosophies.

Although a theoretical minimal system which has no inertial elements appears possible, a two-star orientation-supervised platform of small weight, size, and moderate gyro performance is preferred for this discussion. The inertial unit is indicated for the following reasons:

(a) *Attitude control*

(1) Engine orientation for feasibility in maneuvers and for energy minimization;

(2) Antenna direction and stabilization;

(3) Optical instrument (when employed) stabilization and access;

(4) Relatively instantaneous source of impulse (integrated acceleration) feedback for engine control;

(5) Reconnaissance requirements.

(b) *Guidance*

(1) Boost-out guidance requirements;

(2) In midcourse and terminal thrusts of appreciable duration, or for the more futuristic case of a complementary impulse plus continuous thrust configuration, the gravity vector discrimination scheme requires registering all accelerations not results of the attraction of gravity fields;

(3) Accelerations owing to unpredictable perturbations (solar radiations, thrust perturbations, etc.);

(4) Precise target approach control.

C. Configuration A

(1) Functional Elements. This spaceborne configuration, at the moment, is felt to be most favored. It consists basically of a moderately advanced base line command system with equipments which consist of the following major functional units:

(1) Doppler transmitter and receiver system;

(2) Reference oscillators;

(3) Command and telemeter data link;

(4) Inertial platform and stellar supervisor;

(5) Attitude stabilization and propulsion control;

(6) Line-of-sight stabilization computer.

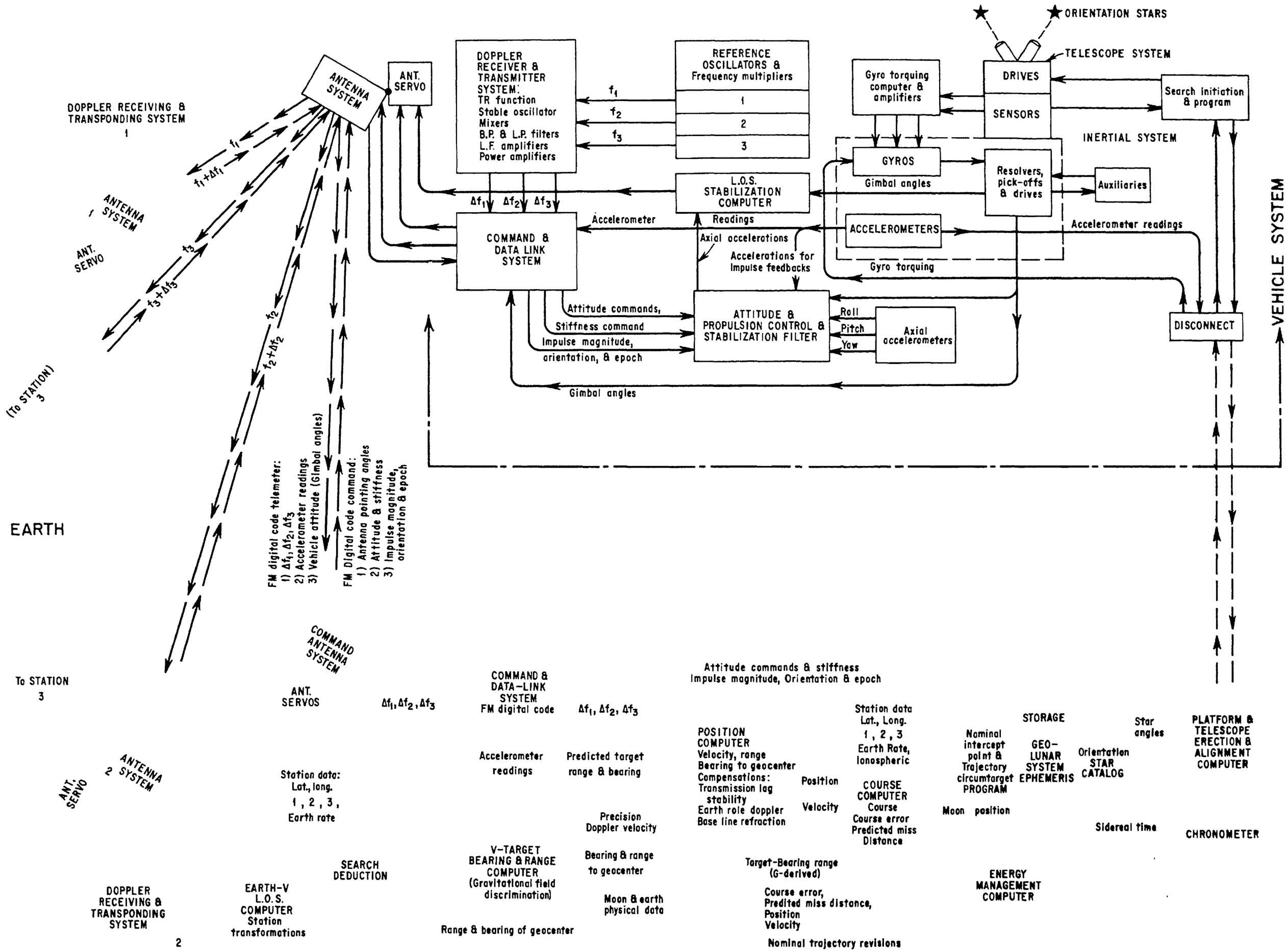


Fig. A13.9 Configuration A, lunar command guidance.

The base-line group is comprised of the following major elements:

(1) Command and telemeter data link;
(2) Base-line antennas pointing angle computer;
(3) Doppler transmitter and receiver systems (3 groups), and directed antenna systems (3);
(4) Vehicle position computer;
(5) Vehicle course computer;
(6) Energy minimization and propulsion management computer;
(7) Chronometer and geo-lunar system ephemeris (optional);
(8) Vehicle-target bearing and range deducer;
(9) Platform and telescope erection and alignment computer;
(10) Auxiliaries.

It should be noted first that Units (6) and (7) are optional. They have been included because, being ground-based, the weighting function of their system penalties is small, while their advantages in effecting great savings of spaceborne propulsion energy and equipment energy (considerations of transit time, reliability, etc.) may be great, as preliminary systems, dynamics, and error analysis indicate. Operations analysis and propulsion economics will ascertain their disposition.

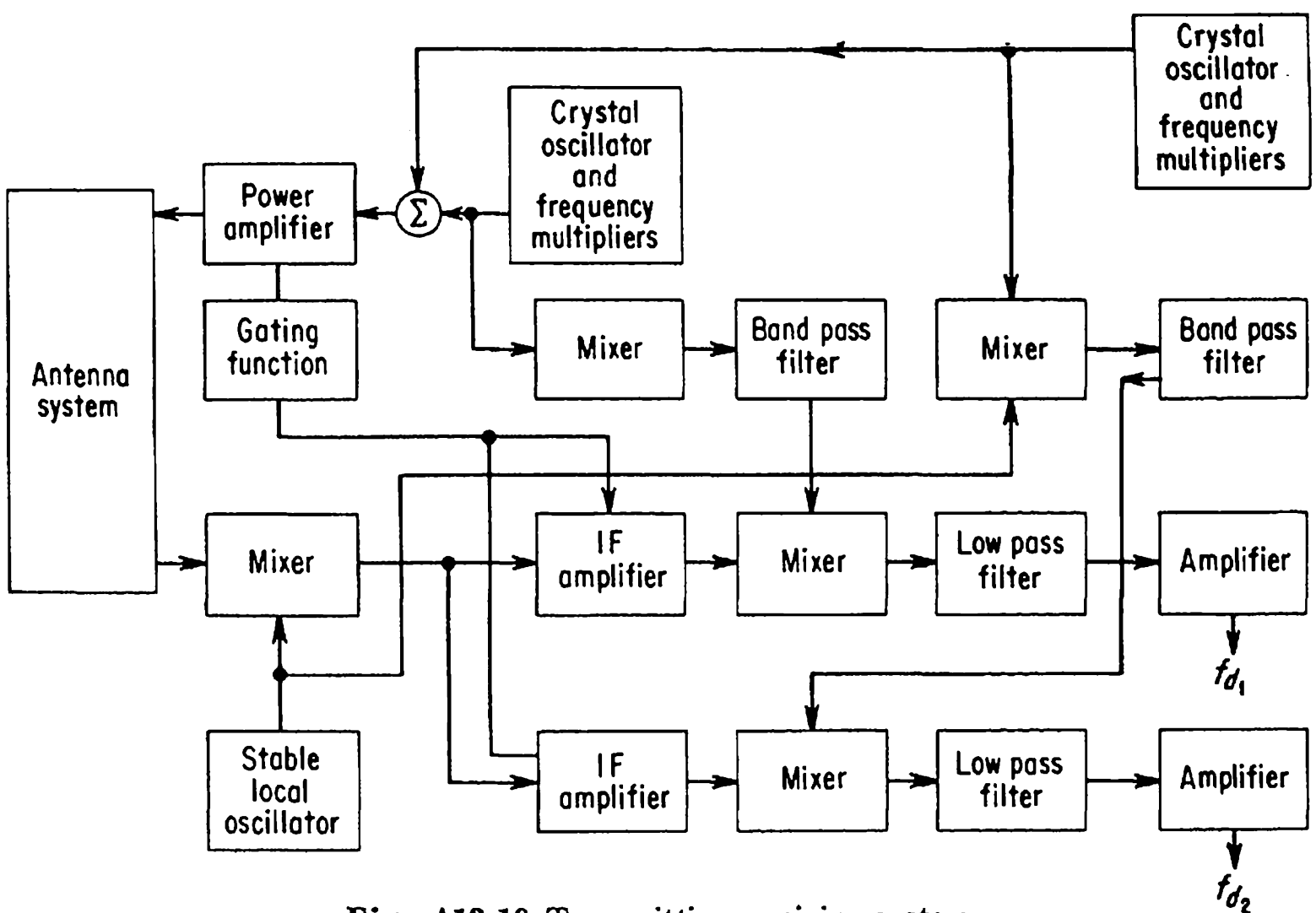


Fig. A13.10 Transmitting-receiving system.

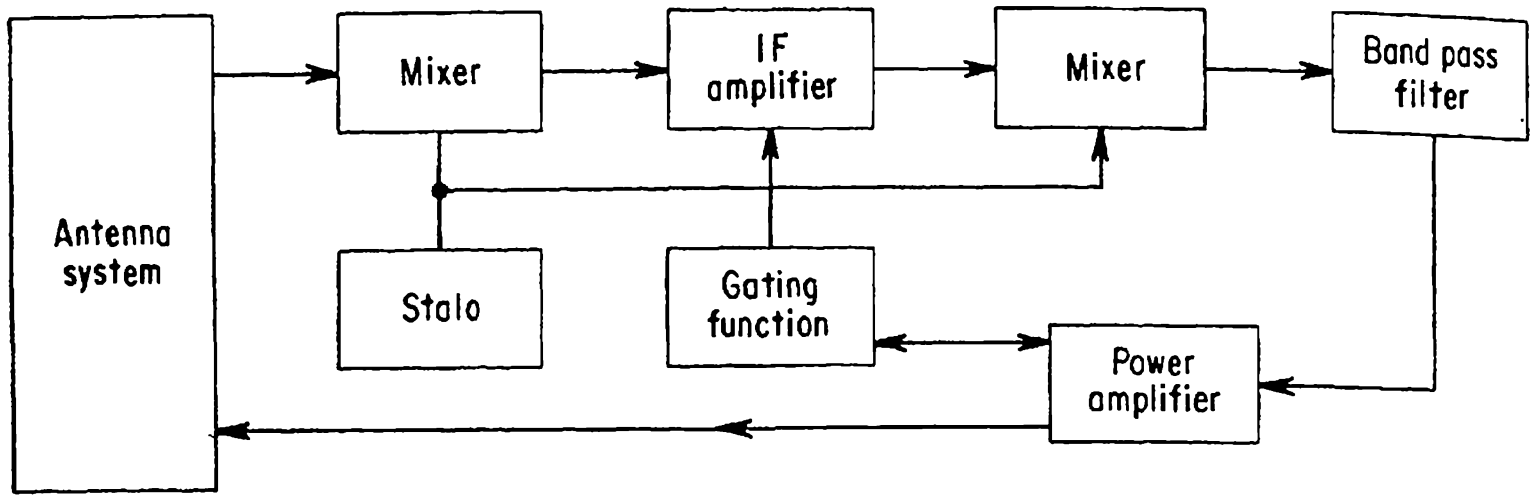


Fig. A13.11 Transponder system.

Secondly, it should be realized that the delineations such as course computer, position computer, etc., do not necessarily mean the existence of separate physical or computational entities, but are merely functional delineations.

The following presentation is confined to implementation and evaluation of the Doppler technique and gravitational discrimination target location schemes; broader treatment of other functions or comprehensive systems analysis will be touched upon only when they are pertinent. For example, the philosophy and analysis of programmed *vs.* flexible trajectory maneuvers, etc., have been, and will continue to be, subjects for study.

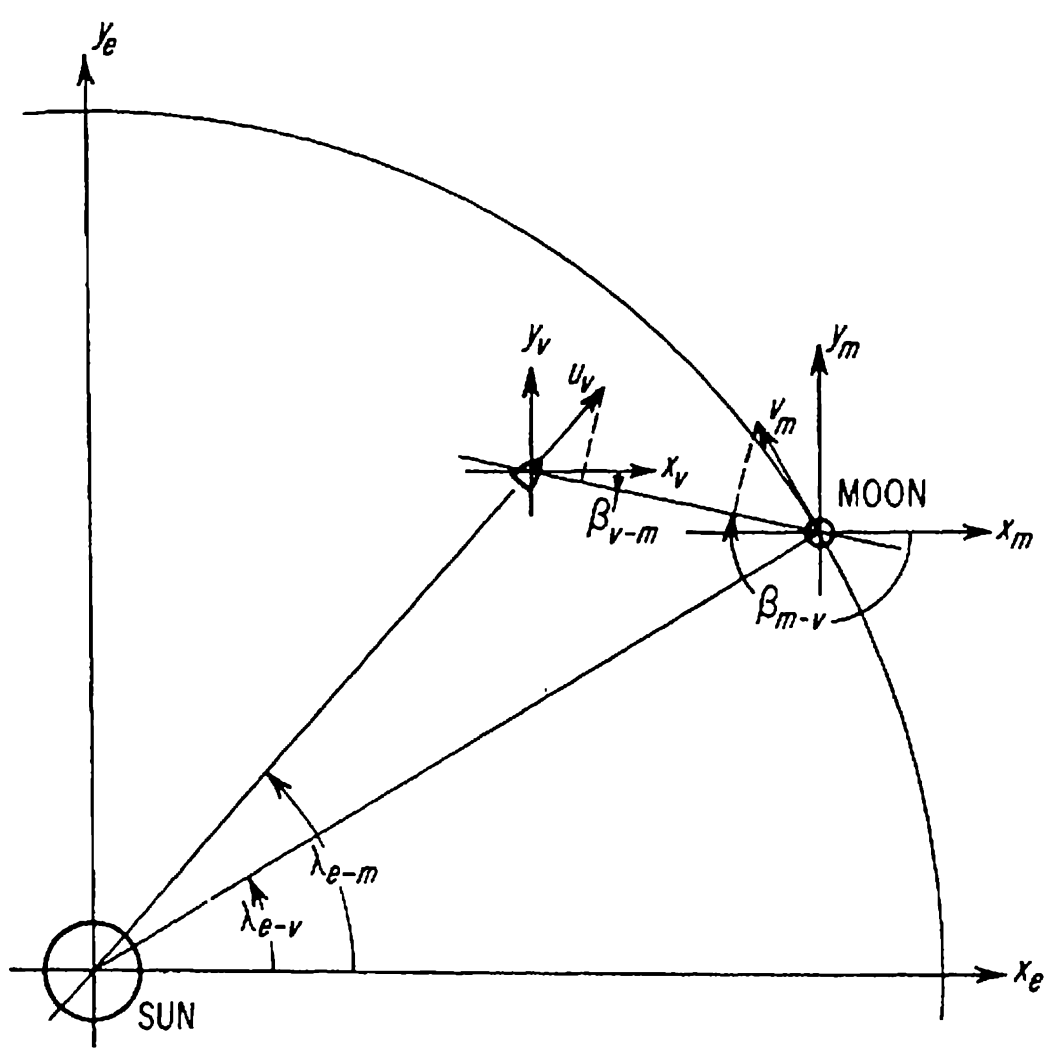


Fig. A13.12 Geometrics for dynamics analysis.

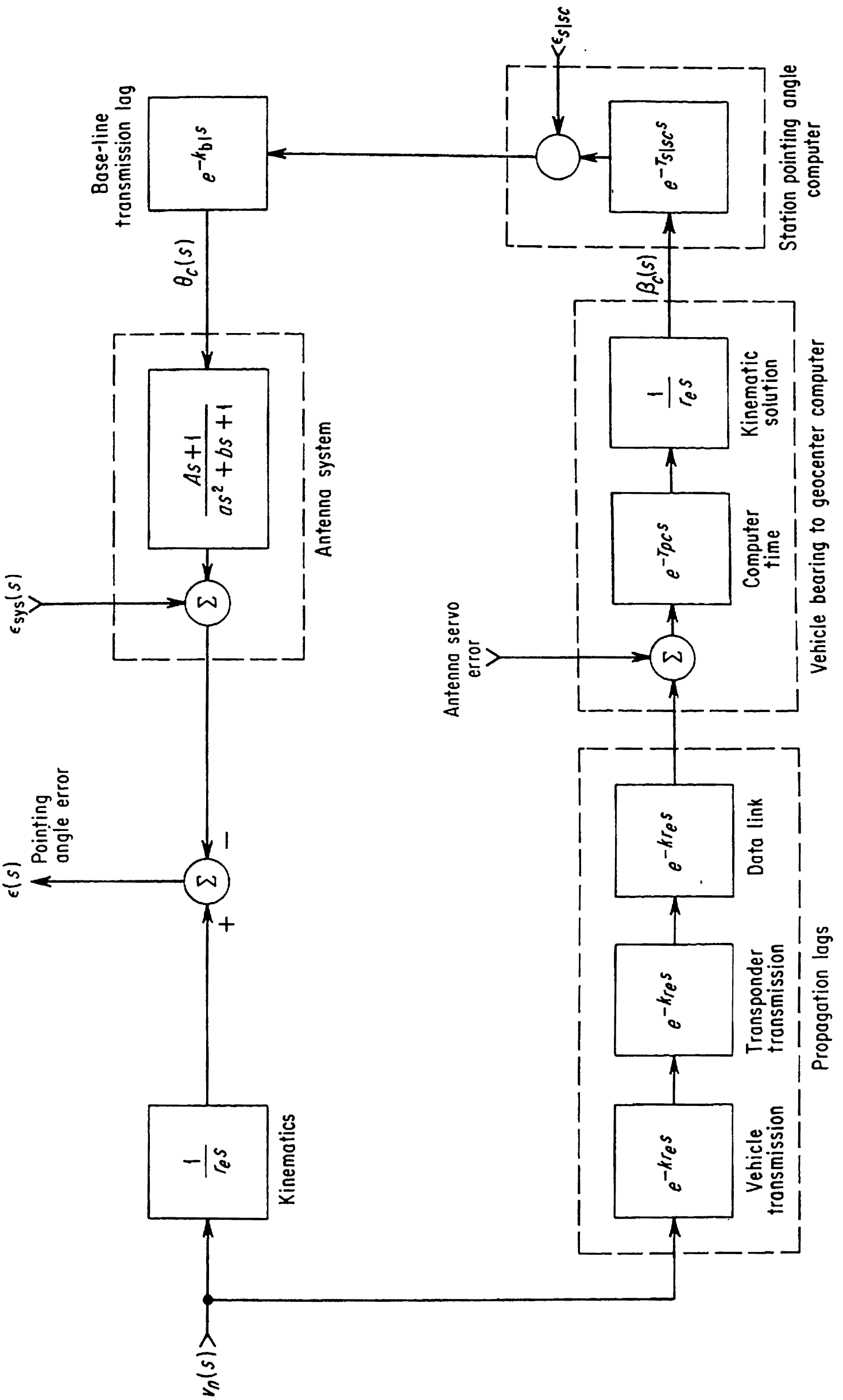


Fig. A13.13 (Linearized) simplified dynamics of antenna-pointing loop command system.

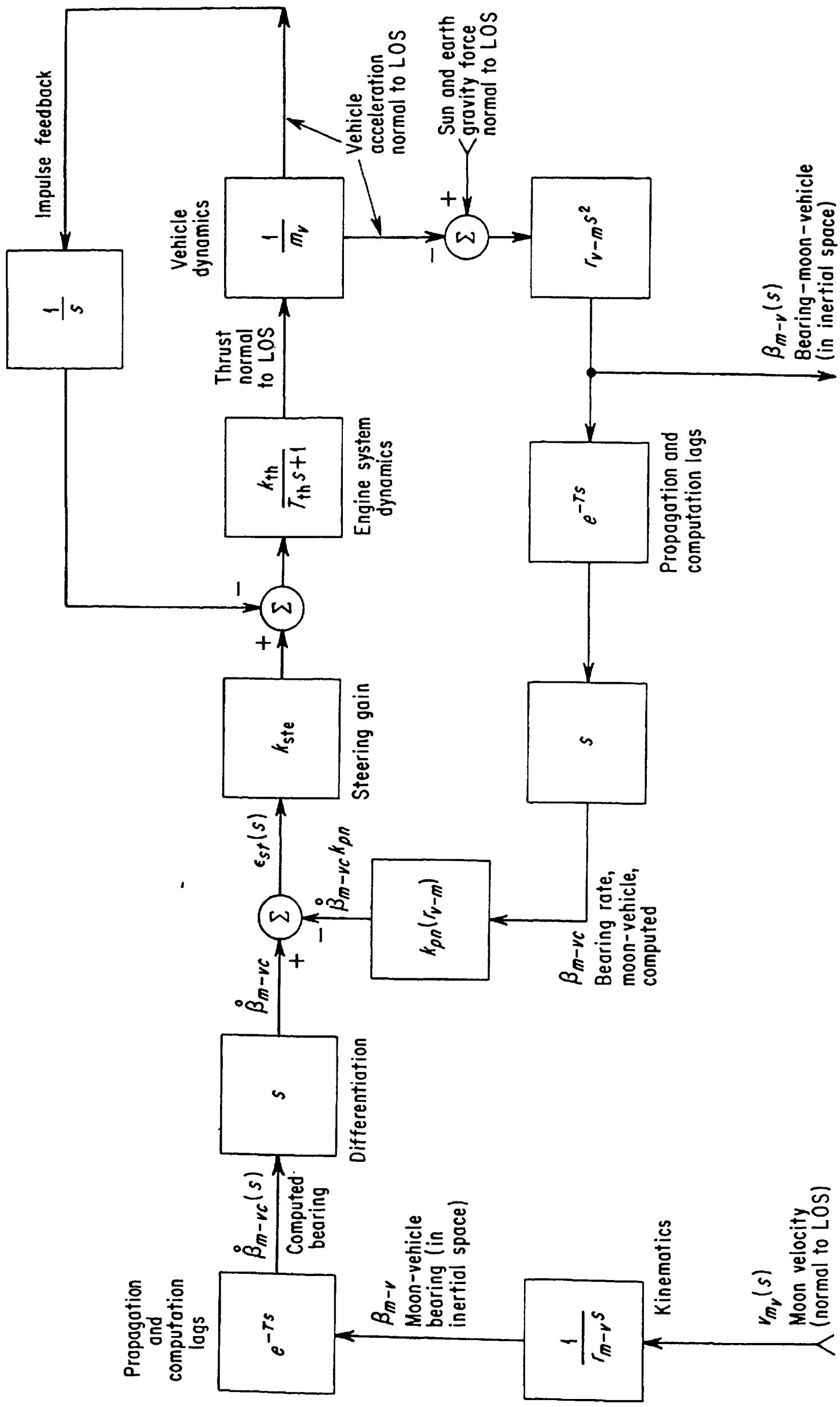


Fig. A13.14 (Linearized) simplified guidance dynamics.

(2) THE DOPPLER SYSTEM.

(a) *Doppler receiver and transmitter system.* A feasible, but not necessarily optimum, system is further delineated in Fig. A13.10, and the foreseeable problem areas are enumerated and a method of attack elaborated where possible. A novel method of precision Doppler spectral frequency tracking and differentiation using digital differential analyzer techniques is presented.

(b) *Doppler transponding system.* Similarly, this system is further delineated in Fig. A13.11. No major problems appear to exist here.

(c) *Doppler antenna system and tracking.* It appears to be possible to effect vehicle tracking without individual station (or group) automatic self tracking capabilities. In a manner described subsequently, vehicle bearing to the geocenter is computed in the position computer; thus, bearing information is operated upon by the earth-vehicle line-of-sight computer, which computes separate pointing angles to each of the three base line stations.

The simplified linearized dynamics of the antenna-directing loop are portrayed in Figs. A13.12, A13.13 and A13.14. Considering attendant transmission lags and pessimistic computation lags of 1 second each and an antenna servo first order time constant of $\frac{1}{2}$ second, the antenna pointing angle/true bearing angle transfer function may be shown to be of the form

$$\alpha_{\text{ant}}^{(s)} = \frac{e^{-Ts}(As + 1)}{(as^2 + bs + 1)} \left[\frac{k}{s^2}\right], \tag{188}$$

where T = total computer and communication lag
k = l.o.s. rate
A, a, b, c = antenna servo system characteristic constants
$k/s^2 = \alpha(s)_{\text{true}}$ = Laplace transform of true vehicle line-of-sight. (189)

Solving Eq. (188) for

$$\alpha_i(t) - \alpha_0(t) = \epsilon_{\text{tr}}(t), \tag{190}$$

it can be shown that the tracking error is in the order of microradians, which is several orders of magnitude below antenna resolution capabilities. Note that the relative dynamic error is independent of the range.

If prediction is required, it could easily be effected in the antenna pointing computer by digital linear (or parabolic) extrapolation.

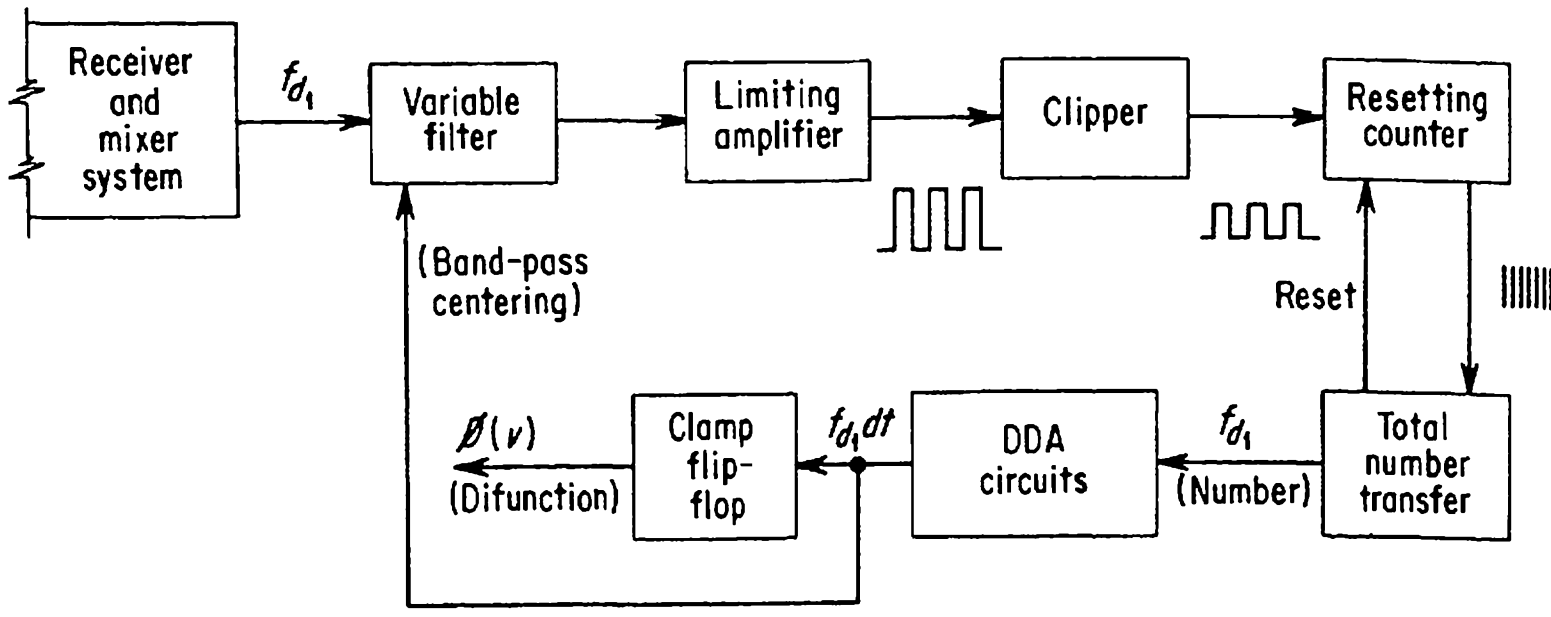


Fig. A13.15 Digital spectral frequency tracker.

Note also that, aside from systematic and computational errors, no "free-floating" dynamic errors exist. If position computation is interrupted, the resumed (and possibly diminished) signals are operated upon by the position computer, and tight tracking is restored.

(d) *Doppler spectral frequency tracker system.* Either c-w or quasi-c-w reception and transmission may be considered. In a c-w system, two spaceborne antennas, one for reception and one for transmission, would be required, and both antennas would be directed. Since achieving adequate pulse repetition rate and filtering seem feasible, the quasi-c-w technique is suggested. Figure A13.15 shows the basic elements which may be contained. An X-band spectrum may be considered.

A digital device which counts the zero crossings of the clock output, and within which a derivative (acceleration) of frequency shift (velocity) can be obtained, is suggested below (see Figs. A13.15 and A13.16). It is estimated that using digital differential analyzer schemes and magnetostrictive delay lines (instead of magnetic drums), an operational accuracy of 10^{-6} is quite feasible. Because of the dynamic limitations of DDA techniques, it is presumed that the frequency modulation of the Doppler shift is not

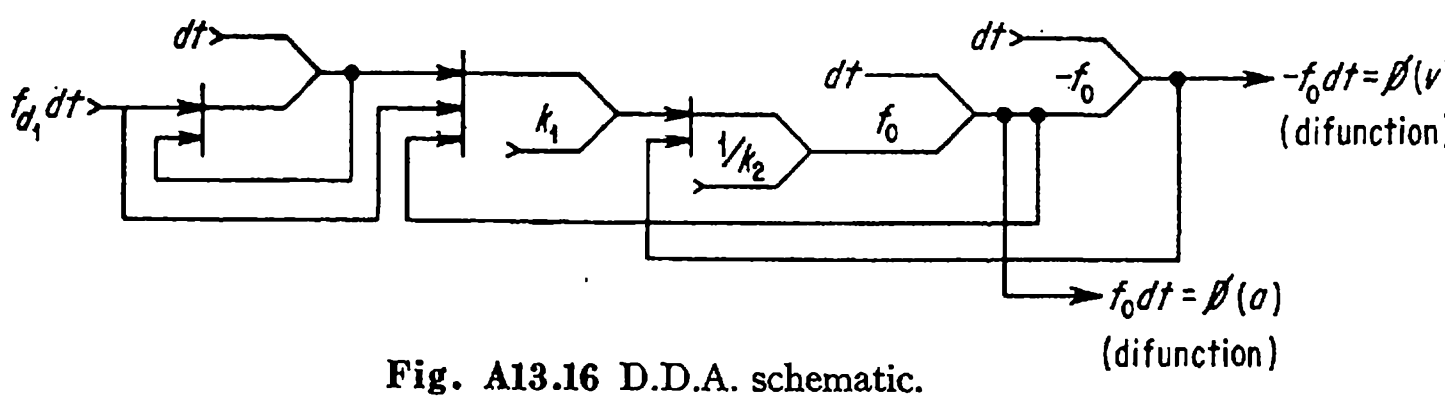


Fig. A13.16 D.D.A. schematic.

high. Certainly the modulations resulting from astrodynamic acceleration profiles, at their worst, will be several orders of magnitude below DDA tracking capabilities.

(e) *Doppler system problem areas.* Several foreseen problem areas, believed to be quite amenable to solution, are enumerated briefly below.

(1) The stability of the reference oscillators must be bounded within some limit, depending upon the required velocity accuracies. In the "reflected" Doppler system, frequency stability of 10^{-8} is more than adequate. However, without sophisticated frequency control, this degree of stability may not be achievable in the vehicle.

(2) The acceptable Doppler receiver signal/noise ratio at extreme ranges may be just within the state-of-art. Receiver sensitivity necessarily must be quite high.

(3) Achieving the required pulse repetition rate (1 megacycle) in a gated power amplifier is barely within the present state-of-the-art; i.e., quasi-c-w, less than $\frac{1}{2}$ microsecond pulse rate at $\frac{1}{4}$ duty cycle will be required.

(4) A lightweight and simple, yet precision, Doppler spectrum frequency tracker is required. Contemporary analog frequency trackers (such as RADAN) comprise an appreciable fraction (37 lbs) of the Doppler electronics (85 lbs), and provide 0.1 per cent operational accuracy.

(3) Position Computer. Having as inputs the vehicle-telemetered Doppler shifts, and transponding station geodetic position and earth rate data, the position, velocity, bearing, and range to the geocenter are computed.

Using digital linear (or parabolic, if necessary) smoothing and extrapolation techniques, transmission lags and earth rate Doppler effects are compensated. The apparent contraction of the base line resulting from ionospheric refraction is also compensated as a function of antenna pointing angles. The incorporation of four base-line stations rather than three may provide more effective resolution of ionospheric effects.

Velocity resolved into the vehicle-moving coordinate system is fed into the target bearing and range deducer.

(4) Target Bearing and Range Deducer. Utilizing position computer-derived bearing and range to the geocenter, Doppler-derived velocity

resolved into vehicle coordinates, telemetered accelerometer readings (when they exist), and moon and earth physical data (with gravitational anomalies), target position is deduced in a manner elaborated previously.

No input to aid in compensation for tellurian or lunar oblateness is indicated for the following reason: the aberration in tellurian gravity force as a function of (equatorial) latitude (declination) and range can be shown to be

$$g'_e(\lambda, r_e) = \bar{j}'_e 0.98 g_e \sin\lambda \cos\lambda + \bar{k}'_e g_e \left(\frac{a}{r_e}\right)^2 \left[i - \left(\frac{a_e}{r_e}\right)^2 \sin^2\lambda\right], \tag{191}$$

where $g'_e(\lambda, R)$ = gravitational force at declination λ and range R,
λ = declination,
g_e = gravitational force at the surface and at the equinoctial,
a = earth radius at the equator,
$\bar{j}'_e, \bar{k}'_e$ = geodetic coordinates,
r_e = range to the geocenter.

At the terminal ranges (r_m) of 40,000 to 1000 miles, the ratio $(a/r_e)^2$ is small indeed; therefore all $\bar{k}'_e$ terms may be dropped or compounded with the $\bar{j}'_e$ terms, and the following assumption may be made for a sufficient degree of accuracy:

$$g'_e = kg_e, \tag{192}$$

where k = correction constant.

It is likely that the accelerometer information will not be used continuously during the flight, except during major midcourse corrections and during powered adaptation maneuvers. A statistical error analysis will reveal the trade-offs between accelerometer bias contamination and utility during mid-course flight.

(5) Vehicle Line-of-Sight Stabilization Computer. This function serves to slave the necessarily remote antenna to the vehicle coordinate reference (inertial platform), and, secondarily, provide crude stabilization during thrust perturbations, operating upon antenna gimbal angle signals and supplementary axial acceleration signals from the attitude control.

The slaving performance will be determined by a trade-off between antenna resolution (which is a function of propagation efficiency, etc.) and the probability of interrupted tracking and difficulties in resuming vehicle location because of diminished received signal.

(6) Stellar Supervised Inertial System. The inertial system consists of lightweight gyros having two or three degrees of freedom, precision accelerometers, and two orientation supervision telescopes. The

gyro torquing computer will resolve quadrature errors from both telescopes, feeding appropriate signals to three (or two) gyroscopes. A search initiation and capture program is indicated in the event of star loss through thrust perturbations and maneuvers, especially during boost-out adaptation phases.

(7) COURSE COMPUTER. This function, operating upon present position and velocity, and nominal and revised objective, computes course, course error, and predicts miss distance.

(8) ENERGY MANAGEMENT COMPUTER. This function analyzes present course error, present velocity, stored target position (then deduced target position at approach ranges), and then revises the trajectory and intercept point to optimize energy minimization and transit time, and insure feasibility. In addition, it will command optimal engine orientation and thrust epochs for both energy optimization and navigation feasibility and error sensitivity. To conserve energy, attitude control stiffness will be relaxed between powered epochs, and appropriate commands will be sent to the stabilization filter.

(9) ATTITUDE CONTROL. The philosophy, techniques, and problem areas of attitude control may be considered a separate subject.†‡§ It is pertinent only to note the system tie-ins and possibilities for usable signal extraction.

D. CONFIGURATION B

Configuration *B* is nearly identical with Configuration *A*, but incorporates an optical disk tracker. The details of such a device are not considered here. Several schemes have been proposed.|| In addition to the problem of discriminating albedo change in a body obscured by shadows and atmosphere, the matter of target oblateness,§ systematic and stochastic components of instrument "noise" limits the accuracy to the order of 10^{-3} in range and bearing, using "averaging" techniques.

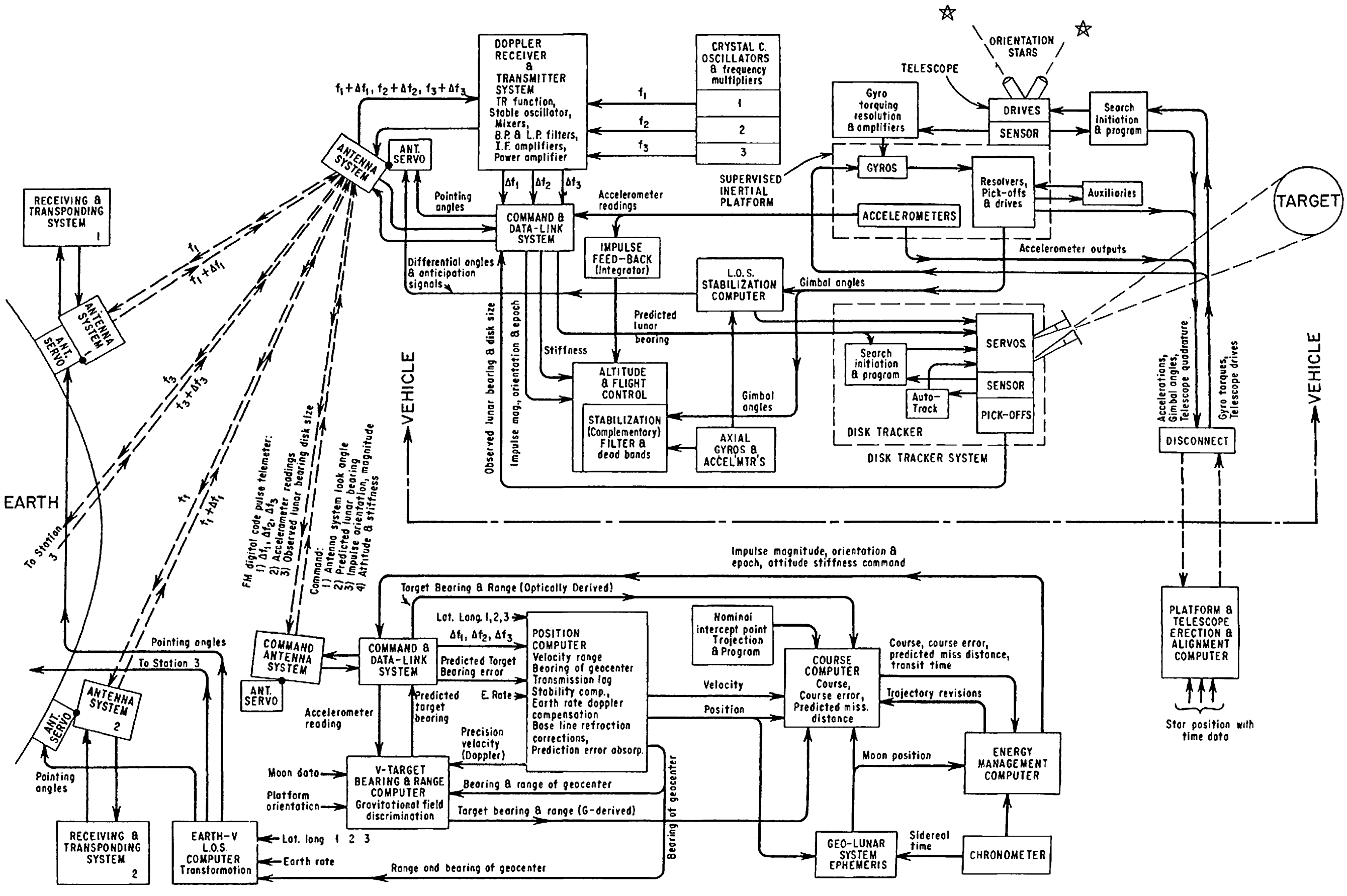
For the lunar operation, this instrument will serve as a redundant

† R. E. Roberson, "A Review of the Current Status of Satellite Attitude Control," San Diego: Space Exploration Regional Meeting, IAS/ARS, August 1958.

‡ ———, "The Problem Areas in Space Navigation," Los Angeles: IAS National Summer Meeting, July 1958.

§ ———, "Optical Determination of Orientation and Position near a Planet," *ARS Presentation*, July 1958, pp. 643–58. *See* Appendix *B*.

|| Lanning, Frey and Trageser, "Preliminary Considerations on the Instrumentation of Photographic Reconnaissance of Mars," *MIT Instrumentation Laboratory Report*, No. R-174 (April 1958).


DOPPLER RECEIVER & TRANSMITTER SYSTEM TR function, Stable oscillator, Mixers, B.P. & L.P. filters, I.F. amplifiers, Power amplifier
CRYSTAL C. OSCILLATORS & frequency multipliers
1
2
3
f_1
f_2
f_3
$f_1+\Delta f_1$, $f_2+\Delta f_2$, $f_3+\Delta f_3$
Δf_1 Δf_2 Δf_3
ANTENNA SYSTEM
ANT. SERVO
RECEIVING & TRANSPONDING SYSTEM 1
Pointing angles
COMMAND & DATA-LINK SYSTEM
Accelerometer readings
IMPULSE FEED-BACK (Integrator)
Differential angles & anticipation signals
L.O.S. STABILIZATION COMPUTER
Gimbal angles
Predicted lunar bearing
Stiffness
ALTITUDE & FLIGHT CONTROL
STABILIZATION (Complementary) FILTER & dead bands
Gimbal angles
AXIAL GYROS & ACCEL'MTR'S
Observed lunar bearing & disk size
Impulse mag., orientation & epoch
VEHICLE
ORIENTATION STARS
TELESCOPE
Gyro torquing resolution & amplifiers
DRIVES
SENSOR
Search initiation & program
SUPERVISED INERTIAL PLATFORM
GYROS
Resolvers, Pick-offs & drives
Auxiliaries
ACCELEROMETERS
Accelerometer outputs
TARGET
SERVOS
SENSOR
PICK-OFFS
Search initiation & program
Auto-Track
DISK TRACKER
DISK TRACKER SYSTEM
Accelerations, Gimbal angles, Telescope quadrature
Gyro torques, Telescope drives
DISCONNECT
VEHICLE
EARTH
To Station 3
FM digital code pulse telemeter:
1) Δf_1, Δf_2, Δf_3
2) Accelerometer readings
3) Observed lunar bearing disk size
Command:
1) Antenna system look angle
2) Predicted lunar bearing
3) Impulse orientation, magnitude
4) Attitude & stiffness
Impulse magnitude, orientation & epoch, attitude stiffness command
Target Bearing & Range (Optically Derived)
Lat. Long. 1,2,3
Δf_1, Δf_2, Δf_3
Predicted Target Bearing error
E. Rate
Predicted target bearing
Precision velocity (Doppler)
Accelerometer reading
COMMAND ANTENNA SYSTEM
ANT. SERVO
COMMAND & DATA-LINK SYSTEM
POSITION COMPUTER
Velocity range
Bearing of geocenter
Transmission lag
Stability comp.,
Earth rate doppler compensation
Base line refraction corrections,
Prediction error absorp.
Nominal intercept point Trajection & Program
COURSE COMPUTER
Course,
Course error,
Predicted miss. distance
Course, course error, predicted miss distance, transit time
Trajectory revisions
Velocity
Position
PLATFORM & TELESCOPE ERECTION & ALIGNMENT COMPUTER
Star position with time data
Pointing angles
To Station 3
ANTENNA SYSTEM 2
ANT. SERVO
Pointing angles
Moon data
Platform orientation
V-TARGET BEARING & RANGE COMPUTER
Gravitational field discrimination
Bearing & range of geocenter
Target bearing & range (G-derived)
Bearing of geocenter
Moon position
ENERGY MANAGEMENT COMPUTER
RECEIVING & TRANSPONDING SYSTEM 2
EARTH-V L.O.S. COMPUTER Transformation
Lat. long 1 2 3
Earth rate
Range and bearing of geocenter
GEO-LUNAR SYSTEM EPHEMERIS
Sidereal time
CHRONOMETER

terminal referencing source providing range, bearing, and the vertical in the event of blanking out of Doppler reception and transmission from the earth base line during the circumlunar passage.

In addition to reliability considerations (providing a redundant information source), the concept of error minimization by comparing reference point (target or other) prediction and observed reference point position, and then analyzing and operating upon the discrepancy to correct the systematic errors, is made possible. Statistically, it can be shown that considerable error minimization is possible with such a feedback complementary configuration.

After tellurian boost-out, or at some point in cislunar space, the disk tracker may be slewed to the expected moon bearing which is deduced by the gravity field vector discrimination computer, a search and lock-on procedure initiated, and automatic tracking continued through the circumlunar phase.

E. Configuration C.

This configuration has been presented in the event that trade-offs in propulsion economics, reliability considerations, and communications problems have been resolved, and these indicate that a spaceborne system of such computational complexity is tolerable or favorable. A comprehensive study program in operations research, operations and systems analysis, and engineering feasibility would be required.

Note that for the duration of a typical lunar transit, a crystal-controlled reference oscillator, with some modification, can serve a dual function as a chronometer to assist the energy management computer in optimum impulse epoch judgment and control. The nominal program is consulted and revised accordingly.

In general, the functions are identical with those of Configurations *A* and *B*. The load and accuracy demands upon the data link are greatly reduced, however, and the requirement for precision telemetry of the most important information of all, the vehicle-sensed Doppler shifts, is eliminated. The telemetered base-line antenna pointing angle need not be known to extreme precision.

However, for obvious reasons, it would seem desirable to telemeter vehicle-sensed Doppler shifts in order to provide supplemental vehicle location. In conclusion, a *fundamental* and *extensive* analysis will be required to resolve the question of the *sharing* of spaceborne and base line guidance functions.

5. Interplanetary Guidance

Because an adequate base line in general will not be available in remote heliocentric space, reliance upon Doppler-derived velocity and range information (by the matching of atomic clocks) will be required. Until the feasibility and accuracy of observing Doppler shifts in the reference body is established, additional information from an optical transit and/or solar disk tracker, and supplemental information from celestial emphemerides will be required.

The theory of observation requirements to determine position has been discussed to some extent.†‡ Several combinations of reference bodies are possible,† the optimum choice of combinations are subjects for evaluation and are not pertinent to this note. Although position determination is possible without ephemerides (or ephemerides plus chronometer), error analyses† indicate that an ephemeris and chronometer should be used.

Pertinent considerations are the additional utilization of the Doppler system's reference oscillator as a chronometer, providing accuracy which is quite sufficient. Of all the various choices of planet-sun-star combinations available, the selection of the earth and two stars during tellurian hyperbolic escape, one star and the sun for midcourse guidance, and two stars and the sun for terminal guidance, may be considered.

Many possibilities exist. Using the gravity field vector deduction scheme, the inertial platform orientation may be slaved to two stars; the range and bearing to the earth can be computed by the use of three base-line stations, and the range and bearing to the sun deduced, or the bearing and range to the sun could be obtained, from a disk tracker and the bearing and range to the earth deduced.

The aforementioned sun or earth position deduction scheme will permit slewing of the disk tracker to the expected position, and complementary signal filtering after lock-on. Because of this, and the initial precision Doppler base-line information, heliocentric elliptical flight may be started on a much more favorable basis. Base-line information and gravity field deduction may be useful to a range of 500,000 miles. Bearing information to an accuracy of 10^{-4} may be obtained at a range of 20×10^6 miles.

† L. Laramore, "Celestial Observations for Space Navigation," Los Angeles: ARS Summer Meeting, 1958.

‡ R. W. Wheelon, "Position Determination for Midcourse Guidance," Santa Barbara: Inst. of Nav., 14th annual meeting.

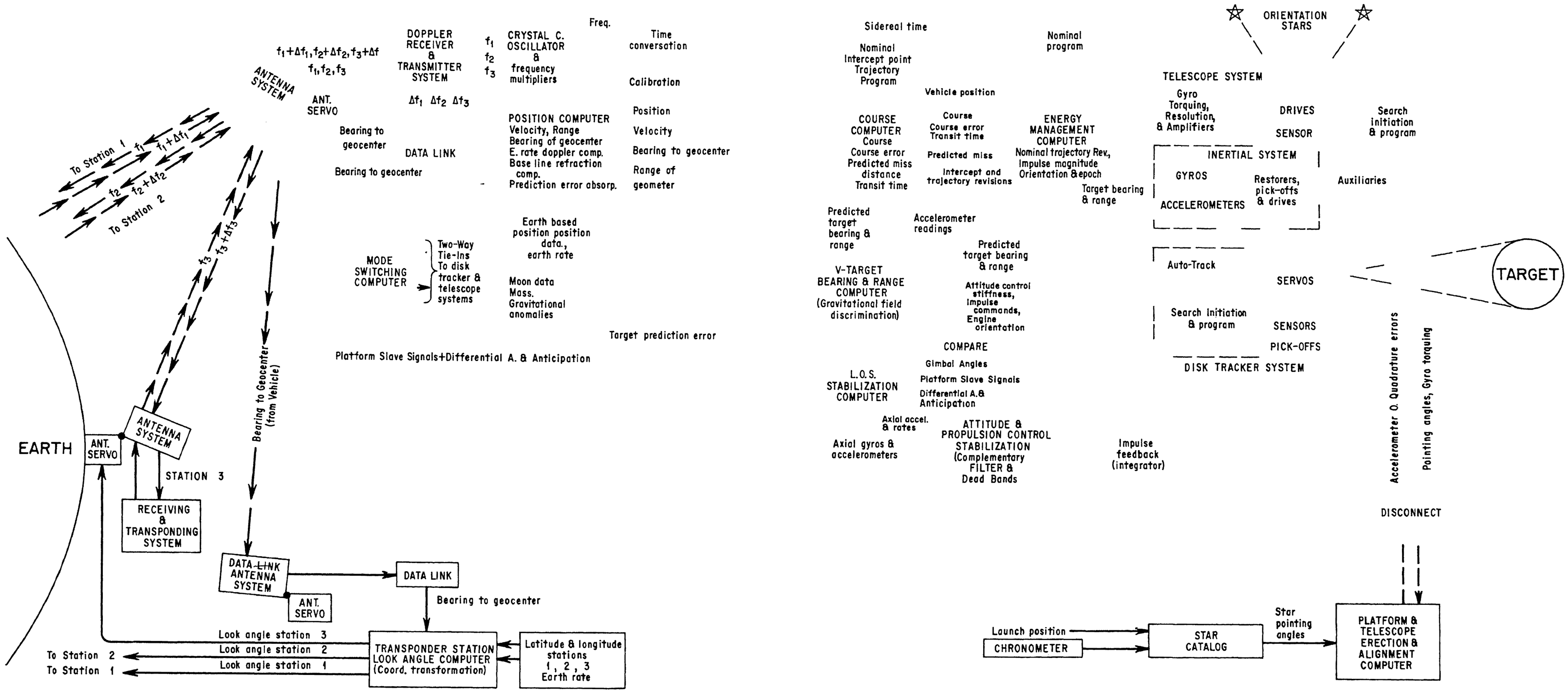
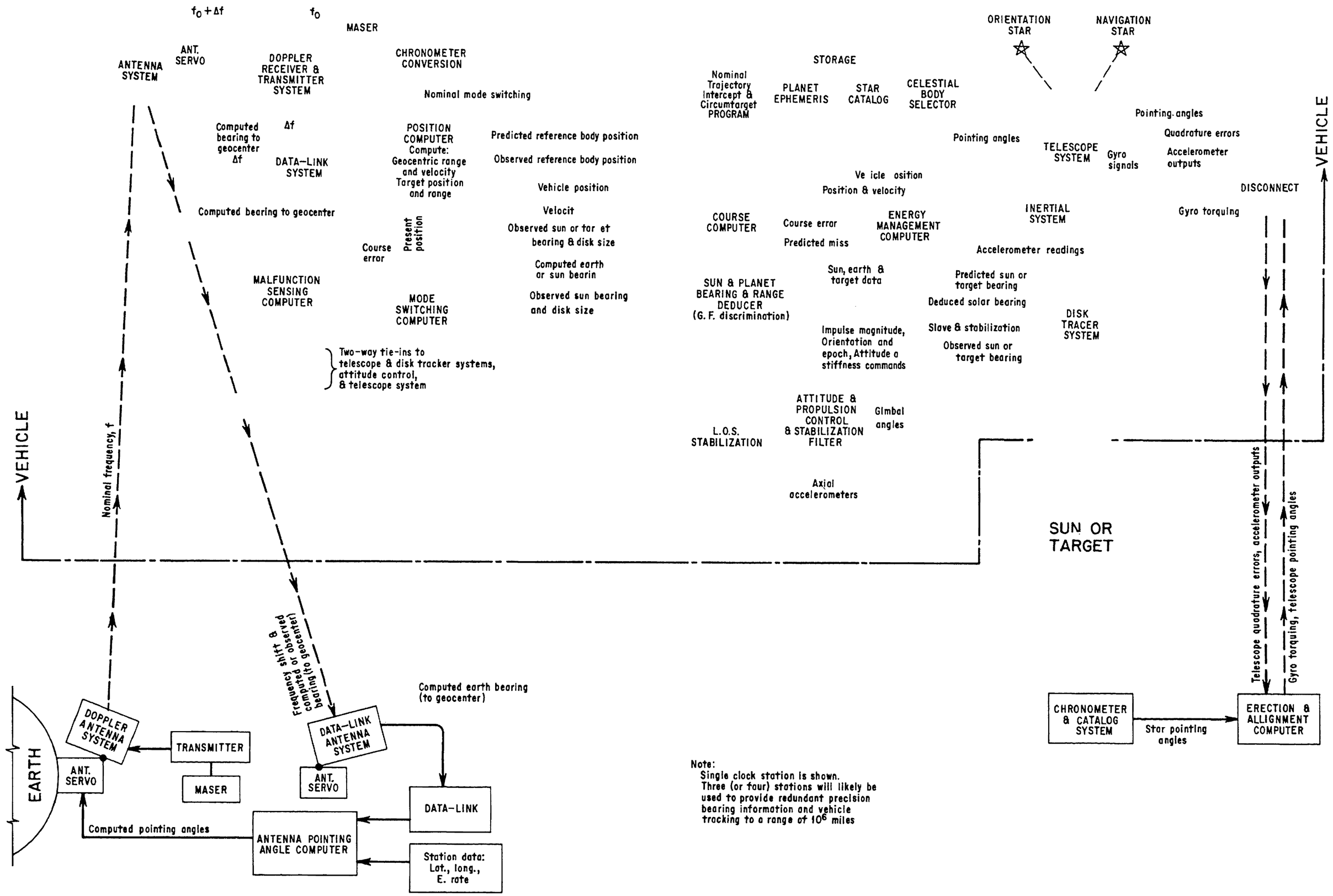


Fig. A13.18 Configuration C, lunar guidance self-contained.



Note:
Single clock station is shown.
Three (or four) stations will likely be used to provide redundant precision bearing information and vehicle tracking to a range of 10^6 miles

Fig. A13.19 Configuration D, interplanetary guidance.

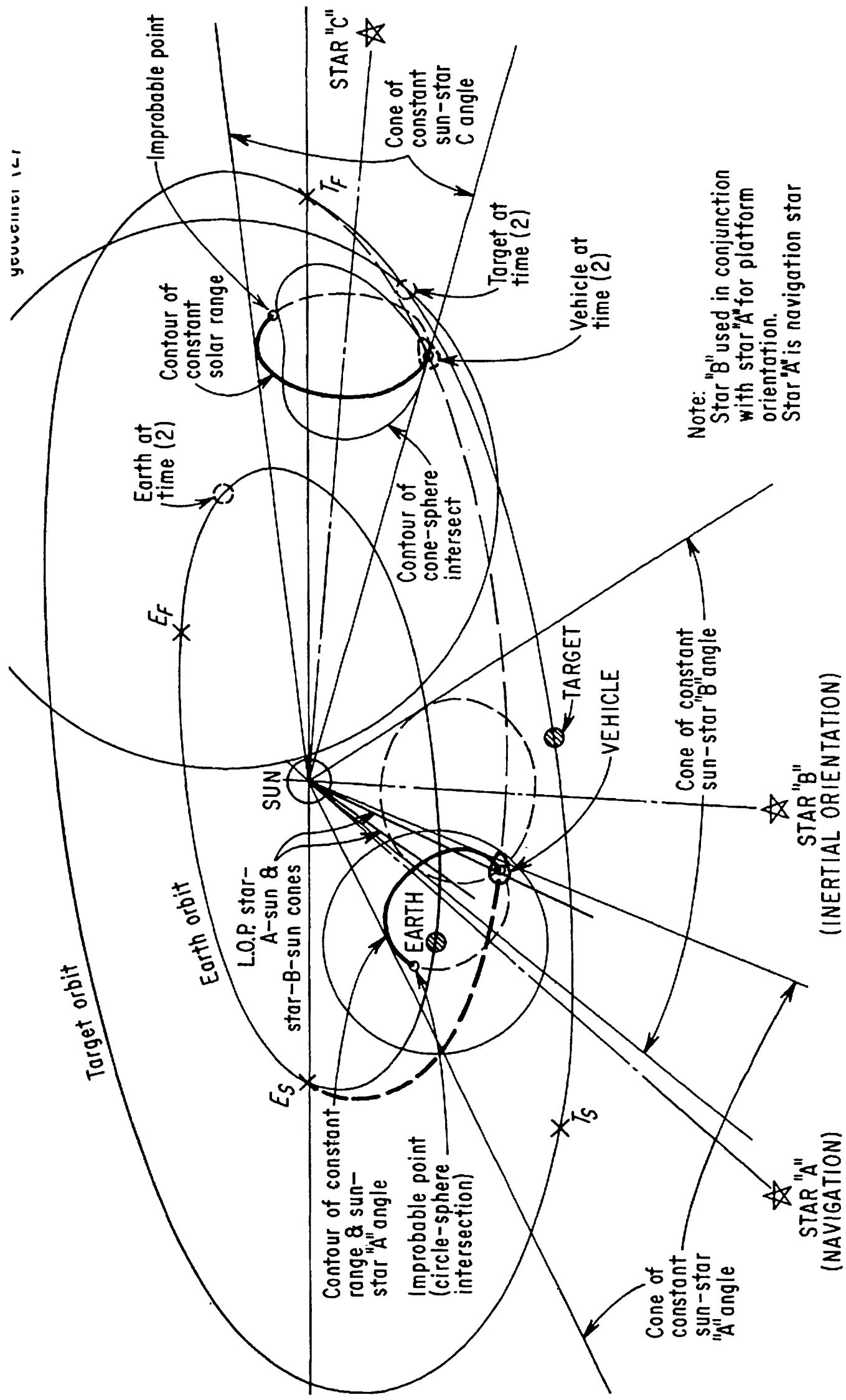


Fig. A13.20 Geometrics of celestial-Doppler astronavigation—interplanetary.

In midcourse (heliocentric) flight, the vehicle will be influenced (in a significant fashion) by the solar field only. Usually, position determination must be accomplished solely by optical transits. However, the advent of precision range information achievable by the clock matching scheme relieves the number of reference body pairs that would otherwise be required; i.e., either combinations of

(1) (a) sun-star
(b) sun-star
(c) sun diameter

(2) (a) sun-star
(b) sun-star
(c) planet-star + time + ephemeris

(3) (a) sun-star
(b) sun-star
(c) planet-star
(d) planet-star } + ephemeris only

The availability of precision range to the geocenter will reduce the number of optical observations from three or four to two, as listed:

(a) sun-star,
(b) sun diameter.

The simplification is illustrated in Fig. A13.20. The sphere of range to the geocenter is intersected by the cone surface of sun-star position which will contain the contour of sun-star-solar range position, which will intersect the sphere of range position at two points. With an approximate knowledge of position by dead reckoning, one of the points may be discarded as impossible. (Star selection and switching as functions of position will be required to ascertain intersection of the cone and sphere surfaces.)

It is a matter of interest to note that in the target activity sphere, target position could be deduced if one additional source of velocity, say from another star radio or visible spectrum, could be obtained, consistent with the general theory of gravitational field discrimination presented earlier. Consider Fig. A13.21.

With the normal components of the tellurian, stellar, and solar accelerations derived from temporally correlated spectral observations and disk tracker, respectively, and with a knowledge of the masses of the (influencing) bodies, the three-dimensional vector bearing, and range to the target can be deduced. This technique may be especially valuable in the case

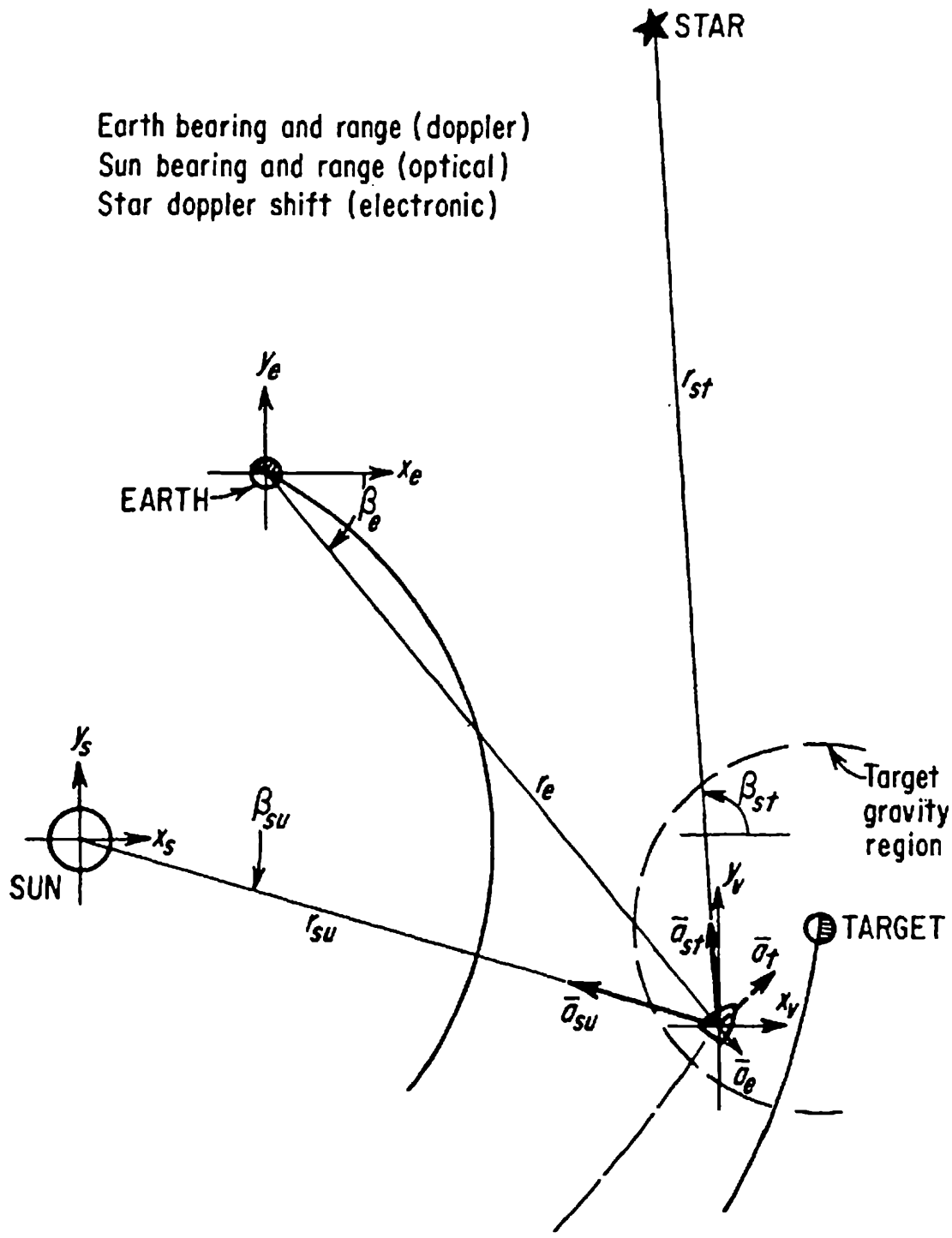


Fig. A13.21 Target-bearing and range deduction.

where the target is obscured by dense atmosphere or shadows. A solar disk tracker, which will not contend with serious body oblateness, may, but utilizing stochastic processes, be expected to have an accuracy in the order of 10^{-4} radians.

Two philosophies of telescope system configurations may be considered: automatic tracking (transit) or bearing prediction. In Configuration *D*, the latter concept is indicated. The planet ephemeris or stellar almanac is consulted on the basis of *indicated* present position, and expected bearing angles are extracted. The telescope-sensed discrepancy will inform the position computer of its error in a continuous feedback process. Some proposed systems have only one optical system which performs both the function of disk tracker and transit.† The evaluation of this technique is not pertinent to this appendix.

† "Preliminary Considerations on the Instrumentation of a Photographic Reconnaissance of Mars," *MIT Laboratory Report*, No. R-174 (April 1958).

Addendum:

Limitations of Contemporary Terrestrial Doppler Navigation Radars

1. Limitations on Existing Navigation Radars

The accuracy of existing navigation systems based upon the use of the Doppler shift of a microwave radar signal is a function of several variables. Among these are:

(1) Topography,
(2) Transmitter stability (both short- and long-time stability),
(3) Local oscillator stability,
(4) Antenna propagation pattern,
(5) Frequency tracker accuracy,
(6) Signal-to-noise ratio,
(7) Heading reference.

(1) TOPOGRAPHY. The errors caused by topography result from the way in which the radar return is modified by the illuminated scatterers. These scatterers are randomly placed within the antenna illumination pattern and vary randomly in size between scatterers. Consequently, the amplitude of the radar return varies randomly in time within the illuminated ground

path, and as a direct result, the Doppler spectrum of the return signal is "noisy." It must be smoothed for a period of time from one to ten seconds in the usual aircraft installation if the standard deviation of the observed spectrum is to be kept to within 10^{-4} of the mean Doppler frequency.

Even after considerable smoothing, there still remain uncertainties owing to topography. In the first place, the velocity determined over a body of water is relative to the surface of the water. Naturally, then, surface currents place undesired velocity biases in the measured aircraft velocity. A second effect of equal importance which is associated with travel over bodies of water is the skewing of the Doppler spectrum. This effect may be as large as 10^{-1} of the mean Doppler frequency and is very difficult to compensate for in the design of the system. The spectrum skewness (in the non-Gaussian sense) is a function of three parameters: (1) the mean "look angle" of the radar antenna; (2) the beam width of the antenna propagation pattern; and (3) the Beaufort number (surface roughness) of the sea condition. Assuming the ability to compensate entirely for causes (1) and (2), it still is doubtful whether the effects resulting from (3) can be entirely eradicated. In present Doppler designs, the total error arising from the three causes above is less than 10^{-3} of the mean Doppler frequency.

It should be noted that the implementation of the Doppler effect proposed herein for the astronautic system does not suffer from any of the effects above since there will be no reflected signal.

(2) Transmitter Stability. The long-time stability of the transmitter and local oscillator affect the Doppler frequency measurement as a percentage error of the "true" Doppler frequency. If the stability is held to 10^{-3} of the nominal radar frequency, the velocity error likewise will be 10^{-3} of the "true" velocity. It should be noted carefully that the astronautic system described here differs from existing systems quite markedly on this point.

The short-time stability of the transmitter and local oscillator contribute to the deterioration of the signal-to-noise ratio of the system, but do not otherwise affect system performance.

(3) Antenna Pattern. The antenna propagation pattern contributes to the system error in three ways. As noted above, a finite beam width leads to a skewing of the Doppler spectrum over bodies of water. In addition, a finite beam width implies a finite time of scatterer illumination. Consequently, a given scatterer yields a Doppler return which has a finite duration in time. Assuming that the return is a sinusoid of fixed frequency

during the time of illumination (τ), it follows that, with respect to energy, it covers a frequency spectrum inversely proportional to the time of illumination. Consequently, the Doppler return signal never can yield a spectrum narrower than $1/\tau$.

The third way in which the antenna propagation pattern contributes to system error is by its lack of symmetry. The shape of the spectrum of the Doppler return literally duplicates the shape of the two-way pattern of the antenna, barring distortions caused by the topography of the earth. Consequently, any first order asymmetry of the antenna propagation pattern will result in a first order Doppler error which, of course, may be removed by equipment calibration if it remains constant.

(4) Frequency Tracker. The frequency tracker contributes errors which are a function of its particular mechanization. In existing equipments, the total error arising from the tracker mechanization represents very nearly the total velocity error of the system. This error usually is approximately 10^{-3} of the true velocity.

(5) Signal/Noise Ratio. The signal-to-noise ratio seldom is a limiting factor in the design of Doppler navigation systems. If it is 10 or better, there is no degradation of system performance, provided the spectrum tracker is designed to track the center of the Doppler spectrum and to exclude all other regions of the energy spectrum.

(6) Heading Reference. At present, the heading reference is the limitation on system accuracy. Existing references are reliable only to about 4×10^{-3} radians. It is unlikely that this error can be reduced to less than 10^{-4} radians.

2. Summary

The theoretical limitation on the performance of existing Doppler systems is the unknown error contribution of the topography over which navigation must take place. Present equipment reduces this error to negligible proportions (10^{-4} or less) over land, but over bodies of water system performance will be degraded in proportion to the ratio of unknown surface currents to the aircraft velocity.

Appendix 14

Applications of Electromyographic Techniques in the Integration of Man-Machine Systems

"...for scientific management has for its object just what labor-saving machinery has for its object, increased output per unit of human effort..."

—FREDERICK WINSLOW TAYLOR

1. Introduction

The objectives of this appendix will be to suggest the applications of electromyography (EMG) techniques and concepts to that portion of an illustrative man-machine system primarily concerned with the flow of information from the man to the machine. In addition, it appears that an unexplored possible communication channel exists from the machine to the man in the form of direct motor point stimulation of the man by the machine.

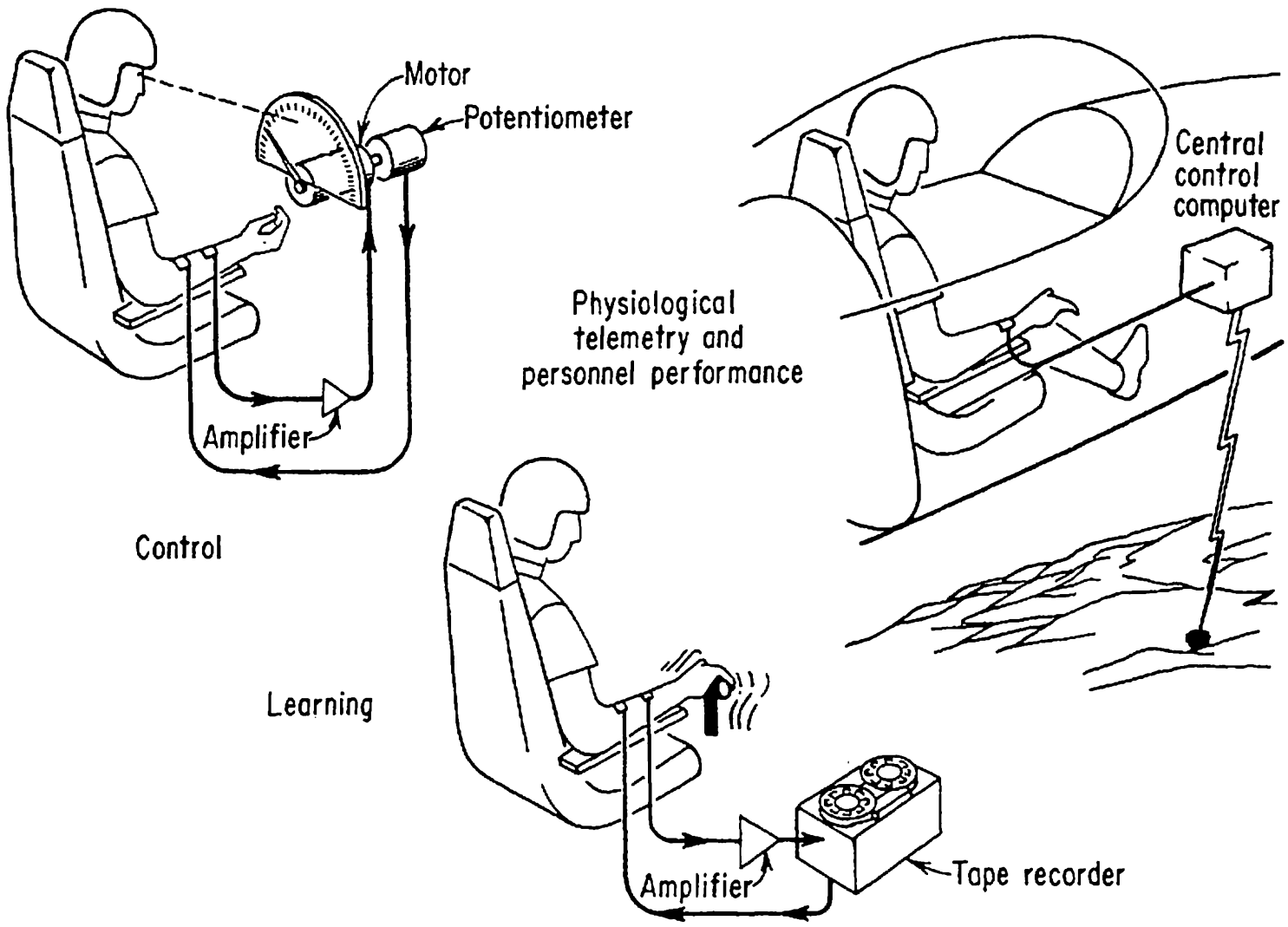


Fig. A14.1 Applications of electromyography.

Three general areas that are evident from requirements studies for such a system are the following:

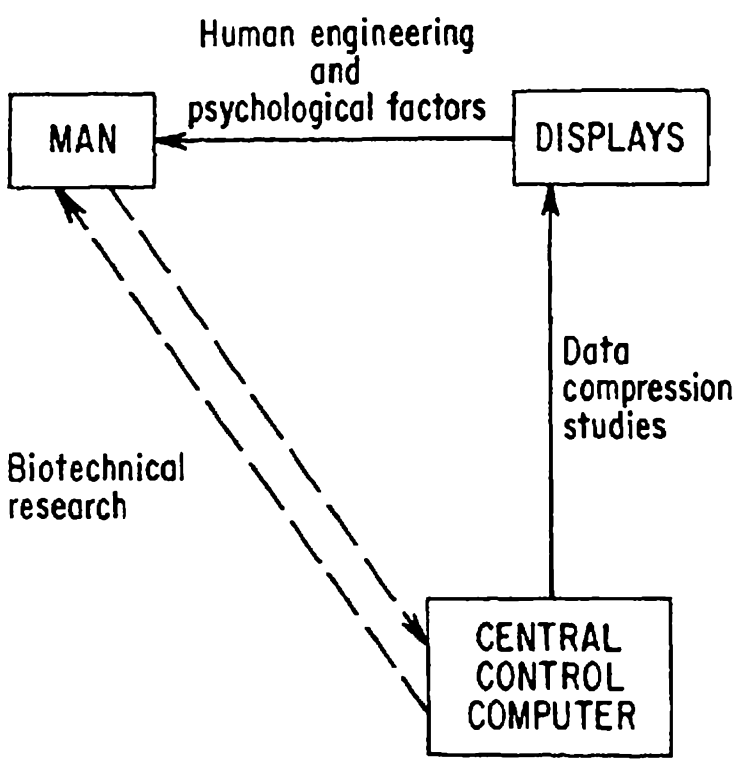


Fig. A14.2 Elements of man-machine systems.

(1) The central control computer concept (unified function approach);

(2) Adequate displays to provide input information to the man for
   (a) Present status,
   (b) Attention-compelling,
   (c) Decision-making;

(3) Proper forms of controls to provide for the man
   (a) Quickening and unburdening,
   (b) Override capability.

Figure A14.1 graphically suggests systems applications within these areas; the relationships among the several areas are as shown in Fig. A14.2.

2. Preliminary Discussion

The different theoretical aspects of EMG control applications can be subdivided into a group of distinct categories.

(1) MAN TO MACHINE. The electrical signals produced in the muscles are used to regulate and control machine performance. Two machine categories are to be considered:

(a) Motors and actuators,
(b) Recording devices.

(2) MACHINE TO MAN. Another communication approach explores a hitherto unexploited input channel to the man. Muscle motor points are artificially stimulated in an effort to induce predictable contraction of muscle tissue from external devices such as:

(a) Transducers and computers,
(b) Playback recordings.

It is possible to induce smooth, predictable contractions of muscle tissue through the use of specifically designed surface electrodes in combinations with phase, frequency, and amplitude-modulated signals.

(3) CLOSED LOOP. A combination of (1) and (2) should make possible a closed loop system with command functions originating in the human brain, electrical signals produced in the muscles fed directly to the machine, and feedback from the machine to the man's muscle. This loop is shown in Fig. A14.3.

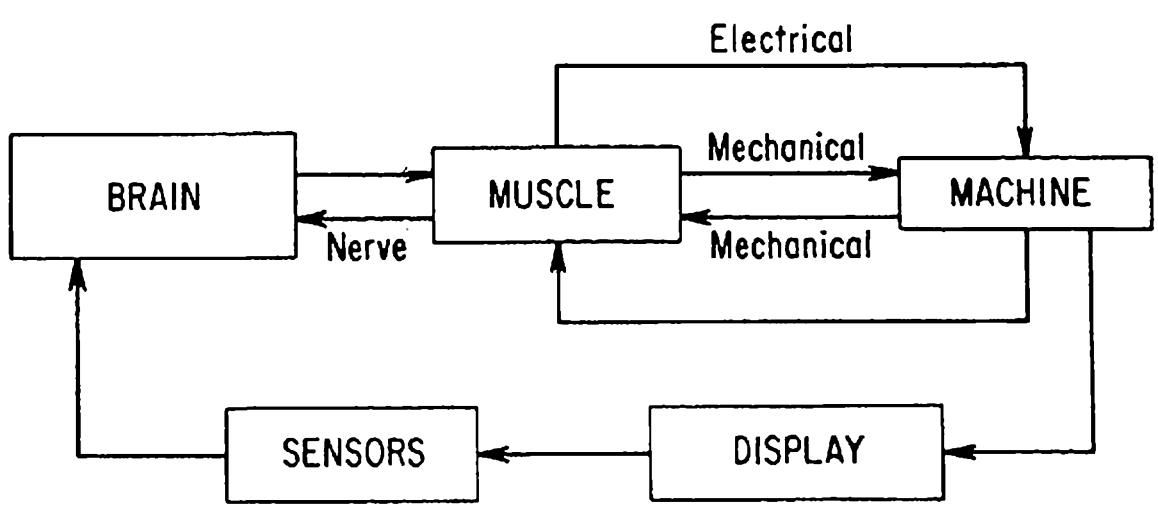


Fig. A14.3 Man-machine loop.

One interesting aspect of the electrical loops going from the man to the machine and returning is that they represent a variable gain power boost system much like power steering in an automobile.

(4) OTHER APPLICATIONS. Two other applications, besides control, are apparent.

(a) *Minimizing trial and error in the acquisition of motor skills.* It is possible that a higher rate of acquiring skills in simplex and/or complex muscular actions (performance) may be provided by a repeated playback from a recorder and the additional necessary equipment into the muscles or groups of muscles involved. Proprioception provides the link to the brain whereby the repeated performance of the action is possibly developed into an acquired skill.

(b) *Physiological telemetry.* Human muscular performance under adverse conditions can be telemetered to a ground station where human performance and machine performance can be correlated. In order to avoid confusion concerning the type of muscular movements, there would be multiplex information with the electromyography to measure or monitor dynametric tensions developed, heat produced, and movements of dials, levers, or switches. The information would be gathered from both the man and machine.

Some background material concerning the above-mentioned clinical aspects of the physiological phenomena is next presented in this discussion.

3. Physiological Aspects of Neuromuscular Activity

The primary action of muscles is contraction. A contracting muscle shortens and pulls. This action is the basis of animal movement and motion. When a muscle is effectively stimulated, a number of very rapid transformations take place. In the muscle, electrical, structural, chemical, and thermal changes occur. To the best of present knowledge, the initiation and propagation of an electrical response (the muscle action potential) constitutes the first event or change following excitation of a muscle.[1,4] In the past the main use of electromyography (the measurement of electrical action potential) has been for clinical or medical purposes. It has been recently demonstrated that the experimental techniques for recording electrical responses of muscles can be utilized for quantitative measurements of the contraction of muscles.[7] This electrical activity of the muscle

occurs prior to the muscle actually shortening (isotonic contraction), or prior to the muscle just developing tension (isometric contraction).

(1) CLINICAL MEASUREMENT TECHNIQUES. Two general methods have been employed in making these measurements, which differ only in the type of information obtained.[9] The first method utilizes needle electrodes, one of which is thrust into the thick part of the muscle. Two other electrodes are placed close by in a neutral position. The second method uses a larger electrode placed on the surface of the skin overlying the appropriate muscle. Since both methods are in general use, each enjoys particular advantages and conversely suffers particular disadvantages. It is desirable to elaborate each method so that a choice of one of these can be justified for the purposes of this proposal.

(a) *Method I: Needle Electrodes.* A needle electrode, insulated in such a manner that only a microscopic portion is exposed, permits measurements to be made of the responses of motor units—a motor unit is a nerve fiber and all connecting muscle fibers. This very localized observation permits clinical evaluation of possible neuromuscular dysfunction. The information obtained on the cellular level is interpreted in terms of discharge rate, discharge amplitude, and time events (asynchronous or synchronous). It is also possible to compare the activity of several parts of a single muscle by using multiple needles. Since these electrodes have to come in actual contact with the muscle fiber or fibers, the skin must be penetrated. Necessary adjuncts to this method are the proper safeguards, such as sterile and sharp needles, clean skin, etc., to be taken against infection. The procedure is usually not painful unless a nerve or nerve trunk is punctured. If muscle activity is to be measured while the muscle is functioning, the deformation of the skin overriding the muscle will be felt and perceived as discomfort, or possibly pain.

(b) *Method II: Surface Electrodes.* If the second method, which involves the use of surface electrodes, is adapted, the skin is not penetrated. This method, therefore, permits long-time and multiple observations. The high resistance imposed by the skin is usually lowered by the application of mineralized pastes like Bentonite or other electrode pastes used in electrocardiography. The responses obtained with surface electrodes are in general smaller because of resistances imposed by the nonactive tissue between source and pickup points. The electrical phenomena is the combined potential of many individual fibers and motor units. Surface electrode pickup is particularly useful for measurements of muscle

electrical responses produced by percutaneous stimulation of the nerve. The magnitude of response, in microvolts, is determined, in part, by the number of activated units.

The total electrical activity from a voluntary contraction of any given muscle can be measured. It has been found, particularly under isometric conditions, with skin surface electrodes, that the tension developed (with a time delay of tenths of milliseconds) is proportional to the average electrical activity in that muscle.[7] For this reason the surface type electrodes better serve the purpose of suggested applications.

(2) EQUIPMENT REQUIREMENTS FOR ELECTROMYOGRAPHIC MEASUREMENTS. To record muscular action potentials, it is standard operating procedure to have apparatus that will (1) detect, (2) amplify these voltages from suitable electrodes, (3) display these voltages on a cathode ray tube for immediate viewing and possible photographing, (4) make these electrical potentials audible through a loud speaker, and (5) record them on magnetic tape for future use.[2,8]

Suitable materials for use as surface electrodes are silver or other inert materials. These are attached to coaxial flexible cables. For clinical use they should be small (1 cm in diameter) and fairly flexible. Some provision should be made to permit movement of the muscle or limb, so that no undue restriction is placed on the muscle being investigated. Of course, during such movement of the muscle, the surface electrodes are to remain in contact with the skin.

There are three such electrodes which may be used for this purpose: (1) an active electrode, (2) a reference electrode, and (3) a ground electrode.[8] The active and reference electrodes constitute the recording, unipolar, or monopolar electrodes. The third electrode is the ground in common with the subject's limb and is led into the apparatus appropriately. The ground in common eliminates the pickup of extraneous and sixty-cycle potentials.

(a) *Detecting and stimulating electrodes for biotechnical systems.* In most biotechnical systems, electrodes are used to bridge the gap between organism and electronic equipment. Used in this manner, they are usually the weakest link in the system. Size, construction, method of attachment, chemical, and electrical parameters are all factors contributing heavily toward success or failure of the detection.

There are two basic types of electrodes, surface and percutaneous, in use today. In general, the surface electrode takes the form of a flat disc

or rectangle of thin metal which is attached to the skin by tape, collodion, or wide strips of rubber. The percutaneous (more commonly called needle electrode) usually consists of a single, very sharp needle which is inserted at a sharp angle slightly beneath the outer layer of skin. This type is primarily used in electroencephalography although it has found application recently in electrocardiography. A modification of the needle electrode is also used in electromyography, in which case, the needle assumes a coaxial form and is of much greater diameter and length than the electroencephalographic type. In the application suggested, only surface electrodes would be used.

When a metallic conductor is placed adjacent to living tissue, interaction between the two, called ionic polarization, may take place. The amount of polarization is directly related to the position of the metal in the chemical activity series. Hence, copper and zinc are poor materials to use for electrodes. Moreover, if a potential difference is impressed between two such electrodes, the polarization will be even more intense.[6]

Many researchers disagree as to the exact nature of this polarization. It is believed that a half cell is formed between tissue fluid and the electrode itself.

There are two basic methods to counteract polarization. The most widely used is that of establishing a chemical boundary between the electrode and the tissue under survey. A metal and one of its salts, (in solution or paste form), such as zinc and zinc-sulphate, or silver and silver chloride, are often used to effect this chemical state. The combination of a metal and its salt forms an electrolytic junction between the tissue and its saline solution. Unfortunately, the chemical (particularly the zinc, zinc-sulphate combination) is injurious to living tissue. The method also suffers from being limited in the length of time it may be used, owing to drying of the salt solution and a resulting increase in polarization, contact resistance, and motion artifacts.[6]

An improved method, which seems to have many benefits of the metal, metal-salt combination and none of its drawbacks, is merely to select one of the less active metals for the electrode. Such metals as gold, platinum, and tantalum have proven to be quite satisfactory when used with a suitable electrode jelly such as EKG-SOL or SANBORN REDUX.

The problem of polarization is of minor importance in this project since EMG signals range from 50 cps to over 1,000 cps. It does become important when investigation centers in the region between d-c and several cps.

As might be expected, the ability of the needle to sense information from

tiny areas to the exclusion of other signals is very great. This ability, termed "selectivity," apparently is determined by the size of the electrode in relation to the area being investigated and its relationship geometrically to that area, that is, the proximity of the electrode to the signal source.

Since, in this proposed application, surface electrodes will be the sole type used, one aim will be to achieve a selectivity with the surface electrode which approaches that of the needle electrode.

It should be noted that the surface area of the electrode is *effectively that area of the skin underneath the electrode which is covered by electrode paste.* Therefore, when high selectivity is of importance, it is well to keep the electrode small and the area covered by paste directly under the electrode. If gross activity is of importance, it is possible to make the electrode as large as a square inch or more. Signal strength or output from the muscles being investigated is inversely related to electrode size. Therefore, for a highly selective system, the requirements are for small electrodes with relatively high amplification. However, for gross EMG-sensing, larger electrodes must be used, and consequently less amplification is required. In producing difficulty, the method of fastening and possible resulting motion artifacts are highly interrelated. It is extremely important that the method of fastening the electrode does not produce any excess tension in the muscle. For example, some experimenters have attempted to utilize pediatric chest cardiographic electrodes of the suction cup type. At first thought, this would appear to be an excellent idea. However, muscular tension produced by the suction cup masks all resting activity. Consequently, the threshold of such a system is relatively high. On the other hand, there is the tendency of some researchers merely to lay the electrodes on the skin using just the surface tension afforded by the electrode paste. Any motion of the subject would tend to cause motion artifacts of a type giving forth high transient activity; these artifacts are often indistinguishable from natural EMG signals. There are many variations between these two extremes.

A newer method, from which good results are expected, is to use a silver or tantalum wire gauze attached to the body through the use of collodion in which silver powder is suspended. This mixture acts as a combined glue and contacting medium, affording great physical strength over long periods of time and a high reliability of electrical contact. Such a combination has been used on test pilots for periods as long as 72 hours and has resulted in excellent physiological tracings. No ill effects have been noted. Size and, hence, selectivity can thus be rigidly controlled.

In the preceding paragraphs, the parameters affecting electrodes in general have been elucidated. It is convenient to point out the differentiation between electrodes used for sensing biological phenomena (at the surface) and those used for muscle stimulation. As a rule of thumb, it can be said that whenever selectivity is important, size must be kept to a minimum. The rule applies to both stimulation and sensing. Location of the electrodes also plays a large part in helping to determine selectivity. However, location is a subject in itself and will be treated as such.

The difficulty (in muscle stimulation) of maintaining current density and simultaneous selectivity has been pointed out by some observers. There are two methods that can be used for stimulating muscles electronically, while still maintaining a good safety factor. The first method is to use a current *limiting* device, which prevents the stimulus current from going above the "safe" level. The other approach is to use a *constant* current device, which maintains the current at some predetermined safe level regardless of load changes.[10]

In the current limiting type, changes in electrode contact, basal skin resistance, or electrode size alter the value of current flowing through the stimulated organism. It is readily apparent, however, that with this method, electrode size has little influence on current density. Thus, in order to achieve stimuli of sufficient (threshold) magnitude, the magnitude of the stimulating pulse must be large.

On the other hand, if the constant current approach is used, the electrode size can be decreased to the point where selectivity is enhanced, because current density increases with a decrease of electrode size.

The subsequent necessary conversion of the electromyographic signals into knowledge that can be utilized in some sort of electronic component is termed operations on the signal.

(b) *Operations with EMG waveforms.* The EMG potentials, as observed at the surface, consist of high amplitude spikes (0.5 to 1.5 millivolts) interspersed with lower amplitude and somewhat slower waves. The frequency bandwidth is approximately 50 to 1,000 cycles per second. The frequency distribution appears to follow a Gaussian distribution.

There are three types of routine operation of primary interest: binary or "on-off" processes, wherein an operation is either performed or is not performed; and two types of continuous flow processes, unidirectional and bidirectional. One variation of the binary type is the stepping function, whereby an operation proceeds to the next logical step upon the presence of an "on" signal. Both continuous flow processes give smooth

control over all ranges, the only difference between them being that bidirectional affords control in both directions from a neutral position.

Three basic parameters of the EMG waveform, amplitude, frequency, and phase must be treated. In addition, there are some applications which conceivably will utilize combinations of all three. We will consider, on these muscle action potentials, two fundamental operations, integration and waveform.

Basically, integration gives an output proportional to muscular activity per unit time. Naturally, it is desired to secure an accurate, continuous integration for all types of signals to be recorded. Previous methods have not been capable of achieving this goal. This is due to the fact that most integrators are built around an RC network. Since any RC network has a finite discharge time, these networks will not follow changes in gross activity accurately, especially when the activity is changing from a high level to a lower level. Another difficulty to cope with in using all types of integrators is that they are not selective as to the signals they will integrate; i.e., they will integrate noise and external artifacts (motion, etc.) together with EMG signals. Such a performance results in an untrue recording of muscle activity.[3,6]

RC integrators are simple and relatively inexpensive to construct. Obviously, since the capacitor in the circuit cannot continue to charge indefinitely, means must be taken to discharge it periodically.

There are two methods of attacking this problem. "Fixed time discharging" utilizes an accurate time reference to activate a relay or gate circuit whereby the capacitor is discharged. It is evident that the output voltage is in the form of a ramp with amplitude proportional to muscle activity per unit time. A major difficulty with this approach is the inherent nonlinearity of the capacitator as it begins to approach the top of its charging curve. The appropriate use of a feedback amplifier and proper adjustment of circuit constants can minimize this difficulty.

"Fixed voltage discharging" utilizes an accurate voltage reference. When the voltage on the capacitor reaches the level of this reference, a discharge mechanism is activated, and one pulse is recorded. The number of pulses is thus proportional to EMG activity. If circuit parameters are correctly chosen, the charging of the capacitor may be limited to the linear portion of the curve, and the difficulty of the previous method is overcome.

If it is desired to obtain a continuous integral, this end can be approximated by rectification to yield the envelope. Averaging methods such as

these yield values proportional to frequency and amplitude; but, as was previously stated, the time constant of the circuit introduces some difficulty. If the time constant is too short (i.e., $\frac{1}{4}t$), the network will do little averaging over infrequent pulses, but will follow the build-up of a rapid series adequately. On the other hand, if the time constant is equal to $2t$, good low-frequency averaging is obtained, but poor response to the short, rapid series of spikes results.[3]

Since the approximation of gross activity is of primary interest, a method whereby the EMG is treated as a modulated carrier and is subsequently demodulated will furnish satisfactory results for continuous flow operations.

Waveform analysis can be accomplished by feeding the EMG through very narrow band-pass filters. In effect, this process, an automatic Fourier analysis of the input signal, would be expensive and cumbersome, but analysis of EMG frequency spectra might yield useful information.

Auto-correlation and cross-correlation methods of analysis, using Gram-Charlier approaches such as have been used by WADC in evaluating EEG waveforms, might be applied to this problem.[11] Equipment is expensive and cumbersome, and evaluation of results would have to be correlated with clinically obtained data.

The methods of operation should be selected on the basis of most appropriate use.

(3) Summary of Electrophysical Observations.[9] The electromyogram, consisting of a series of spike discharges, is a record of electrical activity in a muscle during contraction. Each spike is a voltage wave or action potential caused by the depolarization of a group of muscle fibers or a single fiber. Typical wave forms range in number from a few spikes to polyphasic cluster of 25 or more. The duration of the action potentials from a fiber or fibers is in the order of magnitude of a few milliseconds. The amplitude varies from 25 microvolts to as high as 5,000 microvolts, but the average variation is from 100 to 1,000 microvolts. The pulse repetition rate is 2 to 50 cycles per second. The sounds produced after proper audio amplification are designated rough crackling, disorganized, high, medium, or low rough sounds.

(4) Applications to Complex Problems. The preceding paragraphs have considered techniques of measurements of muscular activities. The following paragraphs will treat these same measurements and their appli-

cations to such complex problems as man's ability to contract muscles to subserve the needs of posture and performance of movement.

An orderly approach to the physiology of muscles involves a consideration of some of the central nervous system functions. A basic function of the central nervous system is to provide a connection between an input (stimulus) and an output (effect). This connection is called a simple reflex arc. The results of a reflex always end in an effect which is contraction of a muscle or secretion of a gland. In all reflex actions, there is a central event (inhibition of facilitation) which modifies the effect.[1,4] These central events are localized by their connections to the neuronal pool of the central nervous system, which means the brain and cord themselves. The complexity of regulation increases as more and more of the central nervous system take part. The end result of the reflex actions is always one of some functional meaning for the individual, and reflex actions occurring one after another in time are coordinated for some purposeful character. Such coordination of neural activity into spatial and temporal patterns which serve useful ends is the essence of the integrative action of the nervous system.[1,4]

In order to effect a purposeful and coordinated goal-directed movement, the muscles of a limb are contracted in such a manner that there is a minimum inherent opposition. This means that a muscle that flexes the forearm, for example, will function together with other muscles that help to flex the forearm, and not with muscles that extend the forearm. Muscles that help each other exhibit *synergism*. Muscles that oppose each other are called *antagonists*. The capability to inhibit the antagonistic muscle is one of the most important events of the phenomena of central conduction. It is a common observation that an arrest of contraction, or slackening of existing tension, is as definite a response of skeletal muscle as is its active contraction. Inhibition of the contracted fibers of a muscle, with resulting relaxation, the fundamental principle for coordinated movements, is known as Sherrington's Law of Reciprocal Innervation.

The principle implies that if one were to compare the total integrated electrical output of a set of opposing muscles, it would be seen that they are 180 degrees out of phase. The fact that these electrical activities are out of phase permits a check on the regulation of any output desired. This concept can be applied in setting up one motor or two motors in opposition to each other.

(5) Muscle Stimulation. In order to have man in the integrated system, one must consider the function of muscle stimulation from the

nervous system or some external source. When a muscle is excited to activity, whether naturally or experimentally, the electrical phenomena produced are essentially the same.[5] A wave of excitation, electrical in nature, causes changes in the surface membranes. In a nerve this is called conduction of the impulse, and in muscles it triggers the actual contraction.

In either natural or experimental stimulation, the resulting electrical activity has the following characteristics:

(a) Low frequency 30 to 120 cps,
(b) Volley duration of 10 milliseconds (approximately),
(c) Variable rest time, so there can be a duty frequency as high as 1,000 cps.

Under natural circumstances, the wave excitation originates in the anterior horn cells (muscle activator) of the spinal cord, or higher, and travels to the peripheral muscle. Every nerve and every muscle, unless deeply covered by other muscles, possesses a small area where it is most easily excitable and where a visible contraction can be elicited with a minimal stimulation. This is a motor point.[5] The results of an electrical stimulation depend on the amplitude, form, and duration of the effective period.

Low-frequency currents providing external stimulation have been used in clinical medicine.[5] The therapeutic employment of low-frequency current is twofold: (1) it furnishes a characteristic procedure known as electrodiagnosis for the recognition of pathological conditions of the motor tract; (2) it furnishes means for the stimulation of weak or paralyzed muscles, a form of electrical muscle exercise valuable in the treatment of injuries and disease.

(6) Muscle Stimulation Apparatus. There are several brands of commercial stimulators on the market. They vary in waveform at the output, as well as the regulation that can be imposed on the amplitude. Typical of such apparatus is the "Myosynchron" manufactured by the Birtcher Corporation.† This apparatus has controls for amplitude, duty frequency, and rest time, as well as controlled modulated cluster shape. Commercial stimulators have been built for specific tasks, such as stimulating the heart after it has been arrested. Artificial stimulators have also been used in cases of respiratory arrest by location of the motor point of

† Birtcher Corporation, 4371 Valley Boulevard, Los Angeles 32, California. *See* Catalog No. 210.

the phrenic nerve in the cervical area. Various authors have maintained that the rate of respiration, as well as the pulmonary ventilation rate, can be carefully controlled for long periods of time, much to the benefit of the patient.[9]

It has been common practice of investigators in the past to use artificial stimulation of muscles in a binary "go" or "no-go" activity. This type of activity represents a simple situation when compared with the objectives entertained in this proposal.

To date, no commercially available stimulator makes use of simultaneous frequency, phase, amplitude, and shifting or changing waveform modulation. It is therefore conceivable that the above factors, combined with multipoint muscle stimulation, will lead to superior methods of artificial muscle control.

4. Psychological Considerations of Behavioristic Learning

The potential of learning, by programming of simplex or complex movements into man, presents stimulating and thought provoking aspects. Although it is a practice to recondition muscles or to prevent atrophy of denervated muscle by electrical stimulation, such repeated stimulation does not constitute learning in the usual sense.

Learning may be defined many ways. Some definitions include "teachableness" and some imply that learning is a "mark of the mind." Some authors do not distinguish between learning and behavior. It is extremely difficult to write an entirely satisfactory definition. In order to avoid the pitfalls of a formal definition, the following provisional definition is offered:

Learning is the process by which an activity originates, or is changed, through reacting to an encountered situation, provided that the characteristics of the change of activity cannot be explained on the basis of native response tendencies, maturation, or temporary states of the organism.

If a behavior sequence matures through regular states, irrespective of intervening practice, the behavior is said to develop through maturation, and not through learning. If training procedures do not speed up or modify the behavior, such procedures are not causally important, and the changes do not classify as learning.

Some definitions or learning avoid the problem of behavior or performance by defining learning as a change in the central nervous system. So

long as this change in the nervous system persists, temporary changes in state, such as fatigue and intoxication, affect performance, but not learning. Learning, therefore, is an inference. Such a definition itself makes use of a type of inference concerning the role of the nervous system in learning. In view of the lack of knowledge of what actually happens inside the subject, it is preferable not to include hypothetical neural processes in the definitions of learning. It is sufficient to state that learning takes place. This position does not deny that what is called learning may be a function of nervous tissue. It asserts only that it is not necessary to know anything about neural correlates of learning in order to know that learning occurs. That playback from a machine to muscles, in order to produce controllable movements, could be an aid to the acquisition of motor skills can be supported by the so-called behavioristic theories of learning.

5. Other Applications

We have considered EMG control techniques primarily in the context of high-performance manned vehicles. Of the other applications which spring immediately to mind, two provide very important refinements of present techniques. These are:

(a) Microhandling,
(b) Precision operation of large scale machinery.

A very appealing example of the latter area is the construction robot. The robot is built in an approximate image of man on a larger scale. A man occupies a portion of the interior of the robot, controls electromyographically the robot's movements with high gain *in analogy to his own*, and utilizes analogous sensory equipment placed on the robot. Man thus becomes a giant, clearing a forest with a few sweeps of his blade.

References

1. Bard, Philip (ed.). *Medical Physiology*. 10th ed.; St. Louis: C. V. Mosby Company, 1956.
2. Dickson, C. J., "Electrophysiological Technique," *Electronic Engineering* (London), 1950.
3. Fink, R. B., and M. L. Scheiner, "The Computation of Muscle Activity from the Electromyogram." Paper delivered at the 11th Annual Conference on Electrical Techniques in Biology and Medicine, November 19, 20, 21, 1958, Minneapolis, Minnesota.

4. Fulton, John F. (ed.). *Textbook of Physiology*. 17th ed.; Philadelphia and London: W. B. Saunders Company, 1955.
5. Glasser, Otto (editor-in-chief). *Medical Physics*. Chicago: The Year Book Publishers, Inc., 1950.
6. Grings, William W. *Laboratory Instrumentation in Psychology*. Palo Alto, California: The National Press, 1954.
7. Immon, Verne T., *et al.* "Relation of Human Electromyogram to Muscular Tension," *Electroencephalography and Clinical Neurophysiology*, IV (1952), 87–94.
8. Marinacci, Alberto A. *Clinical Electromyography*. Los Angeles: San Lucas Press, 1955.
9. Sarnoff, S. J., *et al.* "Effect on Circulation of Electrophrenic Respiration and Positive Pressure Breathing During Paralysis of High Spinal Anesthesia" *Annals of Surgery*, CXXXII (November, 1950), 921–99.
10. Woodbury, Lowell A., *et al.* "Pulsed Stimulator Aids Medical Research," *Electronics*, XXII, Part 2 (August, 1949), 84–85.
11. Zander, Carl A. "Computer Analyses Brain Waveforms," *Electronics*, XXXI, No. 29 (July, 1958), 68–72.

Appendix 15

The Tunnel Display Concept

"... Sight, which embraces space, and, through the medium of light, reveals the existence and color of the bodies which surround us ..."

—BRILLAT-SARARIN

1. Introduction

Solutions to the problems arising in a space order of man-machine relationships appear likely to be of radically different expression from what we are accustomed to in the realm of present day high-performance aircraft. In dealing with the man and the controlling capabilities of man, the designer has entered, long since, upon infirm ground. The man, by virtue of his independent evolution and resistance to alterations, has been handled quite differently, of necessity. The designer in this respect has followed the line or augmentation and complementation of fundamental human equipment; i.e., the objective has been to make the man more-than-a-man through the provision of segmented status presentation and actuation means. The man, however, still must actively participate fully

in the functionings associated with the controlled progress of the craft and with the execution of the mission for which the system was designed.

A major complaint against this "house-keeping" regime for the system controller is that, by its nature—one of compromise—the design substrate for a "natural" evolution of controlling practice and principles is precluded. There are too many instruments, too many controls, too great a dependence upon fallible human judgment. Lesser criticism might be directed toward the heavy expenditure of time and money necessary for training operator personnel to the highly skilled level demanded by high-performance craft.

Thus, design necessity and intractable man have combined to pose a difficult class of problems—problems associated with the provision of display/control means which are sufficiently fundamental in their conceptual and actual structure to serve as the basis for further development with minimum dislocation of the parent system. Such display/control complexes will not come into being easily. Their achievement will demand the full exercise of human ingenuity.

2. The Contact Analog

The antecedents in logic and theoretical design for the tunnel display concept are discernible in a fundamental program undertaken several years ago under the auspices of the Office of Naval Research. In essence, the practical consequences of this program were to influence the design of airplane cockpits. The program was initiated in anticipation of alleviating pilot/navigator stress under unfamiliar and biologically unsuitable environments. A major outgrowth of the program which, though not as yet

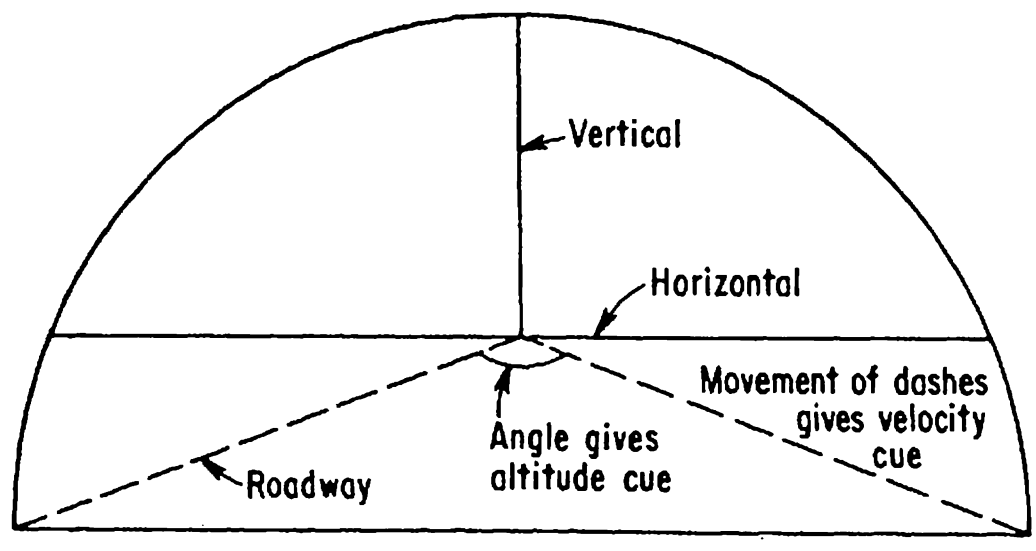


Fig. A15.1 Vertical interim display; Manual (straight/level/command altitude and speed) aspect.

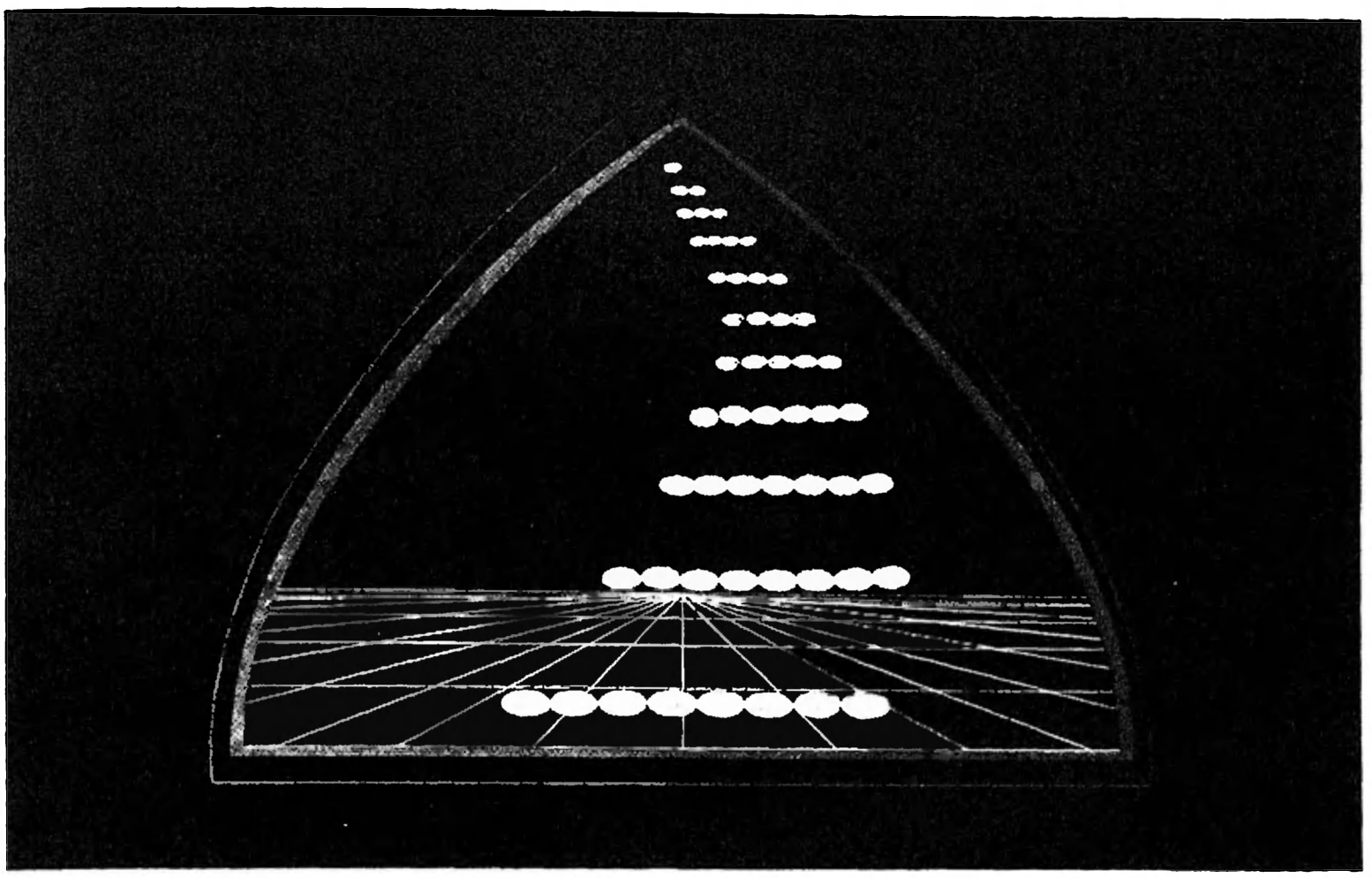


Fig. A15.2 The contact analog.

fully developed, represents a major conceptual advance toward destressing the human subsystem is the integrated cockpit display referred to as the contact analog. A similar but simplified display is the vertical interim display (see Fig. A15.1).

The contact analog, Fig. A15.2, is a form of aircraft instrument display which will incorporate the more important visual cues in contact flight. The display theoretically permits the simulation within the cockpit of an environment which is appropriate for the human organism which it contains. Thus, the windscreen (a flat, transparent CRT) provides a strong cue of internal reference. The relation of the windscreen frame to an integral horizon supplies basic orientation information since the pilot perceives his seat position to be fixed in relation to that of the aircraft. Roll is represented by clockwise or counterclockwise motion of the windscreen frame with respect to the integral horizon. Pitch information is supplied by movement of the windscreen up on the horizon for climb, down for dive.

Perspective in the contact analog is simulated by the apparent convergence of parallel lines at infinity. Systematic changes or gradients in terrain texture are demonstrably useful cues for estimating slant of the earth's surface, altitude, and distance. Distinctive textural differences representing earth and sky are produced; these appear to the viewer to

meet at the surrogate horizon and thus provide an additional aid for the interpretation of spatial organization.

Motion parallax is simulated to provide the basis for estimation of relative distance and velocity. Size and shape of (simulated) familiar objects supply strong visual cues and give orientation information by indicating the position of the pilot in relation to them.

The contact analog artificially presents both environmental and command information through inclusion of an additional feature within the display—"the highway in the sky." The highway in the sky furnishes information for decision-making concerning necessary changes in altitude, attitude, and direction.

3. The Tunnel Display Concept

The tunnel display concept relates to an advanced integrative format for the display of the pilot/navigator (or space-craft controller/occupant) information. In the display hypothesized, the status of system constraints in their entirety is presented to the controller in a visual form perceptually "natural" to him, namely in terms of dynamic alterations of spatio-temporal relationships. These constraints include not only actual

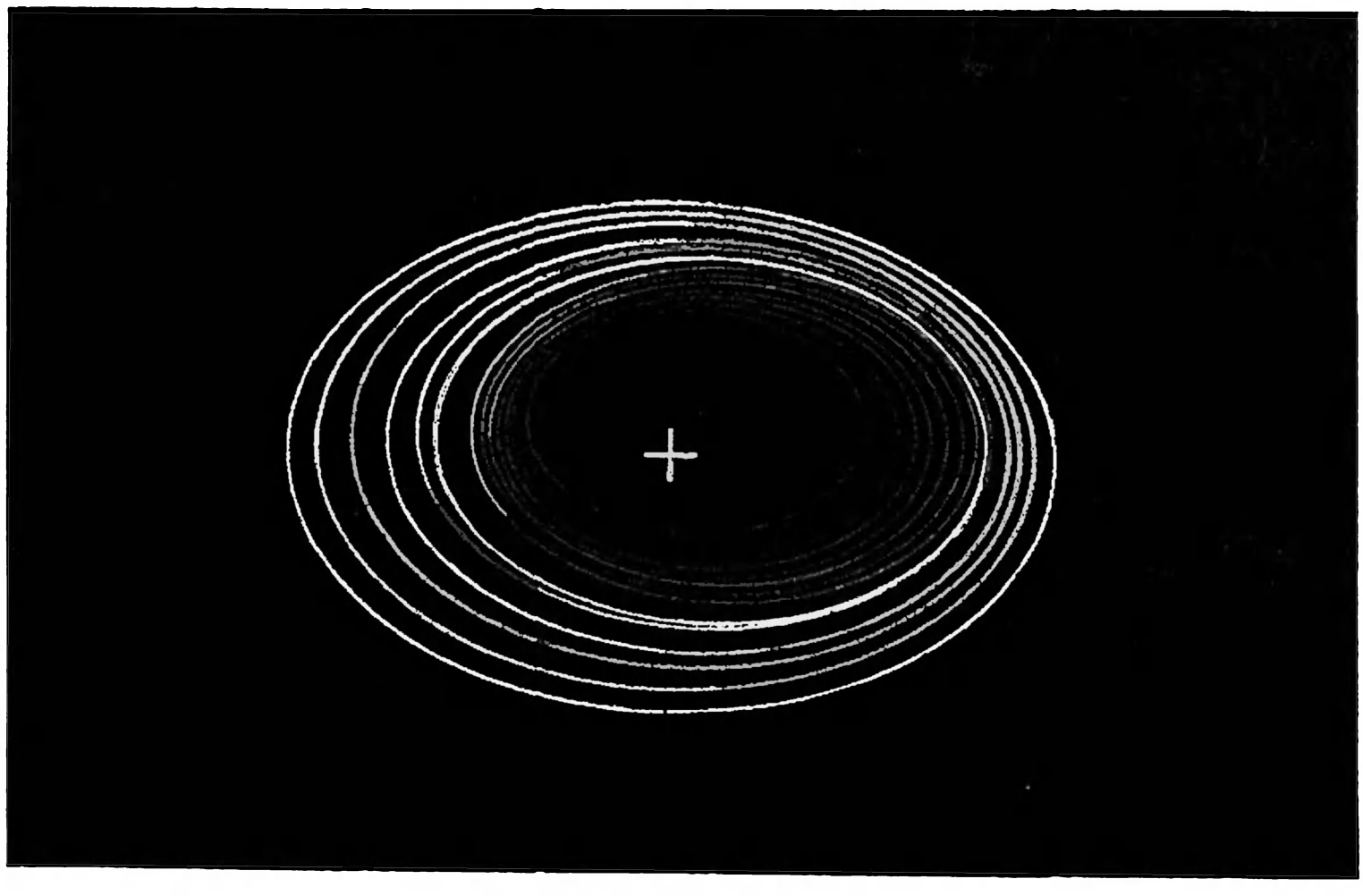


Fig. A15.3 The tunnel display—a transient mode.

spatio-temporal relationships but also spatio-temporal interpretations (made by the control computer) of other constraints, such as the thermodynamic constraints associated with energy-management/trajectory-control.

Figuratively, the controller is operating the conveyance system in a geometric tunnel whose boundaries are defined as the envelopes of classes of admissible operating characteristics. The tunnel constricts, expands, and curves in time in accordance with internal and environmental system parameter changes in transit. An approach to a tunnel boundary indicates a situation going out of control; penetration of such boundary implies abortion of the mission and, perhaps, system destruction. Satisfactory retention of the system within the tunnel implies the capability for completing the mission. The use of such a display format is, of course, intended in conjunction with high capability data-processing, and, possibly, adaptive logical devices which provide adequate quickening and unburdening as well as discrimination among the causes of boundary approach.

One of the major problems to which the tunnel display concept is addressed is that of reducing the variety of information array presentation means (and hence actuator control) to manageable proportions for one individual serving as system controller.

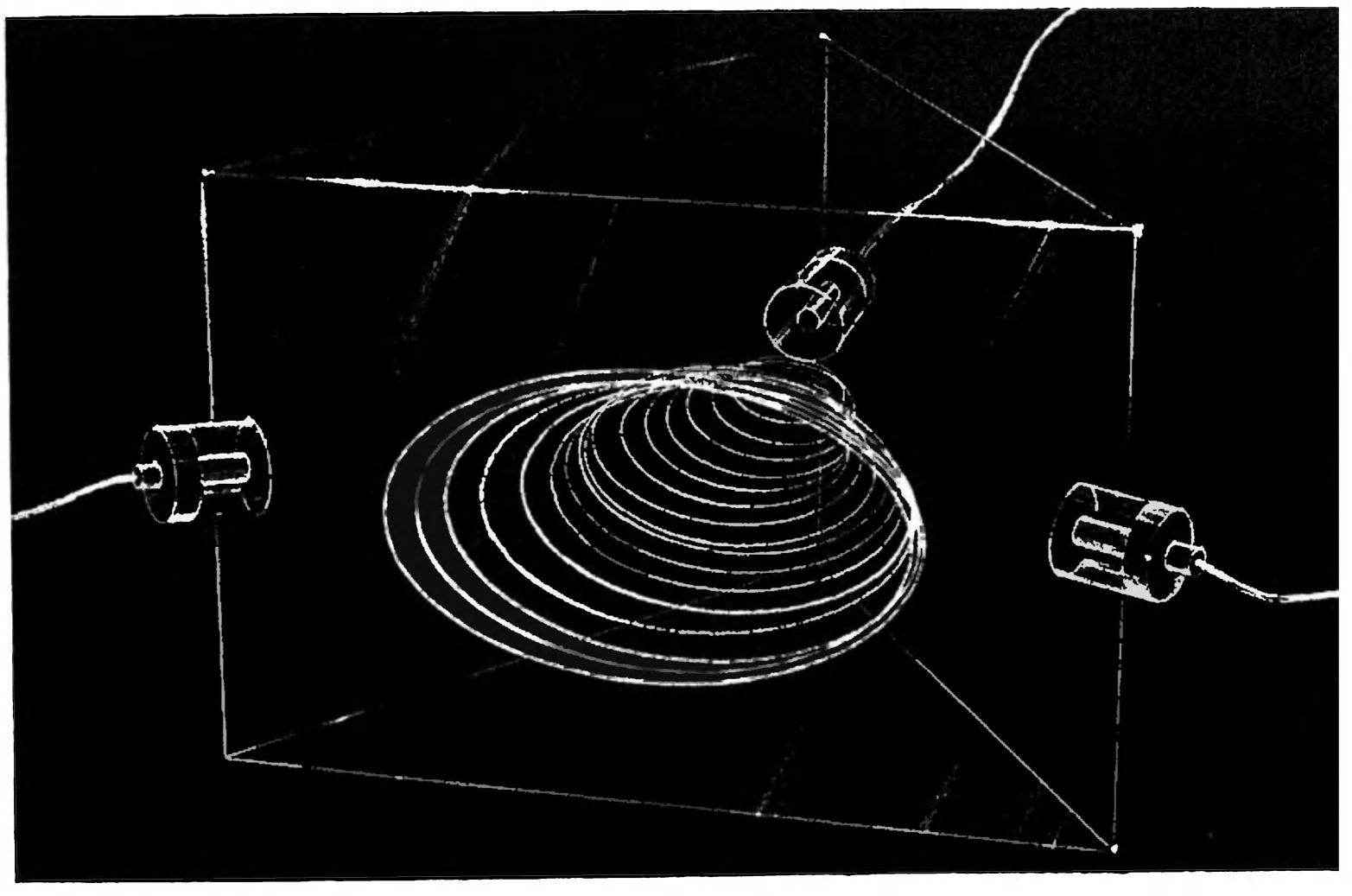


Fig. A15.4 An advanced three-dimensional display, I.

Possibly one of the largest steps which we might take in this direction would be to reduce the perceptible variety of outputs produced by the sensory equipments characteristic of high-performance aircraft. "Perceptible" is a key word here. Its use implies that our approach is one of attention to display format primarily. From the standpoint of the electronic controls designer, the requirement is one of signal conversion to a single form. The single form sought would alter and reduce sensed information irrespective of source and intensity at the point of, or prior to, the display/man interface. These and similar problems of mechanization will receive further attention later in this appendix.

The information input channel which affords the greatest promise for easy entrance within the human organism is visual. We know that modern man receives as much as 85 per cent of his total information concerning his external environment through visual channels. We propose, therefore, to perpetuate convention in that we exploit the unique discriminative properties of the human eye in the design of a display with enhanced dynamic features. There is no theoretical difficulty in presenting the status of system constraints to the controller in a common signal form available to and easily assimilable through the visual channel. Information so presented will be of immediate and "natural" usefulness to the controller, because it will be expressed in familiar terms which, in the common experience of the past, had made possible acceptance and generally beneficial responses to alterations in the environment.

Now, it is true that there are physio-psychological problems associated with such appealing simplicity. The problem in this instance is one expressible in terms of attention to conflicting stimuli received by other inherent channels to the perceptual apparatus of the human organism. One must, in addition, give attention to the situation wherein there is deprivation or serious attenuation of sensory stimuli for the human operator. In the former situation one must cancel or modify conflicting stimuli; in the latter, reactions must be artificially induced at levels above threshold.

Let us now consider the generalized appearance of the tunnel display. Fig. A15.3 is an artist's representation of the appearance of the display in a transient mode. The tunnel, it may be observed, appears to be three-dimensional. Some thought has been given to the means by which this appearance may be achieved. Methods thought to be feasible and psychologically compatible are illustrated by the following figures.

Three-dimensional solutions are receiving active consideration. Figure A15.4 represents an advanced three-dimensional display of considerable

promise. This display combines techniques taken from nuclear instrumentation, particularly gaseous scintillation, physical optics, and computer logical circuitry. The display would be created by a complex of three electron guns, appropriately placed, emitting their charged particles through a small thin window into the visual display area. The confined display area is filled with a noble gaseous scintillator whose individual atoms can be excited to emit light quanta if bombarded by sufficiently energetic charged particles. The acceleration potential imparted by each gun is insufficient to excite the atoms but, when the three beam particles coincide, the atoms in the near vicinity of the conical vertex are excited and will immediately emit light quanta. A small amount of gaseous wavelength shifter is added to the noble gas, which causes the emitted quanta to be in the visible wavelength region. The material used to contain the gas will have an index of refraction of approximately one so as to minimize optical loss and reflections; however, it will have a sufficient thickness to contain all the radiation within its volume. The resolution of this device will be limited only by the mutual focusing ability of the three electron guns. The display complex, while actually coming from several sources, appears to come from a single source. It should be noted that the only form

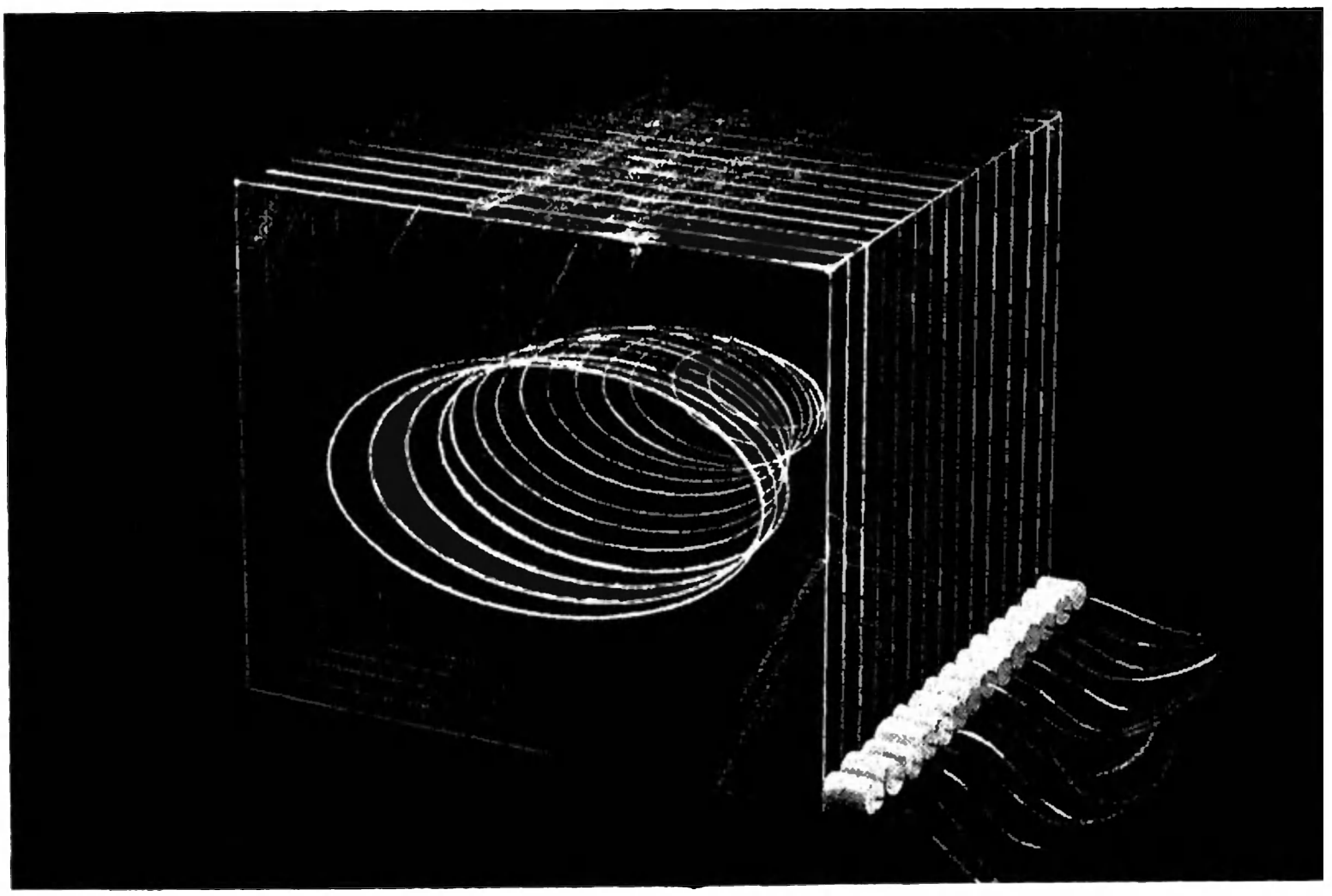


Fig. A15.5 An advanced three-dimensional display, II.

of radio irradiation associated with the display is that of the electron beams which are confined to the guns and the display; thus, there is no additional biological hazard to the viewer. Thought has been devoted to circuitry, computation requirements, and programming aspects for this display—the details of which exceed the scope of this discussion.

Figure A15.5 suggests another possibility for three-dimensionalizing the display. It consists of a horizontal bank of thin transparent CRT tubes. Successive pattern variations, programmable for a given mission, would appear in the apparent distant field of the viewed display; these variations in indication of dynamic alteration of boundary conditions would appear to approach the viewer at a rate positively correlated with actual rate of approach to sensed conditions. A dynamic steroscopic effect would be thereby achieved. Errors of parallax would be minimized through "sandwiching" the CRT's and through use of a reference viewing axis requiring alignment of the observer with fixed reference points integral to the display.

The foregoing schemes have the common motivation of making the display fit into the visual world that the observer has built up through lifelong experience. To provide a stable frame of reference the tunnel is represented as made up of boundaries which, in the simulated near visual field, are definitely located and alterable in their pattern at a rate of change acceptable to the observer. New boundaries are added at the "distant" end of the tunnel as the observer moves through space. These distant boundaries will, as in the real world, approach at a rate distinctly slower than those in the near field. Design rates of change for boundaries throughout the display would be influenced by several factors, namely, elapsed time for the sequence: perception, decision-making, decision-execution, recovery, reconciliation with experience, the particular situation, and the constraints affecting equipment design. In gravity environments an acceptable ratio between the relationship of the craft to its real environment and the displayed relationship would be quite different from that employed in non-gravity, obstacle-free environments. In any case, there are computational capabilities within acceptable rates of assimilation and interpretation by the controller.

Another point to be considered in the area of display compatibility with experience is that of the means by which attention of the controller is to be diverted from the display to computer-selected subdisplay/actuator systems. Probably the most effective and generally economical code would be in the form of words or brief word-groups superimposed

upon the tunnel display, preferably, for reasons to be detailed later, in or near the center of the viewed field of the tunnel.

There are at least two reasons for using words rather than some other kind of coding. The main advantage is that the number of coding steps is unlimited. The second advantage is that the viewer is greatly helped in learning to make use of the code if symbols are used that have been very frequently associated with a variety of responses. Other codes, with the exception of some colors and certain forms do not have this advantage. From early childhood predifferentiation training has been going on for letters and words so that they are easily distinguished from one another. Therefore, the task of associating these stimuli with new responses is made much easier than if the controller has to deal with entirely new stimuli. The results of a large number of experiments provide a great deal of empirical support for this assumption.

Words and brief word-groups, if employed in the context suggested, should be presented in the place most likely to capture the controller's attention. This would be in the center of the display, actually in the tunnel as shown by Fig. A15.6. Upon, or immediately prior to, the appearance of action word(s), "FIRE," "JETTISON CANOPY," "FUEL," etc., a tone would be sounded. This arrangement would have the advantage that

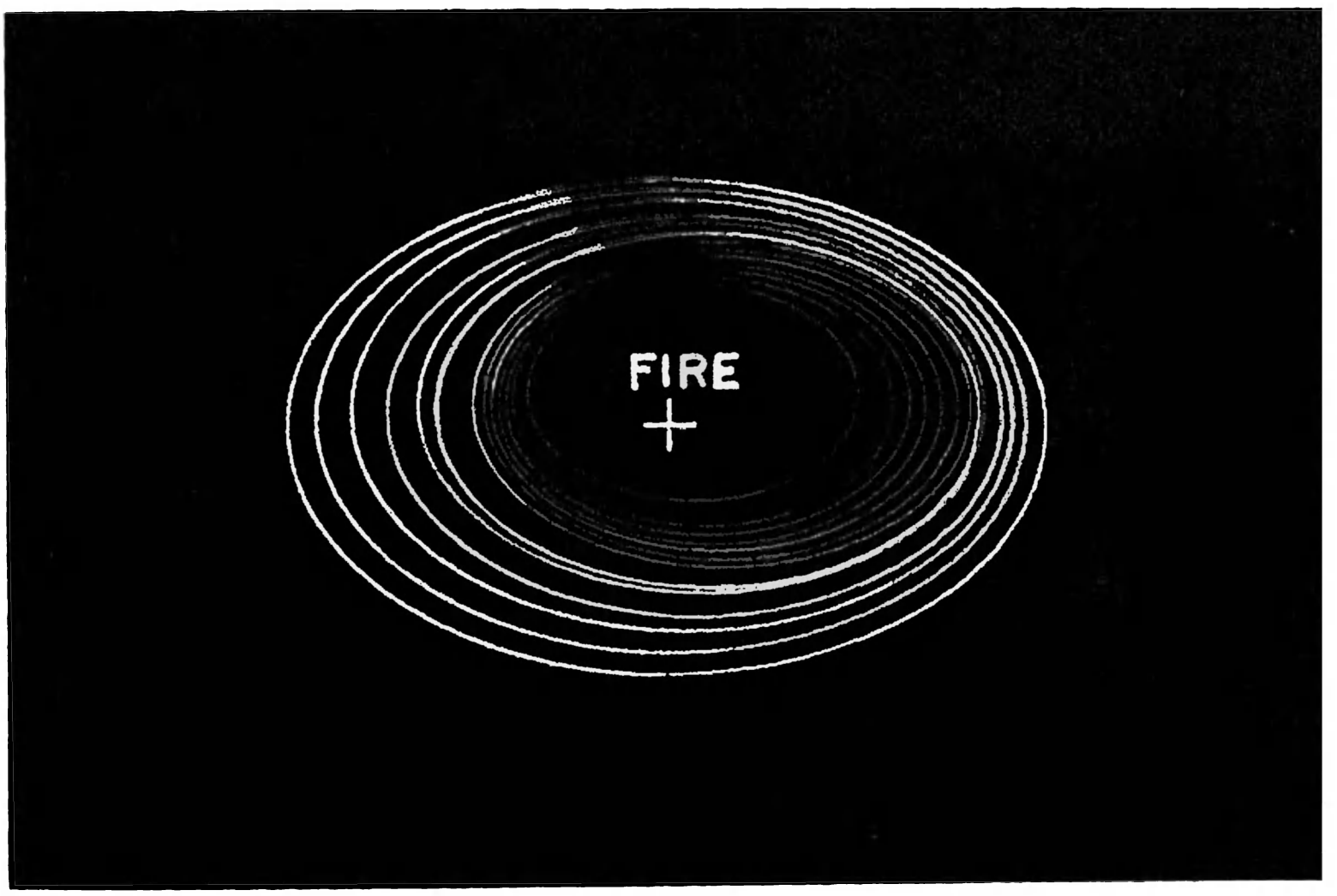


Fig. A15.6 Audio-visual stimuli presentation.

sound is more insistent than visual stimuli. In general, visual stimuli are poorly suited to serve as alerting signals, but ideally suited to serve as directive signals. Experimental investigations indicate that the centering of the alerting signal has at least two advantages. One is that the same stimulus is more likely to be detected sooner if it is near the center than if it is on the periphery of a CRT or CRT-like display. The second advantage is that, under stressed conditions, the observer's field of attention tends to become more restricted if he is monitoring a display.

The tunnel boundaries symbolize analogically a variety of information inputs and provide a perceptually continuous representation of fundamentally discontinuous information as indicated by Fig. A15.7.

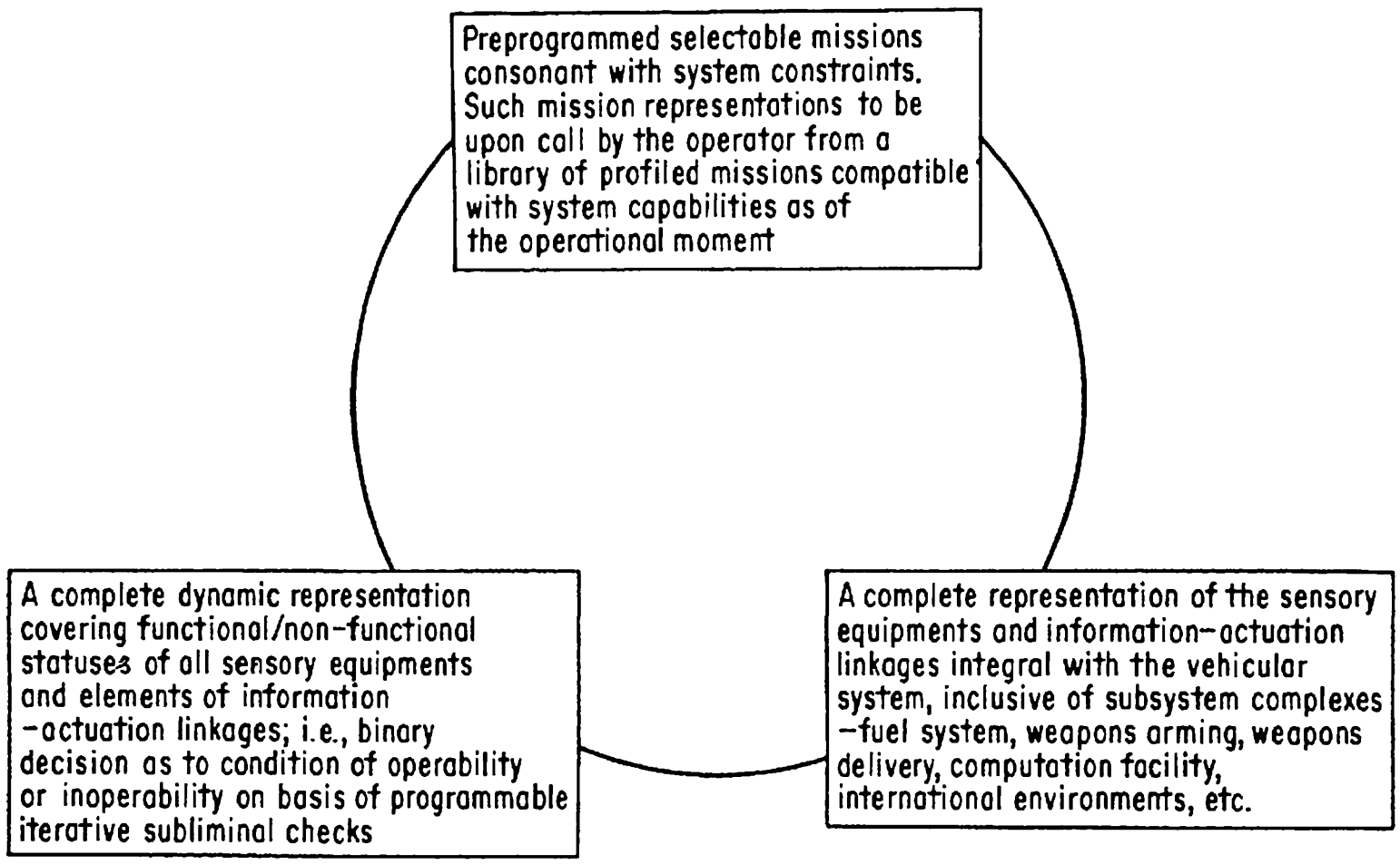


Fig. A15.7 Information supplied analogically by the tunnel display.

Another detail of the design of the display requires clarification. This relates to the manner in which alignment of the observer's visual axis is maintained in proper relationship to the tunnel outlines during critical periods. This may be accomplished as shown in Fig. A15.8. At the center of the face of the display a cross hair pattern is inscribed. Under dynamic conditions of flight the axis of the velocity vector of the craft is as represented by line *bc*, or by its continuation to the eye of the observer, line *abc*. Upon impingement of the representational axis of the velocity vector by a tunnel boundary at point *c*, visual and auditory signal circuits are

Fig. A15.8 Relationship of the visual axis to the display.

actuated as mentioned earlier. The impingement point "blossoms" to an intensity distinctly in contrast with the tunnel boundaries. The controller's attention is now fully absorbed by the display, and he follows a course of corrective or, as the case may be, selective action. Under normal situations of nonimpingement, no perceptible signal would be actuated other than that provided by tunnel boundary movement. Primary sensing of the impingement would be accomplished automatically by means of computer-associated circuits. Design of this portion of the device will be in accord with the special geometrical and optical treatments involved. Automatic sensing of a trend toward tunnel boundary penetration will, through appropriate circuitry, actuate visual and auditory stimuli prompting the controller to corrective action, or possibly to selection of a secondary preprogrammed mission. A further advantage of this scheme for on-call alignment of the visual axis is that when it is most critically important to the controller, a stable appearing reference field is supplied.

We now describe other detailed features of the tunnel display. In keeping with the principle that we should not present more information to the controller than he is capable of interpreting readily and unambiguously, we should perhaps permit a certain latitude in the aberrations of sensory/actuator mechanisms and indicate only those instances wherein there is an indication of trend toward failure. Thus, we might apply the principles

of statistical quality control to sensory/actuator status. We would program, as suggested earlier, subliminal operational checks near-continuously. Assuming the sensor or actuator of interest were at the check-moment, within "acceptable" operational limits, the conditions would not be depicted. In the commercial management vernacular, the system would in this respect be operating under the canon of management by exception.

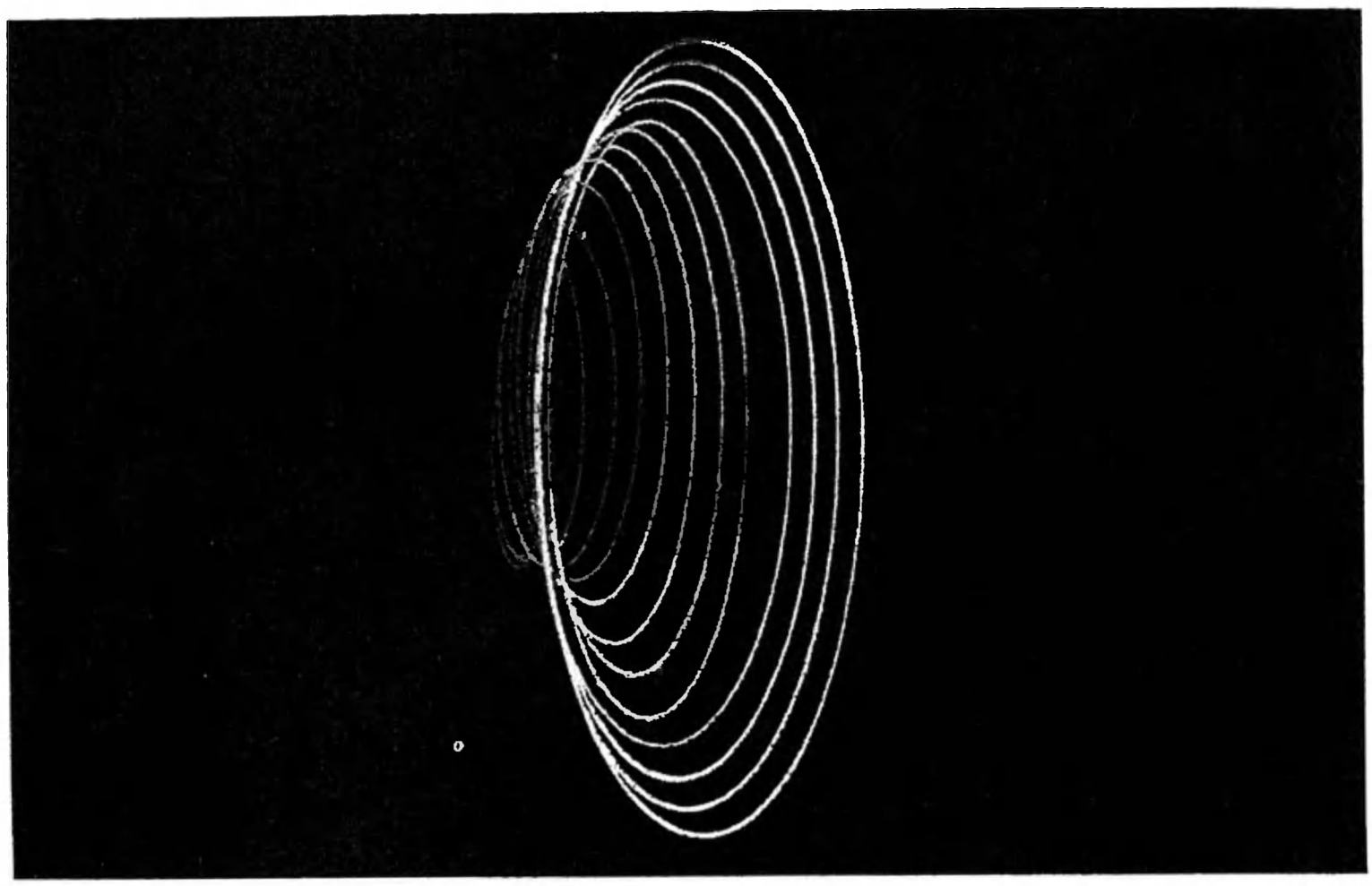


Fig. A15.9 A critical situation, I.

Figure A15.9 illustrates a condition wherein the tunnel has begun to assume a rather pathological form, distinctly elliptical, with the minor axis progressively reduced, and becoming more so at a greater apparent distance in the visual field. From this condition the controller would deduce that he was entering a critical area in his mission. If he were in an aircraft, for instance, he might be following a protective corridor between enemy gun emplacements, or conserving his fuel by following a pass-route through an air-navigable mountainous region. Figure A15.10 depicts a critical astro-navigation aspect; wide horizontal latitude is suggested in the corridor outlines, but there is vertical restriction. The spacecraft may in this instance be following a re-entry pattern.

Before proceeding further it would be well to make clear that subdisplays and subactuator systems will be provided. This implies that we must

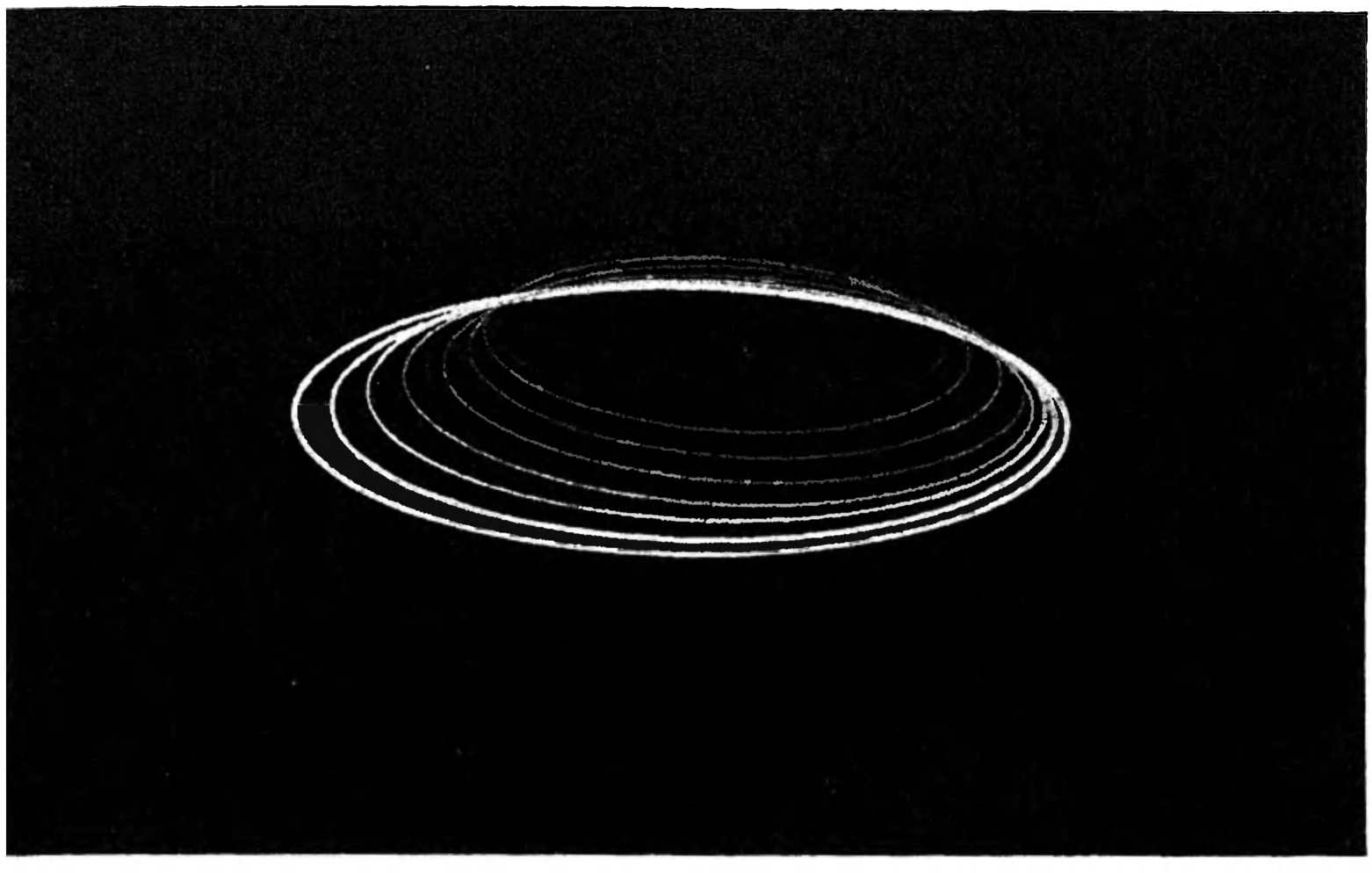


Fig. A15.10 A critical situation, II.

also be concerned with the relationship of the central tunnel display to the discrete and only partially integrated subdisplays and actuator arrays. Attention of the controller under trends of degeneration will be gained through stimulation of two major human input channels—vision and audition. This is desirable in that the controller's attention may be temporarily distracted from the visual display—an auditory signal actuated concurrently with a visual signal would restore his attention to the visual channel. Coding of the visual channel would be such as to direct his attention to the discrete standby display/actuator complex of relevance to the condition signalled. A system design constraint to be observed here would be that those display/actuator arrays which were of greatest priority from the standpoint of system preservation (or possibly of mission satisfaction) would occupy the visual-anatomical envelopes of greatest sensitivity. Those having less urgent relationships would be located in areas of lesser availability. In general such arrays would be distributed in a fashion consistent with the logic of their use.

In keeping with the principle of minimal demands upon the personal economy of the controller, only one highly integrated manually operated control will normally be used in conjunction with the major display. As in the display itself, the control is also a spatio-temporal "interpretation" to the computer of actions necessary to retain an appropriate location in

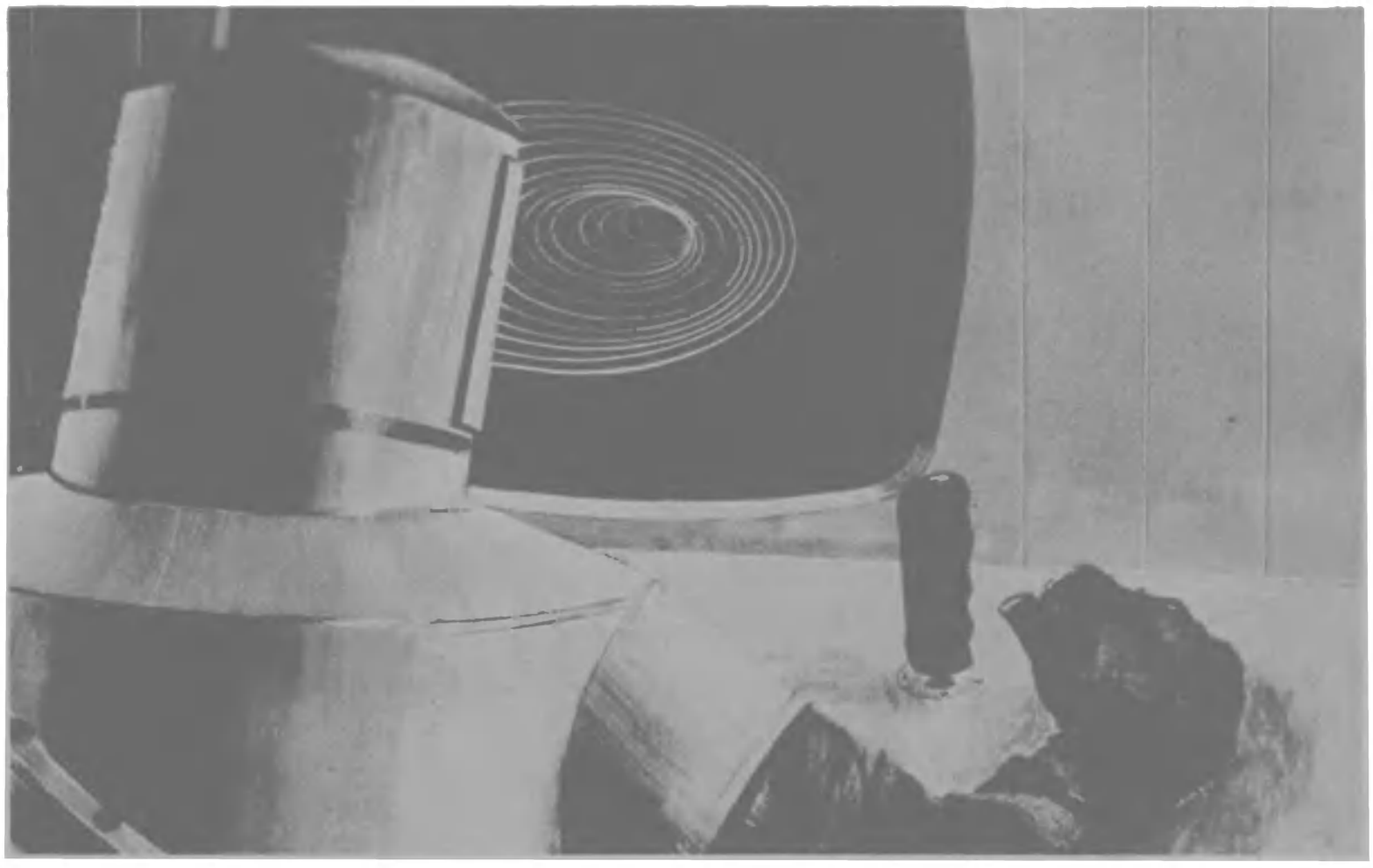


Fig. A15.11 A generalized cockpit layout.

the tunnel. This control as illustrated in the cockpit layout depicted by Fig. A15.11 is not encumbered by encrustations of special purpose buttons, knobs, switches, etc. Its contours fit the hand nicely, and "feel" of the feedback is artificially produced and supplied to create the sensations of response normal to the flight mode in effect. The control responds readily to rolling motions of the hand, wrist, and lower arm. A rest is provided for support and stability of the forearm. Other actuator and operability-checking devices are provided, but these are masked except during periods of use. Such arrays are to be highly integrated within themselves and with the tunnel display as well; their exposure would normally be automatic and sequential as programmed, or as delineated by preprogrammed routines governing unusual or emergency procedures. Under other special operating and in-flight maintenance checking procedures, the masked arrays would be selectively exposed upon call by the controller.

4. An Approach to Implementation

To illustrate one approach to mechanizing the tunnel display, the following hypothetical generation equipment can be considered.

A cathode ray tube using electrostatic deflection along the xy axis may be used for the purpose of providing a display medium. The generation of circles and ellipses on the face of this tube can be produced by establishing specific families of Lissajous figures.

It is only necessary to specify amplitude and phase angle, because all figures of concern are generated by application of the same frequency to both axes of the display tube. For example, a circular section is produced by exciting the two inputs to the tube with the following voltages:

$$X = A \sin \omega t,$$

$$Y = A \sin (\omega t + 90) = A \cos \omega t.$$

The displacement of the center of the circle from the origin of the tube coordinates is accomplished by generating two d-c potentials, x_1 and y_1, which when summed with X and Y will produce the circle with relocated center.

The ellipse is generated by specifying the rotational angle R of the major axis (σ), the magnitude of the major axis (a), and the magnitude of the minor axis (b).

The equations are then:

$$x = (a \cos \sigma - b \sin \sigma) \sin \omega t,$$

$$y = (a \cos \sigma + b \cos \sigma) \cos \omega t.$$

Using a pair of resolvers it is apparent that these equations can be simply mechanized as shown in Fig. A15.12.

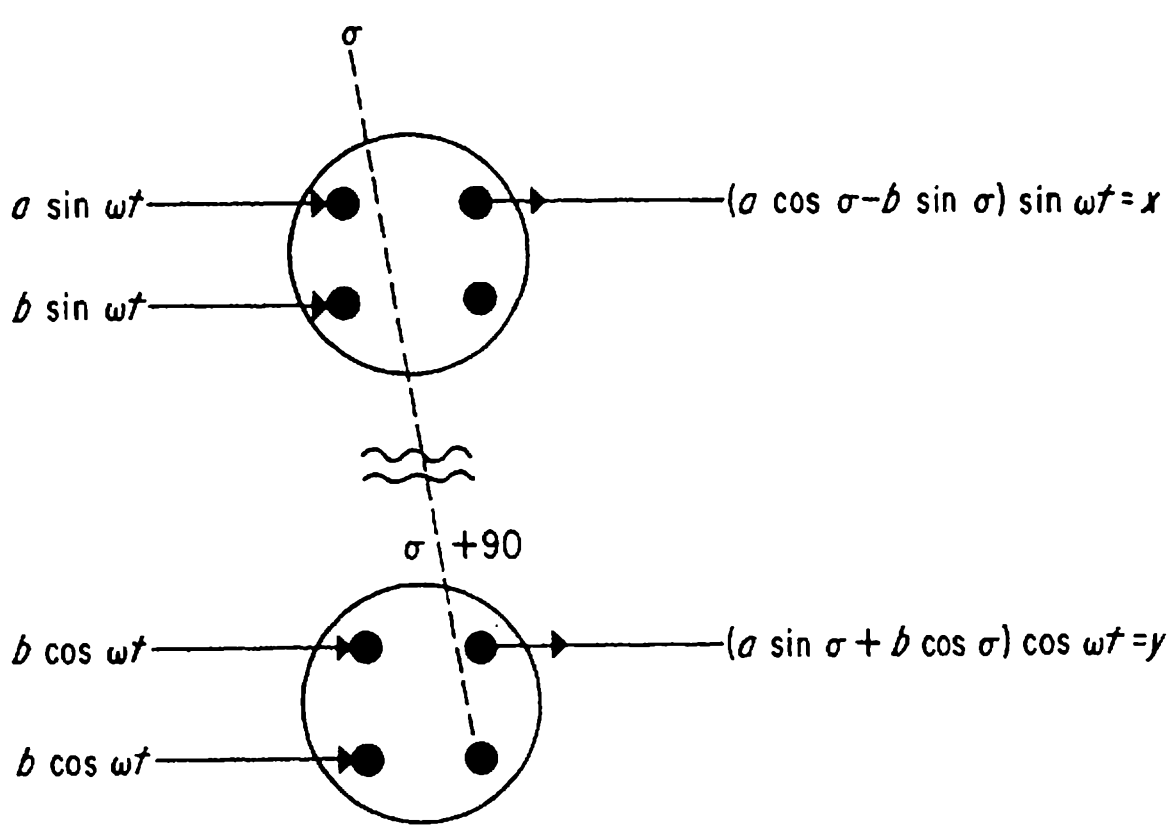


Fig. A15.12 Mechanization equations.

The same method of biasing the ellipse so as to produce an offset ellipse is recommended.

In addition to requiring computed values of a, b, σ, x, and y_2 as prime parameters, a perspective generator is needed which will cause the aforementioned variables to be in approximate perspective as a function of slant range to be the center of an imaginary ellipse in true space. Dynamically then, as the display is generated, a modulation of the true values of a, b, σ, x_1, and y_1 takes place to create the family of ellipses (the circle now being a special case).

One additional item is required. This is a blanking generator which is used to blank out all but discrete portions of the display so that sequences of ellipses rather than a continuous spiral will appear. Figure A15.13 presents a simplified diagram illustrative of this approach.

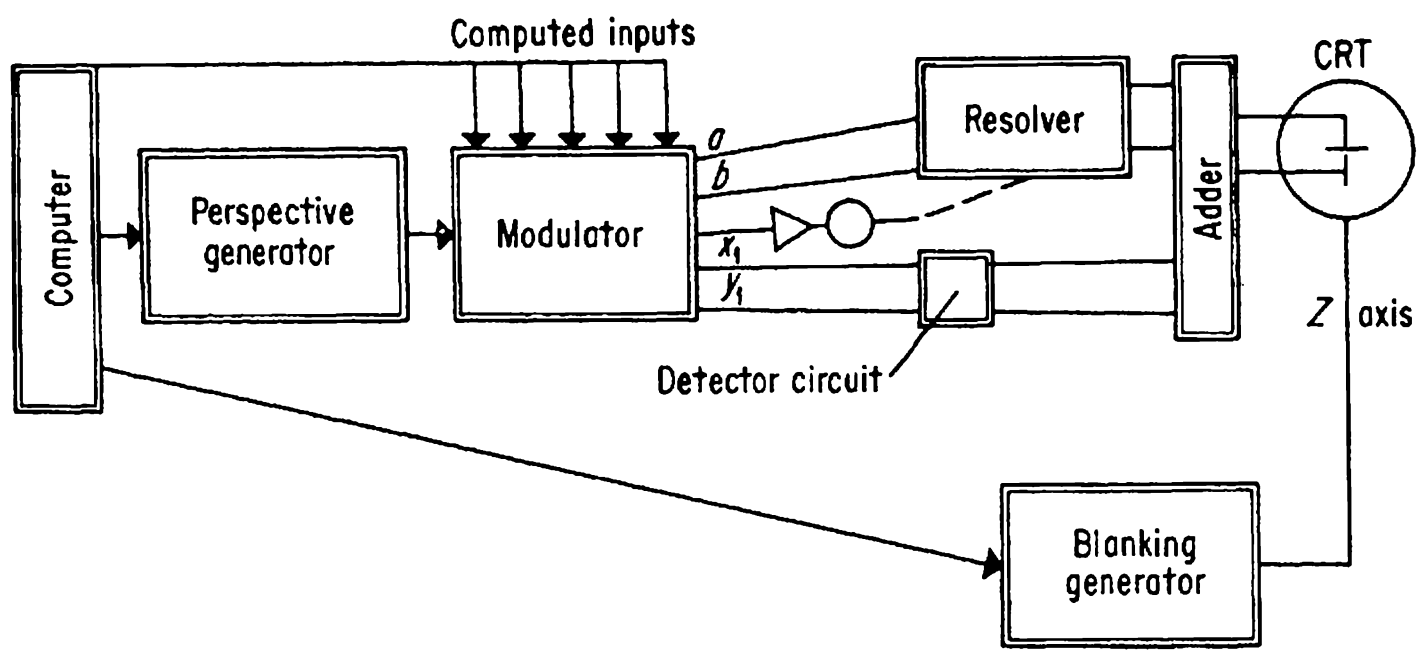


Fig. A15.13 Simplified mechanization diagram.

In summary, the concept advanced here rests upon an urgent concern with human decision-making and execution processes in exceptional environments. Under the terms of the tunnel display concept, system controller behavior is linked firmly with the satisfaction of one mission. The mission may be of primary, secondary, or other order of significance, dependent upon successively integrated statuses of constraints upon the system. Mission profiles in consonance with system limitations are preprogrammed and become effective upon selection by the controller. By so doing, the decisions are made at a level of selection from among alternative, temporarily "prepackaged" missions, *not* at the level of personal integration of system constraints. In effect then, we shall have reduced a great variety of disparate decisions at discrete levels and various degrees of relationship to decisions at the highest level realizable. The functions

of the controller are thus enabled to expand from the contemporary level to levels associated with selection from achievable missions. In terms of warfare, the controller is now a selector of strategies rather than a participant-observer in local combat tactics.

Additional Reading

Arnoult, Malcolm D. "Recognition of Shapes Following Paired-Associates Pretraining in Air Force Human Engineering," *Personnel and Training Research*, eds. G. Finch and F. Cameron, ARDC TR 56–8, (Baltimore, Md.) 1956.

Baker, C. H. "Attention to Visual Displays During a Vigilance Task: I. Biasing Attention," *Brit. Jour. Psychol.*, LIX (1958), 279–88.

Bursill, A. E. "The Restriction of Peripheral Vision During Exposure to Hot and Humid Conditions," *Quart. Jour. Exp. Psychol.*, X (1958), 114–29.

Van Laer, John. "Factors Relevant to the Development of Aircraft Warning and Caution Signal Systems," USN Air Materiel Cent. Lab. Rep., No. 348 (1957), 40 pp.

Index

Accelerometers:
  angular, 297
  coupled, 296
  gravity effects on, 305
  gravity field vector discrimination scheme, 303
  platform-stabilized, 305
  telemetered, 334
Acceptance testing, 104
Accounting:
  automation of, 110, 112
  Social Security system, 77, 113
Action potentials, muscular, 348-358
  integration, 355
  waveform, 354, 355
Actuation means, 6
Adaptive control, 51, 85, 86
Additive processes, 10–11
Agricultural uses of nuclear techniques, 109
Aiding, 81
Albedo change, 335
Algorithmic method, for solving equations, 50
Allowed states, 59–60
Analog computer, 33, 51
Analog function generators, 298
Analog trackers, 320, 333
Analysis:
  applications, 32
  approaches to, 24–26
  capabilities, 28–32
  classification, 26
  constraint, 43
  economic, 40
  error, 319–324, 325
Analysis (Cont.)
  evaluation, 44
  evolution, 44
  extension, 43, 44
  feasibility, 32–37, 43–44
  growth potential, 43, 44–45
  hybrid systems, 26
  mathematical, 26
  operational, 42–43
  operations, 337
  requirements, 26–28
  systems, 325, 337
  waveform, 354, 355
Applications analysis, 32
Army surveillance system, 236–252
  communication channels, 243–249
    code selection, 251
    radio and wire systems, 243–249
    radio *vs.* wire transmission, 247–248
    telephony *vs.* telegraphy, 248
  design, 249–252
  equipment diversity, 237–238
  flow diagram, 250
  information requirements, 238–243
    format diversity, 238, 242–243
  problem areas consideration, 236–243
  temporal data requirements, 237
Ashby, W. Ross, 67
Astronautical systems, 293–301
  need for program, 295–296
  objectives, 294–295
  plan of attack, 296–301
  research and development studies, 300
  scheduling, 300, 301

Astronavigation:
Doppler effect, 302–341
instrumentation, 303
Atomic amplifiers, 318
Atomic clocks, 303, 310, 320–321, 325, 332, 338
Atomistic decision, 76
Autocontrol, 48
Automation:
basic, 110–111, 117–119, 123–124
capability development, 119
economic, 119
machine control, automatic, 117–118
market, 123–124
military, 119
purpose of, 119
reliability, 119
compromise, 81, 113–116, 122–124
major problems, 122–123
planning, 123–124
control for weapons system management, 40
in electric power generation, 121
electronic components of automated systems, 122–123
elimination of moving parts, 123
of flight safety procedures, 268–292
limitations, 113–114
machine tools, 118
organizational, 111–113, 118–122
partial, 122–124
of scientific computations, 112
in telephone system, 121–122
Automobile factory, automated, 122
Autopilot, 48, 298

Baird, K. M., 321
Ballistic computations, 49
Band-pass filters, 355
Basic automation (*See* Automation, basic)
Beam spraying, 87
Beat-frequency coding, 74
Beaufort number, 343
Behavior:
behavioral theory, 125–126
of static decisions, 264–267
behavioristic learning, 358–359
mathematical model, deterministic, 3
randomness, 3
stochastic variables, 3, 6
Bellman, Richard, 139
Bidirectional flow processes, 353
Binary expansion, 7
Binary processes, 353
Biological systems, 84–90 (*See also* Cybernetic systems)
Bionics, 84
Biotechnical systems, 350–353
Birmingham, H. P., 79
Birtcher Corp., 357
Blanking generator, 376
Bookkeeping, double entry, 14
Boost-out guidance, 325, 326
Borel set, 144
Branching operations, 75
Buchheim, R. W., 309

Cables, coaxial flexible, 350
Capabilities analysis, 26, 28–32
Cathode ray tube, 350, 375
Central nervous system:
computer-oriented characteristics, 74
features designed into control systems, 84
functions, 356
muscle stimulation, 356–358
neuronal pool, 356
Charnes, A., 99
Checkout, 113
Check-reading, 79
Chemical industry, automation in, 121
Chronometer, 326, 337, 338
Clincial measurement techniques, 349–350, 351
needle electrodes, 349, 351
surface electrodes, 349–350
Coding:
beat-frequency, 74
code selection, 251
digital computers, 59–62
human participation, 67–68
linguistics, 68
multivalued code, 169
optimal performance, 67
parity codes, 60
transformation processes, 67
tunnel display concept, 368–370, 373
Cojoined operations, 75
Colonization, space, 116
Color hearing, 69
Command system, 308, 309
Commercial systems, 101–102

Communication systems:
blockages, 95–96
radio and wire systems, 243–249
telegraphy, 247, 248
telephone, 247, 248
Complexity, in systems, 54–81, 124
difficulties arising from, 55
exponential decrease in reliability, 58–59
human performance, 69
measures of, 54–55
necessary complexity, 75, 76
in systems engineering, 124
Components, 9–12, 50–51, 122–123
capabilities, 50–51
defined, 11
electronic, of automated systems, 122–123
Compromise automation (*See* Automation, compromise)
Computers:
analog, 33, 51
central control:
for aircraft systems, 17, 48–50
data sources, 169
expansion of examples into matrices (charts), 175–225
information processing, 168–169
for man-machine systems, 166–169, 346
requirements, 165–169
in space vehicles, 112
computations, 55–61, 112, 148
course, 335
crosstalk between, 247
in decision problems, 231
digital (*See* Digital computers)
energy management, 335, 337
equipment, 98–99
gyro torquing, 336
ideal, 33–37
logical circuitry, 367
position in systems heirarchy, 11
remote control programming, 247
system for in-flight safety, 268–292
vehicle line-of-sight-stabilization, 334
Confidence limits, 55
Console, FCO (fire control officer), hypothetical weapons system, 155–164
Constraint analysis, 43
Continuous flow processes, 353–354
Continuum hypothesis, 54–55
Control systems:
man-operated, 79–81
special problems, 82–90
Cooper, W. W., 99
Cost estimation, programming for, 99–100
Cryogenics, 51
Cybernetic systems:
devised, 84–87
natural, 87–90
advantages, 87–90
fabrication-in-the large, 87
self-test, 87
patterned after biological systems, 88–89
thought processes, 84–90
Cyberneticist, 15

Decision space, 141
Decisions:
atomistic, 76
complex, 76
conceptual structure, 76
in-flight emergencies, 270–271
simplex, 66, 76
static, 264–267
behavioral theory, 264–267
environment, 265–266
estimates and criteria, 266–267
vs. dynamic, 265
Description, of the system, 48–53
descriptive model, 48–50, 52
problems, 50–51
simulation, 51–53
Design:
airborne weapons systems, 20–21
army surveillance system, 249–252
central control computer, requirements, 168–169
constraints, imposed, 20–21
evolutionary, 38–45
flow, 38–45
growth potential, 40–44
system matrices, 40, 165–225
freeze, 21, 52
hardware, 136
human value, 17
hypothetical system: Porcupine, 145–164
logical, 12, 132–137, 230–235
Mark I, 38–45
objectives, 20–23
optimal, 16–18
organization, 94
Devised systems, 4, 108

Digital computers:
airborne, 56–57
for astronautical systems, 298
control, 27, 165
counting register in, 7
general purpose, 112
in navigational instrumentation, 32
prediction and extrapolation, 33
simulation, use of, 52–53
Digital control system, 33–37
Digital delay, 35
Digital integrator, 51, 226–229
Digital spectral frequency tracker, 332
Directed set, 129
Directive triumvirate, 102
Disallowed state, 59–60
Display:
advanced, 77, 78–79
compatibility with experience, 368–369
computations, hypothetical weapons system, 153
control elements on FCO console (table), 157–163
conventional, 77–78
digital computer, central control, 165
integrative, 77
interim, 77, 78–79
mapping, 291
military, 77–81
quickening, 79–81
three-dimensional, 366–368
time framework, 77
tunnel display concept, 361–377
unburdening, 79–81
vs. major computation, 170
visual, 366–374
Doppler effect:
astronavigation, 302–341
of electromagnetic signals, 310
error analysis, 319–324
relativistic phenomenon, 310–317
accuracies, 313–314
curves of constant frequency shift, 314–316
radial component of velocity, 316–317
techniques, 317–319
theory, 310–313
in space informational studies, 296
space navigation, 302–341
superiority to radar, 304–305
system considerations, 324–337
lunar guidance, 325–326
Doppler effect (Cont.)
systems matrix, 170–174
theory and techniques, 305–324
time dilation, 312
transition of ordinary to relativistic Doppler, 112
transverse, 312
Doppler radars:
error-noise statistics, 31
navigational, 31–32, 48, 342–344
Doppler radio, 305
Doppler shift, 312, 318, 321, 325, 333, 337, 342
Doppler spectral frequency tracker system, 332–333
Drag information, 271
Duffett, J. R., 253
Dynamic programming:
in decision-making, 231
defined, 138
in logical design, 137
processes, 141–144
components of, 144
optimality, principles of, 142–144
substituted for combinatorial planning, 148–149
terminology, 139–141

Effectiveness, 13–18
Electric power generation, automation in, 121
Electrode pastes, 349, 352
Electrodes:
active, 350
detecting and stimulating, 350–353
ground, 350
location of, 353
needle, 349, 351
pediatric chest cardiographic, 352
reference, 350
surface, 349–353
Electrodiagnosis, 357
Electromagnetic signals, 310, 313
Electromagnetic theory:
Lifshitz-Landau, 51
Maxwellian, 51, 310
Electromechanical checkout devices, 113
Electromyographic techniques:
application to complex problems, 355–356
behavioristic learning, 358–359
equipment required for, 350–353
integration of man-machine systems, 345–360

Electromyographic techniques (Cont.)
microhandling, 359
operation of large scale machinery, 359
robots, 359
Electronic components, of automated systems, 122–123
Electrostatic deflection, 375
Ellis, David, 141
Energy management computers, 335, 337
Equality test, between two functions of time, 231
Error analysis, 319–324, 325
Evaluation analysis, 44
Evolutionary designs, 38–45
Evolutionary planning, 78–79
Exploration, extraterrestrial bodies, 115–116
Extension analysis, 43, 44
Extrapolation, 33–37

Fabrication-in-the-large, 87, 90
Fail-safe operation, of automated systems, 123
Failure modes, 106, 263
Feasibility analysis:
broad-brush considerations, 32
defined, 26
growth potential, 43–44
prediction and extrapolation, 33–37
Feedback:
amplifier, 354
in control systems, 82–84, 120
impulse, for engine control, 326
informational, 70–71
mode-switching, 82–83
open-loop control, 82–83
stability, 83–84
Filters:
band-pass, 355
low-pass, 319
nonlinear, 74
optimum, 31
Fire control officer, 155–164
Fisher, R. A., 25
Flight management program, 115–116
Fourier analysis, 355
Frequency trackers, 321, 342, 344
Functions of time, 230–231

Gabler, R. T., 309
Gaussian distribution, 353
Gestalt perception, 66
Good policy, 138–139, 144
Gram-Charlier approaches, 355
Gravity field vector deduction, 305–309, 338
Gross approximations, 7
Growth potential analysis, 43, 44–45
Guidance:
boost-out, 325, 326
ICBM schemes, 325
interplanetary, 338–341
Q scheme, 325
self-contained *vs.* command, 325
Gyro torquing computer, 336
Gyroscopes, 336

Hahn, Hans, 54
Hardware:
design, 136
philosophy, 136, 231
techniques, 125
Heading reference, 342, 344
Heat detectors, 291
Heat seeking missiles, 291
Heffern, E. C., 309
Heuristic programming, 85, 86
"Highway in the sky," 364
Human engineering, 17, 65, 72, 157, 165, 277–278
Human subsystems:
automatic reactions, 269
capabilities peculiar to humans, 69
computer characteristics of man, 74–75
encoding and decoding, 67–68
human central nervous system, 74
human memory, 75
human perceptual abilities, 66
human problems, 64
human sensory abilities, 65–66
input capabilities, 65
judgment errors, 70–72
man-machine comparisons, 72–77
decision input capabilities, 75
flexibility, 73
judgment and reasoning, 73
memory, 73
perceptual abilities, 73
prediction, 73
sensory ability, 72–73
man-machine interfaces, 77–81
man-operated control systems, 79–81
mental activity complex, 68–69
thought processes, 84–90
verbalizing capacity, 85
Hybrid systems, 4, 26, 70, 104, 108

ICBM guidance scheme, 325
Ideal computer, 33–37
Inertial/radio-fix navigation, 112
Inertial system, stellar, 334–335
Information flow:
  for Mark I design, 39–40
  processing, 16, 168–169
  in system management, 94–96
  weapons system, 150–155
Information transfer, 10, 60
Informational systems:
  requirements for army surveillance system, 238–243
  signal and noise, 28–32
  space-oriented, 293–301
    need for program, 295–296
    objectives, 294–295
    plan of attack, 296–301
    research and development studies, 300
    scheduling, 300, 301
Infra-red detectors, 296
Inherent indifference, 265
Input:
  defined, 3
  devices, 118
  feedback, 82–83
  Fourier analysis of signal, 355
  human capabilities, 65
  information flow, weapons system, 152
  nature of, 4
Input argument, 130
Input space, 130
Instrumentation:
  astronavigation, 303
  nuclear, 367
Integrator, digital, 51, 226–229
  definition, 226–227
  mathematical model, 226–229
  schematic, 227–229
Interplanetary guidance, 338–341
Ionic polarization, 351
Iron industry, automation in, 121
Iteration rate, 27
Iterative method, for solving equations, 50

Kahn, A., 79
Kerr cells, 303
Kinetic energy, 297, 307
Kluging, 44, 78, 112, 124
Kohler's apes, 85
Korean conflict, 21

Lanning, Frey, and Trageser, 335
Laplace principle, 266
Laplace transform, 33
Laramore, L., 338
Learning, defined, 358
Lifshitz-Landau electromagnetic theory, 51
Linear programming, 137
Lissajous figures, 375
Logarithmic spirals, 303
Logic, foundation of new design philosophy, 230–235
Logic circuit, 11
Logical design, 12, 132–137
  problem, 134–136
  terminology, 132–134
Logical operations, 27
Logical units, 298
Logistic concepts, 106
Loop computations, 48–49
Lorentz, H. A., 296, 310, 311
Lunar guidance, 325–326

Machine tools, automation, 110, 118
Malthus, Thomas R., 116
Management science, 65
Man-machine systems (*See also* Compromise automation)
  airborne, 166–169
  allocation of duties, 297
  central control computers for, 166–169, 346
  electromyographic techniques in integration, 345–360
  exploration of extraterrestrial bodies, 115–116
  safety procedures, mechanization of, 268–292
  space-oriented, 295
  tunnel display concept, 361–377
Mapping, 129, 291
Marginal programming, 87
Mark I design, 38–45
  growth potential, 40
  information flow, 39–40
  kluging, 44
MASER (microwave amplification by stimulated emission of radiation), 303, 310, 314–315, 318, 319, 320–321
Mathematical analysis, 26
Mathematical model, 3, 226–229
Maxwellian electromagnetic theory, 51, 310
Mechanization (*See* Automation)
Mellon, B., 99

Memory:
access times, 27
capacity, 112
devices, 60, 87, 299
high capacity random, 298
magnetic drum, 11
man-machine comparisons, 73
serial access, 298
Mental activity complex, 68–69
Microelectronics, 125
Microhandling, 359
Microlock, 31
Micromechanics, 125
Microminiaturization, airborne computers, 57
MIG planes, 21
Milling machines, automated, 117–118
Miniaturization, airborne computers, 57
Missiles:
ballistic, subsystems, 295
firing problems, 235
guidance computations, 41
heat seeking, 291
Modulation, radio and wire systems:
continuous wave, 245
pulse techniques, 245–246
sidebands, 246–247
single-sideband frequency, 249
suppressed carrier, 246
types, AM *vs.* FM, 249
Muller, H. J., 116
Muscle control, artificial, 358

Navigation:
astronautical systems, 296, 298
computations, 155
Doppler radars, 342–344
inertial/radio-fix, 112
space, 302–341
Needle electrodes, 349, 351
Neurotic tendencies, in systems, 120
Newton's law, 5
Noise, 28–32
definition, 29
elimination of, 30
filtering, 30–31
sources, 30
Noise bursts, 253
Noise spectrum, 30–31
Nuclear instrumentation, 367
Nuclear techniques, agricultural uses, 109
Numerical analysis, 26

Office of Naval Research, 362
Operational approach, 91–106
operational check-out, 105–106
system management, 92–102
systems support functions, 103–106
Operations research, 65, 337
Optical system, 341
Optimal design, 16–18
Optimal policy, 138–139
Optimality, principle of, 139, 142–144
Organization design, 94
Organizational automation, 111–113, 118–122
military aspects, 119–120
telephone system, 121–122
Output:
defined, 3
information flow, weapons system, 152
nature of, 4
Output argument, 130
Output space, 130

Partial automation, 122–124 (*See also* Compromise automation)
Peano's successor function, 7
Petroleum and refining industry, automation in, 121
Phenomenological devices, 22, 32, 51, 125
Polarization, ionic, 351
Prediction, 33–37
Production lines, automation of, 112
Programming:
of computers by remote control, 247
for cost estimation, 99–100
dynamic, 137, 138–144
defined, 138
optimality, principles of, 142–144
processes, 141–144
terminology, 139–141
flight management, 115–116
functional, 137
heuristic, 85, 86
linear, 137
marginal, 87
micro-programming techniques, 299
quadratic, 137
Prosthetic surgery, 116

Q guidance scheme, 325
Quadratic programming, 137
Quads, 254–263
Quality control, 104–106
Quickening, 30, 79–81, 113, 346, 365

Radar:
for astronautical systems, 296
Doppler (*See* Doppler radar)
gun acquisition of (flow diagram), 234
introduction of noise, 29
multi-channel tracking, 145
PPI (plot position indicator), 30
Radio-frequency generator, 319
Radio systems, 243–249
Real time systems, 26–27, 28
Redundancy:
bundle-type logic, 299
codes in digital computers, 59–62
cybernetic machine, 86
in guidance schemes, 325
low-level, 59, 253–263
order code, 61–62
sensory, 69
subcomponentry, 88
superposition of response, 87
Regret, 136, 137, 266
Roberson, R. E., 335
Robot, 359

Safety procedures, mechanization, 268–292
Savage, L. J., 264
Schedules, 97–100, 300, 301
Scintillation, 29, 367
Selection, 13–18
Selectivity, 352–353
Self-checking, 59–61
Self-organizing systems, 86
Self-repair, 88, 90
Self-test, 87, 90
Semiconductors, 22, 32, 263, 299
Sensors, 31, 249
Sensory channels:
human perceptual abilities, 66
human sensory abilities, 65–66
man and machine ability, comparison, 72–73
Set theory, 128–129
Sherrington's law of reciprocal innervation, 356
Signal:
bandwidth, 30
bypass, 231
definition, 29
electromagnetic, 310, 313
filtering, 338
and noise, 28–32
Signal-to-noise ratio, 28, 30–31, 322, 333, 342, 343, 344
Simon, H. A., 264
Simplex decisions, 66, 76
Slack, 97, 98
Slaving, 334, 338
Smith, D. S., 321
Social Security System, accounting, 77, 113
Sorting, 66, 250
Soviet Union:
MIG's, 21
prosthetic surgery, 116
Space:
colonization, 116
exploration, 109
manned flight, 126
radiation areas, 32
vehicles, manned, 112–113
Steel industry, automation in, 121
Stochastic processes, 341
Stochastic variables, 3, 6
Storage:
flip-flop, 7
human memory, 75
routine, 251
Suboptimization, 17–18, 21
Subsystems:
additiveness, 10
capabilities, 50–51
and components, 9–12
human, 64–81
multiple navigational sensory, 31–32
projection, 11–12
radio-fix, 48
troop deployment practices, 10
Supersystems, 9–12
Support functions, 103–106
defined, 103–104
documentary material, maintenance, 106
logistics concepts, 106
operational check-out, 105–106
quality control, 104–106
Surgery, prosthetic, 116
Synthesis, 11–12, 90
System management, 92–102
Advisory Committee, 93–94
commercial systems, 101–102
constraints, 100
control of program, 93
cooperation with Advisory Committee, 96–97
delegation and communication, 100–101

System management (Cont.)
directive triumvirate, 93, 94, 102
documentary material, maintenance, 106
information flows, 94–96
operational check-out, 105–106
responsibilities, 94
schedules, 97–100
Systems:
analysis, 24–37, 325, 337
automatic checkout, 113
basic concepts, 2–8, 129–131
classification, 4
communications network, 10
complexity, 54–81
control, 82–90
definition, 3–4, 128–131
description, 4–7, 48–53
design objectives, 22
deterministic, 7
devised, 4, 108
displays, 77–81
hierarchy, 11, 12
hybrid, 4, 26, 70, 104, 108
and missions, 8
natural, 4
neurotic tendencies, 120
one-of-a-kind, 38
real-time, 26–27, 28
reliability computations, 55–61
self-checking, 59–61
self-organizing, 86
simulation, 51–53
states, 3
support functions, 103–106
technology forecast, 107–124
terminology, 128–129
Systems engineering, 8, 12, 124, 125
Systems matrix approach, 40, 165–225

T-33 computers, 268, 292
Tabulation, 98–99
Taylor, F. V., 79
Technology forecast, 107–124
Telegraphy, 247, 248
Telemetering systems, 235
Telephone system, automation in, 121–122, 247, 248
Telescopes, 335, 341
Television, 247, 296
Terminal guidance, 302–304, 305, 325
Tests, 104, 165, 231
Thought processes, 84–90
Three-dimensional display, 366–368
Tunnel display concept, 361–377
cockpit layout, 374
contact analog, 362–364
generalized appearance, 364, 366
information supplied analogically, 370
implementation, 374–377
major problems, 365–366
mechanization equations, 375
relationship of visual axis to display, 371
three-dimensional display, 366–368

Unburdening, 30, 79–81, 113, 346, 365
Unidirectional flow processes, 353
Unified function concept, 346

Value:
expected, 266
human, and system design, 17
scale, 78
in systems concepts, 13
theories of, 15, 126
Value space, 15, 141
Van Allen Belt, 32
Video processing, information flow, 153
Visual displays, 366–374

Waveform analysis, 354, 355
Weapons system, 92–102, 145–164
Weather control and prediction, 77, 109, 113
Wheelon, R. W., 338
Whorf-Sapir hypothesis, 85
Wire systems, 243–249
Wronskian matrix, 257–258

Zero crossings, 222, 332