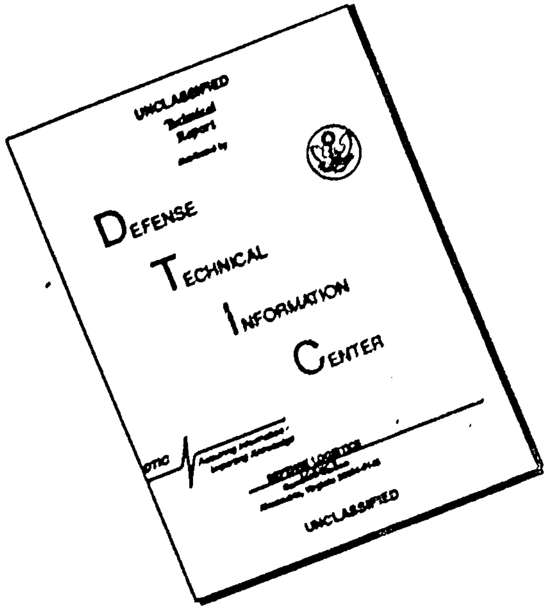
AD-758 977

AUTOMATION IN MANNED AEROSPACE SYSTEMS

Advisory Group for Aerospace Research and Development
Paris, France

March 1973

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

AGARD-CP-114

AGARD-CP-114

AD 753977

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No. 114

on

Automation in Manned Aerospace Systems

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY
ON BACK COVER

AGARD-CP-114

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.114

AUTOMATION IN MANNED AEROSPACE SYSTEMS

Papers presented at the 24th Technical Meeting of the Avionics Panel of AGARD held in Dayton, Ohio, USA 16–19 October 1972.

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Program and the Aerospace Applications Studies Program. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The material in this publication has been reproduced
directly from copy supplied by AGARD or the author.

Published March 1973

629.7.05 : 621-52 : 658.3.04

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St. London. W1P 1HD

THEME

The continuous expansion of aerospace system requirements results in the ever increasing assignment of computing, logic and decision making functions to "on-board" digital computers. The need for man as an integral element in aerospace systems seems likely for the next decade. His role and the proper integration of man and machine for maximum system effectiveness will, however, require periodic re-examination in view of the growing capability of the machine to assume functions previously reserved for man.

This conference considered the current capabilities and potentials for automating manned aerospace systems, and described the "tools" currently available and under development for assigning system functions to man, machine and man and machine, so as to best satisfy aerospace system requirements and constraints.

Four sessions provided broad coverage of this important subject:

1. Definition and Design Considerations examined systematic procedures for the translation of mission into system functions, analysis of the functions to determine feasible means for their implementation, and the assignment of these functions to man and machine, separately and in combination, based upon the respective capabilities of each.

2. Meeting Machine Requirements considered some of the complex system functions generally assigned to men and discussed the machine requirements which must be satisfied in order to successfully automate these functions. The capabilities of current and future "on-board" computers with regard to the performance of logical and decision making functions, learning (adaptive control), flexibility, malfunction detection and compensation real-time control were discussed. The influence of computer architecture, programming languages, computer system organization and the problems of computer integration into the total system, upon automation of aerospace systems was presented. The practical limits to aerospace system automation imposed by system requirements and state-of-the-art computer technology was considered.

3. The Role and Characteristics of Man in Manned Aerospace Systems explored current techniques for describing human performance as a system operator and maintainer. The status of current efforts to quantitatively describe human performance through the use of empirical and theoretical mathematical models was assessed. Qualitative techniques for defining human capabilities in satisfying aerospace mission roles, which are based upon limited experimental data and reflected in human factors guidelines, specifications and practices was described.

4. The Total System described approaches to the selection of a system design in which man and machine complement one another in satisfying the system requirements. Criteria for optimizing the integration of man and machine to obtain required system performance under imposed constraints was discussed. Examples of fully automated and man-machine aerospace systems illustrating many of the topics presented during this technical meeting were presented.

Program Chairman

Dr Edward Keonjian
Chief, Microelectronics and Circuit Design
Grumman Aerospace Corporation
Bethpage, New York
USA

Avionics Panel Chairman

Dr Irving J.Gabelman
Chief Scientist
Rome Air Development Center
Griffiss Air Force Base
New York
USA

Deputy Chairman

Mr Roger Voles
E.M.I. Electronics Limited
Blyth Road
Hayes, Middlesex
UK

CONTENTS

TECHNICAL EVALUATION REPORT

ON

XXIV Avionics Panel Technical Meeting

ON

Automation in Manned Aerospace Systems

16 - 19 October, 1972

by

Edward Keonjian
Grumman Aerospace Corporation
Bethpage, New York, USA

1. Introduction

The meeting was held in the National Cash Register Education Center. Sugar Camp, Dayton, Ohio, USA. The host of the meeting was Wright-Patterson Air Force Base, USAF. The program for the meeting was prepared by a committee chaired by Dr. Edward Keonjian. and consisted of 24 papers. half of which were from Europe. Over 1 persons from seven NATO countries attended the meeting.

On Thursday, 19 October, the participants visited the Missile System Division and Columbus Aircraft, Division of North American Rockwell Corporation in Columbus. Ohio. Enroute to Columbus, a brief tour of the new Air Force Museum at Wright-Patterson Air Force Base was conducted.

The purpose of the meeting was to bring together interested specialists from within the NATO community to discuss the current capabilities and potentials for automating manned aerospace systems.

2. Technical Evaluation - General Remarks

The purpose of this report is to provide a brief summary of the meeting and to indicate where possible developments, conclusions, and areas recommended for further investigation. The subject matter of the meeting covered a wide range of topics which were divided into four general areas:

- Definition and Design Considerations
- Meeting Machine Requirements
- Role and Characteristics of Man in Aerospace Systems
- Total System.

The program did not allow time for a round table discussion in which the relevance of the papers presented could have been assessed and recommendations drawn up as a guide to future work. However, each paper was followed by a brief (5 to 10 minute) discussion, which at times, was very lively, indicating a high degree of interest on the part of the audience.

By a special arrangement with the USAF, two engineers from Saab-Scania Aerospace group of Sweden attended the meeting and were given an opportunity to make an informal presentation on the Electronic System of aircraft model 37VIGGEN. The presentation generated considerable interest and active discussion. As a result, it was decided to include this material in the Proceedings (see paper 8 by B. Sjoberg).

The opening of the meeting was highlighted by the presence of Professor Neil A. Armstrong, of the University of Cincinnati, Ohio. USA. Prof. Armstrong, the first man to set foot on the moon, addressed the meeting with revelations of his personal experiences as an astronaut and shared his view on man-machine relationship in the future space flights. Some film clips from the Apollo program preceeded Prof. Armstrong's address.

The official welcoming address was given by Lt. General James T. Stewart, Commander, Aeronautical Systems Division. WPAFB, Host of the Meeting, followed by Dr. I. Gabelman. Chairman, Avionics Panel AGARD, who described the objectives organization and operation of AGARD.

3. Session Evaluations

Session 1 - Definition and Design Consideration

The session was opened by an introductory paper on man's role in integrated control and information management systems. The paper was accompanied by two short films narrated by the speakers. The paper described a kind of information processing and data management system that could relieve the crew of such tasks as pre flight subsystem checkout and periodic system status checks.

The crew's role in this connection was examined. A prototype generalized display and command technique was described. The paper was an updated version of the work presented by the authors a year earlier at the Institute of Navigation, National Space Meeting at Huntsville, Ala.

The next paper (#2) was devoted to outlining the general guidelines for the design of manned aerospace vehicles. The paper dealt with technical specifications set forth by the Franco-British efforts to insure the safety of the operation of the Concorde aircraft. The paper dealt chiefly with the philosophy, rather than the hardware, of reducing crew workload by a precise study of pilot behavior. The philosophy however was not purely theoretical and was based on a number of simulations and inflight tests on Mirage III, Extendard IV, Caravelle, Boeing 707 and on two test beds: a variable stability Mirage III and a variable Stability N262.

The next paper (#3) dealt with the management approach. and the results of the study conducted by the Grumman Aerospace Corporation on NASA's Space Shuttle. with the primary purpose of reducing cost and cost risk of the program. The final configuration adapted resulted in an operational cost per flight of approximately 11 million dollars.

A case history of man-machine system definition was made in the concluding paper (#6) of the session. The paper described the results of NASA's Space Station (SS) Definition studies completed by North American Rockwell Corporation. The information was summarized in a standard format so that the potential users from NASA and other contractors could critique. discuss and suggest alternatives in a common meeting.

The study resulted in the definition of an Information Subsystem consisting of a unique combination of multi-processing computation, internal data distribution via a digital data bus, crew interfacing via a set of multi-processing display and control consoles, and external data distribution via a combination of VHF, S and K band links.

Session II. - Meeting Machine Requirements

This session was chiefly devoted to the complex system functions generally assigned to man, and to the means for designating these functions to the machine. In this light the session discussed the capabilities of current and future "on-board" computers, including the logical and decision making functions of the computers, learning flexibility, malfunction detection, and compensation real-time control. Papers # 7, 10 and 15 highlighted the session.

In Paper 7, "Automated Techniques for Spacecraft Monitoring" (ATSM), the author addressed himself to the study of the problem of efficient and reliable spacecraft monitoring through automation for future spacecraft. Indeed, a very timely subject. The study program was implemented within the configuration of the Real-Time Computer Complex, which is the core of NASA's Mission Control Center at Houston, Texas. The test bed used telemetry to perform selected flight controller monitoring functions for Apollo missions.

The automated monitoring. described by the paper represented a new concept of mission ground support, whereby system specialists. freed from routine monitoring. could devote their expertise to unprogrammed or contingency situations. Furthermore. placing automated monitoring programs on board future spacecraft could free astronauts from tedious monitoring and routine spacecraft control.

A simulation technique for a specific spacecraft performing a specific mission (the proposed Resuable Shuttle Booster (RSB) stage) was described in paper #10. "Optimum Spaceborne Computer System by Simulation". The configuration which was developed is considered as an optimum with respect to the efficient use of computer hardware. Further. in an environment where ultimate high reliability is a requirement for some programs. this type of model can be used to determine the effect of executing all programs in a high reliability mode. with the attendant advantage of relieving the Executive System of the task of mixed mode scheduling.

Paper #15. "The Experimental Evaluation of Automated Navigational Systems", described certain aspects of automated avionics systems which are being examined in the Royal Radar Establishment Comet Exercise. A Comet 4 aircraft has been re-equipped as a flying laboratory for this work. The installation described proved to be flexible and capable of being modified or extended with a minimum of effort for future experimental navigational systems.

Session III - The Role and Characteristics of Man in Manned Aerospace Systems.

"Current Status of Models for the Human Operator as a Controller and Decision Maker in Manned Aerospace Systems" was the title of the survey paper # 16. The paper dealt with the human operating modeling resulting from simple laboratory experiments on human hypothesis testing and some recent work in statistical decision theory. The author indicated however, that much more experimental work needs to be carried out before an intelligent allocation of tasks can be made. The findings of this paper may become especially useful with the increased complexity of future systems.

Paper # 17, "Manual Landings in Fog", describes the results of 18 fog-flying sorties using a Category II operation terminated by a manual landing. A wide variety of fog structure and visual sequences were illustrated to demonstrate the lack of relationship between the visual segment at high decision heights, the height at which visual contact is first made and the Runway Visual Range measurement.

Performance in fog, compared with clear weather, was worse during day light that at night. Another conclusion of the paper was that shallow fogs are potentially deceptive in tempting the pilot to land in very low Runway Visual Ranges (RVR's). The flying described in this paper is necessarily a small sample and the author sees the logical continuation being an analysis of instrumented Category II landing in service.

Session IV - Total System

Papers in this session described approaches to the selection of a system design in which man and machine complement one another in satisfying the system requirements.

In paper # 23. "Potential Teleoperator Applications in Manned Aerospace Systems", teleoperators have been defined as extenders and augmenters of man. Some examples of potential applications of teleoperators were given among them: long manipulator booms to be mounted on the space shuttle to transfer cargo; free-flying teleoperators capable of acquiring, inspecting, and repairing satellites in orbit; use of teleoperators for remote control spacecraft or aircraft in lieu of man.

A new cockpit concept for a future, one-man, multi-mission fighter aircraft was described in paper # 24, "Man-Machine Considerations in the Development of a Cockpit for an Advanced Tactical Fighter".

The key elements of this new design concept are: multiple, time-shared electronic displays; keyboard and voice command computer input devices; "wrap-around" cockpit arrangement for ease of access to the control-display devices; an integrated total energy command and a system of dependent automation that permits reduced pilot work load during anomalies. The concept was evaluated by a simulation program airing such objectives as the examination of the potential of voice command as a computer input; evaluation of pilot usability of multi-purpose display concept in support of multi-mode operation, and other related objectives. According to the data given in the paper, the inexperienced pilot (less than 300 hours) did at least as well as those with many more hours. The concept presented stimulated a very active discussion from the audience.

In conclusion it was evident that the subject matter of the meeting and a great majority of the papers generated considerable interest on the part of the audience which manifested itself in many active discussions in this successful meeting. It is the opinion of the writer however, that in view of the termination of the Apollo program and the increasing emphasis on total automation and the use of unmanned probes for exploration of the universe, a technical meeting on Automation in Unmanned Aerospace Systems would be an appropriate subject for one of the future technical meetings of the Avionics Panel.

MAN'S ROLE IN INTEGRATED CONTROL AND INFORMATION MANAGEMENT SYSTEMS

by

J.L. Nevins & I.S. Johnson

Charles Stark Draper Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts

I. INTRODUCTION

Display and control techniques for avionics and large process control systems are undergoing dramatic changes as a result of three major factors: (1) demonstrated performance of well-designed systems that include non-redundant airborne processors with a mean-time-to-failure of about 40,000 hours (ref 1.) (2) the design of even more complex systems by applying this kind of component technology to computer systems organized to be fault tolerant and gracefully degradable (ref 2); and (3) availability of flight qualifiable general purpose interactive graphical display/control devices with significant capability and flexibility.

The object of this paper is to discuss (1) display control considerations associated with these new techniques, (2) general purpose displays in contrast with previous techniques, and (3) a prototype interactive display/command design presently implemented featuring a pushplate CRT overlay for command input.

II. DISPLAY AND RETRIEVAL OF STORED INFORMATION

General purpose interactive graphical display and control techniques allow more flexibility in the manner data may be displayed and responded to by the operator. With conventional techniques, specific devices are dedicated to each subsystem and the system designer must decide how best to locate the various displays and controls under varying criteria caused by varying mission requirements (data needed for landing and takeoff not necessarily needed for cruise and navigation) and varying system modes (checkout/nominal flight/emergency).

This fragmented approach to presenting information is largely attributable to technology limitations and reluctance to implement new and perhaps marginally verified techniques. Dedicated devices take up panel space according to their functions and are located according to the relative importance, size, and expected amount of use. As a consequence, cockpits are filled with a proliferation of dedicated devices which may be used but once or twice during the course of a mission. This distribution of devices and associated information we call "horizontal" in structure. That is, the instantly available information lies spread out before the operator, access being limited largely by the user's familiarity with the panel format.

Major criticisms of this format include: (1) limited information can be provided; (2) panel space is taken up equally by devices used 10% of the time as by those used 90%; (3) the user has to sort through a maze of switches and dials to retrieve much of his information; (4) flexibility of display format is minimal; (5) danger of actuating or interpreting wrong device at the wrong time is greater than if unused devices could be "stowed" in some sense. A major desirable feature is that information is immediately available at the flick of the eyes, particularly important during emergencies.

Newer techniques offer an alternative by permitting the information to be stored in a more "vertical" structure, i.e., currently required information is the only data provided; the unneeded data is kept out of sight. This is done by time sharing a device capable of displaying any one of a multiplicity of display formats. These devices have been referred to as "generalized" displays and can help fulfill the objectives of minimizing weight, space, component proliferation (cockpit clutter), and maximizing reliability.

In the effort to generalize displays and controls, great care must be taken to minimize the effort required to extract "stowed" information, particularly in emergency contingencies. These new systems may be less forgiving than the more fragmented approach because the "unneeded data" is stowed. In addition, certain information is to be dedicated in any case, and reasonable algorithms should be developed to determine which data fall in this class. Suppression of the unneeded data can be done in a variety of ways. Determining which technique to use is a critical problem. Primary factors in dedicating devices are frequency of use and required immediacy of response, particularly in crew safety and mission success contexts. Practical solutions appear to include a judicious mixture of "vertical" and "horizontal" information presentation.

Our design activities in developing command-control techniques for complex systems have primarily concentrated on potential space-borne applications. The hardware/software design objectives for this work include:

A. Enable crew performance as system supervisor of the following kinds of tasks.

1) Mission Sequence Initiation
2) Specification of data, constraints, performance options
3) Definition of system objectives
4) Overview of automatic system sequencing, control, and status
5) Reconfiguration of subsystem interrelations
6) Performance of specific backup tasks in case of hardware/software malfunction;

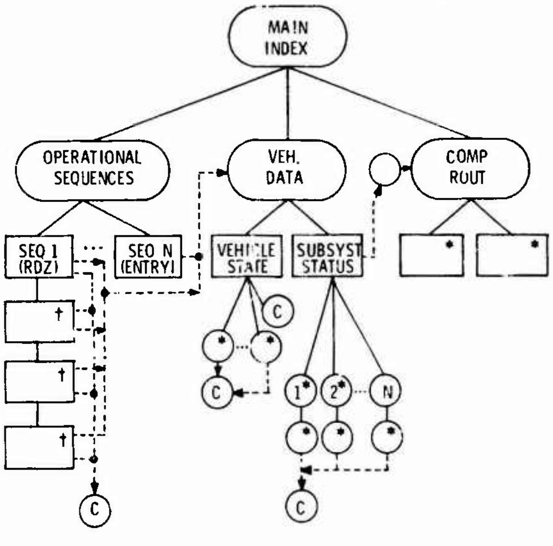
B. Enable significant real time (or updated) changes of displayed information without requiring alteration of the basic software display file length and structure (this places requirements on the design of display generators which drive reformattable display devices);

C. Provide an interactive crew/computer interface scheme which will require a minimal amount of dedicated hardware, which is easy to learn to use, and provides rapid access to all information available in the system.

A major objective is to devise an interface information flow sufficient to enable rapid, accurate crew appraisal of the system state, and to enable the operator to override automatic systems when his heuristic, nonliteral decision processes demand it. The system should be designed to force some minimum crew attention to the vehicle to ensure adequate operator takeover capability as well as effective flight control.

For any airborne computer system in the foreseeable future, there will be strategy limitations. That is, we will not be able to predict all possible events to be encountered (except in the most finite of systems and environments) and we will not be able to optimize decision logic for all options opened to the crew. These limitations are derived largely from limited knowledge in planning an operational algorithm and limits in computer technology (speed and resolution). Hence, data regarding systems' status and environment factors will have to be readily available at all times, and the data will have to be structured so as to maximize the ease of interpretability. This means integrating data such that the crew is assisted in perceiving key relationships among data, but without obscuring other relationships that may not have been foreseen "a priori".

A basic assumption is that the crew is in ultimate control i.e., the system will not take him where he does not want to go. The crew is the "flight controller" and the command-control system is his tool to permit efficient achievement of the mission/task objective(s). The question of how much automation is too much in a given instance can be largely a value judgment, depending on crew confidence in the instrumentation and general philosophy on the degree to which the system should be controlled manually. Hence, flexibility in utilizing automatic and manual modes in these difficult-to-resolve areas is mandatory. The automated functions should be implemented to appear to operate in a rational or reasonable manner. This and reliability are probably the two prime elements in fostering high user confidence in the system.



* Returnable to Main Index or to Calling Display

† Returnable to Main Index
Can call Vehicle Data or special computation routines

Figure 1 General Crew/Computer Interface Structure

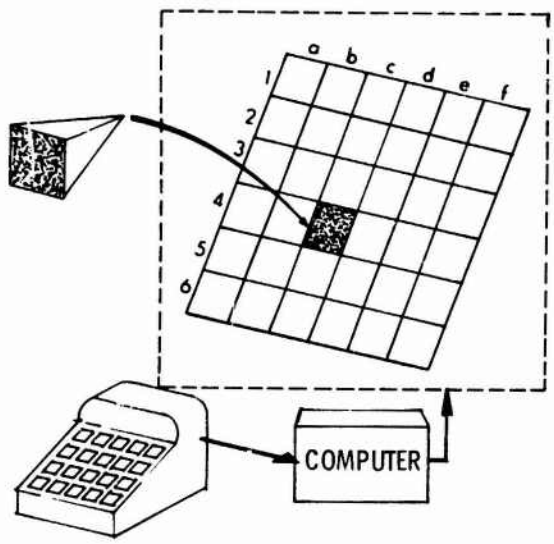


Figure 2 Stored Display Formats

III. POSSIBLE COMMAND-CONTROL TECHNIQUES

A. Data Presentation

For spaceborne applications, we have found it useful to divide the information required by the operator into the following functional units: (1) major mission phase sequences, providing the overview and macro-command capability for mission phase progress, including comparison of alternative solutions provided by the integrated computational system as available; (2) vehicle state, a subset of which will normally be key parameters in cueing the operator to the status of mission sequences; (3) subsystem(s) status and reconfiguration data/control. Some of the information implicit in the latter two categories may not be of sufficient priority to be kept up to date at all times. For this reason, a fourth group of information may be considered—special computation routines which will be available on call to the operator to provide vehicle and systems state information in special circumstances.

Considering the general structure of the operator's information requirements and the nature of its storage (spread about in several subsystems and separate computation compartments), we have chosen an interface scheme as diagrammed in its general form in Figure 1. This is a tree-like structure with links across branches, enabling call of special data from the midst of (and return to) major mission sequences. The operator can recycle to the head of a sequence and to step "forward" through data display sequences. No piece of data is more than four keystrokes removed from a main line display.

Other techniques which have been considered include a simple callable master index frame where all the data pages of interest would be callable by a coded indexing scheme (Fig. 2). Another technique is a two level identifier coupled with arbitrary numeric code for each option or data set (as used in the onboard Apollo programs) (ref 3 & 4). A third technique would be a hierarchical structure incapable of indexing between the lower elements of the tree branches. Such a technique has been used in library retrieval schemes.

B. Command Input

Several alternative input techniques have been considered, including various keyboard arrangements. It was decided to implement a transparent display screen overlay enabling the operator to point at a desired option drawn on the display. This overlay design was motivated by the desire to pursue the concept of maximum crew/display/command integration to its reasonable limits. Other constraints to be satisfied include:

1) Preclude the need for extra equipment to be handled by the operator (e.g., light- or sona-pens);
2) Don't reduce display image quality;
3) Create keyboard "feel" to extent possible;
4) Enable crew operation in pressurized garments.

The overlay "keyboard" consists of an 8 x 8 matrix of focussed light beams at one inch intervals sensed by photodetectors, all of which are mounted on the periphery of the display screen. Breaking a pair of the light beams by pointing at the display and pressing a pushplate about 1/16" toward the screen constitutes a "keystroke" input.

C. General Description of the Prototype

1. The prototype interactive control/display system has been fabricated to be used as a tool for evaluating alternative reformattable command/display techniques. Earlier reports have presented detailed descriptions of the hardware underlying this effort (ref 5 & 6). Briefly the basic components are (Fig. 3):

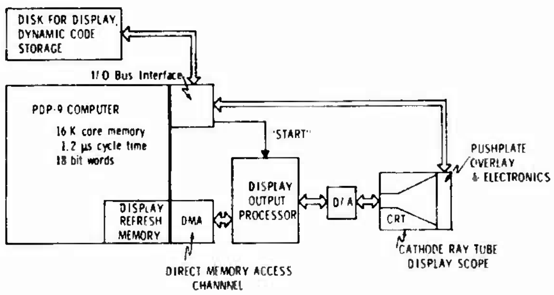


Figure 3 Display/Command Prototype System Elements

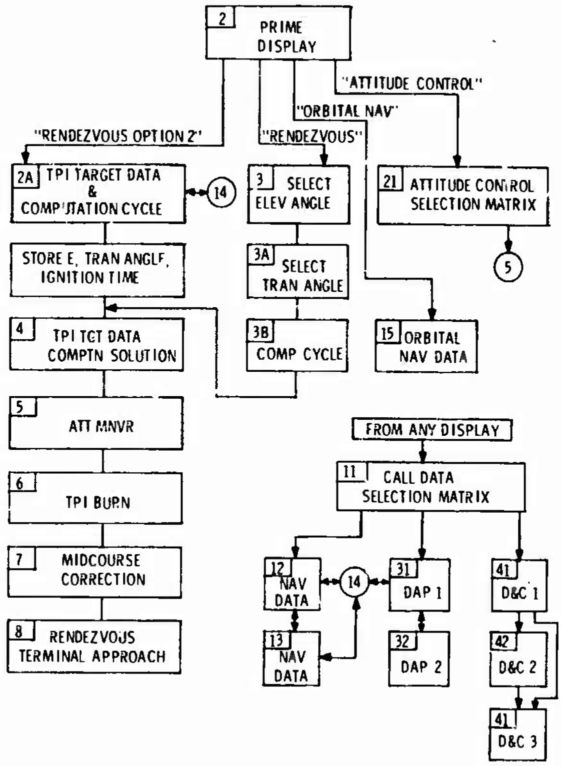


Figure 4 Prototype Display Sequence

a. Display Output Processor (graphics generator), an 8K x 16 bit 2 us instruction time, read only memory list file processor, the output of which defines the beam position and intensity on the CRT. Nested subroutining and indexing and hardware rotation permit direct means to dramatic changes in display formats with minimal coding and significant reduction in execution time. All display elements are created with straight line vectors. Display segments may be altered in size, intensity, by flashing on/off, and rotated.
b. PDP-9 computer (18 bit, 16K, 1.2 us) part of whose memory is used by the DOP for refreshing the display; also stored here are the executive operating program, a simulated environment, and coding to update the displays.
c. A 12" square Kratos CRT (P7 Phosphor) modified with a 3-bit intensity input is the display device. Display formats are 8" square.
d. A transparent CRT overlay pushplate is the input device, described above.

2. The programmed interactive control scheme in its general form is shown in Fig. 1. Specific displays and their interrelationships are diagrammed in Fig. 4. This approach has been taken to satisfy many of the considerations discussed earlier, and to explore possible useful techniques for data display and sequence control in the context of the space shuttle. A more detailed description of the programmed system's salient features follows.

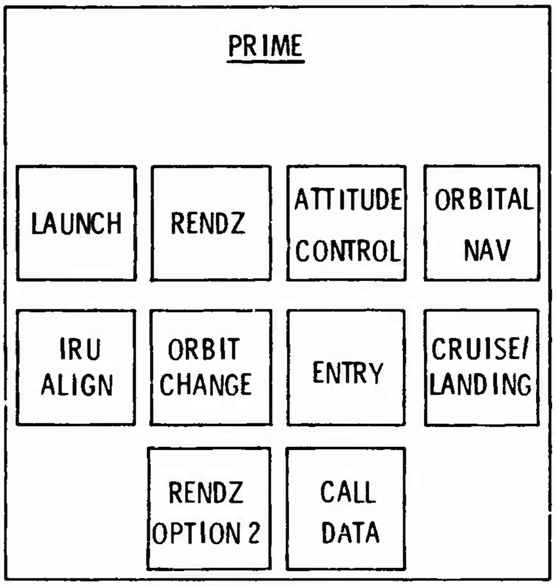


Figure 5

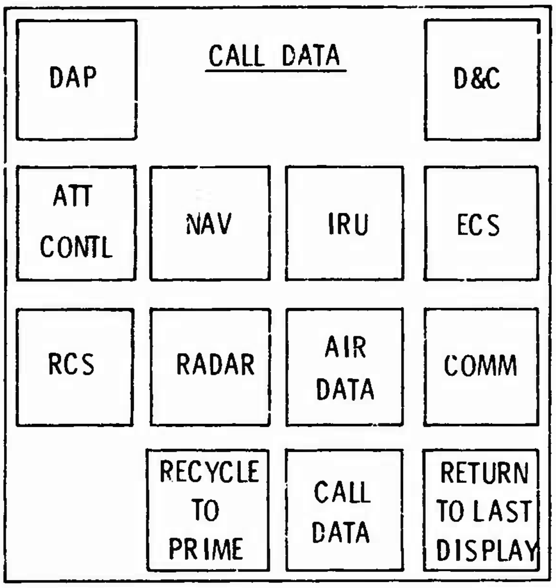


Figure 6

The main index or "Prime" display (Fig. 5) provides a selection matrix of mission phase functional sequences, any one of which the operator may select with a single finger stroke. Options to review vehicle state or to call special computation subroutines, as well as return to the prime display, may be exercised at any time. Having entered a sequence, the subsequent displays provide all available options as the operator progresses through the mission phase. Nominal options are designated by underlining and flashing the square "key" associated with the option.

If the "Call Data" option is selected, another index of options is provided (Fig. 6). This consists of an array of subsystems and functional units under which a library of data is stored in memory. These show fixed and dynamic data, representing vehicle state and/or subsystem(s) status. In all of these operations a miskey by the operator may be corrected with a single rekey stroke.

The Prime or sequence display from which the "Call Data" matrix is called is scaled down and stored in the upper left of the CRT screen. The intent is to keep the operator informed regarding where in a sequence he is, rather than relying on his memory. This is particularly attractive in high stress situations or if the operator must direct his attention elsewhere having entered the Call Data block.

All data are displayed in base 10 only and all formats include word length and decimal point location. More than one parameter is assigned to each display format. For those cases where crew alteration of the data is required, an up/down pointer cursor and numeric calligraphic keyboard have been implemented. The cursor may be moved one parameter at a time with a single keystroke. In those cases where a data category contains too many parameters to fit conveniently in one format, the cursor may be used to "turn the page" when the top or bottom of the current list is passed.

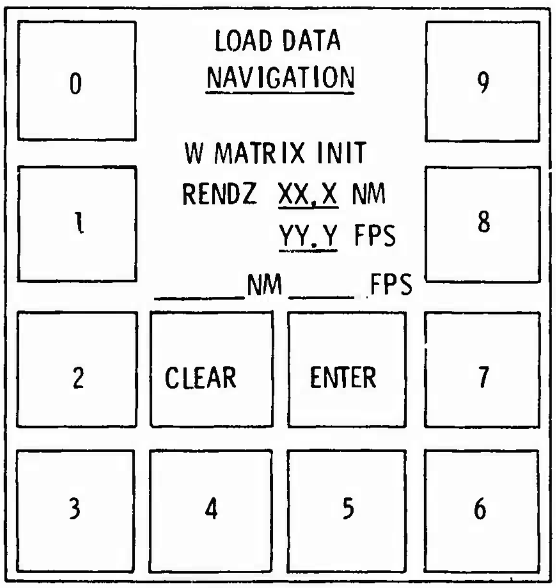


Figure 7

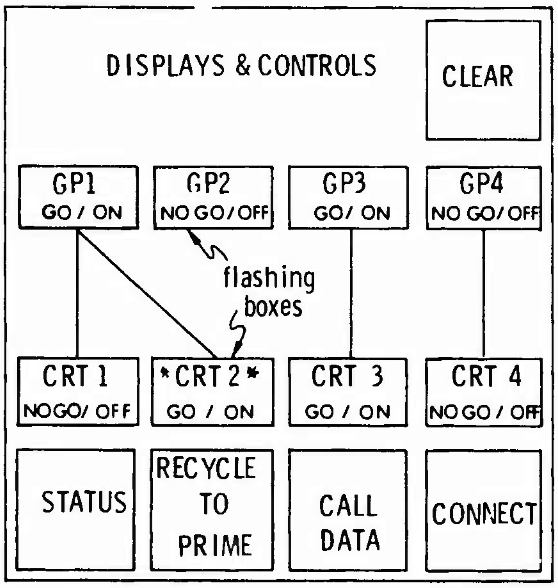


Figure 8

A calligraphic numeric keyboard (Fig. 7) has been implemented to enable alteration of data associated with the Call Data lists and other data alteration. As the data is keyed in, it appears in the buffer zone of this display. Not until "ENTER" is keyed is the computer erasable memory location(s) altered or the display changed on the data page. Miskey may be corrected by hitting "CLEAR".

Also available from the Call Datamatrix is the ability to reconfigure multiple redundant components of a simplified display subsystem. This display/command format (Fig. 8) enables operator control of CRT/graphics generator connections as well as control of on/off status and parameter review with a minimum of pushplate keystroke activity.

This overlay technique provides a simple means of providing a representation of on/off switches. An example reaction control configuration for a digital autopilot similar to that on the Apollo CSM is shown in Fig. 9.

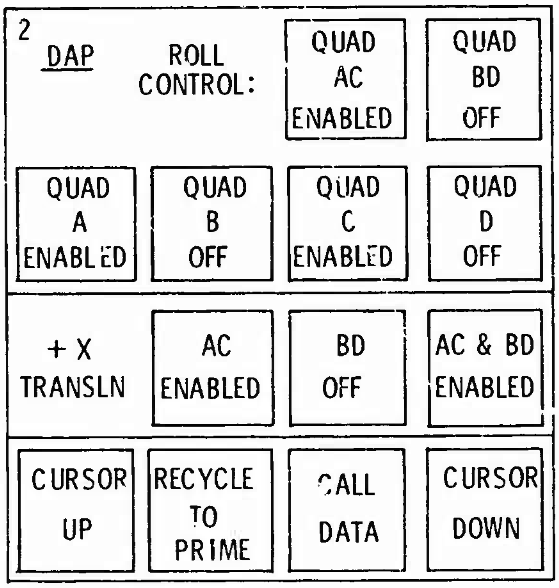


Figure 9

As experience is gained with this system, those functions which could be profitably dedicated to their own input key are being identified. These include "Recycle to Prime" and "Call Data", functions useful at all displays. In an expanded operational context such functions as engine control (on/off) and other crew/mission safety items should be dedicated and probably redundantly hardwired around the central processor(s).

Solid state keys mounted at the periphery of the CRT are another configuration under investigation. This would reduce the number of available options in any given display, but experience to date has not demonstrated the need for all those available with the pushplate overlay. The solid state keys have a superior subjective feel and eliminate the tendency to cover up the option being selected as it is keyed in. A mix of overlay and solid state keys and reformattable solid state keys are other configurations being investigated.

Critique of Present Design

The prototype interactive control scheme as fabricated offers advances over conventional designs including:

1) Direct, obvious relation between the information displayed and choice of executable options.
2) No shift of focus of attention between display of information and execution of choice.
3) Fewer keystrokes to achieve retrieval of data.
4) Data retrieval without using numeric codes as symbols for subsystems, data, sequences, etc.
5) Performance of subsidiary tasks/displays without erasing present main sequence display from CRT screen.
6) Inclusion of data associated with a large number of conventional dedicated display devices.

The general structure of the scheme is sufficiently intuitive that learning problems to date have been minimized. The design can be augmented by including teaching aids as required, or integrating more traditional check-list type material into the displays as desired. Further perfection of the mechanical reliability of the overlay and reduction of viewing paralax caused by the distance between the CRT screen and the pushplate should be pursued, however.

Some of the potential difficulties with this kind of design include:

1) Designer's inability to anticipate all data parameters which might have to be immediately available to an operator in an unforeseen emergency may result in undesirably long sequences to retrieve a piece of data or establish alternative active configuration(s). This can be alleviated perhaps by structuring the data network in a more horizontal format—use of more display screens.

2) Not all users may find a particular sequence to retrieve a data point stored at a lower level very obvious. Probably care should be taken to design the system to enable retrieval of particular bits of information by more than one route through the structure.

3) In unanticipated situations where many parameters are having to be reviewed virtually simultaneously, forcing the operator to switch back and forth among displays may prove unacceptable. This can be alleviated by suitable display device redundancy, which will probably be forced by hardware reliability considerations, in many cases.

4) The possibility exists that even though the time and effort to reach data at the nth level may be small compared with more traditional techniques, it may be more irksome to the user because of perceived remoteness of the data. It may prove that having to look up a code in a book and key it in is more satisfactory subjectively than being led down a 3 or 4 or more level path.

V. OTHER POSSIBLE APPLICATIONS

PROCESS CONTROL

Process control facilities are usually highly automated systems which perform quite precise and rather narrowly-defined tasks. A human monitor/operator is appended to deal with those contingencies which are so diverse and occassional that it is unreasonable to automate the system to deal with them unaided (ref 7).

The human typically scans a handful of critical parameters to verify acceptable performance. The operator is required to perform at a reasonably complex interface and deal with data immediately at hand rather than rely upon repeated practice sessions dealing with the same contingency. He must perform under stress and be able to control the timing of his actions. To deal with boredom, wandering attention, and time delays in familiarizing himself with the current state of an aberrant system, non-static information and new forms of communication need development. This kind of situation in which operators must act quickly on very occassional bases requires communication schemes which will refresh the operator's understanding of his available options and help him judiciously filter out that information which is not immediately relevant.

COMPUTER AIDED INSTRUCTION

The requirement for immediacy of student-system communication is critical and obvious. The goal of minimizing the amount of relatively complex keyboard activity to be learned just to start to use the system is clear. The ability to present all and only those options pertinent to a particular question or subject is highly desirable. The ability to jump readily to another learning sequence as ideas germinate is one of potential attraction. The generalized display/control concept described here goes a long way to fulfilling these desirable characteristics in a self-paced instruction context, although forced paced instruction/testing may be desirable under some circumstances..

MEDICAL PATIENT HISTORY/SYMPTOM SELF-DESCRIPTION

To the extent that the patient can describe his symptoms and past medical history without direct contact with a doctor and without undue time spent and incidence of error, we will have reduced one significant part of the doctor's burden. This application is similar to the computer aided learning situation. The patient is asked a series of basic questions about his health status, the answers to which lead to logical follow-on questions. The doctor may build upon this basic framework with personal communication with the patient.

A touchpoint overlay for the patient's information input eliminates the need to know how to type or use other mechanical devices. The digital storage of the data eliminates problems associated with deciphering of handwriting. Such a system has in fact been tried in a real medical environment with promising preliminary results (ref 8).

Provision for definition of unfamiliar terms, forcing the user to acknowledge understanding of the question and/or key terms are elements which should be included in this and the student learning context.

VEHICLE DISPATCH

Vehicle dispatch operations typically consist of one or more dispatchers responsible for routing and distributing vehicles to need points as a function of constantly-changing real-time demand. The ability to oversee the total operation and to zero in on the status of particular vehicles and neighborhood vehicle distribution densities is required. In order to help dispatchers associated with police, taxi, trucking, fire, and ambulance facilities to communicate effectively with a useful data retrieval system, elimination of keyboards appears to be a particularly attractive characteristic. These operations appear to rely on horizontal data presentation; to the extent they do, computer overlay interface compatibility with large scale displays will be important elements in a useful system design.

SUMMARY

We have not yet addressed ourselves directly to the problems of manual override of automatic functions and strategy limitation consequences in this new context. A full-scale simulation including rather complete implementation of representative subsystems operating in time-critical mission phases with failures introduced is required for a thorough evaluation of such factors. Further effort to quantify the differences between the present and alternative designs is planned. Also, investigations of the applicability of such a pushplate interactive control scheme to other fields of application are being pursued.

REFERENCES

(1) Hall, E., "Reliability History of the Apollo Guidance Computer", MIT Draper Lab Report R-713, Jan.. 1972.

(2) Hopkins, A. L., "A New Standard for Information Processing Systems for Manned Space Flight", MIT Instrumentation Lab Report R-646, September, 1969.

(3) Green, A.I., & Filene, R.J., "Keyboard and Display Program and Operation," MIT Instrumentation Lab Report E-2129, May, 1967.

(4) Nevins, J.L., "Apollo, a Transition in the Art of Piloting a Vehicle," MIT Draper Lab Report E-2476; Navigation, Journal of ION, Fall, 1971.

(5) Holford, S., et. al., "Development of a Generalized, Interactive Display and Control Concept, MIT Draper Lab Report E-2530, July, 1970.

(6) Nevins, J.L., & Johnson, I. S., "Man's Role in Integrated Vehicular Information Management Systems," MIT Draper Lab Report E-2564; Navigation, Journal of ION, Spring, 1972.

(7) Carbonell, J. R., "Artificial Intelligence and Large Interactive Man Computer Systems," Bolt, Beranek, & Newman, Cambridge, Mass.; Nat'l. Tech. Info. Service AD 726-441, July, 1971.

(8) Weksel, W., et. al., "The Automated Medical History," Fall Joint Computer Conference, 1968.

GENERAL GUIDELINE FOR THE DESIGN OF MANNED AEROSPACE VEHICLES

Jean-Claude WANNER

Office National d'Etudes et de Recherches Aérospatiales

CHATILLON - 92 - (France)

Summary

The Franco-British airworthiness autorities have been brought to review the set of technical specifications they intended to require for Concorde in order to insure the safety of the missions of this new transport aircraft. Most of the old rules of thumb generally used for the conventional aircraft appeared as obsolete or unapplicable to a supersonic transport.

In order to guide the definition of these new regulations, a theoretical method was developed for evaluating the reliability of the missions of manned aerospace vehicles.

This method, called E.S.A.U. for "Etude de la sécurité des Aéronefs en Utilisation" (*), is based on an investigation of the way of occurrence of accidents. It has been seen that an accident is due to a set of incidents which can be classified into only three different types. The study of each type of incident, the probability of occurrence which has to be reduced in order to increase the safety, is very useful to help the designer of a new project to choose between possible solutions, taking into account the reliability of the systems, the possible human errors and the flight conditions.

Résumé

Les autorités franco-britanniques de certification ont été amenées à revoir l'ensemble des conditions techniques qu'elles désiraient appliquer à Concorde pour assurer la sécurité des missions de ce nouvel avion de transport.

En effet la plupart des vieilles règles de l'art utilisées généralement pour les avions classiques étaient dépassées ou inapplicables à un avion conçu pour le transport supersonique.

Pour orienter la rédaction de ces nouvelles règles, une méthode théorique d'évaluation de la fiabilité de mission des véhicules aérospatiaux pilotés a été élaborée. Cette méthode appelée E.S.A.U. "Etude de la sécurité des Aéronefs en Utilisation" est fondée sur l'étude des conditions dans lesquelles surviennent les accidents. L'étude de chaque type d'incident, dont nous avons à réduire la probabilité d'apparition afin d'augmenter la sécurité, est très utile pour aider le responsable d'un projet nouveau à choisir parmi plusieurs solutions en tenant compte de la fiabilité des systèmes, des erreurs humaines possibles et des conditions de vol.

(*)(a free translation of E.S.A.U. could be "I.S.A.A.C." for "Investigations on Safety of Aircraft and Crews").

When, at the beginning of the Concorde project, we undertook, with our British colleagues of the Air Registration Board, to set the new requirements of handling qualities adapted for supersonic transport, we thought we should make only a few modification to the BCAR and FAR. Very quickly, indeed, it appeared that the magnitude of the flight envelope, the complexity of the systems and the modern design of the aircraft would lead us to reconsider those requirements altogether and to build a new philosophy of safety based on the study of incidents and accidents.

This philosophy, we called E.S.A.U. in French for Etude de la Sécurité des Aéronefs en Utilisation, which can be translated by I.S.A.A.C.'S, Investigation on Safety of Aircraft and Crews in Service, can be used as a general guideline for designing modern manned aerospace vehicle.

The study of incidents, which may lead the vehicle to an accident or to an interruption of its mission, shows that they can be classed into two categories :

- the incidents which could have been avoided by a modification of the vehicle design, of its technology and of the way it was used and, on the other hand ;

- the incidents which come from the failure of people or material involved in guidance and traffic control.

The incidents of the first type, of which alone, we have to deal with here, occur when a parameter, which characterizes the behaviour of the aircraft or of a part of the aircraft, crosses over a critical value. The origin of these critical values may be aerodynamic (for instance maximum value of the angle of attack), structural (for instance maximum load factor, maximum r.p.m. of the engines), thermodynamic (maximum fuel flow of reheat) and so on.

It is easy to see that a limit may be crossed over after a set of occurrences which can be classified into three categories :

a) The pilot has at his disposal all the controls necessary to maintain all the parameters between the proper limits, but the task is too difficult to fulfil for a human operator, because for instance the frequency of data collection necessary to control the aircraft is too high, or because the pilot does not know very well the relative values of the critical parameter and its limit. Consequently the pilot lets the parameter cross over the limit.

   This type of incident is called Pilotability incident. I had to create a new word, in French as well as in English, because there were no known word for that type of incident.

b) An external perturbation, a gust for instance, or an internal one like a failure, either modifies the value of a critical parameter or modifies the value of the limit itself. For instance a gust increases the angle of attack, an engine failure increases the sideslip angle, a wing flap actuator breakage modifies the limit of angle of attack. This type of incident is called incident due to sensitivity to perturbations.

c) To follow the flight path prescribed by the air traffic control, to avoid an obstacle or to join the flight path after a divergence due to incidents of the two previous types, the pilot has to fulfil a manoeuvre which modifies the values of different parameters. For instance a pitch-up manoeuvre increases the angle of attack and brings it nearer to the limit. This last type of incident is called Manoeuvrability incident (here also I had to create a new word).

   An example will show more clearly how an accident can occur as a result of a set of incidents of the three types.

During an I.L.S. approach without visibility, the stability augmentor systems and the autothrottle having failed, the pilot lets the speed and the altitude decrease and looses fifteen knots and fitfy feet. This is a pilotability incident due to a lack of stability ; the safety margin for angle of attack has thus been already reduced by the loss of speed. Noticing the error in altitude the pilot begins a pitch-up manoeuvre ; this manoeuvrability incident increases again the angle of attack. Andlast a strong gust adds its effect to the two previous increments of angle of attack ; the angle of attack reaches the limit which involves a stall. So a set of incidents of the three types can bring a parameter beyond a limit.

I would like to insist for a moment on the sensitivity to failures. We have to be careful not to confuse the sensitivity to perturbation, in other words, the transient effect of the occurrence of a failure, and the modification of the state of the aircraft after the failure. In this new state the characteristics of the aircraft are not the same as in the normal state, before failure ; so Pilotability and Manoeuvrability can be reduced or Sensitivity to another failure can be increased.

So let us not confuse the transient effect of a failure and the modification of the state of the aircraft due to a loss of a function. For instance the failure of one channel of a multiplex system can give a perturbation to the aircraft at the moment when the failure occurs but does not modify the state of the aircraft, and thence its performance, since the function of the system is still fulfilled by the non failed channels.

As we have seen, the pilot is not involved in the two last types of incident, Manoeuvrability and Sensitivity to perturbations. So the rules that the designer has to follow in order to reduce the probability of accident due to these two types of incidents can be built by mathematical deductions and physical experiments : evaluation of reliability of systems, computation or measurement of the vehicle response to a sudden failure or to a gust, measurement of the vehiclo manoeuvering performance, and so on.

But for the Pilotability incidents the methods are different since it is impossible, for the moment and, I am afraid, for a long time, to represent mathematicaly the pilot.

In order to better understand the pilotability incidents and to evaluate the influence of automatic systems on risks of accident of this type, we have built a model of the pilot, which is not a mathematical model but which intends to detail his way of action on the aircraft.

First let us try to define precisaly the task of the pilot. We have then to give a number of definitions.

We divide the flight into a number of parts we call Phases.

A phase has a general purpose.

For instance the Phase "Climb" has the following purpose :

"From the height of fifty feet after take off fly the aircraft until reaching the cruise altitude, following a given ground pattern".

It is necessary to divide each phase into a number of elementary parts called Sub-Phases, because the task of the pilot is not unique during a Phase.

A Sub-Phase has one elementary purpose ; for instance during the Phase I.L.S. Approach, we can look at the Sub-Phase "final descent", the elementary purpose of this Sub-Phase being:

"Using I.L.S., fly the aircraft in descent, until reaching three hundred feet in good conditions to make a visual landing".

"In good conditions" means here in good position with the correct heading and the right speed. Thus we have to notice that the objective of each Sub-Phase is given with margins taking into account the possibility of performing the following Sub-Phase.

A selected configuration of the aircraft is defined by the position of the different selectors. By selector we mean controls maintained in fixed position during the Sub-Phase. We have to notice that a control can be or not be a selector according to its use ; for instance the pitchtrim and the throttle are selectors during the Take-Off Phase but not during the Approach Phase when they are used in the pilot loop.

A True Configuration is the result of a failure situation on a Selected Configuration.

So for a given Sub-Phase there is one Selected Configuration given in the Flight Manual and a set of True Configurations differing by the number and the type of failures.

The State of the Aircraft is then defined by a Selected Configuration } True Configuration.
a failure situation

a mass of the aircraft

and a mass distribution generally given by the longitudinal position of the centre of gravity.

Parallel to the State of the Aircraft we can define the State of the Atmosphere and, for the Sub-Phases on the ground, the State of the Runway. They are defined by the set of all parameters which can modify the behaviour of the airplane and the behaviour of the crew. For instance, wind, temperature, gusts, clouds, rain, hail and so on. After study we can reduce this set of parameters to only eight for the Atmosphere and five for the runway.

State of the Atmosphere :

- Pressure, Temperature Humidity which act mainly on performance,
- Intensity of Turbulence we can measure by the root mean square of the vertical and horizontal components of gust velocity,
- Temperature gradient,
- Visibility,
- Icing,
- And last, for the Take-Off and Landing Phases, the laws of variation of wind, force and direction, versus altitude.

State of the Runway :

- Length and width,
- Mean slope,
- Profile (in other words, the undulations),
- and last, friction coefficient.

For Immediate Safety the pilot has to observe the aircraft limitations, for instance limitations on angle of attack, in other words, limitations on speed and load factor. He has also to observe operational limitations, for instance height above the ground, altitude or flight level prescribed by the air traffic control and so on.

And last he must not jeopardize the achievement of the following Sub-Phase ; in other words he has to reach the elementary objective of the present Sub-Phase within tolerated margins, position, height, speed, heading and so on.

To reach this double objective, Immediate Safety and Short Term Safety, the pilot uses a Flight Technique as a guide ; this Flight Technique depends on the State of the Aircraft and the State of the Atmosphere during the Sub-Phase. The Flight Technique is generally given in the Flight Manual as relations between the different flight parameters used by the pilot : speed, altitude, attitude angles, angle of attack if provided on the instrument panel.

A Sub-Phase, given by its elementary objective with tolerances, the State of the Aircraft, the State of the Atmosphere and the State of the Runway if necessary, the chosen Flight Technique and the Secondary work define a task.

By secondary work we mean, for instance, radio traffic navigation, reading check-list and so on.

Now having a precise definition of the task of the pilot, taking into account all the factors which influence the flight, we can undertake an investigation of the pilot behaviour.

The data concerning the flight path and the immediate safety, position of the aircraft, attitude angles, angle of attack, speed and so on are provided to the crew in different ways

- some data are directly or indirectly measured and provided on the instrument panel in analog or sometimes in digital form,
- some data are not measured because directly accessible to the pilot, for instance position of the aircraft with regard to the runway in visual landing, linear and angular accelerations,
- and last some data concerning the State of the Aircraft are provided on the instrument display (positions of the landing gear, of the flaps, engines R.P.M. and so on).

All the cues are collected by the different human sensors, which are eyes, ears, arms and legs and the whole body.

We have to notice that the eyes are double sensors : the central vision collects few but precise data and peripheral vision collects numerous but not precise cues. The ears are also multiple sensors : the external ear collects sound and the internal ear collects angular and linear accelerations and direction of the apparent vertical.

It is very important to note that a datum is collected by a sensor, transmitted to the brain and therefore used by the brain if, and only if, the brain gives the corresponding order, in other words calls up the datum.

This remark is very important because it means that collection of different data cannot be made in a simultaneous way ; the brain asks the eyes to look at this instrument, then at another instrument, and after asks the ears to listen to this or that signal ; the scanning procedure, generally learned by training, may be modified by an alarm signal coming from peripheral vision, external or internal ear. But the alarm signal must be more intense when the pilot is more attentive, his work load being high. How many pilots have landed with gearup, without noticing the alarm signal because, the visibility being bad, they were focussing their attention on the view of the runway and other aircraft.

The data being collected by the sensors are transmitted to the brain which, by direct comparison with well known situations or by computations according to programmes stored in the memory, gives two kinds of orders

- an order for action on the controls,
- a call for new data to be collected by a given human sensor.

The first type of orders generates what we call external loops and the second ones generates internal loops. (Fig. 1).

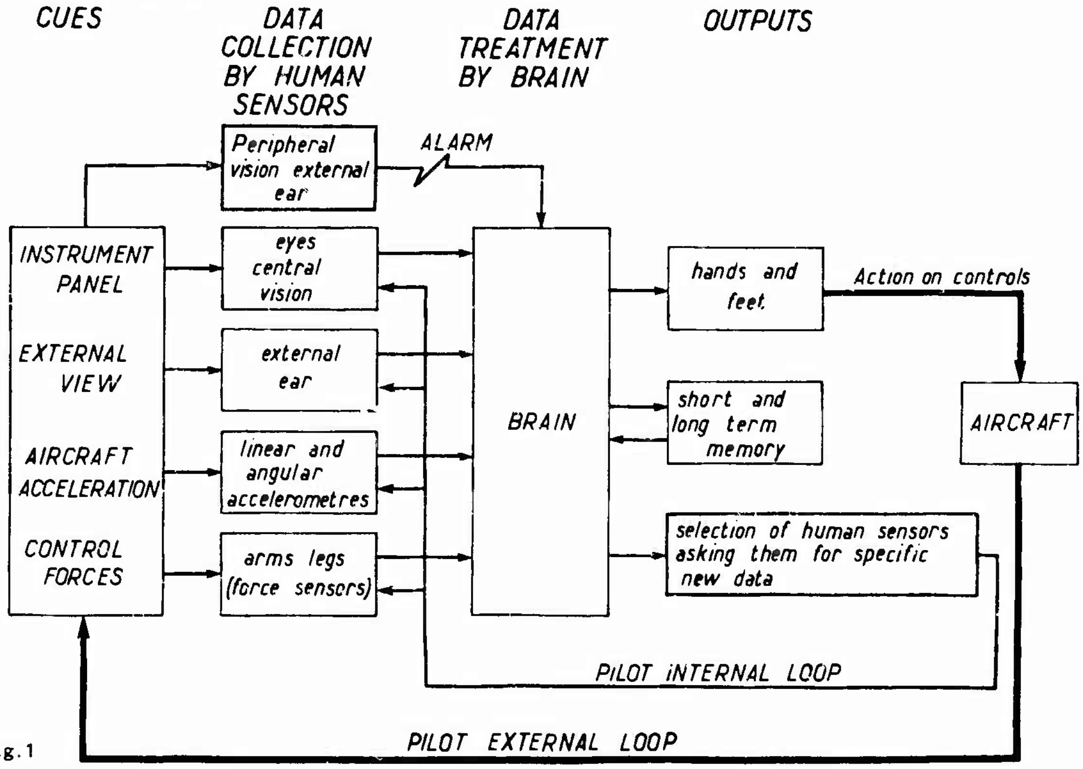


Fig.1

Let us try now to analyse the behaviour of the pilot. The position of the aircraft, given to the pilot by instrument reading or external view, is compared with the position required by the flight technique. For instance the horizontal pointer of the I.L.S. indicator is seen above the central point ; this error is analysed by the brain who chooses the right manoeuvre to perform, in order to correct it. This error analysis demands a sophisticated mental process to transform the reading of relative position of the pointer and the central point into aircraft position error and to elaborate the right corrective manoeuvre, in this case a pitch up manoeuvre with a correct magnitude. This manoeuvre being chosen, the brain, through an internal loop, asks the eyes to pick up data about pitch attitude. The difference between the actual pitch attitude and the chosen pitch attitude is analysed and the brain chooses the right control to be moved and determines the forces to be applied. In our case the pilot pulls up the stick with a force which evidently is not evaluated in pounds, but which the pilot feels to be correct. To perform this last manoeuvre the pilot pulls the stick, asking his arms through another internal loop to transmit force feeling. The motion of the stick is stopped when the pilot feels the predicted force.

So the central brain puts successively in service different loops on asking for new data through human sensors. There are three types of loops. The biggest ones which are relative to parameters concerning the short term safety, in other words, the flight path, position and speed. The second ones are relative to parameters concerning immediate safety, in other words, attitude angles, angle of attack and so on. And last the smallest ones which are the control forces loop. (Fig. 2).

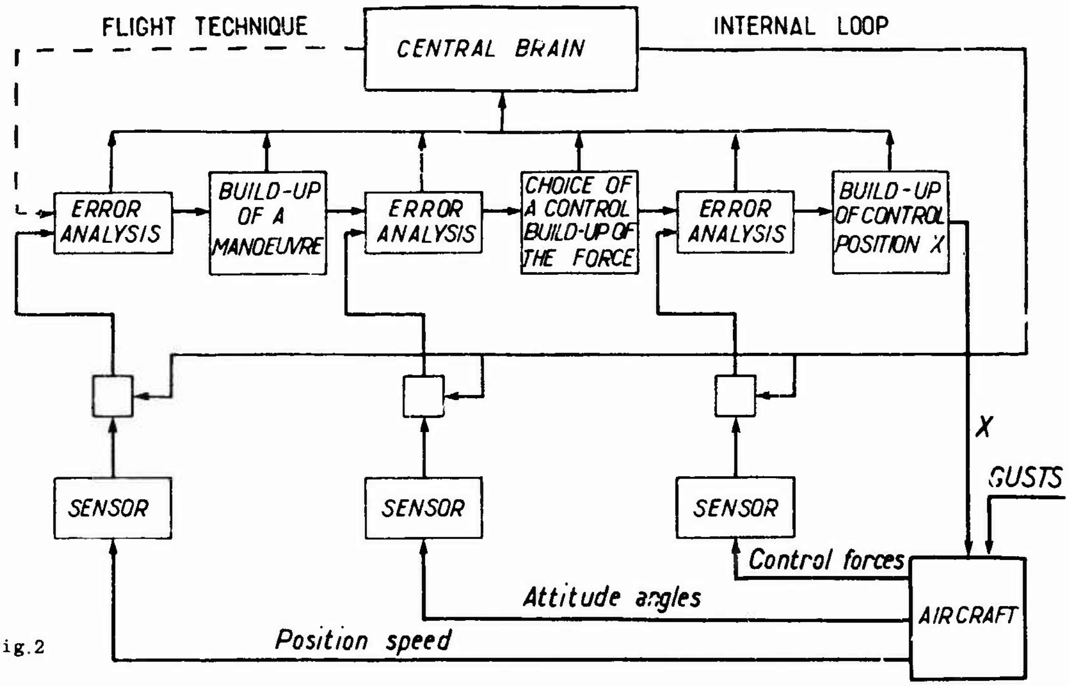


Fig.2

It is very important to note that the entire diagram is in fact much more complex. So it is possible to represent every output parameters, lateral and longitudinal positions of the stick, position of the pedals, of the throttle, of the different trims, of the airbrakes, of the selectors and so on, we may also represent every human sensor and every parmeter concerning the flight. So we could obtain a large number of loops of the three kinds given in the diagram.

Let us notice also that at each moment there is only one loop in service and this is one of the most fundamental differences between pilot and autopilot. The choice of the loop in service is made by the central part of the brain, as represented at the top of the diagram, by an order, through an internal loop, to the chosen sensor to transmit the necessary data.

This model will help us to define the pilot workload and will be a guide when building automatic systems designed to reduce the pilot workload.

The pilot workload, during a given Sub-Phase in fulfilling a task is measured by the number of elementary operations of data collection and treatment described on the diagram.

The immediate consequence of this definition is that we cannot measure directly the workload : it looks quite impossible for the moment to follow in detail the data treatment in the brain.

Another consequence is that it is hopeless to build experimentally a transfer function representing the pilot because there is not one transfer function, even in a very complex form, but a set of transfer functions used successively in an order chosen by the scanning of the different sensors ; this scanning itself depends on the data, on the environnement, on the training of the pilot and so on ; consequently it depends partially on a random phenomenon.

In order to reduce the workload, we have to limit the amount of data that the pilot's brain has to collect and treat in each loop. But we have to be very careful not to reduce the workload in a loop while the workload increases in another one.

The design of the flight director gives a precise example of this type of mistake. The flight director abolishes all the workload due to the two first loops. The pilot has not to analyse the situation of the aircraft, to choose the right manoeuvre nor to watch the attitude angles ; he has only to follow the bars of the flight director ; the loops are reduced to a loop for collection of data given by the flight director and two loops for stick forces. At first view it seems that the workload has been reduced by a large amount. Indeed the pilot, not knowing what the situation of the aircraft is, cannot make any prediction of what will happen in the near future. So he cannot predict what will be the following order of the flight director and then he is obliged to collect the data without interruption. We have transformed the pilot into a pure robot, a bad servomechanism. The workload due to the permanent collection of data is very high since the pilot cannot rest like with a conventional display, with which he is able to make prediction.

It is here the occasion to point out a very important remark : we have to use the pilot's brain in a right way. The human brain must not be used as a pure amplifier, as a servomechanism. Everybody knows that a well trained ape makes a better job than a human pilot does, when the job is a pure robot's job. But the human brain can collect a large number of data, quantitative and qualitative, some of these data being only feelings, can build a model of the situation, compare it with memorised situations and then can take a decision of action even when the case has not been previously predicted.

For instance an automatic system is very good to guide a bomber up to the target, to show the target to the pilot, to place the aircraft in good position for bombing, to guide the bombs and so on. But how to build a servomechanism being able to recognise a red cross of a friendly flag on the target ?

The decision, in this case the decision of bombing, has to be taken by the pilot, his brain being discharged of stupid mechanical jobs and fully opened to not predicted data.

I think that sailors have well understood that point since a long time. The captain takes the decision for modification of heading and speed, but does not handle the helm, the sails or the engine by himself. It is interesting to know that, at least in France, the crew had the same attitude on the big seaplanes we built between 1935 and 1950 ; the captain was in a good position on the upper deck to well understand the situation, having all the necessary information concerning the flight, and there was, if I dare say, somewhere in the plane, a quarter-master who had to maintain heading and climb angle given by the captain.

Not forgetting this very important point, let us look now at the design of a modern cockpit, using not human slaves but servomechanisms.

The most dangerous Flight Phases are the phases near the ground, Approach, Landing and Take Off. Let us see first the possibility of reduction of the workload by cockpit design, during these critical Phases.

In order to reduce the workload due to the analysis of the situation, in other words analysis of the position of the aircraft versus the runway, we shall present the situation not with cross pointers, scales and digit but exactly as the pilot can see it in a visual landing through the windscreen

So a head up display will provide an horizon with heading and slope graduations, and a synthetic runway. Theoretically these informations could be sufficient since during a visual landing the pilot has nothing else except the speed; we shall see that point later; but experience has shown that it is not very easy to make an approach with the right angle of descent. It is easy to maintain the aircraft in the vertical plane of symmetry of the runway but not along the glide path ; indeed the velocity vector is generally in the well known plane of symmetry of the aircraft, but generally also the pilot does not know very well its direction in this plane.

So we shall add a new information in the head up display; the track on the ground of the air velocity vector, which gives the point that the aircraft will reach if the pilot maintains the control in fixed position and if there is no gust, nor wind. The use of air velocity vector instead of ground velocity vector is based on two reasons ; first the experience has shown that the wind corrections are faint and easy to predict; secondly the angle between the air velocity vector and the reference axis is by definition the angle of attack. This last remark has two important consequences: the one measurement of the angle of attack can provide the necessary information to introduce the velocity vector in the head up display. On another hand the angular distance seen by the pilot between the air velocity vector and a mark fixed in the head up, and representing the reference axis of the aircraft, is the angle of attack measured at full scale. Consequently if the mark is fixed in such a position that the angle between the mark and the reference axis is equal to the optimum angle of attack for the approach phase, it is then easy for the pilot to handle the aircraft, by observing the limitations of angle of attack. And last we shall remark that the angular distance between the velocity vector and the horizon is the climb or descent angle γ. (Fig. 3).

Then we provide in the head up the total climb angle γ_t which is

$$\gamma_t = \gamma + \frac{1}{g}\frac{dV}{dt}$$

The total climb angle can be measured by two accelerometers on board and depends on the difference between the thrust and the drag. Therefore it can be handled by the engine throttle and if the pilot, acting on the throttle, puts the symbol of the total climb angle at the same level as the velocity vector, γ_t and γ are equal and the acceleration along the path is null, in other words the speed is constant.

It would be too long to describe in detail how to use the informations so provided in the head up. We shall point out only two remarks.

There has been a large amount of discussion about the comparison of the head up and head down displays, but very often some confusion arose between the two types of displays and the types of informations provided in the head up.

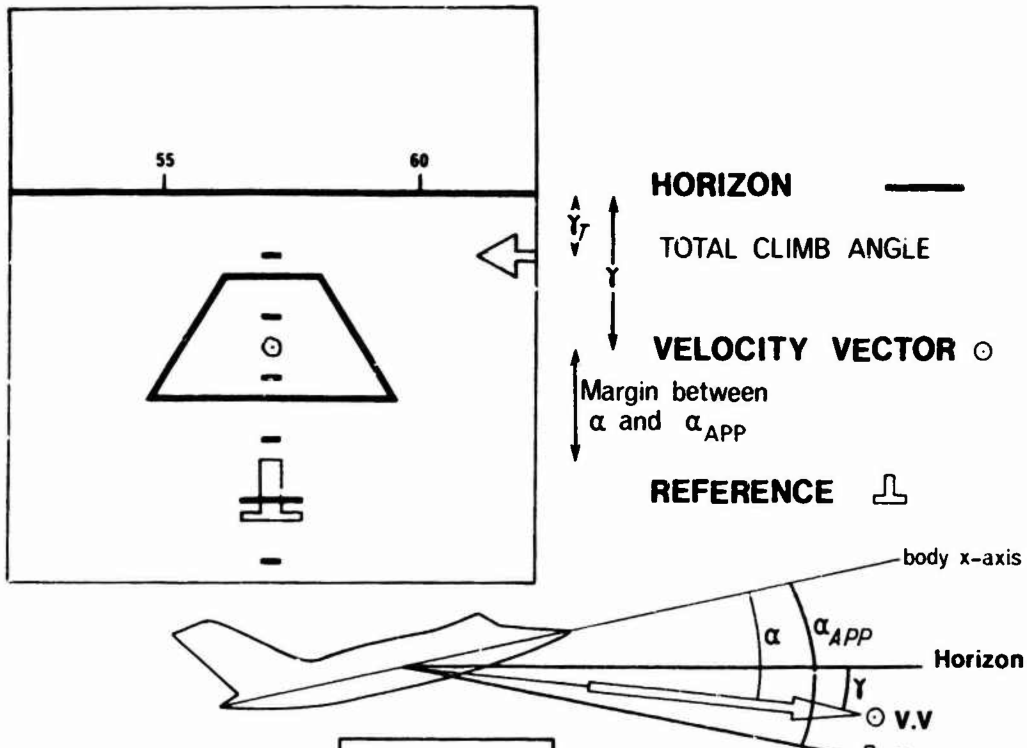


Fig. 3

$$\gamma_T = \gamma + \frac{1}{g}\frac{dV}{dt}$$

Our investigations have shown that when conventional data, like altitude or speed scales, cross pointers are provided in the head up, the pilot cannot at the same time read the data and look at the ground even when the scales are focused at infinity. In some cases during our tests, pilots did not believe that the scales were focused at infinity ; it seems that it is not normal to see a scale at infinity ; so the brain, by reflex, asks the eyes to focus at the normal distance to read a scale and not at infinity, which forces the eyes to focus willingly at infinity and gives this abnormal feeling.

Another very important reason to display, in the head up, the informations related to the external world is based on the following remark. The physiological study of the eyes shows that the eye-balls are maintained in fixed position versus the external world, by a loop using angular and vertical accelerations measured by the internal ear. Every pilot has remarked that during a visual approach with turbulence it is easier to see the runway than the instruments : it is because the eyes are fixed versus the earth in spite of the motions of the aircraft. Consequently if data relative to the external world, like horizon, runway, velocity vector, are provided in the head up, the pilot has no difficulty for looking at the symbols in spite of the vibrations and the motions of the aircraft due to turbulence. This result is exact only if symbols are truly fixed versus the external world. If we provide scales or cross pointers which are fixed versus the aircraft the pilot will meet with difficulties on reading them in turbulence.

On another hand if we provide symbols like horizon and runway in a head down display.the pilot shall have the same kind of difficulty even if symbols are at full scale, because they shall not be focused at infinity.

The head up display as described here will be used either in automatic landings or in manual landings, the pilot handling then the aircraft with a ministick.

If the automatic landing is used daily, which is necessary to reduce the work of airline pilots mainly for short haul missions, it is important to check continuously the autopilot during the landing. The head up display gives all the necessary informations easy to handle and sufficient to make short term prediction : if the velocity vector is on the threshold of the runway at 2.5 degrees under the horizon and if the total climb angle is at the same level as the velocity vector, the pilot knows that nothing very dangerous can happen in the next ten seconds , since flight path modification necessitates application of forces during several seconds. Even in case of sudden failure of the autopilot the pilot has time enough to handle the aircraft by himself, and we know that the informations are sufficient to correctly land the aircraft since we have made a large number of successful true manual blind landings with this type of display.

So with the head up display and the autopilot we have reduced the workload of the pilot since his task is only a watching task, easy to perform.

For manual task we have also to improve the immediate safety loop and the control forces loop.

The conventional stick has been a very fruitful invention about seventy years ago ; it had been then possible to handle directly and together the ailerons and the elevator ; but the stick had to be big enough to enable the pilot to apply the necessary forces counteracting the aerodynamic efforts on the controls.

Now on modern aircraft equipped with servos, the problem is quite different. It is no more a problem of efforts but a problem of flow of information to provide to the controls. Indeed the flow of information coming from the brain and passing through the arm and the stick is certainly lower than the flow of information passing through the fingers. Consequently the stick is no more useful to apply high forces on the controls, takes too much room in the cockpit, hiding a large field of the instrument display, obliges to design a heavy system of artificial forces, sometimes is the origin of pilot-induced oscillations and limits the possible flow of information coming from the brain.

Each of these reasons is sufficient to justify a ministick in the place of the conventional stick. Indeed a ministick is not a reduction of a stick but a device which can be easily handled by the fingers.

Evidently the use of a ministick necessitates to fly by wire ; there is certainly a technical difficulty to obtain a reliable enough system but I think it is mainly a psychological problem which actually blocks its acceptance by the pilots ; we have had exactly the same kind of problem when we accepted to loose the aircraft in case of total failure of the hydraulic system.

Using electric signalling, we can then improve the workload due to the intermediate loop. Instead of acting directly on the controls through pure amplifiers, the ministick will be the input of two autopilots : to a given force on the right or on the left on the ministick will correspond a given rate of change of the bankangle ; to a given fore and aft force on the ministick will correspond a given rate of change of the climb angle. And last a third autopilot will maintain a zero sideslip angle. So without efforts on the ministick the bank angle and the climb angle will be constant and their commanded rates of change will be independant of the flight conditions.

This paper is too short to look at the effect of this new design of the cockpit for the other phases, but it is easy to shaw that the workload is also reduced in these cases.

Nevertheless a last device, and not the least, is necessary in the cockpit.

On the model of the pilot, we have not represented the long-term safety loop, in other words the loop related to the objectives of the Phases. The input of this loop is the position of the aircraft in the mission profile and the outputs are the different flight techniques related to each Phase and Sub-Phase. So we have to present to the pilot all the information necessary for navigation and guidance of the aircraft. This information shall be provided on a cathode ray tube, the type of information and their presentation being chosen by the pilot according to each Sub-Phase. For instance, for the approach phase the pilot will call a map display giving the runway, the airways, the different beacons and the future positions of the aircraft. For en route phases, we will present the flight envelope with the different limitations : stalling speed, maximum speed, maximum Mach number an so on, and the optimum flight conditions computed with the true atmosphere. For the cruise phase : navigation parameters, provided by the inertial system, fuel consumption and computed parameters related to fuel consumption, like estimated flight time, available distance and so on, will be displayed in the head down.

And last it will be possible for the pilot to call the informations normally displayed on the head up to have them on the head down in case of failure of the head up.

I think it would be necessary to spend much more time on this problem to fully investigate all its aspects but our intentions on writing this paper was only to give an idea of our philosophy on reducing workload by a precise study of the pilot behaviour.

Nevertheless this philosophy is not a pure abstraction but is based on a great number of simulations and in-flight tests on Mirage III, Etendard IV, Caravelle, Boeing 707 and on two test beds : a variable stability Mirage III and a variable stability N 262.

References

- John W. Carlson and Jean-Claude L. Wanner, "Comparison of French and United States Flying qualities Requirements", AGARD Conference Proceedings N° 106, Flight Mechanic Panel Meeting at Ottawa Oct. 1971, Paper 1.

- R.K. Bernotat and Jean-Claude L. Wanner, "Pilot Workload", AGARD Conference Proceedings N° 106, Flight Mechanic Panel Meeting at Ottawa Oct. 1971, Paper 18.

- Jean-Claude L. Wanner, "Etude de la sécurité des aéronefs en utilisation", Publication Service Technique Aéronautique 1969.

- Jean-Claude L. Wanner, "Technique de Pilotage et Qualité de vol des avions de la nouvelle génération", Conférence L. Blériot, AFITAE and Royal Aeronautical Society, L'Aéronautique et l'Astronautique N° 31, July 1971.

THE INFLUENCE OF COST AND TECHNICAL RISK ON THE DESIGN OF THE AVIONICS SYSTEM FOR THE SPACE SHUTTLE

HOWARD T. WRIGHT
ENGINEERING MANAGER
SPACE SHUTTLE PROGRAM
GRUMMAN AEROSPACE CORPORATION
BETHPAGE, NEW YORK

SUMMARY

The evolution of the Space Shuttle Program, from its inception to the release of the request for proposal in April of 1972, has been influenced primarily by cost considerations. Various configurations were studied, and cost pre-flights were traded against developmental cost. These studies indicated operational costs between 4.5 million and 15.8 million dollars per flight. The baseline configuration was based on the best competition between development and operational cost considerations. The configuration selected by NASA was a small orbiter vehicle with an external Hydrogen and Oxygen tank and two solid rocket engines. This configuration results in operational cost per flight of approximately 11 million dollars.

During the period of basic vehicle configuration evolution, the avionics system also changed from a fully automatic system with data bus operation and large central computers to a more de-centralized or federated system of more conventional design. This paper presents some of the considerations that influenced the design of the avionics system. The primary motivation was cost; however, it became apparent that cost estimates for the system alone could not be the deciding factor. The risk on the cost estimates proved to be very high for the new fully automatic system and therefore a trend away from new equipments to the use of "off-the-shelf" equipments developed.

1. INTRODUCTION

The evolution of the Space Shuttle avionics system is an ideal topic for discussion at a meeting on the subject of "Automation in Manned Aerospace Systems". Early studies provided for a high degree of automation, not only for automatic failure detection and reconfiguration, but also for a completely autonomous onboard checkout and capability for pre-flight checkout as well as on-orbit checkout. The use of large central digital computers controlled virtually all vehicle functions and the cockpit was almost without control switches. The pilot would control various vehicle functions through a data entry keyboard through which he would communicate his desires to the central computer. The central computer would transmit commands and receive subsystem status information via a data bus system. All power switching functions were remote controlled by the computer data bus.

In early 1971 it became evident that the NASA budget for the Space Shuttle Program would be severely limited. The basic vehicle configuration studies were redirected and the two stage fully reusable Space Shuttle system, as shown in Figure 1 was dropped. Cost and performance studies led to a vehicle configuration with a single orbiter vehicle, two solid rocket motors and an external hydrogen/oxygen tank as shown in Figure 2. In October of 1971, NASA called upon the shuttle participating contractors to study the avionics system with the primary purpose of reducing cost and cost risk. This paper describes the management approach, study process, and results of the study conducted by the Grumman Aerospace Corporation.

2. GENERAL INFORMATION

The Management Challenge

Nothing could be more subjective than the estimate of risk in terms of cost and schedule when it comes to the development of a new and complex avionics system. The optimist can readily demonstrate the efficiency of a highly centralized approach with a data bus and many fully automatic functions and the pessimist can visualize the risk involved with such a system. Our management approach was to avoid the risk of the centralized approach by directing the study of a federated system. For the Space Shuttle Program, this approach has several advantages. First of all, the program plan calls for approximately one year of horizontal flight testing prior to the first vertical flight. A system architecture that is completely centralized would require essentially all of the elements for vertical flight one year earlier than otherwise required and the horizontal flight test program would be dependent upon successful integration of a rather complex avionic system. Our federated approach permits us to provide only that part of the avionics system that is essential for the horizontal flight test phase, thus limiting the risk during the first year of flight testing to equipments that are no more complex than that required for the flight test of a conventional aircraft. This phased approach also permits an opportunity to limit peak annual funding requirements. A second management decision aimed at reducing risks and cost was the direction to use existing hardware wherever possible.

2. GENERAL INFORMATION (Continued)

The Technical Approach

In accordance with the above management direction a technical approach was established. Several major avionic system architectural decisions were required. The following brief summary of these decisions established the technical approach.

a) Digital vs Analog Flight Control Computer

A decision was made in favor of a digital flight control computer for the following reasons.

1) Changes are easier to accomplish. As a result of some recent experience at Boeing on the SST program, it was obvious that the analog approach was a cost risk approach. For example, if high gains are required the stability of the power supplies will require very special design techniques. Since we are going to have redundant flight control computers the output voting with digital computers is much more easily accomplished.

2) A digital computer can accommodate a wide range of performance requirements.

3) A simpler interface results when a digital computer is used since most other equipments in the system are digital.

4) Software development can proceed during the development and early flight test program without the need for hardware change.

b) A second major architectural decision was to specify a dedicated aerodynamic flight control computer. This decision was made in order to decouple the risk of the horizontal flight test program from the orbital electronics system. This approach actually forces software modularization, simplifies test programs and permits the use of existing hardware, all of which contribute to a lower cost risk.

c) The third major architectural decision was to specify a fly by wire system. Studies of a mechanical control cable system for a spacecraft indicated a weight penalty of at least 300 lb. for a mechanical control system. At this early stage of development we could not be certain that the inherent vehicle stability would be adequate to fly the vehicle without a stability augmentation system and therefore it was decided to specify fly by wire for both the aerodynamic and the spacecraft control system.

d) A fourth major decision had to do with the controls and displays. The fundamental tradeoff in this area had to do with the use of existing dedicated flight instruments vs flight instrumentation information displayed on a CRT. This decision was perhaps the most difficult to make and in fact, resulted in a compromise. There was no question about the availability of qualified dedicated instruments. There was also no question about the degree of flexibility that a CRT display system would provide to the shuttle cockpit display system. In order to resolve this question full scale mockups of the shuttle cabin were employed. At first it was felt that there would be insufficient cabin panel area to accommodate the dedicated instrument system. The detailed mockup, however, revealed that there was sufficient panel space for dedicated instruments for all flight safety functions as well as three CRT's. Figure 3 shows the general arrangement of the cockpit instruments that resulted from this study.

With these major decisions as guidelines, the remainder of the system was established. The following is a description of the system. A top level view of the system is shown in Figure 4.

The communications and tracking subsystem provides tracking, voice, data and TV links. It delivers navaid data, STDN/SGLS state vector updates, and rendezvous ranging to the GN&C subsystem. The latter acquires additional data from a multimode optical sensor (MMOS), air data sensor (ADS), and body-mounted rate gyro and linear accelerometer assemblies (RGA-LAA). In missions which require it, a sensor deployed from the payload bay delivers rendezvous tracking data.

In all flight regimes, primary guidance and navigation functions are performed by a four-gimballed inertial platform (IMU) operating in conjunction with the guidance, navigation and space flight control computer (GNC). Both are triply redundant.

During ascent, the primary system commands the main engines, (ME), the reaction control system (RCS) and the aerodynamic control surfaces via their respective control electronics. In the event of generic failure in the GNC, the aerodynamic flight control computer (ACC) performs these functions in a backup mode.

In orbit, the primary system commands the OMS and the RCS for rendezvous, orbital operations and docking. In the event of failure of the GNC, prior to entry, the ACC is capable of commanding the OMS and RCS for de-orbit. Re-entry is commanded by the primary system and involves a blend of RCS and aerodynamic attitude control. In the event of failure of the GNC, command is picked up by the ACC.

In the aerodynamic regime, the ACC assumes command of surfaces and of the air-breathing propulsion subsystem (ABPS). A generic failure in the triply redundant ACC results in manual takeover via an independent dual redundant analog backup system.

2. GENERAL INFORMATION (Continued)

The Technical Approach (Continued)

During landing, the primary system performs the guidance computations by Kalman filtering of microwave scanning beam (MSB) data. Guidance commands are then delivered to the ACC which controls the vehicle to touchdown and rollout. If the ACC fails generically, the pilot exercises manual control via the backup system, using a landing display driven by the GNC. On the other hand, a generic failure of the GNC allows autoland to proceed in a degraded mode, without benefit of Kalman filtering. Alternatively, the pilot has the option to fly manually using the landing display, augmenting its guidance with direct visual cues.

Flight-critical displays are driven by the GNC, backed up the ACC, or directly by sensors. Others, such as the multifunctional (CRT) displays, are driven by the systems monitoring computer (SMC). The latter has no role in flight-critical functions because its software will change during the course of the program. If flight critical functions were included in the SMC, high software verification costs would be incurred. All flight-critical command and control functions are hardwired except for the ME interface. All newly-developed hardware will contain built-in test (BIT) and most of the candidate "off-the-shelf" equipments are so equipped.

The payload interfaces with the avionics subsystems are shown in Figure 4. They include S-band communications and tracking, electrical power, dedicated software resident in the system monitoring computer, command and control of payload via a mission specialist station and state vector and attitude initialization data from GN&C. The mission specialist can monitor, record, display or downlink data using his console. Switch interlocks and visual cues between the mission specialist console, the payload handling console, and the cargo bay facilitate a safe deployment or retrieval of payloads.

Primary Flight Station

A LM-type flight director attitude indicator (FDAI) provides three-axis data in space and two-axis (pitch and roll) data during entry and aeroflight. The FDAI was selected in trade studies over the use of an all electronic attitude display because it is available, space qualified, and proven in both Shuttle simulation studies and in LM. The gimbals of the FDAI and the flight director error needles are driven by the GNC in the primary mode and by the ACC in the back-up mode. This mechanization provides the required level of redundancy and flexibility, and eliminates the need for a separate gimbal angle sequence transformation assembly. Compact multitape vertical scale instruments provide air data (angle-of-attack, airspeed, and Mach no.) and altitude/range data (altitude, range, altitude rate, and range rate). The altitude/range instrument is a modified-LM component. Drive sources for the horizontal situation indicator include TACAN, ILS, MSB and the GNC. A dual set of dedicated backup entry instruments provide downrange/crossrange data, commanded drag g error. Simulation studies have shown the entry profile can be flown with just these instruments.

The three multifunction (5 in. x 7 in.) CRT displays provide stroke-written alphanumeric GNC data in an interactive dialogue mode; subsystems' status and reconfiguration data; and supplementary flight graphic displays (e.g., entry corridor, abort data, etc.). The CRT uses a rare-earth, high-brightness, long life P-44 phosphor and a matched filter to enhance contrast. The processor provides symbol generation and 60 Hz display refresh. Interactive GNC/keyboard dialog formats are generated in the GNC since they are flight critical. Supplementary flight graphic displays are formulated in the SMC with data from the GNC. Both CRT display and processor are double gasket-sealed, cold plate-cooled, and have built-in test features.

Hardwire C&W annunciators, an audio/visual alert system, malfunction indicators on individual subsystem panels, dedicated subsystem displays, and the malfunction CRT displays provide sufficient diagnostic capability to display vehicle malfunctions.

Performance Monitoring Station (PMS) (Figure No. 5)

A CRT display and keyboard provides subsystem and payload status monitoring via the SMC. This console is a modular reconfigurable design which initially includes DFI during the horizontal flight test program. When operational, it will be configured to provide on-orbit functions such as: ECS compartment status (airlock, etc.), door and hatch status, remote film camera control, vehicle reconfiguration, etc.

Mission Specialist Station (MSS) (Figure No. 6)

The MSS provides monitoring and control functions associated with payloads. It also incorporates a CRT and keyboard which interface with the SMC or an additional payload-furnished computer. C&W displays provide payload alert status. Modular reconfigurable panel space is provided for dedicated D&C's and GFE equipment.

Payload Handling Station (PHS) (Figure No. 7)

The PHS is equipped with a B&W TV system to enhance payload capture and handling operations. The console is designed for one-man operation of the manipulator arm using a combination of direct and TV viewing. A set of attitude and translation controllers with appropriate switching (Vehicle-Off-Manipulator) provide both vehicle maneuvering and manipulator operation.

3. CONCLUSION

The influence of cost and risk has had a marked effect on the design of the Space Shuttle Program. The risk associated with the avionics system development has been reduced to a minimum by use of a conventional design approach that takes advantage of many existing hardware designs.

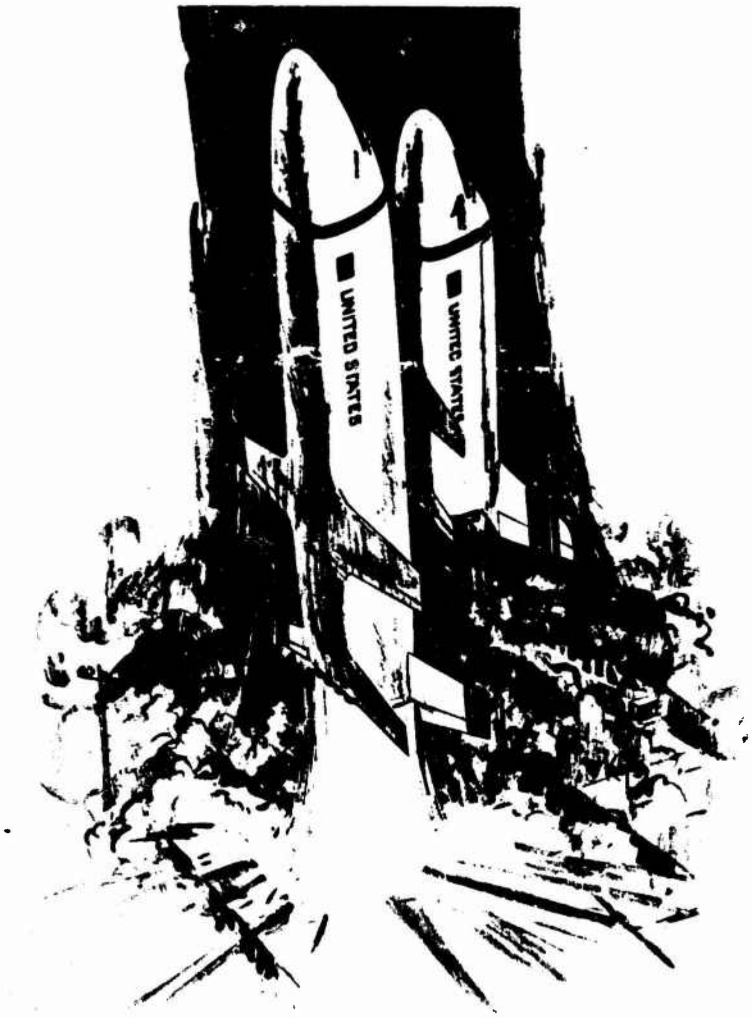


Fig.1 Two stage fully reusable vehicle

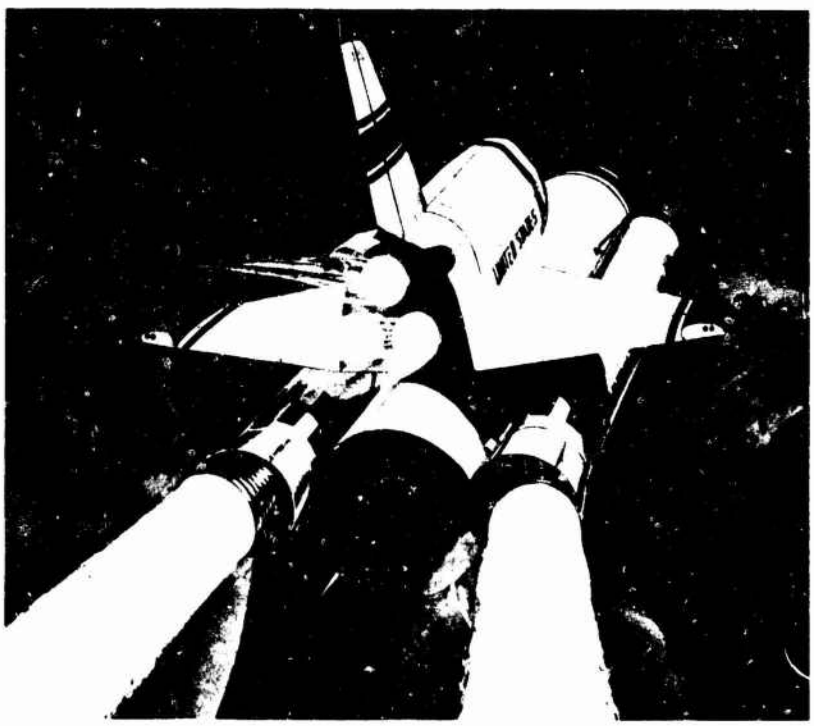
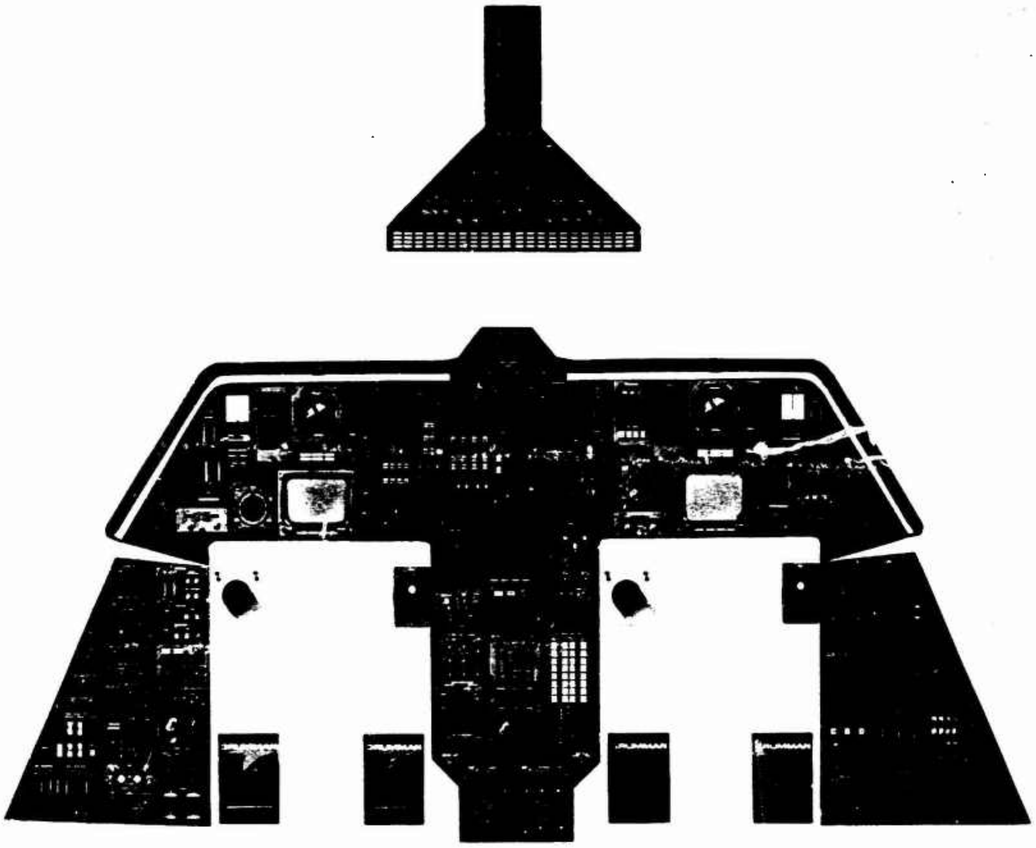


Fig.2 Space shuttle vehicle with external hydrogen oxygen tank and solid rocket motor booster



Displays and controls

FLIGHT CREW DISPLAYS AND CONTROLS

(PROVIDES SINGLE AND DUAL FLIGHT CAPABILITY)

MAJOR D & C SUBSYSTEM FEATURES

- o COMPLETE COMPLIANCE WITH NASA ESTABLISHED REQUIREMENTS
- o A DUAL FLIGHT CAPABILITY FOR EMERGENCY HAND OVER/CREW TASK SHARING
- o A/C & S/C FUNCTIONS INTEGRATED INTO A COMMON FLT. STATION - MINIMUM PNL AREA & WEIGHT
- o ALL CRITICAL D & C'S LOCATED WITHIN REACH OF TWO MAN CREW (INCLUDING CB'S & EC/LS VALUES)
- o LM FLIGHT DIRECTOR ATTITUDE INDICATOR PROVIDING REQUIRED PITCH AXIS RESOLUTION USED FOR BOTH AERODYNAMICS AND SPACE FLIGHT MODES (YAW AXIS LOCKED OUT FOR ENTRY AND AERO MODES)
- o COMPACT MULTI TAPE VERTICAL SCALE INSTRUMENTS USED FOR AIR AND ALTITUDE/RANGE DATA
- o FLIGHT INSTRUMENTS HORIZONTALLY ALIGNED TO PROVIDE A NATURAL CROSS CHECK
- o MULTIMODE HORIZONTAL SITUATION INDICATOR
- o DEDICATED BACKUP ENTRY DISPLAYS PROVIDE DOWN RANGE/CROSS RANGE DATA, COMMANDED DRAG "G" ERROR/FURTHER BACKED UP BY CRT DISPLAYS
- o MULTIPURPOSE CRT DISPLAYS AND KEYBOARDS PROVIDE QUICK ACCESS TO: FLIGHT PROGRAMS, SUB-SYSTEM STATUS, C/O, AND RECONFIGURATION DATA; AND FLIGHT DISPLAY GRAPHICS
  - DISPLAY PROCESSOR GROWTH CAPABILITY FOR ADDITIONAL SYMBOLOGY, LIMITED FORMAT STORAGE, BACKUP ATTITUDE PRESENTATION, INTERFACE WITH TAPE STORAGE UNIT
  - OPTIONAL TV DISPLAY CAPABILITY WITH MINIMAL POWER & WEIGHT PENALTIES PERMITS ALL ELECTRON TUBE DISPLAYS TO BE OF ONE TYPE
  - CRT DISPLAYS ARE PROVIDED WITH A SCRATCH PAD CAPABILITY TO AID DATA ENTRY
  - KEYBOARDS STRATEGICALLY LOCATED
- o SIDE ARM FLY-BY WIRE CONTROLLERS UTILIZED FOR BOTH AERO AND SPACECRAFT PRIMARY AND BACKUP FLIGHT CONTROL FUNCTIONS
  - ROTATIONAL HAND CONTROLLERS (RHC) PROVIDE: PITCH, ROLL AND YAW COMMANDS IN THE SPACE MODE; PITCH AND ROLL CMDS IN THE AERO MODE; ROTATIONAL CMDS ABOUT THE VEHICLE VELOCITY VECTOR DURING ENTRY; BACKUP TVC CMD'S TO THE MAIN AND ORBITAL PROPULSION ENGINE
  - THRUST/TRANSLATION CONTROLLERS PROVIDE: THREE AXIS TRANSLATION CMDS IN THE SPACE MODE: AND ABES THROTTLE CMDS IN THE AERO MODE, CONTROLLERS EQUIPPED WITH AUTO THROTTLE TRACKING CAPABILITY.
  - EACH CONTROLLER IS EQUIPPED WITH A SPEEDBRAKE CONTROL SWITCH
- o RUDDER PEDALS UTILIZED FOR PRIMARY AND BACKUP YAW CONTROL IN AERO FLIGHT FOR NOSE WHEEL STEERING AND FOR VEHICLE BRAKING
- o CENTER CONSOLE THROTTLES USED FOR INDIVIDUAL ENGINE RUN-UP AND THRUST TRIMMING
- o ROLL AND PITCH AERO TRIM CONTROLS LOCATED ON RHC; RUDDER TRIM LOCATED ON CENTER CONSOLE
- o CONVENTIONAL HANDLES USED ON ALL SECONDARY CONTROLS
- o CRITICAL SWITCH POSITIONS MONITORED BY TELEMETRY
- o MODULAR DESIGN PERMITS EASY INSTALLATION OF D & C'S REQUIRED FOR FHF, THE ADDITION OF SPACE HARDWARE FOR FMOF, AND MAINTAINABILITY OF ALL COMPONENTS
- o SAFETY INTERLOCKING IS PROVIDED FOR ALL CRITICAL FUNCTIONS TO PREVENT INADVERTENT OPERATION
- o A MANUAL OVERRIDE CAPABILITY FOR ALL AUTOMATIC FLIGHT CONTROL MODES
- o CREW PROVIDED WITH AC & W ALERT SYSTEM AND A SUFFICIENT DIAGNOSTIC DISPLAY AND COMMAND CAPABILITY TO ISOLATE VEHICLE MALFUNCTIONS AND TO BRING THE VEHICLE TO A SAFE STATE FOLLOWING A FAILURE
- o INCANDESCENT PANEL LIGHTING/RADIOLUMINESCENT DEVICES NOT USED
- o MISSION SPECIALIST STATION DESIGNED FOR QUICK MISSION RECONFIGURATION/WILL ACCOMMODATE 95% OF ALL MISSION REQUIREMENTS ON THE FLT. DECK

Flight crew displays and controls

Figure 3

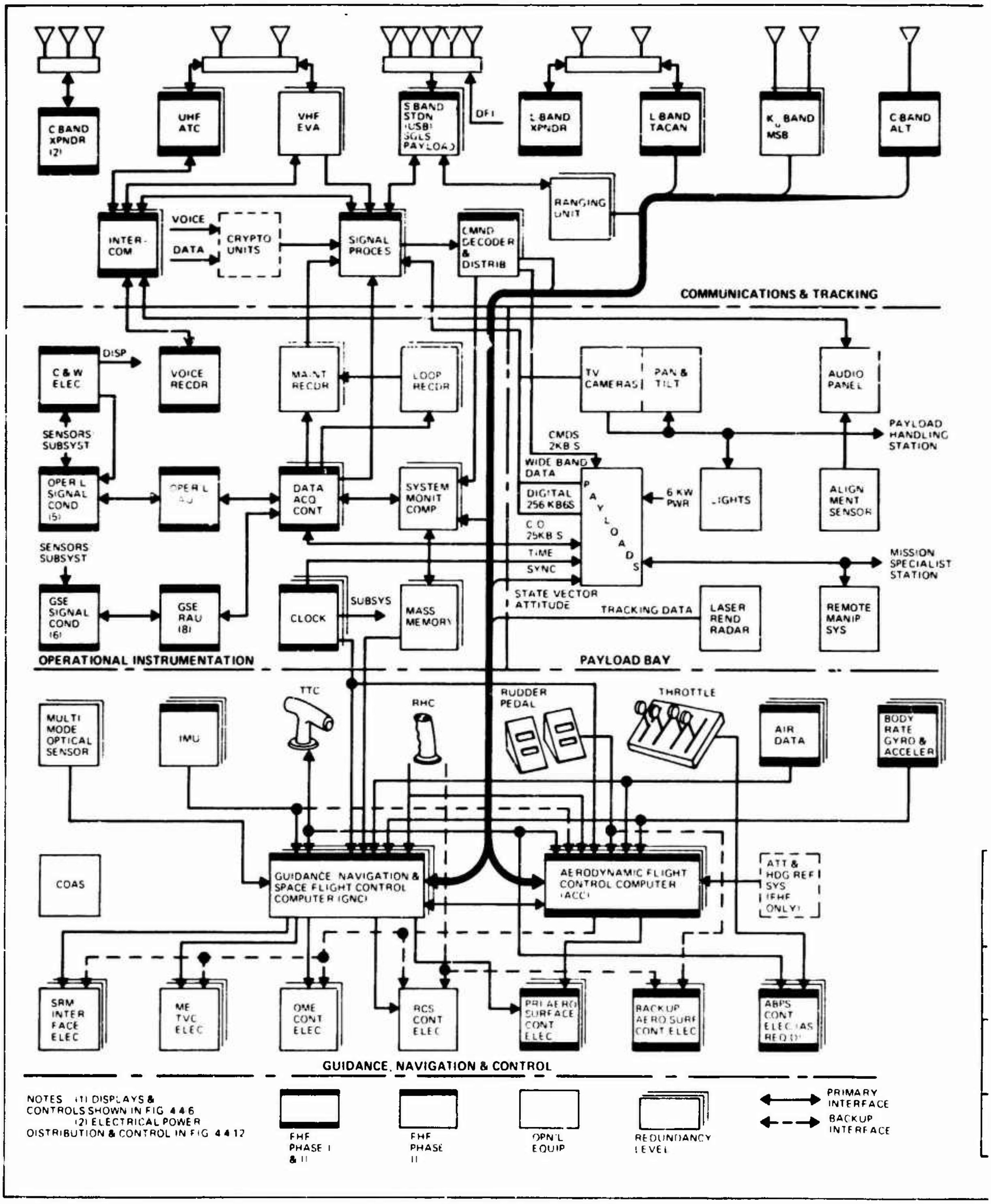


Fig.4 Space shuttle avionics subsystems block diagram

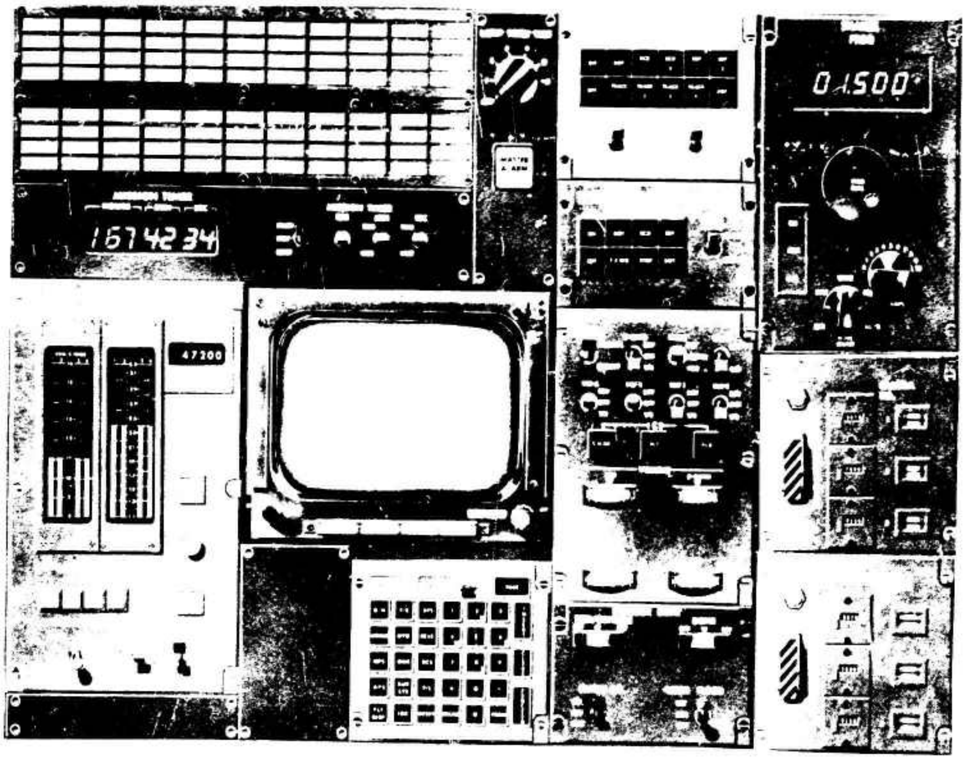


Fig.5 Performance monitoring console

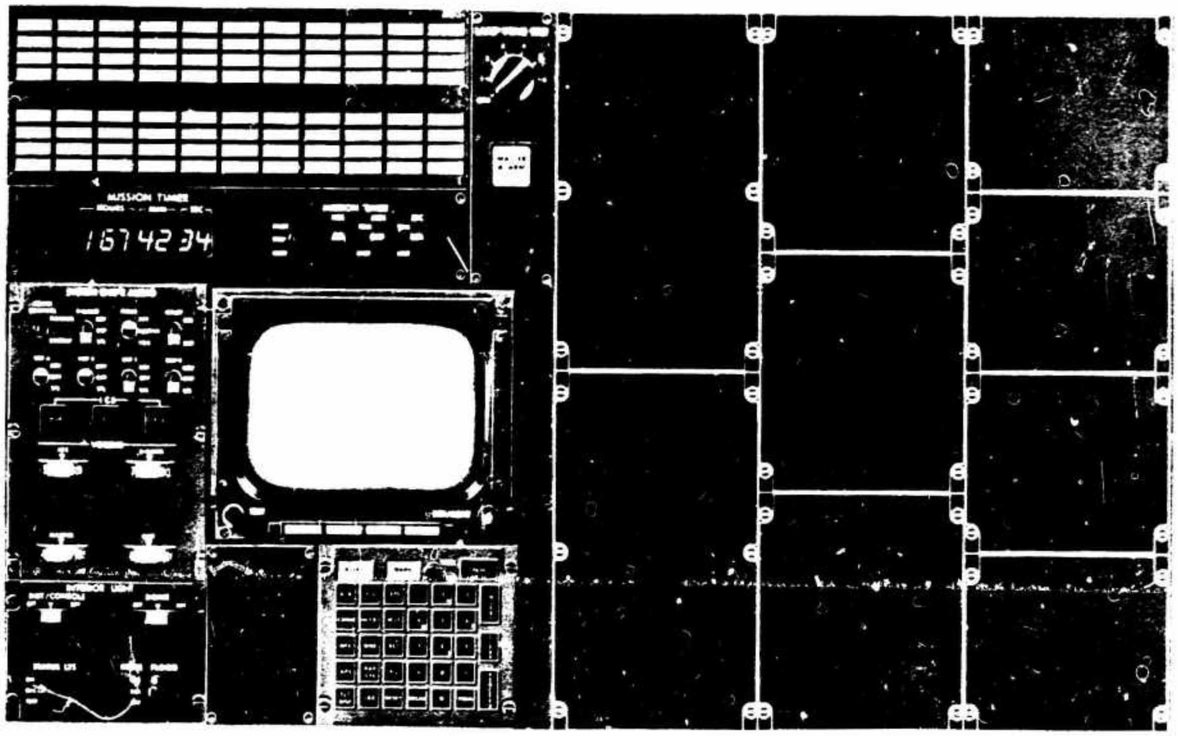


Fig.6 Mission specialist station

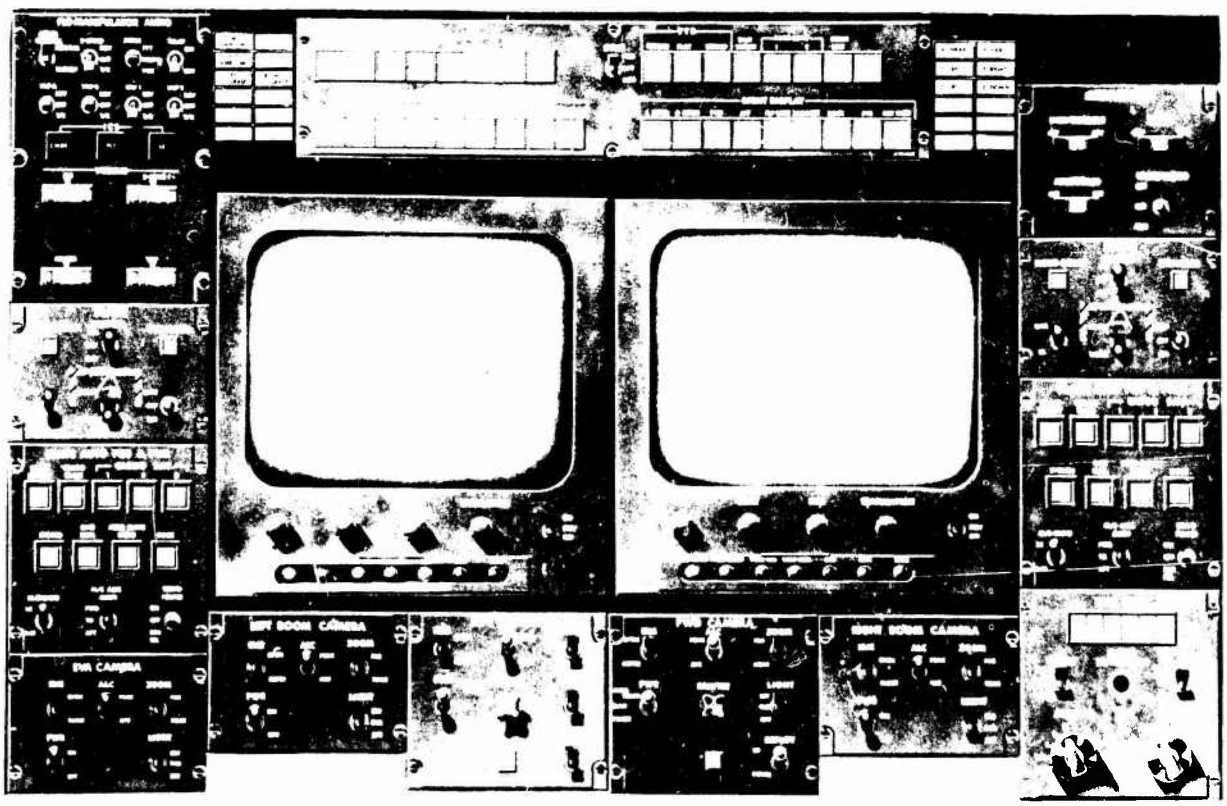


Fig.7 Payload handling station

AUTOMATIC ACQUISITION AND TRACKING METHODS EMPLOYED IN THE JOINT SERVICES IN-FLIGHT DATA TRANSMISSION SYSTEM (JIFDATS)

T. N. Leiboff
Northrop Electronics Division
Palos Verdes Peninsula, Calif.

SUMMARY

In designing the microwave link for the Joint Services In-Flight Data Transmission System, an interesting problem presented itself. How do you acquire and track a "Mach-2" aircraft from a second "Mach-2" aircraft and simultaneously acquire and track a surface terminal with no a-priori knowledge of the aircraft location or altitude? This was part of the system requirements: total system lock-up (ground to relay to sensor aircraft) in less than 90 seconds.

JIFDATS is an all-weather, day-night, multi-sensor, in-flight data transmission system designed for use by all the military services. The initial tests were performed in an RF-4C aircraft, transmission range was extended by use of a second RF-4C as a radio relay. The normal operating mode for JIFDATS is automatic. Except for the usual checkout, servicing and maintenance activities in which personnel take a large part, the only need for personnel functions is to establish the proper conditions for system operation, turn on the system, and monitor the operation to assure continuity of data transmission. In each case though, there is a manual back-up mode for bypassing the automatic features of acquisition and tracking.

A scenario of a "typical" tactical reconnaissance mission is presented showing the various steps taken by the operator in the sensor aircraft, the operator in the relay aircraft, and personnel at the surface terminal during each phase of the mission. It is shown how the relay aircraft automatically acquires the sensor aircraft which is transmitting a low bandwidth signal on an omni-directional antenna, while it rotates its high-gain narrow beam directional antenna. Then how the sensor aircraft locks on to the relay while the relay and ground terminals acquire and track.

1. INTRODUCTION

The Joint Services In-Flight Data Transmission System (JIFDATS) has been developed and flight-demonstrated by Northrop Corporation, Electronics Division in Palos Verdes Peninsula, California under the direction of the Naval Air Systems Command. This program was contracted in October 1969 to meet projected operational requirements through 1980.

1.1 Concept

It is the purpose of JIFDATS to provide the joint military services with day and night all-weather capability for transmission of ground surveillance imagery immediately after acquisition by airborne reconnaissance sensors. This is achieved by high-speed processing of the sensor data, both in the air and on the ground, and an air-to-ground digital microwave link. Figure 60-1 illustrates the JIFDATS concept.

The JIFDATS system whose military designation is Data Transmission System, AN/USQ-49, is composed of three major subsystems: a sensor-equipped aircraft, a mobile surface terminal, and a second aircraft serving as a microwave relay complement. JIFDATS is designed for compatible installation and operation in specified military aircraft and surface terminals appropriate to each service.

The two airborne subsystems have been designed to accommodate a variety of installations compatible with any of the candidate aircraft. For the scheduled flight test of JIFDATS in the Air Force RF-4C aircraft, most of the Sensor and Relay components are mounted in a specially converted fuel pod that is readily detachable from the airplanes. Production designs accommodate internal installation with the aircraft. Figure 60-2 shows the sensor aircraft taking off from Edwards Air Force Base during flight test.

The Surface Terminal Subsystem is similarly amenable to variations in installation and deployment. The mobile land-based units are designed to be airlifted individually by helicopter or collectively in transport aircraft. The surface recording terminal is designed for installation in either trailer-type or truck-mounted shelters, or in shipboard compartments. The antenna terminal employs a self-leveling wheeled platform for land operations and will incorporate ship-motion stabilization for shipboard use.

1.2 Performance

Transmission ranges of up to 270 nautical miles can be achieved directly from the sensor-equipped aircraft, and nearly 500 nautical miles when the relay aircraft is utilized. Moreover, the use of the relay aircraft increases the effective operating range when the sensor aircraft is beyond radio line-of-sight of the ground terminal or flying low-altitude reconnaissance missions.

Of course, the operating range for JIFDATS is constrained by natural phenomena such as radio line-of-sight, reflection from the earth's surface, low grazing angles, and occlusion by intervening terrain. Other factors upon which the range is contingent are the operating frequencies, transmitter power, receiver sensitivity, and antenna beamwidth.

The selectable microwave channels permit up to five separate and independent Sensor-Relay-Surface combinations for transmission on a non-interference basis. The abbreviated 'access time' (elapsed time between sensor output and image availability at the terminal) offers a significant improvement in real-time surveillance missions by reducing the strike reaction time. The access time for high-resolution sensor imagery is approximately 60 seconds for infrared detectors, 90 seconds for side-looking radars, and 180 seconds for photographic cameras. For lower resolution imagery the access time is correspondingly shorter.

The imagery is converted by the Sensor Aircraft for transmission through a digital radio link at microwave frequencies to a Surface Terminal which reconstitutes the imagery on film for immediate viewing and interpretation. For extended-range missions or when radio line-of-sight is occluded by terrain, another JIFDATS subsystem in the Relay Aircraft receives and amplifies the digital radio link outputs from the Sensor Aircraft and re-transmits, at higher frequencies, to the Surface Terminal. The Surface Terminal has dual-band capability to accommodate transmissions from either the Sensor or Relay aircrafts.

The search, acquisition, identification, and lockup between Sensor and Surface or between Relay and Surface is achieved in 90 seconds or less by means of a two-way beacon technique. A high-speed reacquisition capability is included to adjust for any interruption in the link. All acquisition and lockup operations are automatic, with provision for manual override.

The following table lists some of the JIFDATS system performance parameters.

JIFDATS PERFORMANCE

Mission Radius	Up to 500 NM With Relay
Altitudes	Up to 55,000 Ft
Velocities	Up to 1,200 KTS
Aircraft Sensors	IR, SLAR, Photo, Aux. Data
Modulation (Data)	Digital (QPSK)
(Beacon)	Digital (FSK)
RF Frequencies	C-Band and K-Band
Number of RF Channels	5 Each Band
RF Bandwidth	Up to 110 MHz
Video Bandwidth	Up to 50 MHz
Signal-to-Noise Ratio	35 db at Recorder
Photo Quality	Up to 60 lp/mm, 13 Shades
Access Time	1 to 3 Minutes

Figure 60-3 presents a simplified block schematic of the JIFDATS system.

2. DATA TRANSMISSION

Basic to the operation of JIFDATS is the transmission of data from a Sensor Aircraft to a Surface Terminal, either by a direct link or through a Relay Aircraft. The need for the Relay Aircraft arises when the line-of-sight between the Sensor Aircraft and the Surface Terminal is obstructed.

2.1 General Description

In order to ensure uninterrupted flow of reconnaissance data to the Surface Terminal for reconstitution into film imagery, a beacon-tracking link is first established between the JIFDATS subsystems and then maintained throughout the transfer of sensor data down the link. The order in which the Sensor-Relay link and the Relay-Terminal link are established is optional. The probable practice will be for the Relay and the Terminal to link up and track each other while awaiting the Sensor to arrive on station. Actually, a single Relay-Terminal team could serve several sensors in a time-shared scheme.

Transmissions in the direction of the Surface Terminal are defined as 'down link', and 'up link' refers to transmission in the direction of Sensor aircraft with or without the Relay aircraft in the loop.

Following initial acquisition, either air-to-air or air-to-ground, a continuous 'up link' beacon transmission is maintained until the mission is completed. 'Down link' beacon transmissions are employed during all intervals preceding and between transfers of sensor-acquired data to the Terminal. The data transmissions are used in lieu of down-link beacon signals by the receiving station (Relay or Terminal) to track the transmitting station (Sensor or Relay). Beacon signals in both directions are encoded to include station identification, acquisition and lockon information, and reciprocal bearing for the Sensor to expedite mutual lockup with the Relay or the Terminal.

Interference-free operation of the JIFDATS systems in any given area is provided by a unique lockon technique, based on a coded tracking beacon arrangement between the Sensor and Relay aircraft, and between either aircraft and the Surface Terminal. The sensor aircraft surface terminal control panels indicate when complete link lockup has been achieved. JIFDATS provides for complete system lockup without prior knowledge of position data of either aircraft or the surface terminal.

Five data channels with corresponding independent beacon channels permit the operation of a maximum of five JIFDATS systems in one area. The beacon channels, spaced at 10 MHz intervals, are located at the high end of the JIFDATS C- and K-bands.

In order to obtain adequate data quality at the required separation distances between terminals, it is necessary that both airborne and surface terminals make use of high gain directional antennas. Such antennas are characterized by relatively narrow beamwidths. Consequently, they must be oriented to the proper direction (acquisition), and they must continue to point in the proper direction throughout relative geometry changes between the links (tracking).

Because a time limit has been allocated to each link for acquisition, it is necessary to make use of omnidirectional antennas in the aircraft and broad beam antennas on the surface. Each terminal must be capable of acquiring without knowing the location of the other participating terminals. In addition, the system must be capable of rejecting any signal at the working frequencies which are not "authentic" (accompanied by an I.D. code). Finally, the Relay-Sensor link must be capable of acquiring in a hostile environment.

2.2 Surface Acquisition of the Sensor/Relay Aircraft

The Surface Antenna Terminal utilizes a high gain parabolic monopulse tracking antenna together with four pyramidal horn antennas (two for K-Band and two for C-Band) in a sequential lobing tracking configuration (Figure 60-4). The antennas are all mounted on a single pedestal.

To initiate automatic acquisition of the target, the operator energizes the AUTO SEARCH Control Switch. (See Figure 60-5 for operator control panel photograph.) The indicator will illuminate and azimuth search commences at a scan rate of 33 deg/sec. A threshold will be initially established at a level 20 db below the minimum range signal level. The acquisition horn squinted in the direction in which the antenna is moving will be selected. When the AGC signal, which is received from the Down-Converter, exceeds the pre-set threshold, sequential lobing between the two acquisition horns will commence and automatic acquisition in azimuth will be effected. If a target is not encountered during the first sweep, the threshold will be reduced by 20 db for a second sweep. If a target is still not encountered, the threshold will be reduced an additional 20 db, and a third sweep will be made. The three sweeps will be made with the elevation angle set to zero degrees nominally, or to an angle selected by the operator with the elevation Manual Position Control. After azimuth acquisition, the tracking antenna will sweep (at a scan rate of 20 deg/sec) through a preselected elevation angle until elevation acquisition of the target and AUTO TRACK occurs.

If a valid target is not encountered during the three azimuth scans at the first elevation sector, the antenna will be directed up to the second elevation sector. The threshold will be set to a value 40 db below the minimum range signal level and a single sweep in azimuth will be made. If a target is encountered, the antenna will move downward seven degrees (the 3-db point) and then drive upward until the target is encountered and AUTO TRACK occurs.

Should no valid target be encountered during the second elevation sector, the threshold will be set to a value 60 db below the minimum range signal level, the elevation axis will drive up to a third level and the above azimuth search procedure will be repeated.

Three elevation sectors will suffice to cover acquisition over a total elevation sector of at least 35 degrees (and an azimuth sector of 360 degrees). If target acquisition has not occurred, and the switch on the Antenna and Receiver Control Panel is set to 35 degrees, the antenna will drive back to zero degrees (or a manually selected elevation angle other than zero), and the procedure will be repeated until acquisition occurs. If the switch is set to 60 degrees, two additional elevation sectors will be searched prior to repetition of the search procedure.

When azimuth acquisition is effected and the antenna begins to sweep upward, the Monoscan Converter will be activated and the elevation portion of the tracking antenna signal will be switched into the Down Converter and time-shared with the azimuth tracking signal. The resulting "video" signal will be amplitude-modulated by the elevation pointing error. As the target enters the tracking antenna's acquisition cone, the elevation portion of the "video" signal will become larger in amplitude than the azimuth signal; that is, the amplitude of the tracking antenna signal exceeds the amplitude of the acquisition antenna signal. Elevation acquisition will occur and automatic tracking in both axes (AUTO TRACK) will begin. These procedures are diagrammed in Figure 60-6.

There are two alternate methods of search if the aircraft position is roughly known. The first, Sector Search, may be used if the antenna pointing angle is known to be within the pattern of the acquisition horns. The antenna is positioned by the operator to the azimuth and elevation pointing angle. The acquisition horns are sequentially sampled until the aircraft is acquired, then tracking is consummated by the pencil beam antenna.

The second, Spiral Search, is used if the antenna pointing angle is known within ±10 degrees in both azimuth and elevation or it can be selected for automatic reacquisition. The high gain pencil beam is used to acquire the aircraft, beginning from the command pointing angles and searching in an ever increasing spiral pattern. The operator can adjust the limiting size of the spiral cone as well as the amount of over-lap between successive beam sweeps around the spiral pattern. When acquisition is achieved, the surface antenna continues pencil-beam tracking.

Figure 60-7 is a photograph of the Surface Antenna Terminal at the operational test site. The pedestal contains the servo drive equipment and the beacon transmitters and receiver frequency converters are mounted on the counter weight arm.

The antenna terminal can be located as much as 500 feet from the Surface Recording Terminal (SRT) and power generator. The operator's controls shown are duplicated within the SRT where the equipment operators are normally located.

2.3 Relay Aircraft Acquisition of the Sensor Aircraft

The Relay aircraft utilizes directional monopulse antennas (on the top and bottom of the aircraft for each of the two RF bands) for both the acquisition and tracking functions. The upper and lower directional antennas revolve, servo-slaved to each other, at 72 deg/sec with the receiver switching at 50 Hz between antennas. (See Figure 60-8 for a photograph of the Relay Operator's Control Panel.) The Relay's beacon transmitter is turned on and connected to the lower antenna.

Reception of the proper I.D. code enables AGC detection. When the antennas next cross one of the four airframe reference markers, a peak-detecting signal store is enabled. The antennas continue in their first 360-degree revolution with the receiver switching between the upper and lower antenna. At the end of the first revolution, the upper and lower antennas are compared and the transmitter and receiver are switched to the stronger antenna. The maximum AGC signal is stored and the antennas/receiver continue into their second revolution "seeking" an AGC signal within an established tolerance of the previously stored peak signal. When the signal is encountered, a brake is applied to the synchro (which has been following the antenna) and thus the azimuth of the peak signal is marked. The antenna servo position loop is closed and the antenna drives back to the corresponding azimuth at which time a servo position null will occur. A last-moment sampling of the upper and lower antennas is made and the transmitter and receiver are connected to the one demonstrating the greater signal. AUTO TRACK is thereby achieved while using the optimum antenna. Once AUTO TRACK is achieved, the antenna true bearing unit is enabled and the information is transmitted to the Sensor aircraft's omni antenna/receiver. The Sensor aircraft slews its directional antenna to the reciprocal antenna bearing, acquires the Relay aircraft and a LINK LOCK indication is generated in both aircraft.

A similar procedure is employed for the acquisition of the Surface Terminal by the Relay aircraft.

2.4 Sensor Aircraft Acquisition of the Relay Aircraft

The Sensor aircraft utilizes omnidirectional antennas for both transmitting a beacon signal and for receiving antenna true bearing instructions from the other participant in its link. One of these antennas is located on the top of the aircraft and the other is located on the bottom. Associated with each of the omni antennas is a directional monopulse tracking antenna. The described antenna complement is shown in Figure 60-9.

The transmitter and receiver are first connected to the upper omni for a period of 18 seconds. If during that time the proper ID code is received, the antenna true bearing information is decoded from the beacon signal and the upper directional antenna is slewed to the commanded reciprocal angle at which time a servo position null will occur. A last-moment signal strength comparison is made between upper and lower omni antennas after which the transmitter and receiver are connected to the directional antenna associated with the "stronger" omni antenna. If, at this time, the AGC signal is present together with the proper ID code, the system disables the reciprocal bearing loop and closes the receiver loop (AUTO TRACK) and LINK LOCK is indicated. When SYSTEM LOCK indication is received, the transmitter is switched to Data Mode operation whenever the Operate Mode has been selected. See Figure 60-10 for a photograph of the Sensor Operator's Control Panel.

If inadequate signal reception occurs during the upper omni antenna dwell period, the transmitter and receiver are switched to the lower omni antenna and the above procedure is repeated, substituting lower antenna functions for previously described upper antenna functions.

If acquisition does not occur during the lower omni antenna dwell period of 18 seconds, the transmitter and receiver are switched back to the upper omni antenna and the initial procedure is repeated. A similar procedure is employed for the acquisition of the Surface Terminal by Sensor aircraft.

2.5 Airborne Directional Antenna

The Directional Antenna operates as a spaced, duplexed transmit/receive, azimuth tracking antenna. The antenna employs monopulse techniques to track a coded signal while simultaneously receiving a beacon or command signal and transmitting wideband data signals. Figure 60-11 is a photograph of a C-Band and K-Band antenna used in JIFDATS.

The directional transmitting and receiving arrays consist of a broadside of 8 rows (C-Band), 12 rows (K-Band) of short monopoles mounted on a ground plane and spaced approximately one-fourth of a wavelength apart. Each row of elements consists of a reflector, a feed, and several director elements forming a Yagi configuration. The elevation beamwidth of the array is determined by the number and longitude spacing of the elements. The azimuth beamwidth is determined by the number of rows of elements and spacing of the elements. The gain or directivity of the antenna is a function of the horizontal radiating area measured in square wavelengths.

The angle of maximum radiation above the horizontal is a function of the phase velocity in the end-fire radiating elements, which, in turn, are functions of the length, diameter and spacing of the monopoles. The taper on the length and the spacing of the monopoles is adjusted to provide optimum azimuth sidelobe levels. The antenna design permits the use of a single-channel output instead of the normal two-channel outputs required in monopulse tracking systems. The array uses stripline-combining techniques. A spoiler is mechanically switched in to change the elevation coverage from 15 to 30 degrees.

2.6 LOCKUP

Upon completion of the Sensor-Relay or Relay-Surface-Terminal linkup, a 'link lock' confirmation is exchanged and displayed on the operator panels. A 'system lock' confirmation is exchanged and displayed when the Sensor-Relay-Terminal or Sensor-Terminal linkup has been satisfactorily completed. The linking up of any pair of JIFDATS subsystems can be achieved automatically within 90 seconds with no prior knowledge of each other's relative or geographic location. If any part of the link is interrupted, a 'link loss' indication is displayed on all operator panels and a high-speed reacquisition routine is executed to re-establish the affected link and reconfirm 'system lock'.

Tracking is continued by a rate memory servo for 10 seconds after link loss prior to initiating the reacquisition cycle. This shortens the communication down time during temporary fades, interference or extreme aircraft maneuvers.

The airborne antenna coverage is adequate to permit data link performance in a wide range of attitude variations of the sensor or relay aircraft. Coverage is achieved by use of upper and lower antennas, and variable beamwidth. In cases of severe sensor aircraft maneuvers, data transfer may be degraded, but tracking can be maintained by use of the sensor's beacon-only mode.

2.7 Manual Acquisition Mode

The JIFDATS acquisition and tracking systems provide a manual override capability for each airborne terminal and the surface terminal. In the airborne systems, the operator selects manual operation by pulling out the Antenna Bearing Controller knob. This action switches the appropriate receiver and transmitter to the pre-selected Directional Antennas. At the same time, the LINK LOCK indicator becomes an ID code indicator. Rotation of the Antenna Bearing Controller provides the operator with a variable slew rate capability in order to orient the antenna to the proper bearing (in the manual mode, LINK LOCK can be assured only when the general location of the signal source is known). In the Relay aircraft, signal strength meters are provided; consequently, reception of an RF signal and its relative strength is indicated. No signal strength meter is provided in the Sensor aircraft due to a lack of panel space; however, if its antenna receives an authentic signal, the ID code indicator lamp will illuminate (manual mode only), thereby providing a basic source of feedback to the operator.

Subsequent returning (pushing-in) of the Antenna Bearing Controller to the automatic mode position will result in:

(a) Immediate automatic tracking if an authentic signal of adequate magnitude is present.

(b) The automatic acquisition sequence if either of the above conditions are not present.

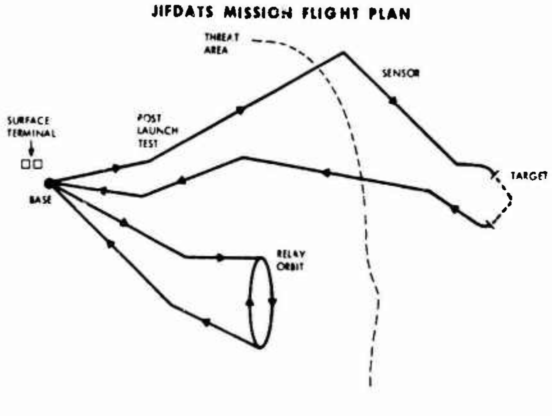
At the Surface Terminal, the Manual position mode of operation may be selected as a secondary operating mode. The antenna's position may be controlled by the Azimuth and Elevation Manual position controls.

The Manual position controls are connected to the 1:1 synchro torque receiver (TR) rotor through an electro-mechanical clutch and a gear train. The gear ratio is 15:1 producing one revolution of the TR rotor for 15 turns of the control. The clutch is energized by 28 volt current supplied by logic whenever the Manual mode is activated.

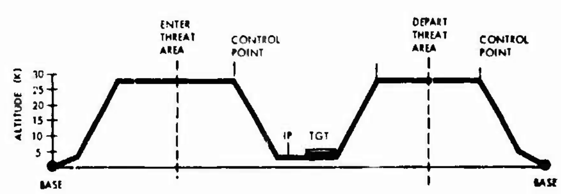
3. TYPICAL SCENARIO

The following is a scenario designed to illustrate the operation of the JIFDATS system. The illustrations below depict a typical JIFDATS mission plan and a nominal flight profile for the Sensor and Relay aircraft. The aircraft launches from the base after preflighting the equipment, climbs out under departure control and conducts a post-launch equipment test. The Relay aircraft acquires the Surface Terminal after entering assigned orbit. The Sensor aircraft then initiates acquisition of the Relay aircraft just prior to entering the threat area. After system link-lock is established, the Sensor aircraft turns toward the target area and descends for the run, leveling out at the Initial Point (IP); makes the run using the KS-87 camera as the sensor; climbs out after the run and returns to base. Tactically, the aircraft could make additional runs on other targets to the limit of fuel or film.

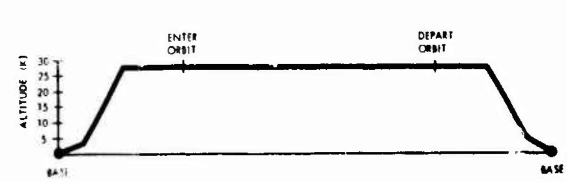
JIFDATS MISSION FLIGHT PLAN



SENSOR AIRCRAFT NOMINAL FLIGHT PROFILE



RELAY AIRCRAFT NOMINAL FLIGHT PROFILE



In preparation for the mission, the sensor aircraft is serviced and equipment preflighted. These activities are conducted by the flight line crew and may be additionally checked by the aircrew just prior to or during the aircraft preflight inspection and checkout.

Since a photo mission is scheduled, the ground servicing crew performs the following operations:

- Inspects and services the In-flight Photo Processor Scanner (IPPS).
- Loads film.
- Inserts the mission ID code in the status controller.
- Performs an external inspection of the subsystem equipments.

Prior to applying ground power to the a craft and equipment, the operator's control panel and circuit breakers are checked for proper positi ing (protective mode). After power is applied, the system is warmed up by placing the operator's Mode Select Switch to STANDBY. When system READY is indicated, the system is placed through a preflight test cycle by selecting TEST on the Mode Select Switch. When the TEST IN PROGRESS indication extinguishes, the system is GO and the preflight is complete. At that time, the system is placed to STANDBY. In the event the system should malfunction, a FAIL indication would be presented on the control panel and examination of the BITE panel in the pod will provide information as to the location of the fault. Corrective action may then be initiated and a new test cycle conducted after repair or replacement.

The Relay aircraft preflight activities are similar to, but not as extensive as, the Sensor aircraft preflight functions. In this case, ground servicing prior to conducting system check-out consists of inserting the mission ID code in the Status Controller and conducting a visual inspection.

The preflight check-out is started by verifying that the system is in a protective mode, by inspecting the operator control panel and circuit breaker panel. After ground power is established, the system is warmed up by selecting the STANDBY mode on the control panel. When the READY light illuminates, the system is returned and placed through a preflight test cycle by selecting the TEST mode. Upon completion of the test, the system is returned to STANDBY in preparation for launch. If a malfunction is indicated, the BITE panel will provide information as to its location so that immediate corrective action can be initiated.

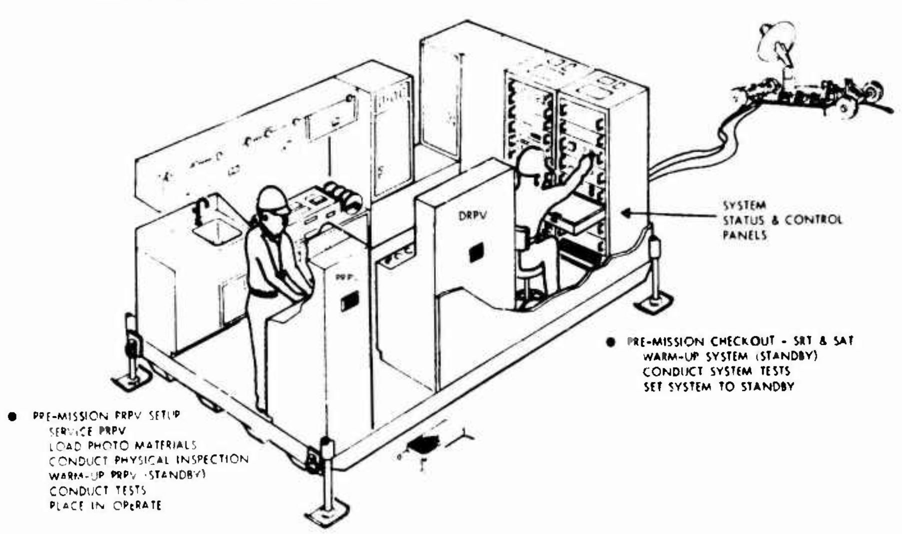
The Surface Terminal pre-mission preparations include those activities necessary to initially prepare the Antenna Terminal and Recording Terminal for receipt and recording of transmitted reconnaissance data resulting from a JIFDATS tactical mission.

The Surface Terminal operation crew first conducts a mechanical equipment checkout prior to applying power to the system. This check consists of the following:

- Verify that equipment and circuit breakers in Surface Recording Terminals (SRT) are in protective mode.
- Enable power supply.
- Inspect power, coaxial and communications cables.
- Conduct visual inspection at Surface Antenna Terminal (SAT).

After the mechanical inspection has been completed, power is applied to the SAT and SRT and the SAT is placed in STANDBY mode. After warmup is completed, the Tracking Antenna and Data Link equipment is placed through the test cycle. Coincident with these activities one of the operators initiates pre-mission servicing of the Photo Recorder Processor Viewer (PRPV). After verifying that servicing is complete, which includes film and expendable supply loading, power is applied to the recorder and GO status is verified. With the SAT in STANDBY mode and the recorders in OPERATE mode, the remaining pre-mission activities are completed, which includes checking mission schedules, supplies and routine housekeeping. At the completion of these activities, one of the operators notifies the appropriate operations center that the terminal is ready for tactical mission transmission as per schedule.

SURFACE TERMINAL PRE-MISSION PREPARATION

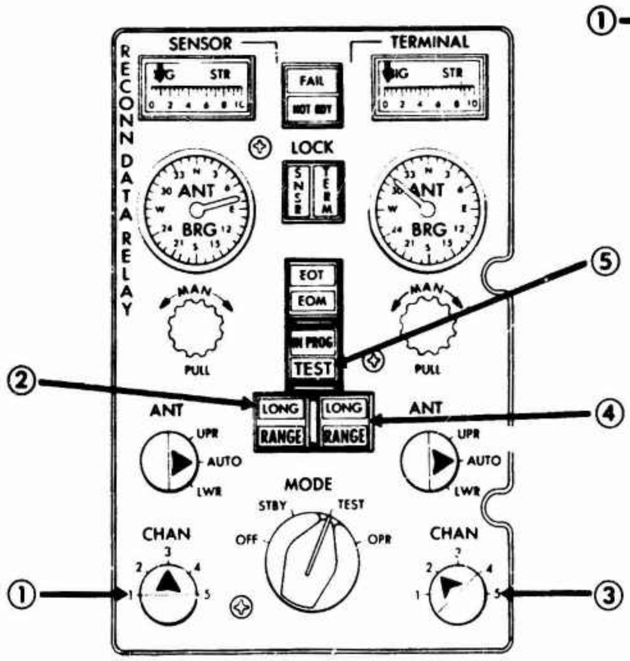


Tactical mission operations start with the Sensor and Relay aircraft takeoff, departure, climb and initial cruise. All preflight and pre-mission operations have been completed and all systems GO. The Sensor aircraft has leveled off after climb to cruise altitude and is ready to initiate a post-launch test of equipment.

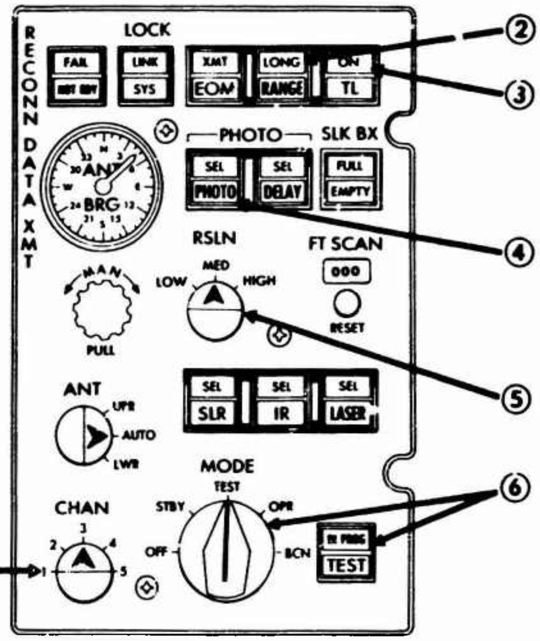
To test and verify the system status, the operator sets up the control panel as illustrated:

1. Selects Channel 3 for transmission.
2. Selects RANGE mode - LONG.
3. Enables Tracking Link.
4. Selects PHOTO mode.
5. Selects MEDIUM Resolution Mode for data transmission.
6. Selects TEST mode; initiates the test by depressing the TEST switch; and monitors TEST IN PROGRESS indicator and FAIL indicator. When IN PROGRESS indicator extinguishes, the test cycle has been completed and the system is GO.

RELAY OPERATOR'S CONTROL PANEL SETUP



SENSOR CONTROL PANEL SETUP



After the Sensor aircraft is launched, the Relay aircraft takes off and initiates a climb to cruise altitude. After reaching cruise altitude, post launch activities are initiated.

At this time, the Relay aircraft operator sets up the control panel in preparation for conducting the first mission and in-flight system test. As illustrated on the control panel, the following selections are made:

1. Selects Channel 3 for reception of sensor aircraft transmissions.
2. Selects sensor range mode - LONG.
3. Selects Channel 2 for transmission of data to the Surface Terminal.
4. Selects Surface Terminal Range Mode - LONG.
5. Selects TEST Mode; initiates the test by depressing the TEST switch, and monitors TEST IN PROGRESS Indicator and FAIL Indicator. When the test has been completed, the TEST IN PROGRESS indicator extinguishes.

Subsequent to the aircraft departures and in accordance with the mission schedules provided, the operators proceed to set up the equipment for acquiring the Relay aircraft and subsequently recording the reconnaissance data.

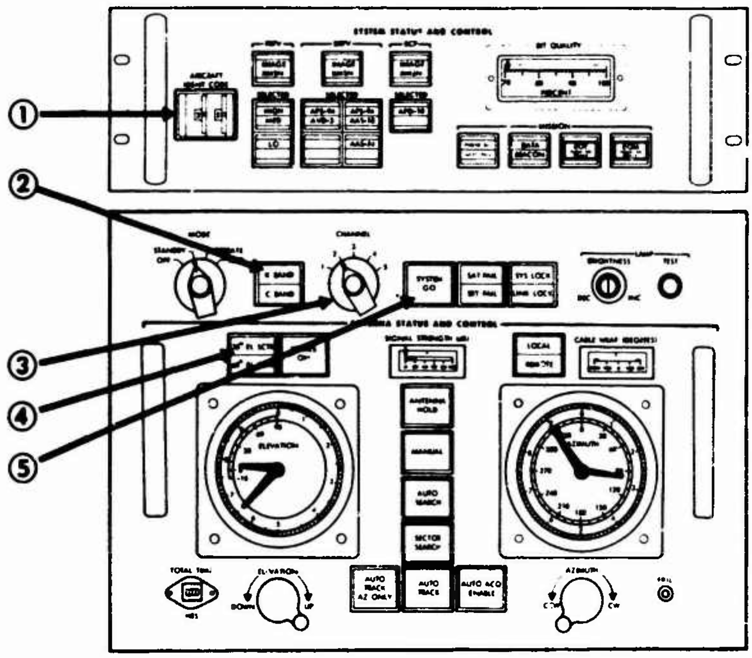
To set up the Surface Terminal equipment for acquisition of the Relay aircraft and receive data, the following selections are made on the control panels located in the SRT.

1. The first mission ID code of 35 is set on the thumbwheels.
2. The K-Band select switch is depressed.
3. Selects Channel 2 for reception of data from the Relay aircraft.
4. Selects 35° EL SCTR for the antenna.
5. Verifies that system is GO.

At the completion of these activities, the Terminal is now ready to be placed in operational status to acquire the Relay aircraft at the scheduled time.

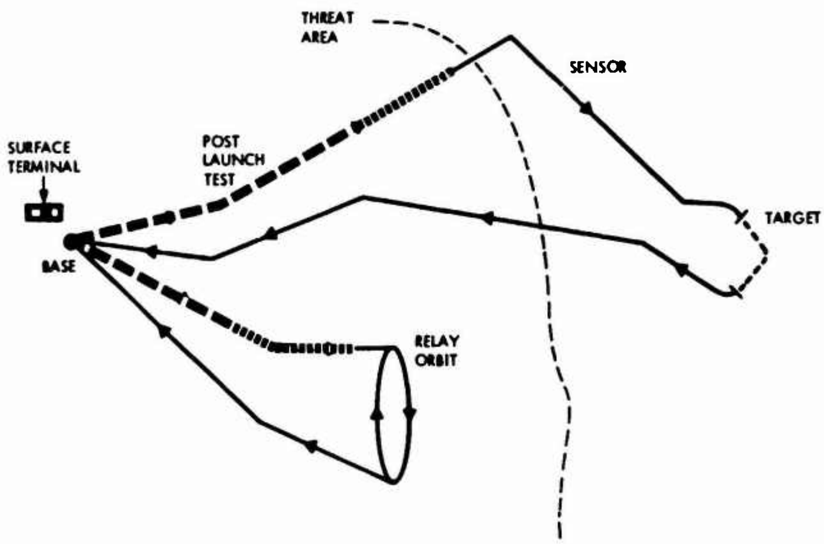
A drawing of the Terminal Control Panel is shown on the following page.

TERMINAL CONTROL PANEL SETUP



As the Sensor aircraft approaches the threat area and the Relay aircraft approaches the orbit area, the Surface Terminal initiates acquisition operations. The results of this activity will establish the Relay/Terminal Link-Lock.

JIFDATS MISSION FLIGHT PLAN



The Surface Terminal operator, in accordance with the mission schedule time, initiates link-acquisition with the Relay aircraft as follows:

1. Sets Mode Switch to "OPERATE".
2. Depresses "AUTO SEARCH" switch.

After enabling the system, the operator verifies that the following indicators are illuminated:

- K-Band
- SYSTEM GO
- 35° EL SCTR
- LOCAL

Antenna begins to search through 360° and periodically the elevation angle changes.

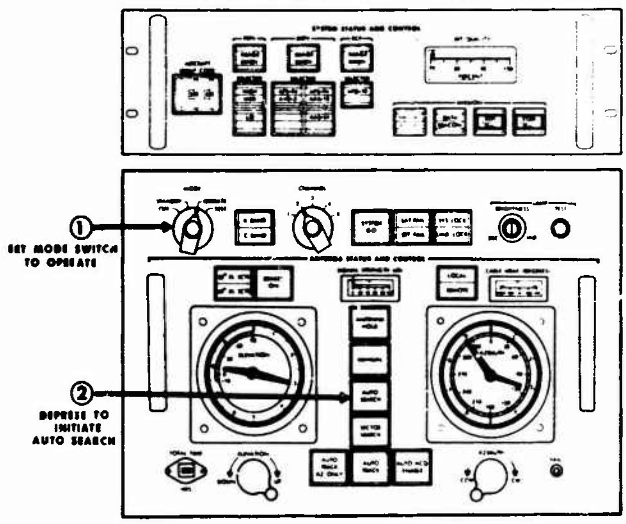
The Relay operator, in accordance with the mission plan, initiates link acquisition with the surface terminal as follows:

1. Sets mode selector switch to OPERATE.
2. Verifies that ANTENNA BEARING indicator is rotating.

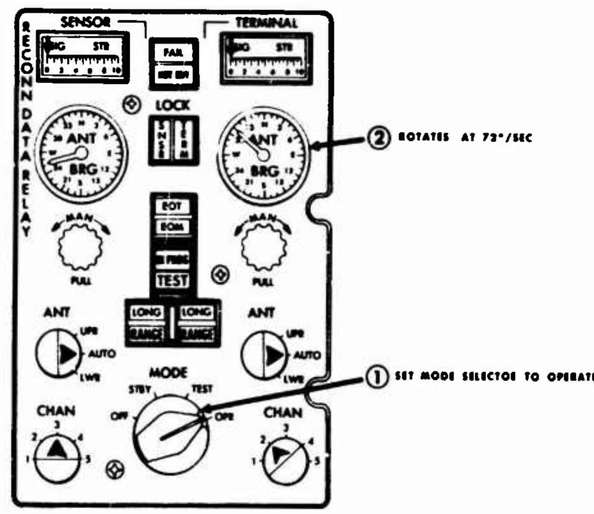
After enabling the system, the operator verifies that the following indicators are illuminated and modes selected:

- Range LONG (Sensor and Terminal)
- Antenna - AUTO
- Channel 3 for Sensor
- Channel 2 for Terminal

TERMINAL CONTROL PANEL (LINK ACQUISITION ENABLED)



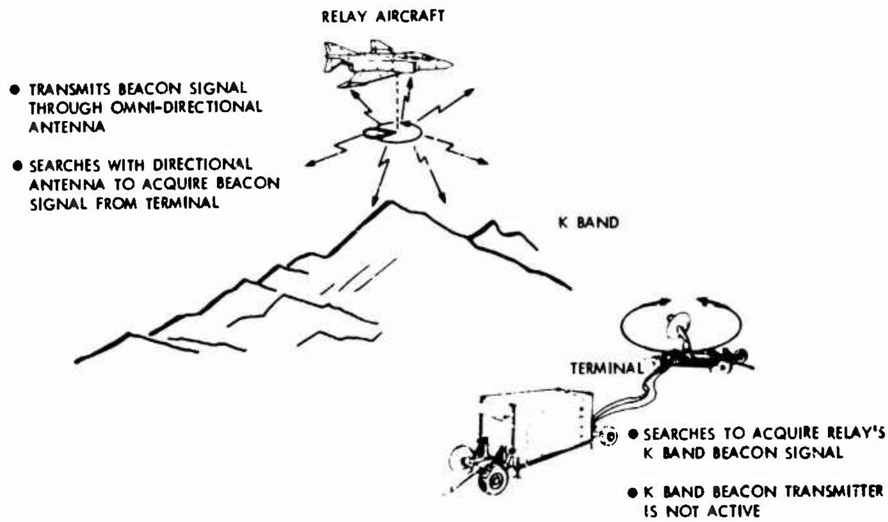
RELAY OPERATOR'S CONTROL PANEL (OPERATE INITIATED)



With both the subsystems enabled and active at this time, the following sequence of events takes place automatically:

- The Relay starts transmitting a beacon signal through the OMNI antenna.
- The Terminal Antenna starts searching for the Relay Beacon signal.
- The Relay aircraft directional antennas start searching for a beacon signal from the ground; however, this beacon has not been enabled.

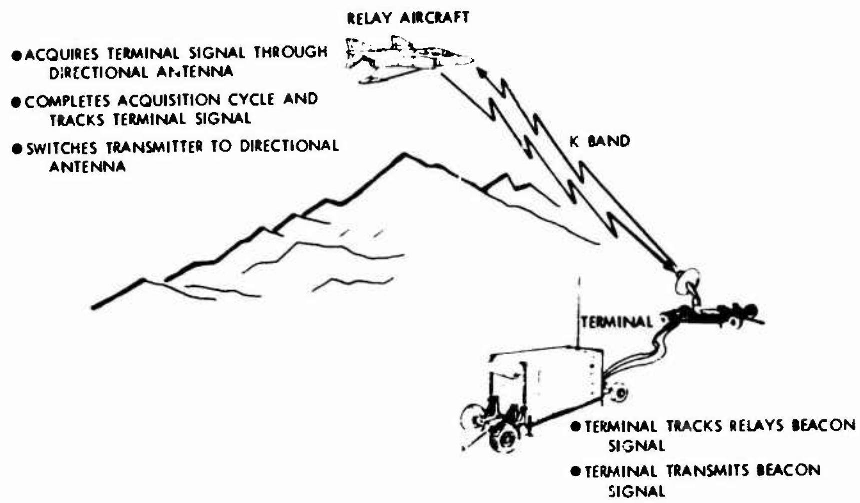
RELAY-TERMINAL LINK ACQUISITION INITIATED



The Surface Terminal Antenna first acquires the Relay aircraft beacon signal on the acquisition horns, decodes the ID and causes the antenna to start tracking the acquired signal. After positioning (pointing) the antenna in the correct azimuth and elevation, the antenna switches automatically to a pencil beam and the surface terminal beacon is enabled.

The Relay aircraft meanwhile is searching for the beacon signal on the directional antennas. When the beacon signal from the ground is acquired, the ID is decoded and the directional antennas stabilize in azimuth and elevation angle. At this time, the transmitter and beacon signal is switched to the directional antenna that has acquired the ground signal. Link-lock has now been established.

RELAY ACQUIRES TERMINAL (LINK LOCK)



While the acquisition cycle is being accomplished automatically, the Relay and Surface Terminal operators need only to monitor the process on their control panels. At the completion of acquisition, the operators will be presented with the following indications.

RELAY

1. Signal Strength Meter will advance to an optimum position.
2. Antenna Bearing Indicator will stabilize in azimuth.
3. Terminal Link-Lock will illuminate.

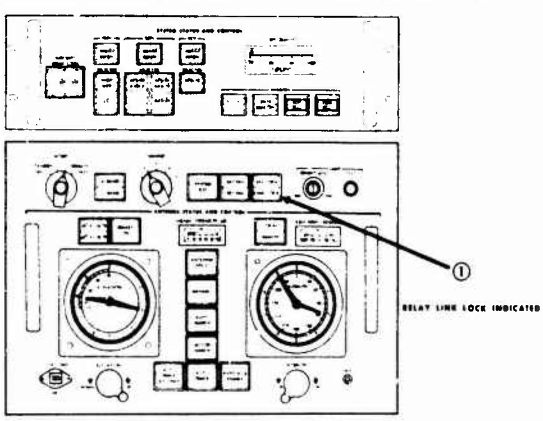
TERMINAL

1. Relay Link-Lock will illuminate.

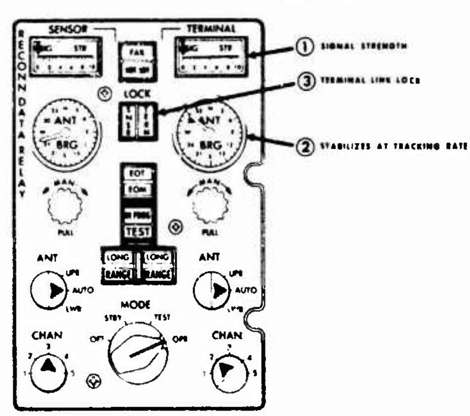
In addition, the signal strength meter will advance to an optimum position; the antenna azimuth indicator will nearly stabilize; the AUTO TRACK and AUTO ACQ ENABLE indicators illuminate; and the BEACON INDICATOR illuminates providing information that beacon data is received. In addition, the following indications are also illuminated:

- K-Band
- System GO
- 35° EL SCTR

TERMINAL CONTROL PANEL
RELAY ACQUIRES TERMINAL (LINK LOCK)



RELAY OPERATOR'S CONTROL PANEL
(RELAY ACQUIRES TERMINAL)

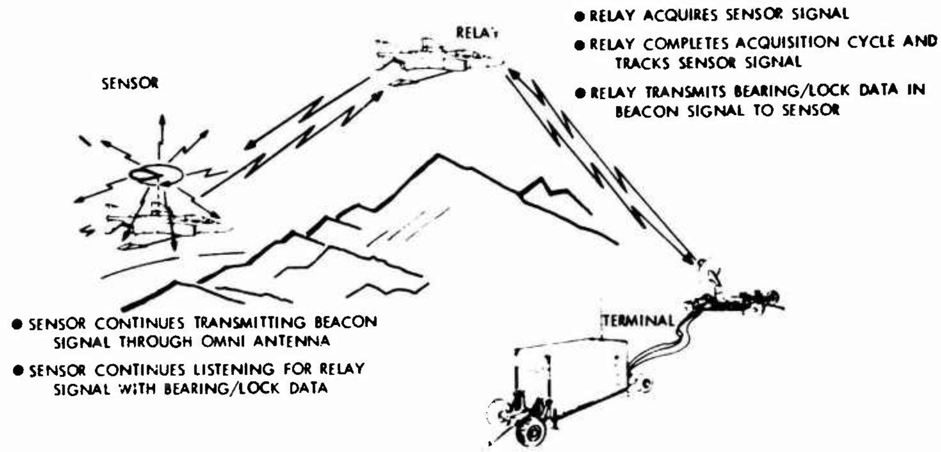


Just after the Relay and Terminal have established link-lock, the Sensor operator notes that he is scheduled to acquire the relay and sets the Mode Select Switch to OPERATE. This action enables the system for automatic acquisition of the Relay aircraft as follows:

- The Sensor starts transmitting a beacon signal through OMNI antenna.
- The Relay is searching for the Sensor beacon signal with the directional antennas.
- The Relay acquires the Sensor beacon signal; decodes the ID and completes the acquisition cycle; and begins tracking the Sensor signal.
- The Relay transmits bearing/lock data on the up-link beacon signal to the Sensor.

- The Sensor receives the Relay bearing/lock data through the omni antennas.
- The Sensor directional antennas respond to the bearing data provided and point in the direction of the Relay aircraft. At this time, the Sensor is ready to switch beacon transmission to the Directional Antennas.

RELAY ACQUIRES SENSOR



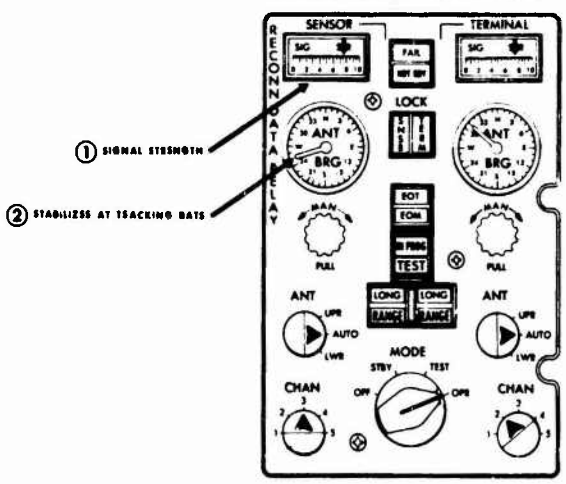
During this acquisition operation the relay operator need only monitor the Control Panel for indications. As soon as the relay has acquired the sensor signal and started tracking, the following indications are present:

1. The signal strength meter will advance to an optimum position.
2. The antenna bearing indicator will stabilize to indicate tracking of the sensor beacon signal.

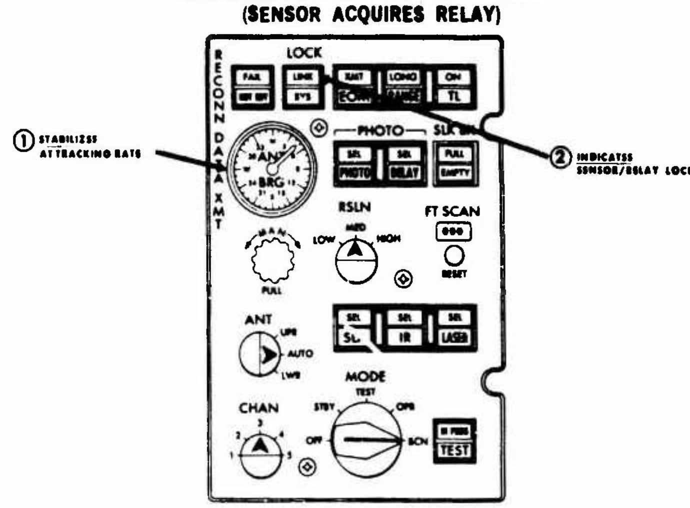
With the system locked between the Sensor and Relay aircraft, the following indications will be presented to the Sensor operator.

1. The antenna bearing stabilizes at tracking rate.
2. The LINK indicator illuminates.

RELAY OPERATOR'S CONTROL PANEL (RELAY ACQUIRES SENSOR)

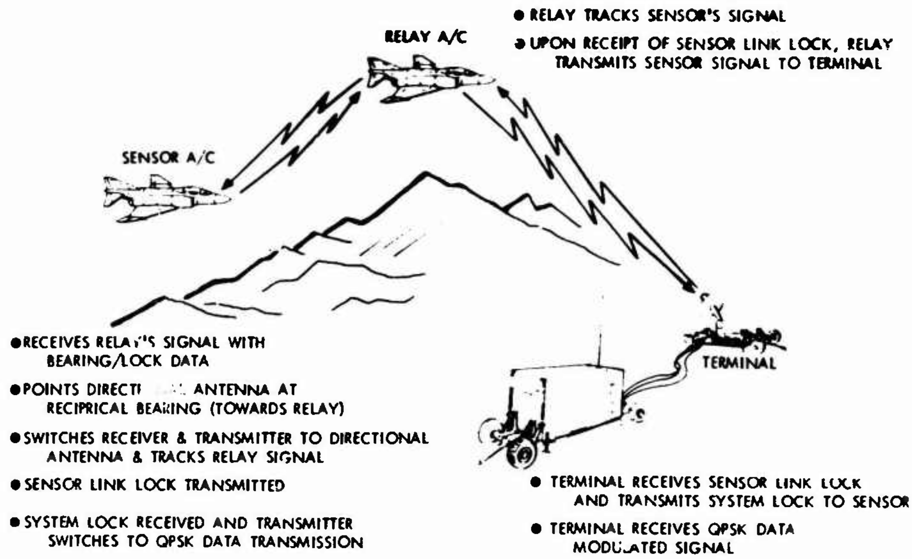


SENSOR OPERATOR'S CONTROL PANEL (SENSOR ACQUIRES RELAY)



Upon receipt of the bearing-lock data from the Relay, the Sensor responds as previously described. When the Directional Antennas on the Sensor aircraft have been pointed at the Relay, the Sensor switches the receiver and transmitter to them and starts tracking the Relay beacon signal. At this point, the Sensor and Relay are tracking each other and the Sensor transmits a link-lock signal to the Relay aircraft. Upon receipt of this signal, the Relay transmits the Sensor link-lock signal to the Surface Terminal. When the Terminal receives the Sensor link-lock signal it generates a system-lock signal and transmits it up-link to the Sensor aircraft. With the system locked so that data can be transmitted, the terminal begins receiving a QPSK modulated signal.

SENSOR ACQUIRES RELAY (SYSTEM LOCK)



With the link established between the Sensor and Relay aircraft, and the Relay aircraft and Surface Terminal, a system-lock has been established. At this time, the following indications are present on the operator's panels:

SENSOR

1. Link illuminated -- Sensor to Relay path established.
2. SYSTEM illuminated indicating a data path to the Surface Terminal.

Additional indicators illuminated are:

- PHOTO SELECT
- TL ON
- Slack Box Empty
- Range - LONG

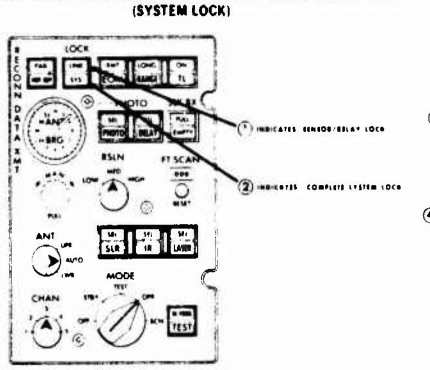
RELAY

1. Signal strength at terminal beacon signal carrier.
2. Lock indication with terminal.
3. Signal strength of sensor beacon carrier signal.
4. Lock indication with Sensor.

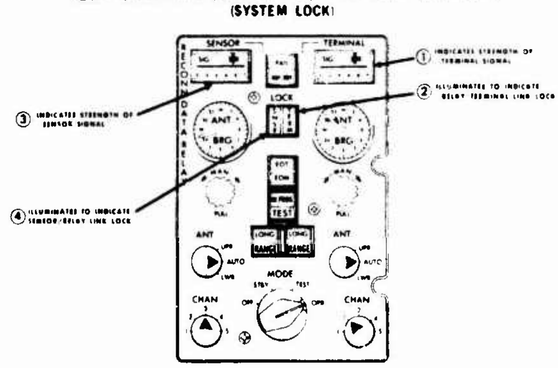
An additional indicator illuminated is:

- Range - LONG for both sensor and terminal

SENSOR OPERATOR'S CONTROL PANEL LINK LOCK INDICATIONS (SYSTEM LOCK)



RELAY OPERATOR'S CONTROL PANEL LINK LOCK INDICATIONS (SYSTEM LOCK)



System lock indications on the Surface Terminal control panel are as follows:

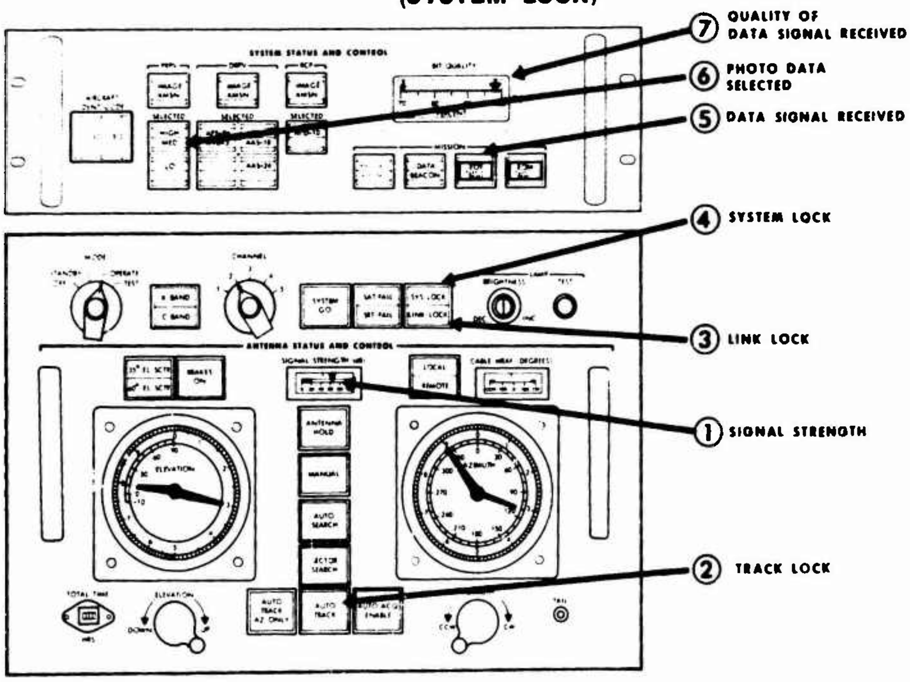
1. Signal strength of Relay beacon carrier.
2. TRACK LOCK illuminated indicating that the terminal is tracking the Relay.
3. LINK LOCK illuminated, indicating a data path between the Relay and Terminal.

5. DATA illuminated, indicating that the Sensor has switched over to DATA mode.
6. MEDIUM resolution, indicating photo sensor selected.
7. BIT quality, indicating the quality of the data signal received.

In addition, the following other indicators are illuminated:

- K-Band
- System GO
- 35° EL SCTR
- LOCAL

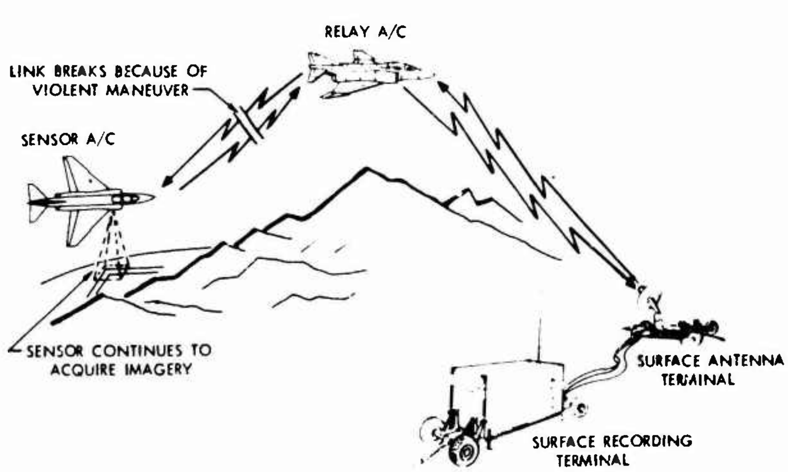
SURFACE TERMINAL CONTROL PANEL LINK LOCK INDICATIONS (SYSTEM LOCK)



The Sensor aircraft is now approaching the target area and is maintaining LINK LOCK with the Relay aircraft which is flying a holding pattern. At this point, the Sensor aircraft makes a rapid descent, leveling out at 2000 feet altitude over the Initial Point. Over the IP the sensor operator verifies system status and the selections made on his control panel. At the camera ON point, the KS-87 camera is activated on the reconnaissance panel and the JIFDATS equipment responds accordingly.

When a violent maneuver is executed, the limits of the antenna beamwidth in the vertical plane can be exceeded for a short interval of time. An explanation of the effects of the link break follows:

SENSOR - RELAY LINK BROKEN



When the pilot makes a violent maneuver and the Sensor/Relay link is lost, the Sensor operator is given the following indications on his control panel:

1. The SYSTEM and LINK indicator extinguishes.
2. The ANTENNA BEARING indicator continues to track at the previous rate for 10 seconds, then automatically switches to the acquisition rotation rate. Also, the FOOT SCAN meter stops advancing.

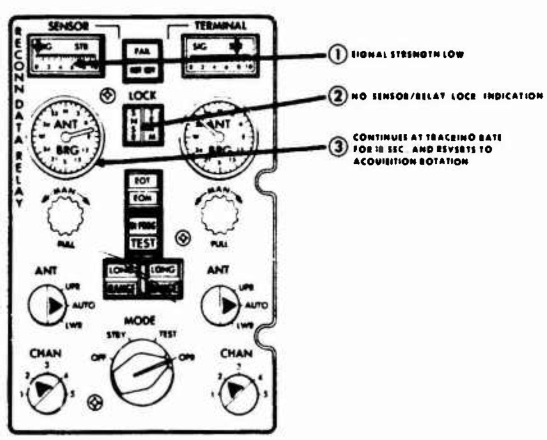
When the link was broken, the scanner stopped advancing film and data transmission terminates. The processor, however, continues to feed film into a secondary slack box until the link is automatically re-established. If the link is not established in 10 seconds, and this probably would happen, the Sensor would re-acquire the Relay using the procedures previously described.

The Relay operator, monitoring his control panel, notices the following indications when the link is broken:

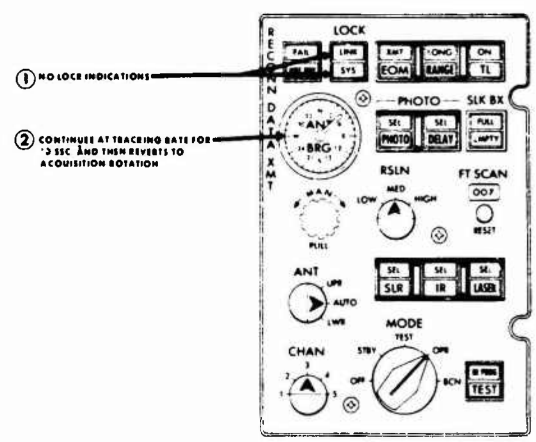
1. The SIGNAL STRENGTH meter drops until the link is re-established.
2. The Sensor LOCK light extinguishes until reacquisition.
3. The ANTENNA BEARING indicator continues to track for 10 seconds and then rotates at acquisition rate until it reacquires the Sensor beacon signal.

Since reacquisition is an automatic function, the operator has only to monitor his panel until the link to the Sensor is re-established.

RELAY CONTROL PANEL LINK LOSS INDICATIONS

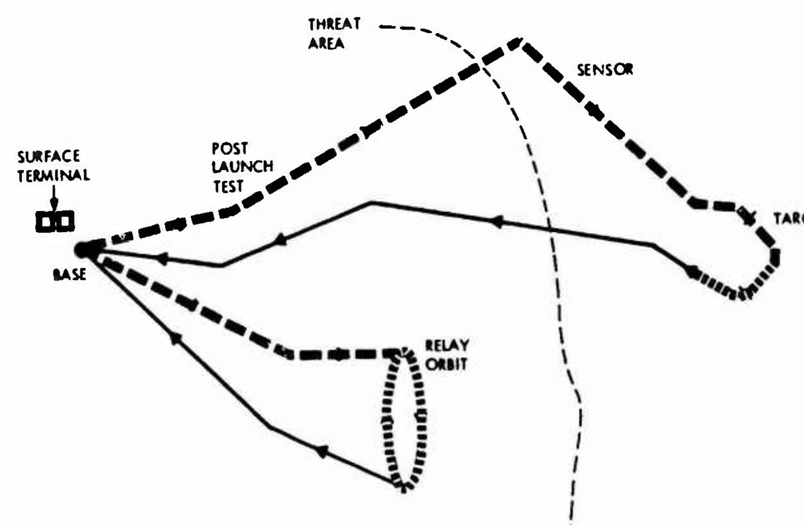


SENSOR CONTROL PANEL LINK LOSS INDICATIONS



During the momentary link break, the camera continues to acquire film and the reconnaissance run is completed with no further incidents as the link is re-established in a few seconds. Immediately after completing the run, the Sensor aircraft initiates a climbing turn, withdrawing from the target area and heads for base after reaching cruise altitude.

JIFDATS MISSION FLIGHT PLAN



When the reconnaissance aircraft completes the run over the target, the Sensor Operator shuts down the camera on the Reconnaissance Control Panel. Shutting the camera OFF, stops the film feeding into the IPPS. However, the IPPS still has film to process and scan for data transmission. Therefore, his only duty for the next several minutes will be to monitor the control panel for indication that processing and scanning is complete.

The Surface Terminal automatically recognizes that the PRPV operate command has disappeared when the last frame has been scanned by the IPPS in the Sensor aircraft.

This action, in turn, causes the Terminal to generate an End-Of-Target (EOT) signal which is transmitted to the Relay aircraft and displayed. The PRPV continues to process reconnaissance data until the last image recorded has passed the viewer. At that time the terminal PRPV transport will stop.

When all the film has been processed, scanned and the data transmitted, the sensor operator verifies this fact by noting the following:

- Slack Box EMPTY indicator illuminates.
- FEET SCAN Counter stops advancing.

At this time, the operator depresses the EOM indicator switch. This action causes a signal to be transmitted to the Surface Terminal and, in turn, it is transmitted to the Relay aircraft. If no more targets are scheduled, the operator sets the Mode Selector Switch to STANDBY and proceeds to Base.

Upon receipt of the EOM signal from the Sensor aircraft, the Surface Terminal operator down loads the JIFDATS reconnaissance imagery and prepares to deliver it to the Imagery Interpretation Facility for rapid target detection and other exploitation.

If the Interpretation Facility is in the immediate vicinity, the film will be quickly handcarried between shelters. If, for some operational reason, the Surface Terminal must be remotely located, vehicle couriers may be used to that extent necessary.

The mission results given in this illustration may be considered typical of route reconnaissance type sorties. Of interest is the fact that more than 50 feet of film can be made available for rapid target detection and command use approximately 12 minutes after the aircraft ended its reconnaissance run. For a point target requiring cover of only 5 feet of film, the time would be considerably less.

4. CONCLUSIONS

Among the advanced state-of-the-art technologies employed in several areas of JIFDATS are: 1) the digital modulation-multiplexing technique which extends the range of transmission, increases the rate of data transfer, and accommodates a wide variety of sensor output characteristics; 2) the design of the airborne directional antennas which have high performance characteristics for their small size; and 3) the film processing technique in the airborne photo processor-scanner set as well as the two surface recorder-processor-viewers, in which high resolution imagery is reproduced in near-real time.

This paper has attempted to describe, in general terms, the JIFDATS system and more specifically, the automatic acquisition and tracking techniques employed in pointing the directional antennas of the two aircraft and the ground terminal. At the time of its writing (June 1, 1972) the operation of this system has been verified by several months of flight testing at Edwards Air Force Base, California.

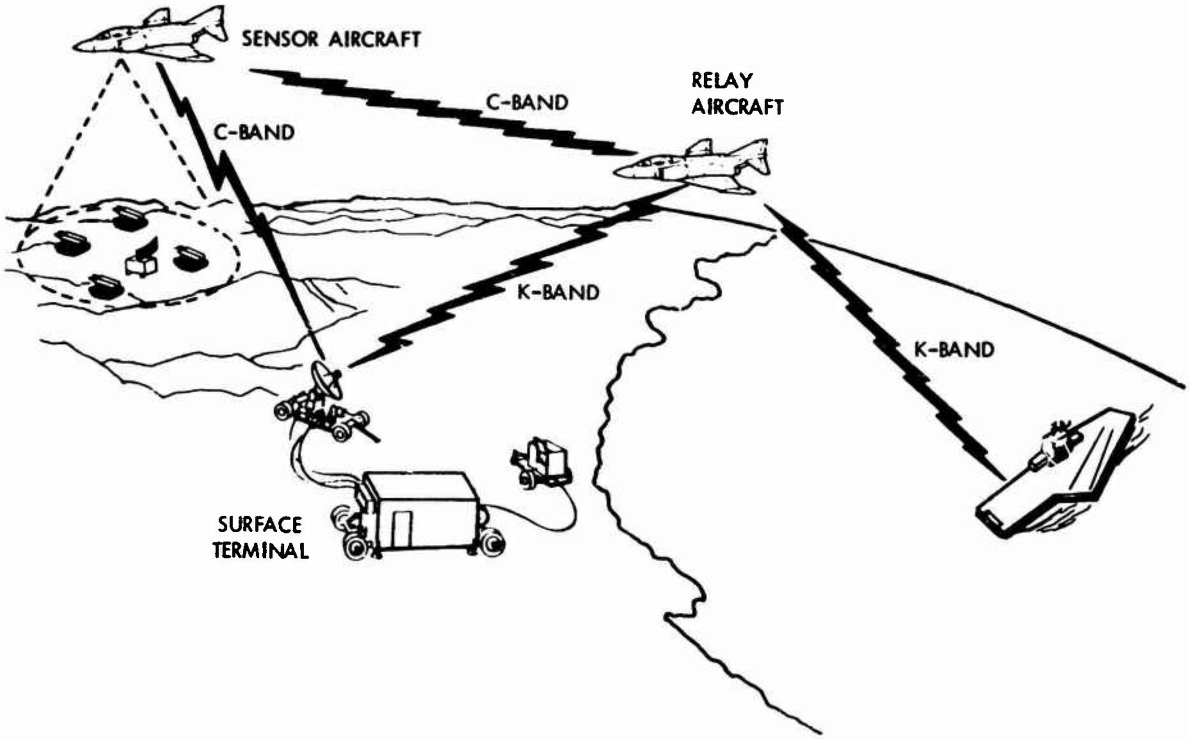


Fig.1 JIFDATS data transmission system

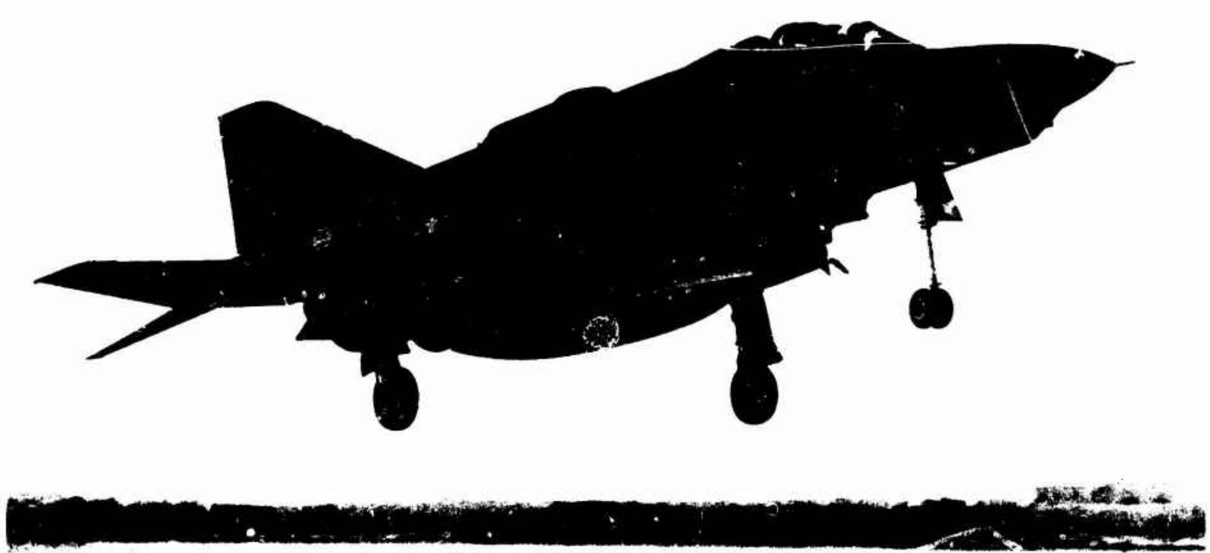


Fig.2 RF-4C sensor aircraft – flight test

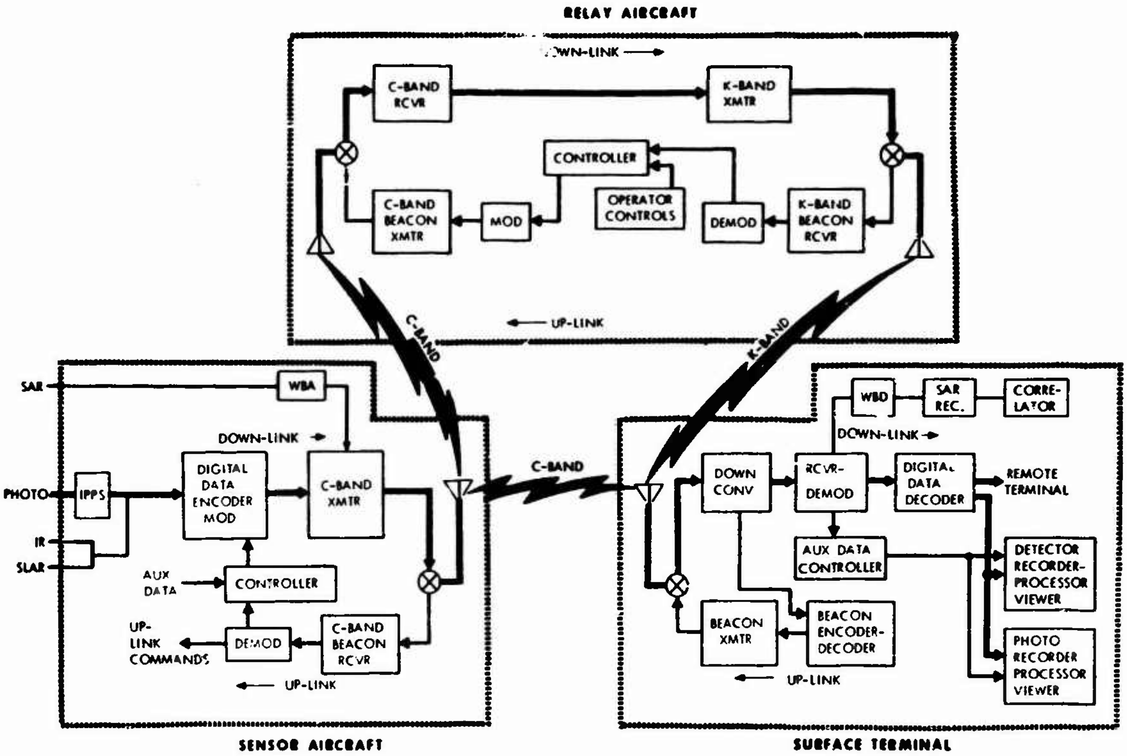


Fig.3 Data transmission system block diagram

CHARACTERISTICS	FREQUENCY BAND	BEAMWIDTH -3 db	GAIN -3 db	SIDE LOBES	TRACK METHOD
TRACKING ANTENNA	C K	2.3° (CONE) 0.69° (CONE)	32 db 44 dB	17.5 db	MONOSCAN
ACQUISITION ANTENNA	C K	AZ 11° EL 15° AZ 5.2° EL 10.75°	20 dB 24 dB	AZ 30 db EL 12.5 db	SEQUENTIAL LOBING

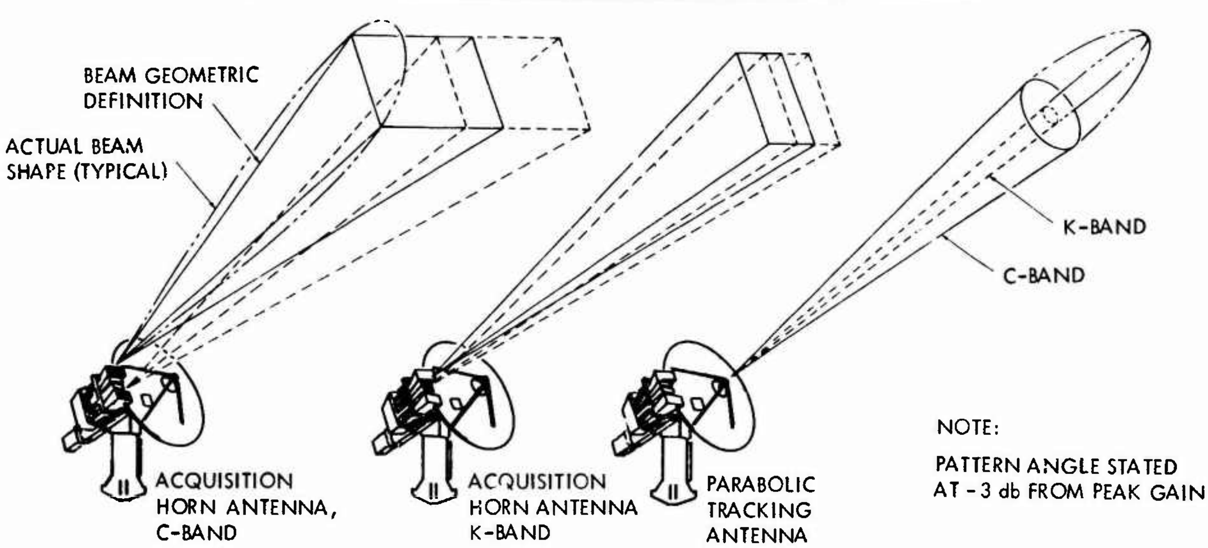


Fig.4 Antenna patterns – surface terminal

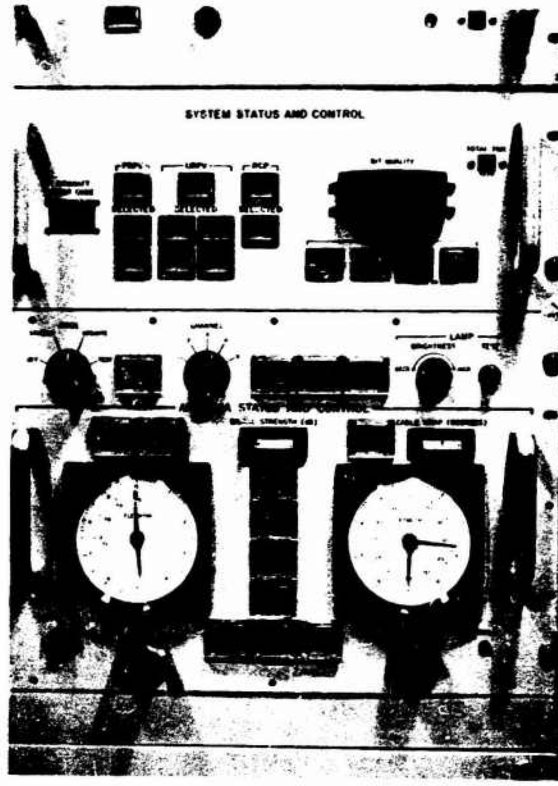


Fig.5 Surface terminal operator's control panel

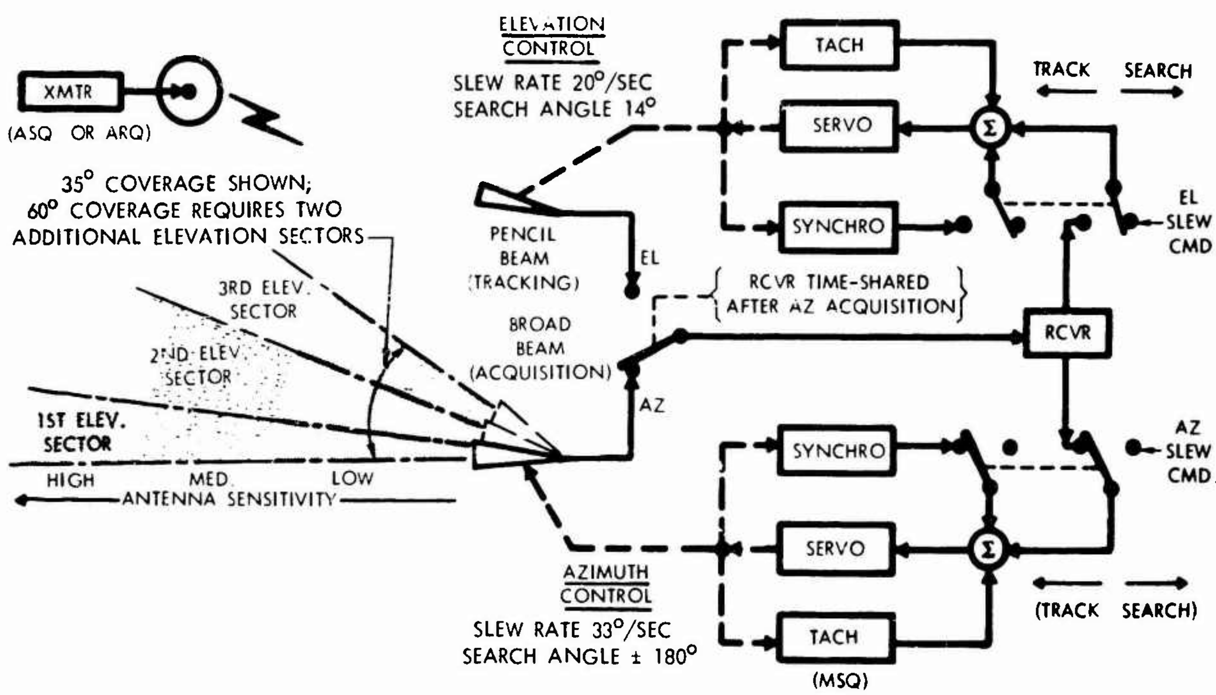


Fig.6 Acquisition – surface acquires sensor or relay

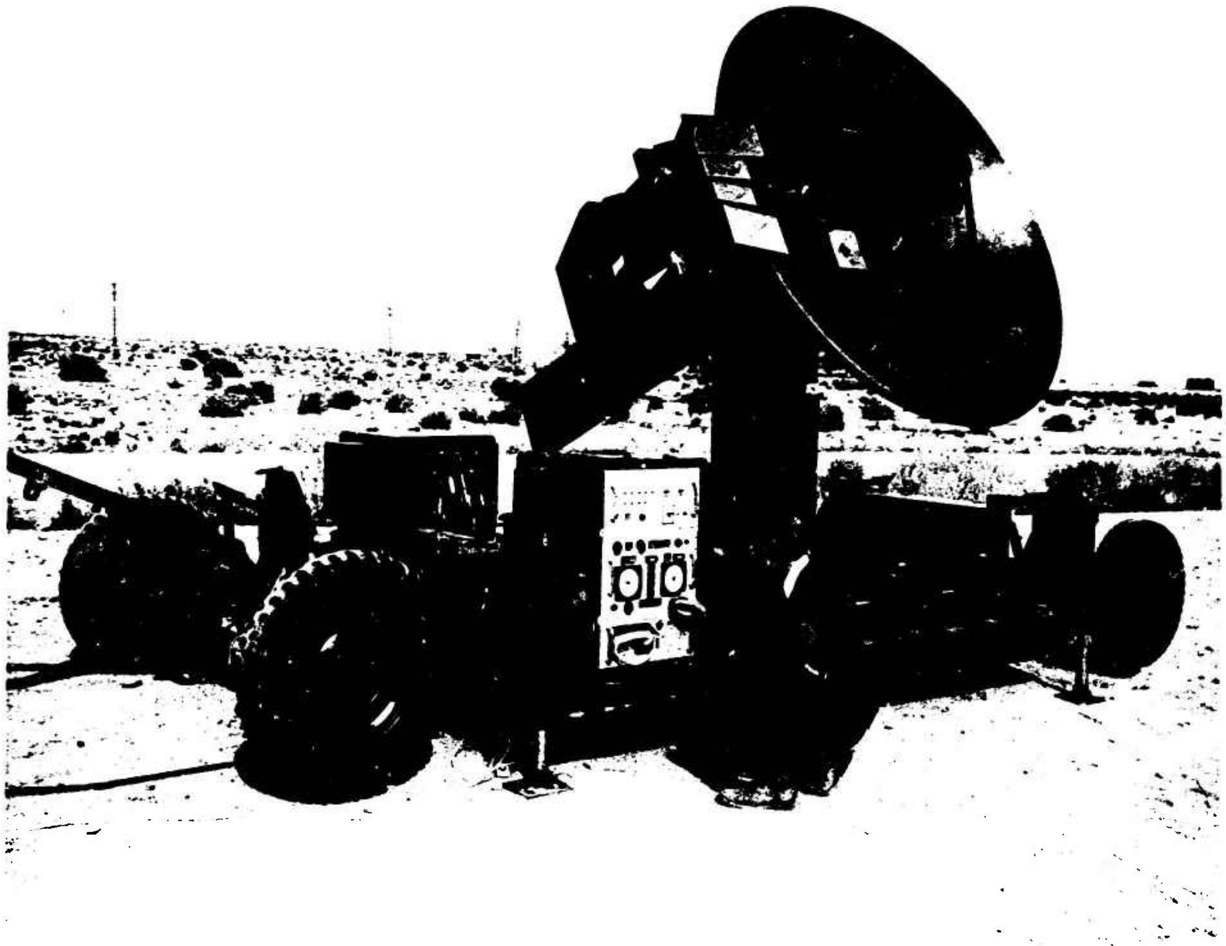


Fig.7 Surface antenna terminal at test site

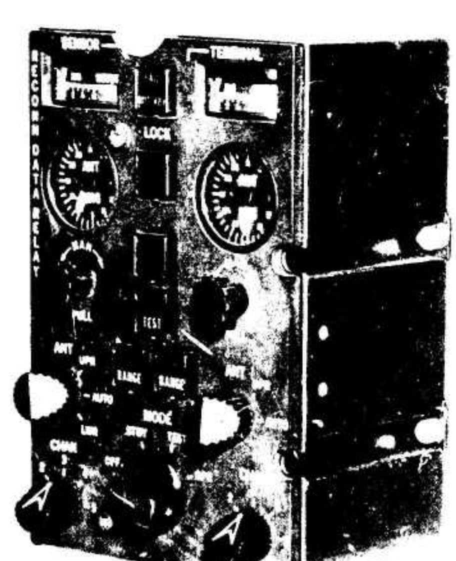


Fig.8 Relay aircraft operator's control panel

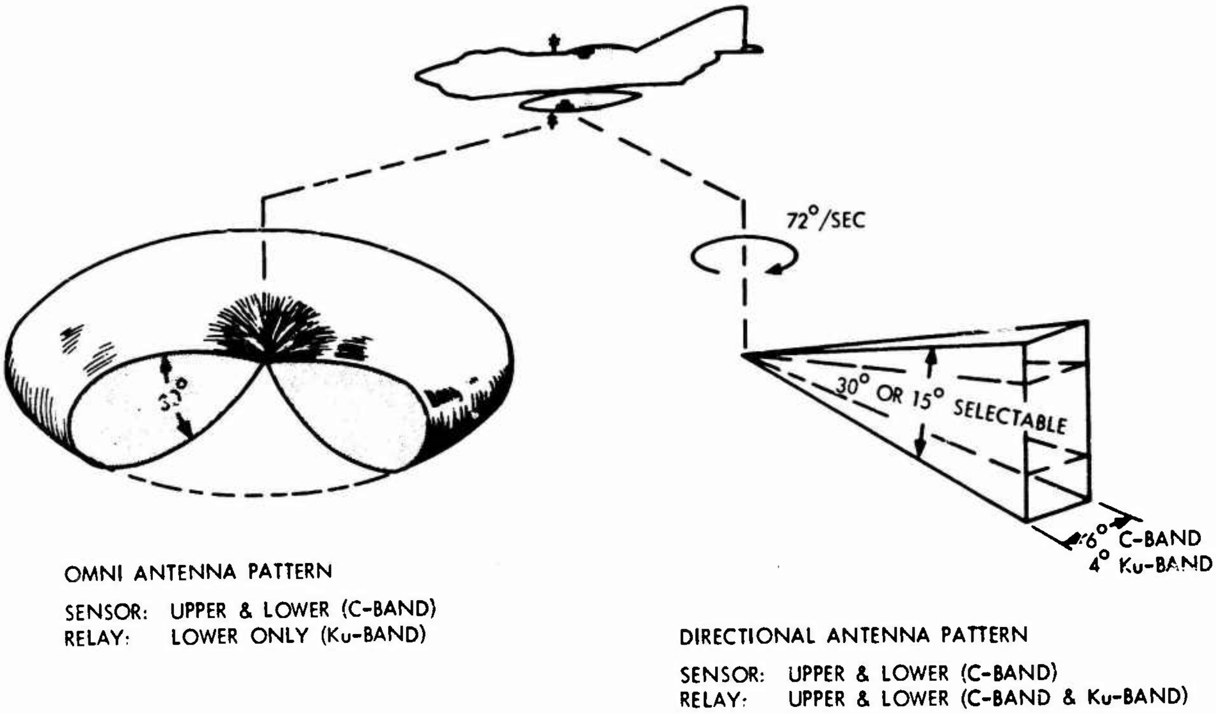


Fig.9 Antenna complement

Fig.10 Sensor aircraft operator's control panel

Fig. 11 HEDATS – [illegible] band and K band airborne directional antennas

DETERMINATION OF AN OPTIMAL TRAJECTORY IN THE PRESENCE OF RISK

A. Tiano, P. Dagnino, M. Piattelli

Laboratorio per l'Automazione Navale
Consiglio Nazionale delle Ricerche
Genova
Italy

SUMMARY

Let us consider a controlled dynamic system, displacing within an assigned space, where "r" moving targets are contained.

The purpose of this report consists in choosing an optimal control sequence transferring our system from an initial point to a preset terminal point. In our case, the optimal trajectory is the one which, complying with some safety constraints imposed by the targets, minimizes a given cost function. Assuming that the system may be supplied with periodical information about the motion of the targets, we can determine a numerical algorithm utilizing a dynamic programming procedure. This procedure is applied to two practical problems:
- marine anticollision aided by computerized radar systems in the presence of N targets;
- determination of an optimal evasion strategy in the presence of cyclonic disturbances.

The solution of the former application is subordinated to the presence on board of a computer interfaced with a radar, while the latter application is based on the availability of periodical radio weather messages.

1. INTRODUCTION

The development of transports by sea and the consequently increased shipping density have aggravated the problem of the prevention of collisions at sea.

Although ships have been equipped with modern radar systems and with different devices intended to facilitate the manual plotting, the collision danger is not yet completely averted; on the contrary, according to recent statistics, the frequency and cost of collisions at sea have increased remarkably, and this is mainly due to the larger size of the involved ships.

Furthermore, the traditional plotting presents the following limitations: the changes of course and speed of a ship are often assessed with a remarkable delay and therefore the necessary precautionary measures are taken too late to have sure prospects of success; the number of the contemporarily considered targets is very small owing to the laborious work required for the manual tracking of each of them; the check of the validity of a preset evasive manoeuvre requires a laborious plotting.

In order to obviate such difficulties, a number of anticollision systems have been developed involving the use of a computer interfaced with a radar. The computer aids the radar operator elaborating in real time the kinematic characteristics of a certain number of targets which are auto-tracked by the system.

The computer alerts the crew of collision danger by an appropriate alarm and points out the threatening target.

The operations that can be performed by such computerized anticollision systems are:
- auto-tracking of the targets by manual or automatic inizialization;
- automatic determination of the targets'position, speed and course;
- prediction of collision danger, deduced from the above information;
- eventual suggestion of evasive manoeuvres;
- check of the evasive manoeuvres.

Such systems, however, still present serious limitations: first of all, anticollision automatic systems supply the operator only with an informative picture of the situation, but they do not give any suggestion, in case of collision danger, about the tactical geometry or the strategy to be adopted in order to formulate a manoeuvre plan. Therefore, the operator must select a safe manoeuvre experimentally, and this imposes a remarkable mental strain on him, just when he finds himself under the psychological pressure of the impending danger.

The manoeuvre plans suggested by the operator are checked by the computer, which, however, tests only whether the immediate danger situation is about to be overcome or not. In the worst situations.a manoeuvre that had been formulated only on the ground of the threat assessment of a particular moment, may lead to the impossibility of reaching the preset destination or, anyhow, to a more threatening situation than the escaped one.

In short, the above mentioned serious limitations may be summarized as follows:
- the operator is demanded to perform an excessively laborious and hard work;
- the anticollision problem is faced only by an immediate tactical procedure and not by a strategic method.

The approach proposed in this article tries to obviate such difficulties, since:

- it presents a manoeuvre plan, with qualitatively optimal characteristics, to the operator who has only to take an acceptance decision;
- it calculates, in real time, the whole manoeuvre sequence extrapolated in the future as long as to reach the destination point, in other words, it presents a complete and constantly updated manoeuvre strategy.

2. STATEMENT OF THE PROBLEM

Let us consider own ship as a controlled dynamic system, whose state vector at time t, $\underline{\xi}(t)$, is described for every $t \in [t_o, +\infty)$ by the differential equation:

$$\frac{d\underline{\xi}}{dt} = \underline{f}\left[(\underline{\xi}, \underline{u}, t)\right] \qquad (2.1)$$

where the state vector $\underline{\xi} = (\alpha, \beta)$, constituted by the ship's position coordinates, belongs to a given sea region $X \subset R^2$.

The control vector $\underline{u} = (\nu, \vartheta)$, which is supposed to be constituted by the ship's speed and course, belongs to a suitable control set U, which takes the ship's characteristics into account.

The explicit form of Eq. (2.1) can thus be written componentwise:

$$\begin{cases} \dfrac{d\alpha}{dt} = \nu(t) \cos \vartheta(t) \\ \\ \dfrac{d\beta}{dt} = \nu(t) \sin \vartheta(t) \end{cases} \qquad (2.2)$$

Let us suppose that the above system must be transferred from an initial point $\underline{\xi}_o = \underline{\xi}(t_o)$ to a given terminal point $\underline{\xi}_f = \underline{\xi}(t_f)$ and that there is a cost functional $V(\underline{\xi}, \underline{u}, t_f)$ associated to each admissible trajectory $\underline{\xi}(t, \underline{u})$ connecting these points.

Within the same sea region X, there are "r" moving targets, the motions of which are described by equations analogous to Eq. (2.1):

$$\frac{d\underline{\eta}^k}{dt} = \underline{g}^k(\underline{\eta}^k, \underline{q}^k, t) \qquad k = 1,\dots r \qquad (2.3)$$

where the control vector $\underline{q}^k$, at the k-th target's disposal, is supposed to belong to a given control set $Q^k \subset R^2$.

Now, let us associate a risk function of the type:

$$\Phi^k : X \times X \longrightarrow \mathbb{R} \qquad k = 1,\dots r \qquad (2.4)$$

to each admissible trajectory of the system, taking some safety requirements, imposed by the presence of the moving targets, into account.

Supposing that at the initial time t_o we have:

$$\Phi^k(\underline{\xi}(t_o), \underline{\eta}^k(t_o)) \geqslant 0 \qquad k = 1,\dots r \qquad (2.5)$$

then we shall require for every $k = 1,\dots r$ and every $t \in [t_o, t_f]$ that

$$\Phi^k(\underline{\xi}(t, \underline{u}), \underline{\eta}^k(t, \underline{q}^k)) \geqslant 0 \qquad k = 1,\dots r \qquad (2.6)$$

for all $\underline{q}^k \in Q^k$.

In our case, the risk functions Φ^k are assumed to be of the form

$$\Phi^k = d(\underline{\xi}(t, \underline{u}), \underline{\eta}^k(t, \underline{q}^k)) - \varepsilon \qquad (2.7)$$

where ε is a positive number and d is the euclidean distance in $\mathbb{R}^2$; we shall thus have a collision avoidance problem. Therefore, the constraints expressed by Eq. (2.6) mean that, at each time instant, collision is avoided if the distance from each target is not inferior than a preset safety value ε.

The problem we want to solve may be formulated as follows:
"Determine a control $\underline{u}(t)$, based upon observations of the targets' trajectories, which transfers the system (2.1) from the initial point $\underline{\xi}_o$ to a given terminal point $\underline{\xi}_f$, and minimizes a given cost functional $V(\underline{\xi}, \underline{u}, t_f)$, while avoiding collision till time t_f."

From the operational point of view, we have thought it rather justified to adopt the transfer time

$t_f - t_o$ as the cost, since in most cases, this is the same as minimizing the ship's cost.

The solution of this problem within the variational optimal control theory, even in such a simple case, is rather difficult because of the presence of time-dependent constraints, in relation to the time evolution of the above mentioned targets.

In this connection, we mention the solution proposed by Warga [1] within the relaxed control theory and the solution suggested by Friedmann [2] within the pursuit-evasion differential games.

In this article, however, in accordance with the need of a numerical solution, the problem will be expressed as a multistage decision process and will be solved by Dynamic Programming.

Anyhow, we observe that a solution of the problem involving a time-continuous choice of the control law $\underline{u}(t)$, based on continuous observations of the targets' states, may be undesirable from the practical point of view. In particular, the problem has been studied with the view of an immediate operational application aboard ships equipped with computerized anticollision radar systems.

3. COLLISION AVOIDANCE AS MULTISTAGE DECISION PROCESS

Let us reformulate our problem in terms of a multistage decision process.

The proposed method allows us to obtain a numerical solution by means of an heuristic adaptation of the Dynamic Programming algorithm, which takes the risk functions, associated to the moving targets, into account. Therefore, let us discretize the ship's possible routes which are obtained by joining the initial to the terminal point. For this purpose, we utilize the grid shown in Fig. 1, which consists of a finite number of possible crossing points.

We consider the columns of the crossing points as stages of the process and associate a state variable x_n to the n-th stage; x_n is supposed to take only a finite number of values which are represented by the crossing points of the n-th column.

In this way, the set of the possible states belonging to stage n is:

$$X_n = \left\{x_n^1, \dots x_n^{k_n}\right\} \qquad n = 0, \dots N \qquad (3.1)$$

where k_n denotes the number of the possible crossing points of the n-th column. The process lasts N+1 stages.

The geometrical structure of the grid, which depends on the presence of natural obstacles and also on computing opportunities, will thus be completely defined as follows:

$$\begin{array}{l} x_o \in X_o = \{x_o^1\} \\ \quad\vdots \qquad \vdots \\ x_n \in X_n = \{x_n^1, \dots x_n^{k_n}\} \\ \quad\vdots \qquad \vdots \\ x_N \in X_N = \{x_N^1\} \end{array} \qquad (3.2)$$

Supposing that at time t_n own ship is at state x_n, a decision must be taken about the value of the state x_{n+1} we want to reach at the subsequent stage.

It is convenient to reduce the number of the possible decisions by introducing a transition operator Γ_{n+1}, which, for every fixed value x_n^i of the state variable x_n, defines a set of possible values for x_{n+1}, given by:

$$\Gamma_{n+1} x_n^i = \left\{x_{n+1}^{i_1}, \dots x_{n+1}^{i_j}, \dots x_{n+1}^{i_n}\right\} \; ; \quad 1 \leq i_1 \leq i_n \leq k_{n+1} \qquad (3.3)$$

Of course, the following relation must be verified

$$X_{n+1} = \bigcup_{i=1}^{k_n} \Gamma_{n+1} x_n^i \quad ; \quad n = 1, \dots N-1 \qquad (3.4)$$

Reciprocally, if x_{n+1} is given, x_n must belong to a certain set $\Gamma_{n+1}^{-1} x_{n+1}$, consisting of those states from which x_{n+1} may be reached.

We shall apply the term "policy from state x_o to state x_N" to any sequence of possible states such that:

$$x_n \in \Gamma_n x_{n-1} \cap \Gamma_{n+1}^{-1} x_{n+1} \qquad (3.5)$$

The set of all possible policies thus consists of the set of all discretized routes joining the initial to the terminal point.

To each pair (x_n, x_{n+1}), in accordance with the constraints expressed by Eq. (3.3), let us associate a value of the transition cost $V_{n+1}(x_n, x_{n+1})$, represented by the time elapsed to perform the transition itself.

We shall assume that every transition occurs at a constant speed, which must be not greater than the maximum speed allowed by the sea conditions and by the currents between states x_n and x_{n+1}. Therefore, the constant speed value will be:

$$\nu = \nu(x_n, x_{n+1})$$

The cost of the transition is then:

$$V_{n+1}(x_n, x_{n+1}) = \frac{d(x_n, x_{n+1})}{\nu(x_n, x_{n+1})} \tag{3.6}$$

where $d(x_n, x_{n+1})$ is the distance between the two states.

Since the cost is represented by time, we can utilize the additive property, and therefore we can associate a total cost represented by

$$\sum_{n=0}^{N-1} V_{n+1}(x_n, x_{n+1})$$

to any policy.

Bearing the kinematic aspect of our problem in mind, the equations describing the ship's motion during any transition from a state x_n to another state x_{n+1} are of the type:

$$\begin{cases} \alpha(t_n+\tau) = \alpha_n + \nu(x_n, x_{n+1}) \cos \vartheta_n \tau \\ \beta(t_n+\tau) = \beta_n + \nu(x_n, x_{n+1}) \sin \vartheta_n \tau \qquad \tau \in [0, V_{n+1}] \end{cases} \tag{3.7}$$

where (α_n, β_n), $(\alpha(t_n+V_{n+1}), \beta(t_n+V_{n+1}))$, are respectively the position coordinates of the crossing points x_n and x_{n+1}, t_n is the time instant at which x_n is reached and ϑ_n is the heading determined by the track joining x_n to x_{n+1}. Thus, each decision corresponds to the choice of a constant value $(\nu(x_n, x_{n+1}), \vartheta_n)$ of the control vector, on the time interval $[t_n, t_n + V_{n+1}]$ during which the transition occurs.

Furthermore, such decisions, based on observations of the targets' positions, must take the safety constraints, expressed by Eq. (2.6), into account.

For this purpose, let us assume that the targets are periodically observed at some given time instants $t'_m = t_o + m \cdot \Delta T, (m = 0, 1, \ldots)$.

Supposing to know the Eqs. (2.3), which describe the targets' movements for each admissible control vector $\underline{q}^k \in Q^k$, we can, on the basis of the latest m-th position observation, determine for every subsequent time t, a set of reachable points $R^k_{t'_m}(t)$, consisting of the positions where the k-th target may be found at all time instants $t > t'_m$ (see Neustadt [3], and Sugino [4]).

These sets define the forbidden areas, delimited by moving contour lines, from which we must keep, during every transition, at a distance not inferior than the preset safe value ε. We shall deal later with a practical way for easily computing the reachable sets under the hypothesis of radar observations.

Let us associate a value of the risk $\Phi^k_{n+1}(x_n, x_{n+1}; t_n)$ to the n-th transition, represented by:

$$\Phi^k_{n+1}(x_n, x_{n+1}; t_n) = \min_{\tau \in [0, V_{n+1}]} \left[d\left(\underline{\xi}(t_n+\tau), R^k_{t'_m}(t_n+\tau)\right) - \varepsilon \right] \tag{3.8}$$

$$k = 1, \ldots r$$

where $\underline{\xi}(\cdot) = (\alpha(\cdot), \beta(\cdot))$ is the ship's position vector describing her transition from x_n to x_{n+1} according to Eq. (3.7), and t_n is the time instant at which such transition begins.

The evaluation of the risk, expressed by Eq. (3.8), connects the minimum distance from the k-th target's reachable set, at which the ship will find herself while travelling from x_n to x_{n+1}, to the safety radius ε. The minimum value of the distance exists under fairly general hypotheses on Eqs. (2.3) and on control sets Q^k, which assure the compactness of the reachable sets and their continuity with respect to time (see Neustadt [3]).

The safety constraints (2.6) require that

$$\Phi^k_{n+1}(x_n, x_{n+1}; t_n) \geq 0 \qquad ; \forall k = 1, \ldots r \tag{3.9}$$

The problem we are dealing with consists thus in determining the optimal policy $(x_0,\dots x_N)$ from the initial state x_0 to the terminal state x_N, which minimizes the travel time and satisfies the constraints (3.9) for each $n = 0,\dots N-1$.

Owing to the dependence of the targets' reachable sets $R^k_{t'_m}(t)$ on index m, we note that the case $m = 0$ corresponds to the OFF-LINE determination, while the case $m > 0$ corresponds to the ON-LINE implementation.

4. COMPUTERIZED RADAR SYSTEMS FOR COLLISION AVOIDANCE

The above introduced targets' reachable sets are of the form:

$$R^k_{t'_m}(t) = \left\{ (\alpha^k, \beta^k) \in X; \quad \begin{array}{l} \alpha^k = \alpha^k_m + \int_{t'_m}^{t} w^k(\tau)\cos\psi^k(\tau)d\tau \\ \beta^k = \beta^k_m + \int_{t'_m}^{t} w^k(\tau)\sin\psi^k(\tau)d\tau \end{array} \quad ; (w^k,\psi^k) \in Q^k; \right\} \qquad (4.1)$$

$$k = 1,\dots r$$

where, according to Eqs. (2.3), $\eta^k(t_m) = (\alpha^k_m, \beta^k_m)$ is the k-th target's position vector observed at time t'_m, and $q^k = (w^k,\psi^k)$ is its control vector, which is constituted by the pair speed and heading.

It is worth noting that the use of the reachable sets, obtained from (4.1) is quite unprofitable from a practical point of view, since such sets, by varying q^k on Q^k, may grow over and over, as time goes by, generating thus too large forbidden areas.

In order to eliminate this inconvenience, we shall make some simplifying assumptions on the targets' motion, basing ourselves on the availability on board of a suitable computerized radar system, which may supply us with complete information about the targets' kinematic parameters.

Computerized anticollision radar systems (which must be regarded as a part of more general integrated navigation systems, like the one shown in Fig. 7) are obtained interfacing a radar with a computer, which performs automatic calculations of the targets' kinematic parameters. Such calculations, which consist mainly in converting range and bearing into position, speed and heading, are based on filtering techniques which combine recursively radar measurements at successive times, in order to reduce the random errors affecting such measurements. Once the targets' kinematic parameters have been determined, estimates of their present and future positions can be made with accuracy, if their equations of motion are known a priori.

For this purpose, we shall assume that the targets' motion, during the time interval ΔT between two subsequent observations, is described by equations of the type:

$$\begin{cases} \alpha^k(t'_m + \tau) = \alpha^k_m + w^k_m \cos\psi^k_m \tau \\ \beta^k(t'_m + \tau) = \beta^k_m + w^k_m \sin\psi^k_m \tau \end{cases} \qquad \begin{array}{l} k = 1,\dots r \\ m = 0,1,\dots \\ \tau \in [0,\Delta T] \end{array} \qquad (4.2)$$

where $\alpha^k_m, \beta^k_m, w^k_m, \psi^k_m$ are the values of the kinematic parameters at time t'_m.

Furthermore, let us suppose to know, according to radar statistics, the corresponding errors $\Delta\alpha^k_m, \Delta\beta^k_m, \Delta w^k_m, \Delta\psi^k_m$, affecting the observations.

Eqs. (4.2) do not mean but predicting the targets' motions at successive time instants, on the basis of the most recent observation and of the assumption of uniform rectilinear motion.

This assumption is satisfactory, according to the normal operational praxis, as may be inferred from the following practical considerations:

a) the uniform rectilinear motion is the normal operational condition of merchant ships and therefore is the most likely condition;
b) the time interval ΔT between two subsequent observations is chosen in such a way that, within its duration, the transients due to course variations, which are quite determining for anticollision, can be considered as extinguished.

The targets' reachable sets $R^k_{t'_m}(t)$ (we had better call them predicted sets) are obtained by extrapolating the uniform rectilinear motion to all time instants $t > t'_m$ and by contemporarily taking the errors, affecting the measurements of the kinematic parameters, into account.

Since the position estimate is inferred from the estimate of speed and heading, let us associate, to each moving target, an angle amplitude $2\Delta\psi^k_m$, whose vertex lies in the position (α^k_m, β^k_m) at time t'_m, and whose axis coincides with the heading direction ψ^k_m.

This assumption, together with the consideration of the error Δw^k_m on the speed measurement, leads us to suppose that, for each $t > t'_m$, the k-th target lies, with **probability 1**, within the reachable **set** constituted by the intersection between an angular and a circular region (see Fig. 2):

$$R^k_{t'_m}(t) = \left\{ (\alpha^k, \beta^k) \in X \; ; \; \begin{array}{ll} \alpha^k = \alpha^k_m + v^k \cos\psi^k (t - t'_m) & v^k \in \left[v^k_m - \Delta v^k_m, \; v^k_m + \Delta v^k_m \right] \\ \beta^k = \beta^k_m + v^k \sin\psi^k (t - t'_m) & \psi^k \in \left[\psi^k_m - \Delta\psi^k_m, \; \psi^k_m + \Delta\psi^k_m \right] \end{array} \right\} \quad (4.3)$$

Accordingly, the probability that the target, at time t, is out of such a set, is assumed to be zero.

5. APPLICATION OF THE PRINCIPLE OF OPTIMALITY

We determine now an optimal OFF-LINE policy for the collision avoidance problem, by means of a dynamic programming procedure.

First of all, we observe that, in the absence of targets, the possible transitions of the system from one state to another of the following stage are controlled only by the geometric operator Γ_{n+1} which acts on the generic state x_n of the n-th stage according to Eq. (3.3).

In the presence of a risk connected with the time evolution of the targets, it is opportune to introduce a risk operator $I_{n+1}(t_n)$ acting on the states reachable from x_n, and depending on the time t_n at which this state is reached.

This operator, owing to the safety constraints (3.9), acts explicitly as follows:

$$I_{n+1}(t^i_n)\,\Gamma_{n+1}\,x^i_n = \begin{cases} \left\{ x^{i_1}_{n+1}, \dots, x^{i_j}_{n+1}, \dots, x^{i_n}_{n+1} \right\} & \text{if } \Phi^k_{n+1}(x^i_n, x^{i_j}_{n+1};\, t^i_n) \geq 0 \quad \forall k=1,\dots r \;\; \forall j=1,\dots n \\ \left\{ x^{i_1}_{n+1}, \dots, x^{\hat{i}_j}_{n+1}, \dots, x^{i_n}_{n+1} \right\} & \text{if } \Phi^{\bar{k}}_{n+1}(x^i_n, x^{\hat{i}_j}_{n+1};\, t^i_n) < 0 \quad \text{for some } \bar{k} \end{cases} \quad (5.1)$$

Therefore, the operator $I_{n+1}(t_n)$ inhibits the possible transitions from a fixed state x^i_n to those reachable states x^{ij}_{n+1} such that, for one target at least, the value of the corresponding risk is negative.

The transition risks $\Phi^k(x_n, x_{n+1};\, t_n)$ are obtained, as we have seen in Section 3, by evaluating the reachable sets $R^k_{t_o}(t_n+\tau)$ at each time instant, during the transition time interval $[t_n, t_n+V_{n+1}]$.

We shall suppose that the value of t_n is represented by the time associated to the subpolicy transferring our system from x_o to x_n in an optimal way (that is in a minimum time). We assume $t_o = 0$ as the initial condition.

For each stage, let us now define a set N_n of forbidden states, which is constituted by those crossing points which cannot be reached anyhow.

The sets N_{n+1} of forbidden states belonging to each stage subsequent to the initial one, can be determined recursively forward, once the set N_n of forbidden states of the preceding stage is known, and utilizing the risk operator $I_{n+1}(t_n)$.

For this purpose, to each state $x_{n+1} \in X_{n+1}$, let us associate the set $Y_n(x_{n+1})$, which is constituted by those not forbidden crossing points of the previous stage from which x_{n+1} can be reached under safety conditions (see Fig. 3). It thus follows that

$$Y_n(x_{n+1}) = \left\{ x_n \in (\Gamma^{-1}_{n+1}\, x_{n+1} - N_n) \,/\, I_{n+1}(t_n)\,\Gamma_{n+1}\, x_n \cap \{x_{n+1}\} \neq \emptyset \right\} \quad (5.2)$$

$$n = 0,\dots N-1$$

and that

$$x_{n+1} \in N_{n+1} \Longleftrightarrow Y_n(x_{n+1}) = \emptyset \quad (5.3)$$

Obviously, the initial state x_o is not forbidden, i.e. it must be $N_o = \emptyset$.

Now we can solve our problem utilizing Bellmann's iterative equation, for each not forbidden state of every stage:

$$f_{o,n+1}(x_o, x_{n+1}) = \min_{x_n \in Y_n(x_{n+1})} \left[f_{o,n}(x_o, x_n) + V_{n+1}(x_n, x_{n+1}) \right] \quad (5.4)$$

$$x_{n+1} \in (X_{n+1} - N_{n+1})$$

$$n = 0,\dots N-1$$

where $f_{o,n}(x_o, x_n) = t_n$ is the minimum temporal cost associated to state x_n, with the initial condition $f_{o,o}(x_o, x_o) = 0$.

Under the hypothesis that $X_{n+1} \supsetneq N_{n+1}$, for every $n = 0,\dots N-1$, Eq. (5.4) will allow us to determine the optimal policy $(x_o,\dots x_N)$ and the associated minimum total cost $f_{o,N}(x_o, x_N)$.

Now, let us see, step by step, how the algorithm proceeds (see the flow chart of Fig. 4).

We shall begin by considering all the transitions from the initial state x_o to each state x_1 of the first stage, and by computing the corresponding costs $V_1(x_o, x_1)$. For each transition, two cases may occur:

$$Y_o(x_1) = \begin{cases} x_1 \\ \text{or} \\ \emptyset \end{cases} \qquad (5.5)$$

according to whether state x_1 is reachable or forbidden. In the former case, we shall have $f_{o,1}(x_o, x_1) = V_1(x_o, x_1)$, while in the latter case, we shall have $x_1 \in N_1$.

The knowledge of the set N_1 of forbidden states, and of the minimum costs $f_{o,1}(x_o, x_1)$ associated to not forbidden states, is utilized at the second stage, with the aim of determining the sets $Y_1(x_2)$ for each $x_2 \in X_2$; from such sets, it is then possible to determine, according to Eq. (5.3), the set N_2 of forbidden states of the second stage, and, according to Eq. (5.4), the minimum costs $f_{o,2}(x_o, x_2)$ associated to not forbidden states.

The procedure goes on as long as to reach the terminal state x_N.

Of course, for the existence of the solution, it is required that no stage should be wholly constituted of forbidden states. Should this occur, it is necessary to change own ship's speed and/or the geometrical structure of the grid, and to perform new computations.

The suggestions with which the operator is supplied, as a result of this first calculation, will be: "keep a given heading for a given time at a given speed".

Actually, the calculation of the minimum times associated to the crossing points of the optimal policy, must be performed, taking own ship's manoeuvrability into account. For this purpose, a "steering function" is utilized which, from the knowledge of own ship's speed and steerage angle, allows us to compute the manoeuvre time. Such time is equally subdivided between two subsequent tracts.

The approximations on manoeuvrability are verified by computing, at the end of the program, a tolerance δT on time (i.e. the maximum delay within which every crossing point of the optimal route may be reached under safety conditions).

6. ON-LINE IMPLEMENTATION OF THE DYNAMIC PROGRAMMING PROCEDURE

The optimal off-line policy for the collision avoidance problem must be periodically checked according to the radar information flow about the motion of the targets and to the monitoring of own ship's position. In other words, at each time instant $t'_m = t_o + m\,\Delta T$, when information on the kinematic parameters of every target are available, a decision must be taken whether to leave the optimal precomputed route unchanged or not. In the latter case, the entire optimal control sequence is recomputed, by using the computer in an on-line mode.

Such on-line implementation proceeds according to the scheme shown in Fig. 5. The on-line program, the flow chart of which is shown in Fig. 6, is composed of the following steps:

1 - Own ship's actual position, speed and heading are monitored. The grid has been constructed in such a way that the shortest transition time is much longer than the radar updating interval; after the ship has left the initial point and is following the precomputed route, as information arrive, it is necessary to compute the distance to go to the next optimum crossing point and the new corresponding arrival time. Supposing that own ship is travelling between x_n and x_{n+1}, in the case of a remarkable deflection from the precomputed optimal route, we must update the optimum times $(t_{n+1}, \ldots, t_N)$ associated to the subsequent crossing points $(x_{n+1}, \ldots, x_N)$, by adding them the delay with which the next crossing point x_{n+1} is reached. Then, according to whether or not such delay is less than the previously computed temporal tolerance δT, we go to step 2 or step 3.

2 - We check if every target has behaved as predicted. Should this not occur, we go to step 3; otherwise the precomputed optimal policy remains unchanged as long as new updated information are received.

3 - In case, owing to some untolerable delay, the optimal times are shifted forward or a target has behaved differently from predictions, a dangerous interference may occur on the ship's subsequent transitions. In the first case, go to 4, otherwise store the updated kinematic parameters of the targets.

4 - If the discrepancy allows us to end the actual transition under safety conditions, we must compute a new optimal policy beginning from the crossing point x_{n+1}, otherwise we rely on the manual operator who, after overcoming the dangerous situation, will rely again upon the automatic operator.

Finally, we observe that this type of on-line implementation is operationally valid if recomputations of the optimal policy do not occur too often. For this purpose, the introduction of reachable sets associated to the targets and of a temporal tolerance associated to the optimal policy will prove particularly useful.

7. COMPUTER PROGRAMS FOR COLLISION AVOIDANCE

Following the dynamic programming approach,presented in the previous sections, FORTRAN computer programs have been prepared and tested on land on the CII 10070 computer of the "Centro di Calcolo" of Genoa University.

At sea, the programs will be tested by the experimental equipments installed on the M/s "Esquilino", during a four months' voyage that will begin about November 1972. The shipboard computer is a Selenia GP16, marine type.

After these tests, the programs will be translated into Assembler Language, as to adapt them to an IBM System Seven (MSP7), which is installed on the first Italian superautomated ship "Lloydiana". The characteristics of the IBM S/7 seem particularly fit for such on board applications. In fact, applications of dynamic programming to real procedures generally require a high speed working memory,and computation times often result too long. In our case, the duration of the on-line implementation program (see Fig. 6) and of the main program (see Fig. 4) must be shorter than the time interval between two data acquisitions from the radar system (about 2-3 minutes).

The CII 10070 System has a core requirement of about 5 k words and a processing time of about 1 minute for the main program, and about 0.20 minutes for the on-line implementation, when considering 40 targets and 22 stages.

In order to obtain these processing times, a Filtering Subroutine is employed by the main program which excludes the targets that do not interfere at all with the discretized ship's routes, from the subsequent computations.

8. NAVIGATION IN THE PRESENCE OF TROPICAL REVOLVING STORMS

The above proposed procedure for the prevention of collisions at sea may be applied to another navigation problem, where the risk is represented by tropical revolving storms. We shall deal in particular with the tropical storms of the South and Middle Indian Ocean, a brief description of which will be supplied in Section 8.1. The problem consists in determining a minimum time evasion strategy in the presence of tropical revolving storms. For this purpose, in Section 8.2, we shall briefly present the modifications that must be made in the previously proposed procedure, on account of the different evolution of these phenomena.

8.1 MODEL OF TROPICAL REVOLVING STORMS

Several attempts, made in the past, to handle the meteorological information concerning these storms from a statistical point of view, have led to scarce results. In particular, the results are largely negative, if one wants to determine cycles, trends and space correlations (see Jordan and Ho [5]).

For our practical application, we have restricted the problem to the South and Middle Indian Ocean tropical storms.Cyclones usually originate between 8° and 20° latitude South, after which, the storm travels to the eastward. The point of recurvature depends on the high pressure areas existing in the ocean. However, some cyclones do not recurve, some others are most irregular and occasionally a small complete loop may be included in a track.

A careful analysis by Kume [6] , which is based on the typhoon reconnaissance flights carried out, after the war, by U.S. A.F. and Navy, allows us to subdivide the cyclone trajectory into four idealized parts (see Fig.11).

Part A: Origin. It is caused by the formation of small low-pressure areas, which are difficult to be forecasted.
Part B: Westward motion. The cyclone moves prevailingly between westward and southwestward. The direction and the speed of the movement are quite stable. The speed value is normally between 8 and 12 knots.
Part C: Point of Recurvature. The point of recurvature corresponds to the position where the western side of a high pressure area is reached by the cyclone, and its westward movement changes to an eastward movement. Normally, it is easy to forecast the latitude of such a point, at least a few days before the cyclone reaches it, but on the contrary, it is rather difficult to forecast its longitude.
Part D: Eastward motion. Before the point of recurvature it is very difficult to forecast which direction the cyclone will take. During this part, the cyclone is generally accelerated and the greater the acceleration is, the greater the curvature becomes.

8.2 MINIMUM TIME STRATEGY IN THE PRESENCE OF CYCLONES

The determination of a minimum time evasion strategy in the presence of tropical revolving storms cannot be solved in terms of optimal weather routing, since these storms are rather anomalous phenomena. As in our case, safety is the prioritary exigence, we deem it preferable to deal this problem analogously to the collision avoidance problem.

Let us suppose that own ship must travel from an initial point A to the terminal point B, according to a reference course, and that she is periodically supplied with information concerning:

(a) - position and speed of the storm eye and diameter of the storm field;
(b) - position and values of oceanic high pressure zones;
(c) - weather forecasts.

In the absence of tropical revolving storms, a minimum time weather routing may be determined by means of a stochastic dynamic programming procedure (see Zoppoli [7]).

The meteorological conditions may be treated deriving their probability densities from radio weather messages (c) and from monthly statistical data (Pilot Charts). The result of the procedure is the minimum time routing, which is assumed to be the reference route.

In the presence of a cyclone, first of all we have to compute with which of its parts own ship may interact, assuming that the storm travels at the maximum speed on the shortest path between its actual position and own ship's future position. On the basis of the (a) information, we can know which stage of its life the cyclone has reached and its probable track. From these data the cyclone reachable set is computed and the evasion strategy is determined, adopting the same procedure as in the collision avoidance problem.

Let us briefly see step by step how the evasion strategy may be determined.

1. A grid of the possible crossing points is constructed, the geometrical characteristics of which depend on the distance between own ship and the cyclone. Since the fundamental purpose consists in evading the cyclone, the grid has not one terminal point, but a whole column of possible terminal points which are ranged on an orthodromic arc, perpendicular to the above defined reference route. The distance between two subsequent stages of the grid corresponds to the distance covered by the ship during the time interval (6 hours) between two subsequent updated information about the cyclone.

2. The cyclone reachable set is computed according to its life stage and to the (a) information. In Part B of the cyclone track, the motion is assumed to be rectilinear uniform. The indetermination of the recurvature longitude fixes the angular uncertainty of the set. In Part D, we increase the angular uncertainty and consider the North-Eastward track as the most probable (see Fig. 12).

3. The safety radius extends, according to the diameters of the storm field, as far as to obtain that own ship meets a wind strength not greater than Beaufort force 6 ÷ 7. The safety radius is measured from the forbidden storm field and takes the difference between the dangerous and the navigable semicircles into account. Therefore, the distance from the reachable set of the cyclone must not be considered symetrically. In our case, the dangerous semicircle lies on the left hand side of the cyclone motion vector. Thus, on the left of this area, we shall assume a safety radius greater than the one imposed on the right.

4. The cost associated to each transition is still time. It depends on the sea conditions with reference to the minimum time weather routing procedure. It is assumed that the ship's speed is the maximum speed allowed by the sea conditions.

5. As in the collision avoidance procedure, an off-line strategy and an on-line implementation, based on periodical information, are determined. The on-line program includes a tactical procedure also in order to minimize the danger when the storm field is very near. This tactical method is based on the above dangerous and navigable semicircles, and utilizes safety sectors as recommended by the practical rules.

The above programs are about to be tested on land by simulation, and they will be tested on board during the "Progetto Esquilino II" in the next year.

REFERENCES

Kaufmann A., Cruon R., 1967, "Dynamic Programming", Academic Press, New York.

1 - Warga J., 1971, "On a Class of Pursuit and Evasion Problems", Journal of Differential Equations, n° 9.

2 - Friedmann A., 1970, "Existence of Value and of Saddle Points for Differential Games of Pursuit and Evasion", Journal of Differential Equations, n° 7.

3 - Neustadt L.M., 1963, "The Existence of Optimal Controls in the Absence of Convexity Conditions", Journal of Mathematical Analysis and Applications, n° 7.

4 - [illegible]gino N., 1969, "On Sufficient Pursuit-Evasion Strategies", Proceedings of the First Symposium on Differential Games.

5 - Jordan C.L., Ho T.C., 1962, "Variations in the Annual Frequency of Tropical Cyclones 1886-1958", Monthly Weather Review, April.

6 - Kume T., 1954, "The Idealized Fundamental Model of the Typhoon Trajectories and the Latitude of Recurving", Proceedings of the Unesco Symposium on Typhoons, Tokyo.

7 - Zoppoli R., 1972, "Minimum Time Ship Routing as N-Stage Decision Process", in course of publication on Journal of Applied Meteorology.

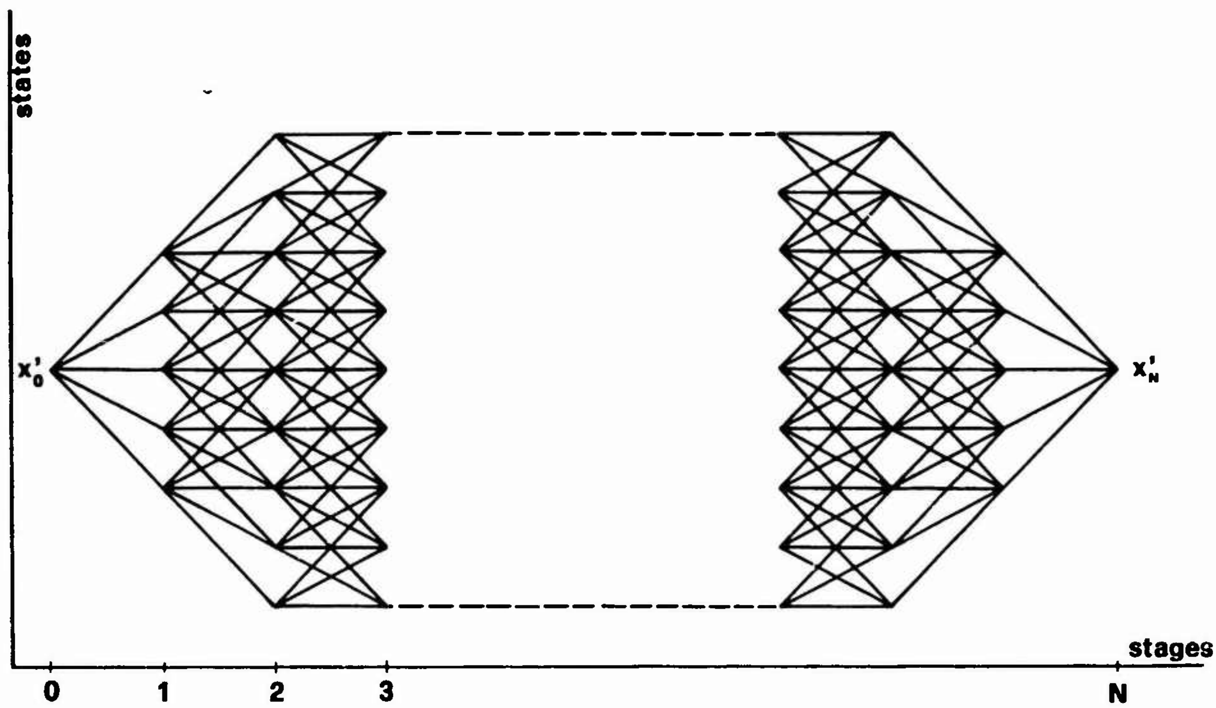


Fig.1 Grid of possible crossing points – the possible transitions are represented by full line

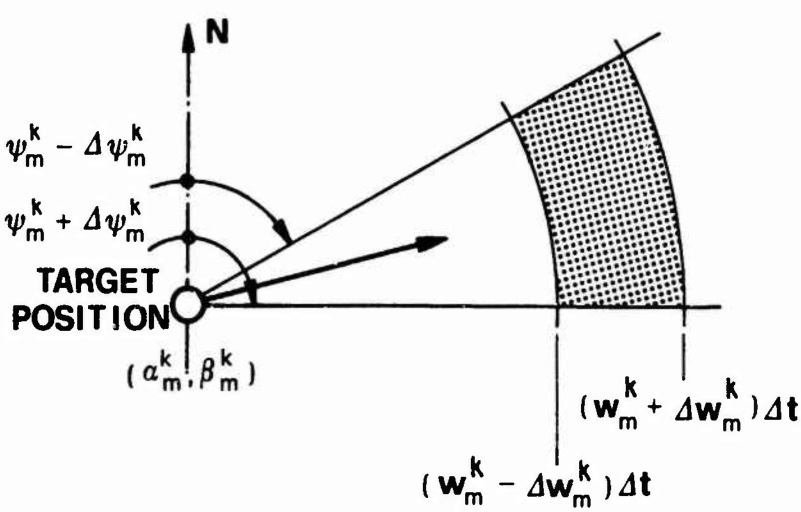


Fig.2 Reachable set

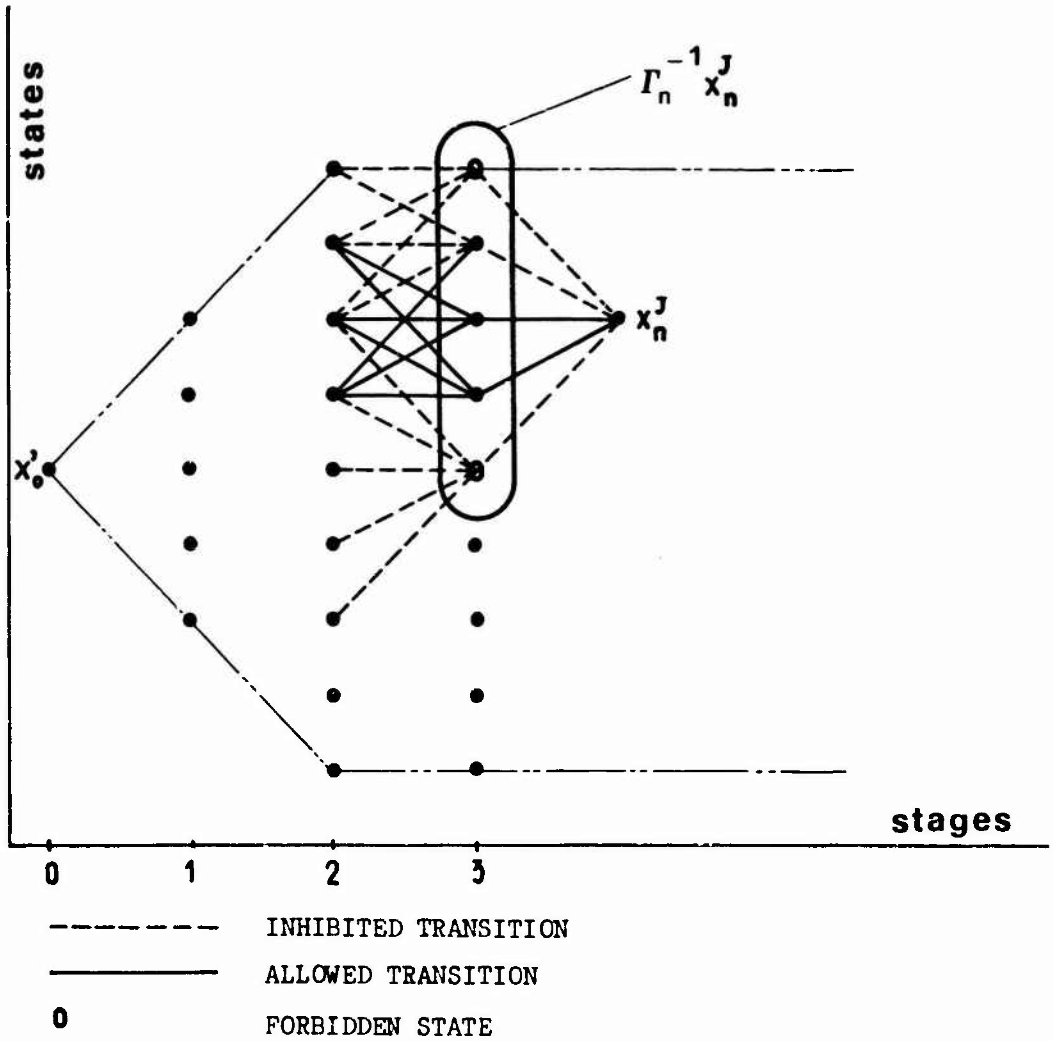


Fig.3 Allowed transition in the presence of risk

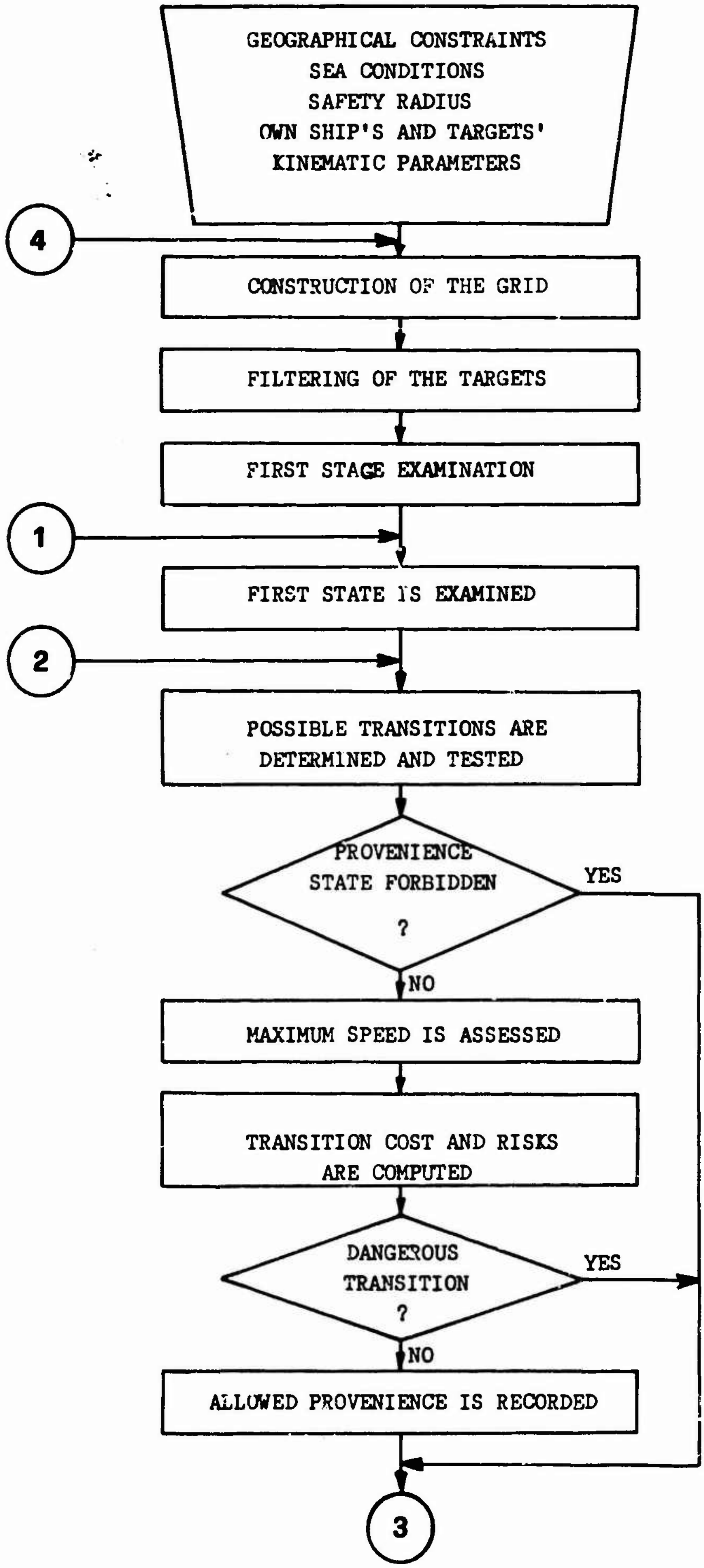


Fig.4(a) Main program flow chart

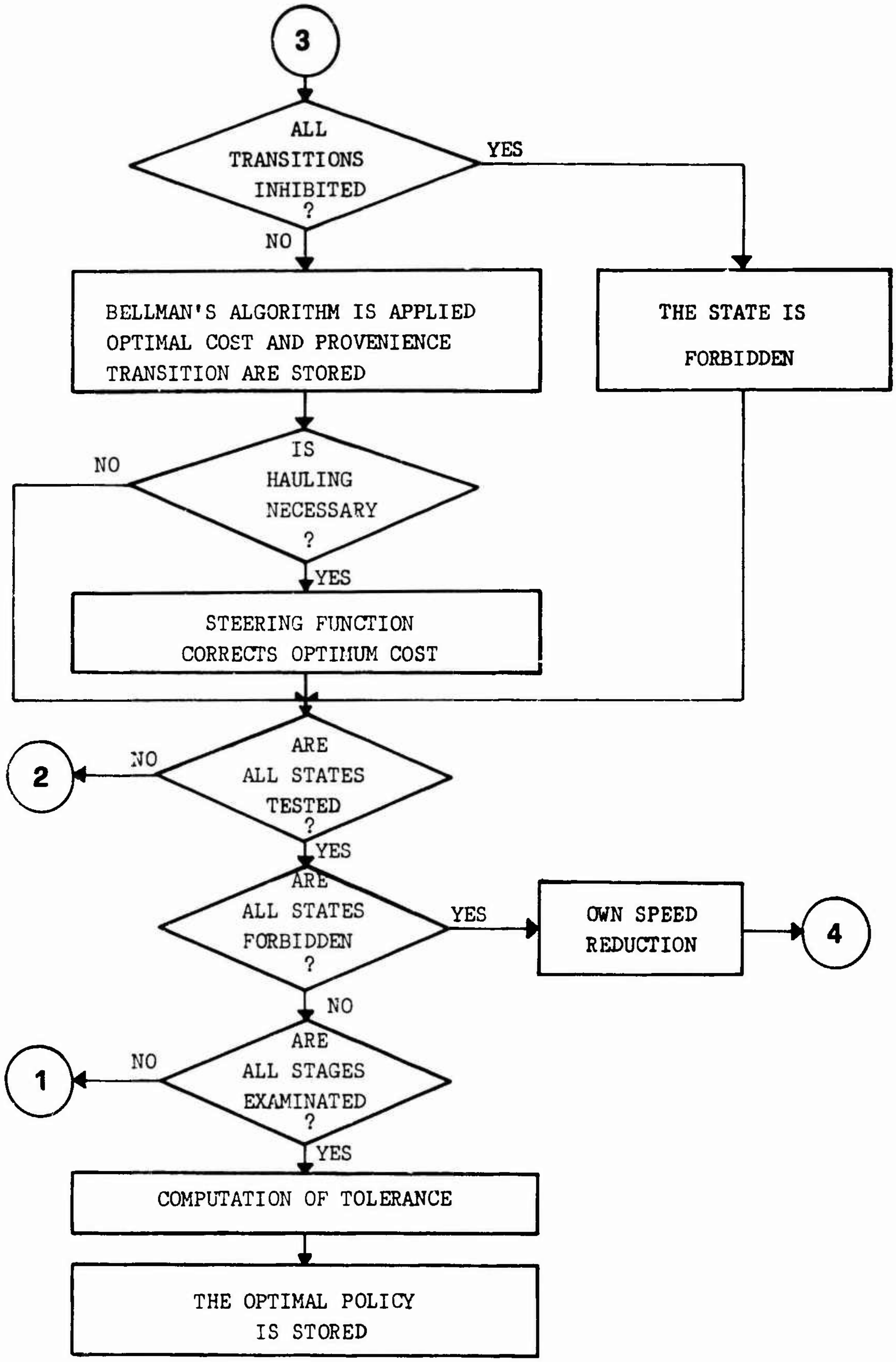


Fig.4(b) Main program flow chart

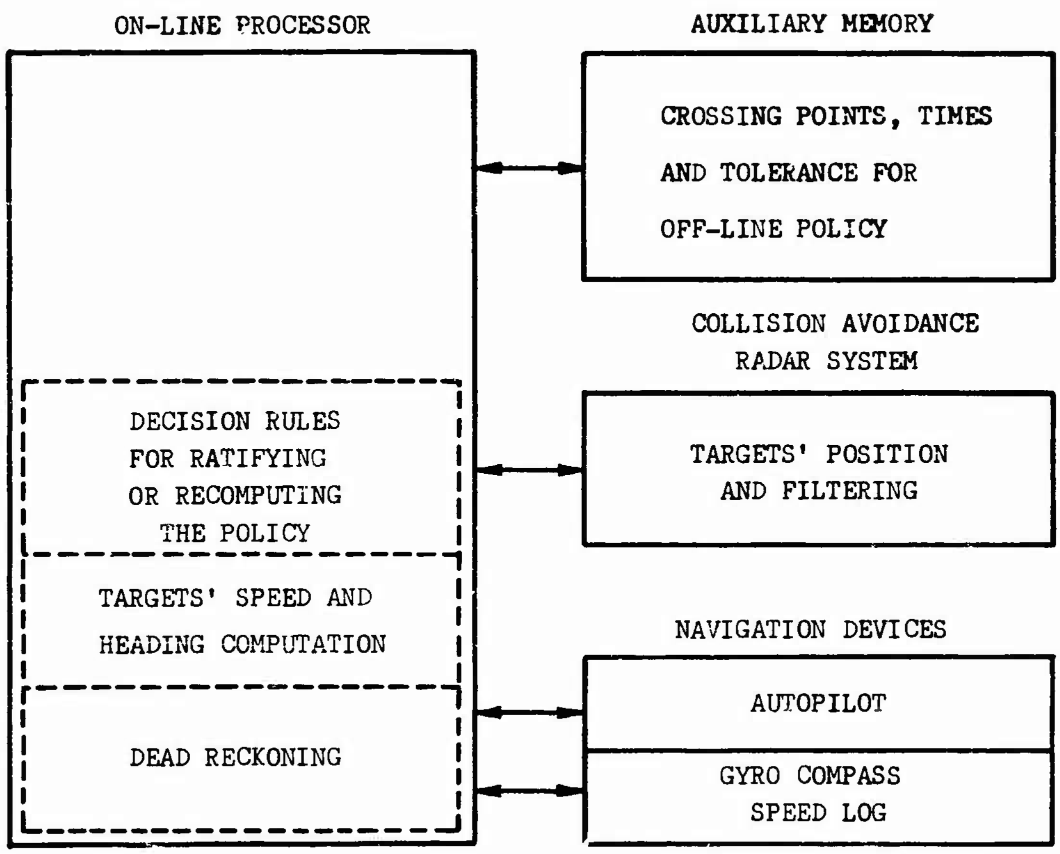


Fig.5 On-line implementation diagram

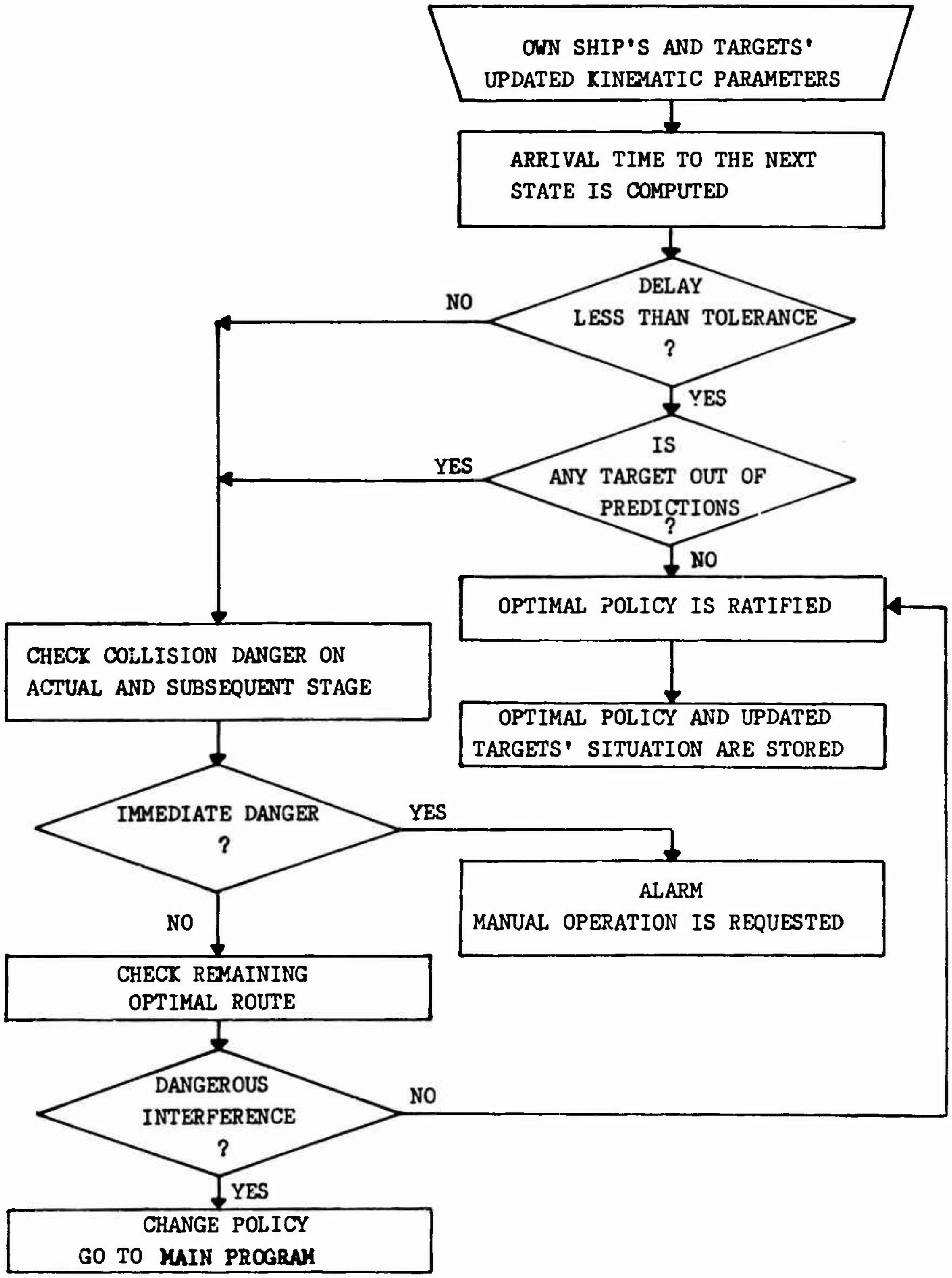


Fig.6 On-line implementation flow chart

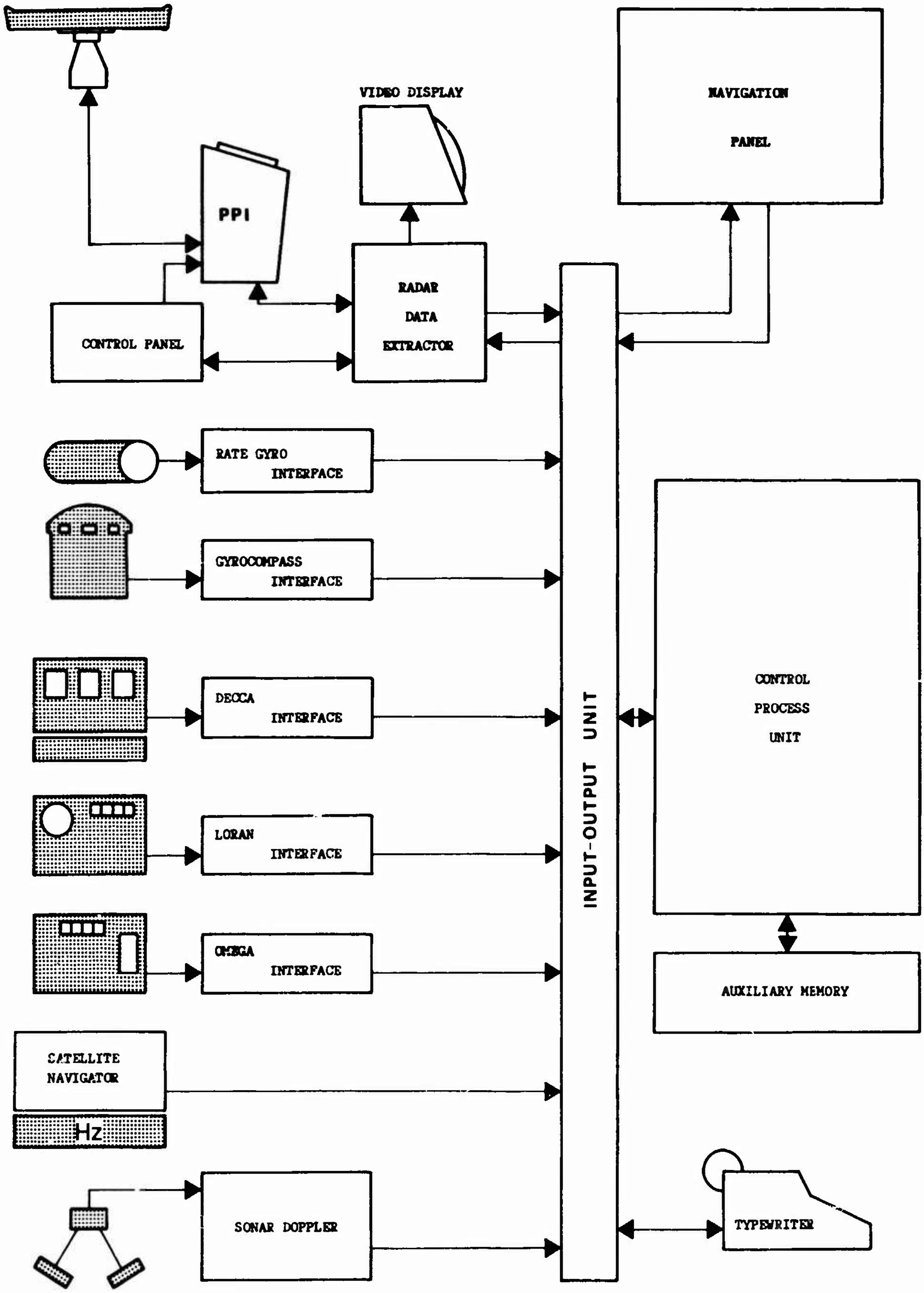


Fig.7 Navigation and anticollision integrated system

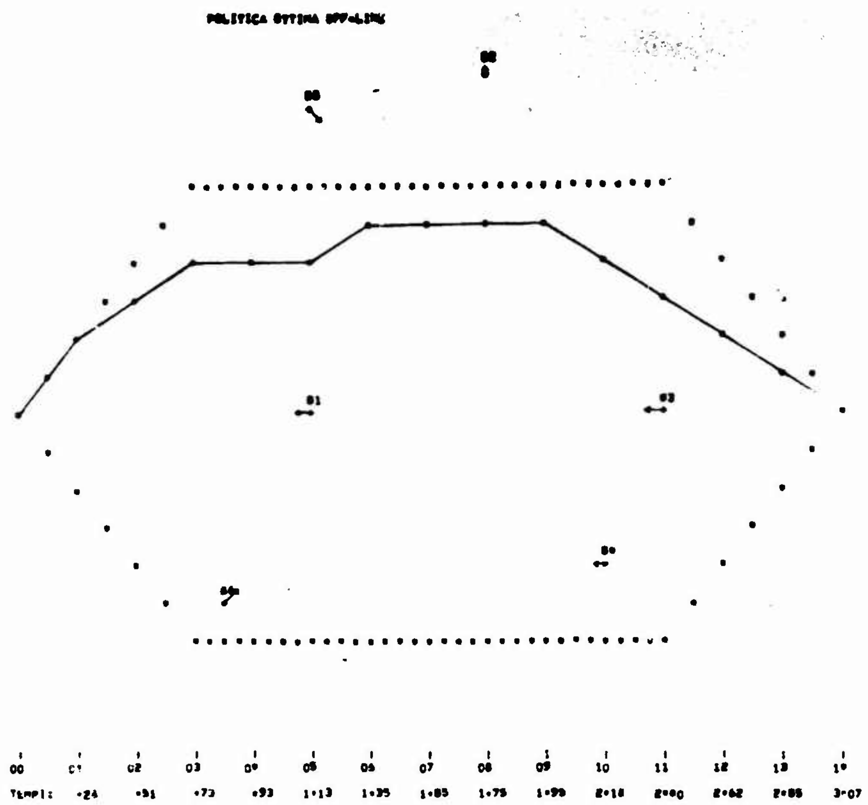


Fig.8 Example of an optimal off-line policy in the presence of 6 targets

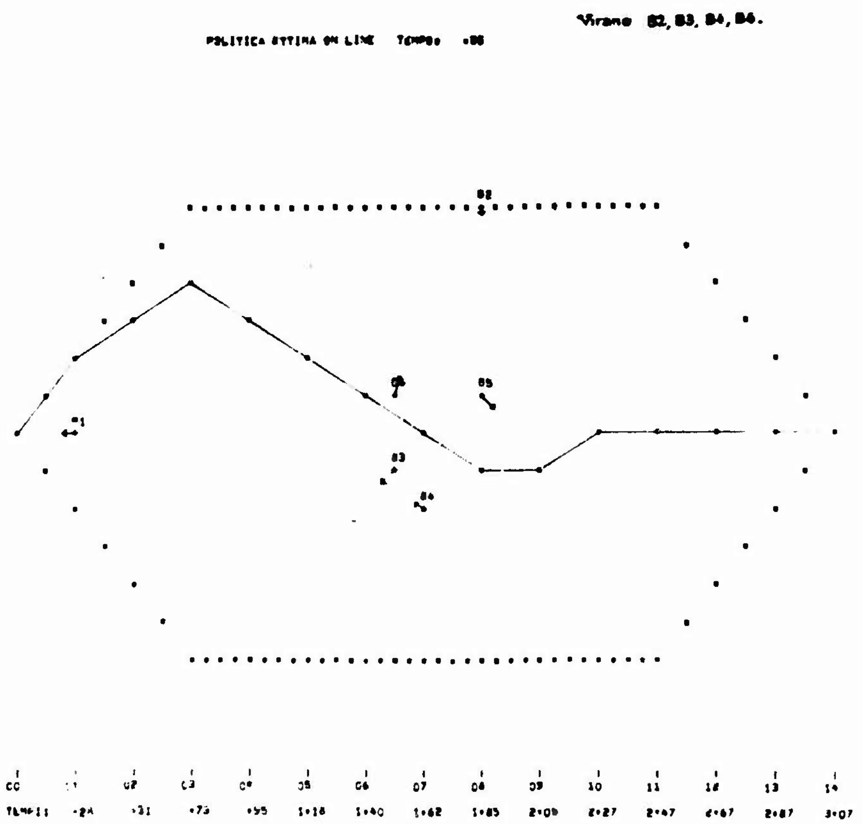


Fig.9 Referring to the example of Figure 8, targets B_3, B_4, B_6 modify their course, i.e., they do not move according to a rectilinear uniform motion, as assumed in Figure 8. A new optimal policy is computed in real time

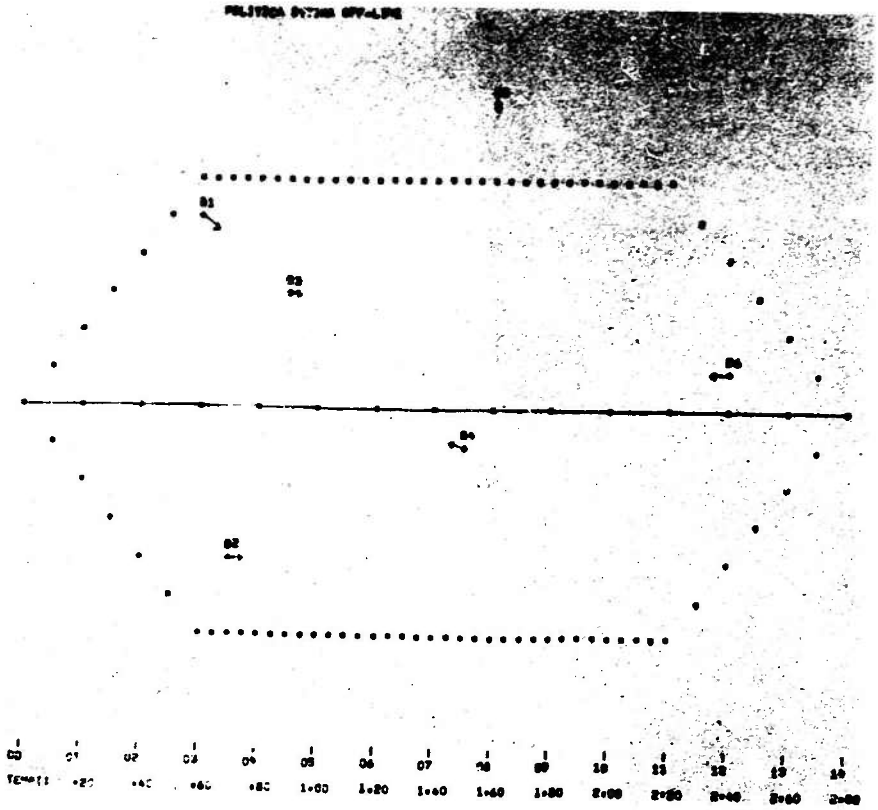


Fig.10 Another example. Who would have ever said so?

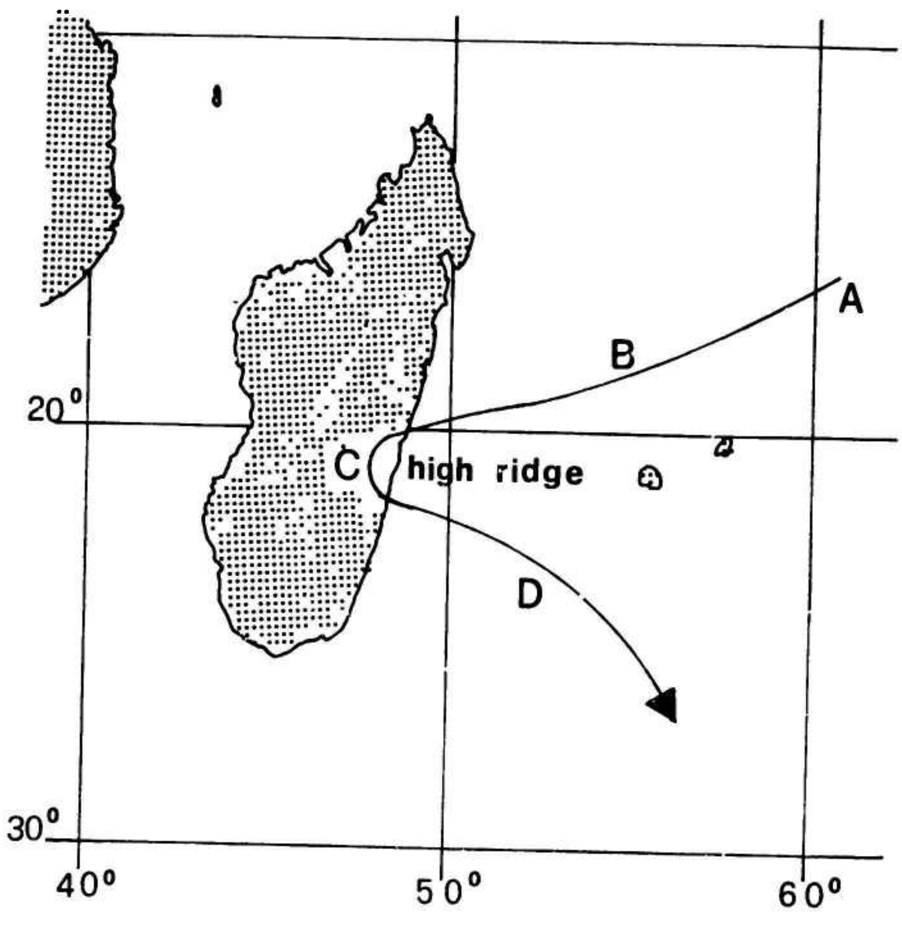


Fig.11 Typical track of a sub-tropical cyclone

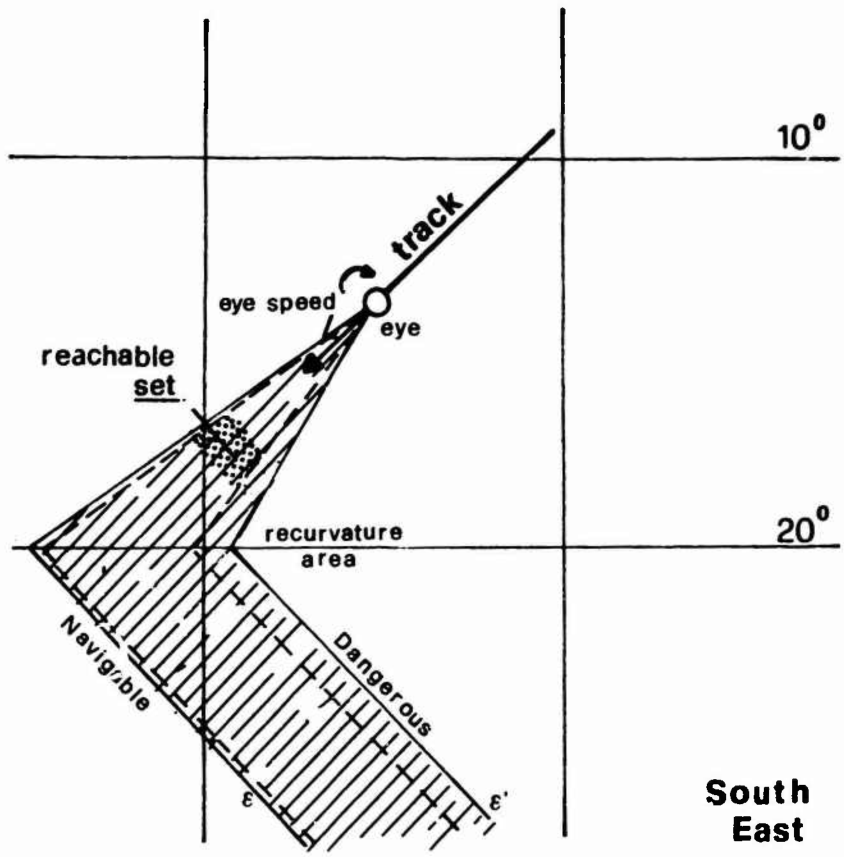


Fig.12 Cyclone reachable set

SPACE STATION INFORMATION SYSTEM REQUIREMENTS - A CASE HISTORY OF MAN-MACHINE SYSTEM DEFINITION

Author: C. R. Gerber
Project Engineering - Information Subsystem
Space Station Engineering
Space Division - North American Rockwell

ABSTRACT

The NASA Space Station (SS) Definition studies completed by Space Division, North American Rockwell (SD) incorporate a multiplicity of automated supporting functions to enhance the useful work capability of very few men. The SS Information System is the means by which the men interface with all subsystems, space experiments, other vehicles and ground support facilities and personnel. It is therefore a driver in determining what program and mission objectives can be satisfied.

The approach utilized by SD evolved a relatively simple procedure to rapidly identify the basic system requirements for a large number of space operations by classifying the operations. Each class was expanded to greater specifics until "drivers" could be identified. Various levels of equipment capabilities (data rates, number of channels, etc.) were postulated, and tentative "scenarios" developed to relate the crew actions and the equipment capabilities to the selected operation. The information was summarized in a standard format so that the potential users (crew and ground support representatives and discipline experts (Communications, Data Processing, Software, Displays and Controls, Instrumentation) from NASA and other contractors could critique, discuss and suggest alternatives in a common meeting. In this way, the expertise from those who would be most involved with the operating system was integrated to develop the Information System performance requirements.

The performance requirements were subjected to a sensitivity analysis by identifying the automation support functions provided by the on-board Information Subsystem to other subsystems, both onboard and external. Certain assumed interactions were found to generate unrealistic burdens (scale factors) upon the ISS, and adjustment of the performance "requirement" was necessary.

The study resulted in the definition of an Information Subsystem consisting of a unique combination of multi-processing computation, internal data distribution via a digital data bus, crew interfacing via a set of multi-purpose display and control consoles, and external data distribution via a combination of VHF, S, and K band RF links.

1.0 INTRODUCTION

The NASA Space Station was conceived as an orbiting platform to conduct manned operations for scientific, engineering and other applications uniquely served by that location. The crew of twelve was to be relieved of routine housekeeping actions for the supporting subsystem functions; (automation) to be independent of ground control in proceeding with their tasks (autonomy), and to be able to conduct a multiplicity of tasks concurrently. The vehicle was to be assembled in orbit using "modules" delivered by the Shuttle Orbiter. FIGURE 1 indicates the construction and allocation of the several modules.

Aside from the on-orbit assembling of separately-launched modules, the spacecraft systems (Life Support, Thermal, Stabilization, Power and Propulsion) are not uniquely different from those used in previous manned vehicles, with the exception of the Information Subsystem. The Information Subsystem includes all communications (internal and external), all data handling, and a large central computer by which the automation and autonomy characteristics are accomplished. In a manned vehicle this subsystem also has the displays and controls that are the work stations for the crew.

It is necessary to distinguish between the Information Subsystem on the spacecraft and the total Information Management System. The Information Management System is defined as a flexible combination of men and equipment, utilizing operations and software to acquire, process and distribute data which enables man to control and support the station, logistics vehicles, launch vehicles, ground stations and experiments. The Information Subsystem within the spacecraft in cooperation with other Information Subsystems (space and ground) and their respective crews, constitute the total Information Management System.

2.0 ANALYTICAL APPROACH

FIGURES 2 and 3 list the functions that are allocated to the on-board and ground Information Subsystems. The on-board functions cover a wide spectrum, and at this point are not defined in terms of hardware, software, crew actions — and this is what must be done to define the Information Subsystem.

This paper describes work performed by the Space Division, North American Rockwell, for the NASA Manned Spacecraft Center under Contract NAS9-9953 for the preliminary definition of a manned space station.

The usual approach to design definition requires operations analysis, function allocation, equipment assignment and performance definition. An operation is a task to be accomplished, such as "Rendezvous and Dock Logistics Vehicle"; each task is segmented into smaller tasks with greater details. Functions are allocated to the several design areas; the designers identify the type of equipment available, and evaluate the expected performance needed to provide the function. This is an iterative process, and may occur on several levels concurrently as well as consecutively. FIGURE 4 indicates these relationships and FIGURE 5 shows that the Information Subsystem Requirement (ISR) is the link between the operations/functional tradeoffs and the equipment performance/capability tradeoffs. The ISR is the unique document tool that was used to "short-cut" the usual system definition approach.

At the initiation of this program it was estimated that about 3000 operations would be identified (See FIGURE 6) To analyze so many operations by the usual approach would have required more resources (men, time and funds) than were available. Fortunately, past experience in space programs allows a method of reduction into three classes: a rather large group that have been performed adequately by existing systems; a smaller group that can be performed with some modifications to existing systems, and an even smaller group that are new to space operations. This last group, which requires the full treatment, can then be compared and combined with the second group for further analysis.

The procedure actually used is shown in FIGURE 7, the ISR Procedure Flow. The Space Station Information System requirements identified by operations analysis were classified in an index (ISR Tree) as in FIGURE 8. Then, by the above procedure those "critical" ISR's were identified (FIGURE 9). For each of these an ISR form was completed (FIGURE 10) to postulate that for one specific operation these specific equipment capabilities would be needed; in this manner the unmanageable total problem was reduced to bite-sized chunks.

However, even at this point there were insufficient resources to conduct equipment tradeoffs; in lieu of time-consuming analysis an approach was conceived to form a working group, limited in size, to evaluate and critique the postulated ISR. The working group was composed of experienced personnel from each of the functional areas which will eventually be involved in the operate of the space station: the Ground Operations (FOD, LCC), the Ground Support Operations (KSC, GSFC), the Technology Support (ISD, ESD), the Flight Operations (FCOD). Each ISR was presented, discussed, modified and eventually disposed of by constructing an entry in the Space Station System Specification.

3.0 DESIGN RESULTS

The several entries in the Space Station System Specification were totalized, categorized and sized by concurrent design analysis. FIGURE 11 is a listing of automatic processing functions, categorized by spacecraft subsystem, that were to be supported by the data processing assembly portions of the Space Station Information Subsystem. Each entry in this table was expanded on a detail sheet to further identify the functional requirement in terms of processing rates, algorithms, etc. The totalizing included all the automatic computation requirements.

In certain cases the speed of reaction and the sheer magnitude of data points forced a reassessment of the requirement, or a reassignment of that function to a different element. For instance, the monitoring, control and backup of some 2000 solid-state circuit breakers (SSCB) proved to greatly load the data bus and central computer; this function was divided into modular groups and reassigned to small mini-processors incorporated within the Electrical Power Subsystems, leaving only the master supervision processing within the central computer.

Similar impact analyses were made relating the communications and display and control assemblies. The resulting preliminary design of the Modular Space Station on-board Information Subsystem has four major assemblies as indicated in FIGURE 12. Some of the major performance parameters are listed in FIGURE 13. The Information Subsystem is housed in two identical control centers, each with a large central computer in a multi-processor, multi-programmed configuration (see FIGURE 14). Either control center is capable of running all spacecraft operations; one center is nominally designated as the Operations Control Center; the other is a Backup Operations Control Center and is nominally designated to support Experiment Operations. Each center includes one or more Operations Console work stations. The computers are linked to each other and to all subsystems and experiment sensors by a redundant, time-shared high speed (10 Mbps) data bus. At certain locations one or more remote processing units (mini-processors) perform local processing/control tasks, and report to and accept commands from the central processor via the data bus. At other locations several small consoles are provided for crew access to the data system for off-line tasks; for example, the station commander's console is provided in his quarters for planning, scheduling, etc; another is provided at the galley for food inventory records. A unique feature is the Portable Console Unit that is used for subsystem trouble-shooting, maintenance and repair operations. The control centers and crew are linked to other spacecraft and the ground by radio, primarily through a Tracking and Data Relay Satellite System (TDRS). The Data Processing Assembly (DPA) configuration is selected for incremental modular growth as well as redundancy needs, and consists of two large and several small computers to provide the total capability required.

THE NASA DMS WORKING GROUP

Permanent Members

W. E. Miller	Headquarters	
S. H. Nassiff	MSC	Flight Crew Operations
K. J. Bobko	MSC	Astronaut
L. M. Pringle	MSC	Data Systems
R. Kosinski	MSC	Communications
R. S. Sayers	MSC	Co-Chairman (Info Mgmt)
R. G. Rose	MSC	Chairman
T. D. Keeton	MSC	Software
C. A. Beers	MSC	Co-Chairman (Flt Operations)
H. Rosengerg	MSC	Communications
C. R. Gerber	NR-SD	

Occasional Participants

B. Beasley	MSFC
G. Hall	MSFC
P. F. Barritt	GSFC
R. D. Smith	KSC
A. A. Tischler	NR-SD
C. L. Gould	NR-SD
C. W. Roberts	NR-SD
Subcontractor Representatives	NR-SD

Other NASA Contractors

- Boeing
- Bendix
- TRW

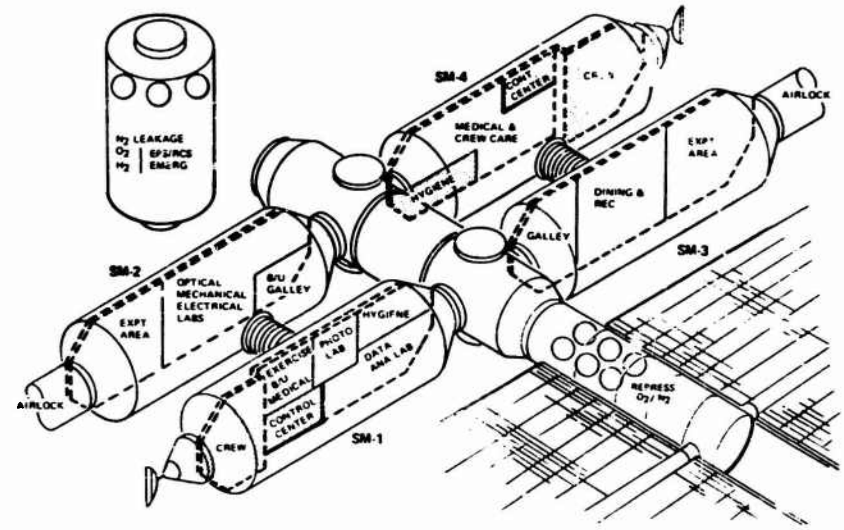


Fig. 1 MSS functional allocations

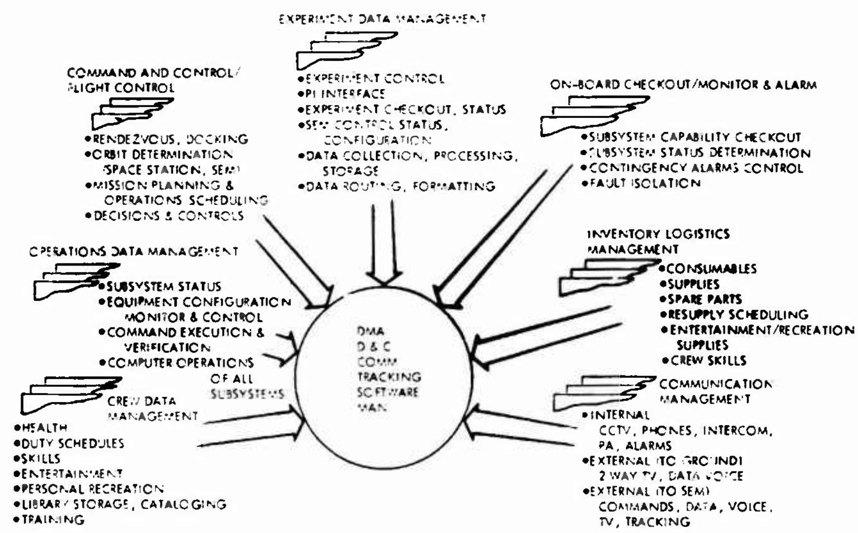


Fig.2 On-board information management subsystem functions

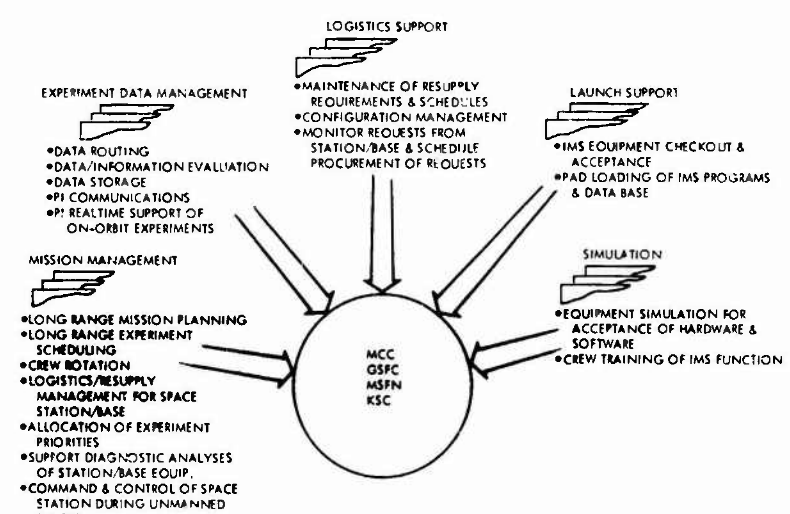


Fig.3 Ground IMS functions

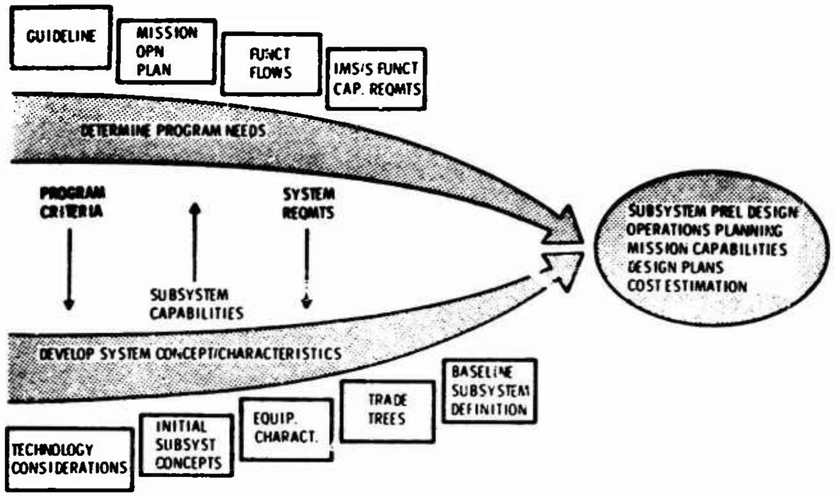


Fig.4 Approach to definition and subsystem selection

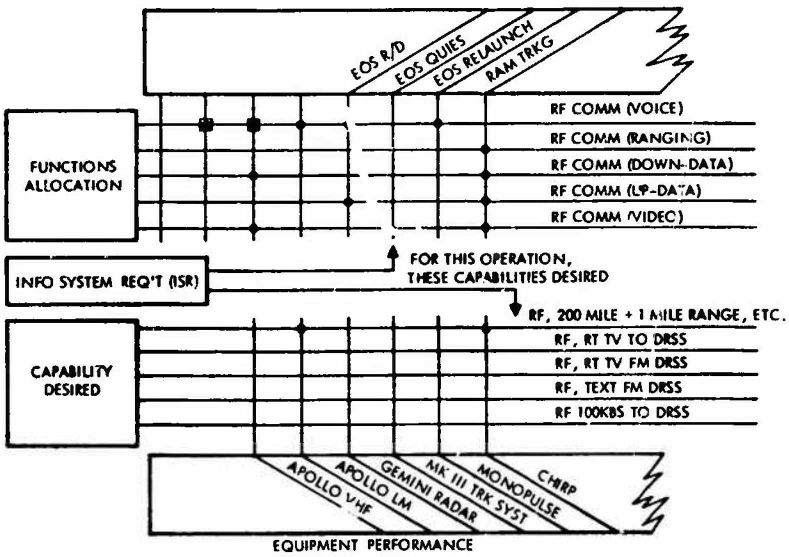


Fig.5 Mission operations

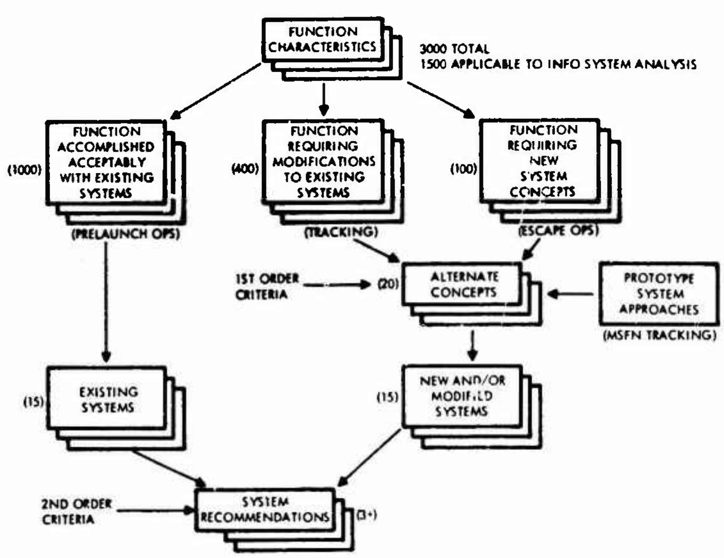


Fig.6 Information system study approach

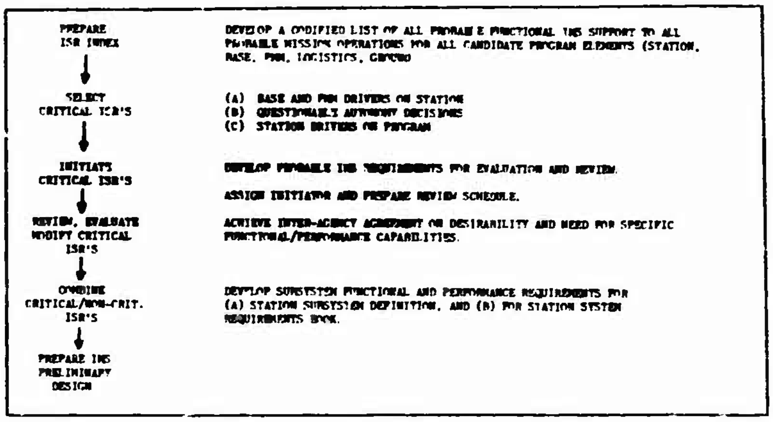


Fig.7 Information system requirement procedure flow

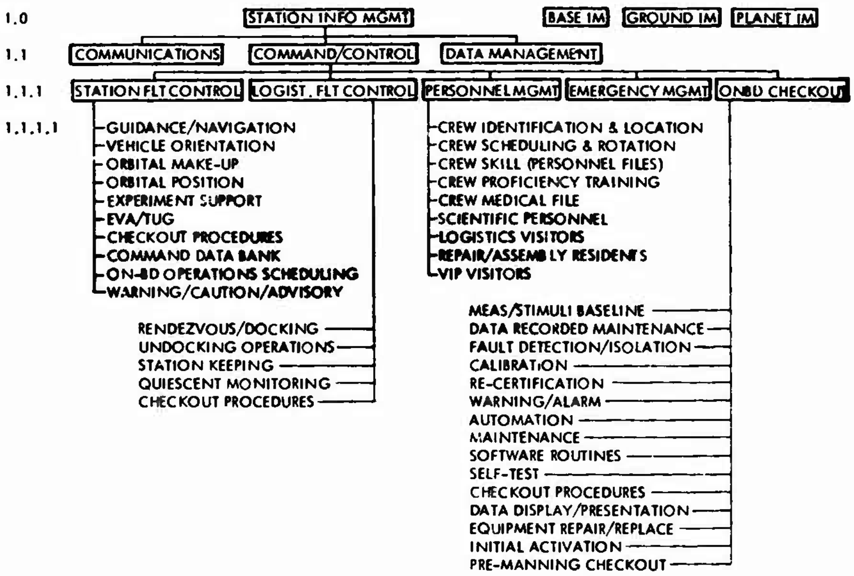


Fig.8 ISR tree (Index)

- PROGRAM DRIVERS (11 - THIRD LEVEL)

  HAVE MAJOR IMPACT UPON STATION/BASE PROGRAM

- CRITICAL (35 - THIRD/FOURTH LEVEL)

  CONTAIN SKELETAL/ESSENTIAL IMS FEATURES

- AUTONOMY (30 - THIRD LEVEL)

  PROVIDE BASIC DEFINITIVE CRITERIA ON FLIGHT VEHICLE DEPENDENCE/INDEPENDENCE

- AUTOMATION (8 - THIRD LEVEL)

  RELIEVE CREW OF MECHANICAL "HOUSEKEEPING" TASKS.

Fig.9 Critical ISR selection

PROGRAM ELEMENT :	Base	SEQUENCE NO.:	A-1
SUBSYSTEM :	Info. Mgmt.	DATE :	10 June 1969
FUNCTION :	Monitor & Alarm	ORIGINATOR :	C. Gerber
OPERATION :	All	ASSIGNED :	- - -
CONCURRENT ISR's:	- - -		
RESPONSE DESIRED:	Incorporate in Baseline Configuration	DATE DUE :	- - -

ABSTRACT:

All spacecraft subsystems status and configuration shall be monitored continuously to detect malfunctions of equipments. Depending on the nature of the malfunction, crew actions will be required within a certain time. Four categories are defined: Emergency, Warning, Caution, and Advisory.

REQUIREMENTS:

1. Emergency: operating subsystem function and/or performance level are such that generate conditions hazardous to the crew or equipment. Automatic adaptation to a "safe" subsystem configuration may be required.
2. Warning: operating subsystem performance deviates from pre-established tolerance limits, where prolonged out-of-tolerance operation would lead to failure of the function. Automatic "fail-safe" action may be required.
3. Caution: operating subsystem performance reserve, or expendable reserve may limit the degree of capability.

(cont. on p. 2)

REMARKS:

Enclosed is a copy of the AAP/OWS Caution and Warning Design Criteria for reference.

DISPOSITION:

Fig.10 Information system requirements

G&C	ECLSS	EPS	RCS	STRUCTURES	CREW	ISS
. Attitude Determination . Navigation Determination . Maneuver Determination . CMG Control •. RSC Control . Experiment Module Update . Shuttle Alignment . Terminal Rendezvous . Berthing	•. Pumpdown and Repressurization •. CO_2 Management •. Electrolysis Control •. O_2 Partial Pressure Cntl . Humidity & Contamination Control •. Circulation & Temp. Control •. O_2/N_2 Control •. Active Thermal Control . Humidity & Urine Recovery Control . Wash Water Recovery . Food Management •. Special Life Support	. S.A. Boom Control . S.A. Pointing Control . S.A. Inverter Control . Fuel Inverter Control . Primary Power Bus Control . Secondary Pwr. Bus Control •. Fuel Cell Cntl . SSCR Control . Differential Current Measurement . Lightning Cntl	. Nitrogen Qty Balance . Hydrogen Gas Control •. Thrust Valve Control . Oxygen Gas Control	. Berthing	. Real Time Medical Data Acquisition . Non-Real Time Medical Data Acquisition . Medical Data Analysis	. Internal Communications Cntl •. External Communications Cntl . Tracking Control •. Command & Message Generation •. Displays & Cntls •. Subsystem Ops. . Planning & Scheduling . Logistics Inventory Control . Information Storage & Retrieval . Mission Analysis & Assessment . Record Management . Remote Terminal Operate •. OPCO (M&A & F.I. Related to Critical Functions is Critical) •. CP Executive

• Critical Functions

Fig.11 DPA functional processing list

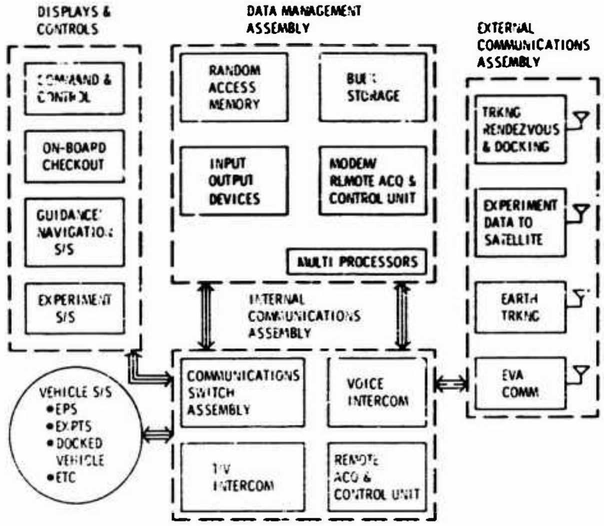


Fig.12 Space station IMS/S hardware concept

DATA PROCESSING ASSEMBLY

- MULTI-PROCESSING SYSTEM 5 PROCESSORS
- 360K WORDS RAM
- $3X10^{12}$ BULK STORAGE
- $2X10^{6}$ "EQUIVALENT ADD" OPERATION/SECOND
- 36 BIT WORD

COMMUNICATIONS

- PRIMARY - RF
  6.5 MHZ BASEBAND INCLUDES COLOR TV, TLM/COMMAND, 500KBPS DATA, 3 VOICE CHANNELS
- ALTERNATE-S-BAND OR K-BAND
- S-BAND & HF TO DEPLOYED MODULES, SHUTTLES, ETC.

TRACKING

- S-BAND - MONOPULSE
- RANGE RATE ACCURACY - 1 FT/SEC UP TO 400 FT/SEC RANGE RATE
- RANGE - 50 FEET TO 400 NAUTICAL MILES
- RANGE ACCURACY - 50 FEET

DISPLAYS AND CONTROLS

- CRT - 30 HZ DISPLAY RATE/3200 CHARACTERS
- HARD COPY
- MICROFILM - 16 MM/100 FRAMES/SEC CASSETTE TYPE 4000 FRAMES/CASSETTE
- PORTABLE DISPLAYS

Fig.13 IMS/S performance characteristics

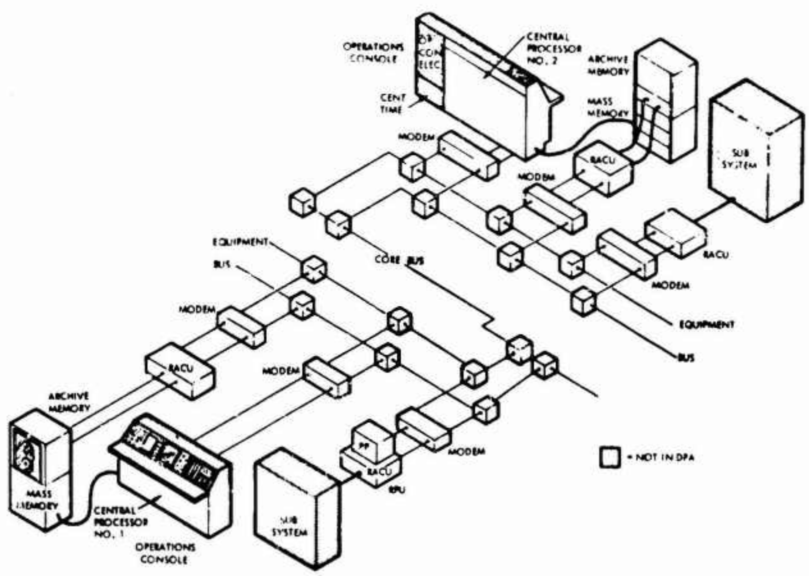


Fig.14 Data processing assembly (DPA)

AUTOMATED TECHNIQUES FOR SPACECRAFT MONITORING

H. Richard Segnar
International Business Machines Corporation
Federal Systems Division
1322 Space Park Drive
Houston, Texas 77058

ACKNOWLEDGEMENTS

The material presented in this paper was gathered during the Automated Techniques for Spacecraft Monitoring (ATSM) study conducted by IBM and NASA. The paper itself is primarily a condensation of the Final Technical Report of that study. The author gratefully acknowledges the efforts of those who participated in the study: without their help this paper would not have been possible. A special thanks to L. Wright, D. Prewit, and G. Evans of IBM.

SUMMARY

The Automated Techniques for Spacecraft Monitoring study conducted by IBM and NASA at the NASA Manned Spacecraft Center (MSC) in Houston, Texas addressed the practicality of efficient and reliable spacecraft monitoring through automation. The automated monitoring study involved implementing and evaluating a test bed program in the manned space flight ground support system. The test bed program automated selected flight control functions and demonstrated software techniques for automatically monitoring spacecraft systems, initiating malfunction diagnostic procedures, and aiding management of spacecraft consummables. Evaluation of this test bed provided insight into the feasibility and advantages of automated monitoring.

The feasibility of implementing automated spacecraft monitoring depends on four factors — sufficient computer resources, suitable monitoring function definitions, adequate spacecraft data, and effective and economical test systems. The advantages of automated monitoring lie in the decision-making speed of the computer and the continuous monitoring coverage provided by an automated monitoring program. Use of these advantages introduces a new concept of spacecraft monitoring in which system specialists, ground based or onboard, freed from routine and tedious monitoring, could devote their expertise to unprogrammed or contingency situations.

INTRODUCTION

The increased complexity of future spacecraft hardware and the necessity to monitor more spacecraft systems has created a need to automate as many flight control functions as possible. The study addressed the question of how to achieve efficient and reliable spacecraft monitoring through automation by conducting the Automated Techniques for Spacecraft Monitoring study. The objective of the study was to assess the practicality and economy of automating flight controller systems monitoring and evaluation functions as part of the ground support system and as part of onboard monitoring in future spacecraft.

To study automated monitoring first hand, a test bed program was implemented within the configuration of the Real Time Computer Complex (RTCC), which is the core of NASA's Mission Control Center at Houston, Texas. The test bed program automated selected flight controller functions and demonstrated the effectiveness of software techniques to automatically monitor spacecraft systems, initiate malfunction diagnostic procedures, and aid management of spacecraft consummables. Evaluation of the test bed's effectiveness provided data for formulating assessments of the feasibility and advantages of automated monitoring.

The remainder of this paper discusses the test bed program and the assessments formulated during the study. The discussion of the test bed program includes descriptions of the test bed environment, the Apollo spacecraft systems monitored by the test bed, and program techniques used. Assessments of automation are divided into assessments of automation in a ground support system and an assessment of automation onboard future spacecraft.

ATSM TEST BED PROGRAM

The ATSM test bed program was implemented within the configuration of the RTCC. Current RTCC programs accept data from both onboard and ground based systems, process the data, format the results into displays, and relay it to flight controllers sitting at monitoring consoles. The information is provided real time — almost instantaneously — so the flight controllers can determine the status of onboard systems, the condition of the astronauts, the position and velocity of the spacecraft at any desired time, and the effect of planned maneuvers.

The test bed program was implemented as an extension of the current Telemetry subsystem supporting Apollo spacecraft. The Telemetry subsystem provides information about the various vehicle systems required throughout a spaceflight. As data is received, it goes through several steps of processing including converting the transmitted data to engineering units (volts, pounds per square inch, etc.), computing additional values, and formatting the data for display devices. Typical display devices include digital television displays (information formatted on cathode ray tubes), chart recorders (data plotted against time on paper charts), and event lights (lights indicating occurrences such as "cabin temperature exceeds 80° Fahrenheit"). Developing the test bed program as part of the Telemetry subsystem permitted utilizing much of this existing processing logic.

Output of the test bed program consisted of digital television displays. Displayed quantities included specific parameter values, prose messages, and binary status indicators (e. g., GO or NO GO). The displays were formatted to facilitate rapid determination of the onboard system status and quick detection of system malfunctions.

The test bed processing can be divided into two categories: (1) the automated flight control monitoring functions and (2) the subsystem control processing. The four automated monitoring functions are: Control Systems Evaluation, SPS/G&N Preburn Checklist Monitor, SMRCS Leak Detection and Isolation, and Water Management.

Each automated function addresses a unique area of monitoring responsibility. Hence, each function operates independently except for Control Processor interfaces. The Control Processor performs services that are common to more than one function.

The procedures used to evaluate the test bed programs, the onboard systems monitored by the test bed, the test bed programs themselves, and the evaluation of each program's effectiveness are described in this section.

Test Bed Evaluation Procedures

To assess the effectiveness of the ATSM test bed programs, the programs were demonstrated in the RTCC under simulated mission conditions. The programs were executed on an IBM System/360 Model 75 computer configured for Apollo 14 mission support. The operation of the test bed programs was observed in the Mission Operation Control Room (MOCR) and the Staff Support Rooms (SSRs) at the Mission Control Center.

Test bed demonstrations were attended by NASA flight controllers, flight support personnel, and programming personnel. Following each demonstration, the observers evaluated the effectiveness of each program, identified desirable program modifications, and assessed automated monitoring techniques. The results of these discussions were recorded for subsequent use in formulating the study's conclusions.

ATSM Control Processor

The ATSM Control Processor manages system initialization, digital television display requests, and ATSM subsystem data control. System initialization consists of allocating sufficient core storage to buffer the telemetry input data and cueing the Apollo Mission System control processor to initialize the Apollo mission support programs. Digital television display requests for ATSM displays cause the Control Processor to cue the appropriate monitoring function to format and output data to the designated TV channel. ATSM subsystem data control consists of interrogating incoming telemetry data for changes in the telemetered parameters. When changes are detected, the Control Processor cues the monitoring programs and passes pointers to the latest telemetry data.

The concept of a single Control Processor proved an effective way to eliminate duplicated logic and simplify system design. Centralizing the display request and data control logic eliminated the duplication of code that would exist if each monitoring program performed these functions independently. Furthermore, the Control Processor performed all the ATSM subsystem interfaces with other mission support programs thus simplifying system design and implementation.

Control System Evaluation

Guidance and control functions are performed on the Apollo spacecraft by the Guidance, Navigation, and Control System (GNCS) and the backup Stabilization and Control System (SCS). The GNCS consists of three subsystems — the inertial subsystem, the computer subsystem, and optical subsystem. The Control Systems Evaluation test bed program automates monitoring some of the functions of the inertial subsystem (ISS) and the computer subsystem.

The inertial subsystem is composed of an inertial measurement unit (IMU), a power and servo assembly, and three coupling data unit (CDU) channels. The IMU provides an inertial reference with a gimbaled, three-degree-of-freedom, gyro-stablized platform. The alignment of the platform is commanded by the CMC which sends digital commands to the CDU. The CDU converts the commands to analog signals which drive the IMU to the desired orientation.

Any change in spacecraft attitude is sensed by comparing the spacecraft attitude with the alignment of the inertial referenced platform. Resolvers mounted on the gimbal axes of the platform provide signals which represent the gimbal angles. The CDU converts these analog signals to digital pulses for the CMC. The CMC compares these angles with the CMC desired angles, and if the angles differ, error signals are generated. The error signals can be used to generate commands for spacecraft guidance systems.

Mounted on the stable platform are the pulsed integrating pendulous accelerometers (PIPAs), which sense changes in spacecraft velocity. Any acceleration or deceleration results in output signals which are representative of the magnitude and direction of the velocity change. The output signals are transmitted to the CMC, which uses the information to update spacecraft velocity data.

The Command Module Computer (CMC) performs guidance functions by executing internal programs using predetermined trajectory parameters, attitude angles from the inertial channels of the CDU, velocity changes from the PIPAs, and commands from the crew to generate control commands. The navigation function is performed by using stored star-landmark or star-horizon data, optics angles from the optics channels of the CDU, and velocity changes from the PIPAs in the execution of navigation programs.

The Control Systems Evaluation (CSE) program monitors and reports the status of the IMU and automates some of the flight controller procedures for evaluating and controlling operation of the IMU. The status of the IMU's hardware is determined from CSM and CMC data. The IMU's performance is evaluated by comparing CMC and SCS attitude and rate measurements. The IMU components monitored by the program include the three-degree-of-freedom, gyro-stablized platform, the CDUs, and the PIPAs.

The program first determines the basic operational mode and submodes of each of the three IMU components. Once the component mode has been determined, the program verifies the mode by checking parameters associated with the mode. If the verification parameters do not reflect the expected value, a message is displayed explaining the verification failure.

The program also determines the operational status of the three components and indicates the IMU's ability to support the three spacecraft control functions — attitude reference, attitude control, and thrust vector control. The status of these six items is represented on the display by status words — GREEN, AMBER, or RED — which indicate satisfactory, uncertain, or critical conditions.

Finally, the program evaluates a series of binary expressions to determine if any procedure should be initiated to modify the operational status of the IMU or to improve its performance. The procedures to be followed are indicated by messages on the program's D/TV display.

The Control Systems Evaluation program performs IMU monitoring functions very effectively. Each processing cycle the program evaluates the status of the IMU based on 113 CMC and CSM parameters. Little or no manual intervention is required to operate the program. The program's output — action and warning messages for the flight controllers — represents the highest level of ground support monitoring feasible.

NASA flight controllers expressed a great deal of confidence in the CSE program's output. There are two primary reasons for this high level of confidence. First, the CSE program devotes 15 percent of its executable code to validating the input data. Input parameters are checked against predefined limits based on the physical capabilities of the onboard system. Parameters which fail this off-scale check are not used by the program. Parity checking is performed on other key CMC parameters before they are used by the program. Use of some parameters is delayed for several sample periods until a parameter's value can be verified by succeeding data samples.

The second factor contributing to the high level of confidence is the use of mode verification logic. After the program has determined the mode of a given IMU component, it attempts to verify the mode by evaluating the status of other IMU components as well as data from the SCS, which is the backup guidance, navigation, and control system. If the mode is verified, the word "VERIFIED" is displayed; if unverified, "UNVERIFIED" is displayed along with the name of the parameter which failed the verification check.

The action and warning messages and the IMU state messages result from evaluating a series of binary expressions. The binary expressions proved to be a very effective technique to record the flight controller's decision making processes. The binary expressions, however, were very difficult to test because each combination of binary elements must be verified. A binary expression which contains four elements requires 4^2 or 16 sets of test data. Thus, using binary expressions does simplify program definition, but it does not lend itself to rapid program checkout.

SPS/G&N Preburn Checklist Monitor

The Service Propulsion System (SPS) is a non-throttleable, pressure fed rocket engine mounted on the service module of the Apollo spacecraft. Thrust is applied through the spacecraft center of gravity by orienting the gimbaled engine mount. As consumables are depleted during a maneuver, the engine nozzle is gimbaled, readjusting the thrust vector to account for center of gravity changes. Commands for this thrust vector control (TVC) can be initiated by the onboard guidance systems or the astronauts.

The guidance and control functions onboard the Apollo spacecraft are performed by the GNCS and the SCS. Rotational and translational attitude and rate sensors provide data which is integrated and conditioned into control commands to the propulsion systems. Guidance and control functions may be performed automatically, primarily through commands from the CMC, or manually by the crew.

Spacecraft attitude control is provided by the Service Module Reaction Control System (SMRCS). The SMRCS is normally used immediately prior to SPS burns to perform an ullage manuever, which is a small X-axis translation. This action accomplishes fuel settling before SPS ignition.

The SPS/G&N Preburn Checklist Monitor program monitors guidance and navigation (G&N) aspects of SPS maneuvers controlled by the CMC. Preburn checklist parameters are tested against the desired configuration and various pitch and yaw gimbal angle measurements are tested against critical lower and upper limits. This checklist monitoring is segmented into three phases based on an ignition timeline. Each phase represents somewhat different processing requirements and monitors up to 62 CSM and CMC measurements.

The second program logic section collectively inspects the status of analog and discrete parameters to detect and identify system failures. This system monitoring involves evaluating "failure equations" (binary expressions) during thrust off, plus-X-axis translation, and thrust on activities. The detectable system failures include hardware, instrumentation, or instrument faults.

The SPS/G&N program monitors a maneuver when requested if the timeline constraints are satisfied. Program status and monitor results are indicated on two digital television displays. On the primary display, program status, checklist processing definitions, and checklist configuration status are indicated. On the secondary display, time-tagged messages identifying any system failures are driven. All parameters are validated so that a violation condition must be noted on four consecutive program executions (normally 4 seconds) before it is indicated. All violations indicated on the displays are retained for historical reference until removed by phase changes or manual request.

Implementing the SPS checklist function provided insight into automated checklist monitoring. The

SPS program systematically monitors astronaut actions and critical maneuver parameters, freeing the flight controller of this responsibility. In addition, the program readily analyzes limit violations and configuration anomalies and detects and identifies system failures.

It should be noted, however, that the ATCM SPS program does not completely automate verification of the preburn checklist. Many of the discretes which the CMC monitors onboard are not retained and downlinked. The program also does not monitor the downlinked readouts of the astronauts' display select keyboard (DSKY) commands to the CMC. Thus, the program monitors only a subset of the onboard procedures, although the switches and analogs that are monitored are especially critical parameters.

Flight control personnel were not able to fully evaluate the effectiveness of the system failure identification logic. The program successfully demonstrates the ability to perform system malfunction detection and analysis. However, two weaknesses are suspected in the failure equation definitions. First, some equations may be too simplified. By not considering all pertinent factors, the equations sometimes become erroneously conclusive. Second, since precise vehicle responses to certain situations are not always known, the equations are defined to detect the suspected response.

Sufficient analysis of the failure equations was restricted by two circumstances. First, realistic fault data is not readily available. Second, a significant period of "fine-tuning" would be necessary to identify the exact failure equations limit values.

Several modifications have been identified which would strengthen the overall effectiveness of the SPS automated checklist monitoring program. These modifications consist primarily of implementing more useful display formats, simplifying user interfaces, and reducing required program storage.

The use of event lights which readily indicate program and system status should be considered. The lights would indicate whether the programs were enabled and whether checklist anomalies or system failures had been detected.

Flight controller analysis and interpretation time could be reduced by implementing a monitoring display which readily identifies checklist anomalies and system failures. Individual columns would permit indication of current analog violations, current event configuration violations, and historic violations. Once a checklist violation no longer exists, the indication would be moved to its respective historic column. System failures would be time-tagged and similarly logged on another portion of the display. Implementating a reset (clear) PBI would then be useful for cueing removal of the historic checklist and system failure messages. Figure 7-1 shows a suitable monitor display format.

If the new monitoring display were implemented, a feedback display to indicate the checklist processing definitions should be considered. For each analog, the display would indicate the limit delta and in which phases the parameter is enabled for testing. If the delta is centered about the gimbal trim angles rather than zero, an additional indicator would be provided. For each configuration discrete the display would indicate the test value for each enabled phase.

Parameter characteristics should be restricted to the test condition and whether enabled or disabled for testing during each checklist phase. Also, the capability to override known faulty input data should be provided. Finally, manual entries can be simplified by (1) eliminating those quantities used only for display, (2) basing ullage information on a single, digital autopilot modes entry, and (3) defining limits with a single, delta value.

SMRCS Leak Detection and Isolation

The SMRCS consists of four functionally identical, physically independent systems located 90 degrees apart around the forward portion of the service module periphery. Each system, called a "quad," includes four rocket engines mounted on the outer surface of the service module and a propellant storage and delivery system inside. The propellant delivery system uses pressurized helium gas to transfer fuel and oxidizer to the engine combustion chambers.

The propellant storage and delivery system consists of a single helium storage tank, two fuel storage tanks, two oxidizer storage tanks, plumbing lines, and flow control valves. (Refer to Figure 7-2.) The helium storage tank contains pressurized helium gas that can be isolated from the remainder of the system by two independently operated helium isolation valves. High pressure helium is reduced to the desired working level as it flows through pressure regulators downstream of the helium isolation valves. Regulated helium pressure is directed to the fuel and oxidizer tanks through check valves to prevent reverse flow of propellant vapors or liquid. Regulated helium entering the propellant tanks exerts pressure on the propellant storage bladders forcing fuel and oxidizer through the propellant distribution lines to the engine. The propellant flow can be stopped by closing the propellant isolation valves between the propellant storage tanks and the engines. Propellant flow from the secondary fuel tank can also be stopped by closing the secondary fuel pressurization valve above the secondary fuel storage tank.

The state of each quad can be determined from transducer readings. The pressure, temperature, and pressure-to-temperature ratio of the pressurized helium gas are measured by transducers in the helium tank. The regulated helium pressure is measured by the helium manifold pressure transducer. Fuel and oxidizer pressures are measured respectively by manifold pressure transducers.

The SMRCS Leak Detection and Isolation program automatically monitors the four SMRCS quads for leaks. The program detects and confirms leaks, aids in locating leaks, and aids in determining the status of quads based on the monitoring of 42 spacecraft measurements. The program compensates for faulty data or failed transducers by validating all input values. Manual controls allow the flight controller to alter the logic flow whenever necessary.

The quad monitoring routine compares each pressure measurement in each quad to lower limits to check for pressure drops. When a drop in pressure is detected, the program checks to determine if the drop can be accounted for by use of the quad (jet firings). If a pressure drop can be accounted for by quad use, the program resets the lower limit and exits. If a pressure drop cannot be accounted for by quad use, the program indicates a leak and cues the leak isolation routine.

The leak isolation routine searches for the leak by requesting that valves within the quad be opened and closed in a prescribed sequence. As valves are closed, each isolated area is monitored for further drops in pressure. When the leak is located, the program displays the location of the leak and recommends the best valve configuration to conserve the leaking material. The program also computes and displays the actual leak rate and time-to-depletion.

Since the Apollo spacecraft can be operated only by the crew and there is no telemetry indication of SMRCS valve positions, the leak search routine compensates for this condition by interfacing with a ground controller. When a crew action is required, the program displays the desired action and suspends leak search processing. The ground controller then makes the desired request via voice loop, waits for crew confirmation, and enters a PROCEED command. The leak search routine assumes the requested action has been performed and resumes processing.

The overall effectiveness of the SMRCS Leak Detection and Isolation program was assessed by evaluating the significant features for degree of automation, system impacts, and man/machine communications. Significant program features include leak detection, thruster activity determination, leak location, and time and quantity predictions.

The leak detection capabilities of the SMRCS program are effective and useful. The program detects leaks of any magnitude rapidly and accurately. Leak detection is automated to a high degree requiring a minimum of human intervention.

Location of leaks is fast and accurate. However, the lack of onboard sensors increases the man/machine communications required, thereby reducing the overall level of automation. Interfacing through the flight controller to manipulate onboard valves does permit effective leak isolating and requires little flight controller time.

Suggested improvements are primarily concerned with thruster activity determination logic, transducer shift detection, status indicator output, and action recommendations. The improvements would reduce the possibility of erroneous results, increase the degree of automation of the program, and improve the usefulness of the program.

The degree of automation of SMRCS monitoring could be increased by using a status indicator for each quad. The indicator would illuminate when the program detects a leak, cueing the flight controller to monitor leak isolation processing on the SMRCS MONITOR display. The "off" condition would signify that no problems existed in the quad. Thus, continuous monitoring of the display would not be necessary.

The program could detect more problems if it monitored all pressures in the quad concurrently. Leaks currently are detected by initially monitoring only the helium tank pressure, since all leaks will eventually cause a drop in helium tank pressure. Monitoring the manifold pressures, for example, would enable the program to detect transducer shifts and very slow manifold leaks. This logic would allow more complete system analysis during contingency situations.

The possibility of erroneous leak indications could be reduced by monitoring the commands from the rotation hand controllers or the translation hand controllers to detect manually initiated SMRCS maneuvers. The program could suspend leak detection during the maneuver and reset lower limits following the maneuver. This logic would compensate for pressure drops resulting from jet firings during the maneuver.

Action recommendations following location of a leak should incorporate the status of all four quads instead of just the leaking quad. Recommendations as to how to configure the entire SMRCS (all four quads) would increase the degree of automation and would be more useful to the flight controllers and the crew.

Water Management

The water system consists of a potable subsystem and a waste subsystem interconnected to permit proper distribution of water. The system includes a 36-pound capacity potable tank, a 56-pound capacity waste tank, distribution lines, flow control valves, and an overboard dump facility (refer to Figure 7-3). Valve configurations are controlled by onboard switches.

Water from the fuel cells is accumulated in the potable tank while the tank is less than full and the potable tank inlet valve is open. If the potable tank is full or the potable inlet valve is closed, the water is diverted to the waste tank. Water for crew consumption flows out of the potable tank to the food preparation unit and the drinking water unit.

Waste water, which is the excess potable water from the fuel cells and water collected by the cyclic accumulators (cabin dehumidifiers), accumulates in the waste tank. This water supplies the glycol evaporators during periods of evaporative cooling. During most mission phases the evaporators are not used and excess water must be dumped overboard. The dump valve is opened by crew command or automatically when the waste tank is full. Automatic dumping is undesirable because the small amounts dumped sometimes result in rapid freezing of the dump lines. Crew control of the dump valve permits dumping in a steady stream at a predetermined time, and, since a dump can disturb spacecraft trajectory, permits dumps to be coordinated with spacecraft maneuvers.

Telemetry sensors indicate the waste and potable tank quantities in percent of total tank capacity. The dump nozzle temperature is monitored to prevent freezing of dump lines.

The Water Management program automates management of waste and potable water. The program displays current amounts available and estimated amounts available at selected future times based on five spacecraft measurements. Additional estimates can be requested by manual entry. Manual overrides to telemetry inputs are provided to compensate for faulty transducers or lack of data.

The program computes and displays the amount of potable water available currently and at the next three planned meal times using the potable quantity sensor reading, fuel cell production, crew use, and potable tank inlet valve position. Meal times are entered manually.

The amount of waste water available currently and at the next three planned meal times is computed

and displayed using the waste quantity sensor output, current amounts of waste and potable water, cyclic accumulator production, fuel cell production, crew use, evaporator use, potable tank inlet valve position, and waste tank inlet valve position. The program also estimates when the waste tank will be 85 percent and 100 percent full to aid in scheduling water dumps.

Additional predictions of times and quantities are performed upon manual request. Manual entry of a waste or potable quantity cues the program to estimate when that quantity will be available. Entry of a future time cues the program to estimate the waste and potable quantities available at that time.

The program attempts to estimate the average crew consumption rate and the average fuel cell production rate by monitoring crew use and fuel cell production. The results are used to refine the initial estimates of crew consumption and fuel cell production rates. More accurate estimates of these rates would provide more accurate predictions of waste and potable quantities.

The estimates of future quantities are accurate, useful, and require minimal fligh controller interface. The computations indicating when the waste tank will be 85 percent and 100 percent full are especially useful when coordinating water dumps with spacecraft maneuvers.

The capability of the program to refine initial estimates of fuel cell production and crew use is not effective. The granularity of the data is not sufficient to determine crew use, and periods of missing or faulty data degrade the fuel cell production estimates. Although they reduce the level of automation, manual overrides for these quantities are more accurate and require fewer system resources.

The capability to obtain additional time and quantity estimates is useful and does not significantly effect the system resource requirements. These estimates aid in projecting conditions past the last meal time or between planned meal times, thus increasing the degree of water monitoring automation.

Program output would be more useful if presented in a combination plot and digital format. The display would show actual and predicted quantities plotted against time and would present readouts of plotted quantities in digital form. A display of this type would help the flight controller determine the accuracy of the predictions over a long period, analyze long term system trends, and make rapid analysis of future conditions.

ASSESSMENTS OF AUTOMATION IN THE GROUND SUPPORT ENVIRONMENT

The study assessed the practicality and economy of automated monitoring in the ground support system from data gathered throughout the ATSM study. Data was gathered from technical documents describing Apollo hardware and flight controller monitoring techniques, from technical meetings with NASA flight controllers, and from implementing and evaluating the ATSM test bed program. The assessments we formulated took into consideration the following items:

- Degree of automation
- Resource usage and size of overall system
- Complexity of overall system
- Man/machine communications
- Confidence levels attained.

The degree of automation measured how completely and effectively the automated program performed the monitoring function. Resource usage and overall size measured the resultant system impacts to central processing unit (CPU) utilization, to high speed and bulk storage requirements, and to any required peripheral devices or other computer resources. Both the intricacy of developing program logic and operational difficulties were considered in evaluating the extent and impact of user interfaces. Finally, confidence level assessed the accuracy and completeness of the program outputs.

The assessments of automation in the ground support environment are discussed in this section. The assessments are presented in the chronological order of automated monitoring program development, including the problems of definition and design, instrumentation and data validation considerations and testing and evaluation considerations. Within each subheading, the assessments are presented in order of significance, the most significant first. Finally, the overall assessment of the feasibility and benefits of ground support automation is discussed.

Definition of Flight Controller Functions

Automating flight controller functions begins by defining flight controller activities. The eventual benefits derived depend upon the quality of these definitions. Thorough, complete definitions permit developing more efficient and more useful programs. During development of the test bed program, several assessments of flight controller function definitions were formulated and should be considered in future automation activities.

1. Complete and detailed definitions of flight controller functions must be developed.

   Definitions must present all aspects of monitoring the onboard system — all the factors a flight controller considers, the order in which the flight controller considers these factors, and their relative importance. Formulating these definitions requires thorough knowledge — derived from actual design and operation — of the onboard systems. Expertise must include how the system is to be used, how the system will react under given conditions, and the potential failures of the system. Knowledge of current spacecraft systems should be provided by flight controllers or astronauts who have observed and monitored the systems. Knowledge of new systems such as the Space Shuttle should be provided by the engineers who design and build the systems.

   Knowledge of onboard systems must be combined with a thorough understanding of flight controller activities to formulate useful definitions. Expertise in flight controller techniques must include knowledge of current monitoring methods, air-to-ground interfaces with onboard systems, and a detailed knowledge of flight controller analysis techniques. This knowledge can be provided only by experienced flight control personnel.

Function definitions must include provisions for program actions if undefined situations arise. Since most flight controller functions are complex, there are likely to be some situations which are extremely difficult to program. Rather than precluding an effort to automate all functions, these situations should be considered in the function definitions.

2. Several qualified persons should collaborate on flight controller function definitions.

   Since flight controller functions are complex, better definitions will result when several experienced flight controllers or engineers contribute to the effort. Additionally, flight controllers use slightly different methods. More methods are available to choose from if several knowledgeable persons are involved.

Design of Automated Monitoring Programs

1. Primary and secondary output indicators should be designed to simplify the monitoring, analysis, and interpretation of onboard system status.

   The primary automated program output should be indicators which readily reflect the system state. For instance, a three colored light — clear, amber and red — could indicate whether the inertial measurement unit status is satisfactory, unknown, or disabled. Such indicators eliminate the need for constant monitoring of digital television displays.

   The secondary output should be prose descriptions of system anomalies and recovery procedures. These would be highly selective digital television displays which permit rapid absorption of data with minimal analysis and interpretation. Only currently significant information would be presented, and the messages would be simple and concise. In addition to simplifying monitoring, these displays require less system expertise to comprehend. The data presentation displays used in the current support system would be monitored only in contingency situations.

   It should be noted, however, that prose displays require a significant amount of storage to contain predefined messages. This storage represented 5 to 10 percent of the overall ATSM test bed size.

2. An automated monitoring program should be designed so that changes can be easily implemented.

   Precisely defining a total flight controller function is difficult. Changes in the initial definition of the automated monitoring program must be anticipated. Hardware configuration changes will have more impact on automated monitoring programs than they do on current data presentation programs. Designing the program so that changes can be easily made will decrease the implementation time for developing and testing program modifications.

3. Automated monitoring programs should extensively use process control tables.

   Process control tables are tabular specifications of actions to be performed by the executable routine. The tabular elements consist of control codes, processing constants, and pointers to related information. An executable routine scans these tables, element by element, and responds to the indicated action. This technique divorces the execution logic from the specific processing descriptions.

   Process control tables are especially important to automated monitoring programs for two reasons. First, these tables make processing modification easier, since the tables may be revised or replaced without disturbing the executable routine's logic. New and different operations also may be more easily included in the executable code. Second, program testing is easier when process control tables are used. Less executable code requires verification, and each possible logic path must be checked only once. The process control table logic often can be verified independently. When specific tabular processing requirements change, only the process control table must be verified. Thus, the effort both for initial development and testing and for subsequent changes is minimized using process control tables.

4. User interfaces are required in a ground support environment.

   Manual overrides are required to compensate for missing or bad data. When the onboard system sensor coverage is insufficient to provide all the needed monitoring inputs, these measurements must be entered manually. Similarly, when sensors fail, either the user must provide valid inputs so that the automated program outputs will be correct or additional program logic must be provided to compensate for the missing quantities.

   User interfaces are required to initiate actions recommended by automated monitoring programs since the ground support computer software generally does not directly interface with the onboard hardware systems. The highest level of ground support automation possible is the external indication of recommended actions. Such information must be monitored by the flight controller, relayed to the spacecraft, and executed by the astronauts.

Spacecraft Data Collection and Verification

The data available to automated monitoring programs is especially important since a primary goal is complete automation requiring minimum user interfaces. Specific ground support data collection and verification considerations are discussed below.

1. Sample rates may limit the effectiveness of an automated monitoring program.

   Because spacecraft data transmission is limited, all samples of all onboard measurements are not sent to the RTCC. Sample rate reduction of Apollo spacecraft data between Manned Spaceflight Network (MSFN) remote sites and the RTCC is but one instance. For example, certain measurements of 50 samples per second, collected onboard and transmitted to the remote site, are reduced to ten samples for Communications Command and Telemetry Systems (CCATS) input and one sample per second for RTCC input.

Since a higher sample rate is not available, accurate results from some automated programs are limited. For example, the reaction control system program which accumulates jet ON time sums only whole seconds since jet ON/OFF indications are available only approximately once per second. This degrades most performance analysis programs which require precise calculations.

2. Automated monitoring of onboard systems requires sufficient sensors to determine system changes.

The degree of automation is determined by the placement of sensors within an onboard system. Since sensor coverage on the Apollo spacecraft is not adequate for completely automated monitoring of some systems, limitations must be compensated for by man/machine interfaces. For instance, if a leak detection program recommends closing a valve as part of an isolation activity and no sensor is available to indicate the valve position, a manual entry is required to inform the monitoring program when the valve is closed. Similarly, within a checklist monitoring routine only those switch configurations that are reported by sensors can be automatically monitored. In the Apollo potable and waste water system, actual crew usage is not measured and therefore must be estimated. Thus, the degree of automation is a direct result of sensor coverage.

3. Data validation is mandatory for an automated monitoring program.

Since data validation logic compensates for or ignores faulty data such as transmission errors and hardware failures, it is the most important confidence factor in program monitoring. However, data validation significantly increases program implementation time and computer resources by requiring a large amount of designing, coding, and testing effort. Several validation techniques were implemented in the test bed program. These techniques include verifying the measurement with succeeding samples, testing the measurement against a preselected range which can be dynamically adjusted to reflect spacecraft operation, and checking associated parameters for overall consistency of measurements. Within the test bed program this data validation logic represents approximately 15 percent of the overall size.

4. Automated monitoring is limited by the granularity of the data received.

Data received in the RTCC provides for only 256 data points. Thus, the data appears to change by discrete amounts instead of smooth curves, and the granularity is sometimes too large to achieve the desired degree of automation. For example, a routine to refine the estimated crew water usage based on potable and waste water measurements was ineffective in the test bed program. One pulse code modulation (PCM) unit represents so much water (0.2 pound) that crew usage must accumulate significantly before the refinement routine can recognize any change.

Program Testing and Evaluation

Testing and evaluating the automated monitoring programs of the test bed program is very difficult. The experience gained during the study provided insight for formulating the program testing assessments.

1. Current data generation capabilities are not sufficient to economically test automated monitoring programs.

Data generation systems currently available for testing automated monitoring programs are either uneconomical for initial levels of testing or do not supply adequate test data. The following paragraphs briefly discuss the requirements for automated monitoring program test data, the requirements of an economical testing system, and why current testing systems do not meet these requirements.

Data to test automated monitoring programs must be realistic — parameter values must closely approximate onboard quantities and must change in the same manner and by approximately the same amount as the quantities being simulated would change. Parameters must also change concurrently with other parameters to accurately simulate onboard activity. For instance, when the inertial measurement unit (IMU) in an Apollo spacecraft torques, nine separate telemetry parameters change as the platform moves. Hence data to test an IMU monitoring program must include a case where the nine measurements change concurrently to simulate an IMU torque.

The economical requirements for a testing system are difficult to define; however, a good indication is the number of computers required. The most economical testing system would require only one computer. Testing automated monitoring programs is not as economical when additional computers are required.

2. Economical methods of generating test data for automated monitoring programs should be studied.

Program reliability depends upon complete and thorough testing and evaluation. Further study of testing automated monitoring programs is necessary to identify new and better methods of generating test data.

3. Programming man hours and computer hours must include provisions for reprogramming and retesting.

Critical evaluation of a program's usefulness should be performed during the program testing phase. When this evaluation was performed during the ATSM test bed program demonstration, IBM found that automation of all four flight controller functions could be significantly improved by implementing suggested program modifications. For instance, the SPS/G&N Preburn Checklist Monitor was dramatically improved by changing only a few predefined parameter limits. The SMRCS Leak Detection and Isolation program was much more useful when instructions to enable complete reinitialization of the program at flight controller command were added.

Flight controller functions usually involve analyzing complex factors to make decisions. It is difficult to include all factors and to evaluate the relative importance of each factor in a program. Therefore, the first attempt at automation of a function usually results in gross approximations of the flight controller's task. However, most programs can be refined if sufficient time is allowed for program redesign and retesting.

4. The ratio of testing hours to programming hours for automated monitoring programs is greater than for current ground support programs of comparable size.

During the testing phase of the test bed program, it was found that testing automated monitoring programs required considerably more man hours than testing current telemetry programs of the same same size. The difference in effort is illustrated by comparing the SPS/G&N Preburn Checklist Monitor, which is one of the test bed programs, and the Apollo 14 Universal Plot Processor. The Universal Plot Processor is a telemetry program which plots parameters against time and provides for extensive dynamic definition of plot characteristics.

The ratio of testing hours to programming hours is .55 for the automated monitoring program and .32 for the current ground support program.

	SPS/G&N Preburn Checklist Monitor	Universal Plot Processor
Program size in bytes	6274	8284
Man months to program (design and code)	6.0	6.3
Man months to test	3.3	2.0
Test-to-program ratio	.55	.32

The primary reason for the difference was the time spent preparing test scripts. Scripts for testing automated monitoring programs must represent very precise and complex conditions. Furthermore, analysis of onboard activities to be simulated by the test data is difficult and time consuming.

Overall Assessment of Ground Support Automation

Two primary questions must be asked in assessing the practicality and economy of ground support automation: (1) Is implementing automated monitoring programs feasible? (2) Do the benefits justify the effort?

Implementing automated monitoring is feasible in a ground support environment if the required resources are available. Since the demands on systems resources increase as the level of automated monitoring increases, the primary consideration is computer resources. The processing performed in the current ground support system is primarily data presentation. The next three levels of automation — status monitoring, malfunction diagnosis, and action recommendations — each cause increased system demands. Each level also represents greater difficulty in defining the flight control function and in designing and testing the automated monitoring program.

Alternatives are available for controlling computer resources. Some flight controller functions can and should be implemented in off-line systems to relieve the burden on the real time computer. Most aspects of consumables management functions and post-performance analysis functions are off-line system candidates. They are concerned with long-term trends, future activity planning, and detailed analysis of past occurrences. This type analysis and decision-making does not require as rapid response as other functions and does require considerable storage for historic samples. Thus, these functions can more effectively be implemented in off-line support systems.

As previously identified, computer processing load can be controlled by manual interfaces and supervisory processors. Manual interfaces permit direct flight controller initiation or termination of programs, whereas a supervisor can control load based on program priorities, mission phases, or other criteria.

The second major consideration is that sufficient data must be available. To achieve complete automation of an onboard system, both adequate sensor coverage and adequate sample rates must be provided. The inadequacies of either can in some instances be compensated for by user entry of the needed data, thereby increasing the user interface requirements. Nevertheless, the overall degree of automation is primarily dependent upon the direct availability of spacecraft data.

The third and fourth major considerations are function definitions and test data generation. Flight control functions must be precisely and completely documented in sufficient detail before automation is feasible, and economical test data generation capabilities must be developed to support initial levels of program testing.

Ground support automation is desirable because the computer system can assume much of the workload and free flight control personnel for higher order tasks. First, the automated monitoring program can monitor a total system in a shorter time span than a flight controller, it can often analyze system status faster, and it can consider more possibilities. The flight controller, on the other hand, can consider only the most probable situations when performing status or malfunction analysis with a given time constraint.

Second, an automated monitoring program can provide continuous monitoring over extended periods without being effected by the tedious nature of monitoring. This task cannot be done as effectively by an individual. Therefore, the monitoring program is more likely to readily identify any anomalies which may occur.

Finally, an automated monitoring program reduces the amount of flight controller time required to monitor a system. The simplified high levels of output — status indicators and selective prose displays — reduce both the amount of information to be monitored and the analysis and interpretation required; therefore, a flight controller can monitor more systems concurrently and can readily evaluate current systems or systems that are even more complex.

A new concept of mission ground support can evolve through the use of automated monitoring programs. Suitable definition of high level outputs would permit fewer flight controllers to monitor routine mission activity. The particular system specialist, no longer responsible for this nominal mission monitoring, would provide support for unprogrammed or contingency problems.

ASSESSMENT OF AUTOMATION ONBOARD FUTURE SPACECRAFT

To assess the practicality of placing automated monitoring programs in an onboard environment, information was gathered from experience with previous ground support systems, experience and knowledge gained from implementing the ATSM test bed program, and from contacts with other groups participating in future mission studies. Information from these three sources was applied to an onboard environment and used to assess onboard automation.

Assessing the practicality of integrating automated monitoring programs onboard future spacecraft involves weighing the resources required for such an effort against potential benefits. Onboard automation depends upon the availability of onboard resources, the most important of which is the onboard computer's capabilities. As the level of automation increases, the computer assumes more of the burden of monitoring and controlling the spacecraft. This requires more CPU time, data storage, and input and output interfaces as the size and complexity of the programs increase.

Methods have been identified to enhance the use of onboard computer resources. When ground support of spaceflight is still maintained, functions such as consumables management and post-performance analysis could be automated in the ground support system. These functions do not require instantaneous decisions and often do require large amounts of storage for historical data. Additionally, measurement samples could be recorded onboard and periodically "dumped" to the ground system to avoid some of the current sample limitations.

Manual controls and supervisor programs are two other significant methods for controlling computer resource use. Manual controls for initiating and terminating processing would allow the flight crew close control of the computer load. A supervisor program to monitor and control the processing loads could establish processing priorities during heavy CPU usage to ensure that the most critical processing was performed first. Thus if sufficient computer resources were not available, these two techniques would facilitate at least partial automation of onboard monitoring functions.

Another resource requirement for onboard automation is sufficient sensor coverage of onboard systems. Sensors must measure enough parameters at a high enough sample rate to allow thorough and accurate determination of system state and performance. This data must be communicated to the program in a timely and efficient manner to best automate system monitoring. Lack of sensor coverage can sometimes be overcome by manually entering parameters not measured by a sensor; however, manual inputs reduce the potential level of automation and increase the burden on the astronaut.

If the goal of total onboard automation and resulting spacecraft autonomy is to be achieved, the hardware and software interfaces must be modified to permit software control of hardware. Current hardware and software interfaces permit only monitoring of hardware activities. If the software does not also have hardware control capabilities, the astronaut monitoring and interface requirements would be comparable to flight controller ground support responsibilities.

However, before implementing any programs to monitor and control hardware, the capability of the hardware to monitor itself should be considered. Switching to redundant backup systems in the event of primary system failure might be more practical than inplementing software to control the system. The ability of some hardware to detect a fault in itself and notify the computer of a suspected anomaly could be used to reduce the size of the fault detection logic in the program.

Onboard automation has several advantages over ground support automation. Most of the advantages of onboard automation arise from the proximity of the hardware and software. Closer integration of hardware and software facilitates data collection, decreases software-to-hardware response time, and enables a higher level of automation.

Programs placed in an onboard environment can maintain constant contact with the data source. The necessity of long range data transmission with its associated sample rate reduction is also eliminated. These data implications result in constant and more complete analysis of onboard systems, quicker detection of problems, and faster diagnosis of anomalies. Constant receipt of data enables the software to maintain a complete history of system performance and provide accurate analysis of system trends.

The level of automation of flight monitoring and control is potentially higher onboard than in the ground support environment. Concurrent design of software and hardware would permit the most appropriate trade-offs and would ensure sufficient interfaces. Potentially, software automation could free the astronaut from tedious monitoring and routine spacecraft operation.

REFERENCES

1. IBM, 8 August 1970, "ATSM Test Bed Program Functional Specification."
2. IBM, 11 November 1970, "ATSM Test Bed Program System Design Specification."
3. IBM, 1 June 1971, "ATSM Final Technical Report."
4. NASA, 15 January 1970, "Apollo Operations Handbook Block II Spacecraft, Volume 1," SM2A-03-BLOCK II - (1).
5. NASA, 15 December 1969, "Apollo Operations Handbook Command and Service Module, Volume 2," SM2A-03-BLOCK II - (2).
6. NASA, 19 April 1968, "Apollo Training Block II CSM Propulsion Subsystem."
7. NASA, 1 October 1969, "Command/Service Module Handbook, CSM 108 through CSM 111," FC007.
8. NASA, 20 October 1970, "CSM Malfunction Procedures," NASA Document SICB3210081-341.

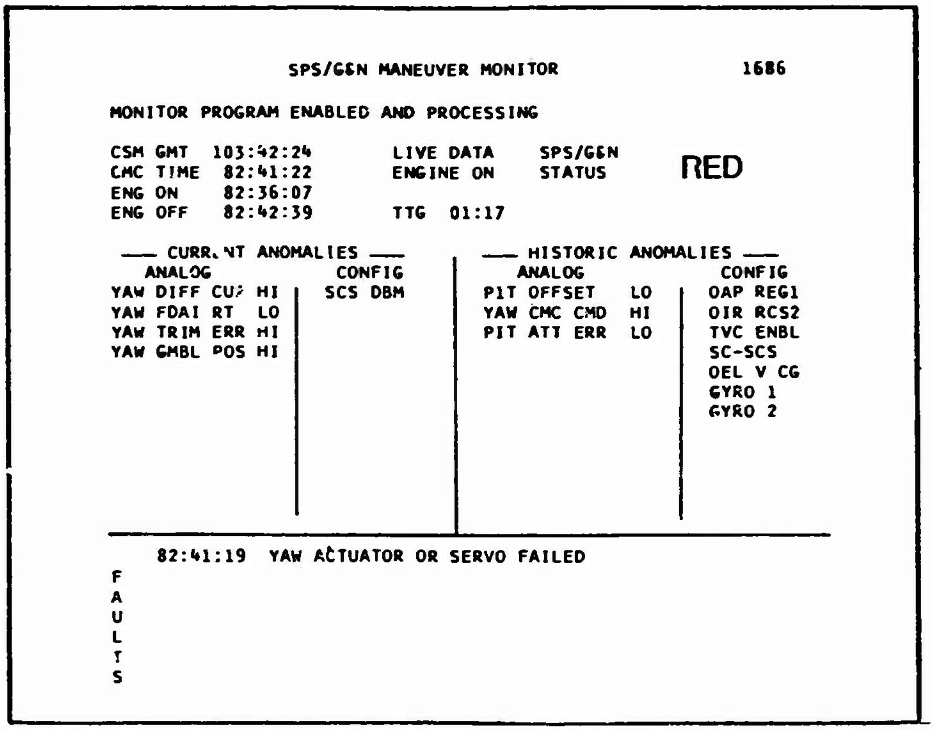


Fig.7-1 SPS/G & N display

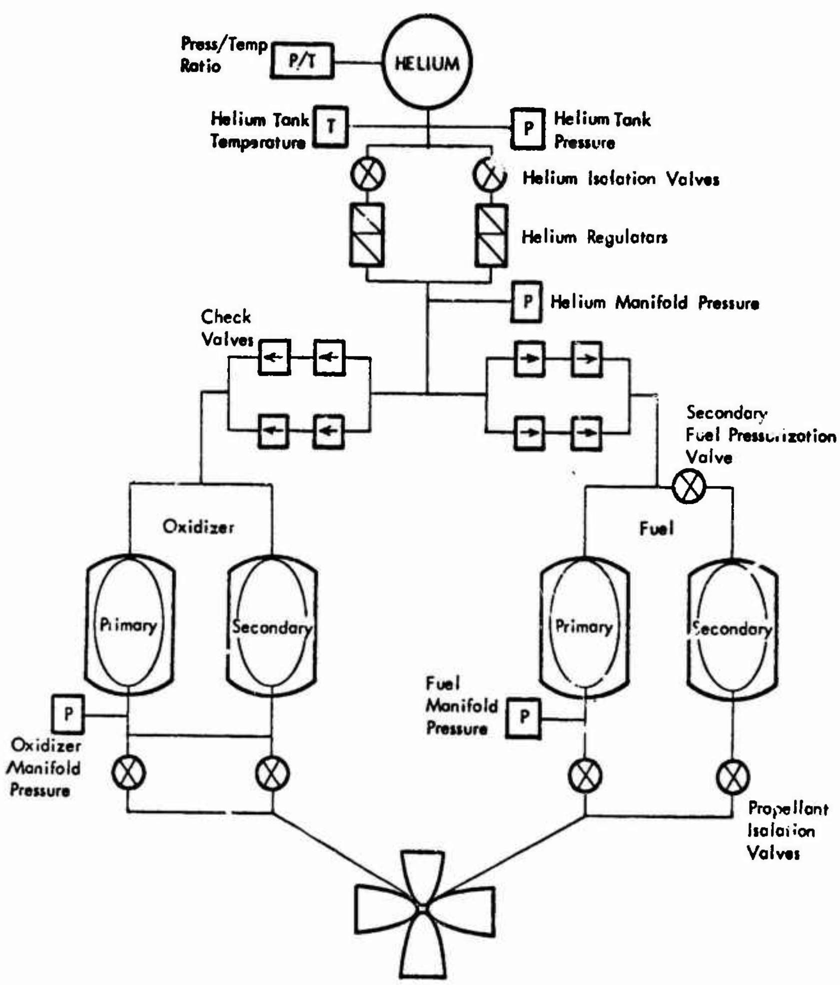


Fig.7-2 SMRCS quad

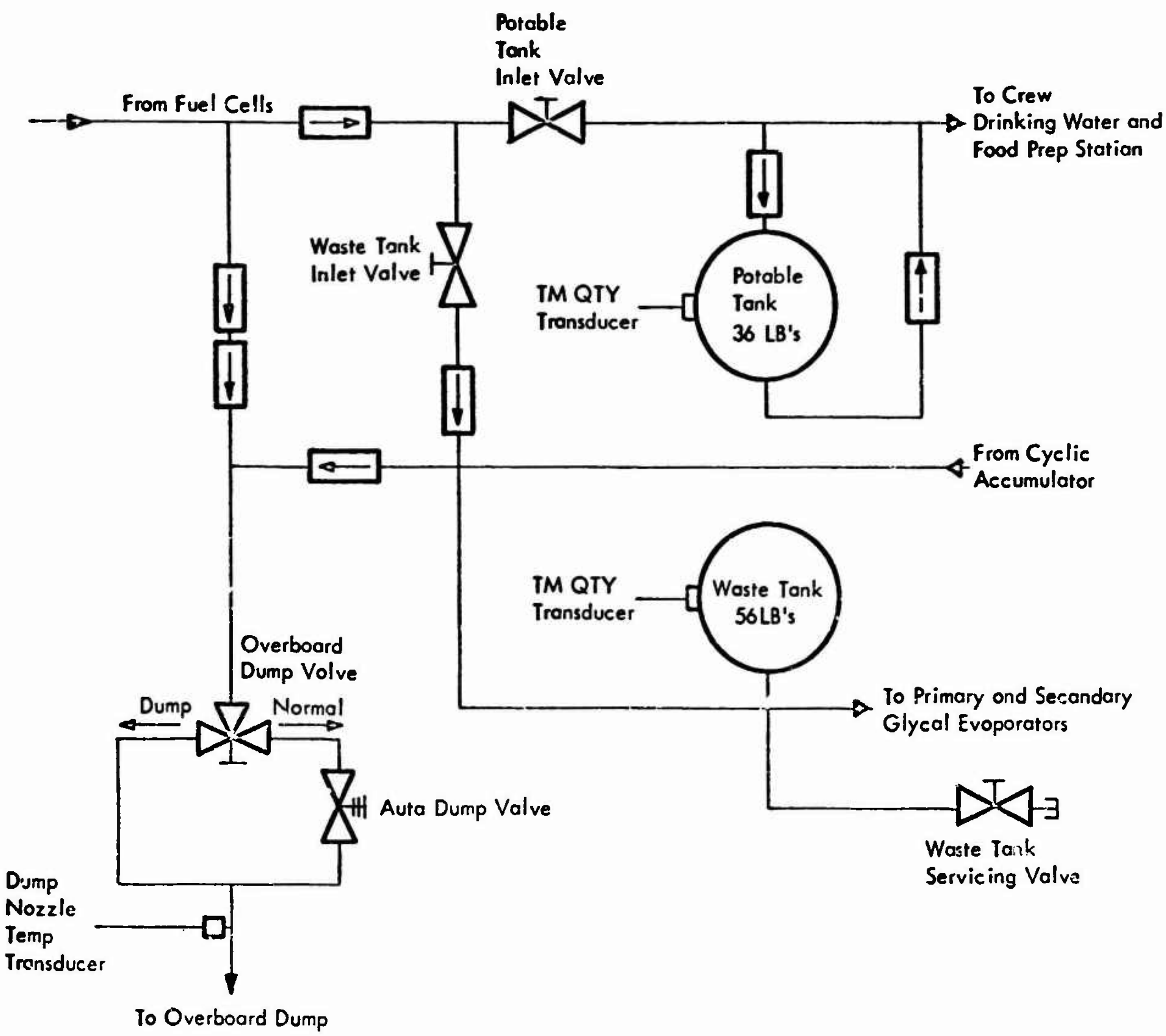


Fig.7-3 Water system

SOME DEVELOPMENT TRENDS IN THE INTEGRATION OF ELECTRONIC SYSTEMS IN THE SWEDISH AIRCRAFT 37 VIGGEN

by

Bengt Sjöberg
Systems & Avionics Department
Aerospace Division
SAAB-SCANIA AB
Linköping
SWEDEN

SUMMARY

The Swedish 37 VIGGEN Aircraft is being developed in several versions and the electronic systems of the attack version and the later fighter version are compared and some development trends are discussed. An increased rôle of the central computer is recognized as well as a trend towards digitalization of several subsystems.

1. INTRODUCTION

The 37 VIGGEN aircraft has been developed by SAAB-SCANIA, Linköping, for the Swedish Air Force. Using the same basic airframe and engine three versions of the aircraft will be produced for attack, reconnaissance and fighter combat roles respectively. The attack version is presently under series delivery to the Swedish Air Force with the reconnaissance version to follow. The fighter-interceptor version is in its early stage of development and intended for series delivery in the late seventies. A common trainer version is also being produced.

The aircraft is a single-seat STOL aircraft mainly recognized by its canard configuration, see fig. 1.

The design of the electronic systems of the different versions is carried out in close cooperation between SAAB-SCANIA as the main contractor, the Swedish Air Materiel Department and the different equipment contractors.

The main object of this paper is simply to show some development trends in Swedish, military avionics, with special attention to the use of digital equipment.

2. THE ELECTRONIC SYSTEM OF THE ATTACK VERSION

The electronic system of this version was defined in the time period 1962-64 and among the decisions taken was to generate a main part of all tactical combat functions by use of a central, digital computer, see fig 2. Also included was an electronic head-up display with a mainly analog wave form generator. The flight control system was independent of the central computer and so was a conventional set of primary and standby cockpit instruments.

The signal communication with the digital computer was mainly analog with A/D and D/A time-shared converters in the computer I/O-unit. The system further comprised a conventional, two-gyro reference platform with mainly analog electronics, an analog air data computer, a fixed, navigation Doppler sensor equipment, a microwave ILS, a ground mapping and attack radar with analog electronics, an analog automatic flight control system, a fixed accelerometer package with analog signal generation, a radio navigation equipment an a radar altimeter, to mention the more important parts of its mainly analog, electronic sensor system.

To a limited extent, digital communication with the central computer was used.

The displays, including the electronic HUD and the radar indicator, were driven with a conventional, analog technique.

The central computer had the following essential data:

• memory:	ferrite core RAM
• memory volume:	16K X 13 bits
• data word length:	26 bits
• instruction word length:	13 and 26 bits
• cycle time	2,8 µsec.
• add time	5,6 µsec.
• multipl. time	23,8 µsec.
• weight of power unit, memory unit, processor unit and I/O units:	~ 60 kg
• power consumption:	600 W
• I/O signals:	~ 70 DC analog, 35 x 13 bits parallell binary

The number of system variables in the attack version computer programs, that is variables which carry main information of a quantitative or logical nature between the different program blocks, is about 700 with the number of program blocks about 30.

3. THE ELECTRONIC SYSTEM OF THE FIGHTER-INTERCEPTOR VERSION

The fighter-interceptor version is now being designed very close to ten years later than the attack version, During this time electronic systems have moved towards digital solutions. The use of general purpose computers is no longer felt as pioneer work as was the case in 1962 when the attack version digital computer was introduced. Figure 3 illustrates with double contours where mainly digital equipments are now being used for the fighter system.

The general system layout remains from the earlier version with a central, general purpose digital computer for the generation of main tactical functions. The air data computer is now digital and so are important sections of the automatic flight control system including autothrottle functions. The gyro platform has been replaced by an inertial platform which is tied to the central computer for the generation of the navigation and velocity vector functions. Furthermore the display generation has become digitalized and now drives three electronic displays including an electronic map display. The radar also has a digital control and data processing unit with digital interface with the display generator. Several tape equipments are tied to the digital processors.

Each of the different digital programmable processors that have been mentioned above communicate with the central computer via time-shared, serial binary channels.

A digital data communication link from the ground control system is tied to the central computer. The central computer has grown due to increased demands from the intercept functions. Some of its main data are:

• memory:	ferrite core RAM
• memory volume:	48 K x 16 bits
• data word length:	32 bits
• instruction word length:	16 and 32 bits
• cycle time:	1,9 μsec.
• add time:	3,8 μsec.
• multipl. time:	7,2 μsec.
• total weight:	27 kg
• power consumption:	500 W
• I/O signals:	~ 35 DC analog (Inputs)
	46 x 16 bits serial binary
	32 bits parallell binary
	2 parallell binary 8 bit buses

The number of system variables has now grown to about 1500 and the number of program blocks to about 40.

We can thus recognize two trends in the digitalization process. Firstly we can see an increased use of digital techniques in several subsystems, where a programmable processor portion is combined with special purpose circuitry to generate special functions.

Especially where high processing rates are needed like in the display generator or in the radar data processing, local, special purpose computers are coming into use. Secondly the demands on the central computer have increased due to the need of more, integrated system functions. The signal communication system between the CDC and the rest of the electronic system is mainly serial binary. The conversion has to a large extent been moved out to the sensor equipments or to special adapter units which are located close to the places of use.

4. RELIABILITY, MONITORING AND TEST FUNCTIONS

4.1 Reliability relationships

The situation concerning the total failure rates of the more central parts of the electronic systems in the attack version versus the fighter version is shown in figures 4 and 5. It is seen from fig 4 that the central computer reliability in the attack version was well matched to the reliabilities of the main sensor and display equipments. It is also evident that the basic aircraft with the engine and the main supply systems give a very large contribution to the overall relibility.

In the fighter version, fig 5, we can see certain reliability improvements as more digitalized equipment is introduced. The central computer has improved due to improved component technique and due to an increased use of digital signal communication, in spite of a capacity and performance increase of about 3 times relative to the CDC in the attack version. Although the signal communication system has been modernized largely to serial binary transmission, it still retains its contribution to failure rate, partly due to the increase of information flow, partly due to the high number of components that are still needed in each channel. The digital ADC gives a definite improvement, and so does the introduction of the inertial platform, partly due to the elimination of repeater servos, and partly due to the generation of inertial navigation functions in the CDC. The radar MTBF has not been improved, nor have the electronic displays. The latter have been expanded in the fighter version to 3 displays from two in the attack version. The two radars can hardly be easily compared as they are designed for quite different rôles.

It can be seen from fig 5 that the basic aircraft is an even more dominating contributor to the fighter version total failure rate.

4.2 Monitoring and reversionary modes

The ability of digital equipments to perform selfmonitoring is used as much as is reasonable in the fighter version together with monitoring of the type used in analog equipments already in the attack version. In an attempt to give the pilot a reasonable amount of reversionary modes which he can learn to handle, three main levels of system operation have been defined, see fig 6. The highest level represents all tactical functions and systems at full performance. The basic level contains the central computer in operation and has retained a certain tactical ability. The back-up level is designed to make it possible to fly the aircraft safely back to its base and does not depend on the central computer.

Redundancy techniques are applied mainly in the automatic flight control system where a two-channel, fail-safe redundancy is used for sensors and critical parts of the control system.

It can be seen from fig 6 how the main part of the electronic system has been grouped in 3 very different failure rates for the 3 levels.The supply systems likewise through redundancy can show a definite step to the back-up level. The engine and the primary flight control system are large contributors in all three function levels and cannot be divided like the electronic system. This should be a question of concern to the designers of engines, fuel systems and mechanical-hydraulic systems.

4.3 Built-in test

With the introduction of more digital equipment there is a growing trend towards built-in test functions. In the attack version centralized test for fault localization and performance is performed by use of an external "test vehicle". This is a costly and time-consuming process and the fighter version will see a larger portion of built-in test functions, controlled by the central computer in cooperation with the other digital equipments, and which will to a large extent be independent of elaborate external equipment.

5. CENTRALIZED VERSUS DEDICATED COMPUTER CONCEPTS

The discussion concerning centralized versus dedicated computers deeply engages people in the avionics field in Sweden as well as elsewhere. We are of the opinion though, that it is not a question of either - or, but rather a question of both. Reliability investigations naturally show that in general simple functions get their reliability decreased if they are integrated in a larger computer, whereas more complex functions get their reliability decreased if they are divided between several small computers.

Hence simple functions that are essential for flight safety will continue to be mechanized with independent, small computers, in some cases with the use of redundancy techniques to achieve fail-safe or fail-operational function. An example of this is the flight control system. Another reason for the application of small, separate computers is the need for very high iteration rates, often combined with the use of considerable amount of special purpose circuitry.

Examples of the latter trend are the digitalization of the electronic display generator and of the radar. Of importance here also is the management of the equipment development, where a functionelly natural boundary should be defined so that the volume of communication with the rest of the system is reasonably small.

When we come to the general tactical functions of a modern, single-seat military aircraft such as navigation, intercept and vectoring functions, fire control, integrated display and control functions etc. we believe in the use of one, central computer in order to keep the total reliability up. Another reason is the high degree of integration of functions and mode logic which is necessary if a single man shall be able to operate the system in combat situations, and which requires a high volume of communication between the different functional blocks. We can for example recall that the number of quantitative an logical variables communicating between the CDC program blocks in the fighter version is about 1500. (In a certain sense we may consider that the single brain and limited sense system of the one man is matched by the single, central computer with its integrated functions adapted to an optimized communication with this one man through a highly integrated display and control system).

More sophisticated weapons are satisfactorily controlled by a central computer, due to the relatively low weapon reliabilities, whereas simple weapons may be given a back-up level mode independent of the main part of the electronic system including the CDC, in order to retain a certain tactical self-defense capability in cases of electronic system failures. It is thus the authors opinion that we will see both a trend towards increased, dedicated digitalization in sybsystems, especially where flight-safety is involved and hence often redundancy techniques have to be used, as well as a trend towards centralization of navigation and tactical functions which require a high degree of integration especially in single seat aircraft. The dedicated approach should be used with caution, though, and only where special reasons are proven to exist, in order to keep the total component amount low. Getting the aircraft off the ground with full performance at a low maintenance cost should remain an important object.

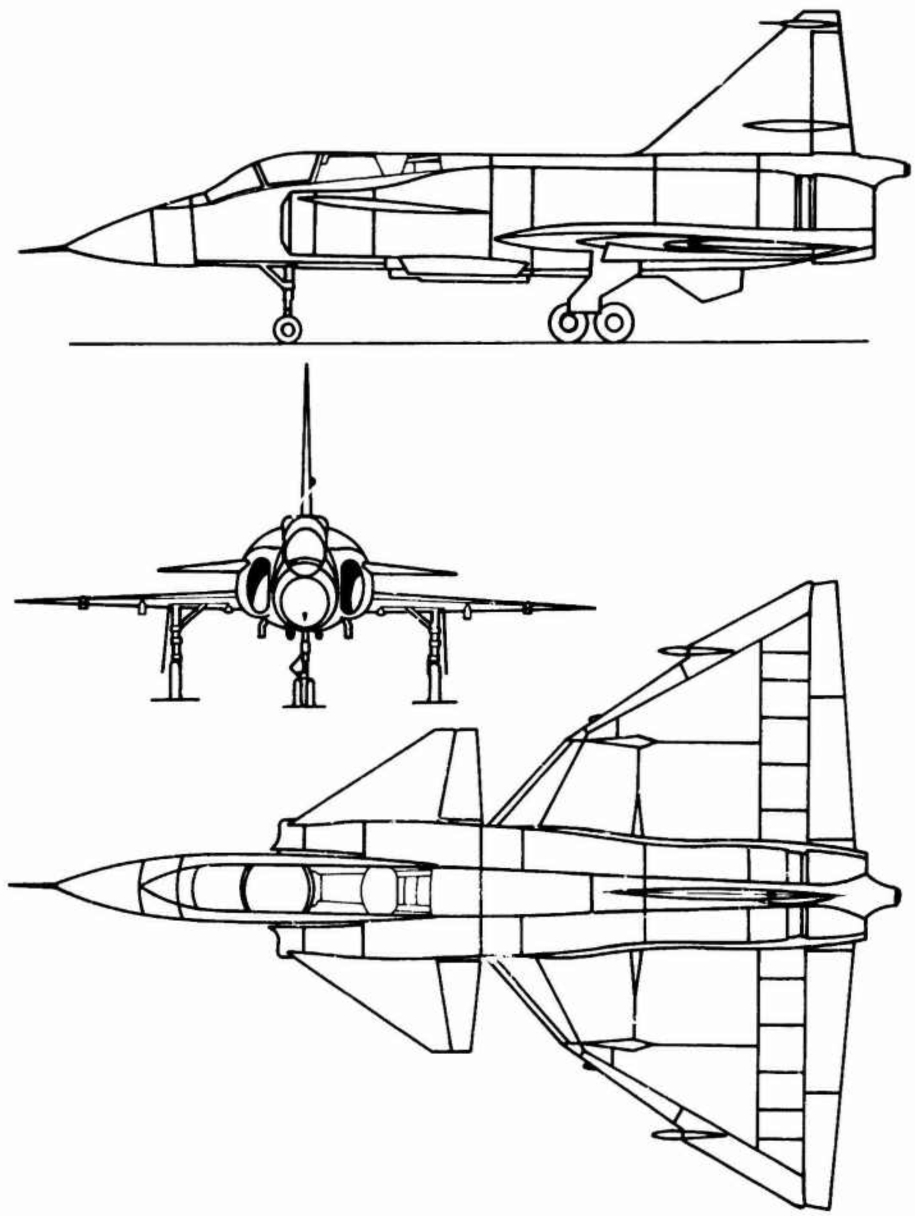


Fig. 1. The Aircraft 37 VIGGEN

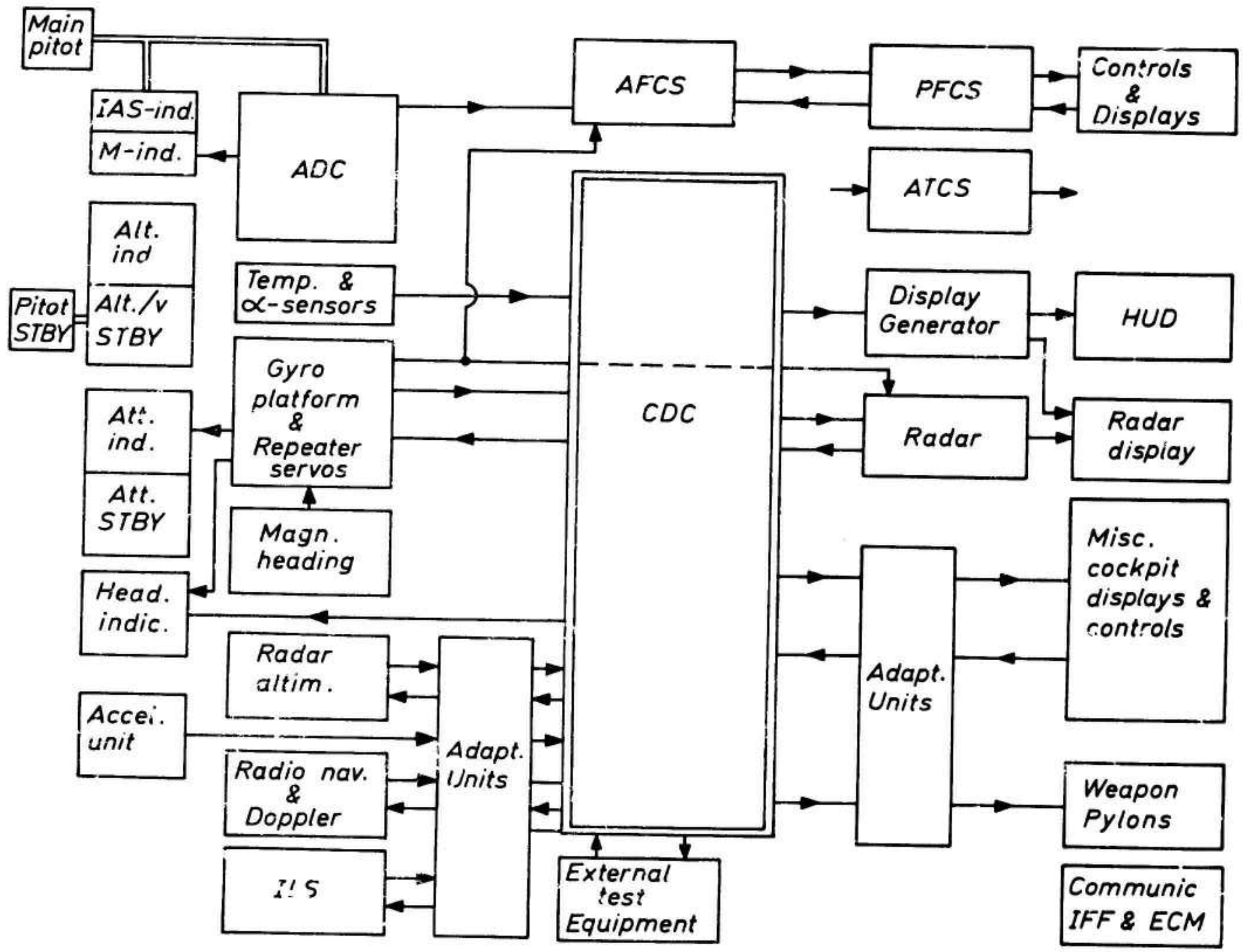


Fig. 2. Attack Version Electronic System

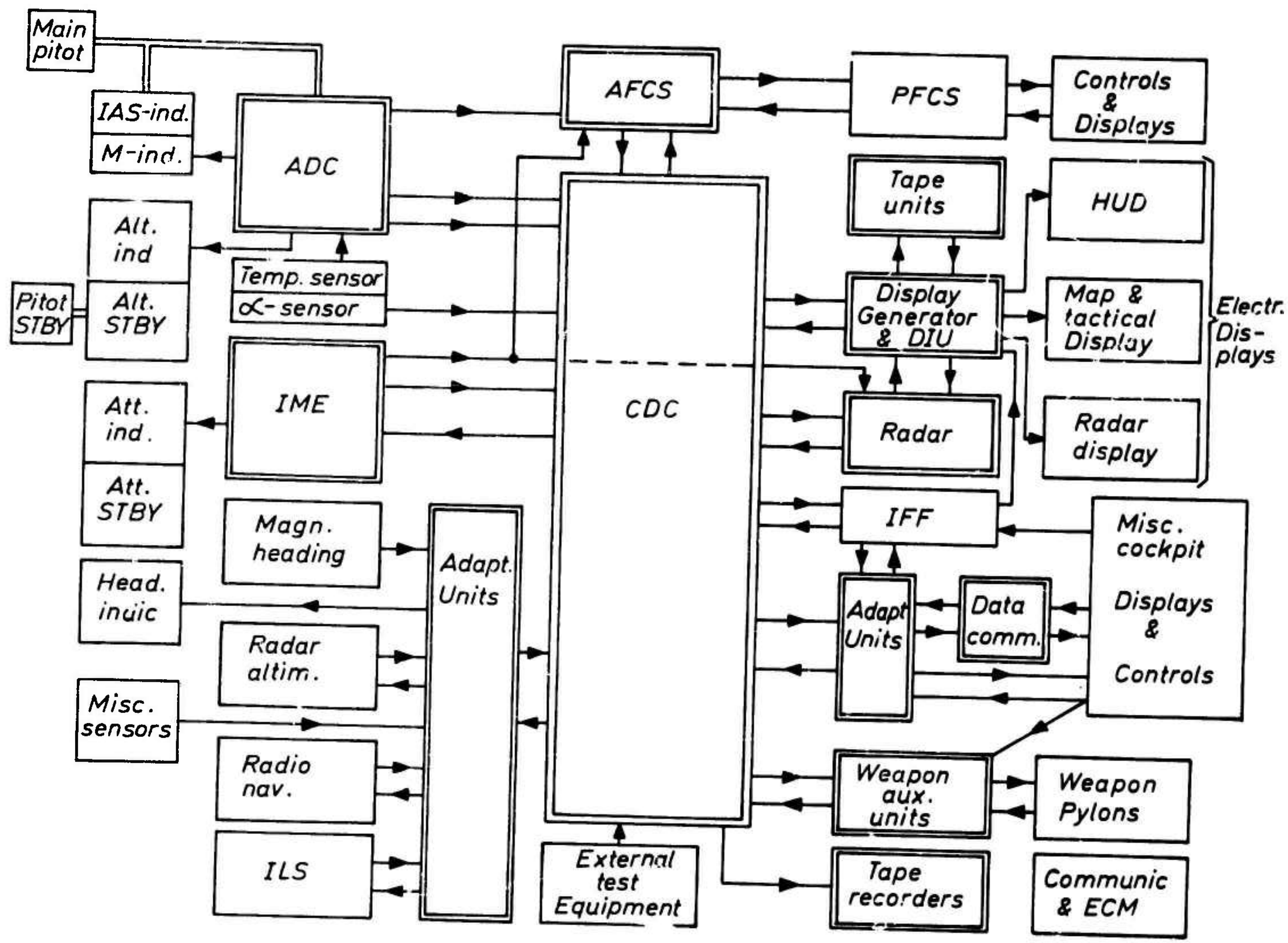


Fig. 3. Fighter Version Electronic System

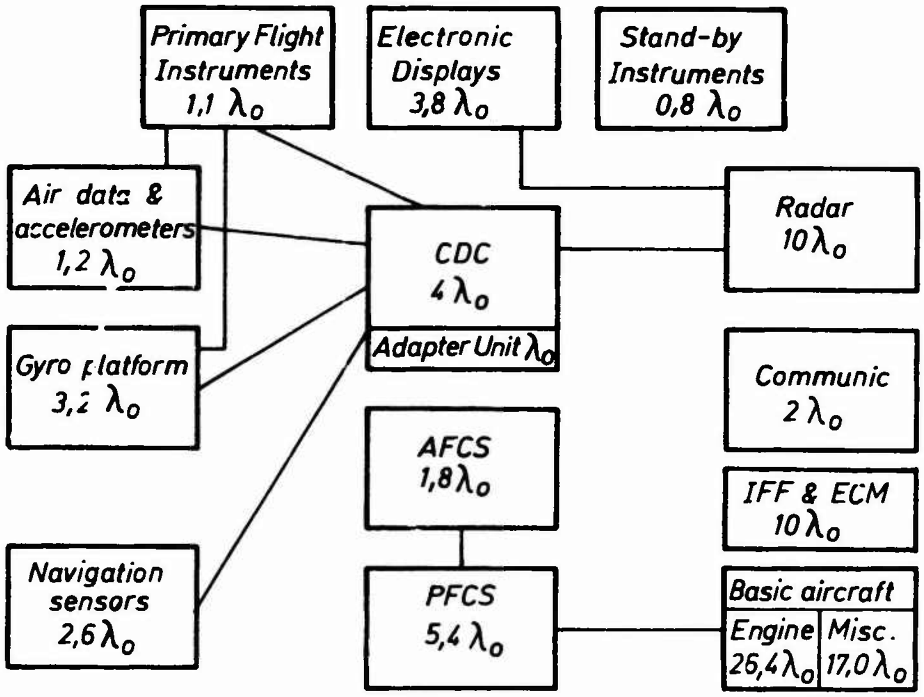


Fig. 4. Attack Version. Relative Failure Rates

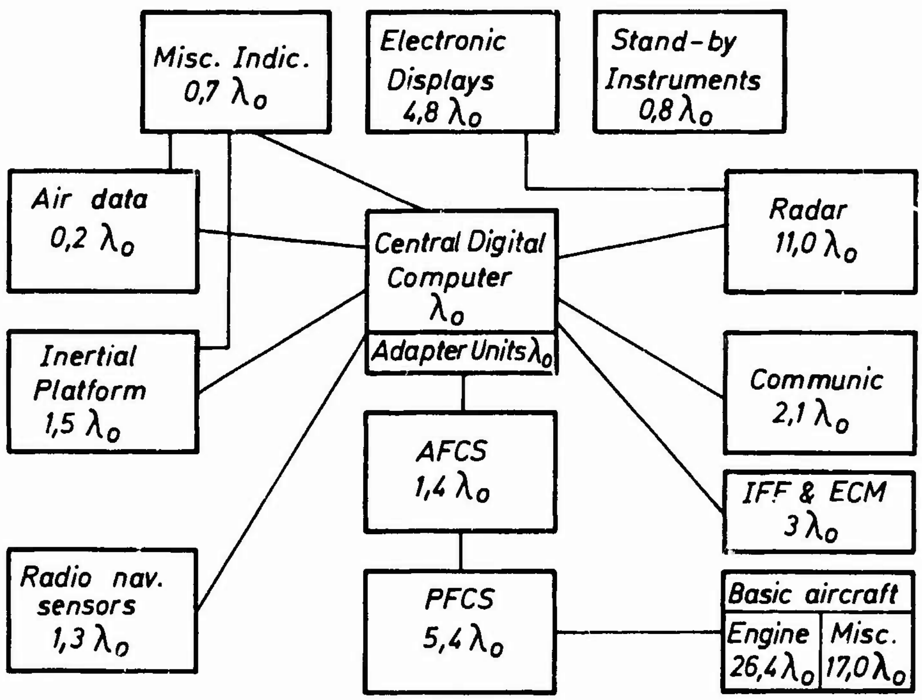


Fig. 5. Fighter Version. Relative Failure Rates

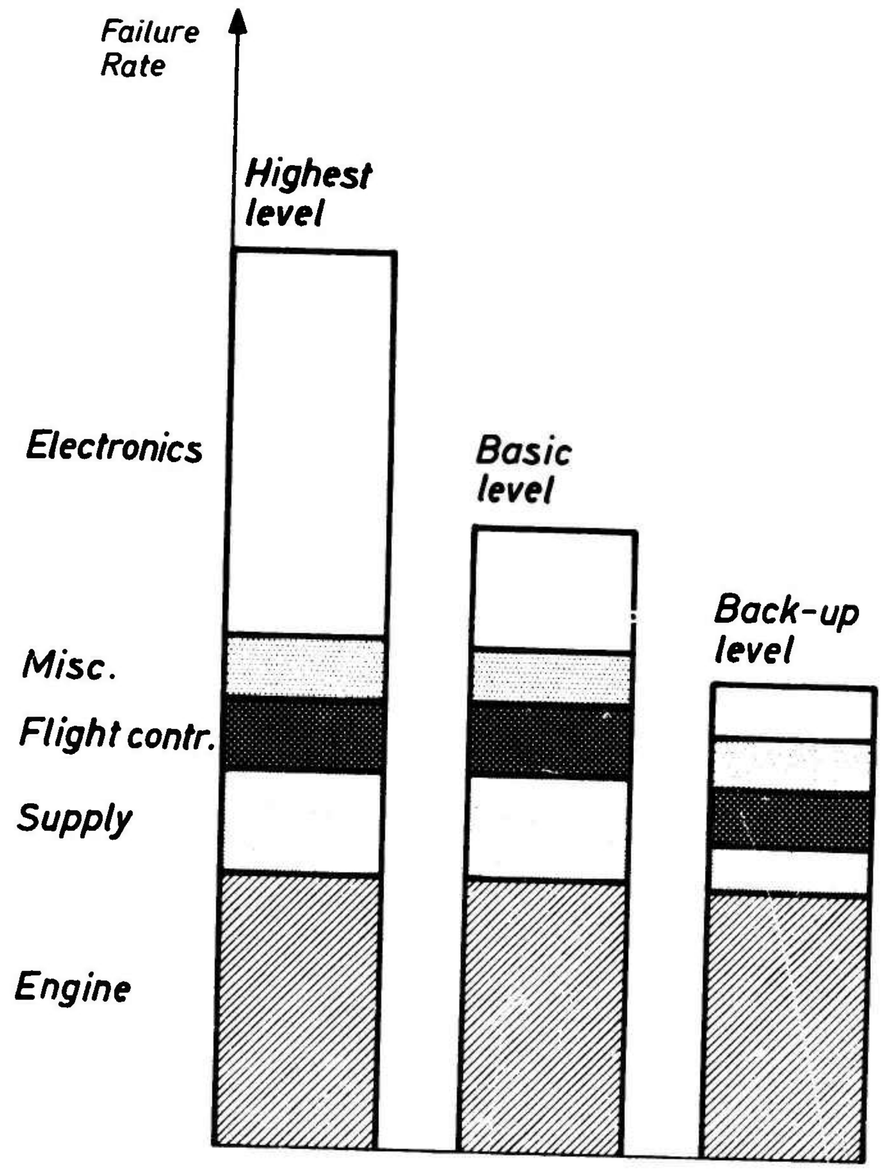


Fig. 6. Failure rate and function levels

MULTILOOP ATTITUDE CONTROL SYSTEM FOR A SATELLITE WITH FLEXIBLE BOOMS

R. DI LORENZO & E. DE BERNARDIS

AERITALIA, Centro Elettronico Avio
CASELLE TORINESE
ITALY

SUMMARY

A class of momentum exchange devices control configurations has been considered, namely that which provides a momentum quite larger along one body axes rather than along the other ones.
The general equations of a satellite controlled in such a way have been used, in order to provide a control system which is independent from the particular devices used (double gimballed flywheel CMGs, etc....); these equations have been modified in order to take into account that the satellite has a couple of flexible booms (for instance, very large roll-out solar arrays etc.).
A simple multiloop controller has been designed for such equations, and it is shown that to adapt it to each particular actuators configuration it is only necessary to design three very conventional inner control loops.

Finally, a simulation of the full flexible system has been made, using FORTRAN V, with reasonable numerical values of the satellite dynamics parameters, where it is shown that a controller designed considering rigid the whole satellite results either in instability or very degraded pointing accuracy.

1. INTRODUCTION

The majority of papers published, in the authors' knowledge, on satellites with flexible appendages (solar arrays, antennas, etc.) may be classified broadly in two main classes: a) development of models and b) stability analysis of the uncontrolled system.
For point a) see for instance ref. 3, 6, 9, and for point b) see ref. 4, 7, 8.
For what concerns the design of an attitude control system of a satellite of this type (referred to, heretofore, as "flexible satellite"), generally the procedure depicted in fig. 1 is followed:
a control system is designed considering rigid the satellite's appendages, then a simulation is made where the appendages are modeled as flexible; if the simulations results are still acceptable the A.C.S. designed in such a way is retained.
When the flexible appendages are very large, generally, owing mainly to coupling effects between the satellite's axis, a large loop gain results in instability, and a small loop gain results in very degraded pointing accuracy (see ref. 1).
This fact points out the need of designing a controller specifically for the satellite's dynamics with the appendages considered flexible, as they actually are (see ref. 2).
Attempts in this direction are very few todate (see ref. 2, 5, 6,); a new approach is proposed herein.

2. MODELING OF THE SATELLITE

We will consider here a very simple satellite configuration, i.e. the one shown in fig. 2 where:

xyz = principal axes of inertia of the undeformed satellite

XYZ = axes through the satellite center of mass, fixed in space

φ, θ, ψ = Euler angles, pitch, roll and yaw

p,q,r, = Satellite main body angular velocity components referred to the principal axes of inertia.

The flexible solar arrays are modeled here as shown in fig. 3 (see ref. 2): a point mass may have linear two-degrees-of-freedom movements with respect to a material frame which can only rotate with respect to the x axes. All the movements are harmonic with some damping, assumed, for simplicity, proportional to the velocity.

The dynamics of the satellite, when momentum exchange devices do not act within it, can be described in terms of Lagrange's equations of motion (see ref. 10) as follows:

$$\frac{d}{dt}\frac{\delta T}{\partial p} - r\frac{\delta T}{\delta q} + q\frac{\delta T}{\delta r} = M_x$$

$$\frac{d}{dt}\frac{\delta T}{\delta q} - p\frac{\delta T}{\delta r} + r\frac{\delta T}{\delta p} = M_y$$

$$\frac{d}{dt}\frac{\delta T}{\delta r} - q\frac{\delta T}{\delta p} + p\frac{\delta T}{\delta q} = M_z$$

$$\frac{d}{dt}\frac{\delta T}{\delta \dot\alpha_y} - \frac{\delta T}{\delta \alpha_y} = C_y\dot\alpha_y - K_y\alpha_y \qquad \frac{d}{dt}\frac{\delta T}{\delta \dot\beta_z} - \frac{\delta T}{\partial \beta_z} = -C_z\dot\beta_z - K_z\beta_z \tag{1}$$

$$\frac{d}{dt}\frac{\delta T}{\delta \dot\alpha_z} - \frac{\delta T}{\delta \alpha_z} = -C_z\dot\alpha_z - K_z\alpha_z \qquad \frac{d}{dt}\frac{\delta T}{\delta \dot\alpha_\rho} - \frac{\delta T}{\delta \alpha_\rho} = -C_\rho\dot\alpha_\rho - K_\rho\alpha_\rho$$

$$\frac{d}{dt}\frac{\delta T}{\delta \dot\beta_y} - \frac{\delta T}{\delta \beta_y} = -C_y\dot\beta_y - K_y\beta_y \qquad \frac{d}{dt}\frac{\delta T}{\partial \dot\beta_\rho} - \frac{\delta T}{\delta \beta_\rho} = -C_\rho\dot\beta_\rho - K_\rho\beta_\rho$$

Where T is the kynetic energy of the system; M_x, M_y and M_z are the torques (external and/or jets) applied to the main body of the satellite; Ky, Kz, K_ρ the spring constants and C_y, C_z, C_ρ the damping factors, as shown in fig. 3. In fig. 3 are also shown the free coordinates $\alpha_y, \beta_y, \alpha_z, \beta_z, \alpha_\rho, \beta_\rho$ which describe the positions of the movable parts with respect to the main body. Dot indicates differentiation with respect to time.

I_x, I_y and I_z are the principal moments of inertia of the main body, and I_ρ is the moment of inertia of the movable material frame with respect to the x axes.

If we consider α_y ...... β_ρ small, equations (1) become:

$$I_x\dot p + qr(I_y - I_z) - I_\rho(\ddot\alpha_\rho - \ddot\beta_\rho) = M_x \qquad I_y\dot q + rp(I_z - I_x) + Q\ell(\ddot\beta_z - \ddot\alpha_z) = M_y$$

$$I_z\dot r + pq(I_x - I_y) + Q\ell(\ddot\alpha_y - \ddot\beta_y) = M_z \tag{2}$$

$$\ddot\alpha_y + \frac{C_y}{Q}\dot\alpha_y + \frac{K_y}{Q}\alpha_y = -\dot r\ell \quad ; \quad \ddot\beta_y + \frac{C_y}{Q}\dot\beta_y + \frac{K_y}{Q}\beta_y = \dot r\ell \quad , \quad \ddot\alpha_z + \frac{C_z}{Q}\dot\alpha_z \; \frac{K_z}{Q}\alpha_z = \dot q\ell$$

$$\ddot\beta_z + \frac{C_z}{Q}\dot\beta_z + \frac{K_z}{Q}\beta_z = -\dot q\ell \quad ; \quad \ddot\alpha_\rho + \frac{C_\rho}{I_\rho}\dot\alpha_\rho + \frac{K_\rho}{I_\rho}\alpha_\rho = -\dot p \quad ; \quad \ddot\beta_\rho + \frac{C_\rho}{I_\rho}\dot\beta_\rho + \frac{K_\rho}{I_\rho}\beta_\rho = -\dot p$$

where: Q = value of the point mass

1 = distance of the point mass from the satellite center of mass

Assume p, q and r very small; therefore neglect their products. In such a case:

$$p \simeq \dot\varphi \qquad q \simeq \dot\theta \qquad r \simeq \dot\psi \tag{3}$$

Moreover, if:

$$\zeta = \alpha_y - \beta_y \quad ; \quad \eta = -\alpha_z + \beta_z \quad ; \quad \xi = \alpha_\rho + \beta_\rho \tag{4}$$

the dynamics is described by the decoupled set of equations:

$$\begin{cases} I_x\ddot\varphi + I_\rho\ddot\xi = M_x \\ I_\rho\ddot\xi + C_\rho\dot\xi + K_\rho\xi = -2\ddot\varphi I_\rho \end{cases} \qquad \begin{cases} I_y\ddot\vartheta + Q\ell\ddot\eta = M_y \\ Q\ddot\eta + C_z\dot\eta + K_z\eta = -2\ddot\vartheta Q\ell \end{cases} \tag{5}$$

$$\begin{cases} I_z \ddot{\psi} + Q\ell \ddot{\zeta} = M_z \\ Q\ddot{\zeta} + C_v \dot{\zeta} + K_v \zeta = -2\ddot{\psi} Q\ell \end{cases}$$

It can be concluded that the dynamics around each bo[illegible] [illegible]es is described by a set of equations of the type:

$$A\ddot{\sigma} + B\ddot{\gamma} = M \qquad ; \quad C\ddot{\gamma} + D\dot{\gamma} + E\gamma = -F\ddot{\sigma} \tag{6}$$

Where σ is the angular displacement of the main body with respect to an inertial reference, γ the one of the flexible arrays with respect to the main body, A ...... F suitable coefficients, and M the control torques applied to the main body. If we let:

$$x_1 = \sigma \;;\quad x_2 = \gamma \;;\quad x_3 = \dot{\sigma} \;;\quad x_4 = \dot{\gamma} \tag{7}$$

(6) take the form:

$$\begin{aligned} \dot{x}_1 &= x_3 \;; \qquad x_3 = A_1 x_1 + A_2 x_2 + A_3 M \\ \dot{x}_2 &= x_4 \;; \qquad x_4 = -A_4 x_1 - A_5 x_2 - A_6 M \end{aligned} \tag{8}$$

Where A_1 ....A_6 are suitable coefficients.

It may be verified, by means of a continuous controllability test (see ref. 11), that this linearized system is controllable.
In particular, this means that it is possible to find a time history for M (the torque applied to the main body of the satellite), which is able, starting from arbitrary values of displacements and vibrations, to force the system to an established position in space, with zero rate and with vibrations completely damped out.

3. VALIDATION OF THE MODEL

Following Liking & Fleisher (see ref. 6), modeling very accurately, i.e. with a whatever large number of masses and springs, the flexible appendages of a satellite, the dynamics of the system for each axis can be described by the block diagram shown in fig. 4, where:

I = inertia of the rigid part (main body)

T = applied torque

α = attitude angle

σ_j = natural frequency of the j-th mode

ρ_j = damping factor of the j-th mode

S_j = weighting factor of the j-th mode

For reasonable values of the parameters, it can be shown (see ref. 1) that S_1 is much greater than S_2, S_3 ....
It is then generally possible to neglect in fig. 4 all the feedback blocks except the one which represents the first vibration mode.
If this is done, the transfer function for one axis becomes:

$$\frac{\alpha(s)}{T(s)} = \frac{s^2 + 2\rho_j \sigma_1 s + \sigma_1^2}{I_s^2 \left[s^2 \left(1 - \frac{S_1^2}{I}\right) + 2\rho_1 \sigma_1 s + \sigma_1^2 \right]}$$

formally identical to system (8).

It must be pointed out that the assumption of a damping force proportional to the velocity is not completely correct. However it is possible to choose ρ in such an optimum way as to make the effects of an histheretic damping equivalent to those with a damping proportional to the velocity.
This is left for future investigations.

4. CONTROL WITH MOMENTUM EXCHANGE DEVICES

When an actuator is used which provides a momentum H, with respect to the body axes, the forces it exerts on the satellite are (see ref. 2a and 12):

$$M_x = -\dot H_x + rH_y - qH_z \;; \quad M_y = -\dot H_y + pH_z - rH_x \;; \quad M_z = -\dot H_z + qH_x + pH_y \tag{9}$$

where H_x, H_y and H_z are the components of H along the body axes.

Usually the actuators configuration employed provides a momentum which is quite larger along one axis (say x) than along the other ones; as p, q and r are considered small, it is possible to rewrite (9) as follows:

$$M_x = -\dot H_x \;; \quad M_y = -\dot H_y - rH_o \simeq -\dot H_y - \dot\psi H_o \;; \quad M_z = -\dot H_z + qH_o \simeq -\dot H_z + \dot\vartheta H_o \tag{10}$$

where H_o is the nominal value of H_x, the largest component of H.
If (10) is substituted into (5), the set of equations describing the flexible satellite is the following:

$$\begin{cases} I_x\ddot\varphi + I_\rho\ddot\xi = -\dot H_x \\ I_\rho\ddot\xi + C_\rho\dot\xi + K_\rho\xi + 2\ddot\varphi I_\rho = 0 \end{cases} \tag{11}$$

$$\begin{aligned} &I_y\ddot\vartheta + H_o\dot\psi + Q\ell\ddot\eta = -\dot H_y \qquad && Q\ddot\eta + C_z\dot\eta + K_z\eta + 2\ddot\vartheta Q\ell = 0 \\ &I_z\ddot\psi - H_o\dot\vartheta + Q\ell\ddot\zeta = -\dot H_z \qquad && Q\ddot\zeta + C_y\dot\zeta + K_y\zeta + 2\ddot\psi Q\ell = 0 \end{aligned} \tag{12}$$

In such a way the system breaks into two linear ones, the former describing only the behaviour around the x axis, and the latter describing the behaviour of the coupled y and z axes.
Using Laplace transform, from (11) and (12):

$$\varphi = -\frac{1}{sF_x}H_x \tag{13}$$

$$\vartheta = \frac{-sF_zH_y + H_oH_z}{s^2F_yF_z + H_o^2} \qquad \psi = \frac{-H_oH_y - sF_yF_z}{s^2F_yF_z + H_o^2} \tag{14}$$

where:

$$F_x = I_x \qquad F_y = I_y \qquad F_z = I_z \tag{15}$$

if the satellite is considered rigid, and:

$$\begin{aligned} F_x &= \frac{s^2(I_x - 2I_\rho) + s\frac{I_x}{I_\rho}C_\rho + \frac{I_x}{I_\rho}K_\rho}{s^2 + s\frac{C_\rho}{I_\rho} + \frac{K_\rho}{I_\rho}} \\ F_y &= \frac{s^2(I_y - 2Q\ell^2) + s\frac{I_y}{Q}C_z + \frac{I_y}{Q}K_z}{s^2 + s\frac{C_z}{Q} + \frac{K_z}{Q}} \\ F_z &= \frac{s^2(I_z - 2Q\ell^2) + s\frac{I_z}{Q}C_y + \frac{I_z}{Q}K_y}{s^2 + s\frac{C_y}{Q} + \frac{K_y}{Q}} \end{aligned} \tag{16}$$

if the satellite is considered flexible, as it actually is.

As the problem of designing a control loop for the decoupled x axes, described by (13), is quite simple and classical, it will not be considered here. We will design a multiloop controller for system (14), describing the coupled motion of the y and z axis.

5. MULTILOOP CONTROLLER DESIGN

The synthesis method described in ref. 14 is used to design a multiloop controller for system (14). Such a method is referred to, in the literature, as "decoupling via feedback", as it enables to control independently the state variables. In our case it means that a command on ϑ does not affect ψ and viceversa.

We will not attempt to describe the method here, and give only the synthesis result, shown in fig. 5. A_y and A_z are the resulting open loop transfer functions for the y and z axis respectively; to have a suitable closed loop response they have been chosen:

$$A_y = A_z = \frac{\omega_n^2}{s^2 + 2\delta_n \omega_n s} \tag{17}$$

with:

$$\delta_n = 0.7 \qquad \omega_n = 0.4 \tag{18}$$

F_x, F_y and F_z are represented by (15) if the satellite is rigid, and by (16) if it is flexible.

6. SIMULATION RESULTS

A simulation program, using FORTRAN V, has been written, using the following set of reasonable numerical values (see ref. 2):

$I_y = 1500$

$I_z = 1800$

$\frac{C_y}{Q} \frac{C_z}{Q} = 2.5 \quad 10^{-2}$

$\frac{K_y}{Q} \frac{K_z}{Q} = 4 \quad 10^{-2}$

$l = 4$

$Q = 7$

$H_0 = 100$

To show the effects of flexibility on a controller designed considering rigid the solar arrays, two types of simulations have been made:

a) using (15) in the controller and (16) in the satellite dynamics (referred to heretofore as "rigid controller" case)

b) using (16) both in the controller and the satellite dynamics (referred to heretofore as "flexible controller" case)

In figures 6 and 7 an acquisition is shown (step input with ϑ_R =1.0 and ψ_R = 0.8), for the flexible and rigid controller case respectively; with the flexible controller the response is exactly what is expected and the satellite is pointed correctly, while with the rigid controller persistent oscillations remain after the acquisition.
In figures 8 and 9 the satellite behaviour during normal mode operations is shown. ϑ_R and ψ_R are kept to zero, and a constant disturbance torque equal to 10^{-4} Nw . m is supposed applied to the satellite y axes. Fig. 8 shows the satellite behaviour in the flexible controller case, ϑ and ψ exhibiting steady state oscillations, together with the time history of H_y and H_z. The pointing accuracy is lower than $2.8 \cdot 10^{-4}$ degrees.

Figure 9 shows instability for the rigid controller case.

7. CONCLUSIONS

In the simulation I_y and I_z are higher than $2Ql^2$; the opposite case mak · a zero with positive real part appear in some fo the controller transfer functions; in such a case the scheme used in this paper is no more directly applicable, but, using reference 14, it shoud be possible to modify it in order to achieve the same results.
Actually, in fig. 5, H_y and H_z are electrical signals and not momentums; thus, when designing the real attitude control system, it will be necessary to draw two inner control loops to include the actuators, as shown in fig. 10. The transfer function characteristics of these inner control loops, as well as those of the sensors used, can be considered in the design of the controller; this would not be very easy to do using already developed methods, like WHECON or GRASP (see ref. 15 and 16),

which apply, moreover, only to the double gimballed flywheel actuator configuration.
The results presented in this paper should point out two basic facts:

a) To take into account flexibility is in the majority of cases much easy, as very complicated models are not necessary.

b) When designing an attitude control system with momentum exchange actuators for a flexible spacecraft, flexibility should be taken into account from the beginning, as an attitude control system designed for a rigid spacecraft could result unstable or very degraded from the pointing accuracy point of view.

ACKNOWLEDGEMENTS

The author's are gratefull to C.E.A. Direction for giving permission of publication, as well as to Dr. Franco Gnavi who made it possible to make the work both from the operational and background competences point of view.
Acknowledgements are due also to Mr. W. Carpegna for the outstanding execution of the simulations.

REFERENCES

1. Gamba M. & others, November 1971 : "A.C.S. of the European TV Broadcasting Satellite Study", a study carried out by FIAT Centro Elettronico Avio for ESRO.

2. Di Lorenzo R. & others, 1971: " A Study of the Control & Stabilization Aspects of an Advanced Telecommunications Satellite with Flexible Antennae and Solar Arrays", to be published in ESRO's Contractors Reports series.

3. Deat J.E., July 1970 : "Dynamical Equations of Nonrigid Satellites", AIAA Journal, p. 1344

4. Rossi L.C., Michelini R.C. &Ghigl-azza R., August 1969 : "Attitude Dynamics & Stability conditions of a Non-Rigid Spinning Satellite", The Aeronautical Quarterly, p. 223.

5. Karnopp D., November 1967 : "Optimal Control of Maneuver - Induced Vibration in Flexible Aerospace Vehicles", AIAA Journal, p. 2017.

6. Likins P.W. & Fleisher G.E., March 1971: "Results of Flexible Spacecraft Attitude Control Studies Utilizing Hybrid Coordinates", J. Spacecraft, p. 264

7. Nelson H.D. & Meirovitch L., December 1966 : "Stability of a Nonsymmetrical Satellite with Elastically Connected Moving Parts", the Journal of SAstronomical Sciences.

8. Meirovitch L., July 1970 : "Stability of a Spinning Body Containing Elastic Parts via Liapunov's Direct Method", AIAA Journal.

9. Ojalvo J.U. & Newman M., July 1970 :"Vibration Modes of Large Structures by an Automatic Matrix Reduction Method", AIAA KJournal, p. 1234

10. Lagrange J.L., 1953 : "Mécanique Analytique", 3ème Edition, Vol. II, p. 212, Mallet Bachelier.

11. Zadeh L.A. & Desoer C.A., 1963 : "Linear system theory", Mc-Graw-Hill

12. Greensite A.L. : "Attitude Control in Space", NASA CR 831

13. Sévely Y., 1969 : "Systèmes et asservissements linéeres échantillonés", Dunod.

14. Quazza G., 1965 : "Sistemi a più variabili regolate" in "Problemi attuali di teoria dei controlli automatici", edited by C.N.R., Rome.

15. Lyons M.G., Lebsock K.L., Scott E.D., August 1971 : "Double Gimballed Reaction Wheel Attitude Control System for High Altitude Communications Satellites". AIAA Paper 71-949.

16. Mork H.L., August 1971 : "Synthesis and Design of a Gimballed Reaction Wheel Attitude Stabilization Package (GRASP). AIAA Paper 71-950.

17. Fearnsides I.I., July 1970 : "Use of Control Moment Gyros for the Stabilization of a Spinning Satellite", AIAA Journal, Vol. 8, No. 7, p. 1365.

18. Gilbert E.G., February 1969 : "The decoupling of multivariable systems by state variable feedback", SIAM J. Control, vol. 7,

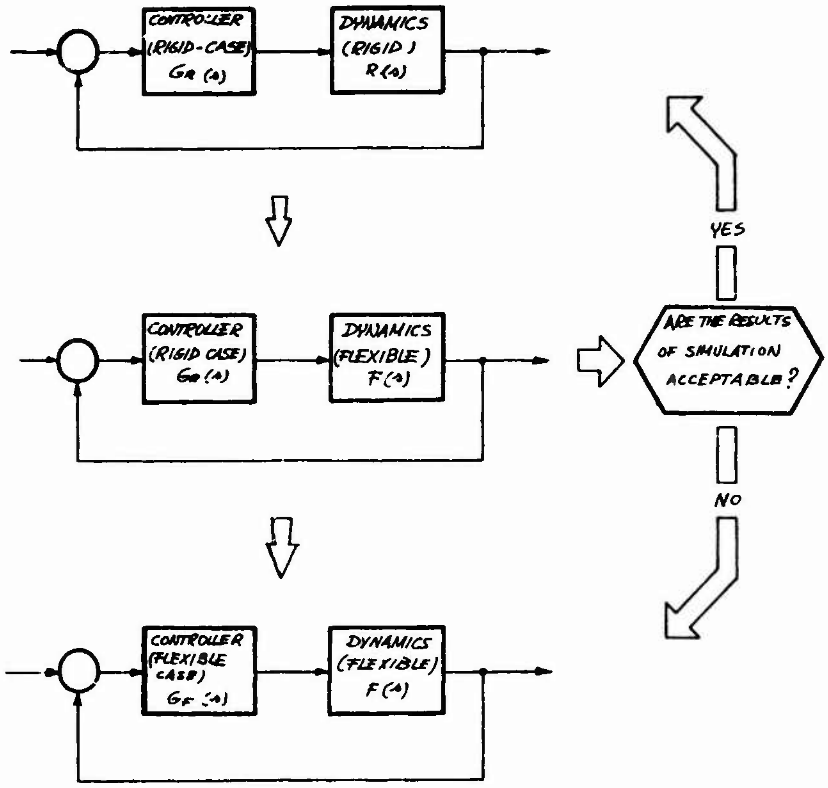


Fig.1 Procedure used to design an A.C.S. for a satellite with flexible appendages

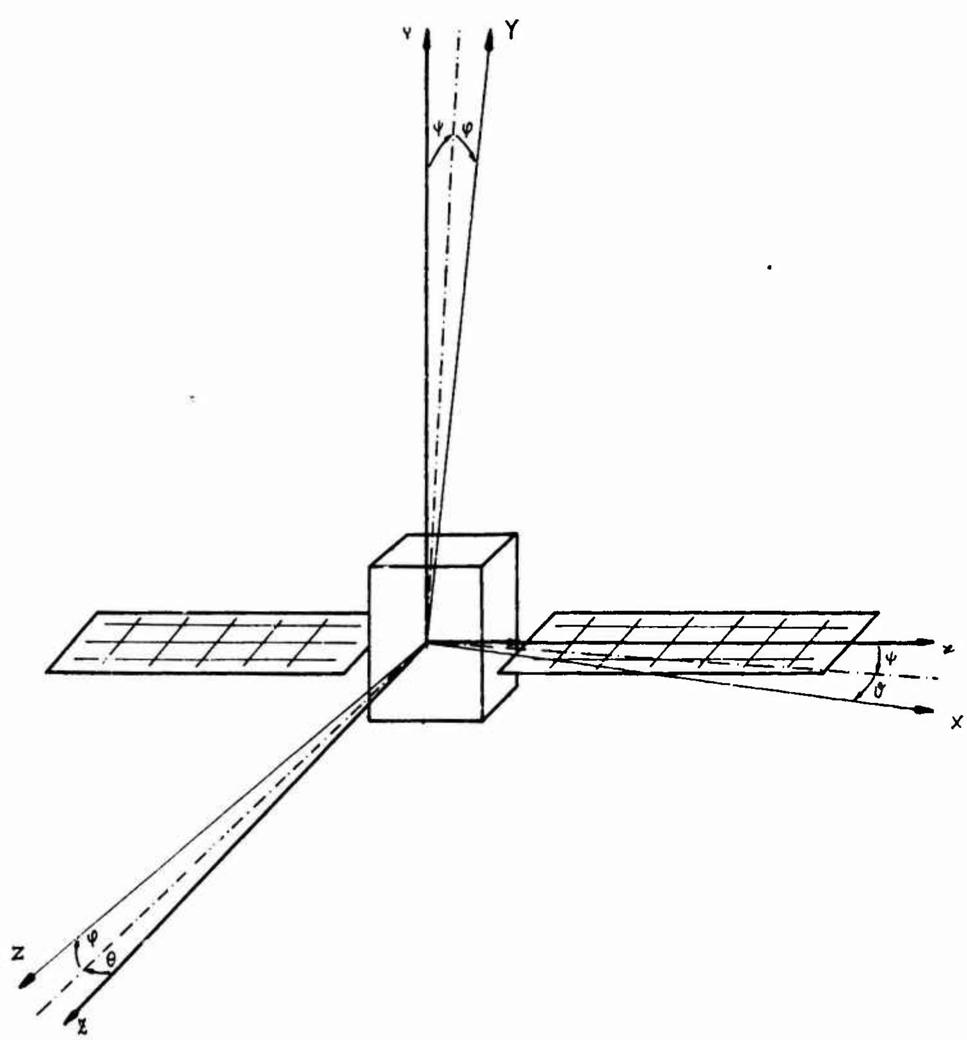


Fig.2 Satellite configuration

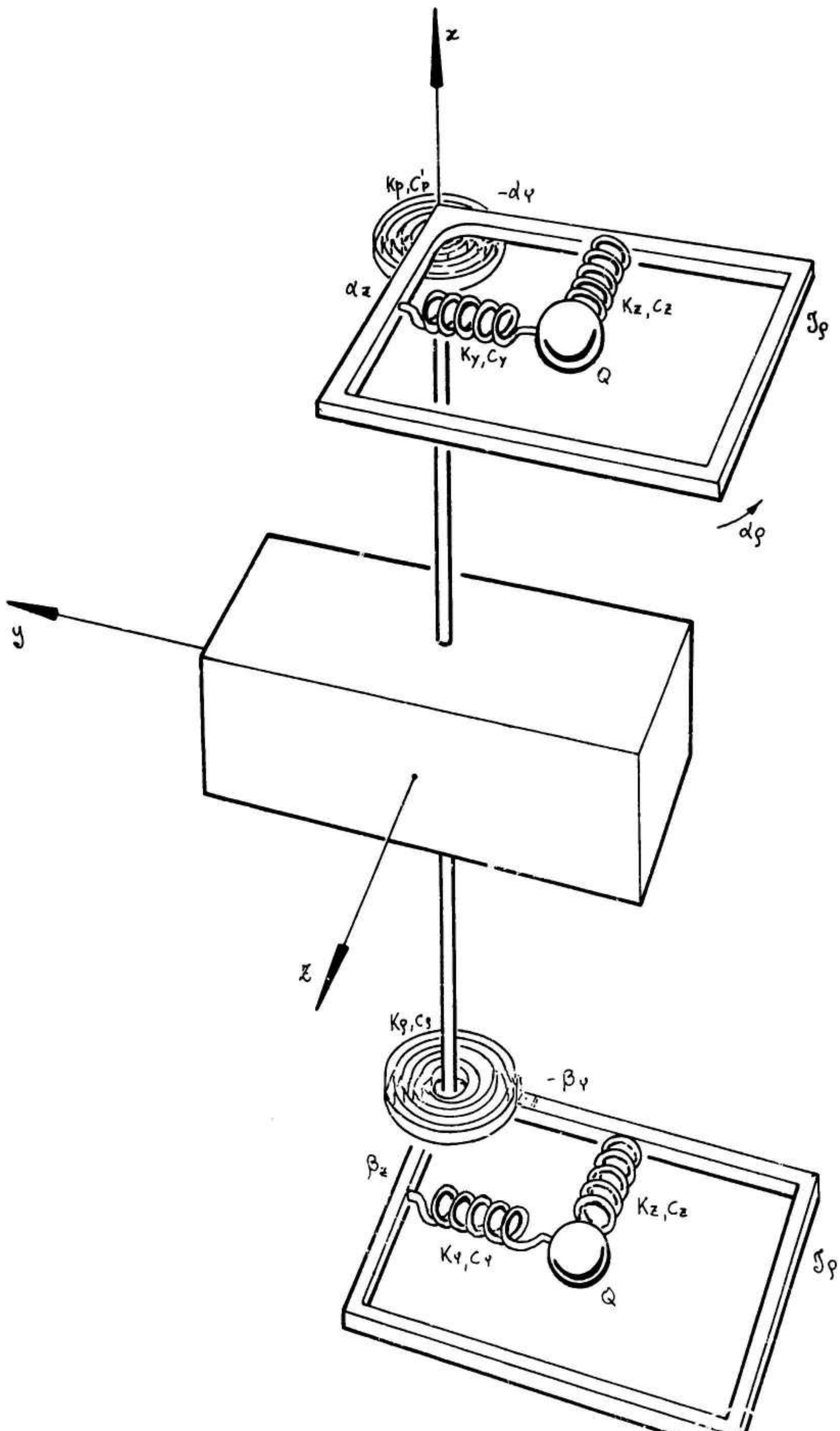


Fig.3 Dynamic model of flexible array

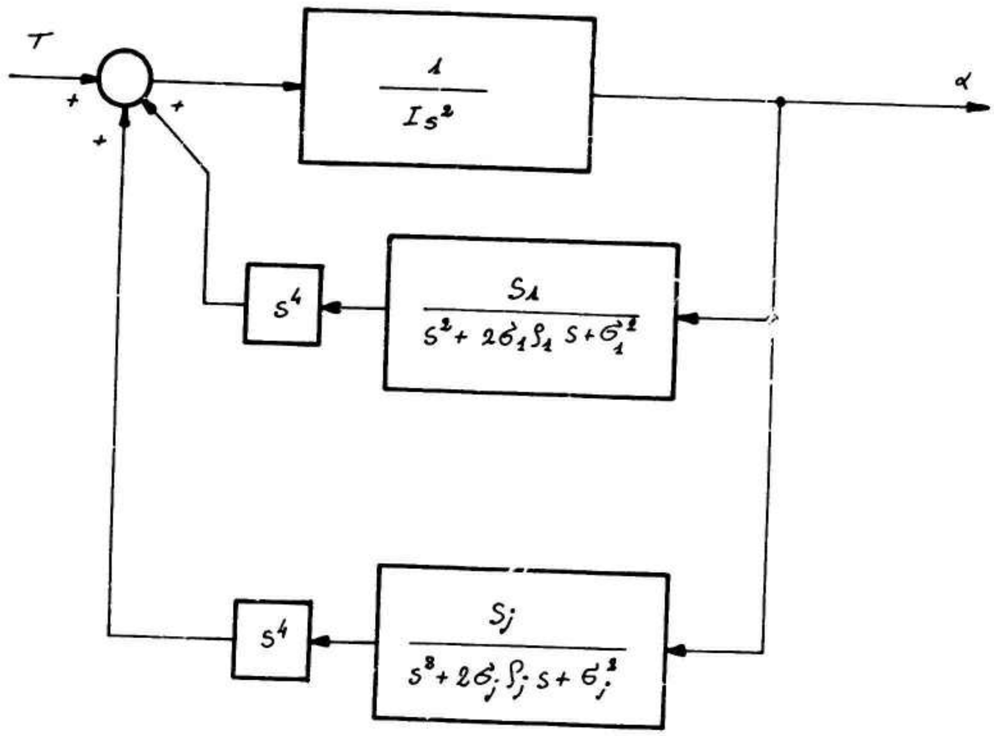


Fig.4 Likins and Fleisher model of a flexible satellite single axis

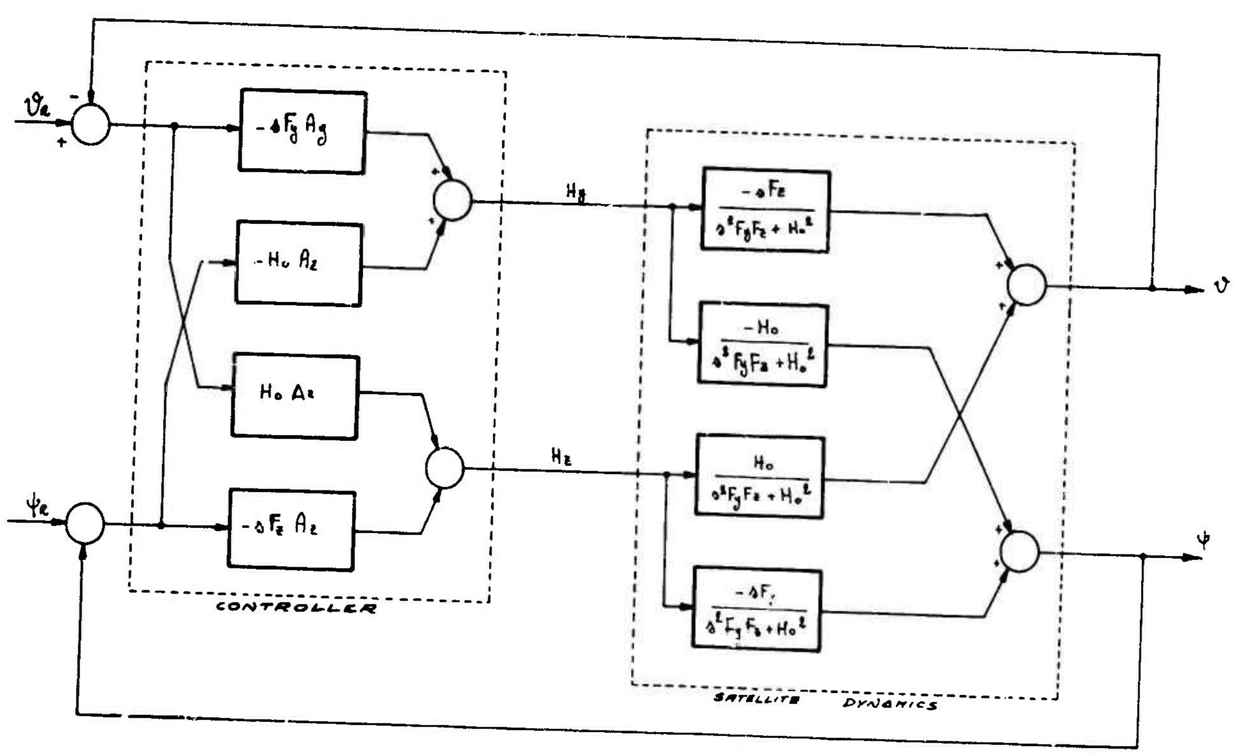


Fig.5 Multiloop controller synthesis for coupled Y and Z axis

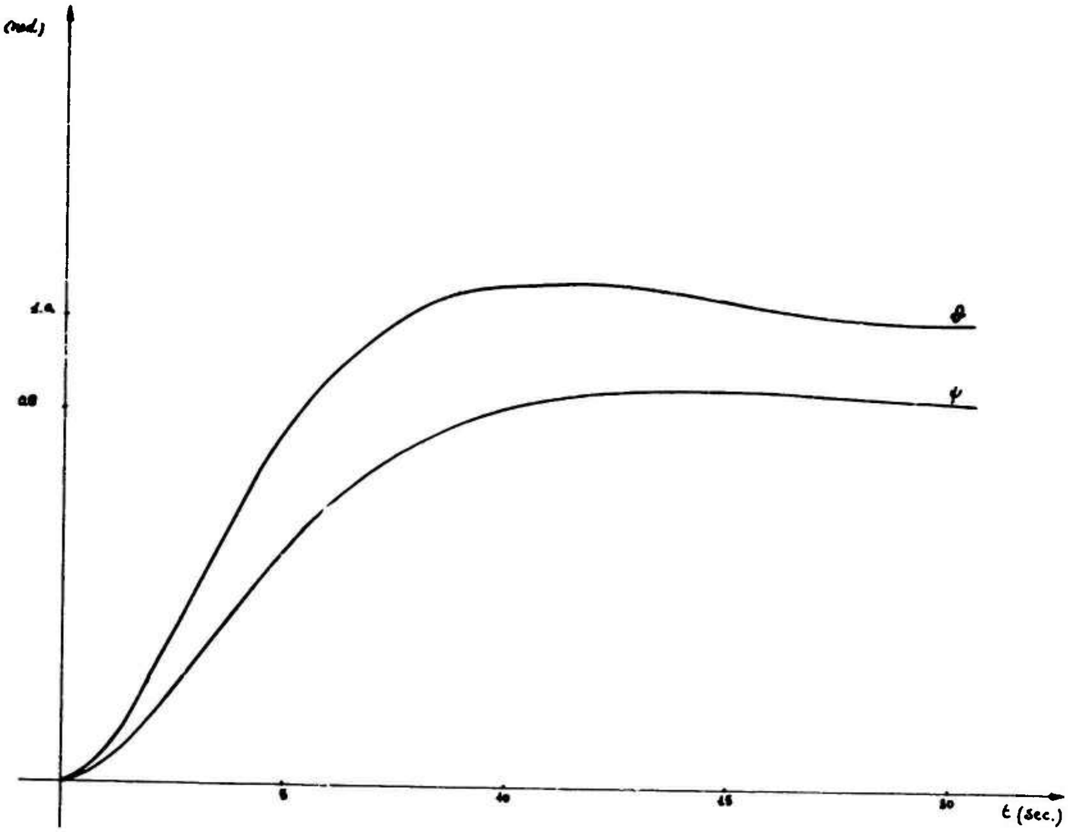


Fig.6 Acquisition with flexible controller

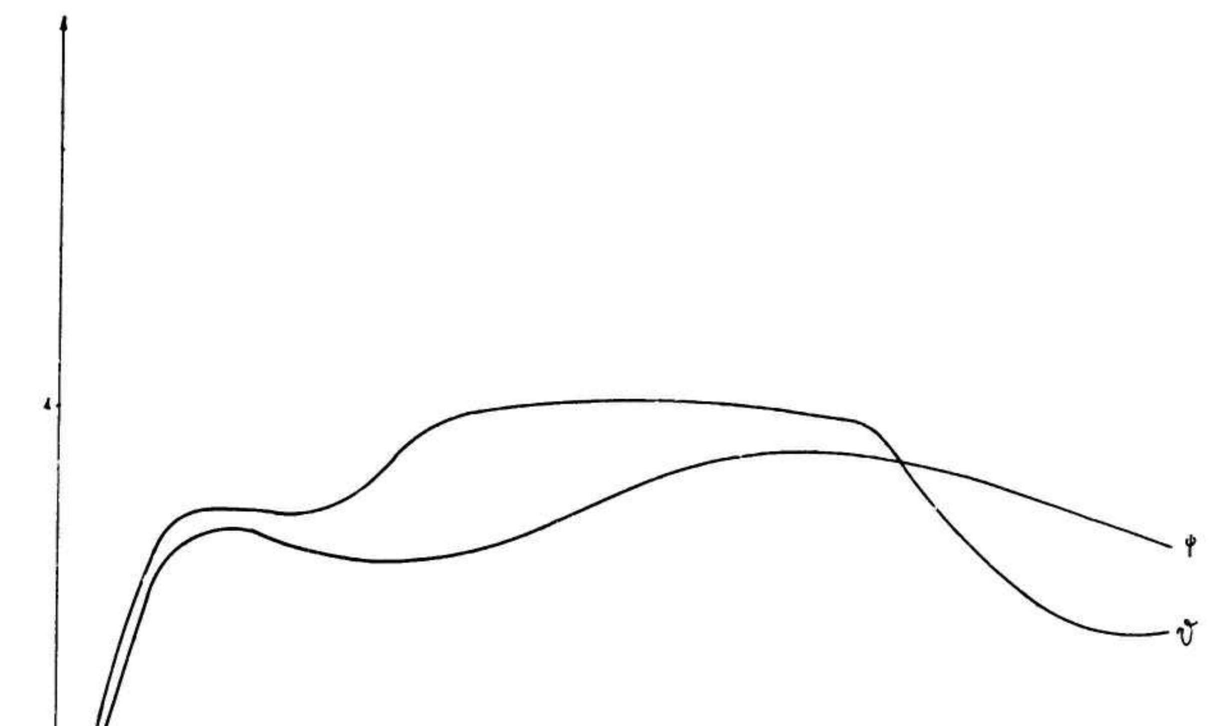


Fig.7 Acquisition with rigid controller

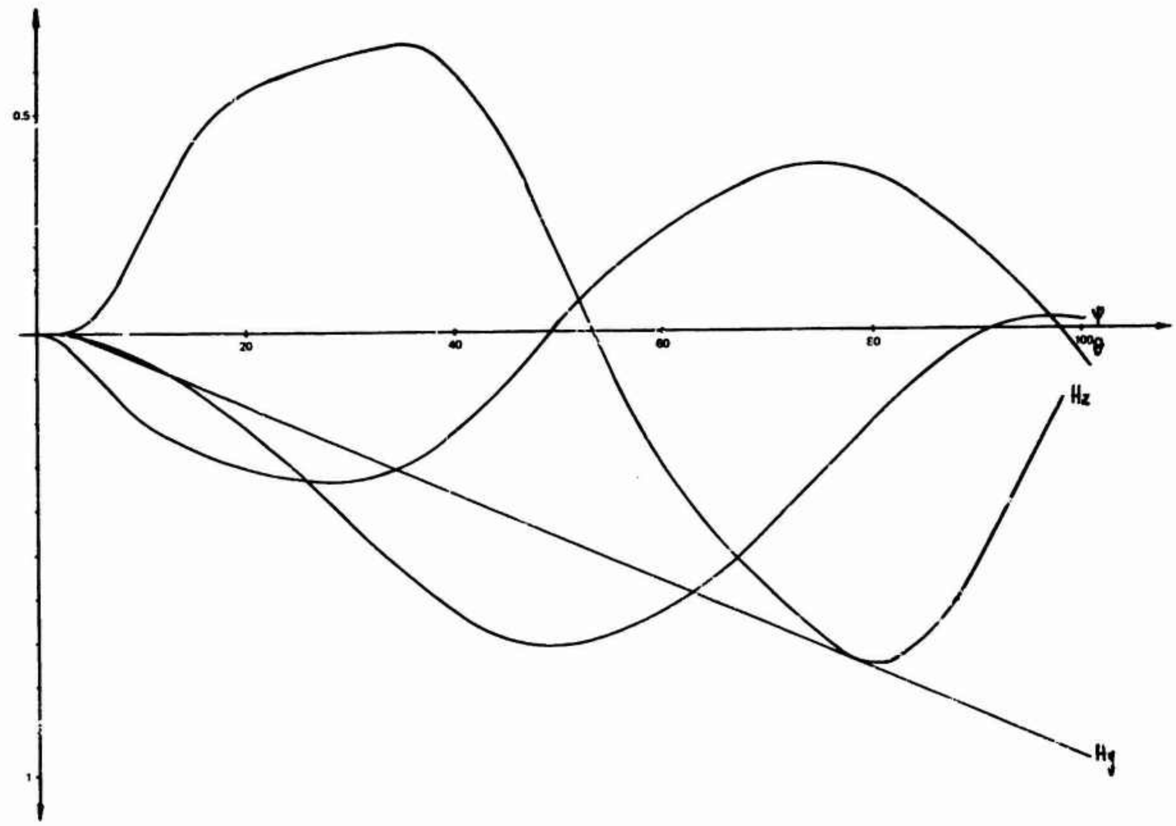


Fig.8 Normal mode operation with flexible controller

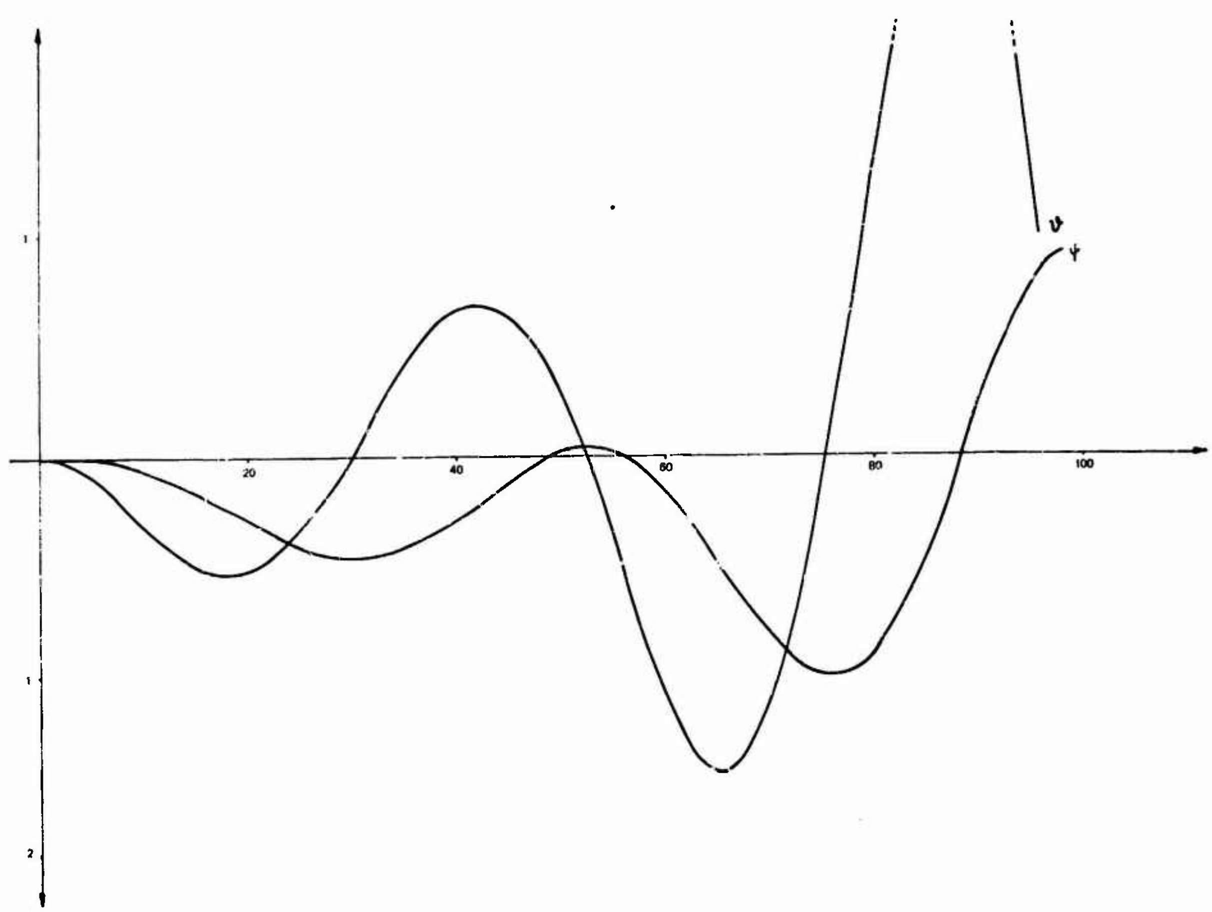


Fig.9 Normal mode operation with rigid controller

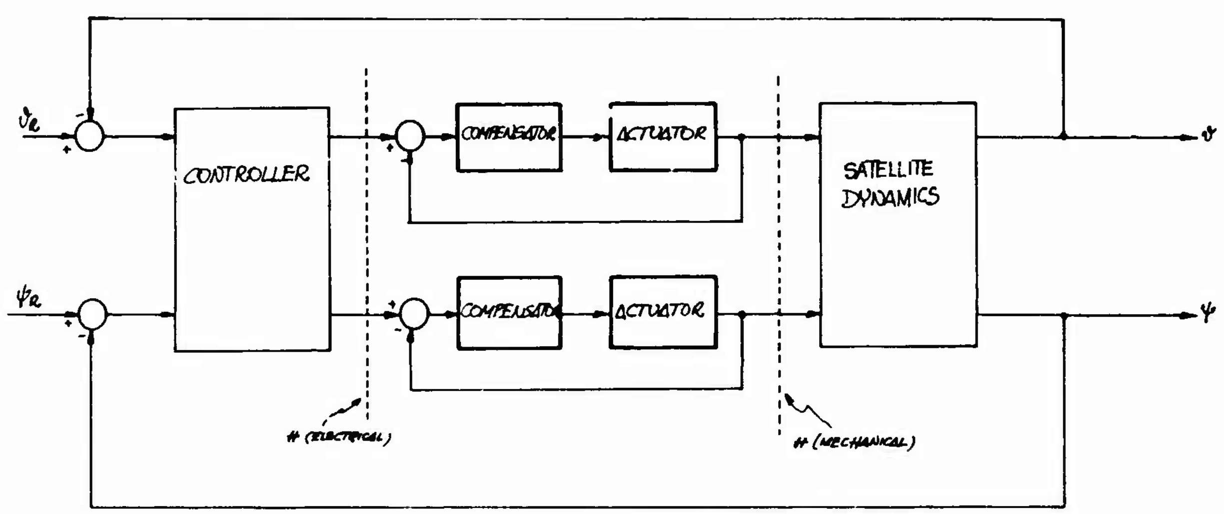


Fig.10 How the actuators can be introduced

OPTIMUM SPACEBORNE COMPUTER SYSTEM DESIGN

BY SIMULATION

E. Williams
Computer Sciences Corporation
8300 Whitesburg Drive, S.
Huntsville, Alabama USA

H. Kerner
Marshall Space Flight Center
Huntsville, Alabama

J.E. Weatherbee
Computer Sciences Corporation

D.S. Taylor
Computer Sciences Corporation

B. Hodges
Marshall Space Flight Center

SUMMARY

A deterministic simulator is described which models the Automatically Reconfigurable Modular Multiprocessor System (ARMMS), a candidate computer system for future manned and unmanned space missions. Its use as a tool to study and determine the minimum computer system configuration necessary to satisfy the on-board computational requirements of a typical mission is presented.

The effectiveness of a simulation model of a computer system as a design tool depends on the accuracy and fidelity of the definition of the specific data processing load. Hence the simulation techniques are described in relation to a specific spacecraft performing a specific mission, namely the proposed Reusable Shuttle Booster (RSB) stage.

While ARMMS is not a candidate for the RSB computer, this application was chosen because of the availability of a detailed description of the RSB data processing workload in which the computational requirements of the various RSB subsystems have been identified by mission phase and are used as parametric inputs to the deterministic model of the complete data processing system.

The paper describes how the computer system configuration is determined in order to satisfy the data processing demand of the various Shuttle Booster subsystems. The configuration which is developed as a result of studies with the simulator is optimal with respect to the efficient use of computer system resources.

1. INTRODUCTION

The advent of third generation computing concepts has created problems in the selection or design of a computer to process a given workload. For example, multiprogramming enables the computer system to process more than one task concurrently, allowing more efficient utilization of system resources such as central processing unit and input-output subsystems; this concept, however, has made the selection or design of a computer system no longer the relatively simple task it once was.

A great amount of effort has been expended on this problem of matching a modern computer system to its data processing workload resulting in many different approaches such as simulation, mathematical modeling, etc. (NIELSEN, N.R., 1967), (GAVER, D.P., 1967). In order to accomplish such an optimized match, however, detailed knowledge of both the proposed system and workload is required.

This paper presents an approach to the optimization of the Automatically Reconfigurable Modular Multiprocessor System (ARMMS) to a well-defined data processing workload. ARMMS is presently being designed by Astrionics Laboratory of the Marshall Space Flight Center, Huntsville, Alabama, and is addressing the anticipated requirements for both higher computing capacity and reliability which may characterize spaceborne computers in the late 1970's to mid-1980's. ARMMS is intended to achieve both of these objectives through a highly modular computer architecture which can be configured as a multiprocessor for maximum computing speed or as a triple modular redundant (TMR) system with standby spares for extremely high reliability. Moreover, the configuration will be dynamic in that it will be possible to change the configuration in real time as needed by various mission phases or events.

The workload used in this study was that for the Reusable Shuttle Booster (RSB) Stage. While it is recognized that ARMMS is not a candidate for the RSB computer, this application was chosen because of the availability of a detailed description (Univac Report, 1971) of the RSB data processing workload. This combination of the ARMMS system description and the RSB workload description is used to illustrate the optimization procedure through a method of digital simulation.

The RSB data processing workload is typical of most aerospace applications in that it consists of a number of repetitive tasks which iterate at various rates. In the RSB application, however, many of these tasks must operate in a high reliability, i.e., TMR, mode. Hence, the situation exists where such a task will use three CPU's for a few milliseconds and then release them for use by other tasks which may be required to operate in the simplex mode, i.e., single CPU, for a further few milliseconds. These tasks will update system data elements, which in turn will be used by other executing program modules.

In an environment of such complexity, simulation of the system is imperative in order to determine the values of system resources such as number of CPU's, width of data busses, number of input/output channels, etc.

The remainder of the paper describes such a simulation model together with the results of a series of simulation runs in which various parameter values were established for the ARMMS system in its hypothetical role as the on-board computer for the Reusable Shuttle Booster.

2. SYSTEM DESCRIPTION

2.1. ARMMS Baseline.

The basic ARMMS configuration is shown in Figure 1. It consists of a Central Processing Element (CPE) composed of a number of central processing units (CPU's) which execute programs contained in a random access memory (RAM) backed up by bulk storage. External subsystems communicate with the computer via a number of Input/Output Elements (IOE's). Various data busses interconnect the system modules, as shown in Figure 1.

The system operates under the control of a dedicated executive module, the Block Organizer and System Scheduler (BOSS).

The number of CPU's, size of RAM, number of input/output elements and the widths of the various data busses will be dictated by the data processing and reliability requirements of each particular mission.

2.2. Workload Description.

A detailed data processing workload analysis is essential in order to perform a realistic design of the associated data processing system. Such an analysis has been performed for the Reusable Shuttle Booster (Univac Report, 1971) and was used to define the requirements of an on-board computing system. This workload description consists of the identification of a number of discrete mission time-phases during which specific program modules are executed and interact with specific data elements. Each program module is defined in terms of the amount of memory required, the number and type of instructions executed, the input and output data rates and the program iteration rate. The program modules interact with a number of data elements, each of which is defined by its size, the update iteration rate, and function, i.e., whether it is a source of or a destination for data associated with the program modules and external subsystems.

<u>Shuttle Booster Mission Phases</u>

The various phases of the Shuttle Booster Stage have been identified and are shown in Figure 2. Note that this timeline refers to a reusable booster stage which is intended to re-enter the earth's atmosphere after each launch, then cruise and land like a conventional aircraft.

<u>Program Module Description</u>

A sample program module description is shown in Figure 3. The program module described is called BIGP (Boost Iterative Guidance Program) and is defined in terms of its total number of instructions, the number of long and short instructions executed in a normal iteration, the required data space, the number of times the program executes per second, the reliability requirement, and the mission phases during which the program executes.

This description allows an exact allocation of memory space to be made to each module, consisting of an instruction area and a data area. Also, knowing the CPU execution speed, the amount of CPU time per iteration may be found directly from the total number of instructions executed per iteration.

The reliability requirement determines whether one or three CPU's are required each time the program executes.

<u>Data Element Description</u>

A data element is defined by the number of times it is used per second, its size in bits, its sources of updated data and the program elements and/or subsystems which use its data.

A number of such data elements are described in Figure 4: for example, ABEFU and BPAYAC are both single source/single destination, ABEFU being updated by the Navigational Subsystem NAV and used by program module SDSU, while BPAYAC is updated by program module BIGP and is used by the Flight Control Subsystem FCS. Other examples are shown in Figure 4 for single source/multiple destination, multiple source/single destination and multiple source/multiple destination elements.

<u>Data Base Description</u>

The above information concerning program modules and data elements is combined into a single data base entry for each executing program. A sample entry for BIGP is shown in Figure 5.

3. SIMULATOR DESCRIPTION

3.1. General Description

The simulator to be described allows the interaction of the various program modules, data elements and external subsystems to be analyzed. A portion of this complex interaction is shown in Figure 6, where the basic ARMMS system of Figure 1 has had superimposed upon it some of the interacting program modules and data elements described above. Figure 6 shows the program module BIGP in execution at a rate of 1344 short and 376 long instructions two times per second. BIGP requires 707 words of instruction space and 97 words of data space. It requires updated information from data elements INEVAP and TPINCT which are themselves updated at rates of 32 and 2 per second respectively. BIGP supplies information to data element BPAYAC, which in turn is used by the Flight Control Subsystem (FCS).

Figure 6 represents a snapshot of an instant in time; when BIGP finishes a particular iteration it releases its CPUs for use by another program module(s). Should any program module fail to iterate at its required rate, due say to priority conflicts with other modules, the simulator will flag the event as real time violation. It is the object of the simulation to have all program modules and data elements executing at their predefined rates within a minimum system under the constraint of no real time violations.

3.2. System Simulator

The system simulator is a program which implements the functions described in Figure 6. Its flowchart is shown in Figure 7.

Each program has a pre-assigned priority based upon its iteration rate and reliability requirement. The priority assigned to each program is given in Table 1, where a high figure implies a high priority; the letters T and S refer to the TMR and simplex modes of operation, e.g., 16/T describes a program which iterates at a rate of 16 times per second in the TMR mode.

Itn. Rate/Mode	64/T	64/S	32/T	32/S	16/T	16/S	8/T	8/S	4/T	4/S	2/T	2/S	1/T	1/S
Priority	14	13	12	11	10	9	8	7	6	5	4	3	2	1

Table 1. Program Priority Scheme

From the iteration rates shown, it is obvious that each program has a deadline time which must be met in order to avoid a real-time violation. Hence for no real-time violations, programs having priorities 13 or 14 must execute every 1/64 sec., those with priorities 11 or 12 every 1/32 sec., etc.

When a program's deadline time arrives, a check is made to see if the previous iteration of that program is complete: if it is, the program is entered in the CPU queue; if not, a real-time violation (RTV) flag is set. When a program completes execution, it is re-entered into the CPU queue.

Programs are selected from the CPU queue in order of their priority, higher priority programs having the ability to preempt executing programs of lower priority. Execution takes place for a time determined by the number of long and short instructions to be executed per iteration. This execution allows data elements and/or subsystems to be supplied with updated data.

After each execution iteration, the processor(s) are released for use by other programs of lower priority which are currently resident in the CPU queue.

3.3. Interface Simulator

As was stated in Section 3.2., the duration of each iteration cycle of each executing program module is determined by the number of short and long instructions executed per iteration. Examples of short instructions are ADD, SHIFT, JUMP; while MULTIPLY, DIVIDE are typical long instruction types. The execution times for each instruction is specified for the ARMMS CPE and hence, by assuming a Gibson mix (SMITH, J.M., 1968) for all programs, an average execution time may be determined for both long and short instructions. Referring to Figure 1, the basic ARMMS configuration, which represents a system having a number of CPU's interacting with a random access memory consisting of a number of modules, a situation can arise where more than one CPU will try to access a given memory module. When this occurs, interference takes place with a resultant delay in the execution of one of the conflicting programs. In an environment such as this, therefore, it is no longer valid to determine the execution time of any given program on the basis of average instruction execution time alone.

In order to take the effect of memory interference into account, it was necessary to simulate the interference at the CPE/RAM interface. This was done by means of the Interface Simulator.

The flowchart of Figure 8 illustrates the function of this simulator: it is initialized by specifying all input parameters such as number of CPU's, number of memory modules, data bus widths, number of instructions fetched per access, etc.

The Interface Simulator operates under the assumption that there is always a non-zero set of instructions awaiting execution by each CPU; this represents a worst case condition. Note from Figure 8 that an instruction fetch and data fetch are not initiated for all instructions executed: for example, an instruction fetch is not always necessary if a multiple instruction fetch is incorporated in the simulated CPE; similarly, a data fetch is not required for every instruction executed, e.g., JUMP. Values of functions defining the fetch/no fetch ratios are supplied as inputs to the simulator. This fetch/no fetch ratio is a function of the system being simulated and the data fetch/no fetch ratio is a function of the instruction mix; in the present simulation, a Gibson mix was assumed.

The Interface Simulator computes the total execution time for both long and short instructions by simulating the execution of a large number of instructions. This number, which is supplied to the simulator as input, is presently set at 3000, giving a variance of less than 1 percent on the computed times.

4. SYSTEM SIMULATION

4.1 Introduction.

This section describes how the simulation models were used to determine values for those parameters necessary to specify the design of a minimum computer system capable of performing the data processing for all mission phases of the Reusable Shuttle Booster.

In the work to be described, certain system parameter values were fixed due to constraints other than data processing throughput; in addition, certain assumptions were made. These are:

(a) The maximum number of tasks which can process concurrently is three. This figure was derived from consideration of the load placed upon the executive module, BOSS.

(b) There are two one-way busses for communication from the CPE to RAM, and another two connecting RAM to the CPE. These busses are time shared by the three instruction streams assumed under (a); two busses were considered necessary for the purpose of reliability.

(c) The processing speed of the CPE is fixed: This arises due to the fact that ARMMS is an extension of the Space Ultrareliable Modular Computer (SUMC) which is presently under development at MSFC, and it is anticipated that a version of SUMC will be the processor module of the ARMMS system.

(d) The RAM cycle time is assumed to be 750 n.sec: This is available using present-day passive memory technology and is close to the memory speed proposed for ARMMS. Destructive readout is assumed.

(e) There is no triplication of stored data for the TMR operation: All memory transfers will be parity checked and the data fanned out into three CPU's where TMR is required.

(f) A single address instruction format is assumed; in general, one operand is fetched from RAM and the other from a bank of general registers located in each CPU.

(g) The bus transfer time is 100 n.sec.

(h) The time taken to assign a task to a processor or pre-empt an operating task is 100 μ.sec.

(i) The basic machine word length is 32 bits.

The simulation models were then used to determine:

(a) The effect of multiple instruction fetches per memory access;

(b) The number of CPU's;

(c) The number of RAM modules;

(d) The widths of the system busses.

The approach taken was to inspect the workloads for each mission phase and select the one which appeared to place greatest demands upon the computer system. Mission Phase 18 appeared to be the most stringent in its requirements and was selected as the workload on which the initial investigations were performed. This phase was characterized by the requirement that all tasks were required to operste in the TMR mode, constraining the system to operate with a multiple of three CPU's.

4.2. Determination of Number of CPU's

Since the actual average long and short instruction execution times are functions of the system configuration and the resulting memory/CPE interface conflicts, a set of runs was made in which the execution times were varied over a wide range and any real-time violations noted. Figure 9 shows the results of such a set of runs as simulated for three CPU's; it can be seen that, if the long instruction execution time exceeds 5 μsec for a 1 μsec short instruction execution time, or 3 μsec for a 1.5 μsec short instruction execution time, real time violations occur. The corresponding times for the SUMC processing unit are known to be of the order of 6 and 2 μsec, respectively, and, hence, a single TMR processor set is inadequate for the task of processing the Mission Phase 18 workload.

Figure 10 shows a similar chart for six CPU's, i.e., two TMR sets. It can be seen thst this configuration will readily process the Mission Phase 18 workload under the constraint of the SUMC processing times.

Note that these tests were carried out on a system having a single TMR input/output subsystem and a single bit CPE-IOE bus width. These parameter vslues were chosen after it was established that they imposed no constraints on the system due to the relatively light amount of I/O activity to and from the external subsystems.

Hence, it is concluded that the ARMMS system will require six CPU's in order to meet the data processing and reliability requirements of Mission Phase 18.

Having established the need for six CPU's, the Interface Simulator was then used to determine the effect upon instruction execution times of

(a) varying the number of RAM modules

(b) multiple instruction fetch

(c) variation of the width of the RAM/CPE busses

All these tests were performed with relation to Mission Phase 18 in that two instruction streams were simulated representing the two TMR sets required by that Mission Phase.

Also, an analysis of the proposed instruction set as mapped into a Gibson mix revealed that 16 percent of all instructions required no D-bank access, and this figure was used in all experiments performed.

The RAM/CPE busses were assumed to be 32 bits wide.

4.3 Determination of Number of Memory Modules

A set of tests was carried out on the Interface Simulator to determine the effect, upon instruction execution times, of varying the number of RAM modules assigned separately to the storage of instructions and data. Table 2 shows the instruction execution times obtained from these tests as a function of the number of I-bank and D-bank modules.

I-Bank Modules D-Bank Modules	4 8	8 16	16 32
SIT (μsec) LIT (μsec)	2.02 6.20	2.01 6.19	1.99 6.16

SIT = Short Instruction Time
LIT = Long Instruction Time

Table 2. Instruction Execution Time as a Function of Number of Memory Modules

It can be seen from Table 2 that the amount of memory interference is small when greater than 12 memory modules are used. The total memory requirement for Mission Phase 18 is approximately 32K of I-bank and 64K of D-bank; hence, Table 2 represents the effects of varying module size from 8K to 2K.

From the standpoint of reliability, it is unlikely that an 8K module would be used since it represents too large a "throwaway" module. Hence, the 4K module was chosen and used in all further studies to be described.

4.4 Multiple Instruction Fetch

The design contractor of the ARMMS system (Hughes Aircraft Company) has recommended the concept of a small local store within each CPU. A major function of such a local store would be to retain a small number of previously executed instructions so that, should a branch backwards within this bound occur, an instruction fetch from main memory is not required. After a detailed analysis of an extensive set of aerospace programs, Hughes state that a saving of 4 percent in the number of instructions accessed in main memory is achieved for an eight instruction retention capability. While this has a negligible effect upon instruction execution times, an eight words instruction retention capability will handle approximately 35 percent of all branch instructions (Hughes Report, 1972). The remaining 65 percent of the branch instructions are jump-ahead type, and were assumed to represent 10 percent of all other instructions.

These figures were used as the basis for a set of experiments to determine the effect of multiple instruction fetch upon instruction execution times. Table 3 shows the results of these experiments for 2, 4 and 8 instructions fetched per memory access.

Instructions per Fetch	1	2	4	8
SIT (μsec) LIT (μsec)	2.01 6.19	1.63 5.75	1.40 5.57	1.29 5.45

SIT = Short Instruction Time
LIT = Long Instruction Time

Table 3. Instruction Execution Time as a Function of Number of Instructions per Fetch

The greatest percentage improvement is observed in going from one to two instructions per fetch, and this improvement was used in the next set of experiments as a trade-off against reduction in RAM/CPE bus widths.

Note that a two-instruction fetch implies reading a double word from memory on each memory access.

4.5. Determination of RAM/CPE Bus Widths

From Figure 10 it can be seen that the 6 CPU configuration can meet all real-time requirements with execution times of up to 3 μsec for short instructions and up to 8 μsec for long instructions. Table 3 shows that the two instruction fetch produces execution times of 1.63 μsec and 5.75 μsec, respectively, for short and long instructions when using 32-bit-wide RAM/CPE busses. Instruction execution speed is traded off against RAM/CPE bus width in Table 4 which summarizes the results of a set of tests where the RAM/CPE bus widths were held equal but reduced from the 32-bit baseline value down to a width of 4 bits.

RAM/CPE Bus Widths (bits)	32	16	8	4
SIT (μsec) LIT (μsec)	1.63 5.75	1.91 6.03	2.48 6.62	3.63 7.68

SIT = Short Instruction Time
LIT = Long Instruction Time

Table 4. Instruction Execution Time as a Function of RAM/CPE Bus Widths

These results show that the RAM/CPE busses can be reduced to a width of 8 bits and still satisfy the real-time requirements of Mission Phase 18.

Hence, based upon the computational requirements of Mission Phase 18 together with the constraints stated in Section 4.1., the optimum configuration of the ARMMS system is

- 6 CPU's
- 8 x 4K RAM modules for instruction storage
- 16 x 4K RAM modules for data storage
- 2 instruction fetch per memory access
- 8 bit RAM/CPE busses
- single TMR Input/Output Element (IOE)
- single bit CPE/IOE bus

The above configuration was then used in the simulation of the data processing required by Mission Phase 9, the one assessed to be next most demanding upon data processing resources after Phase 18: Phase 9 is characterized by having mixed mode execution, i.e., both simplex and TMR.

The instruction execution times produced by the Interface Simulator were 2.53 μsec and 6.73 μsec, respectively, for short and long instructions. These times were derived under worst case conditions where a three instruction stream environment was assumed for the whole of Mission Phase 9. Since this Phase is mixed mode, two instruction stream operation will occur part of the time when two TMR programs are executing concurrently.

Using the above instruction execution times, all Phase 9 programs were simulated without producing any real-time violations.

The simulation of Mission Phase 9 was repeated with the same configuration, but this time under the assumption that all programs executed in the TMR mode. Instruction execution times of 2.48 and 6.62 μ sec, respectively, were used as derived for the case of two instruction streams. Again, no real-time violations occurred.

Checks on several other Mission Phases revealed no real-time violations when using the two TMR set configuration; hence, for the Reusable Shuttle Booster application, all programs should execute in the TMR mode, thereby relieving the Executive Module BOSS of the task of scheduling for a mixed mode operation.

5. CONCLUSION

A deterministic simulator has been described and its application as a design tool illustrated by the hypothetical example of defining a minimum Automatically Reconfigurable Modular Multiprocessor System (ARMMS) which can process the data pertaining to the Reusable Shuttle Booster mission.

It has proved possible to define the required ARMMS hardware in terms of

(a) number of CPU's

(b) number of RAM modules; and

(c) width of the system busses.

This definition was based upon a data processing load description which was broken down into a number of mission phases, each mission phase being defined in terms of the executing program modules and associated data elements. The program module description consisted of the total number of instructions, the number of long and short instructions executed per normal iteration, the required data space, the number of times the program executes per second, and the reliability requirements. Each data element was defined by the number of times it is used per second, its size, and its sources of updated dsta and the program elements and/or subsystems which use its data.

It is concluded that, provided the data processing requirements can be defined to the above level of detail, a minimum hardware configuration can be derived. Further, in an environment where high reliability is a requirement for some programs, this type of model can be used to determine the effect of executing all programs in a high-reliability mode, with the attendant advantage of relieving the Executive System of the task of mixed mode scheduling.

ACKNOWLEDGEMENT

The authors wish to acknowledge Dr. J.B. White, the ARMMS project coordinator at MSFC, for his continued cooperation and encouragement during the course of this work.

REFERENCES

GAVER, D.P., 1967, "Probability Models for Multiprogramming Computer Systems," Journal ACM, Vol. 14, pp. 423-438.

HUGHES AIRCRAFT CO., 1972, "Design of a Modular Digital Computer System," prepared under Contract NAS8-27926.

NIELSEN, N.R., 1967, "The Simulation of Time Sharing Systems," Communications of the ACM, Vol. 10, pp. 397-412.

SMITH, J.M., 1968, "A Review and Comparison of Certain Methods of Computer Performance Evaluation," The Computer Bulletin, Vol. 11, pp. 13-18.

UNIVAC Report No. PX 6550-1, 1971, "Space Shuttle Booster Data Management System (DMS) Requirements Analysis", prepared under Contract NAS8-30186.

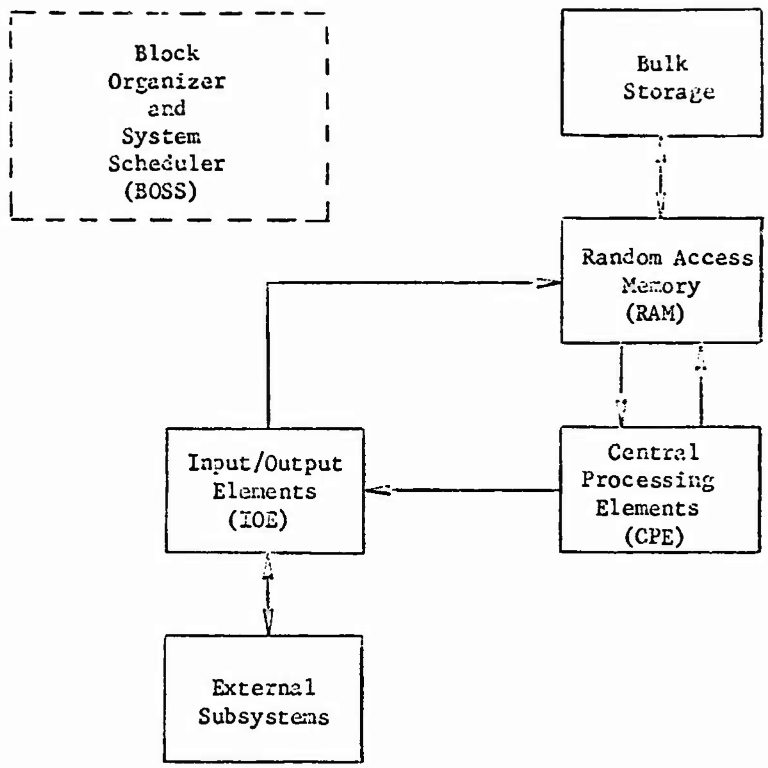


Fig.1 Basic system configuration

0	BEGINNING OF PRELAUNCH CHECKOUT
1	START UP OF THE VEHICLE ELECTRICAL GENERATORS
2	INTERNAL ELECTRICAL POWER ACHIEVED
9	NEAR END OF PRELAUNCH CHECKOUT (GO-INERTIAL)
10	BEGINNING OF LAUNCH
11	INITIATION OF PITCHOVER
18	END OF BOOST SIGNIFICANT ATMOSPHERE
19	BEGINNING OF THRUST TERMINATION
20	BEGINNING OF COAST
30	BEGINNING OF REENTRY
32	ENTRY INTO SIGNIFICANT ATMOSPHERE
34	END OF SURVIVAL PHASE
40	BEGINNING OF CRUISE
50	BEGINNING OF LANDING
51	INITIATE LOWER LANDING GEAR
52	LANDING GEAR DOWN
60	TOUCHDOWN
61	END OF TAXI
70	FERRY TAKE OFF
71	LANDING GEAR UP
80	BEGINNING OF BOOST ABORT
81	END OF BOOST ABORT
82	PILOT INITIATION
83	PROGRAM COMPLETION

Fig.2 Shuttle booster mission phases

PROGRAM NAME	BIGP
TOTAL INSTRUCTIONS SHORT INSTRUCTIONS EXECUTED LONG INSTRUCTIONS EXECUTED	707 1344 376
CONSTANT (24-32 BIT) CONSTANT (12-16 BIT) VARIABLE (24-32 BIT) VARIABLE (12-16 BIT)	0 21 8 60
ITERATIONS/SEC	2
REDUNDANCY	YES
INITIAL SCHEDULING PHASE DESCHEDULING PHASE	18 20

Fig.3 Sample program module

DATA ELEMENT	ITERATION RATE	SIZE (BITS)	UPDATED BY	USED BY
ABEFU BPAYAC	16 2	576 32	NAV BIPG	SDSU FCS
INEVAP	32	192	SD64	SD32 SD16 SD02 BGFP COGP MPST
AILSDO	1	4	AILC AILM	NAV
OCTRUL	4	7	EGSP TCMP TGSP	EGS EGCP

Fig.4 Sample data element description

PROGRAM NAME	IBANK	DBANK	SINST	LINST	PRATE	TMR	INPUT	IRATE	OUTPUT	ORATE
BIGP	707	97	1344	376	2	1	INEVAP	32	BPAYAC	2
							TPTNCT	2		

Fig.5 Sample data base entry

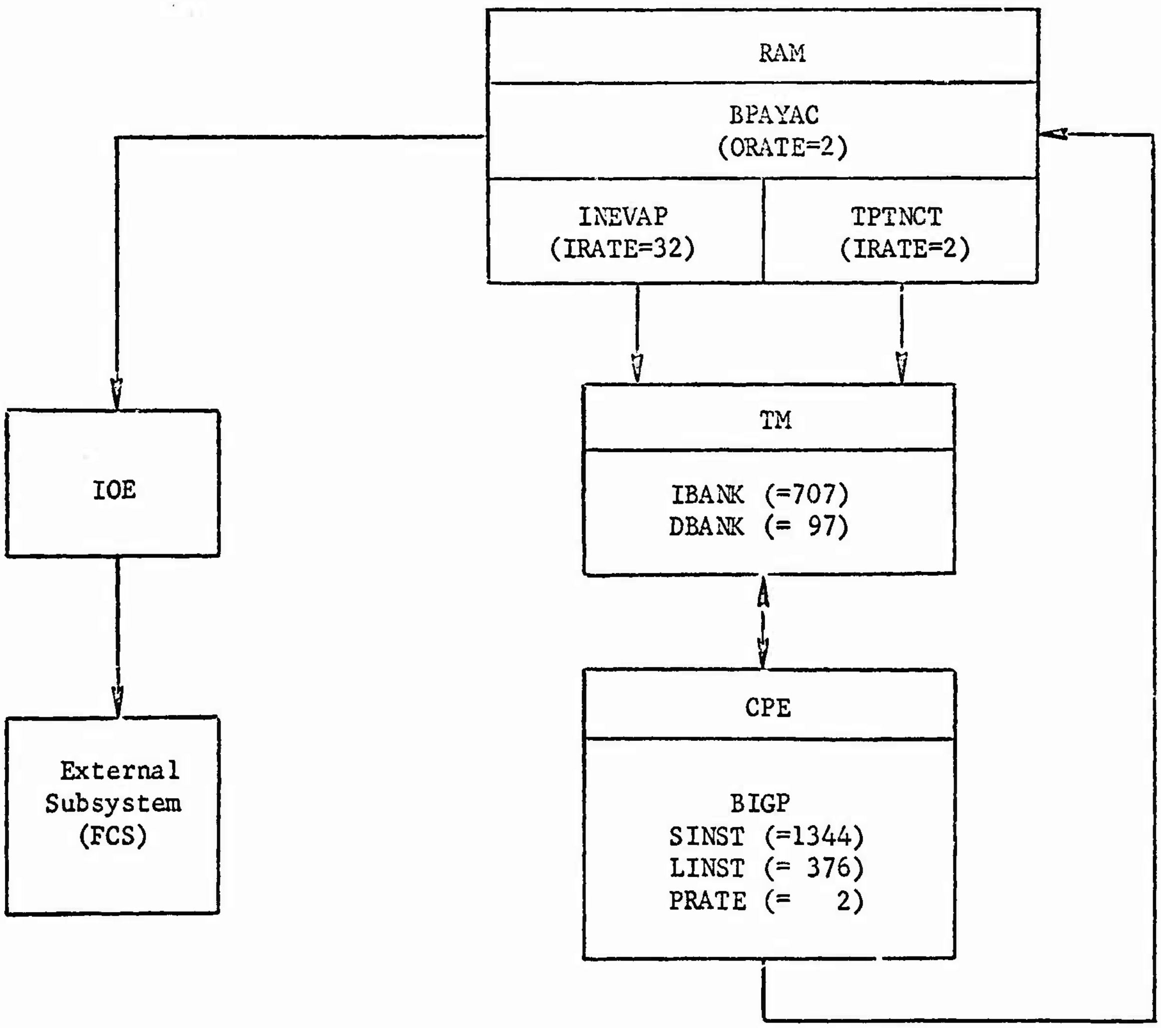


Fig.6 Portion of ARMMS baseline data flow simulation

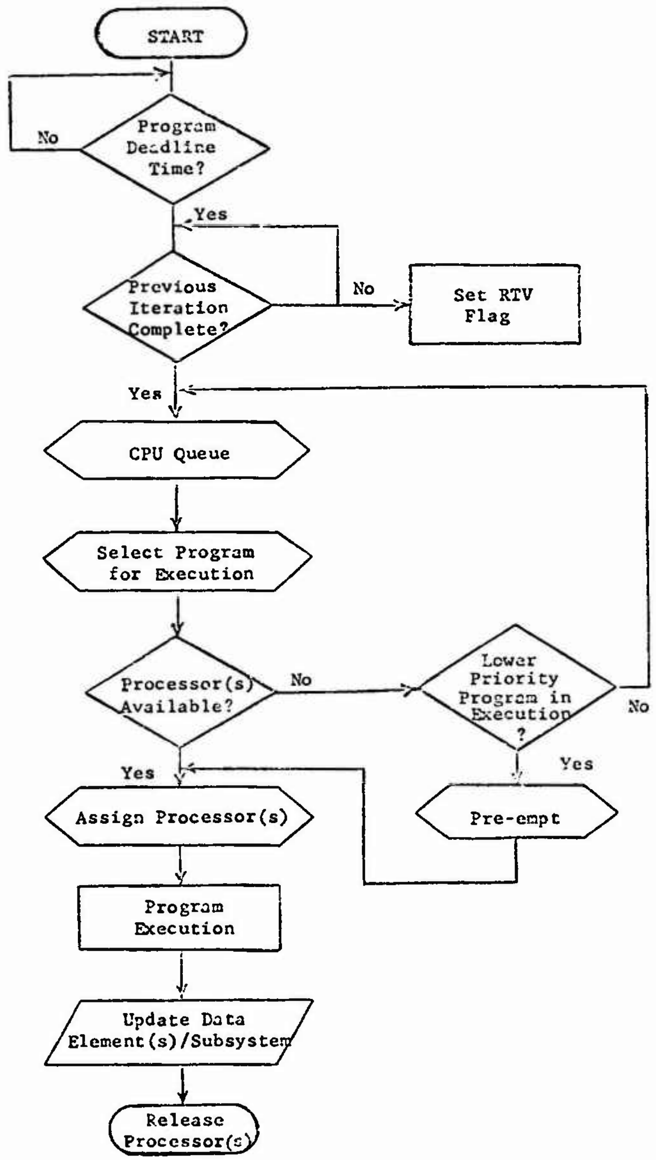


Fig.7 System simulator flowchart

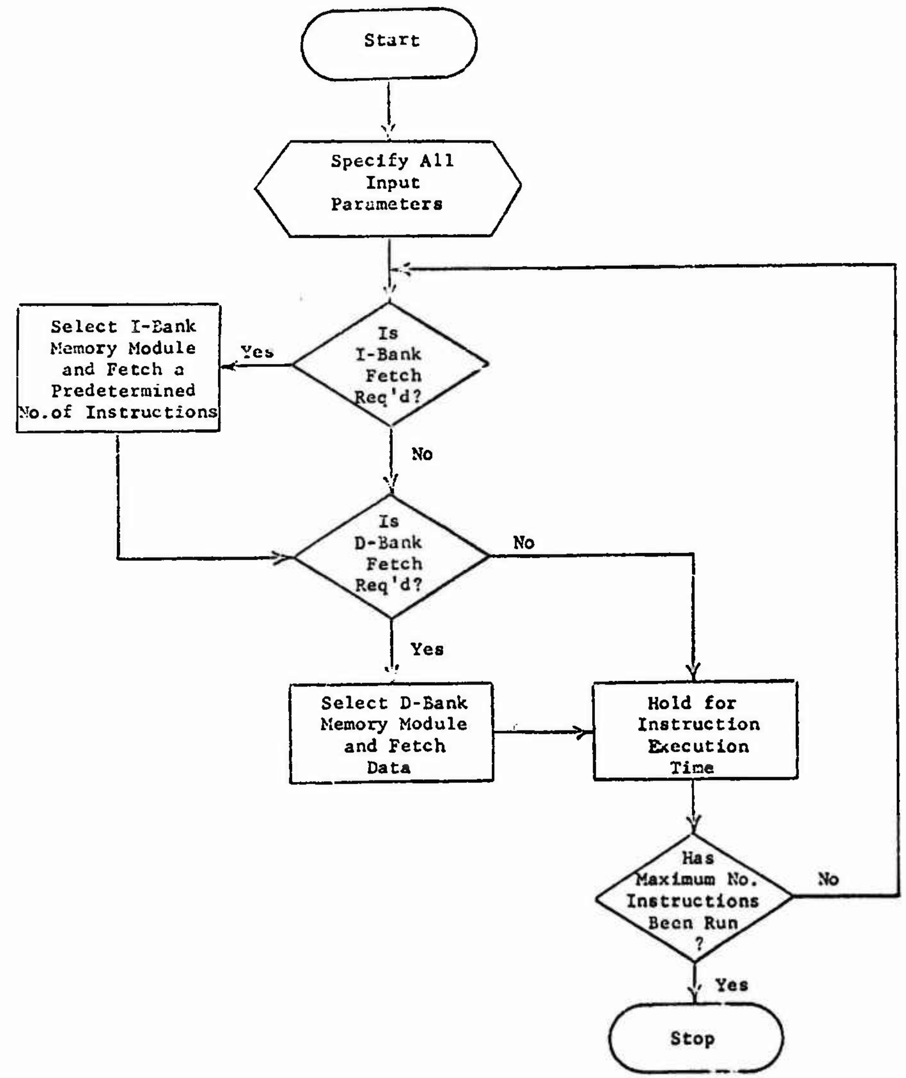
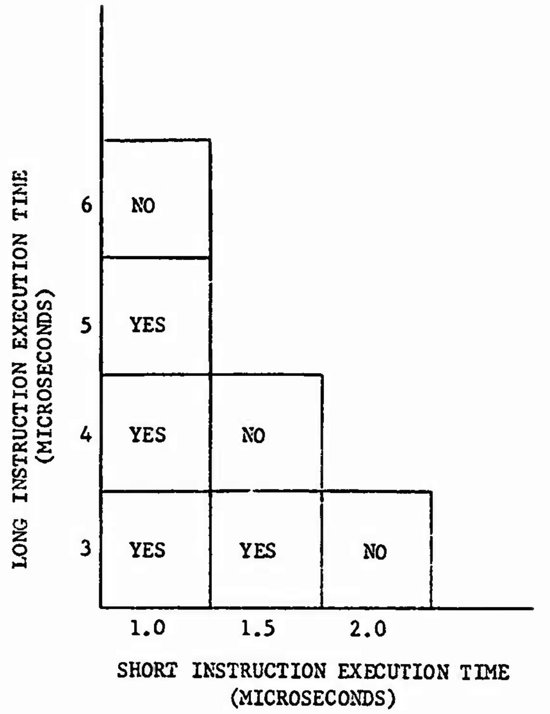
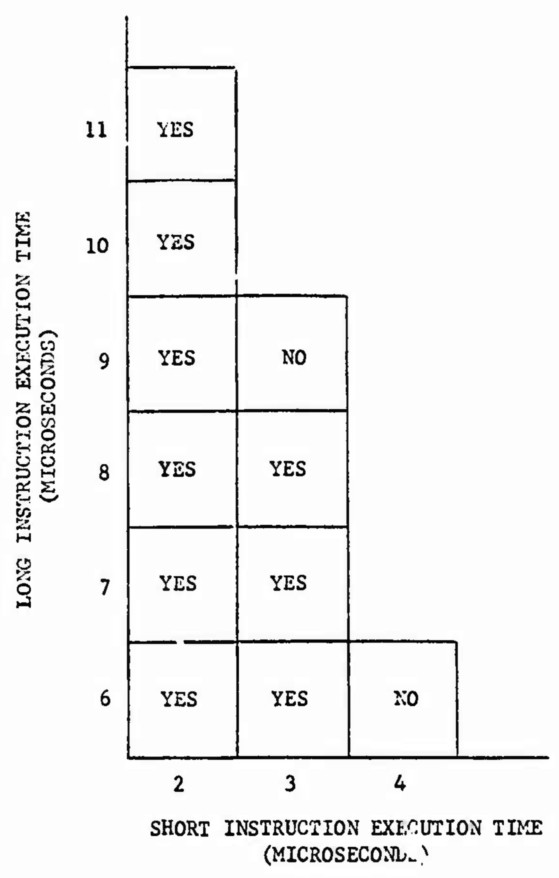


Fig.8 Interface simulator flowchart



YES - All program deadlines met
NO - Some program deadlines not met

Fig.9 Results from system simulator with 3 CPU's



YES - All program deadlines met
NO - Some program deadlines not met

Fig.10 Results from system simulator with 6 CPU's

EXTENSION OF SIMULA 67 FOR PROCESS CONTROL

Juliusz H. Kardasz
ISTITUTO DI ELABORAZIONE DELL'INFORMAZIONE DEL C. N. R.
56100-Pisa
Italy

SUMMARY

This paper presents an extension of SIMULA 67 towards process control. The extension is prepared with an idea of using it to control large systems of interconnected devices where, besides the fundamental control activities, the necessity for real-time simulation arises, in order to define a future behaviour of the system. This extension combines both characteristics of procedural and fill-in-the-blank (format oriented) languages. The programming requirements for process control applications are discussed and a comparison is made between some algorithmic languages with respect to the degrees in which they meet these requirements. This discussion shows that SIMULA 67 requires the introduction of less new concepts than other languages in order to be extended for process control. These new concepts include first of all the "interface with a process" which is introduced by an external class to be implementation defined. A procedural language is used for constructing the body of the system, composed of procedures and classes. Using the facilities of Simula 67 for the preparation of problem-oriented languages, the particular applications can be treated by the introduction of parameters for a fill-in-the-blank language based on the body of the system previously introduced.

1. INTRODUCTION

In this paper an extension of SIMULA 67 towards process control is presented. The idea of extending this language was motivated by a conviction that there is a need in process control environment for a language, which is powerful enough to be suitable not only for very simple types of process control loops, but also for cases of control tasks comprising a great number of interdependent control loops together with a possible variable arrangement of process and control devices.

The requirements which should be met by such a language are discussed in Section 1 and some comparisons of a possibility to meet them by algorithmic programming languages (ALGOL 60, SIMULA 67, FORTRAN, PL/I) are given. In the literature, various extensions of other algorithmic languages, and especially ALGOL 60, FORTRAN and PL/I are to be found; therefore a brief account of these works is given in Section 2. Data reported both in Section 1 and in Section 2 indicate that SIMULA 67 (Dahl, O.J., 1968), being itself a very powerful system, requires a minor amount of new concepts than other programming languages in order to be extended for process control. This is why this paper is directed towards indicating the philosophy of using SIMULA 67 for process control rather than towards introducing large amounts of new concepts. Section 3 contains a discussion of the requirements reported in Section 1 in the light of the possibilities of SIMULA 67, as well as the extensions which are to be added to SIMULA 67 in order to meet these requirements. Section 4 reports an organization of application programmes for process control, expressed in a suitably extended SIMULA 67.

Two features of SIMULA 67 are extensively utilized in this work. The first of them is the concept of "selfextendibility", and the second is the concept of "external class". The selfextendibility of SIMULA 67 meets the fundamental requirement of process control languages to be both procedural and fill-in-the-blank. The body of a system is written in the procedural language SIMULA 67 and contains a number of procedures and classes. These procedures and classes may be changed and updated when a need arises for new statements and declarations. The body of the system is expressed as a class in SIMULA 67 which is called "process control". This class being expressed by standard SIMULA 67 features is implementation independent. A problem-oriented language enables the user to program a specific control system configuration and the corresponding control task, by defining parameters for procedures contained in the class "process control". This problem-oriented language has a fill-in-the-blank character and is composed of a "declaration part" and a "generation part". In the declaration part, the structure of a particular control system is defined by the user; and in the generation part, detailed characteristics of the control devices are introduced.

The process interface is expressed by the implementation-dependent external class "process interface". This class provides process measurements from sensors, control signals for effectors, activation of alarm devices and cooperation between multiple peripheral devices.

This extension can use the classes and procedures of a recently defined, a SIMULA 67 based, language for the simulation of the dynamic behaviour of chemical plants (Kardasz, J.H., Molnar G., 1971); such a combination facilitates the creation of software packages for process control, taking into account the results of a previous simulation.

2. PROCESS CONTROL PROGRAMMING REQUIREMENTS

In order to examine the necessary extensions to SIMULA 67, the process control programming requirements are summarized in Table 1. These requirements are compared with the existing programming structures of the following languages: ALGOL 60, SIMULA 67, FORTRAN and PL/I. This table shows clearly that SIMULA 67 requires fewer extensions than other languages in order to be used for real-time process control. All the subsets of five out of the nine fundamental sets of requirements, as well as a great number of other subsets, are satisfied by SIMULA 67.

Table 1: A comparison of process control programming requirements with a possibility of meeting them by general purpose programming languages without extensions.

Process control programming requirements	SIMULA 67	ALGOL 60	FORTRAN	PL/I
1	2	3	4	5
1. <u>General requirements</u>				
1.1. Algorithmic capability. This is required in a degree similar to scientific computation.	sufficient	sufficient	sufficient	sufficent
1.2. Easy decomposition. Indispensable for modular approach in order to decompose large control systems into small elements.	good	no	no	fair
1.3. Application language capability. A possibility of introducing new concepts.	good	no	no	fair
1.4. List processing capability. Indispensable for treating systems composed of number of elements.	good	no	no	yes
1.5. String handling.	good	option	poor	yes
1.6. Selfextendibility of programme structure.	flexible	fair	no	fair
1.7. Selfextendibility of data structure	very good	no	no	yes
1.8. Dynamic memory allocation.	yes	yes	no	yes
1.9. Quasi-parallel processing.	yes	no	no	yes
1.10. Remote-access for variables.	yes	no	no	poor
2. <u>Data Acquisition and Direct Control Functions</u>.				
2.1. Scanning of sensors. Provides the interfacing required for communication between the process and computer.	no	no	no	no
2.2. Converting analog variables. The analog input signals are converted by the computer to digital values(usually in engineering units) and stored in storage locations.	no	no	no	no
2.3. Alarms. Checking the process varia-	no	no	no	no

1	2	3	4	5
bles against preset limits for normal and safe process operation.				
2.4. Logging. Informing the operator about the process behaviour by writing or displaying appropriate messages.	no	no	no	no
2.5. Closed-loop set point control.	no	no	no	no
2.6. Direct digital control.	no	no	no	no
2.7. Equipment control.	no	no	no	no
3. Optimizing control functions				
3.1. Steady-state performance optimization for processes for which reliable models are available.	yes	yes	yes	yes
3.2. Determination of safe trajectories for operating point change.	yes	yes	yes	yes
3.3. Control and determination of start-up and shut down sequences.	yes	yes	yes	yes
3.4. Dynamic performance optimization.	yes	yes	yes	yes
4. Adaptive control functions				
4.1. Mathematical model identification and modification.	yes	yes	yes	yes
4.2. Security sensing and evaluation.	yes	yes	yes	yes
5. Management information system				
5.1. File handling.	yes	no	no	yes
5.2. Information retrieval.	no	no	no	no
5.3. Maintenance of operating standards.	yes	no	no	yes
5.4. Scheduling. Production planning and scheduling based on sales, demand, or raw materials.	yes	no	no	yes
5.5. Reporting daily production records for individual units and for the entire plant.	yes	no	no	yes
5.6. Development of economic reports for plant management.	yes	no	no	yes
5.7. Text processing.	yes	optional	poor	yes
6. Special Software features				
6.1. Introduction of Assembly type programming.	impl.dep.	impl.dep.	impl.dep.	impl.dep.
6.2. Interprogramme communication.	yes (by files)	no	no	yes
6.3. Rich supply of data types and structures including different character codes, strings, labels, lists, etc., for the variety of programming tasks which arise in operator communication, file handling, executive programme, and control programmes like the DDC system.	yes	no	no	yes
6.4. Ability to reference the hardware features of the system.	no	no	no	no
6.5. Compatibility with the language structure of the special purpose control languages.	yes	no	no	no
6.6. Relative location of the programme and data are known at compile time or only at the object time.	no	no	yes	mixed
6.7. Possibility of treating different hardware configurations.	no	no	no	no
6.8. Ability to test programmes when the system is on-line.	no	no	no	no
6.9. Communication elements: device description, channel rates, code format, mes-	no	no	no	no

1	2	3	4	5
sage recognition, header analysis, validation, acknowledgement, logging, storing, header composition, transmission, polling, interruption.				
6.10. Resolution of conflicting demands: scheduling algorithms, priority levels, dynamic priorities, queue management, queue specification.	no	no	no	no
6.11. Time reference: absolute time, elapsed time,time limits, real clock, pseudo-clock, sequential ordering, precedence.	no	no	no	no
6.12. Handling of real-time analog/digital signals.	no	no	no	no
6.13. Handling external attention signals (interrupts).	no	no	no	no
6.14. Tayloring of programme segments (tasking).	yes	no	no	yes
6.15. Multiprogramming facilities.	no	no	no	yes
6.16. Machine and configuration independence of language.	yes	yes	yes	yes
6.17. Memory protection.	no	no	no	no
6.18. Compiling of the programme when computer is in control (not language dependent).	no	no	no	no
6.19. Possibility for a compiled programme to be tested and debugged while the computer is controlling the process.	no	no	no	no
6.20. Common data base.	yes	no	no	yes
7. <u>Special system functions</u>				
7.1. Activation-deactivation of single alarm points or entire blocks of alarms.	no	no	no	no
7.2. Activation-deactivation of specific control loops.	no	no	no	no
7.3. Changing of control loop parameters.	no	no	no	no
7.4. Changing of coefficients in analog conversion equations.	no	no	no	no
7.5. Changing the state of contact inputs.	no	no	no	no
7.6. Activation of central computer programmes and other special process programmes.	no	no	no	no
7.7. Printout or display on specified peripheral devices.	impl.dep.	impl.dep.	impl.dep.	impl.dep.
7.8. Determination of control actions by means of decision tables of logical variables describing external events.	no	no	no	no
7.9. Specification statements to enable the user to describe the physical properties of objects such as motors, valves, and control loops.	no	no	no	no
7.10. Activation-deactivation of specific devices (pumps, motors, valves etc.).	no	no	no	no
8. <u>Hardware interface</u>				
8.1. Input of data to programme and output of results in the traditional sense.	yes	impl.dep.	yes	yes
8.2. Conversational communication between operator and computer through console or typewriter.	no	no	no	no
8.3. Handling of multiple peripheral devices.	yes	impl.dep.	yes	yes
8.4. Handling of peripheral-device failures.	no	no	no	no
8.5. Ability of handling process input/output.	no	no	no	no

1	2	3	4	5
8.6. Queueing of conflicting requests for i/o devices.	no	no	no	no
9. Other requirements				
9.1. Curve fitting of miscellaneous process data.	yes	yes	yes	yes
9.2. Development and testing of new process models.	yes	yes	yes	yes
9.3. Evaluation of design performance.	yes	yes	yes	yes
9.4. Simulation of proposed process changes and expected process behaviours.	yes	difficult	difficult	yes
9.5. Execution of routine engineering calculations.	yes	yes	yes	yes
9.6. Statistical data reduction.	yes	yes	yes	yes
9.7. Debugging and testing new programmes.	yes	yes	yes	yes

3. TYPICAL EXTENSIONS OF ALGORITHMIC PROGRAMMING LANGUAGES FOR PROCESS CONTROL

Several authors have created extensions to existing algorithmic programming languages in order to adapt them for process control tasks. These extensions were essentially made by providing the possibility of using new special purpose process-control-oriented concepts without increasing the general power of the basic languages. A list of the main features thus introduced is given below, with an indicative purpose and without any detailed description (which can be found in the references).

3.1. ALGOL 60 (Gertler, J., 1970)

a) "variable classes" and "variable forms" for handling process variables, b) special standard functions: sampling, time-sequencing, differentiation and integration, checking, control, positioning, c) "latent parameters" and "substituting parameters" for simplified call of standard functions, d) multiprogramming facilities, e) interrupt system.

3.2. FORTRAN (Roberts, B.C., 1968)

a) ENCODE/DECODE feature (formatted read/write of core buffers), b) zero and negative subscripting of variables, c) expression evaluation in DO parameters, I/O lists etc., d) extended format features, e) variable subroutine returns, f) REPEAT WHILE and REPEAT FOR statements.

3.3. FORTRAN (Hochmeyer, R.E., et al., 1968)

a) reentrant library, b) BYTE statement, c) RELATIVE statement, d) ASSEM statement, e) common data storage, f) MONITOR, g) on-line debugging capability.

3.4. FORTRAN (Mensh, M., 1968)

a) CONNECT INTERRUPT statement, b) CONNECT CLOCK statement, c) CONNECT TIMER statement, d) overlapping of input/output and computation, e) assigned dynamically programme priorities, f) DEVICE statement, g) ability to transfer information between named files in bulk storage and arrays in core.

3.5. PL/I (Boulton, P.I.P., 1970)

a) attribute ANALOG, b) attributes: REFERENCE ACCESS, COMMAND ACCESS, PERIOD ACCESS (time expression), INTERRUPT (interrupt expression, c) HISTORY, d) TIME, e) INTERRUPT, f) programme segment interrelation, g) PRIORITY, h) STATUS, i) COMPLETION, j) EXCLUDE prefix, k) EXCLUSIVE prefix.

4. EXTENSION OF SIMULA 67 FOR PROCESS CONTROL

4.1. Extension of syntax

In order to be able to treat real-time process control programming requirements the following syntactical extension is introduced:

<type declaration>:= <type><simple Algol variable>{,<simple Algol variable>}... |ext <type>
<simple Algol variable >{,<simple Algol variable>}...
<type> := real | integer | Boolean | character | text | ref (<process identifier>)
<simple Algol variable> := <VARIABLE identifier>

Every declaration implying ext is defined by implementation, and the user can use an external variable as if it were declared by the declaration suggested by the extension shown in the above syntax. All the external declarations are included in the class "process control" during implementation.

External definition defines the types for external variable which are means for communicating with external devices. The values of an external variable, of real, integer, Boolean, character or text type are defined by the input devices if it represents "process input". The output for process effectors is defined by the programmer. In the case ext ref the reference value is always defined by the programmer, but the external device can resume the object referenced by it as a result of an interrupt.

All the implementation dependent features are included in the class "process interface" which enters the class "process control". Obviously, any object reference assigned to an ext ref variable must be generated from a subclass of "process interface".

4.2. System defined class "process control"

All the necessary non syntactical extensions of SIMULA 67 are introduced inside the system-defined class "process control". This section contains a list of the particular classes and procedures forming class "process control" together with a discussion of those features of SIMULA 67 which do not require any extension, as they are introduced automatically by the SIMULA 67 common base. The order of the discussion is the same as in Table 1.

4.2.1. General requirements. All of these requirements are met by SIMULA 67. There is no need for the introduction of new concepts.

4.2.2. Data acquisition and direct digital control function. No one of these requirements is met by SIMULA 67 and some new concepts should be introduced. All the required new concepts can be introduced easily by using the general SIMULA 67 framework. For the introduction of these concepts, the concepts of class and of procedure are used.

4.2.2.1. Scanning of sensors. Provides input of both analog and digital signal values, and their updating, at fixed intervals. A data gathering function scans the set of sensors which interface the process, to read the various variables. This activity can be expressed in SIMULA 67 by the procedure "scanning".

```
procedure scanning;
    ref (sensor) sensed value;
    ref (process interface) clock, procin;
    comment "procin" and "clock" are declared in the class "process interface" and stand
    for external process input variables and for the time clock;
    begin
    if clock true do
    begin
    L:Y:-sensor list;
    for Y:-Y.succ while Y =/= none do
    begin
    Y.sensed value := procin ( ... );
    go to L;
    end;
    end;
    end;
```

4.2.2.2. Conversion of analog variables. The analog input signals are converted by the computer to digital values, usually in engineering units, and stored in storage locations reserved for the engineering units table. Functional programmes use only the engineering unit values of the process variables from the engineering units table. For converting analog variables, the procedure "conversion" is used.

```
procedure conversion;
     ref (sensor) sensed value, conversion coefficient, measured value;
     ref (process interface) clock;
     begin
     if clock true do
     begin
     L:Y:-sensor list;
     for Y:-Y.succ while Y =/= none do
     begin
     Y.measured value := Y.sensed value x Y.conversion coefficient;
     go to L;
     end;
     end;
     end;
```

4.2.2.3. Alarm scanning. The alarm scanning programme is used to check the process variables against preset limits for normal and safe process operation. An alarm list specifies the actions to be taken by the control system on abnormal conditions. All analog inputs and calculated variables such as ratios, efficiences, etc. may be scanned by the programme. Depending upon individual system design, limit checking may be performed before or after conversion of the readings to engineering units. There is also a possibility to use more complicated process characteristics for alarm checking, as for instance the balances upon some process units. The procedure "alarm scanning" defined below is based only on checking the process variables against preset limits.

```
procedure alarm scanning;
     ref (sensor) measured value, alarm value;
     ref (process interface) clock;
     begin
     if clock true do
     begin
     L:Y:-sensor list;
     for Y:-Y.succ while Y =/= none do
     begin
     if Y.measured values Y.alarm value do "alarm actions";
     end;
     end;
     end;
```

Procedures based on other checking principles can be written in a similar way.

4.2.2.4. Logging. This operation provides information on process behaviour for the operator. There are two possible types of logging: at defined periods of time, or after an alarm based on the results of an alarm checking. The logging procedure essentially types out the results of measurements or states of devices on the operator's typewriter.

```
procedure logging:
     begin
     .....
     comment there is the specification of texts and data to be typed;
     .....
     end;
```

4.2.2.5. Closed-loop set point control. This is a computer control of an individual loop.

4.2.2.6. Direct digital control.

4.2.2.7. Equipment control. The computer modifies setpoints and controls parameters for the control loops at specified time intervals or when process variables exceed preset limits. In equipment control calculations are designed so as to control a single piece of equipment within an overall process.

The points 4.2.2.5., 4.2.2.6., 4.2.2.7., are discussed jointly because they use the same elements of the general extension of SIMULA 67. The following extensions are introduced:

The action of sensing elements and their performance is expressed by the class sensor. This class contains all the information concerning sensors in the process loop and in an overall layaut of the plant. This information contains data on: process units at which the sensors are attached, nature of measured variable (e.g. temperature), etc.

```
class sensor (sensed value, conversion coefficient, measured value, alarm value, inp);
     ref (sensor) succ;
     real sensed value, conversion coefficient, measured value, alarm value;
     begin
     if sensor number == none then sensor list :- this sensor
     else sensor number.succ :- this sensor;
     sensor number :- this sensor;
     end;
```

Characteristics of effector elements and their performances are expressed by the class effector. The class effector contains all the information concerning actions and connections of effectors in the plant. This information contains data on: process units at which the effectors are attached, and their way of acting on the process.

```
class effector (inp, out, control value);
     ref (controller) out;
     real control value;
     begin
     .....
     end;
```

The actions of controllers which determine the control actions are represented by the class "controller". This class contains information on connections of control loops with the appropriate sensors and effectors and data on the types of process control actions to be performed by the controllers. The following control actions are used in this extension of SIMULA 67: proportional, integral and derivative either simple or combined. Other control modes may be introduced by the user in a similar way by creating ad hoc classes.

```
class controller (inp, out);
     ref (sensor) inp; ref (effector) out;
     begin
     .....
     end;
```

Particular types of controllers are:

```
controller class inoffcontroller (maxvalue, minvalue);
     real maxvalue, minvalue;
     begin
     out.controlvalue := if inp.sensed value <maxvalue or inp.sensed value> minvalue
     then 0 else sign (maxvalue - inp.sensed value)
     end;

controller class analcontroller (setpoint, devnull, CP, CD, CI);
     real setpoint, devnull, CP, CD, CI;
     begin
     real oldreg, tegral, setdev;
     oldreg := setpoint;
     setdev := inp.sensed value - setpoint;
     tegral := tegral + setdev x dt;
     out.control value := devnull + CP x setdev + CI x tegral + CD x (inp.out.control
     value - oldreg)/dt;
     oldreg := inp.sensed value;
     end;
```

The closed-loop set point control forms then a special class prefixed by the class controller.

```
controller class closed-loop set point control;
     begin
     .....
     end;
```

The direct digital control forms an another class prefixed also by the class "controller". These two classes are used alternatively. Anyway, both of them remain in the system and their effective use is determined at the generation stage.

```
controller class direct digital control;
     begin
     .....
     end;
```

The above discussion shows that all the necessary concepts for data acquisition and direct control functions can be expressed by the SIMULA 67 facilities. For programming them, no additional requirements are necessary. It should be noted, however, that data from the sensors and signals for the effectors should be provided by special, implementation created facilities. This point requires an appropriate computer process interface which in some sense is independent of the language definition.

4.2.3. Optimizing effect of the control function. The computer is used to optimize process efficiency, productivity, product distribution, or product quality according to the given mathematical models of a process. Linear programming, search techniques, simulation methods or statistical methods are used to define the operating conditions for process optimization.

SIMULA 67 contains all the facilities necessary for programming different algorithms used for optimization. It should be pointed out that, in this case, it is possible to construct appropriate procedures which define the control values according to the user defined goal function, taking into account the measured variables. For the multivariable control, the programming in SIMULA 67 is more convenient than in other languages.

4.2.4. Adaptive effect of the control function. The computer is used to update models at periodic intervals to take into account declining catalyst activity, fouling of heat exchangers, or decreased performance of separation equipment. This function too is easily programmed in SIMULA 67 without the necessity of introducing new concepts.

4.2.5. Management information system. This is a very important function of control computer systems. This function can be completly fulfilled by SIMULA 67. File handling is here an especially important function as it is necessary for creating and maintaining files which contain measured data and other information on the behaviour of the process and its present characteristics.

4.2.6. Special software features.

4.2.6.1. Introduction of Assembly type programming. This point depends in large scale on implementation.

4.2.6.2. Interprogramme communication. The structure of a programme in SIMULA 67 facilitates such an interprogramme communication. By introducing programmes in the form of classes or procedures, the appropriate coordination of all necessary programme segments is only a matter of main programme organization.

4.2.6.3. Data types and structures. The data types and structures of SIMULA 67 seem rich enough for process control applications, and there is no need for other data types and structures but for the introduction of measured data and control signals.

4.2.6.4. Ability to reference hardware features. This ability is important in order to achieve the required efficiency. For example, control of the executive system when an error occurs is critical for the DDC programme, and the system programmer must be able to reference error indicators, registers, interrupts, status indicators and other hardware variables. All the necessary hardware features are declared inside the <u>ext</u> <u>ref</u> (process interface). The user, then, can generate the object which is resumed at each occurrence of an interrupt from the device.
Example: Assuming the introduction by the implementation of the following declaration:

<u>ext</u> <u>ref</u> (hardware feature) register A;

the user can generate:

register A :- <u>new</u> hardware feature (⟨list of parameters⟩);

<u>comment</u> now any interrupt coming from the device "register A" resumes the object generated by the class "hardware feature".

4.2.6.5. Compatibility with special purpose control languages. This is easily achieved in SIMULA 67 because of the selfextendibility of this language. This point is further discussed in detail in Section 5.

4.2.6.6. Information about relative location of the programme and data. This point can be treated in the implementation phase.

4.2.6.7. Different hardware configurations. There is no essential obstacles to introduce it during the implementation according to what was said at point 4.2.6.4..

4.2.6.8. Ability to test programmes when the system is on line. This is very difficult to introduce using SIMULA 67. It will probably not be implemented.

4.2.6.9. Communication elements. They should all be defined during the implementation.

4.2.6.10. Resolution of conflicting requests. This problem is treated by assigning every request an appropriate priority, which shall cause the resuming of the objects upon the occurrence of external interrupts, in the same way as for the treatment of hardware features discussed in Section 4.2.6.4..

4.2.6.11. Time reference. This is indispensable for scanning the sensors. The "time" reference is introduced in the ext ref (process interface), and can be referenced by the user.
Example: Implementation introduced declaration

    ext ref (time reference) clock;

the user can generate:

    clock :- new time reference (<list of parameters>);

Each time the interrupt occurs from the device clock it resumes the object generated by the class "time reference" which can update the internal system clock by the time elapsed since the last interrupt occurred.

4.2.6.12. Handling real-time analog/digital signals. The way of handling them depends on implementation, and the respective software is introduced in "process interface".

4.2.6.13. Handling external interrupts. The external interrupts cause the resuming of objects created by the user according to what was said in Section 4.2.6.4..

4.2.6.14. Interrelating programme segments. The very structure of SIMULA 67 programme permits to introduce it without difficulty.

4.2.6.15. Multiprogramming facilities. They should be introduced during the implementation. Anyway it is rather difficult to accomplish this under SIMULA 67.

4.2.6.16. Machine and configuration independency. It is an intrinsic feature of SIMULA 67 that programmes in this language and its extensions are machine independent.

4.2.6.17. Memory protection. It should be guaranteed by the implementation.

4.2.6.18. Compiling programmes when computer is in control. It seems extremely difficult to introduce this feature in the present extension of SIMULA 67.

4.2.6.19. Testing and debugging while computer is controlling the process. It will not be introduced.

4.2.6.20. Common data base. The common data base is very easily to be introduced. The file handling system of SIMULA 67 is sufficient for this purpose.

4.2.7. Special system functions. These include fill-in-the-blank type programmes for updating and expanding the basic system programmes, as well as alarm action programmes developed to provide specific actions based on random events occurring during the process. The actions initiated by such programmes are both process and operator-oriented. These functions can be introduced into fill-in-the-blank. (format type) programme which in essence is the application programme in SIMULA 67. This point shall be examined later in some detail. Another point is the implementation of such hardware depending functions as activation of alarms, control actions etc. These points will not be discussed here.

4.2.8. Hardware interface. It is completely dependent on the particular type of hardware, and for this reason it is a task for the implementation part of the work.

4.2.9. Other requirements. SIMULA 67 is quite fit to fulfil these requirements. All of them can be expressed by SIMULA 67 formalisms without any necessity of extensions. Especially, ly, its power for such applications as simulation of proposed process changes and expected process behaviour is remarkable. For this purpose, other extensions of SIMULA 67 can be introduced into this system. In this way, it is possible to perform a control according to the results of process simulation. This is especially important for the cooperation of many control loops working also in a hierarchical way.

5. ORGANIZATION OF APPLICATION PROGRAMMES

5.1. General organization

An application programme written in this extension of SIMULA 67 is composed of two parts. The first part is written in a procedural language and forms a system programme which is common for a large class of process control applications. This part contains also implementation dependent features expressed by the ext ref (process interface). The second part can be called an application programme for a particular application and has the characteristics of a fill-in-the-blank (format type) language. This part contains declarations and generations of particular objects. This way of proceeding is very convenient and derives from the selfextendible character of SIMULA 67.

5.2. Implementation dependent features

According to the previous considerations, the implementation dependent features are introduced in the system programme by the ext ref (process interface). Implementation implies essentially the following declarations:

```
ext ref (process interface) register A, clock, ...;
```

In this way other implementation-defined features can be declared.

5.3. System programme

A system programme is built in the form of a "class" which contains the procedures and classes discussed earlier. All the other necessary facilities are also introduced into this class. The structure of a system programme is as follows:

```
class process control;
     begin
     ref (sensor) sensor list, sensor number, Y;
     ext ref (process interface) ..........;
     class sensor ........................;
     class effector ......................;
     class controller ....................;
     controller class closed-loop set point control .....;
     controller class direct digital control ............;
     controller class inoffcontroller ...................;
     controller class analogcontroller ..................;
     process interface class hardware feature ...........;
     process interface class procin .....................;
     process interface class procout ....................;
     process interface class time reference .............;
     .................................
     procedure scanning ...........;
     procedure converting .........;
     procedure alarm scanning .....;
     procedure logging ............;
     ...........................
     end;
```

5.4. Application programme

An application programme is prefixed by "process control", which enables the application programme to use all the system programme discussed above. This class is composed of two parts: the "declaration part" and the "generation part". It has the character of a fill-in-the-blank programme in the sense that for each particular application the user should introduce only the particular data concerning the nature of control, connections of devices and values of different coefficients. An example of application programme is as follows:

```
process control
    begin
    comment declaration part;
    ref (sensor) sensor1, sensor2, sensor3,........;
    ref (effector) effector1, effector2, effector3,.......;
    ref (controller) contr1, contr2  contr3,.............;
    ..................
    comment generation part;
    sensor1 :- new sensor ( <list of data for sensor 1>);
    sensor2 :- new sensor ( <list of data for sensor 2>);
    sensor3 :- new sensor ( <list of data for sensor 3>);
    ...............
    effector1 :- new effector ( <list of data for effector 1>);
    effector2 :- new effector ( <list of data for effector 2>);
    effector3 :- new effector (<list of data dfor effector 3>);
```

```
.............
contr1 :- new controller ( <list of data for controller 1>);
contr2 :- new controller ( <list of data for controller 2>);
contr3 :- new controller ( <list of data for controller 3>);
.............
clock :- new time reference (<list of data for clock>);
register A :- new hardware feature (< list of data for register A>);
.............
end;
```

6. CONCLUSIONS

The present extension of SIMULA 67 for process control applications meets all programming requirements for this type of application. The minor syntax extension achieved by the introduction of ext type enables to program all the applications, essentially in the same way as in normal SIMULA 67 programming. This syntax extension should be implementation-defined.

The core space needed for SIMULA 67 compiler and efficiency of produced object code remain an open question. The existing SIMULA 67 compilers for non real-time applications require more space than ALGOL 60 or FORTRAN compilers. This point cannot be disregarded, but on the other hand this fact is practically not too important because of the increasing storage capacity in the new process computers. As to the efficiency of object code, it can be stated that some existing SIMULA 67 compilers provide object codes sometimes more efficient than normal ALGOL 60 compilers - and ALGOL 60 was extended for process control.

ACKNOWLEDGEMENTS

The friendly and stimulating discussions with Mr. G. Molnar and Mr. N. Wolkenstein are warmly acknowledged.

REFERENCES

Boulton P.I.P., Reid P.A. "A Process-Control Language" IEEE Trans. on Computers, vol. C-18, pp. 1049-1053.

Dahl O.J., Myrhaug B., Nygaard K. "SIMULA 67- Common base Language" Norwegian Computing Centre, Oslo, 1968.

Gertler J. "High-level programming for process control" The Computer Journal, 13, pp. 70-75, (1970).

Hochmeyer R.E. "CDC 1700 FORTRAN for process control" IEEE Trans. on Ind. Electr.& Control Instr., vol. IECI-15, N° 2, Dec. 1968, pp. 67-70.

Kardasz J.H., Molnar G. "A High-Level Structure Oriented Simulation Language for Chemical Plants" Proceedings of DISCOP - IFAC Symposium on Digital Simulation of Continuous Processes, Győr, Hungary, Sept. 6-10, 1971.

Mensh M., Diehl W. "Extended FORTRAN for Process Control" IEEE Trans. on Ind. Electr. & Control Instr., vol. IECI-15, N° 2, Dec. 1968, pp. 75-79.

Roberts B.C. "FORTRAN IV in a Process Control Environment" IEEE Trans. on Ind. Electr. & Control Instr., vol. IECI-16, pp. 61-63, (1968).

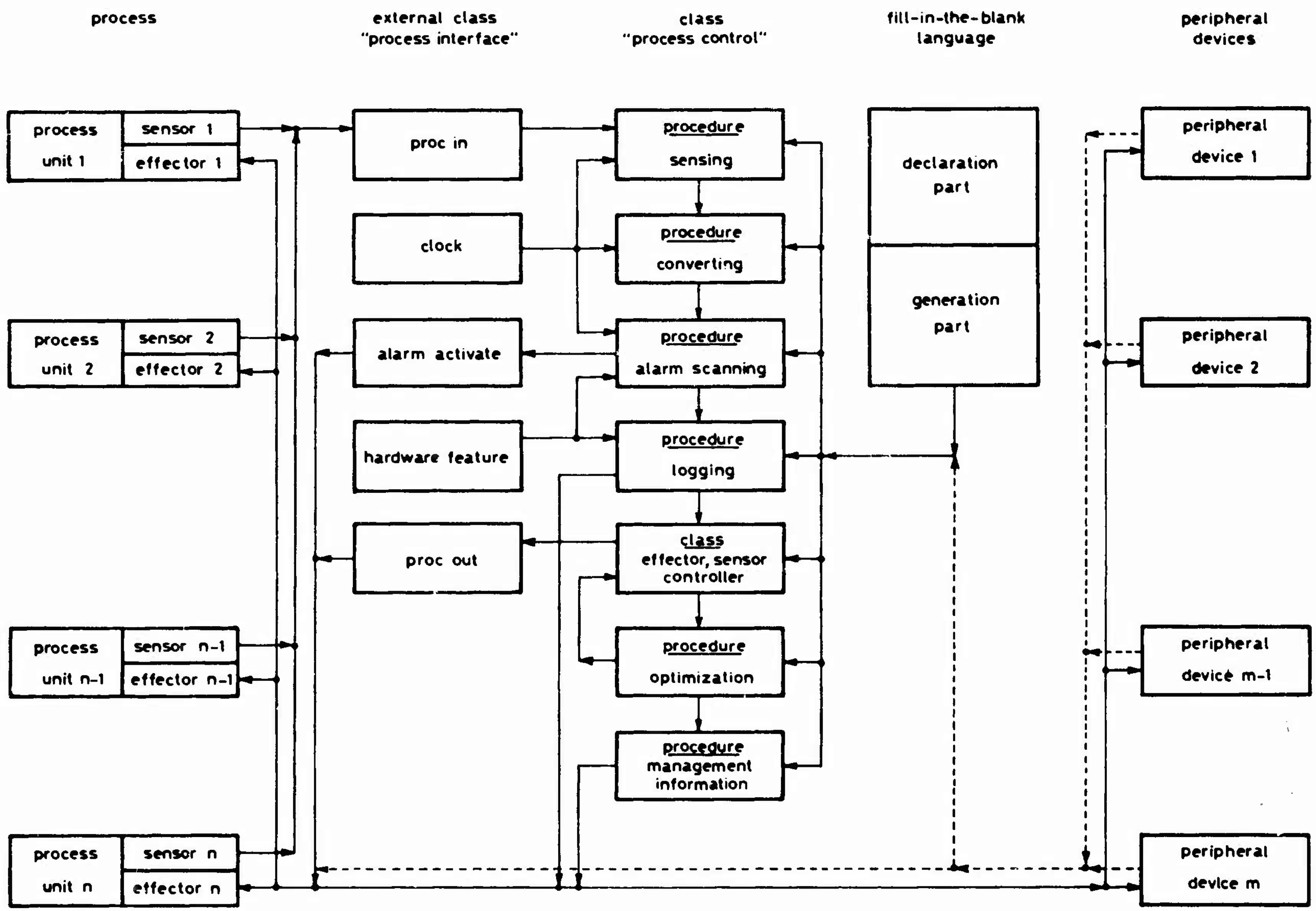


Fig. 1 - Structure of process and software interaction in the extension of SIMULA 67 for process control.

SYSTEMS PERFORMANCE MONITORING FOR ADVANCED MANNED SPACECRAFT

T. V. Chambers
NASA Manned Spacecraft Center
Houston, Texas 77058

SUMMARY

Much effort has been expended in the last few years in defining optimum system mechanizations for advanced manned spacecraft. Several studies have proposed automation of the onboard system management task, with functions such as system status monitoring, configuration management, and redundancy management being accomplished under computer control. An experimental system was used in the laboratory to investigate hardware and software requirements for accomplishing these onboard system management functions. No significant technology problems were revealed. Projections of the onboard processing required for an all-up vehicle, however, indicate that considerable resources would be required to accomplish the overall integration and software verification tasks.

A performance monitor system is proposed for the space shuttle. This system provides support to the flight crew in the management of all onboard systems but does not perform critical switching functions during the flight phase.

1. INTRODUCTION

Responsibility for the monitoring and management of the onboard subsystems in previous manned spacecraft of the United States of America has been shared by the flight crew and the ground-based Mission Control Center. Extremely critical measurements either were available to the flight crew in the form of the normal operational displays or were accessed by a caution and warning system, which alerted the flight crew to anomalous behavior of the onboard equipment. A far greater number of measurements were telemetered to ground stations, where flight controllers were able to track the operation of the spacecraft continuously and, if necessary, to call upon considerable resources of subsystem experts, subsystem data, and previously compiled computer programs to carry out diagnostics, fault isolation, and trend analyses.

The primary objective of the total monitoring effort is, of course, to enable initiation of corrective action before an anomaly can propagate to a catastrophic failure and to provide information on which to base a decision to continue the normal mission, to substitute an alternate mission, or to abort.

For several advanced vehicles, the provision of a more sophisticated onboard capability to support the flight crew in these tasks appears warranted. Such a provision would reduce the dependency on the ground for systems management and, ultimately, would support the objectives of lower costs during the operational stages of these programs.

2. ADVANCED VEHICLES

The National Aeronautics and Space Administration (NASA) space activities in the decade following the Apollo and Skylab Programs are oriented toward (1) unmanned or man-tended exploration and observation of the universe, the solar system planets, and the earth, and (2) manned applications and research activities in earth orbit. A vital system for supporting both types of activities is the reusable space shuttle transportation system for delivery of payloads to and from earth orbit. Research and applications module (RAM) payload carriers are the modular manned laboratories and man-tended observatories for use with the earth-orbiting space shuttle and, ultimately, space station systems for accomplishing these objectives (figure 1).

To date, space exploration has been achieved through the use of expendable launch vehicles and spacecraft. For continued space-flight development in the exploration and exploitation of near and far space, an economical, reusable logistics system is required. The objective of the Space Shuttle Program is to provide a low-cost, economical space transportation system. The NASA objectives require that the development costs and the operational costs must be minimized to the greatest extent possible (reference 1). To achieve low operational costs, reductions must be made in the facilities and personnel supporting flight operations. An onboard system that will support the crew in all system, redundancy, and mission management matters and, thus, reduce the dependency on ground support will influence operational costs favorably. In addition, such a system, used in conjunction with carefully designed ground-support equipment, can reduce facilities and personnel required for vehicle checkout significantly. By streamlining the vehicle checkout and reducing turnaround time, a smaller vehicle inventory can support the same traffic model.

In addition to the requirement for onboard monitoring of the basic shuttle system, equipment for handling the shuttle payloads also is required (reference 2). The sortie RAM is a pressurized module designed to operate when attached to the shuttle orbiter while accommodating a wide array of experiment equipment and providing the subsystem resources essential for the conduct of the experiment. The sortie RAM also provides support for a crew of two mission/payload specialists, who would be transported to and from orbit in the shuttle orbiter. The module is approximately 15 feet long and has a nominal outside diameter of 14 feet. An interface adapter assembly is provided at the forward end to mate with the shuttle orbiter.

On reaching the on-orbit phase, the RAM carrier system and experiments are checked out in preparation for the nominal 6-day experiment period. During this on-orbit phase, detailed knowledge of experiment anomalies, degradation, and overall performance should be made available to the flight crew to ensure a cost-effective mission.

Another type of payload proposed for the shuttle is the free-flying RAM. After boost to orbit by the shuttle, the free-flying RAM is deployed from the cargo bay, and payload activation and checkout are begun. Again, a requirement exists for the acquisition of detailed subsystem and experiment status data that are processed for ready assimilation by the flight crew before separation is attempted.

The on-orbit participation of man is provided to capitalize on his unique abilities for planning, evaluation, recognition, selection, control, and editing that cannot be preprogramed readily or achieved remotely. In addition, on-orbit service, update, maintenance, and repair of free-flying observatories by periodic manual tending allows long operational and scientific life of these facilities.

Full participation of the flight crew in these tasks requires that the subsystem management duties are organized in such a way that they represent a reasonable proportion of the crew total workload. Accordingly, systems have been proposed for these vehicles to not only carry out subsystem performance monitoring tasks under the control of an onboard computer but, acting upon the information acquired, to perform all configuration, redundancy management, and routine maintenance operations in an entirely automatic manner (references 3 to 6).

In an attempt to gage the development risk associated with the hardware and software implementation of such systems, two divisions at the NASA Manned Spacecraft Center (MSC) (Houston, Texas) have investigated typical problems in the laboratory. An experimental quad-redundant partial guidance and control (G&C) system was assembled by the MSC Guidance and Control Division. One objective was to investigate mechanization problems associated with performance monitoring and redundancy management. The system basically consisted of a simplex flight control computer and a simplex system management computer communicating with redundant sensors and actuators over a quadruple data-bus system. Full details of this system can be obtained from references 7 and 8.

3. ONBOARD SYSTEMS MANAGEMENT BREADBOARD

Another investigation, conducted by the MSC Information Systems Division (ISD), consisted of the integration of representative partial spacecraft subsystems with a data management system (DMS) employing sophisticated display and control (D&C) techniques. Although the DMS and the D&C subsystem were nonredundant, certain redundancy features were incorporated in the partial subsystems to investigate and to demonstrate

(a) System performance monitoring

(b) Redundancy switching

(c) Configuration management

(d) Limited onboard checkout

Of particular interest was the relationship between these functions and the D&C subsystem. A primary objective of the demonstration was to provide automatic control for relieving the flight crew of tedious tasks where the functions were of a routine housekeeping nature and, in isolated cases, where the switching time criticality ruled out manual response times.

The overall system architecture desired was a general case in which the computational function of subsystems with minimal requirements would be handled by a central data management computer whereas other subsystems with specialized or demanding requirements would be serviced by a local processor or subsystem management computer. However, practical considerations such as the availability of hardware and personnel, the difficulties of the integration process, and, in particular, the necessity to break down the software problem into sections manageable by the individual organization elements of the ISD led to the choice of a decentralized architecture.

The basic subsystems comprising the onboard systems management (OSM) breadboard are

(a) The DMS

(b) The D&C subsystem

(c) The electrical power subsystem (EPS)

(d) The environmental control subsystem (ECS)

(e) The instrumentation subsystem (INS)

Each of the last four subsystems incorporates a local processor or subsystem management computer as shown in figure 2, with the DMS computer controlling the data bus and converting subsystem data to engineering units.

3.1. Onboard Systems Management Breadboard System Functions

The overall OSM breadboard is able to demonstrate, to some degree, the following computerized functions.

(a) System performance monitoring - The monitoring procedure is passive and is accomplished by reading test points and by comparing their values to a series of limits stored in the computer. When a parameter is found to lie outside a limit, an indication is transmitted to the caution and warning (C&W) display.

(b) Redundancy switching - To provide a degree of redundancy, certain functional paths within the subsystems were duplicated. Controls were provided to inject faults into the main functional paths. Detection of these faults by the continuous status monitoring operation resulted in the computer taking corrective action to switch in a redundant path. The D&C subsystem indicated the corrective action taken.

(c) Configuration management - In this context, system configuration management means the control and monitoring of the configuration in accordance with programed time lines and operating procedures. During the initialization sequence, subsystems were automatically configured to the "normal" configuration after selecting alternate configurations to check out redundant paths. In addition, by command, selected configuration changes (e.g., "landing gear down") were simulated by the combined systems.

(d) Onboard checkout - Normal initialization of the system included configuration switching to all redundant paths for detection of any malfunctions by the surveillance or passive monitoring routines. However, the system also is capable of the scheduled testing of all parameters necessary to verify equipment integrity before operation. In this mode, each measurement can be accessed and displayed in engineering units for comparison with a written operational checkout procedure. It is probable that, in a realistic vehicle system, certain functional paths would require the insertion of stimuli to ensure complete checkout coverage. This capability was not provided in the OSM.

3.2. Onboard Systems Management Breadboard Description

3.2.1. Data management system

The inclusion of the DMS in the OSM breadboard has a number of distinct advantages in terms of providing the capability for evaluating DMS concepts during dynamic operation of the OSM breadboard. A major concern of all data management schemes is coordination of the man-machine interface. The crew must be provided with information on the condition of the total system as well as information on the condition of the individual subsystems and their components. As a result of this requirement, the DMS software includes a structured D&C processing data base that ensures rapid and accurate retrieval of specific data in which the crew has indicated a current interest. In addition, this data base is capable of bringing critical information to the crew's attention without solicitation. These forced displays do not preempt selection by the crew; they augment it.

A second major concern of all data management schemes is the coordination of the subsystems that manage the local required functions. As implemented on the OSM breadboard, the DMS provides executive control of all data-bus traffic through a software polling scheme. Each subsystem, in a predetermined sequence, is granted a time period in which to conduct its bus communications. The subsystem sequence is executed periodically and is alterable dynamically to allow subcommutation or supercommutation of subsystem-bus access based on mission requirements. In contrast to a hardware time-slot generator, the individual subsystem is provided with a surrender mechanism, which allows use of any part or all of the allocated time period. The data traffic between subsystems and with the DMS computer is formatted and customized to satisfy the individual subsystem needs.

In an evaluation project such as the OSM breadboard, it is useful to be as realistic as possible. Use of a ground-based computer for the DMS allows testing of DMS concepts, but a more convincing demonstration can be made with a flight computer. The IBM 4 Pi EP computer used on the OSM breadboard as the DMS computer originally was designed for the U.S. Air Force Manned Orbiting Laboratory. Although this computer is not fully flight qualified, it is flight packaged and represents a typical spacecraft hardware configuration. As a result, the program design and programing conventions have provided realistic experience that is relevant to advanced spacecraft projects (figure 3).

3.2.2. Display and control subsystem

The D&C subsystem is the major monitoring and control point of the OSM breadboard for the system operator. It exists as the only operating man-machine interface for the OSM breadboard. Although data routing and management tasks are performed by the DMS, all results of these operations (those immediately available to the operator) are generated and displayed in the D&C subsystem. Breadboard control also is initiated at this subsystem. Control of subsystems is limited to those external commands that the other subsystems are designed to accept. The D&C subsystem does not control the internal operation of other subsystems because each subsystem includes a local processor, internal control software, and internal control hardware to accomplish the assigned tasks.

In performing the control and monitoring functions for the OSM breadboard, the D&C subsystem

(a) Receives commands from the operator

(b) Formats the commands for data-bus transmission

(c) Transmits the commands to the DMS under DMS control

(d) Receives data from the DMS

(e) Sorts data according to D&C subsystem component destination

(f) Retrieves static data from mass storage in accordance with received instructions

(g) Generates English-language messages for display on the cockpit cathode-ray tubes (CRT's)

(h) Routes DMS input data directly to the cockpit display devices

All commands from the D&C subsystem to the other subsystems first must be sent to the DMS. The routing and the implied meaning of messages are determined by the DMS through previous operator-commanded levels of operation and subsystem control. After establishment of the level of control and the subsystem to be controlled, the D&C commands are routed to the intended subsystem for execution.

The cockpit arrangement of the display devices is shown in figure 4. Five display devices are used as the input and output means of controlling and monitoring the system. The alphanumeric reconfigurable control panel (ARCP), the power initialization panel, the matrix read-out panel, the numeric keyboard, and the C&W panel constitute the implementation of the input/output devices.

The OSM breadboard was developed and implemented to operate on a "management by exception" concept, by which each subsystem would operate without direct and continuous monitoring by the DMS. However, all failures or error conditions sensed or corrected (or both) by a subsystem will be reported to the DMS as a warning condition. When this warning is passed on to the D&C subsystem, an indication will appear on the C&W panel. When this warning occurs, the operator may wish to determine the exact status of the warning condition within a particular subsystem. This concept of subsystem control, including operator actions, is illustrated in figure 5. The operator observes a warning on the C&W panel for the configuration status recorder (CSR) and calls up additional monitoring capabilities for the subsystem with the ARCP. The subsystem then responds with the condition that initiated the warning to the DMS. The operator then takes corrective action by replacing the tape, by cleaning the heads, or by taking other required corrective measures to remove the warning. Next, the operator requests status and gets a "SYSTEM GO" if corrections are complete. The warning indicator is removed automatically upon receipt of a GO indication. The operator has the option of selecting and troubleshooting each subsystem in turn by using the previously mentioned sequence of operations.

3.2.3. Simulated subsystems

The overall OSM breadboard included partial simulators of three onboard systems to provide a representative task for exercising the DMS and the D&C subsystem. The EPS and the INS are discussed briefly. The ECS simulator is discussed in more detail to illustrate the degree of automation demonstrated.

3.2.3.1. Electrical power subsystem

The EPS of the vehicle is assumed to consist of redundant distributed 28-volt buses supplying both real (demonstration system) and dummy loads. Supplies of 115-volt, 400-hertz alternating current (ac) also are redundant, and the inverter itself is modular in that each phase is packaged separately and can be replaced by a switched-in standby in case of failure. Both direct-current (dc) and ac systems rely exclusively upon solid-state circuit breakers and power controllers to provide compatibility with the computer control of the system and to benefit from the weight saving made possible by remote operation of the circuit breakers.

The subsystem management computer is a 16-bit digital computer processor fabricated from standard MOS-LSI circuits and designated the P703 microprocessor. The P703 processes more than 80 signals from the EPS in accomplishing its four main tasks of powering up, surveillance, load-fault correction, and redundancy switching.

The primary mode of the P703, the surveillance mode, is entered automatically at the completion of the power-up routine. In this mode, the preprocessor continuously monitors all measurement points for indication of a tripped power controller (PC) and for out-of-tolerance readings on analog data.

Identification of a tripped PC by the P703 will result in a change to a routine that will attempt to reset the PC or that will take other measures to reconfigure the system into a safe condition. If the load is equipped with a redundant PC, the P703 switches the load to another PC within 1 millisecond. The P703 then resets the tripped PC and applies it to the load. If the PC successfully resets and takes up the load, the redundant PC is reset, and the program returns to the surveillance mode.

The DMS sequentially polls each standard digital interface (SDI). In this particular case, the EPS-SDI will respond with a data word indicating the loss of one level of redundancy. As a result, the DMS will set up a C&W indicator in the crew station to provide the crew with a display of the EPS status. Figures 6 and 7 show the EPS equipment.

3.2.3.2. Instrumentation subsystem

(a) Transducer signals - The INS contains a device, the sensor instrumentation processor (SIP), that is capable of sampling transducer measurements from other OSM breadboard subsystems. The function of this subsystem is to produce signals for the DMS simulating those that might exist in actual space vehicle instrumentation or others that may not be part of any definitive subsystem. At the present, eight instrumentation channels from the central timing equipment (CTE), as well as 14 internal SIP measurements, are being processed. The SIP, as presently configured, has a maximum capacity of 72 signal channels and logic for a maximum capacity of 128 signal channels.

(b) Signal processing - The basic SIP output is addressable upon request by the subsystem processor, and each discrete channel is a unique voltage level from one signal source. This output is connected to a modified DMI 620 computer analog controller with a 128-channel interface capacity. The signal processing at this point involves level-shifting logic, control logic, and analog-to-digital conversion.

(c) Central timing equipment - The CTE provides, through the DMS, the time of day and frequency standards for the OSM breadboard. The CTE is an engineering model of a rubidium spacecraft atomic timing system. The time code word provided to the DMS computer is updated once per second. Time in milliseconds is then provided by the DMS. The CTE does not use the data bus because of the high requirements on signal data rates and periodic accuracy. However, the operating status of the CTE is determined by an analog signal interface with the SIP, which does operate by way of the data bus.

(d) Configuration status recorder - The CSR records on magnetic tape any properly identified messages on the data bus, specifically the messages with the control-word bit 16 set to 1. The subsystem supports the OSM breadboard as a flight recorder might serve in an operational aircraft or spacecraft. Data transfers from any OSM breadboard subsystem to the DMS must be of the single-word type in the OSM breadboard. Input sensed by the CSR will be integrated into a stored signal sequence with 500-millisecond timing signals from the DMS. In normal breadboard operation, the CSR is passive and is controlled by DMS commands until a recorder fault is detected, when an indication is sent to the DMS.

(e) Recording - The recording element of the CSR is an Ampex TM-11 tape unit operated with a DMI tape controller device. The TM-11 records with seven-track heads on standard 1/2-inch magnetic tape at densities of either 556 or 800 bits/in. and at tape speeds of 120 in/sec.

3.2.3.3. Environmental control system description

The ECS, as represented on the OSM breadboard, consists of the local processor (designated Micro 500 processor), the electronic circuitry configured to simulate a cabin atmosphere circulation loop, a water coolant loop, a fluorocarbon coolant loop, a processor/standard digital interface unit (SDIU) adapter, and a display panel. The simulator model is a portion of a representative environmental control and life support system. The display panel visually indicates system status and can be used to insert malfunction signals into the simulator. These elements interface to the other elements of the OSM breadboard through an SDIU, which is connected to the main serial data bus. The ECS is connected approximately 400 feet from the data management computer with which it interacts to provide full control of the simulated cabin atmosphere and of the water coolant and fluorocarbon coolant systems.

The cabin environment is controlled automatically by the local processor in the normal operating mode. However, the DMS can reconfigure or control the system as a result of crew commands through crew input devices. The crew also may require the ECS to perform specific checkout routines and to report status. The local processor will report certain off-nominal conditions through the C&W system and will report full status when it receives a status request from the crew by way of the DMS.

3.2.3.4. Equipment configuration

Figure 8 is a block diagram of the major elements of the ECS simulator as implemented on the OSM breadboard. The major components of the ECS are

(a) The processor/SDIU adapter

(b) A Micro 500 processor

(c) The processor/simulator interface

(1) Measurement cells

(2) Stimulus cells

(3) A local serial data bus

(4) The cabin environment/ECS simulator

(5) The simulator display panel

Processor/SDIU adapter - The processor/SDIU adapter provides a parallel interface for the Micro 500 processor with the SDIU and, thus, with the data bus. A half duplex channel is provided for the transfer of data and commands from the data bus to the processor and for the transfer of response and status data from the processor to the data bus.

In addition to the input and output data channels, input and output interrupts and control discretes are conditioned for use by the SDIU and the Micro 500 processor during data transfer. The adapter is basically a data transfer device and does not provide any buffer storage.

Micro 500 processor - The Micro 500 processor serves as the ECS local processor. It provides all local computations required by the system and responds to commands received from the data bus and to data received from the simulator.

The Micro 500 processor is a 2048-word (16 bit) computer with limited software instructions. In addition to the standard instruction set, which uses the control and arithmetic section of the processor, there are five hardwired functions that are activated directly through an input/output (I/O) channel and do not use the Micro 500 software.

The hardwired functions are

(a) START - Starting address for a memory load sequence.

(b) LOAD - Load data word into Micro 500 memory.

(c) DUMP - Return specified Micro 500 memory cell content to DMS.

(d) BYPASS (measure) - Return value of specified data point to DMS.

(e) BYPASS (simulate) - Apply DMS specified stimuli to ECS simulator.

The first three of these functions are used in the initial loading of the Micro 500 software from the central computer by way of the data bus and subsequent testing of the load. This is the technique used to initialize and bring the ECS on line. The last two functions were designed to allow DMS access to the individual measurement and control points in the ECS without using the local processor.

The processor provides local control of the cabin atmosphere, fluorocarbon, and water systems, based upon feedback it receives through the measurement system, and controls all normal (nonlocal processor bypass) communications with the DMS. The processor receives measurement data by way of four local serial data buses from two analog and two discrete measurement cells. Output commands from the processor to the ECS valve actuator are transmitted over four local serial data buses to the stimuli cells.

Measurement/stimulus cells - The analog measurement system is basically a report-by-exception system in that it provides a data acquisition capability that is non-computer controlled (in normal operation), and this feature decreases the load on the system processor below that which would occur in a computer controlled scan system. The measurement cells scan the analog measurement points for an out-of-floating-tolerance variation from the last measured value. If an out-of-floating-tolerance condition is detected, an external interrupt is sent to the processor, initiating data input for subsequent processing. Measurements are input to the local processor on an exception basis, which is controlled by a hardware comparator. Measurements are read by the computer only after the comparator has sensed an out-of-tolerance change in a variable and has transmitted an interrupt to the processor. The autonomy of the measurement system and its report-by-exception mechanization lowers the overall local processor load below that of a computer controlled scan-and-read-all-measurements system. The event measurements are handled similarly, except that any change in the state of an event generates an interrupt. Fixed-limit checks on analog measurements are made by processor software.

All measurements taken from the simulator are the electrical analogs of system variables. No sensor hardware is included in the simulator.

The stimuli cells are used to apply discrete commands to the various control points in the ECS simulator. These discrete commands originate in the processor as a result of ECS control algorithms, are transmitted to the stimuli cells by way of the local serial data bus, and are applied to the simulator as simulated valve-actuation commands.

ECS simulator - The ECS simulator is an electrical simulation of cabin atmosphere contamination and temperature, of the heat exchanger and decontamination loops, and of the actuators and sensors used in controlling the system. The processor measures 27 simulated variables in the simulator and commands actuation of simulated valves in the simulator for control. In addition to the control inputs from the processor, several failures and a boundary condition (radiator away from sun; radiator to sun) may be input to the simulator from manual switches on the display panel.

The simulator is restricted to mechanization of (1) cabin atmosphere contamination and the purification system for the atmosphere and (2) the temperature of the cabin atmosphere and the spacecraft water and heat exchanger system used for temperature control. These two mechanizations are dynamic loops that are monitored and controlled by the control system within the Micro 500 processor software. Figure 9 shows the display panel with schematic representations of the dynamic loops as implemented on the OSM breadboard. The left side of the display panel depicts the loop that removes CO_2 and water vapor from the cabin atmosphere, whereas the right side of the panel depicts the loop that removes waste heat from the spacecraft. The simulator automatically provides for systematic increases in cabin air humidity and cabin air temperature. These systematic increases provide a continuous stimulus to the control system and demonstrate its operation in a dynamic situation.

The ECS simulator, as modeled on the OSM breadboard, is composed of a number of functional elements. The following list defines these elements and their functions in the loop.

(a) Desiccant bed - Removes water vapor from the cabin atmosphere

(b) LiOH filter (2 each, A and B) - Removes carbon dioxide from the cabin atmosphere

(c) Air compressor (2 each, A and B) - Drives cabin air through LiOH filters and desiccant bed

(d) Cabin heat exchanger - Exchanges cabin atmosphere heat with water

(e) Coldplate network - Provides chilled water to avionics systems for heat dissipation

(f) Water accumulator - Provides residual supply of water on inlet side of water pumps

(g) Water pump (2 each, A and B) - Forces water through system

(h) Fluorocarbon/water heat exchanger - Transfers waste heat from water to fluorocarbon

(i) Fluorocarbon accumulator - Provides residual supply of fluorocarbon on inlet side of fluorocarbon pump

(j) Fluorocarbon pump - Forces fluorocarbon through system

(k) Radiator - Disposes of waste heat from fluorocarbon to space

(l) Air cycle heat exchanger

(m) Water boiler - Decreases water temperature

(n) Ground coolant - Provides chilled water to system when on the ground

(o) Water chiller - Cools drinking water

The hardware simulator provides for two continuous maintenance functions, the desiccant bed regulation and the cabin air temperature regulation. The cabin air humidity is increased systematically until the desiccant bed humidity reaches 80 percent. The cabin air flow then bypasses the desiccant bed while the water vapor is vented overboard. When the desiccant bed humidity drops to a preprogramed level, the overboard vent is closed, and the cabin air is again routed through the desiccant bed. The cabin air temperature is increased by 4° F from its previous value. When this level is reached, the water is allowed to flow through the cabin heat exchanger to remove cabin air heat. When the cabin air temperature is reduced to its proper level, the water flow again bypasses the cabin heat exchanger. These processes are continuous during the simulator operation. The simulator contains three loops: cabin air, water coolant, and fluorocarbon; all of these loops are dynamic in the sense that they have changing analog parameter and control points.

The fluorocarbon loop simulates the transfer of heat from the water coolant to the fluorocarbon and the dissipation of the fluorocarbon heat into space by way of a radiator. Included in the fluorocarbon loop are a water/fluorocarbon heat exchanger, a fluorocarbon pump, a fluorocarbon accumulator, and a radiator. Fault insertion devices are included in the fluorocarbon loop of the simulator to simulate a fluorocarbon loop leak or the radiator to/away from the sun.

Display panel - The display panel, as depicted in figure 9, provides a graphic presentation of the function and status of the fluorocarbon, water, and cabin air systems. In addition to the illumination of the flow paths, there are lighted pushbuttons to indicate discrete failure insertion and meters to indicate system variable values. The following meter indications are provided on the panel.

(a) Water boiler outlet temperature, 30° to 70° F

(b) Desiccant humidity, 0 to 100 percent

(c) Differential pressure LiOH filter A, 0 to 10 inches H_2O

(d) Differential pressure LiOH filter B, 0 to 10 inches H_2O

(e) Crew CO_2 partial pressure, 0 to 20 millimeters Hg

(f) Cabin temperature, 40° to 90° F

(g) Compressor outlet pressure, 0 to 20 psia

(h) Heat exchanger outlet temperature, 30° to 70° F

(i) Radiator outlet temperature, -120° to +120° F

(j) Heat exchanger inlet temperature, 20° to 120° F

(k) Fluorocarbon accumulator pressure, 0 to 300 psia

(l) Water pump outlet pressure, 0 to 100 psia

(m) Water accumulator pressure, 0 to 100 psia

A key feature of the ECS simulator, as implemented in the OSM breadboard, is the capability to induce failures into the system. These system failures are inserted from the display panel. Ten pushbutton switches and two manually set potentiometers with the following functions are provided for local fault and system boundary condition insertion.

(a) Radiator to sun

(b) Radiator away from sun

(c) Water leak activation pushbutton plus leak rate potentiometer

(d) Fluorocarbon leak activation pushbutton plus leak rate potentiometer

(e) Water pump failure

(f) LiOH filter A contaminated

(g) LiOH filter A CO_2 saturated

(h) LiOH filter B contaminated

(i) LiOH filter B CO_2 saturated

(j) Air compressor failure

3.2.3.5. Automated ECS management

In addition to the automation of the ECS process control loops previously described, system management functions are also under computer control.

<u>System status monitoring</u> - The ECS simulator, working with the DMS, provides three distinct levels of status monitoring. One level involves monitoring of the various measurement points by the measurement system with reporting to the local processor on an exception basis. Another involves local processor access to the measurement points as a result of

(a) An exception report

(b) A local processor requirement

(c) A status request from the DMS

Another involves reporting by the local processor to the DMS (and thus to the crew) as a result of

(a) A C&W condition

(b) A data request from the DMS resulting from a crew command

These three levels of status monitoring distribute the load between the measurement system, the local processor, and the DMS in such a way as to adequately advise the crew of status without overloading the higher level processor and without saturating the crew with trivial status data.

<u>Onboard checkout and malfunction analysis</u> - The ECS simulator local processor, working with the DMS, provides a full capability for checkout of the valve and sensor portion of the ECS simulator. The necessary status monitoring logic and software is implemented in the DMS, the ECS simulator, and the D&C system to allow the crew to perform checkout and troubleshooting functions from the cockpit. All fault conditions inserted at the display panel can be isolated by the crew. The ECS simulator, as implemented, does not provide the capability for onboard checkout of the local processor or adapter.

<u>Redundancy switching and configuration management</u> - Mechanization of the ECS simulator demonstrates that the status tracking, management decision, and configuration switching functions necessary to manage the three dual redundant elements of the ECS simulator can be accomplished by way of computer control. The primary burden for redundancy and configuration management on the ECS simulator lies with the local processor; however, status data from the subsystem is available to the DMS for management from the DMS with crew assistance. Although the magnitude of the management problem is not comparable to that required for a total data management system, the ECS simulator implementation shows that this function can be performed effectively in a system having many other functions to perform and with interaction between a local and central processor.

<u>Onboard maintenance and malfunction correction</u> - The ECS simulator provides the necessary redundancy and control to demonstrate the concept of automatic onboard maintenance. The local processor, through the measurement system, can sense failures or out-of-tolerance conditions and can issue commands to reconfigure for operation through a redundant loop or for performing cyclic functions such as venting of desiccant water overboard. A manual maintenance feature was included in the ECS simulator by providing for manual replacement of the simulated LiOH filters.

<u>Status recording</u> - The OSM breadboard provides for recording of any bus traffic if the record bit is set in the output control word. The ECS simulator can, therefore, have any of its output recorded by the CSR subsystem. In addition, the DMS can set the record bit on any command it sends the ECS simulator and have that message recorded. The DMS/local processor/CSR concept provides considerable capability for recording status data. The actual logic that controls recording is a function of the software in the DMS and the local processor.

3.2.4. <u>Onboard systems management breadboard software summary</u>

The OSM breadboard software development was affected by several factors. The most important of these was the diverse number of computers used. The second was the development of subsystem software by unique organizations, which required efficient communication and thorough, yet flexible, documentation. Finally, the availability of software development tools, which ranged from a comprehensive set for the IBM 360/44 to a rudimentary loader for the P703, was a constraining factor.

Although five separate computers were used in the OSM breadboard, the available computing power was not fully used. The use of the two prime computer resources (kilo-operations (kops) and memory) is shown in figure 10. This figure shows the available capability and the used capability for each resource. The large use of the memory on the IBM 360/44 and the Micro 500 was occasioned by the use of high-level language coding in the IBM 360/44 and by the limited memory size in the Micro 500. It is interesting to note that the DDI 520 was shared by two subsystems, which posed some interesting software development and timing problems.

The software was developed by the organizational entity responsible for the user subsystem. The software for each computer except the IBM 360/44 was prepared in assembly language. However, the 4 Pi EP software was assembled and could be executed on the IBM 360/44. The compatibility of these two computers was very useful in this respect. The software for the IBM 360/44 was prepared in FORTRAN to execute under the standard IBM operating system. Although assembly language programs would have been more efficient, the necessity to quickly modify D&C software formats to be manipulated made it imperative that a high-level language be used to meet the schedule. Specialized routines such as the display language assembler were available as a software development tool.

The software for each computer is characteristic of the functions each computer is required to perform. The IBM 360/44 as the D&C processor was primarily devoted to the man-machine interface. The IBM 4 Pi EP software was devoted to control of the bus traffic, to message interpretation, and to generating display format requests. The software of each subsystem was devoted to status monitoring, checkout functions, redundancy switching, and subsystem functions. In addition, the software of each subsystem was capable of handling all bus traffic for which it was responsible.

4. DISCUSSION OF SHUTTLE STUDIES

Although many lessons were learned from the evaluation of both the quad-redundant G&C system and the OSM breadboard, the experience has indicated that no significant technology problems exist in mechanizing fully automatic systems capable of detecting subsystem malfunctions and of reconfiguring redundant components to provide a serviceable path. Such systems also would have the inherent capability to control the configuration of the vehicle and the moding of the subsystems automatically in accordance with the requirements of the mission phase, and to relieve the crew workload by carrying out routine maintenance operations.

In some of the advanced vehicles discussed earlier, systems of this type would appear to be necessary because of the periods of unmanned operation (e.g., modules of the space station before orbital assembly) on unmanned free-flying shuttle payloads. However, the shuttle phase B studies contracted by NASA (references 9 and 10) coupled with the experience gained in the breadboarding have provided some insight into the possible management problems and the costs of implementing such systems in a vehicle such as the space shuttle.

The two MSC breadboards represented only a very small portion of the onboard equipment that might constitute the total complement of systems in the space shuttle. Only partial redundancy was implemented in the subsystem components, and no measurement points were monitored in the computer and the D&C hardware. Extrapolations of the software requirements for these limited systems into the requirements for an all-up shuttle vehicle suggest that extremely capable computer systems with a large total main-frame memory requirement would be essential. Although possibly not outside the range of up-to-date machines and the power and weight budgets for the onboard computer system, the requirements in themselves are indicative of an extremely demanding management task in providing for the massive system integration role and subsequent hardware/software verification. It must be recognized that the computer control of critical command and control paths of nonavionics systems creates difficult problems in the verification process.

The multiple identical redundancy technique usually proposed for these critical paths may be an adequate solution to the problem of hardware unreliability but is susceptible to the problem of undetected generic software errors failing all paths simultaneously. The probability of this type of software failure can be made very low by using extensive verification techniques, but these techniques will require the early provision of nonavionics system hardware to the avionics integration laboratory.

The integration of avionics hardware in several past programs has been somewhat demanding in terms of manpower, facilities, and time. The addition of nonavionics systems to this integration process can be expected to increase these requirements considerably.

In keeping with the necessity within the very austere shuttle funding availability to avoid any undue cost or schedule risks, it has been decided that all flight critical commands for the shuttle vehicle should be hardwired and automatic system reconfiguration employed only where dictated for dynamic reasons. Exceptions would be made only where the benefits to be gained clearly outweigh any potential risk.

A computerized performance monitor system that does not provide switching of critical control paths but has the capability to support all mission phases including ground operations is proposed for the space shuttle.

5. PROPOSED SHUTTLE PERFORMANCE MONITOR SYSTEM

The performance monitor system supports the maintenance of vehicle and subsystem status throughout all phases of the space shuttle operation. The following are defined as the principal functions of the performance monitor system. Additional functions such as alternate mission planning and malfunction analyses may be implemented in the performance monitor system as found necessary and feasible, but they do not currently dictate hardware design.

(a) Passive subsystem monitoring

(b) Vehicle configuration support

(c) System status recording

(d) Onboard checkout

In carrying out these functions, the performance monitor system will acquire data from the subsystem and perform the data processing required to present information to the flight crew by way of the onboard displays. The measurement parameters also will be provided in the correct format for transmission on the digital data spacecraft-to-ground link to support flight operations and ground checkout.

Subsystem performance will be monitored continuously by comparing selected parameters with limits stored on board. Anomalous behavior will be indicated to the flight crew and recorded onboard for future support of maintenance operations. Additional recording capability provides for storage of line-replaceable-unit status, engine, voice, time correlation, and crash data.

Critical measurement parameters will be designated for C&W functions. The primary C&W functions will be hardwired, and all necessary processing and display will be accomplished by dedicated hardware. Backup C&W will be provided by the performance monitor system through computer assistance.

Information stored in the performance monitor system computer memory can be displayed to assist the crew in selecting proper system configuration for normal, optional, and contingency modes. The performance monitor system also will monitor consumables for quantity remaining, rate of consumption, and projected future requirements.

The status of all switches, breakers, and subsystem modes defining the selected current configuration will be maintained at all times.

Stimuli and commands necessary to support prelaunch checkout will be introduced into the vehicle subsystems by means of an onboard checkout command decoder. The response of the subsystem stimulated during prelaunch checkout will be monitored by means of pulse-code-modulated downlink to the ground checkout station.

5.1. System Block Diagram

A block diagram of the performance monitor system is shown in figure 11. The main components of this system are

(a) The performance monitor computer

(b) The stored program processor

(c) Remote acquisition units (RAU's)

(d) The maintenance recorder

(e) The loop recorder

(f) The voice recorder

(g) The crash recorder

(h) The C&W logic unit

(i) The checkout command decoder

(j) The mass memory unit

6. CONCLUDING COMMENTS

As can be seen from the preceding description, a considerable capability and growth potential is envisaged for the shuttle performance monitor system. Although complete independence of ground-support facilities is neither anticipated nor considered desirable, it is expected that the performance monitor system can, in the operational stages of the program, develop an onboard checkout capability consistent with the original shuttle goals. Similarly, as the program matures, the performance monitor system can be expected to assume more onboard system management tasks as mission experience and requirements dictate and, thereby, to reduce both flight-crew workload and real-time ground-support requirements.

REFERENCES

1. ANON., 1972, "Space Shuttle Program Request for Proposal 9-BC421-67-2-40P," NASA Manned Spacecraft Center.

2. WITHEE, W. W., 1972, "Research and Applications Modules (RAM) Phase B Study, Executive Summary," report GDCA-DDA72-009, Convair Aerospace Division of General Dynamics.

3. ANON., 1972, "Information Management Advanced Development," reports SD72-SA-0114-I through SD72-SA-0114-V, volumes I through V, North American Rockwell Space Division.

4. PETRIE, David M., 1967, "Automatic Flight Management of Future High-Performance Aircraft," paper presented at the 4th Annual Meeting and Technical Display, AIAA (Anaheim, California).

5. ANON., 1970, "A Fault-Tolerant Information Processing System for Advanced Control, Guidance, and Navigation," report R-659, MIT Charles Stark Draper Laboratory.

6. HOPKINS, A. L., Jr., 1971, "A Fault-Tolerant Information Processing Concept for Space Vehicles," paper 67-847 presented at the IEEE Transactions on Computers.

7. LEE, W. Thad, 1972, "Experience Gained in Programing for Redundant Systems," report MSC-06800, Internal Note MSC-EG-72-9, NASA Manned Spacecraft Center.

8. PROCH, George E., 1972, "Guidance and Control Division Redundant Flight Control System Breadboard Design Report," report LEC/HASD 643D.21.961, Lockheed Electronics Company, Inc.

9. ANON., 1971, "Space Shuttle Program Phase B Report," report MSC-03321, SD71-140, volumes I through XII, North American Rockwell Space Division.

10. ANON., 1971, "Space Shuttle Program Phase B Final Report," report MDC-E0308, volumes I through III, McDonnell Douglas Corporation.

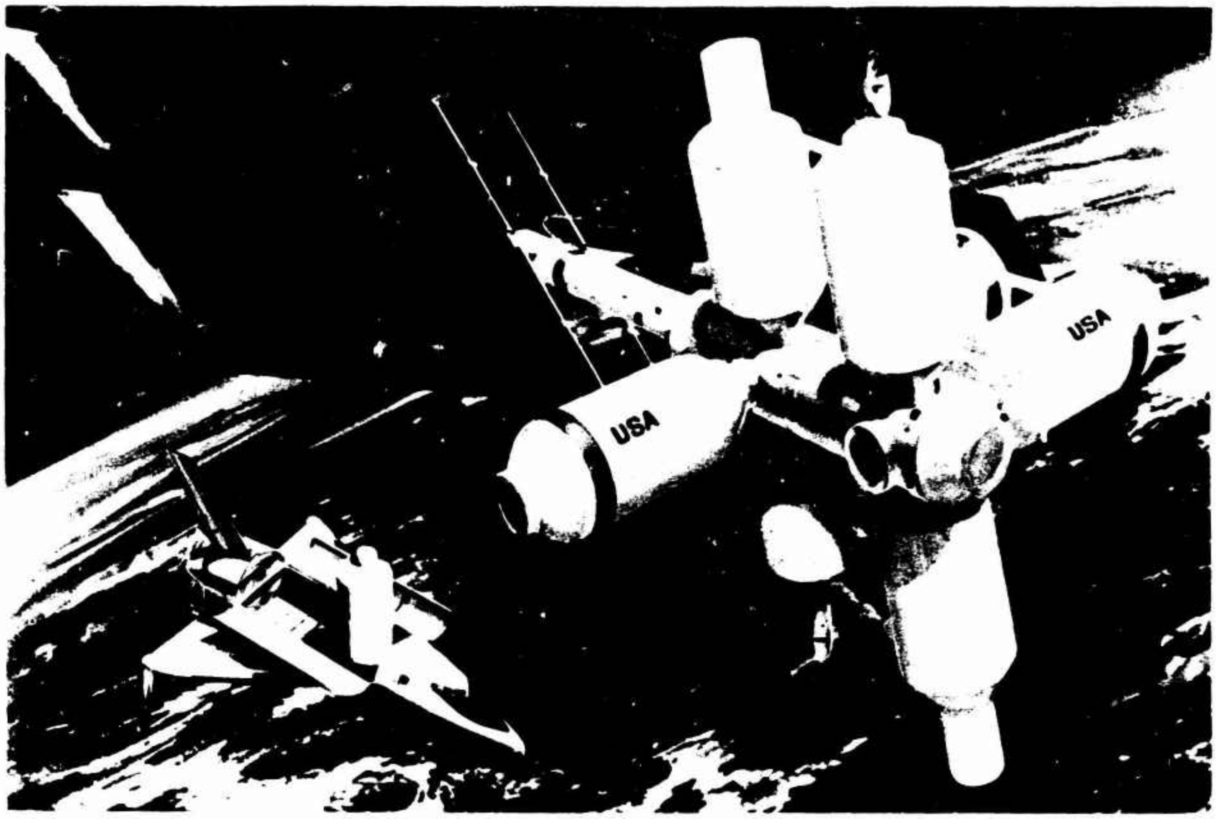


Fig.1 Research and applications module

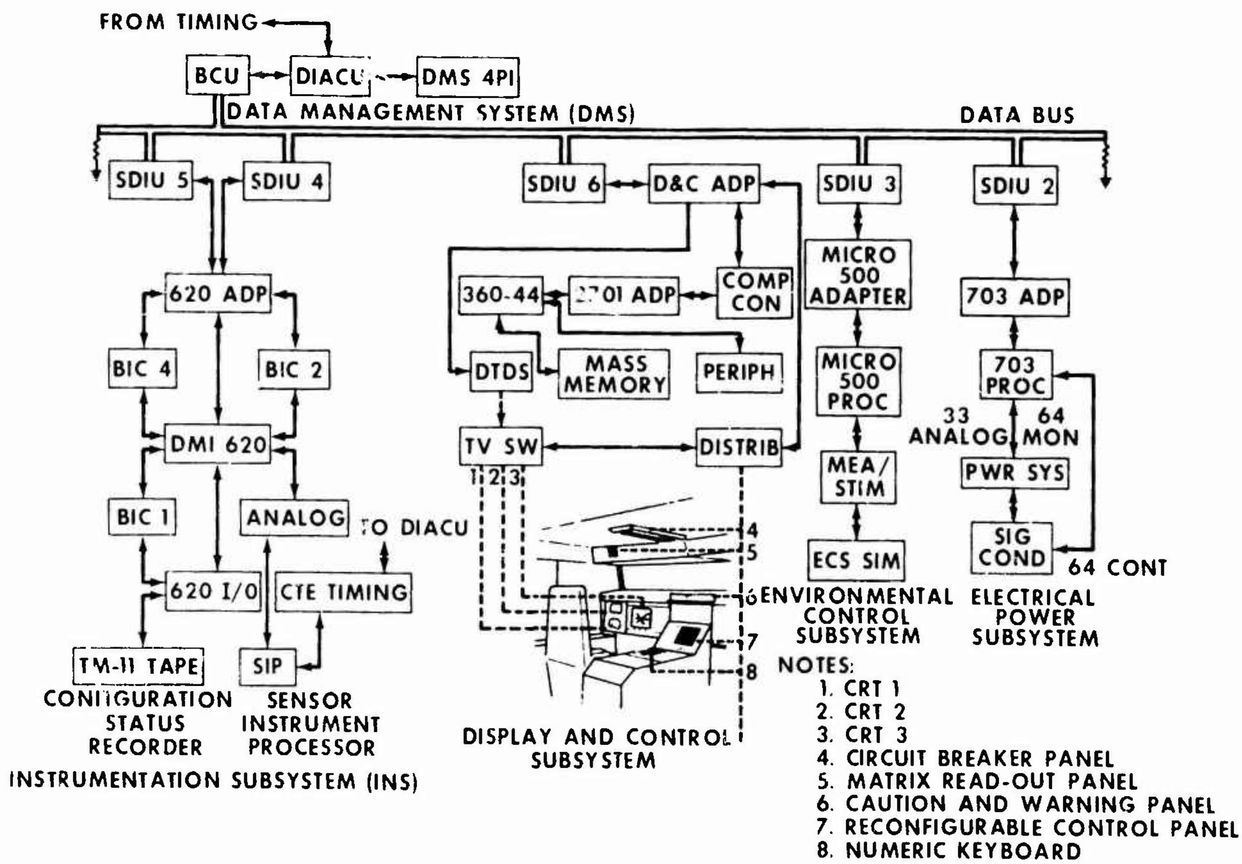


Fig.2 Onboard systems management breadboard

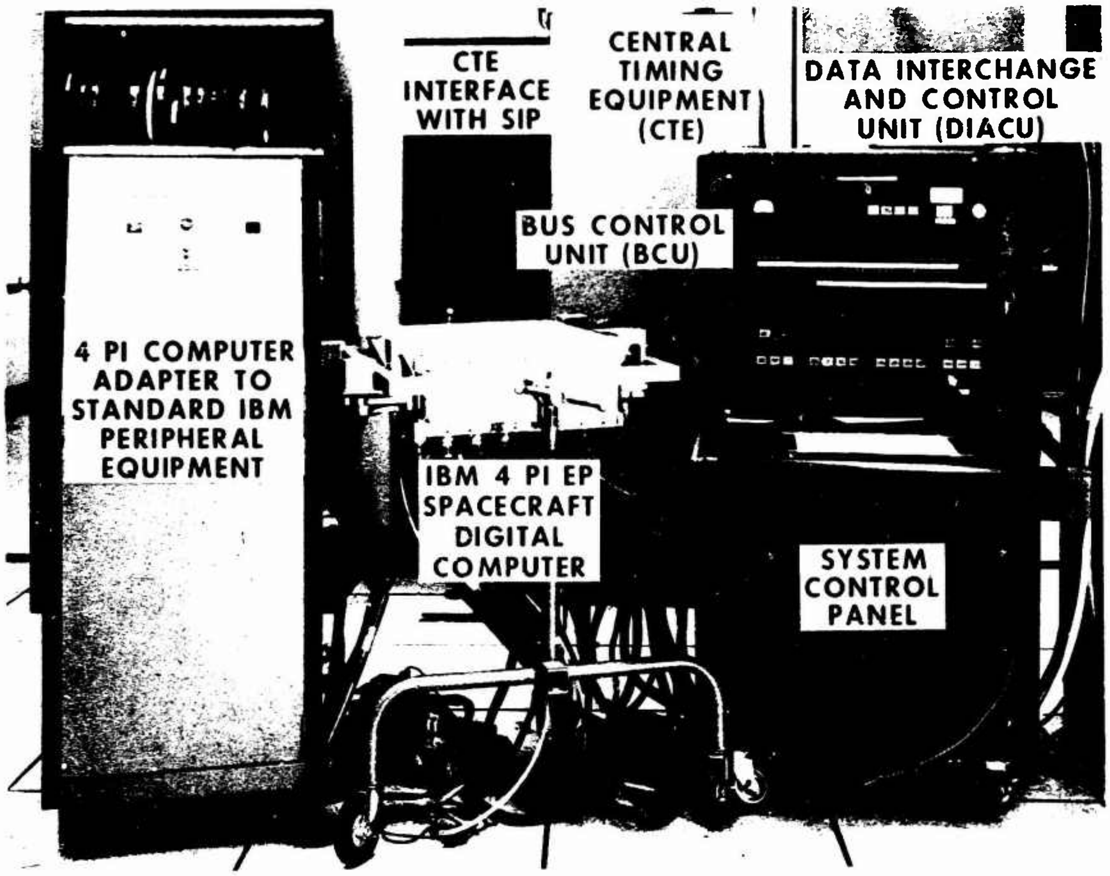


Fig.3 Data management system computer data bus and timing equipment

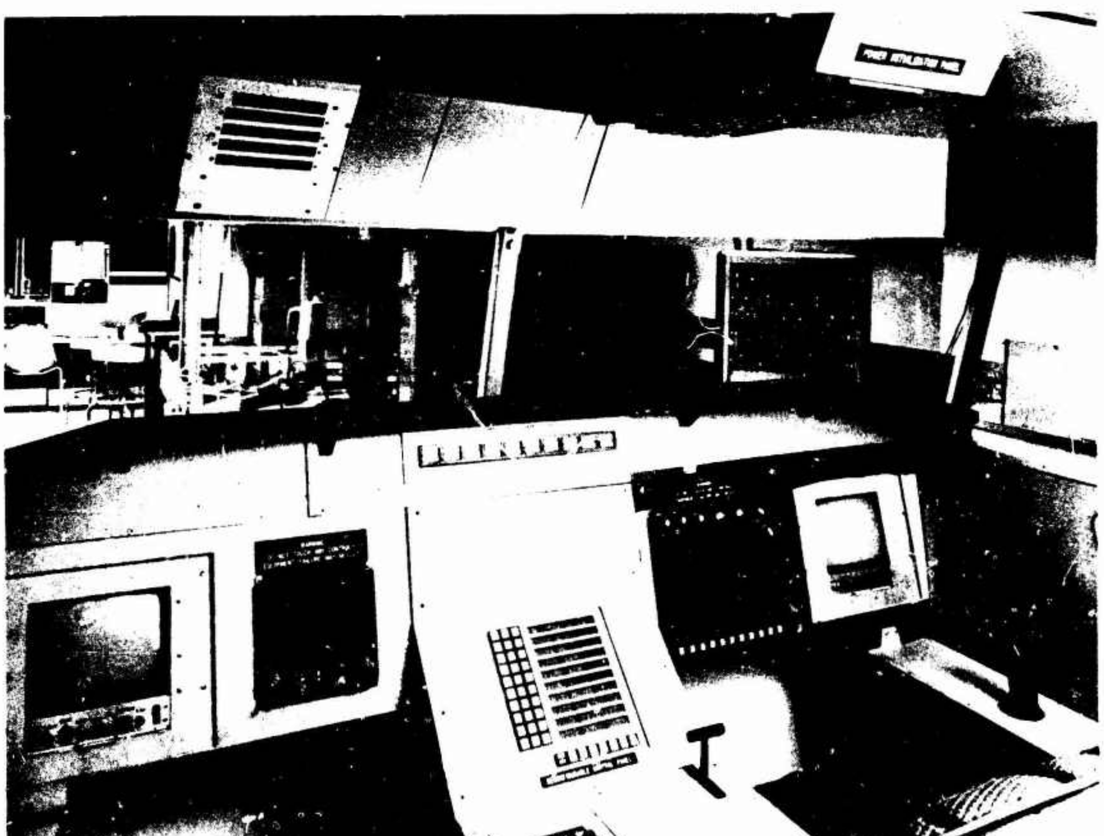


Fig.4 Display and control subsystem display devices mounted in evaluator cockpit

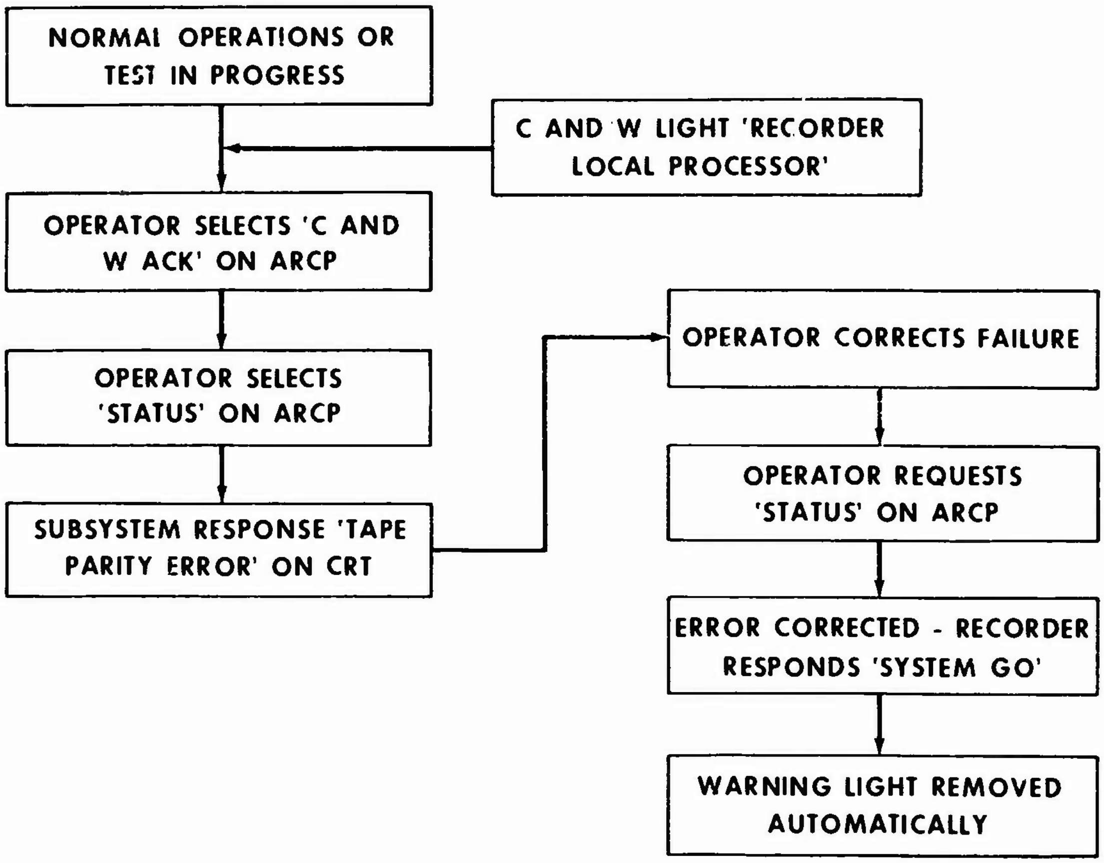


Fig.5 Display and control troubleshooting procedure flow chart

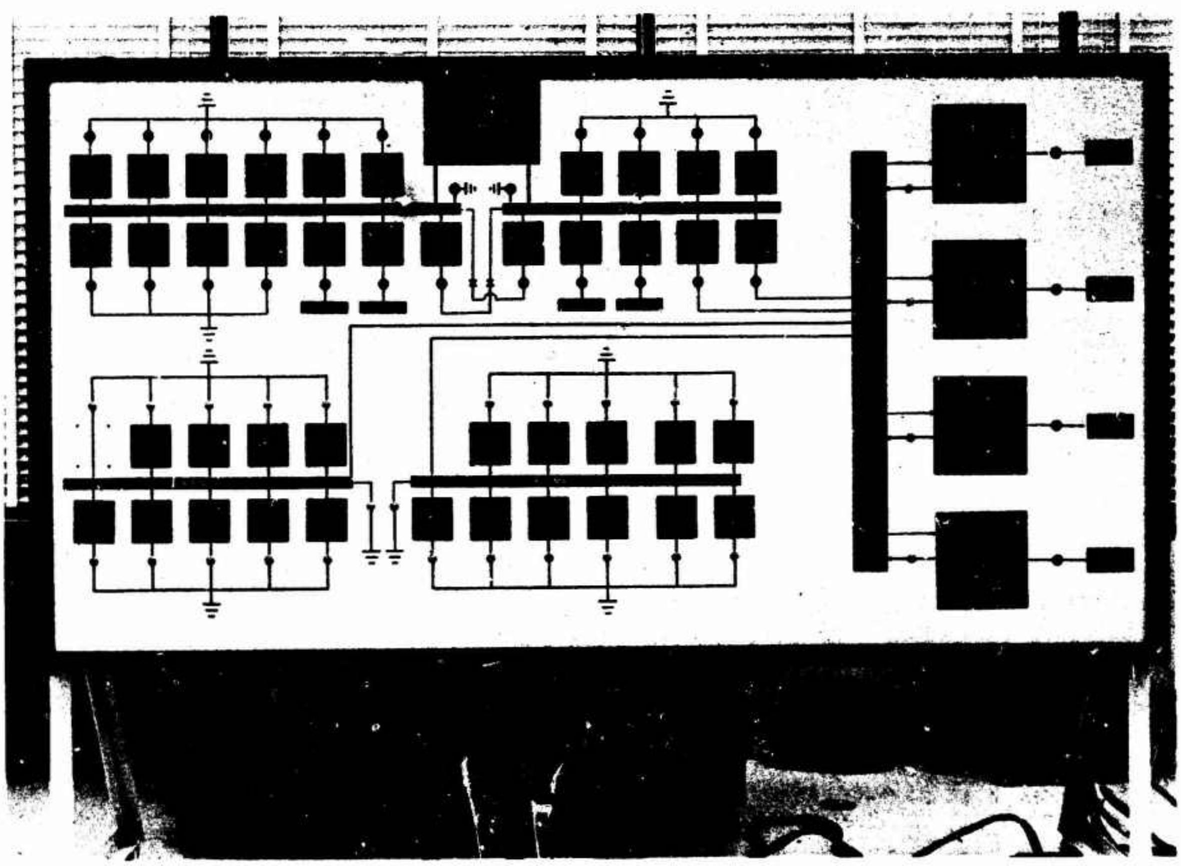


Fig.6 Electrical power subsystem display panel

Fig.7 Electrical power subsystem/processor and panel

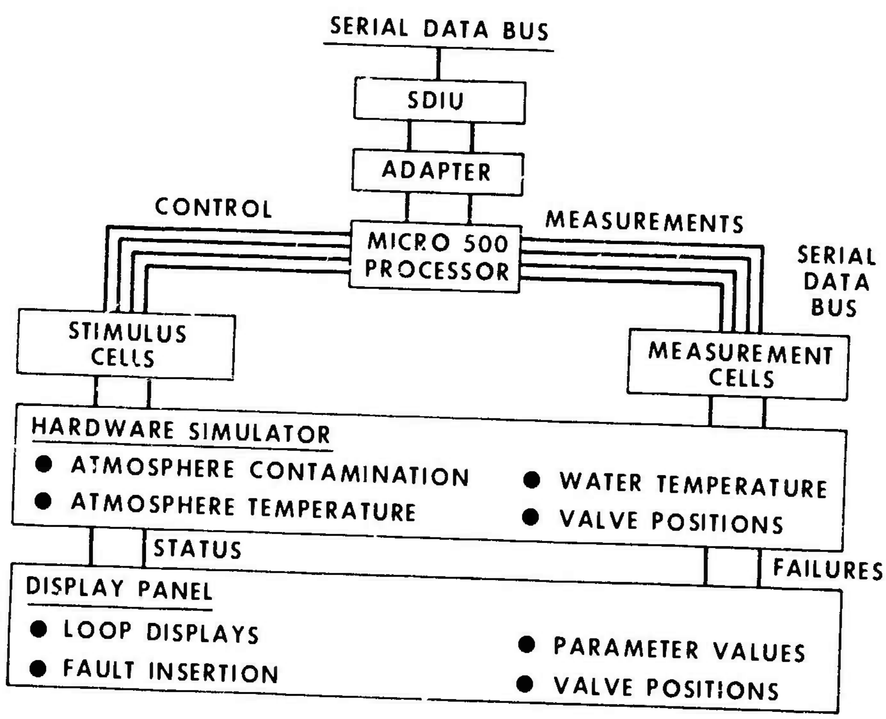


Fig.8 Environmental control subsystem functional diagram

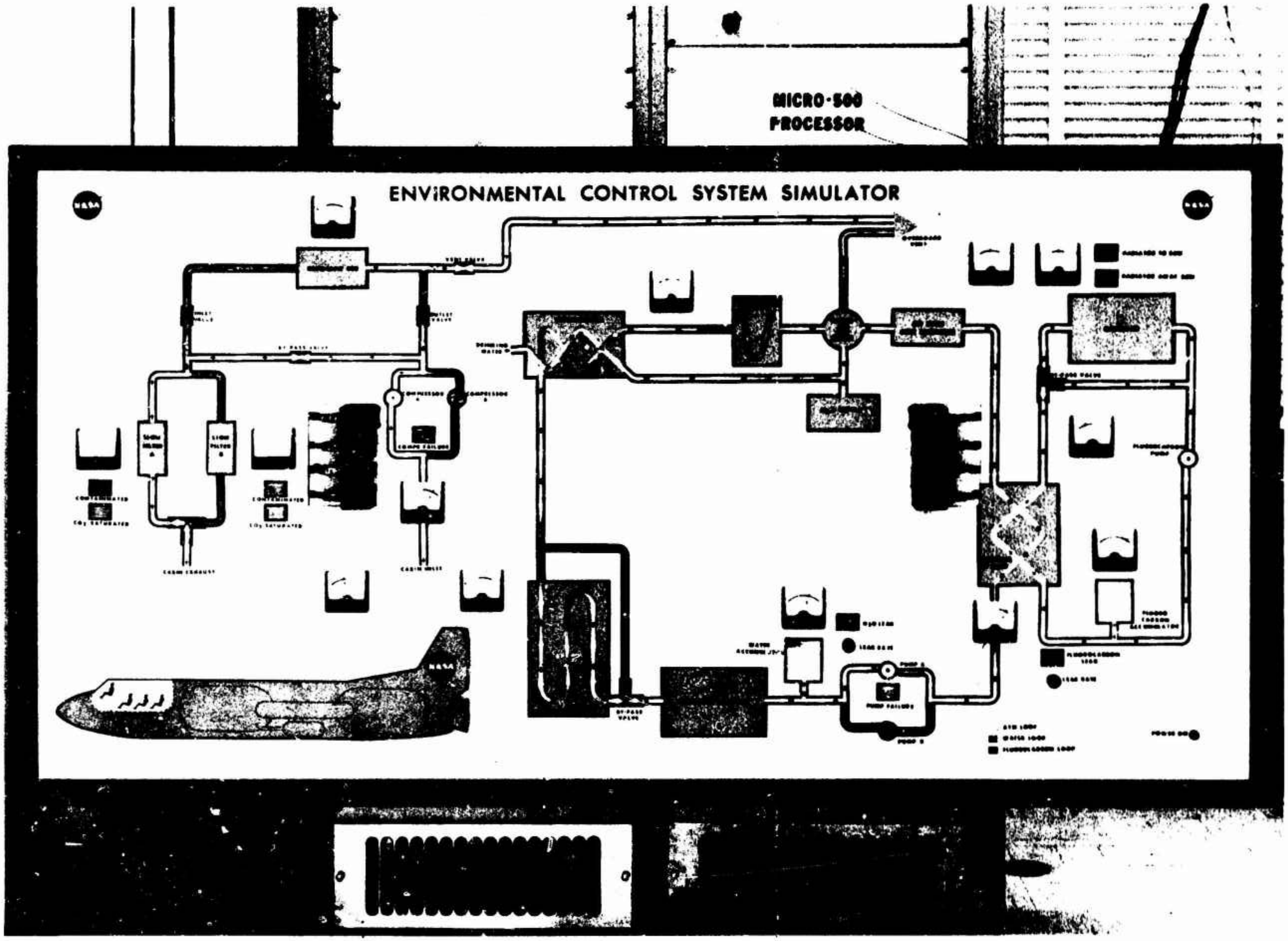


Fig.9 Environmental control subsystem simulator/panel

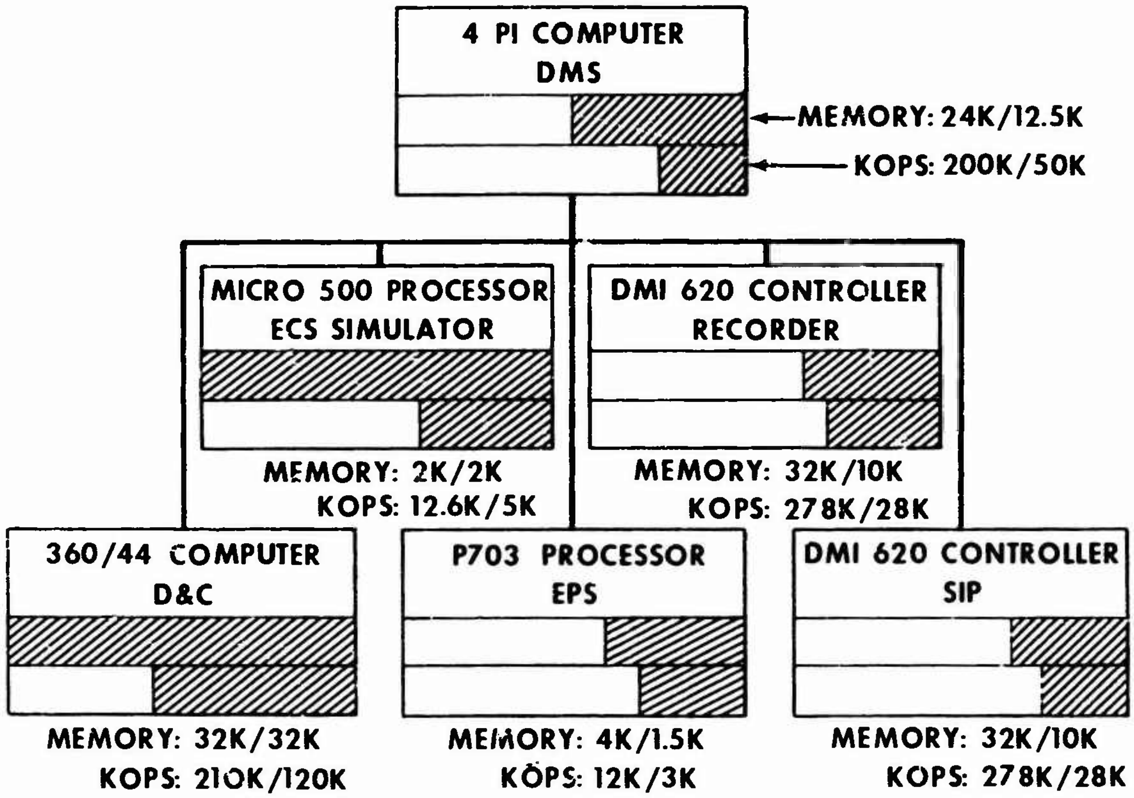


Fig.10 Onboard systems management breadboard computer utilization

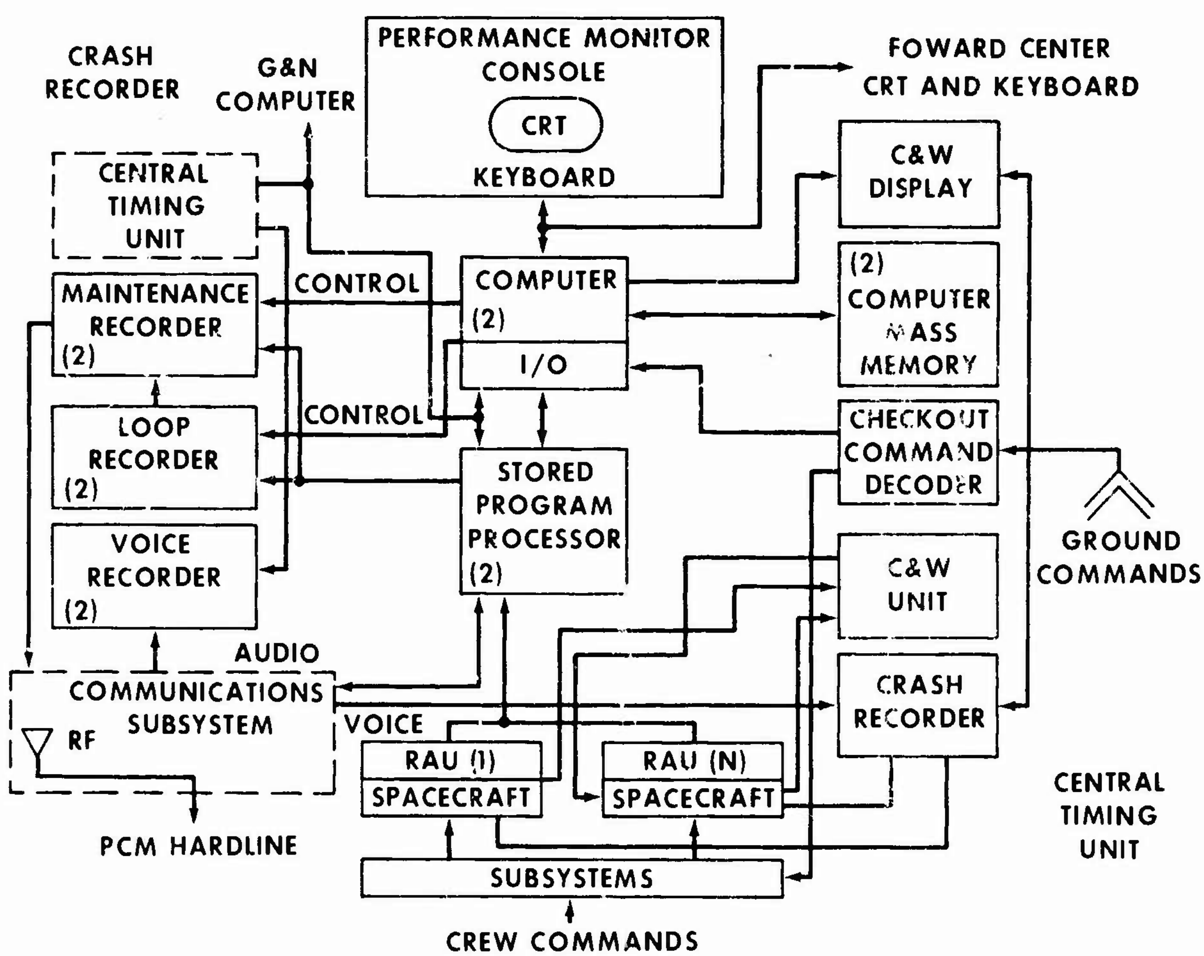


Fig.11 Performance monitor system

A GENERAL PURPOSE COMPUTER FOR SPACEBORNE APPLICATIONS

S. BOESSO and R. GAMBERALE
SELENIA S.p.A.
Km. 12.400, Via Tiburtina, 00131 Rome
ITALY

SUMMARY

This paper describes a modular expandable Computer System studied by Selenia for ESRO (Contract ESTEC 1115/70/AA) for a wide range of possible space missions.
The main goals in designing the computer have been maximum flexibility and reliability, minimum weight and power consumption (using proven materials and technologies) and growth capability to fit future mission requirements.
The study resulted in a stored program, 16 bit, parallel machine, with microprogrammed control which allows full arithmetic and logic capability.
Input-output includes program-controlled and direct-memory-access channels.
A plated-wire memory, expandable up to 64 K words, is used for programs and data.
The basic configuration, with a 4 K memory, has a power consumption around 10 W, a weight of about 5 Kg. and allows a probability of success higher than 0.9 for a one-year mission.
The hardware has been studied in modular form in order to allow different system configurations, including the possibility of multiprocessor arrangements.
Simulator and assembler programs (to be run on a IBM 360/50) have been developed during the course of the study.
This paper presents the main points of trade-off for system design and gives a description of the basic Computer units at the functional block level.
An engineering model of the Computer is presently under development.
The Computer Study was carried out by a Selenia Team including, beyond the authors, the following people: A. Bellini; S. Bottalico, M. Canevari, G. Cicerchia, B. Di Marco, P. Fanchiotti, F. Marcoz, R. Somma.

1. INTRODUCTION

The Computer described in this paper was studied under an ESRO (European Space Research Organization) contract, having as aim the definition of a general-purpose machine for use on board a spacecraft.
The original requirements, applied to a basic configuration with a 4K memory, were the following:

- power consumption: around 10 watts
- weight: around 5 Kg.
- volume: around 5 liters
- probability of failure: less than 10% for 12 to 18 months in orbit
- word length: 16 bits
- technological study and hardware definition to be based on the use of a well proven, presently available technology.

Although the machine was originally intended for use on unmanned satellites, to perform various and different tasks, its characteristics and, in particular, its expandability, make it attractive for a much wider range of applications, as those required in manned spacecraft.
The computer is organized around a Transfer Unit (TU), which acts as a traffic controller of the flow of data to and from the main memory.
The units connected to the TU may be, in principle, any combination of CPU's and Direct Memory Access (DMA) channels.
Requests of memory access are honored according to a predermined priority.
The simplest machine which can be devised, and which is suitable for small spacecraft, has a single CPU and a 4 K by 16 bit plated-wire memory, which can also be directly accessed via the TT & C (Telemetry, Tracking and Command) system (Fig. 1). This configuration, however, can grow up, still using the basic building blocks, to a multiprocessor system. In all cases maximum memory capacity is

64 K by 16 bits. The Memory may be organized as one single block, shared by all CPU's or DMA's or divided into more banks, each bank being shared by all CPU's or DMA's (fig. 3). As the CPU has a microprogrammed control, its instruction repertoire can be changed, if needed, without hardware modifications, by simply replacing the control memory; therefore a multiprocessor system can be devised where each CPU is optimized for a particular task by means of a dedicated instruction set. Whenever required by reliability criteria, spare units (CPU, TU, Memory) can be provided, which would be kept normally off, being switched in only in case of failure of one of their counterparts.
An analysis of the possible tasks of this computer is outside the scope of this paper, but, just to mention one application, it could be employed for the data management of a Research and Applications Module (RAM), where the computer would be devoted to control, monitor and check all subsystems and experiments and drive a display system for interface with the RAM crew.
The characteristics of the computer building blocks will be described in this paper and design trade-offs which have led to particular choices will be justified. The instruction set which will be given is that of the basic CPU, for which a simulator and an Assembler program exist.

2. GENERAL CHARACTERISTICS

Table 1 summarizes the main functional characteristics of the computer. Parallel operation was selected in order to obtain the speed required by the tasks. A 2.8 µs add time can be considered a good figure for a low-power computer.

TABLE 1

On-board Computer (OBC) Characteristics		
- Type	General purpose, stored program	
- Operation	Binary, parallel	
- Word length	16 bits	
- Arithmetic	Two's complement, fixed point	
- Control	Microprogrammed	
- Memory	Expandable up to 64 K	
- Input-Output	Program-controlled and Direct Memory Access channels (DMA)	
- Program Interrupt	16 Expandable request lines	
- Instruction set	Full arithmetic and logic capability	
	Shift/Rotate provision	
	Inter-register transfers	
	Conditional/Unconditional branching	
- Addressing	Single-address instructions, fixed page, index, indirect, immediate	
- Speed	1.4 µs cycle time	
- Typical Execution Times	Load/Store	2.8 µs
	Add/Subtract	2.8 µs
	Multiply	28.0 µs
	Divide	51.8 µs

3. CENTRAL PROCESSING UNIT

3.1. Design Considerations

The system study of the CPU was carried out taking into account on one side the operational requirements derived from an analysis of a certain number of typical satellite tasks and on the other side the hardware constraints.
Because of the stringent requirements in weight, volume, power consumption and reliability, every system decision taken during the study had to be checked against the physical constraints.
This was particularly true for the choice of a microprogrammed control, that is presently applicable because of the availability of fast, low power and highly integrated ROM's.
A first design approach based upon a gradual degradation of CPU performance in failure condition, taking advantage of the intrinsic features of a microprogrammed control, seemed to be attractive, as, for example, in case of a register failure, the related functions could be performed by another one if the microprogram were properly changed.
This would give the CPU the capability of a gradual degradation of the performances apart from the other benefits deriving from the inherent flexibility of such a control type.

A preliminary examination showed, however, that many difficulties arise when a reprogrammable Control Memory is considered.
The use of a semiconductor ROM (Read Only Memory) is not suitable because its content cannot be changed during flight.
On the other hand the use of the Main Memory as a back up storage for the microprogram would have the disadvantage of an increase of power consumption and execution time, because an additional Memory access for every 16-bit control word would be required.
Considering two control words (32 bits) per microstep and 3 microsteps in the average to implement an instruction, this solution would have produced an increase of the execution time and of the power consumption of about a factor of 6, which is not acceptable for this application.
The use of a Random Access Control Memory, reprogrammable from the ground was also considered as a suitable back-up in case of a failure within the ROM. This solution, however would have presented all the problems connected with the information loss in case of a power failure.
For the above reasons, the selected approach has been the design of a very simple CPU, without internal redundancies,which can be backed up by another identical unit if required by the reliability constraints.
The redundant CPU will be kept off and will be switched on from the ground, under detection of a failure signal, in place of the failed unit.
The CPU comprises three main blocks (see fig. 4):

- Operating Unit
- Control Unit
- Program Interrupt Unit

exchanging between them control and condition signals.
The minimum number of registers compatible with good machine performance has been provided.
Communication with the external world is provided by:

- data input, data output and address buses connected to Memory via the Transfer Unit
- data input, data output and address buses which connect the CPU to the external world via Input/Output interface circuits.
- program interrupt requests coming either from peripherals or DMA: 16 interrupt lines are provided with possibility of masking 12 of them.

3.2. Instruction Set

For the definition of the set of instructions the following approach has been adopted:

A certain number of different computer tasks were considered such as experiment management data compression and formatting, attitude control, telemetry and telecommand functions, housekeeping and checkout, etc.
These tasks were thought to be representative of a wide range of space missions without limiting the use of the computer to special purpose applications only but, on the other hand, without imposing too heavy requirements which would have resulted in a computer out of the design constraints.
The tasks were analyzed in terms of "elementary functions" for which the related recurrence frequency was evaluated.
The functions were then grouped into three classes:

- high-recurrence elementary functions: transfer, arithmetic, loop control
- medium-recurrence elementary functions: interrupt routine control, shift, test
- low-recurrence elementary functions: I/O functions, engaged routines control, test on flags, operations on sub-fields of memory words, increment, return jump,

A first set of instructions was tentatively studied, trying to optimize the execution of those functions having the highest recurrence frequencies. Sample programs were then written using this instruction set, for a significant number of the above mentioned tasks, and these programs were checked against the total memory occupancy and the recurrence frequencies of each instruction. As a consequence the set was modified by: deleting some instructions which were scarcely used, replacing sequences of instructions which were frequently used with one instruction only. The effectiveness of the modifications was eventually tested by rewriting the same sample programs and rechecking them.
This set was found acceptable.
The instruction use a 16 bit word. They have been subdivided in the following three groups.

A) Memory-Reference Instructions

This group contains 29 instructions which refer to an operand stored in Memory. If the dictated operation involves two operators the second one is taken from a machine register.

The format is the following:

15 … 11	10	9	8 … 0
OP	I	X	AF

where:

OP Operation Code
I, X Indirect, Index Indicators
AF Address Field

Each five-bit operation code identifies one out of the 29 instructions of the group. Two codes are reserved for the input-output instructions while the code "all zeros" defines the non-memory-reference instructions.
The two bits I, X define the addressing mode: direct, indirect, direct with indexing, indirect with indexing.
The memory is addressed according to a fixed page organization. This means that the 16-bit memory address is obtained using AF as line address inside the page and deriving the page address from the 7 most significant bits of the Program Counter.
This is the operand address if direct addressing is specified (I = 0, X = 0); if indirect addressing is requested (I = 1), the operand address is fetched from memory at the location specified by the address already determined. If indexing is specified (X = 1), the operand address is modified by adding to it the contents of the Index Register.
Only one Index Register is provided. A trade-off was made between the use of one or two index registers, applied to the handling of data contained in a two-entry array; the average gain in speed of execution and memory occupancy, obtainable with use of two index registers was found not to be such as to justify the increased hardware complexity.
In table 2 the memory reference instructions are presented. The meaning of the symbols is as follows:

M	indicates the contents of a memory location at the address specified by the Address Register
→	indicates transfer
A, B	are two arithmetic registers
X	is the index register
R	is the remainder of a division
+, -, x, ÷	indicate the four arithmetic operations
AND, OR, EXOR	indicate the logical "and", "or" and "exclusive or" functions.

Table 2
Memory-Reference Instructions

LDA	M → A	ADB	B + M → B
LDB	M → B	SBA	A - M → A
LDX	M → X	SBB	B - M → B
STA	A → M	MPA	A x M → A, B
STB	B → M	DVA	(A, B) ÷ M → B, R → A
STX	X → M	AND	A AND M → A
JMP	UNCONDITIONAL JUMP	ORA	A OR M → A
JPL	JUMP IF A < O	EXO	A EXOR M → A
JPG	JUMP IF A > O	ADX	X + M → X
JPZ	JUMP IF A = O	INC	M + 1 → M
LNK	SAVE PC AND JUMP TO SUBROUTINE	DEC	M - 1 → M
TXE	SKIP IF X = M	CNG	A → M, M → A
TXN	SKIP IF X ≠ M	LDK	AF → A, with sign extension
ISZ	M + 1 → M, SKIP IF M = O	AAK	A + AF → A, with sign extension
ADA	A + M → A		

Immediate addressing is provided for two instructions, LDK and ADK. In these cases the operand is contained in the address field itself: bit 8 is considered as the sign bit and is extended up to bit 15.

B) Input Output Instructions

These instructions perform the transfer of data between CPU and peripheral devices. Two instructions are provided:

- TIA Trasfer Input data from addressed peripheral to A register
- TAO Transfer data from A register to addressed peripheral or execute an external command. Type of operation is specified by external device configuration.

The format is the same as for the Memory-Reference Instructions, but with AF defining the I/O channel number.
Execution of both instructions is performed as follows:

- the readiness status of the external device is tested by means of a Skip Tell Back (SKTB) line;
- if device is ready, SKTB is at logical zero, data and/or commands are transferred and the next instruction is executed;
- if device is not ready, SKTB is at logical one, no transfer takes place, and the next instruction is skipped.

Data transfer is performed using standard bus-bar which comprises:

- 16 bit parallel output bus
- 16 bit parallel input bus
- 9 bit parallel address bus, so that there can be up to 512 external addressable points, where for external point we mean either a register or a flip-flop or a gate in a peripheral.

The main advantage of this solution is flexibility in view of future system requirements. Widely different peripherals can be handled without need for CPU modification.

C) Non-Memory-Reference Instructions

These instructions either operate on data stored in machine registers or execute internal commands to modify the status of the computer.
They can be subdivided as follows:

- Shift operations (6 instructions)
- Inter-Register transfers (7 instructions)
- Internal commands (12 instructions)
- Others (6 instructions)

The format is as follows:

15 – 11	10 – 4	3 – 0
0 0 0 0 0	SOP	SN

where:

SOP Secondary Operation Code
SN Shift Number or internal command identification

SOP is used as the starting address of the execute microroutine in the control memory.

In table 3 the Non-Memory-Reference Instructions are presented. In particular, it should be noted that:

- Algebraic shift means that the vacated positions, on the left side of the word being shifted, are filled with the sign bit.
- Logical shift means that the positions, vacated by the word being shifted, are filled with zero's.
- Circular shift, or "rotate", means that the bits shifted out from one side are shifted into the vacated positions at the other side.

Table 3
Non-Memory-Reference Instructions

SRA	SHIFT RIGHT A ALGEBRAICALLY	HALT	HALT
SLA	SHIFT LEFT A	CCY	RESET CARRY
DRA	SHIFT RIGHT A, B ALGEBRAICALLY	COF	RESET OVERFLOW
DLA	SHIFT LEFT A, B	SCY	SET CARRY
SRL	SHIFT RIGHT A LOGICALLY	SOF	SET OVERFLOW
RRA	ROTATE RIGHT A	SCC	SKIP ON CARRY AND RESET
DXZ	IF X $\neq$ 0: X - 1 $\rightarrow$ X; IF X = 0: SKIP	SOC	SKIP ON OVERFLOW AND RESET
TAB	A $\rightarrow$ B	EPI	ENABLE PROGRAM INTERRUPT
TAX	A $\rightarrow$ X	DPI	DISABLE PROGRAM INTERRUPT
TBA	B $\rightarrow$ A	EMI	ENABLE MEMORY INTERRUPT
TBX	B $\rightarrow$ X	DMI	DISABLE MEMORY INTERRUPT
TXA	X $\rightarrow$ A	LMR	A $\rightarrow$ MASK REGISTER
TAE	A $\rightarrow$ B, SIGN A $\rightarrow$ A	LSR	A $\rightarrow$ STATUS REGISTER
TBZ	B $\rightarrow$ A, O $\rightarrow$ B	SSR	STATUS REGISTER $\rightarrow$ A
CLB	O $\rightarrow$ B	NOP	NO OPERATION
CMA	COMPLEMENTED A $\rightarrow$ A		

The Status Register is a 16 bit register which includes:

- 12 bit Mask Register, used for masking program interrupt requests
- CY, CarrY flip-flop; - EPI, Enable Program Interrupt flip-flop;
- OVF, OverFlow flip-flop; - EMI, Enable Memory Interrupt flip-flop.

3. 3. Operating Unit

It is the Unit performing all the arithmetic and logical functions on data. It exchanges data and address with Memory and external devices, as well as control signals and conditions with the Control Unit and the Program Interrupt Unit inside the CPU.
The block diagram is shown in fig. 5.
The characteristics are summarized below:

- 6 registers
  - A Upper accumulator
  - B Lower accumulator
  - X Index register
  - PC Program counter
  - DR Data register
  - AR Address register
- Parallel operation
- Arithmetic/logic: Add, Subtract, AND, OR, Exclusive OR, Complement Left/Right Shift
- Versatile interconnections
  - between registers
  - to/from memory
  - to/from input/output
- Internal shifting capability for A and B registers to speed up multiplication and double-length shifting.

3. 4. Control Unit

It is the portion of the CPU devoted to the generation of the control signals for the operation of the other units. It is microprogrammed. This means that the control signals are derived from Ready Only Memory (ROM). Two 16-bit ROM locations form a microinstruction. The hardwired control logic of the traditional approach is replaced by a ROM, a highly integrated component yielding:

- flexibility, to tailor instruction set to mission
- easy fault diagnosis, because of simpler logic
- uniformity of hardware, that allows use of LSI.

The detailed characteristics are shown in table 5.

Table 5
Characteristics of the Control Unit

Type	Microprogrammed
Control Word organization	Encoded
Control word length	32 bits (two 16 bit words)
ROM capacity	256 locations by 16 bits
Micro-instruction time	1.4 μsec

The micro-instruction format is shown below:

31 29	28 26	25 19	18 16	15 13	12 11	10 7	6 0
AS	BS	FS	MS	CS	DS	COND	NA
Selection of registers to be connected to ALU inputs		Function Selection for ALU	Memory Control	Selection of register to be connected to ALU output	Auxiliary Functions	Condition Selection	Next Address

Fig. 6 represents the block diagram of the Control Unit.
The operation sequence of the computer is shown in fig. 7. Every block of the flow diagram, which is conventionally called "State", corresponds in fact to a microroutine stored in the Control Memory. Four states can be distinguished:

NORMAL FETCH STATE

If no interrupt request is detected, the instruction is fetched from the Memory at the address specified by the Program Counter, and the Program Counter is incremented by one.

INTERRUPT FETCH STATE

If an interrupt request is detected, an instruction is fetched from a fixed location in Memory specified by the Program Interrupt Unit, and no increment of Program Counter occurs.

MODIFY STATE

If needed, the operand address is modified according to the specified addressing mode (indirect, indexed, indirect with indexing).

EXECUTE STATE

The instruction is executed in one or more steps, using the control signals derived from the proper microinstructions, fetched from the Control Memory.
At the end of the execute state, the address where to fetch the next instruction is put into the Address Register AR.
To start Computer operation, a Start Command is given as a program interrupt, either from the ground, or from on-board equipment.

3.5. Program Interrupt Unit

It is the portion that gives the CPU the capability to dynamically modify the sequence normally stated by the program stored in the Memory, in reply to an external Interrupt Request.
The Program Interrupt Unit performs the following functions in order to process the interrupt requests:

- it stores program interrupt requests from peripherals
- it filters the requests according to a mask
- it resolves priority conflicts
- it generates an Interrupt Address corresponding to the highest-priority request.

The basic Program Interrupt Unit can process 16 interrupt requests.
The system expandability is in blocks of 4 request lines. Expansion is achieved by adding one or more 4-input cards to the existing circuitry.
Fig. 8 represents the functional diagram of the unit.

4. TT & C INTERFACE UNIT

It consists of two blocks:

4.1. DMA Unit for ESRO PCM Standard Command (CMD-DMA)

It has two functions:

A) Transfer of information from ground to Memory, for program and data loading, which is performed in the following steps:

1. By a first command, the 8 most significant bits of memory address are loaded into the Unit.
2. Each of the following commands (24 bit words, serial) contains the 8 least significant bits of the memory address and the 16 bits of the word to be loaded into memory.
3. If more than 256 (2^8) words are to be loaded, steps 1 and 2 are repeated.

B) Execution of computer generated commands, which is performed in the following steps:

1. The command is loaded into the unit by the CPU under program control.
2. The Unit sends the command to be executed, serially to the Command Decoder.

4.2. DMA Unit for Telemetry (TLM-DMA)

It has two functions:

A) Transfer of information from Memory to ground under Command Request, for memory contents verification, which is performed in the following steps:

1. The unit is initialized by Command using two commands which define the initial address and the number of words of memory data block to be transferred.
2. The unit transfers automatically the specified data block from Memory to Telemetry.

B) Telemetry data transmission under Computer control, which is performed in the following steps:

1. The Unit is initialized by the CPU as in A).
2. The Unit transfers data blocks from Memory to Telemetry under Telemetry control.
3. At the completion of block transfer, the Unit alerts the CPU by means of a program interrupt. In this way, the Unit can be reactivated by the CPU for transferring another memory

data block according to Telemetry format and bit rate.

5. GENERAL PURPOSE DMA UNIT

It operates as follows:

1. The CPU, under program control, activates the DMA Unit by defining the initial address of the data block and the number of words to be transferred.
2. The DMA Unit exchanges the block of data between memory and peripheral, by automatic sequence.
3. If desired, the DMA sends a program interrupt to alert the CPU of the end of the block transfer.

6. MEMORY

The Memory stores data and programs. Non-volatility of the stored data in case of power failure was required, and also loss of information during power transients had to be avoided. For the above reasons, a magnetic-type main Memory seems to be the best choice.
A minimum memory size of 4 K was required. Thus, the basic module of memory will have a capacity not greater than this.
Many Memory modules can be added in order to have a bigger Memory, up to a maximum size of 64 K (using a single level of indirect addressing).
In order to reduce power consumption, the Memory contains power-switching circuitry, driven by the Memory Timing, that switches the power ON only when the memory is addressed. Otherwise it remains in standby state with a very low power consumption.
Trade off between core and plated-wire memories was made, considering as parameters the repetition rates of read and write operations, the memory cycle and the ratio between the duration of the "on" state and that of the "off" state. Assuming an average of one memory access every 2 µs with 4 reads every one write and 80% efficiency of power switching, the plated-wire memory was selected because of its lower power consumption, less than half that of a core memory of the same size.

7. TRANSFER UNIT

The Transfer Unit, (TU) controls the access to the Memory Bus according to a predetermined priority. It includes a Transfer Controller (TC), Data and Address Multiplexers and a Memory Protection circuit (Fig. 9). The TU operates as follows: when a unit needs memory access, the corresponding Request line becomes high. The TC resolves conflicts among simultaneous requests and generates an Acknowledge corresponding to the Request having the highest priority.
The unit which has received the Acknowledge may proceed accessing the memory, while all the others, if their Request lines are high, have their clocks inhibited. When a transfer has been acknowledged, the data and address buses of the enabled unit are connected to the corresponding memory buses via two Multiplexers.
At the same time, if a Memory write has been requested, the address is checked against the limits of the protected memory area: if its value falls within the limits, a Program Interrupt is generated and the Memory Write pulse is inhibited.
If some unit is required to be able to write into the protected area (e. g. the TT & C Interface Unit, for Program loading), it can be made to override the memory protection.
The basic TU can handle four units; more blocks can be tied together to expand the system.

8. TECHNOLOGY AND MATERIALS

The choice of electronic components, of materials and of type of packaging was based on a survey on a world-wide production of materials and components suitable for space applications.
As regards the electronic components, a comparison was made among different technologies (Bipolar, MOS, Junction FET) and among different families.
The choice of a technology and of a family has been based on the following parameters:

- possibility of a good degree of integration
- a good degree of confidence on the production lines
- reliability
- a power consumption as low as possible
- a maximum frequency of operation as high as possible
- a good resistance to radiations
- a sufficient confidence that the product will be available in the next 3 to 5 years
- availability from two or more sources, if possible
- suitability for application in space environment.

The bipolar technology has been chosen primarily because of its easy integration and good radiation resistance. The 54L family seems to be the best compromise as regards number of I. C.'s, power and speed; its availability from more sources makes it even more attractive.

In a few cases, where the hardware simplification could be obtained without substantial increase in power consumption, MSI components from the 93L family have also been used.
As regards the choice of the material for the mechanical structure, a magnesium alloy is considered the most convenient, allowing savings of 30% in weight over a structure of aluminium alloy.
A trade-off between a computer realized in a unique box or in different boxes was made, and the conclusion was reached that a unique container is the best solution as it greatly simplifies hardware, reduces noise problems and increases reliability.
The On-Board Computer package is shown in fig. 10 in the configuration with 4 K memory. The box is organized in two parts.
The top part houses the Memory and the bottom part the electronics. As regards the electronics, it was our intention to show that the study goals can be substantially met if a well proven technology, already space qualified, is used; the weight and volume can be significantly decreased, however, if more advanced technologies, yielding higher component densities, are employed.
The modules, or printed boards (see fig. 11), each containing an average number of 12 I. C. 's, are fixed by mechanical compression. Internal connectors are avoided.
The modularity of the Computer is ensured at the level of functional modules (CPU, DMA, TU, etc.) which are contained in an unique box. In this way, different Computer configurations can be achieved using different arrangements of modules, involving only the special purpose design of the container box.
The logic design of the various units, excluding the memory, carried out with 54L and 93L families, results in the following complexity:

- CPU 270 chips, 23 printed boards
- TT & C Interface Unit 63 chips, 6 printed boards
- General Purpose DMA Unit 60 chips, 5 printed boards
- Transfer Unit & Interface circuits 47 chips + 70 double transistors, 10 printed boards

Weight, volume and power figures for Memory have been computed from typical data, taking into account the law of variation of power vs. cycle time supplied by Memory manufacturers. These figures, for a 4 K words x 16 bit module, are as follows:

- Access time 300 nsec
- Max read rate 2.5 MHz
- Max write rate 1.4 MHz
- Power
  - 0.1 Watt standby
  - 2 Watt operating at 5 μsec cycle
  - 5 Watt operating at 2 μsec cycle

  4 to 1 read/write ratio
- Weight 2 Kg. approx.
- Volume 2 litres approx.

The selected machine cycle, i.e. the duration of a microinstruction, is 1.4 μsec.
This value has been determined considering the memory timing and the worst-case logical delay involved in the execution of a micro-instruction.
Assuming that 70% of the machine cycles require memory access, an average memory cycle time of 2 μsec results, and hence a memory power consumption of 5 watts.
The power consumption for the single units is summarized in the following table 6.

Table 6

Unit / Power (W)	CPU	TT & C Interface Unit	General Purpose DMA	Transfer Unit & Interface	Memory
Operation	4.13	0.71	0.43	0.76	5.00
Standby	0.36	0.71	0.43	0.22	0.10

In standby condition the whole CPU, excluding the Program Interrupt Unit, is off, as are the interface circuits; in the Memory, only the control circuits and the data register are on.
The power consumption depends on the configuration, for the basic configuration consisting of:

- One CPU
- One TU
- One TT & C Interface Unit
- One 4 K x 16 bit Memory

The total power consumption is 10.6 Watts, in operation and 1.39 Watts in standby.
Should a longer machine cycle be selected, the average memory cycle time would be stretched accordingly, and the power consumption would decrease as shown in fig. 12.

9. RELIABILITY

On the basis of the hardware configuration, the reliability of the single units was computed; the results are shown in fig. 13.
Taking into account the reliability behaviour of the single units, different back-up redundancies have been considered (see fig. 14) for the basic configuration already defined.

For each back-up configuration, a reliability-versus-time diagram has been plotted. (see fig. 15)
The following assumptions have been made:

- Spare units operate in standby redundancy
- Reliability figures used were obtained considering a duty cycle of 100% for all units

The results of the reliability analysis are summarized in the following Table, columns 5 and 6.

10. PACKAGING

Here below (Table 7) a breakdown is given of weight, volume and reliability of all the six possible configurations.

Table 7

Weight, Volume and Reliability of Different Back-Up Configurations

Configurations with duplicated units	Weight (Kg.)	Volume (litres)	Typical power (W)	Reliability (100% duty cycle)	
				12 months	18 months
1 - No redundancy	5.0	5.5	10.6 at	0.888	0.840
2 - TU	5.2	5.6	100%	0.893	0.844
3 - CPU	5.9	6.3	operating	0.924	0.888
4 - TU+CPU	6.0	6.3	cycle	0.926	0.895
5 - MEM	7.1	7.4	1.39	0.931	0.897
6 - CPU+TU+MEM	8.2	8.4	standby	0.974	0.959

The power consumption is the same for all configurations because all redundant units are switched off. If the OBC does not operate for 100% of the time, it can be put in standby by means of a HLT instruction at the end of a program and restarted by means of a program interrupt. This arrangement can save considerable power.
The mechanical layout of fig. 10 is applicable for the third configuration.

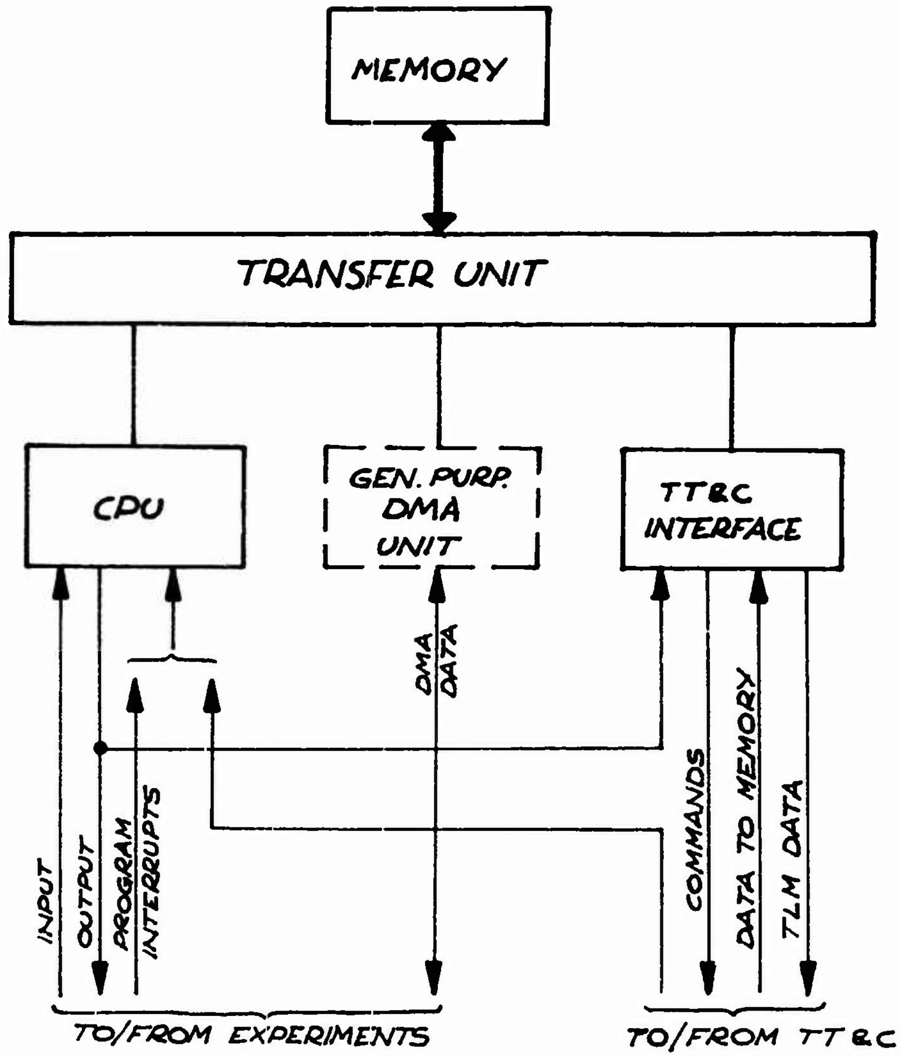


Fig. 1 Block Diagram of the basic, single-processor configuration

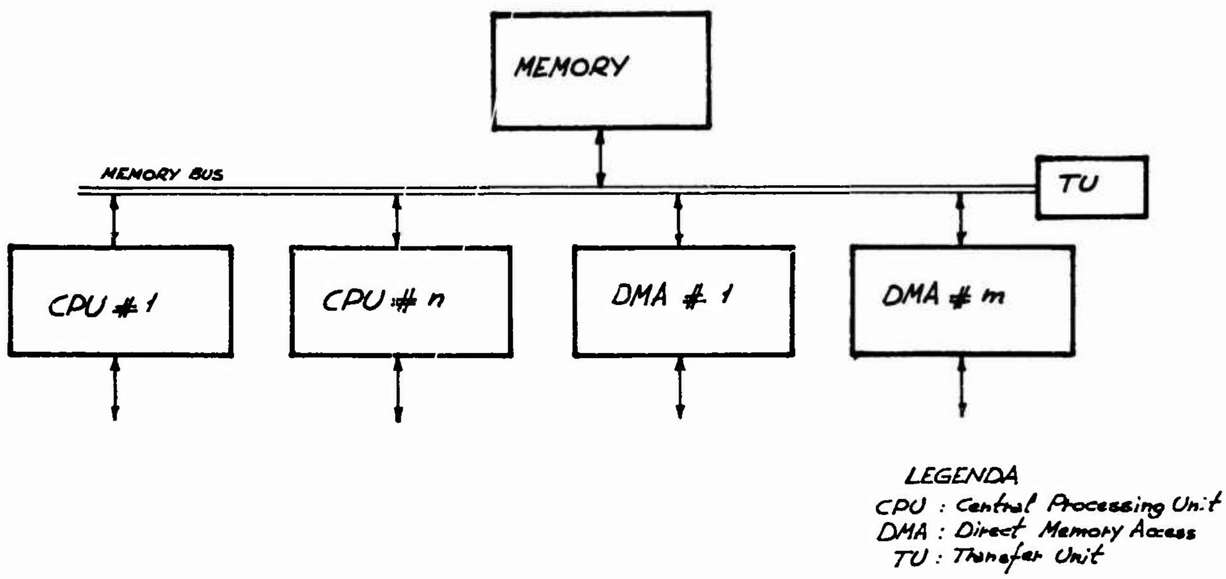


Fig. 2 Multiprocessor Configuration with Common Memory

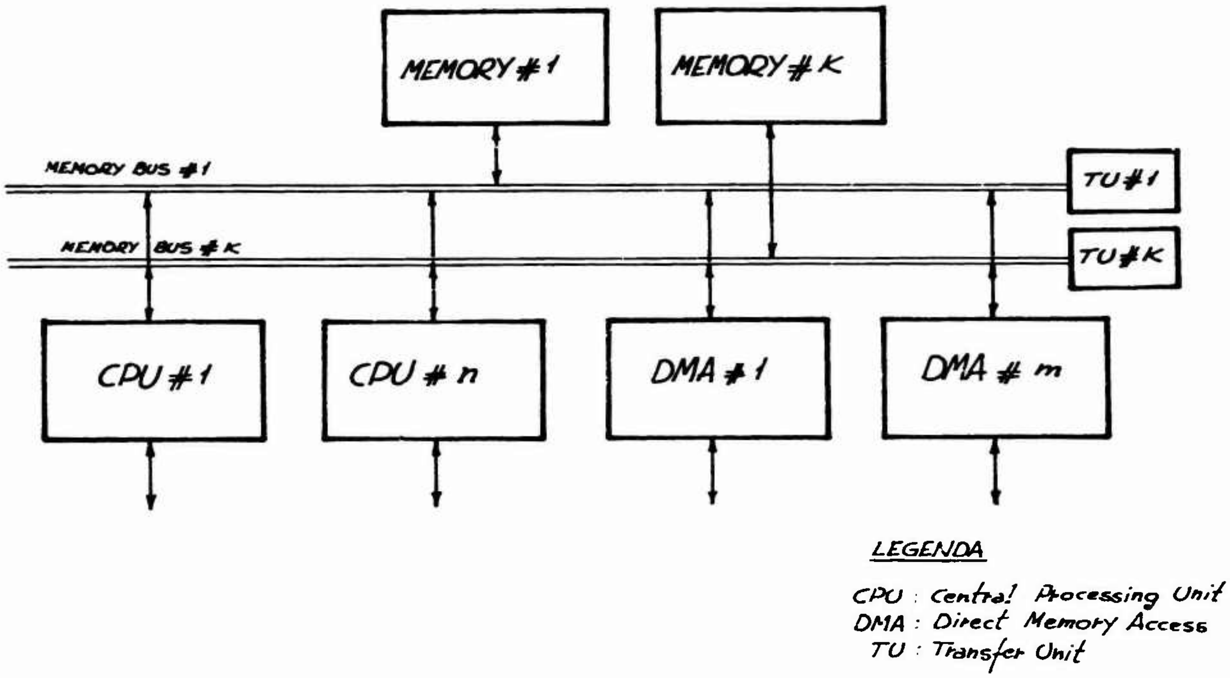


Fig. 3 Multiprocessor with Memory Overlapping

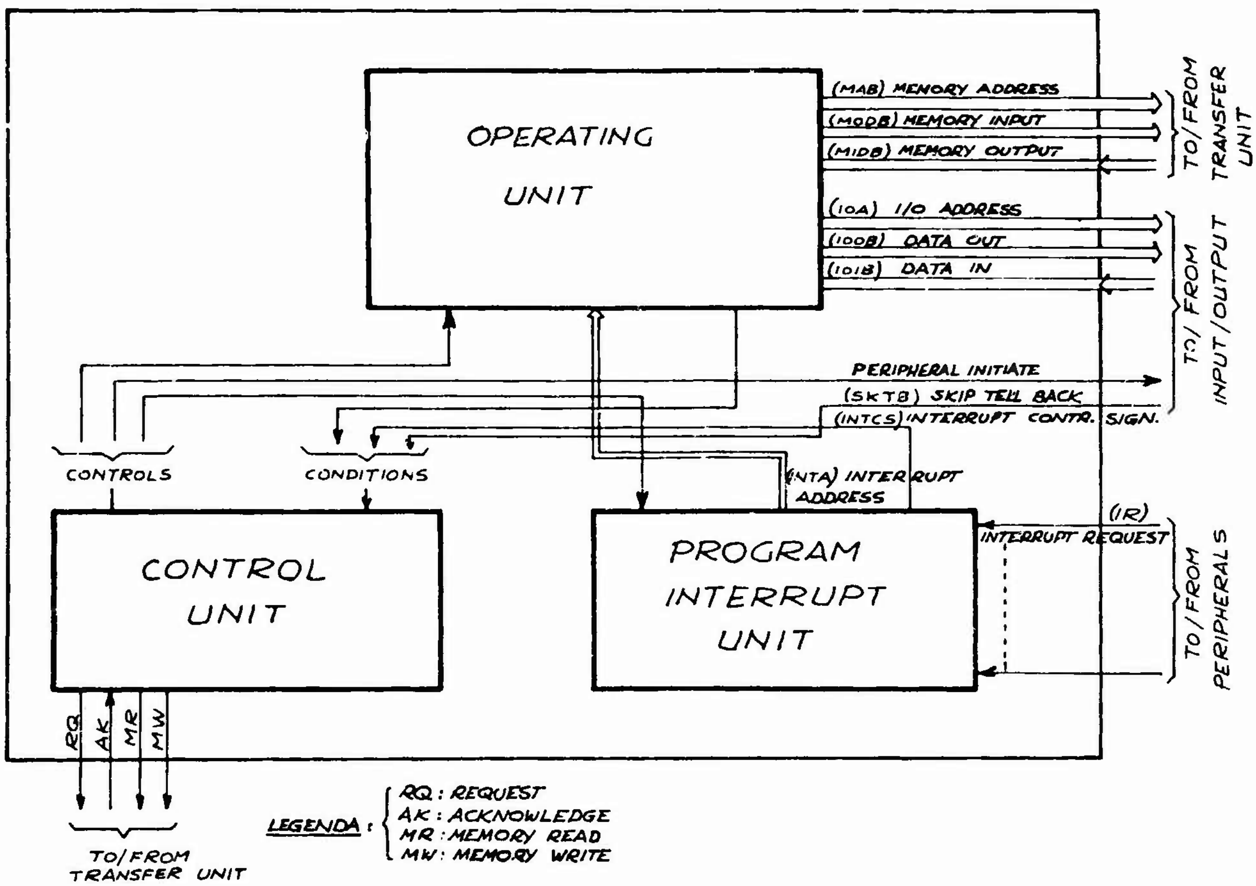


Fig. 4 Central Processing Unit Block Diagram

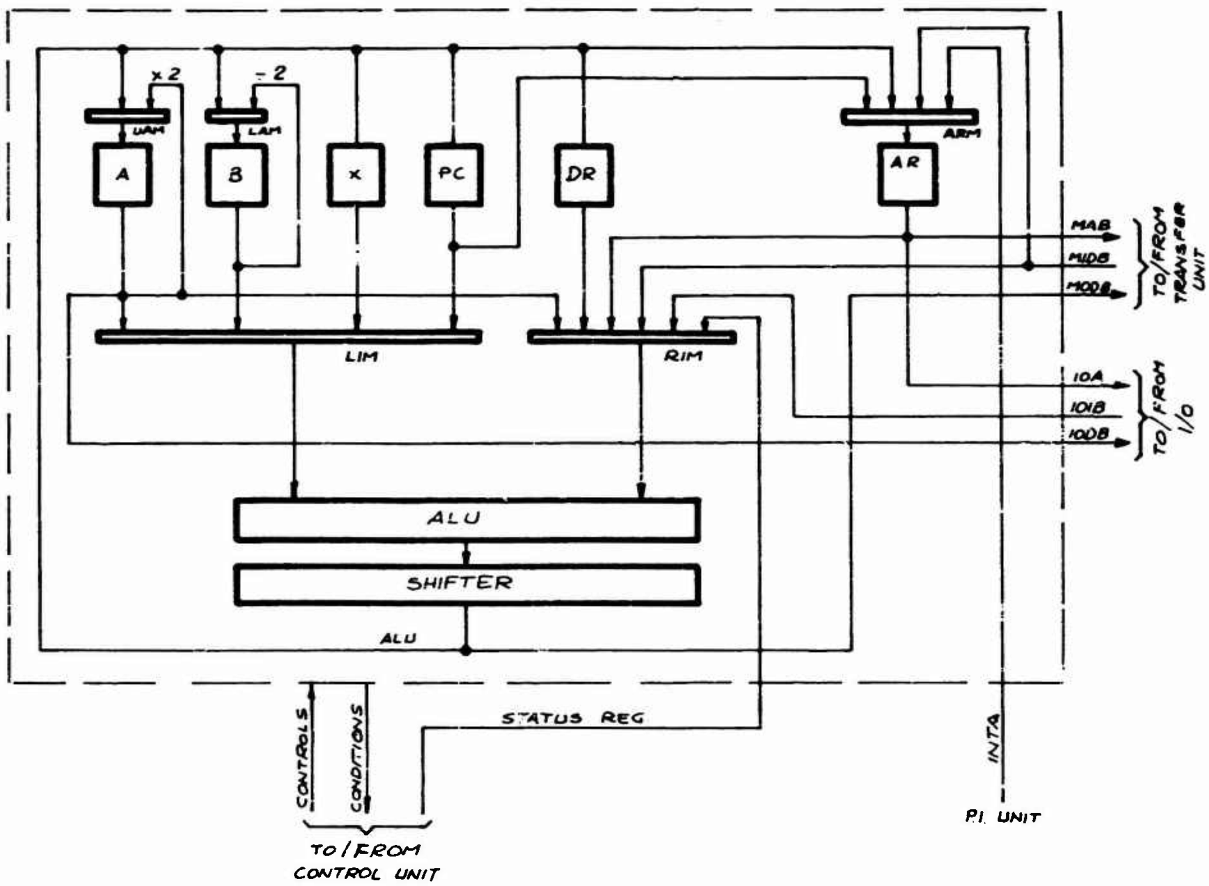


Fig. 5 Operating Unit Block Diagram

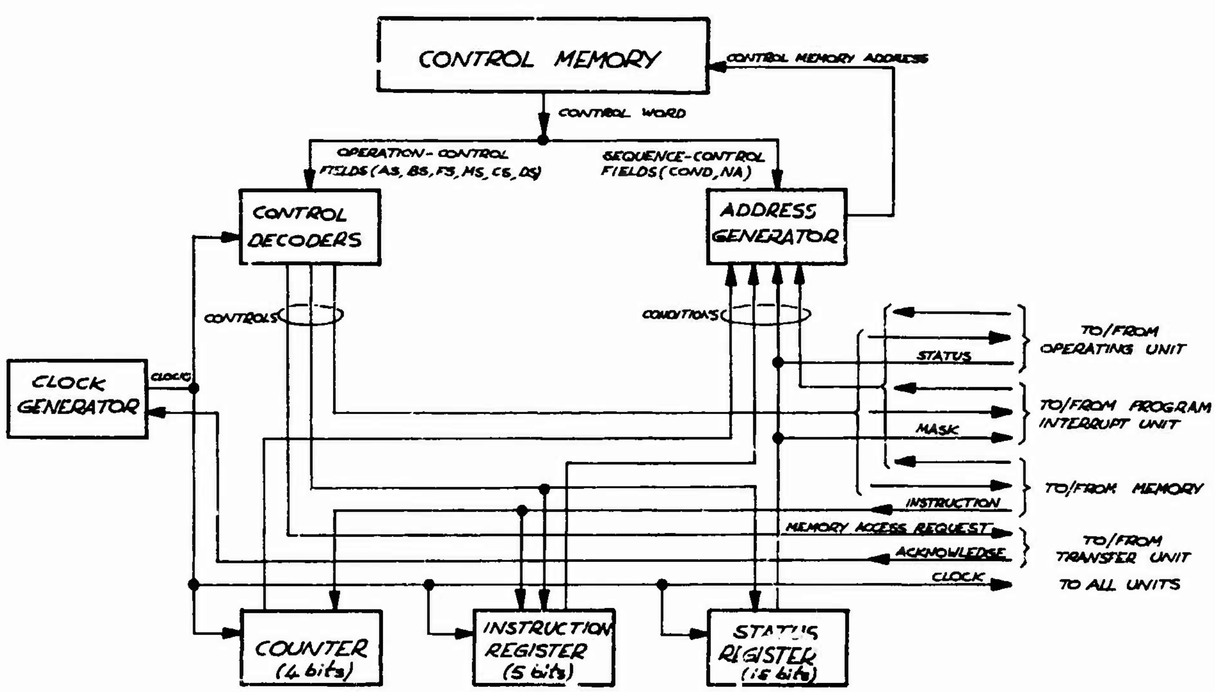


Fig. 6 Control Unit Block Diagram

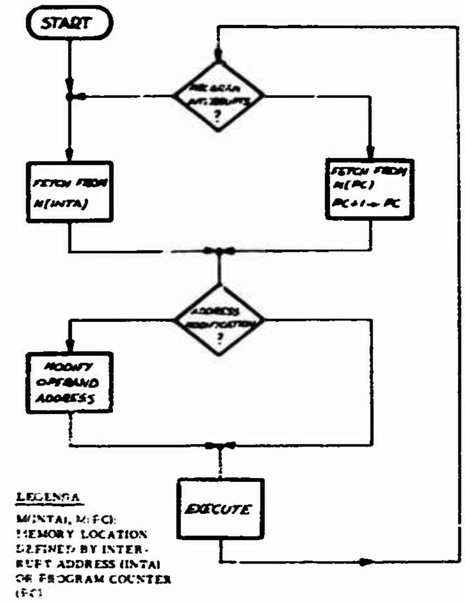


Fig. 7 Operation Sequence

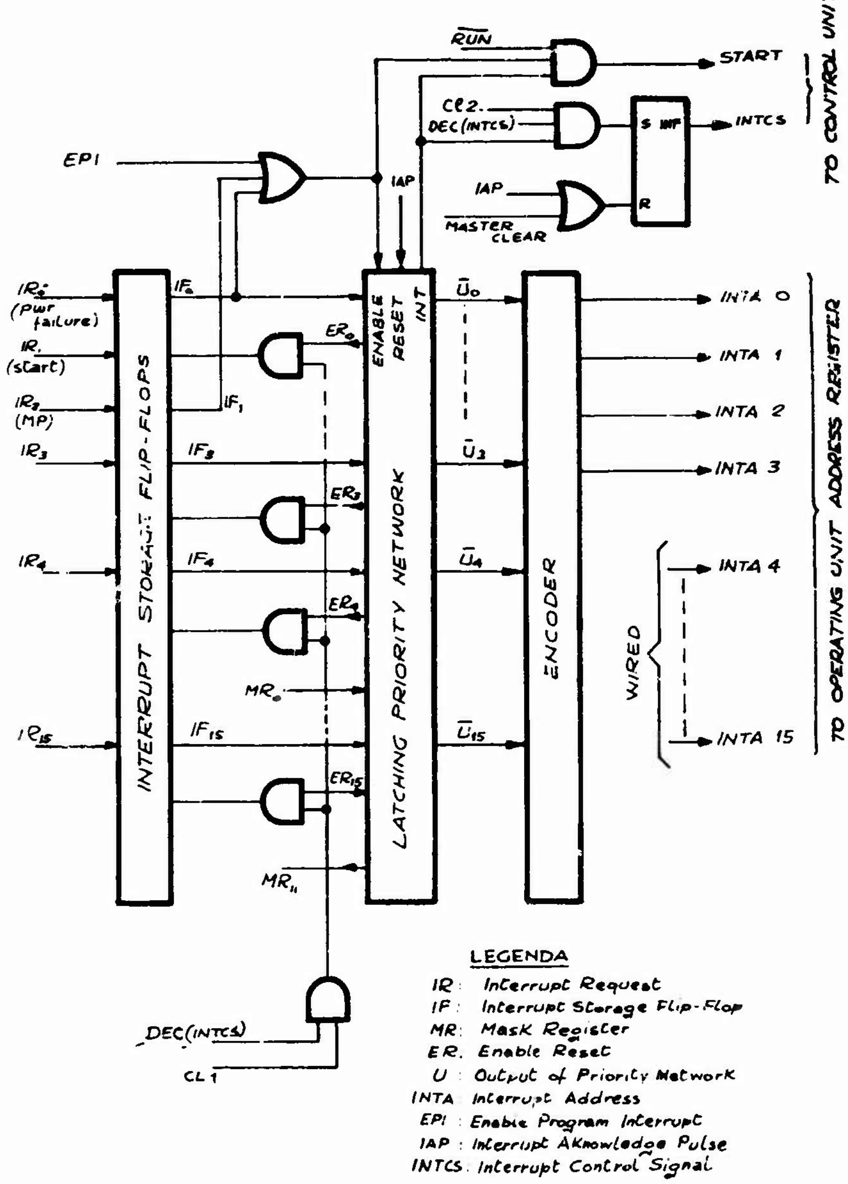


Fig. 8 Logic Diagram of Program Interrupt Unit

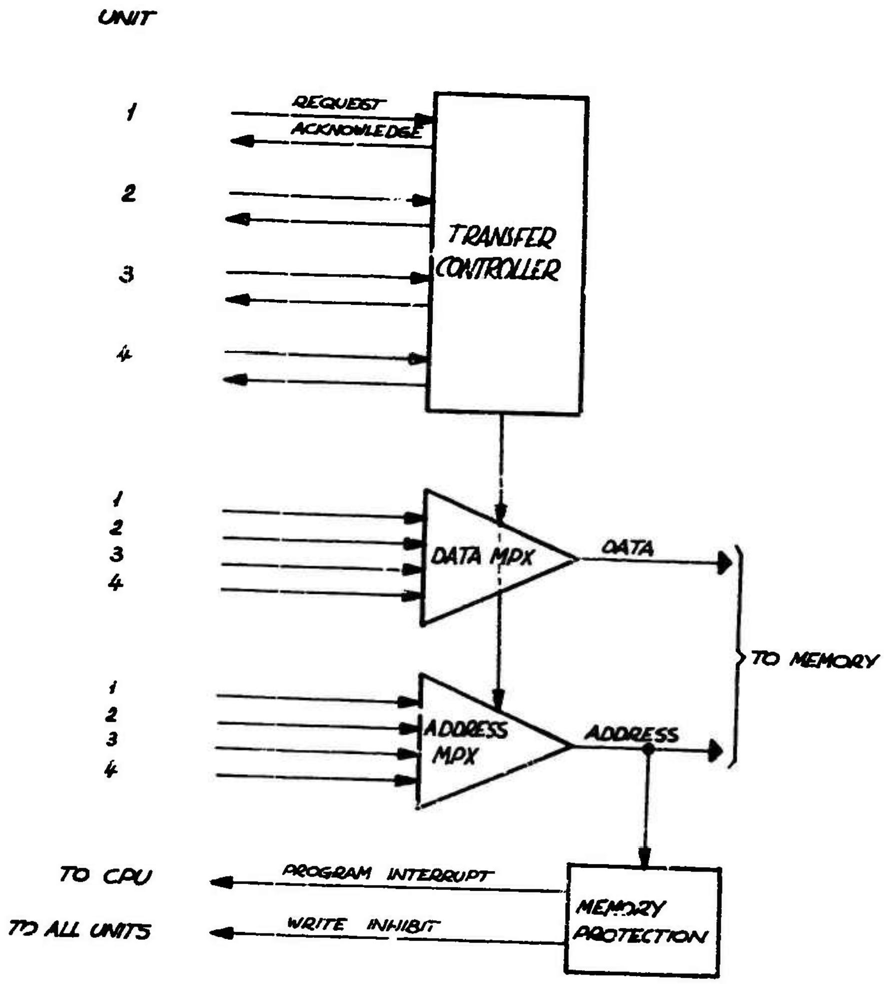


Fig. 9 Block Diagram of the Transfer Controller

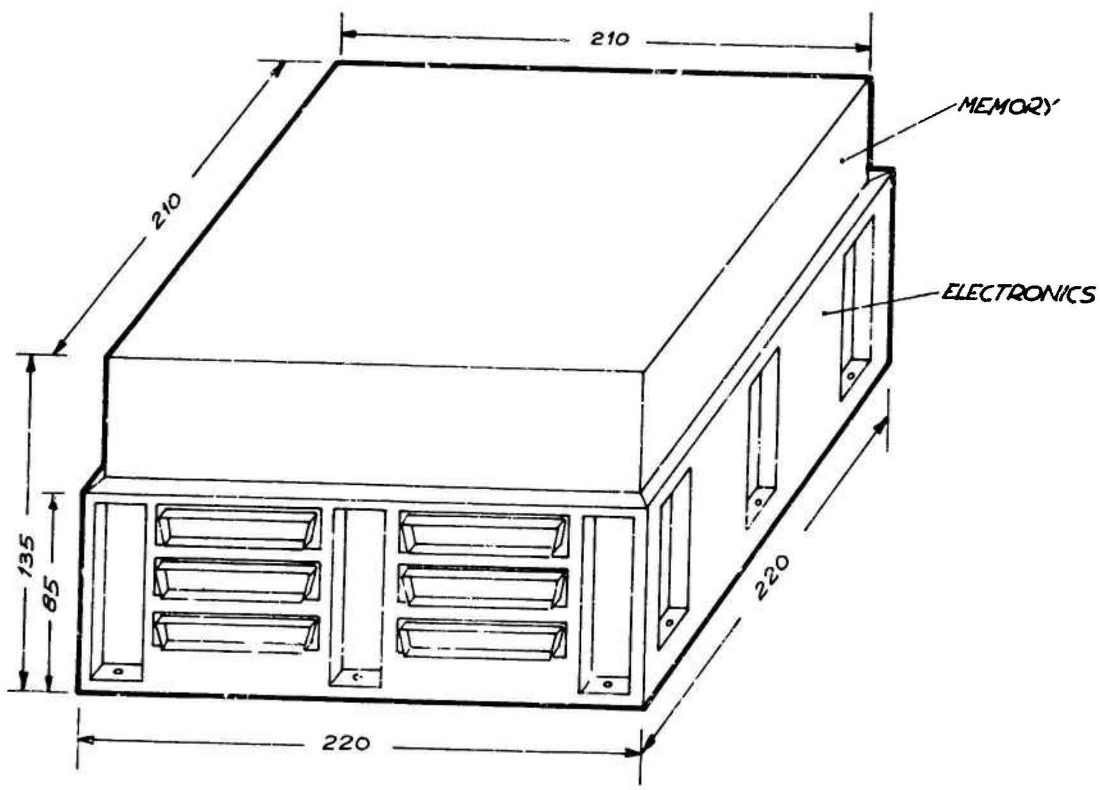


Fig. 10 Outline of Basic Configuration

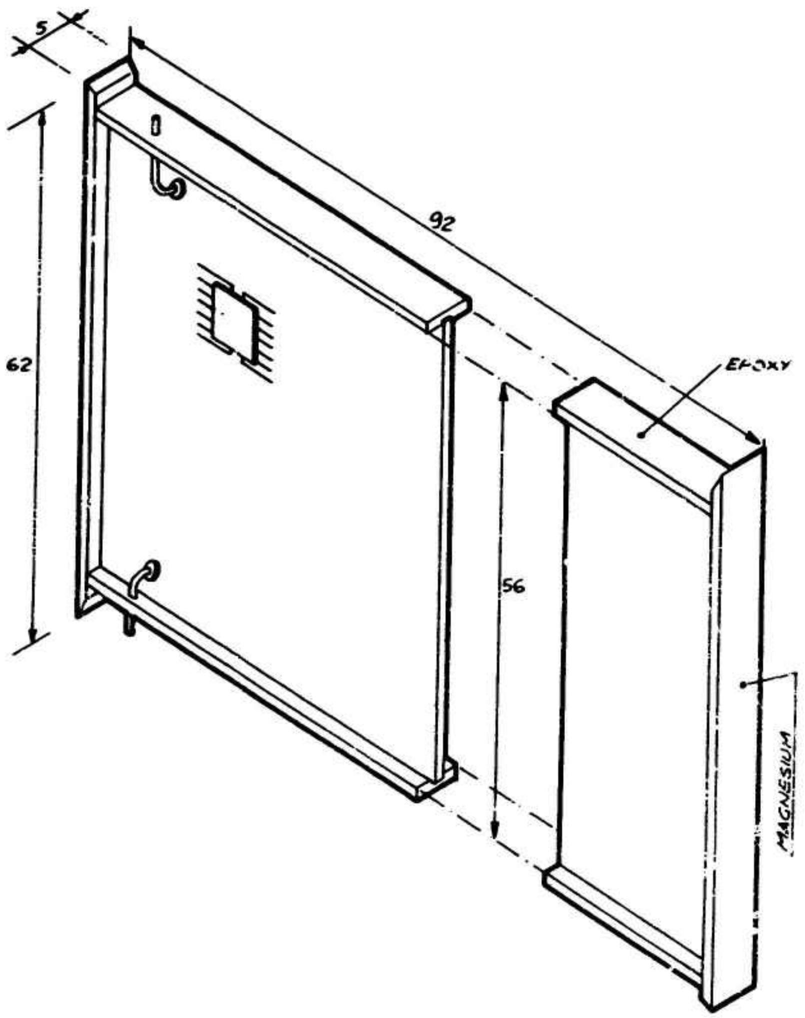


Fig. 11 Printed Circuit Module

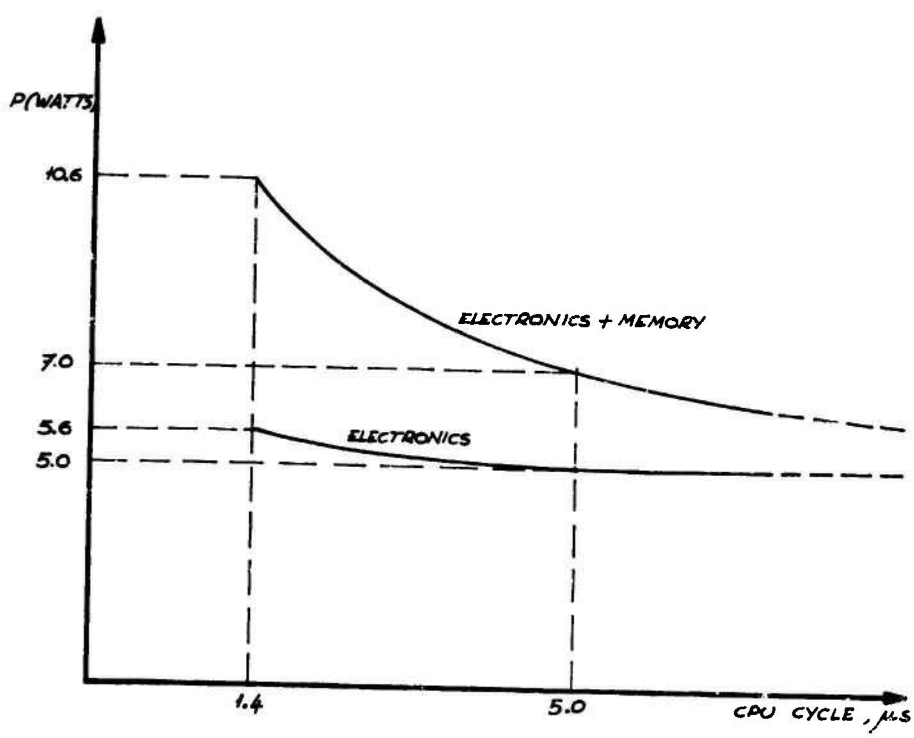


Fig. 12 Power Consumption versus Machine Cycle Time
(Basic Configuration)

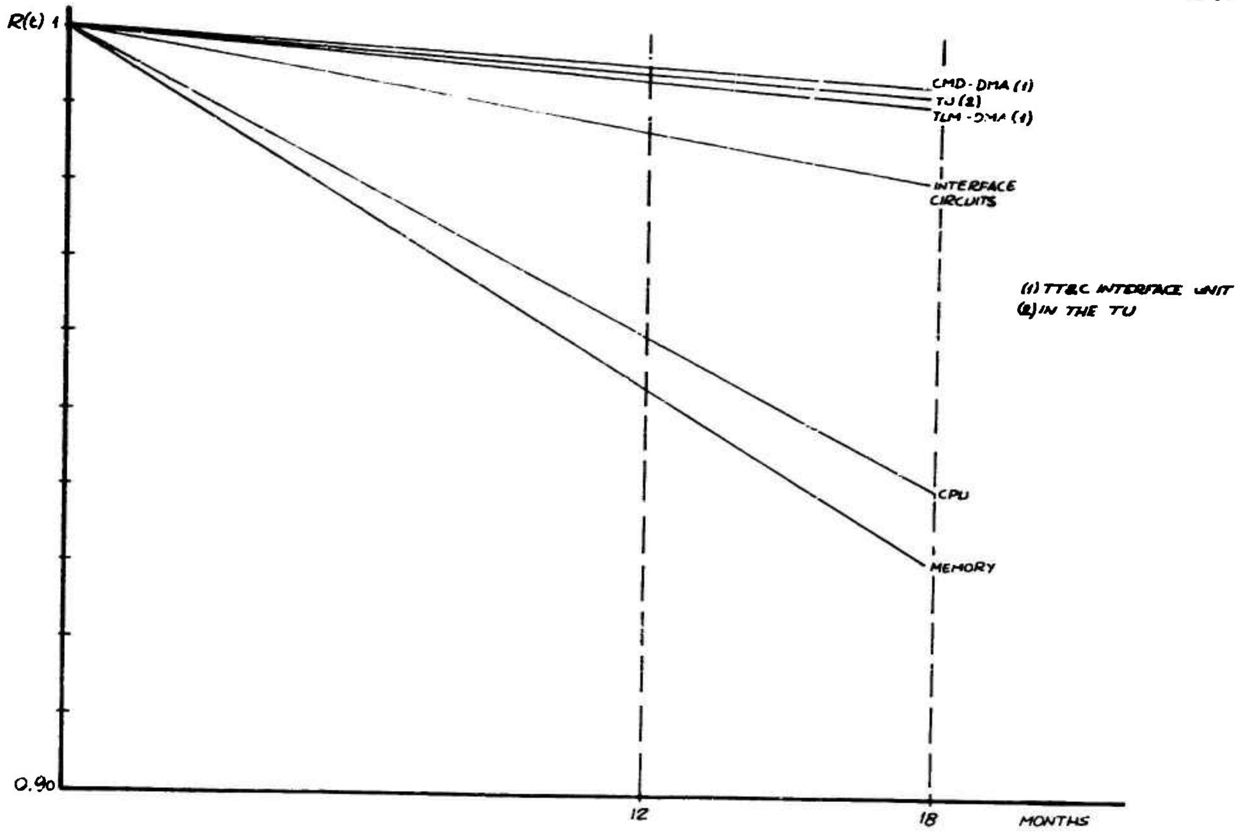


Fig. 13 Reliability of the Single Units

LEGENDA

CPU : CENTRAL PROCESSING UNIT
TU : TRANSFER UNIT
DMA : DIRECT MEMORY ACCESS UNIT
TLM : TELEMETRY
CMD : TELECOMMAND
MEM : MEMORY
INT : INTERFACE CIRCUITS

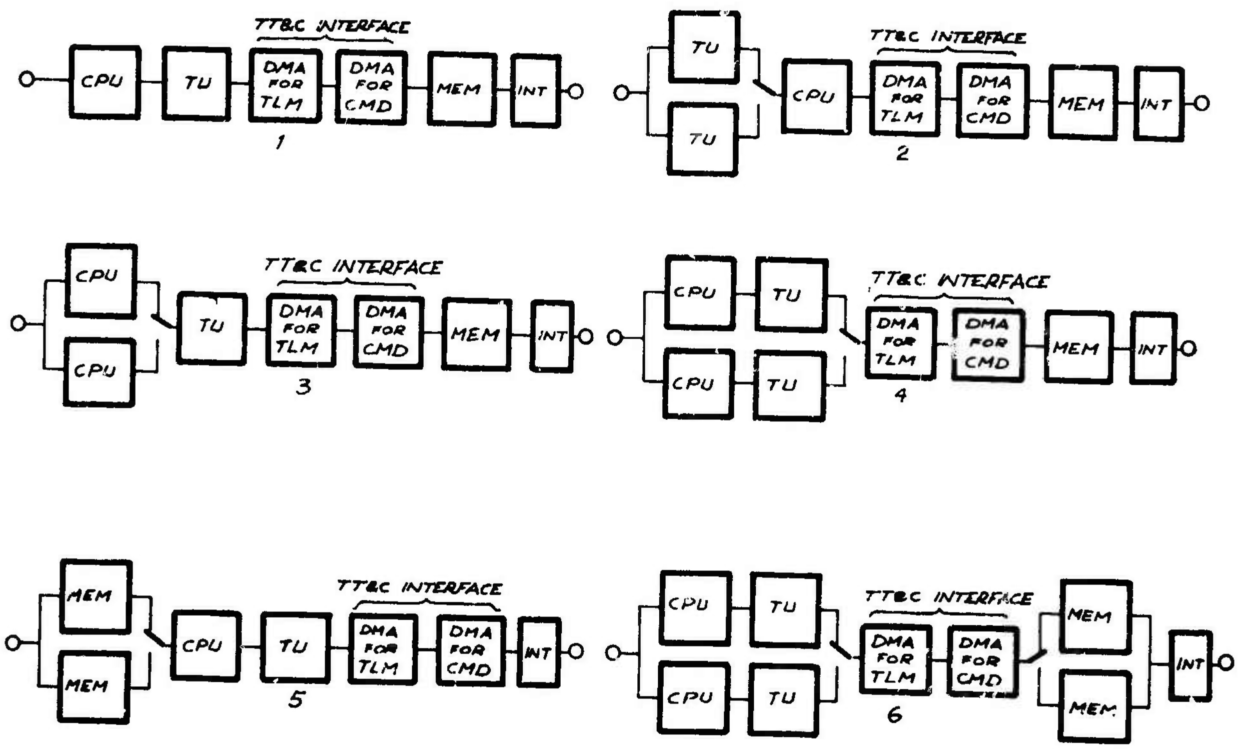


Fig. 14 System Redundancy configurations

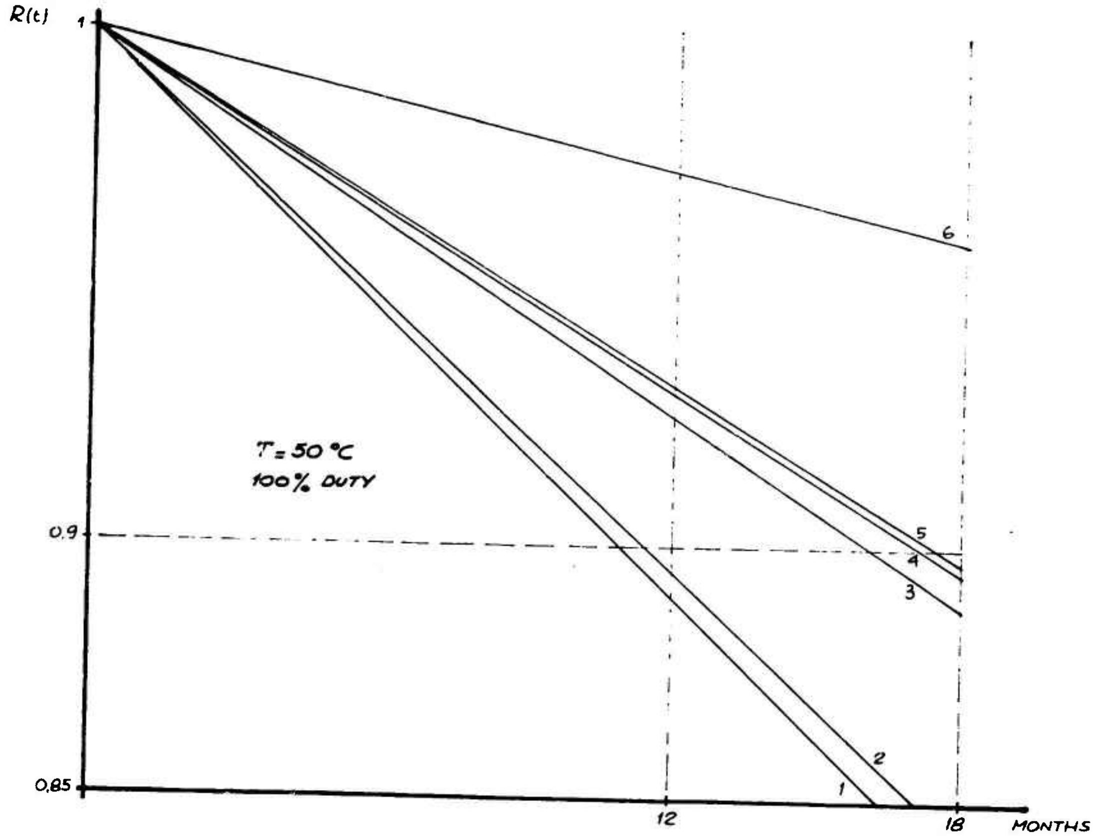


Fig. 15 Reliability of Different Redundancy Configurations

PROCEDE de SURVOL NON INERTIEL

par MM.

P.J. BIGEON Ingénieur en Chef à l'AEROSPATIALE
Ingénieur de l'Ecole Polytechnique
Ingénieur diplômé de l'ENSAR
Professeur honoraire de l'ENSAR

J. LANGLOIS Chef de Département à l'AEROSPATIALE
Ingénieur diplômé de l'ENSAé.

R. BERROIR Chef de Département à l'AEROSPATIALE
Ingénieur de l'Ecole Polytechnique
Ingénieur diplômé de l'ENSAé.

Société Nationale Industrielle Aérospatiale
2 à 18, rue Béranger
92320 CHATILLON
France

INTRODUCTION

1) - Définitions brèves

Système inertiel : Système de guidage utilisant des accéléromètres comme détecteurs principaux et les intégrales simple et double des signaux de sortie des accéléromètres pour déterminer vitesse et position du missile.
Une centrale gyroscopique définit les axes de calcul par rapport aux axes terrestres.

Système non inertiel : La vitesse du mobile est supposée connue par des procédés externes. La position du mobile est déterminée par la connaissance précise du cap du mobile délivrée par une centrale de cap.

2) - Considérations de coût - efficacité

L'étude du coût - efficacité a montré que pour des durées importantes de l'ordre de quelques dizaines de minutes dans l'application envisagée, la précision se dégradait beaucoup plus rapidement pour un système inertiel que pour un système de guidage en cap même pour un vol troposphérique. (Voir figure 1)

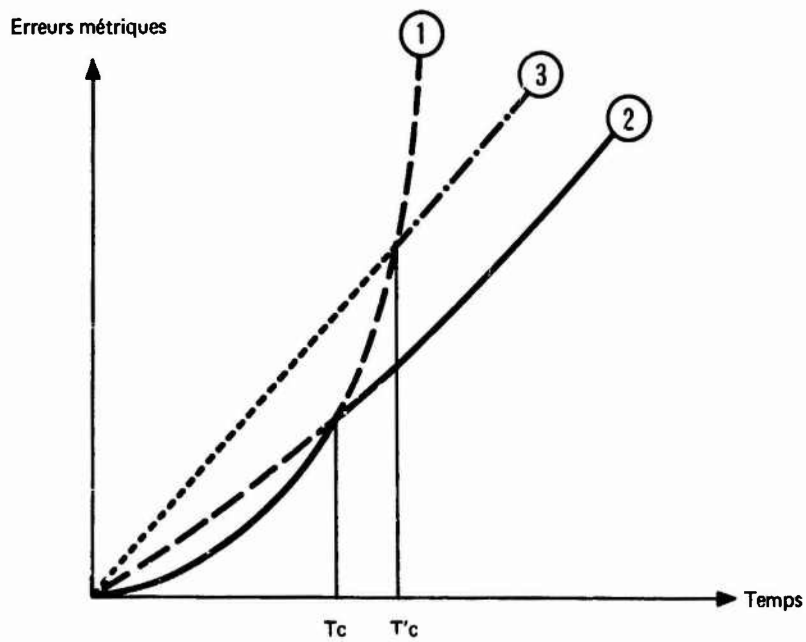


Figure 1.

(1) Centrale inertielle
(2) Centrale de cap sans vent
(3) Centrale de cap en présence de vent

Il en résulte que pour le développement, dans un cadre financier donné, d'un système de survol précis par drone, nous avons été amené à préférer un guidage en cap, associé à une stabilisation barostatique de l'altitude, et à un recalage initial par localisation et télécommande associées à un calculateur de vitesse et position par rapport au sol.
Les erreurs par rapport à une trajectoire assignée sont mesurées en présence de vent, et les corrections ainsi faites sur les éléments initiaux de la trajectoire se répercutent favorablement sur l'ensemble de la trajectoire.

°
° °

Dans une première partie, nous décrirons sommairement les aides au sol et la méthode de préguidage utilisés.

Dans une seconde partie, nous décrirons l'équipement de bord.

En finale, nous présenterons les résultats obtenus.

PREMIERE PARTIE : LES AIDES AU SOL

REALISATION DE LA MISSION

L'exposé ne traite que le cas de la mission type correspondant au survol pendant la phase programmée d'objectifs ponctuels ou d'une zone à surveiller.

Les autres types de missions exécutées en partie sur programme s'en déduisent immédiatement.

Moyens utilisés

Pour réaliser la mission, les moyens suivants sont nécessaires :

a) Une télécommande.

b) Un localisateur interrogeant un répondeur monté sur le missile. Il donne avec une grande précision la distance horizontale et le gisement du missile par rapport à l'antenne.

- un calculateur de route associé au localisateur. Il traite les informations gisement et distance et en extrait la vitesse sol et la route vraie du missile.

c) Un programmateur à bord du missile sur lequel on affiche, avant le lancement, le programme des manoeuvres qui devront être exécutées pendant le vol autonome.

Principe du guidage

La précision du vol programmé dépend essentiellement des conditions initiales d'enclenchement du programme, en particulier de la position et de la route du missile à cet instant.

La méthode consiste à présenter avec précision et dans les conditions requises le missile au point d'enclenchement prévu.

Dans la phase de guidage aller, le missile est aligné sur un axe fixé à l'avance, dit "route d'alignement". Sur cette route a été déterminé le point "d'enclenchement du programmateur" au passage duquel on déclenche la phase de vol programmé.

A - Alignement du missile sur la route prévue.

Pour effectuer l'alignement du missile, l'officier pilote dispose de :
- sa position, sur table traçante,
- sa route vraie, sur un cadran.

La position du missile est fournie par le localisateur. La trajectoire s'inscrit sur la table traçante associée. Le pilote déduit immédiatement de ces informations l'écart de la position du missile par rapport à la route d'alignement qui a, lors de la préparation, été portée sur la table.

La route vraie du missile est fournie par le calculateur. Elle apparaît sur un cadran placé au-dessus de la table traçante. Plus exactement, pour faciliter la tâche du pilote, la lecture donne directement l'écart entre la route vraie et la route d'alignement.

Si la route d'alignement est confondue avec l'axe de tir, l'alignement peut être commencé dès la stabilisation du missile en altitude ; on peut même éventuellement lui donner a priori, la correction de cap nécessaire tenant compte du vent à la rampe.

Dans le cas contraire, on effectue successivement les manoeuvres suivantes :

a) On fait converger le missile vers la route d'alignement suivant une route prédéterminée ;

b) Sur cette route, à une distance déterminée, on déclenche un virage prévu, dit "virage de raccordement", tel qu'à la sortie du virage le missile se trouve à environ un kilomètre de la route d'alignement et converge vers cette dernière sous un angle voisin de 6°. Ce résultat est, en fait, obtenu à 3° près ;

c) Compte tenu de l'angle effectif de présentation par rapport à la route d'alignement, on commande, à une distance déterminée de cette route, le virage correspondant. En sortie de virage, le missile se trouve sur l'axe et sensiblement aligné. Il reste alors à parfaire l'alignement, en principe par une seule correction.

Pratiquement, lors de la préparation, on trace des parallèles à la route d'alignement, situées à des distances correspondant respectivement à des virages de 3°, 6°, 9°. Cette grille permet de déclencher le virage effectif à la distance adéquate, éventuellement par interpolation.

Compte tenu de la constante de temps du calculateur, l'opération d'alignement est effectuée en 30 km. environ.

B - Enclenchement du programmateur

Le programme de vol est affiché sous forme de "temps" sur le programmateur. Le programme ayant été déterminé pour une vitesse sol prévue à l'avance (vitesse propre du missile et vent estimé) il est nécessaire de décaler, sur la route d'alignement, le point d'enclenchement du programmateur en fonction de l'écart entre la vitesse sol prévue et la vitesse sol donnée par le calculateur. Ce dernier paramètre est lu sur un cadran placé au-dessus de la table traçante.
Pour effectuer une lecture précise de la vitesse sol, il est nécessaire de disposer d'un temps de lissage suffisamment long et d'éviter les évolutions importantes du missile.

C - Correction de guidage avant la phase programmée

a) Corrections d'alignement

Durant la dernière phase de l'alignement, on cherche essentiellement à annuler l'écart de route. Ceci a pour conséquence qu'il existe, juste avant l'enclenchement du programme, un écart résiduel en position, le missile suivant alors une route parallèle à la route d'alignement. Cet écart ne dépasse pas quelques centaines de mètres.

Pour corriger l'erreur de passage sur l'objectif qui en résulterait, la méthode diffère selon les deux cas suivants :

- Objectif ponctuel sur la route d'alignement : on donne, avant l'enclenchement du programmateur, un ordre de virage tel que la route finale converge avec la route d'alignement sur l'objectif.

  L'ordre à transmettre est lu directement sur l'"échelle de convergence" portée sur la table traçante et graduée en fonction de l'écart transversal du missile.

  Dans le cas d'une zone de surveillance, on considère le centre de cette zone com[illegible] objectif ponctuel, les écarts de passage aux extrémités seront faibles.

  Cette correction est dite de "convergence".

- Objectif ou zone situés après un virage : il suffit d'enclencher le programmateur [illegible] ue le missile traverse la "droite d'enclenchement" parallèle à la route de survol de l'[illegible]if, droite qui passe par le point prévu pour l'enclenchement du programmateur.

  Le missile suit alors une route parallèle à la route d'alignement. Sa trajectoire est telle qu'à la sortie du virage programmé, il suit la route prévue pour le survol de l'objectif.

  Cette correction est dite des "routes parallèles".

b) Corrections d'enclenchement du programmateur

Le point d'enclenchement du programmateur doit être déplacé en fonction de l'écart entre la vitesse sol réelle et la vitesse sol prévue.

La vitesse propre du missile étant déterminée avec une bonne approximation, l'erreur commise sur sa vitesse-sol estimée résulte essentiellement de la connaissance inexacte que l'on a de la composante axiale du vent. Par rapport au programme prévu, tout se passe comme si, a partir de l'enclenchement du programmateur, l'objectif se déplaçait à une vitesse égale et opposée à la différence entre les vitesses sol réelle d'une part et lue d'autre part.

Il faut alors décaler le point d'enclenchement du programmateur en conséquence.

En pratique on prévoit sur la table traçante, de part et d'autre du point d'enclenchement du programmateur correspondant à la vitesse prévue, des points qui correspondent à des vitesses effectives échelonnées de 5 m/s en 5 m/s. L'instant de l'enclenchement est déterminé par interpolation entre ces valeurs.

PREPARATION DE LA MISSION

Le Poste Central de Tir (P.C.T.) a pour fonction principale de coordonner et de rassembler tous les éléments nécessaires à la préparation de la mission.

L'ordre d'exécution d'une mission de surveillance, transmis au P.C.T., précise :

- l'heure de la mission,
- la nature des équipements de surveillance,
- la définition de l'objectif par les coordonnées géographiques de son centre Xo, Yo,
- la zone de surveillance par sa longueur L et la direction de survol θ o,
- les consignes et les prescriptions particulières de vol.

A partir de ces données, le rôle du P.C.T. est de :

- préparer le plan général de la mission sur la carte,
- déterminer avec précision les éléments principaux de la trajectoire, (route d'alignement, trajectoire sur programme, etc ...)
- donner au camion-rampe l'axe de tir, les éléments d'affichages du programme et l'altitude de vol,
- donner au poste de guidage les valeurs nécessaires à la préparation du calque de vol.

Tracé sur carte de la mission

Le P.C.T. fournit :

- les coordonnées du localisateur et de l'objectif de façon précise,
- les positions possibles de rampe de façon approchée.

Sur la carte (le plus souvent au 1/200 000e pour avoir tous les éléments sur la même planche) le commandant d'unité trace approximativement la mission complète que devra effectuer le missile.

Dans le choix de la trajectoire et tout particulièrement de la route d'alignement, il doit tenir compte de divers impératifs :

- Impératifs tactiques tels que :

  axes de tir possibles,
  zones de survol interdit,
  distance limite d'enclenchement du programmateur, etc ...

- Impératifs techniques tels que :

  alignement supérieur à 30 km,
  trajectoire de raccordement réalisable rapidement,
  parcours simple.

Détermination exacte de la trajectoire

A partir du travail de dégrossissage effectué sur la carte, le PCT se fixe à priori certaines valeurs, par exemple :

- l'orientation de la route d'alignement et la distance du localisateur à cette route,
- la distance du point d'enclenchement du programme à l'objectif ou au début du premier virage,
- l'altitude de vol, le régime réacteur, soit la vitesse propre du missile.

Le P.C.T. dispose alors de toutes les données nécessaires pour établir la trajectoire type.

En règle générale cette trajectoire est déterminée pour un vent nul (vitesse-sol prévue = vitesse propre du missile). Cependant, si le vent moyen est très fort, il sera prudent d'en tenir compte :

- d'une part pour déterminer l'axe de tir,

- d'autre part pour prévoir le point d'enclenchement du programmateur dans une zone convenable.

La Figure 2., ci-après, donne un exemple de trajectoire type.

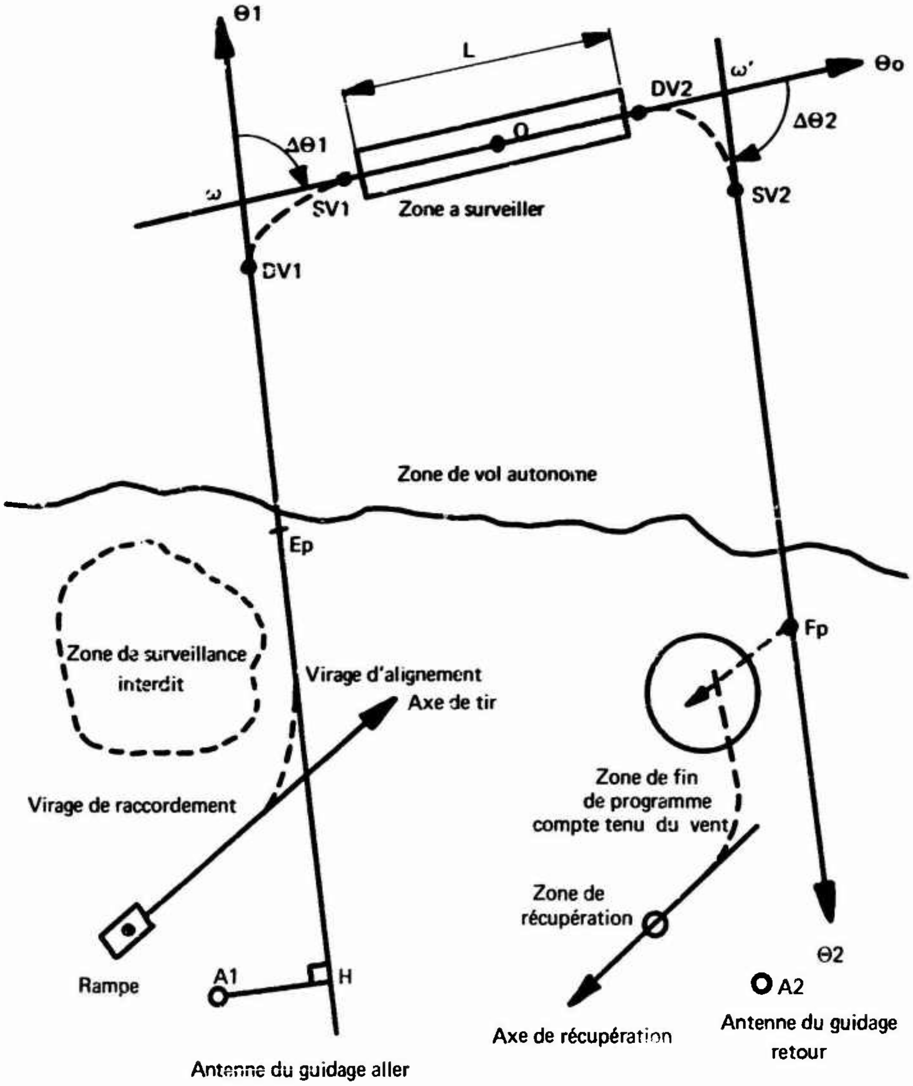


Figure 2. : Détermination de la trajectoire

Les données sont : A (antenne localisateur)
O, D et θ o déterminant la zone à surveiller

indice 1 - aller
indice 0 - survol de l'objectif
indice 2 - retour.

Après tracé sur carte, le P.C.T. s'est fixé :

- la route d'alignement : $\theta 1$
- la distance de l'antenne A à cette route : $\overline{AH}$,
- la distance du point H au point Ep d'enclenchement du programmateur : $\overline{H\ Ep}$.

d'où l'angle de virage $\Delta\theta\ 1 = \theta o - \theta 1$

Compte tenu de la vitesse, des abaques donnent directement les distances

$\overline{Dv\ \omega}$ (Dv = début de virage)
$\overline{\omega\, Sv}$ (Sv = sortie de virage)

d'où les distances

$\overline{Ep\ Dv} = (\overline{Ep\ \omega} - \overline{Dv\ \omega})$
$\overline{Sv\ o} = (\overline{\omega\, 0} - \overline{\omega Sv})$

Les éléments de la trajectoire de retour, qu'il n'est pas nécessaire de déterminer avec précision , peuvent être lus directement sur la carte. Seules les coordonnées du point de récupération doivent être bien connues afin de permettre un atterrissage précis.

Le P.C.T. est alors en mesure de transmettre aux postes de guidage les éléments nécessaires pour la préparation du calque de vol.

Détermination du programme

Pour l'établissement du programme, la procédure est la suivante :

a) On détermine le programme, le vent étant considéré comme nul :

- Pour les segments de droite on prend un temps de parcours égal à

$$\Delta t = \frac{d}{V_p} \quad \begin{cases} d & \text{étant la longueur du segment} \\ V_p & \text{étant la vitesse propre du missile} \end{cases}$$

  pour les virages, on identifie changement de route et changement de cap et on lit directement sur les abaques la durée de l'évolution en virage, de D_V à S_V.

- On se fixe l'instant de démarrage et la durée de fonctionnement des moyens de surveillance.

b) On introduit le vent, au dernier moment, sous forme de corrections.

  Ce vent est le vent estimé dans la zone de l'objectif à partir de sondages effectués dans la zone de lancement.

  La correction principale est celle qui affecte le premier virage. En effet, il n'est plus alors possible d'identifier changement de cap et changement de route. On accroît ou diminue la valeur de changement de cap, d'où la durée du virage pour obtenir le changement de route voulu. Cette correction est donnée directement par des abaques.

  Il est à noter qu'il n'y a pas de correction à faire sur le premier tronçon de ligne droite. En effet mesurant la vitesse-sol, on tient automatiquement compte du vent réel pendant la phase d'alignement.

  On pourrait introduire des corrections pour les éléments de trajectoire (segments de droite, virages) postérieurs au passage sur l'objectif mais il est plus simple de déterminer le décalage, par rapport à la position prévue, du point de fin de programme compte tenu du vent.

La procédure ci-dessus de détermination du programme permet, d'une façon générale, de traiter tous les cas sans documentation particulière.

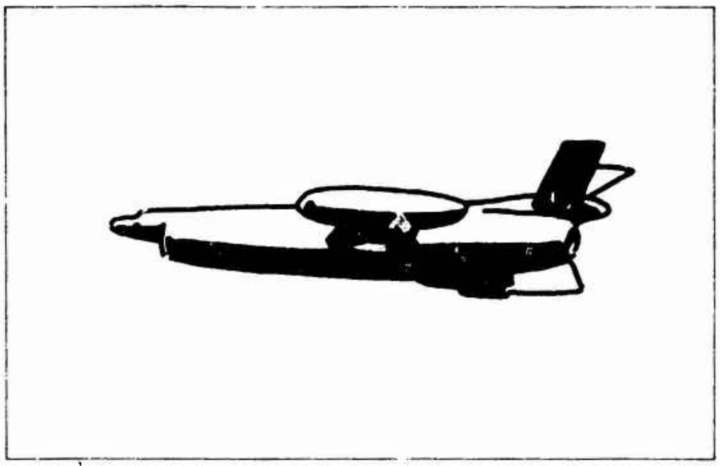


Figure 3.

En utilisation opérationnelle, il est prévu de constituer un dossier de programmes préétablis (durées de ligne droite, ordres de virage, déclenchement et arrêt des moyens de surveillance, etc.). Ces programmes correspondraient à un certain nombre de missions-types, couvrant pratiquement tous les besoins.

Cette procédure opérationnelle autorise une transmission plus rapide et plus discrète du programme à afficher. En outre, celui-ci est enregistré sur une carte perforée qui sert de guide pour la mise en place des fiches sur le programmateur et rend presque impossible l'erreur d'affichage.

Préparation des calques de la table traçante

Chaque poste de guidage reçoit du P.C.T. les éléments à afficher sur le calque de la table traçante.

Ces éléments sont de deux sortes :

- les données essentielles : elles sont directement portées sur le calque en utilisant le localisateur et en se basant uniquement sur la fidélité de cet appareil.
- les données annexes de pilotage : ne nécessitant pas une grande précision, elles peuvent être portées directement avec la règle et le rapporteur sur la table traçante.

a) **Préparation du calque de guidage "Aller"**

Données essentielles :

- la direction de la route d'alignement : $\overline{\theta 1}$
- la distance de l'antenne à cette route : $\overline{AH}$
- la position du point d'enclenchement du programme (Ep) défini par la distance $\overline{HEp}$.

Ces données sont affichées sur la table traçante dans deux systèmes d'axes :

- axes géographiques X, Y :

  Origine = antenne localisateur

  Y (direction du Nord) : vertical, orienté vers le haut
  X (direction de l'Est) : horizontal, orienté vers la droite
  L'affichage se fait normalement au $\frac{1}{100.000e}$

- axes calculateurs ξ, η :

  Origine : point H

  ξ ($\overline{HEp}$) : vertical, orienté vers le haut
  η (direction $\overline{AH}$) : horizontal
  L'affichage se fait normalement au $\frac{1}{100.000e}$ suivant l'axe des ξ et au $\frac{3}{100.000e}$ suivant l'axe des η ; ce qui permet d'avoir une meilleure connaissance de l'écart en position du missile.

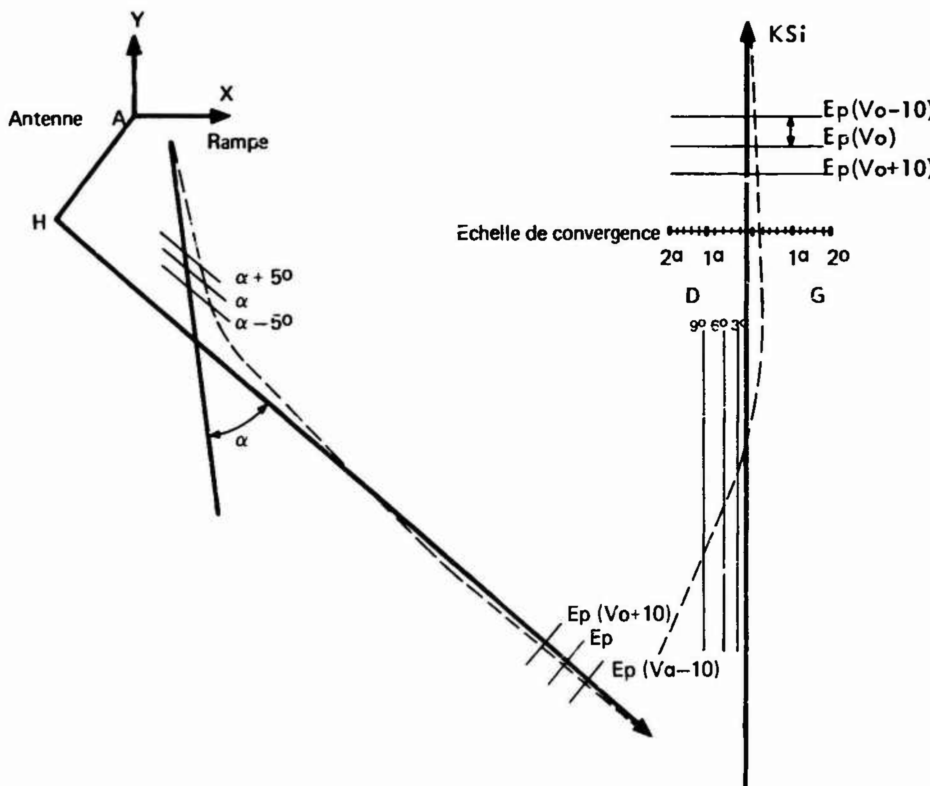


Figure 4. - Préparation de calque de guidage aller

Données annexes du pilotage :

Les autres éléments à porter sur le calque sont :

- dans le système d'axes X, Y
  la position de la rampe,
  l'axe de tir,
  éventuellement la route de raccordement,
  le réseau de droites correspondant au déclenchement du virage de raccordement;

- dans le système d'axes ξ , η
  pour la phase de raccordement ; les parallèles à l'axe des dites 3°, 6°, 9° ;

  pour la phase d'enclenchement du programmateur :
  les droites d'enclenchement correspondant respectivement aux différentes vitesses-sol possibles, éventuellement l'échelle de convergence.

L'affichage de ces éléments est très rapide.

Affichage des éléments sur rampe

Le programme est affiché juste avant la mise en route du missile, l'altitude pendant la montée en régime du réacteur.

EXECUTION DE LA MISSION

Le personnel nécessaire pour assurer l'exécution de la mission dans chaque poste de guidage est le suivant :

- un officier pilote, qui, au vu de la table traçante, décide des évolutions à faire exécuter au missile et des ordres à lui passer,
- un opérateur de télécommande qui, sur indication de l'officier pilote, affiche les ordres à passer sur la boîte de prédétermination et les déclenche au commandement,
- un opérateur qui assure l'acquisition du missile par le localisateur et surveille le bon fonctionnement de celui-ci.

Pour le guidage aller, le pilote est assisté d'un aide pilote dont la fonction essentielle est de lire les indications des cadrans du calculateur et d'en faire le "lissage".

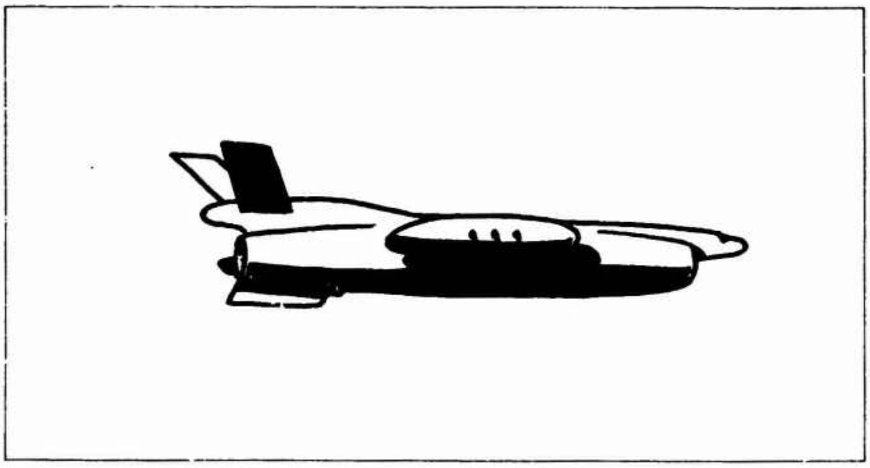


Figure 5.

Guidage "aller"

Le guidage aller comporte les phases successives suivantes :

a) Acquisition du missile au départ

L'antenne ayant été orientée à 20° près dans la direction du missile, l'opérateur, dès que la réponse apparait sur le scope, enclenche la poursuite automatique. L'acquisition ne prend que quelques secondes.

Si la rampe est à portée directe, l'acquisition se fait sur rampe.

b) Raccordement

La route de raccordement est, en général, l'axe de tir.

L'officier pilote, en fonction de l'écart entre cette route et la route d'alignement, déclenche, à la distance convenable prévue et portée sur la table traçante le virage de raccordement.

c) Alignement du missile

En fonction de l'écart de route restant, l'officier pilote déclenche le virage d'alignement au passage de la parallèle convenable (interpolation de la grille 3°, 6°, 9°).

Après ce virage d'alignement, la correction de l'écart résiduel de route n'est possible que lorsqu'on dispose d'une information très précise en provenance du calculateur. La valeur indiquée est correcte après moins d'une minute de vol stabilisé.

d) Corrections

Lorsque l'écart de route est parfaitement stabilisé, le pilote donne l'ordre de virage annulant cet écart, le missile suit alors une route parallèle à l'axe d'alignement.

S'il y a lieu de faire une convergence compte tenu de l'écart en position du missile, le pilote lit directement sur l'échelle de convergence la valeur de l'ordre qu'il doit donner.

e) Enclenchement du programmateur

Compte tenu de la vitesse-sol donnée par le calculateur, le pilote enclenche le programmateur quand le missile franchit la droite d'enclenchement convenable.

-=-=-=-=-=-=-=-=-

DEUXIEME PARTIE : LES EQUIPEMENTS DE BORD

Le SYSTEME de GUIDAGE

Les équipements utilisés pour le guidage sont :

a) le pilote automatique. Il contient les références gyroscopiques. Il élabore les signaux qui, transmis aux gouvernes, permettent de stabiliser le missile. Les références de stabilisation peuvent être modifiées par le programmateur.

b) la sonde barostatique. Elle fournit au pilote automatique le signal de référence nécessaire pour la stabilisation en altitude.

c) le programmateur. Il délivre les signaux de guidage pendant la phase de vol autonome.

PILOTE AUTOMATIQUE

Le pilote automatique élabore les signaux à transmettre aux gouvernes et permet d'assurer la stabilisation du missile en assiette et en roulis. De plus il comporte une surveillance en cap et en altitude. Ces références peuvent être modifiées en vol par le programmateur.

Les références sont élaborées par :

- la centrale gyroscopique incorporée au pilote. Elle fournit le cap Ψ du missile et les angles d'assiette θ et de roulis φ ;
- la sonde barostatique. Elle mesure l'écart d'altitude-pression par rapport à l'altitude affichée pour le vol et transmet cette information au pilote automatique.

Le pilote comprend :

a) la centrale gyroscopique,

b) une chaîne d'asservissement en roulis liée à la chaîne de cap,

c) une chaîne d'asservissement en tangage qui assure la stabilisation du missile autour d'une valeur donnée de l'angle d'assiette θ ; de plus, tout ordre de tangage modifie l'assiette de référence d'une quantité $\Delta\theta$; en outre, un dispositif de surveillance d'altitude permet, s'il est enclenché, d'assurer un vol à altitude constante.

d) une chaîne d'asservissement en cap qui, en l'absence d'ordre extérieur au pilote, maintient constant le cap Ψ du missile : tout écart détecté par la centrale gyroscopique est compensé par une action sur le roulis qui provoque le virage nécessaire. De plus, tout ordre de virage modifie le cap de référence d'une quantité $\Delta\Psi$ qui est affichée dans le pilote : le missile tourne alors de façon que l'asservissement se fasse autour de la nouvelle valeur du cap $\Psi + \Delta\Psi$.

La juxtaposition des sous-ensembles du pilote SX, fournissant les indications nécessaires au guidage du missile est représentée par la figure 6. ci-après.

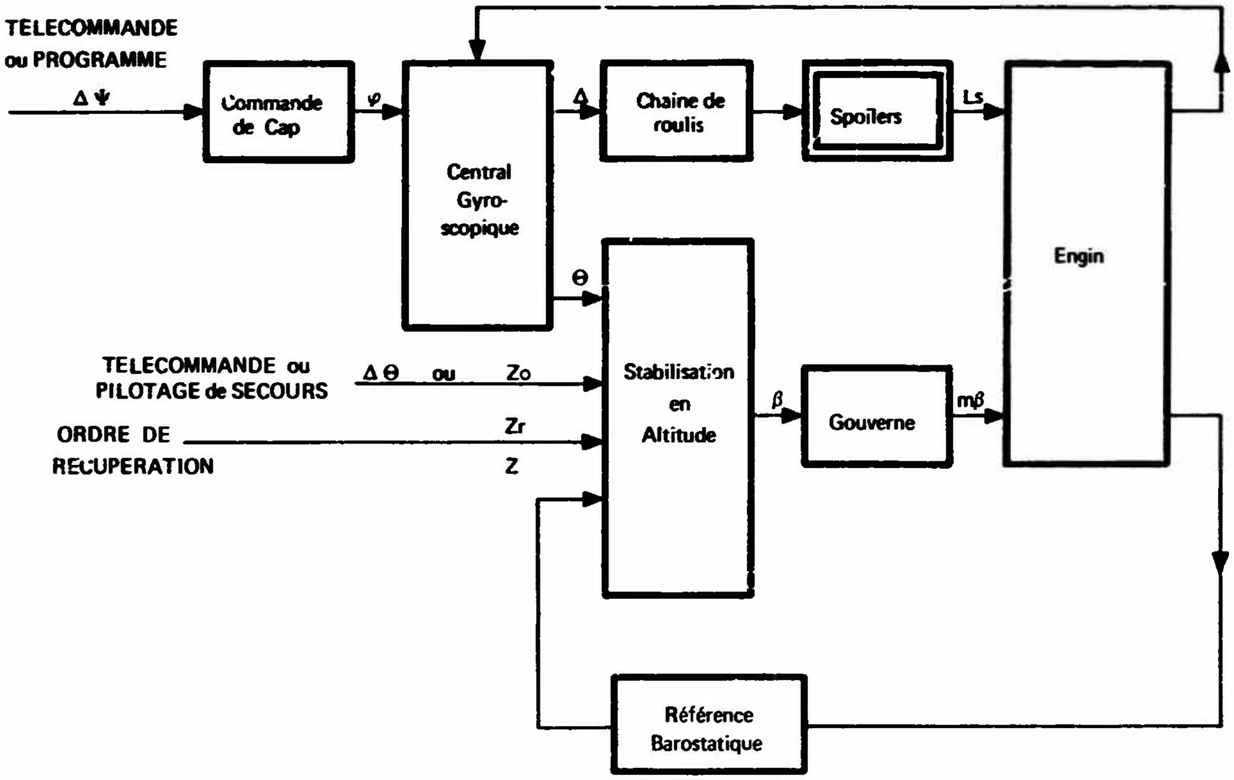


Figure 6. - Chaîne de pilotage du missile

LA CENTRALE GYROSCOPIQUE (Fig. 7)

Cet ensemble est une plate-forme stabilisée à deux gyroscopes (un gyrascope de verticale, un gyroscope directionnel), constituée de la façon suivante : un cadre extérieur stabilisé en roulis supporte deux cadres intermédiaires stabilisés en tangage.

La stabilisation est faite à partir du gyroscape de verticale par des servomécanismes électromécaniques classiques (moteur diphasé 400 Hz et génératrice tachymétrique). La mesure de l'écart entre les cadres et le gyroscope est faite par détecteur inductif. Ces deux cadres intermédiaires constituent chacun, une plate-forme horizontale ; ils supportent, l'un, le gyroscope de verticale, l'autre, le gyrascope directionnel.

Le gyroscope de verticale est surveillé par deux détecteurs à gravité (au mercure) portés par le cadre intermédiaire.

Le gyrascope directionnel est tatalement libre en roulis. La mesure de l'angle de cap se fait par une chaîne de synchradétection comprenant :

- un synchrodifférentiel qui permet de réaliser le calage initial,
- un synchrodétecteur porté par le gyroscope,
- un synchrotransmetteur commandé par un servomécanisme.

Le synchrodifférentiel est également utilisé pour transmettre les ardres de changement de cap. Il est placé à l'extérieur de l'ensemble gyroscopique.

Les perfarmances de la plate-forme sont les suivantes :

- dérive du gyroscape directionnel : $\Delta\Psi < 3°/h$,
- dérive du gyrascape de verticale : en l'absence d'érection : $\Delta\theta$ au $\Delta\varphi <$ 3 degrés en 10 minutes,
- précisian sur la détectian de cap <3 minutes, d'angle
- sortie en tangage limitée en précision par le potentiomètre de détection à 0,3°.

LA CHAINE D'ASSERVISSEMENT EN ROULIS (Fig. 8)

Elle élabore et transmet aux gauvernes les ordres de stabilisation en raulis du missile.

Une embardée en roulis du missile φe , se traduit par un décalage φp de la plate-forme. L'asservissement propre de la plate-farme tend à annuler φp avant même que φe soit annulé. La référence φ réf. donnée par le gyra de verticale est comparée par le détecteur B3 à l'inclinaison φp de la plate-farme. Le signal électrique φ réf. - φp est amplifié et excite le moteur M3 qui entraîne simultanément la plate-forme (retour d'asservissement) et, par un réducteur de rapport 1/1, le rotor du détecteur inductif Di qui élabore la loi de pilotage en cap et raulis.

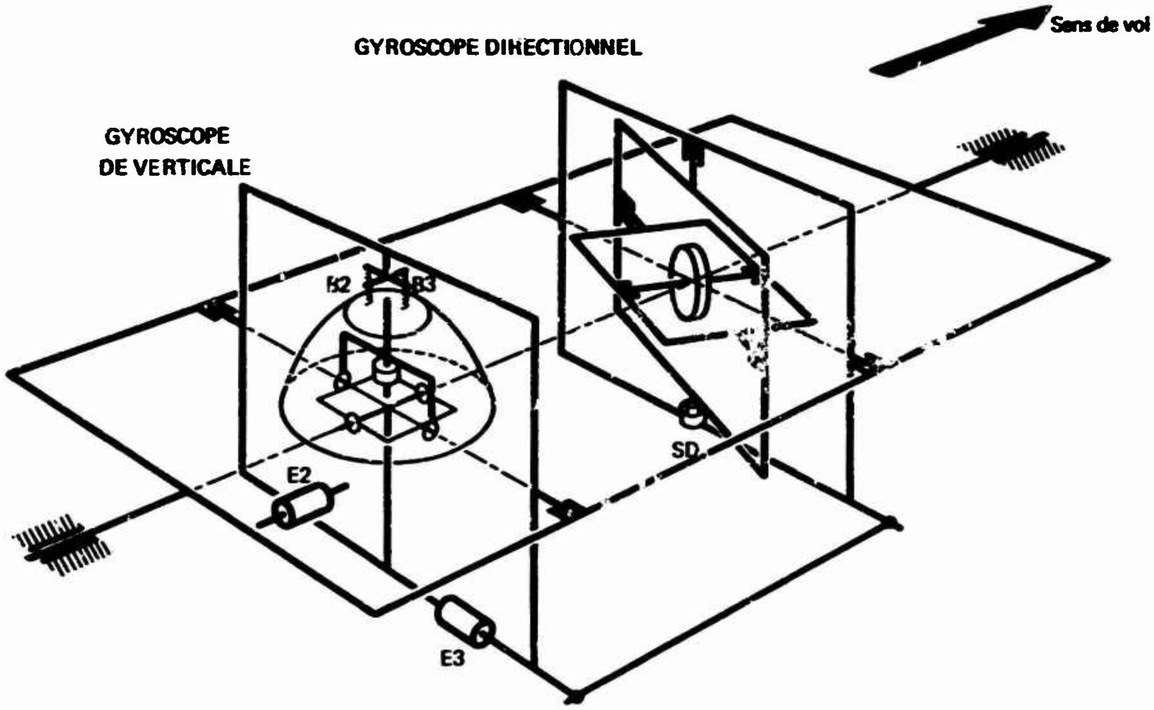


Fig. 7 – PLATEFORME GYROSCOPIQUE
sans les servo-moteurs

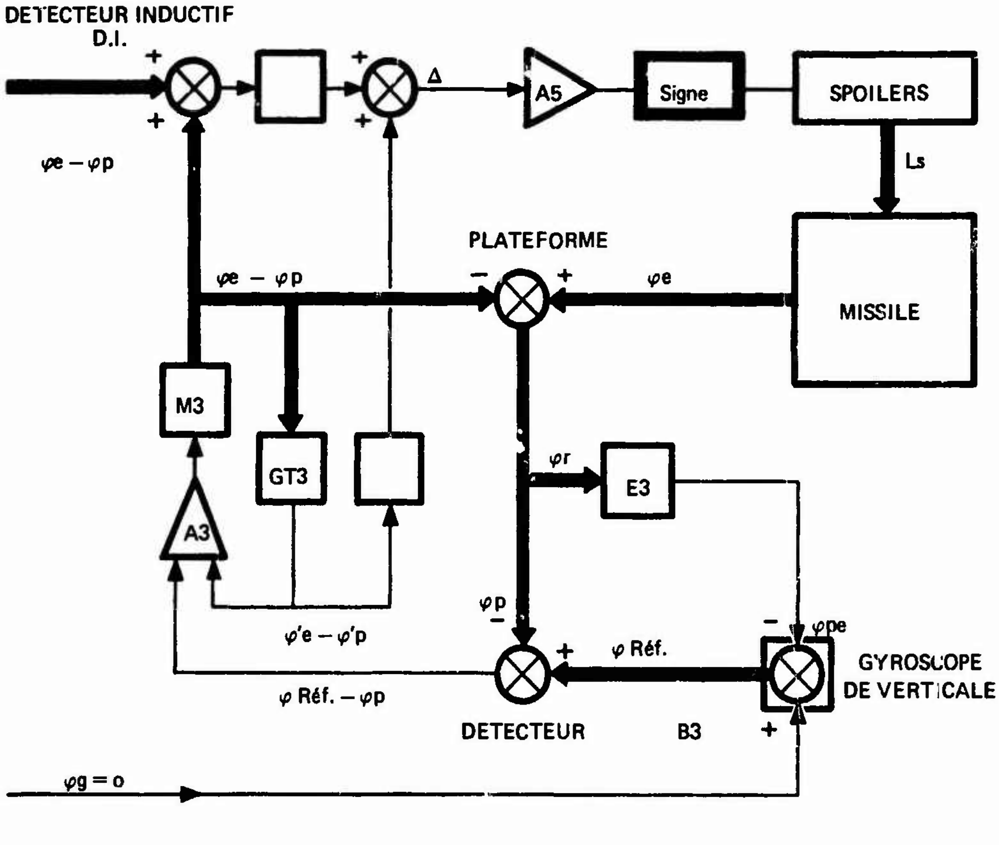


Fig. 8 – CHAINE D'ASSERVISSEMENT en ROULIS

Le signal électrique de sortie de ce détecteur est :

$$\Delta 1 = k_2 \; (\varphi e - \varphi \rho + k' \alpha_T \;)$$

Une génératrice tachymétrique GT3 accouplée à M3 fournit les termes d'amortissement du servo de la plate-forme et du missile. Ce dernier terme vaut :

$$\Delta 2 = k_1 \; (\dot{\varphi} e - \dot{\varphi} \rho \;)$$

D'autre part, le gyro de verticale est asservi à la référence $\theta = 0$, grâce à l'érecteur E 3 à niveau de mercure, monté sur la plate-forme.

Les signaux $\Delta 1$ et $\Delta 2$ sont additionnés et amplifiés pour répondre à la loi de pilotage en roulis :

$$\Delta = k_2 \; (\varphi e - \varphi \rho + k' \alpha_T \;) + k_1 \; (\dot{\varphi} e - \dot{\varphi} \rho \;)$$

et le signal obtenu attaque l'ampli tout ou rien de commande des spoilers.

LA CHAINE D'ASSERVISSEMENT EN TANGAGE (Fig. 9)

La chaîne de tangage correspond à une stabilisation en assiette et à une surveillance en altitude.

A l'intérieur de la centrale tout se passe en assiette θ de la même façon qu'en roulis. Le missile présente une assiette θe , qui tend à être annulée après que soit annulée l'assiette $\theta \rho$, de la plate-forme. Celle-ci est asservie à rester horizontale par le moteur M2 excité par le signal amplifié de l'écart θ réf. - $\theta \rho$ obtenu en comparant dans le détecteur B2, l'assiette $\theta \rho$ et l'assiette θ réf. de référence du gyro de verticale. L'erreur d'asservissement sur θ réf. est éliminée par l'érecteur E 2 à niveau de mercure. La génératrice tachymétrique GT2 assure l'amortissement.

L'angle $\theta i = \theta e - \theta \rho$ est transformé en signal électrique par un potentiomètre et on injecte le signal $K \theta i$ dans l'amplificateur d'addition qui élabore la loi de pilotage en tangage. Suivant les cas de vol, on obtient un braquage de gouverne désiré βd , répondant aux différents cas de pilotage.

Le signal $\beta d - \beta gl + k_2'' \beta' g$ est amplifié, démodulé et vient exciter le moteur M21 de commande du vérin de profondeur qui actionne la gouverne.

LA CHAINE D'ASSERVISSEMENT EN CAP (Fig. 11)

Un écart de cap est détecté par le synchrodétecteur monté sur l'axe de lacet du gyro directionnel.

L'utilisation et le montage du synchrodétecteur directement sur l'axe de lacet, lui-même stabilisé verticalement, suppriment les erreurs de transmission, de cardan ainsi que les perturbations dues au couple.

On obtient la mesure du cap par recopie par un servomécanisme utilisant la chaîne de synchrodétecteur. (Voir figure 10, page 14)

L'ordre de commande de cap αd est élaboré par le servomécanisme de commande du synchrodifférentiel (ampli A 11 moteur M 11, réducteur R 11 et génératrice tachymétrique d'amortissement GT 11).

Il est composé de deux termes :

- l'angle de recalage Ψo
- l'ordre éventuel de changement de cap $\Delta \alpha d$

$$\alpha d = \Psi o + \Sigma \Delta \alpha d$$

et $\Delta \Psi$ est l'écart de cap détecté par le synchro SD :

$$\Delta \Psi = \alpha d - \Psi i$$

Cet écart $\Delta \Psi$ correspond à l'angle de correction $\alpha \Psi$ affiché par le moteur M_1 , qui est écrêté par le limiteur à $\pm 13°$.

Alors, $\alpha_T = \delta \Psi$ pour $\delta \Psi \leqslant 13°$

$\alpha_T = \pm 13°$ pour $\delta \Psi > 13°$

L'amortissement du servomécanisme (voir figure 11, page 14) est assuré par la génératrice tachymétrique GT_1.

Un réducteur mécanique transmet l'angle α_T , dans le rapport k', au détecteur inductif qui élabore la loi de pilotage en combinant les termes de cap et de roulis.

La boucle asservie se referme conformément au diagramme fonctionnel de la chaine de cap.

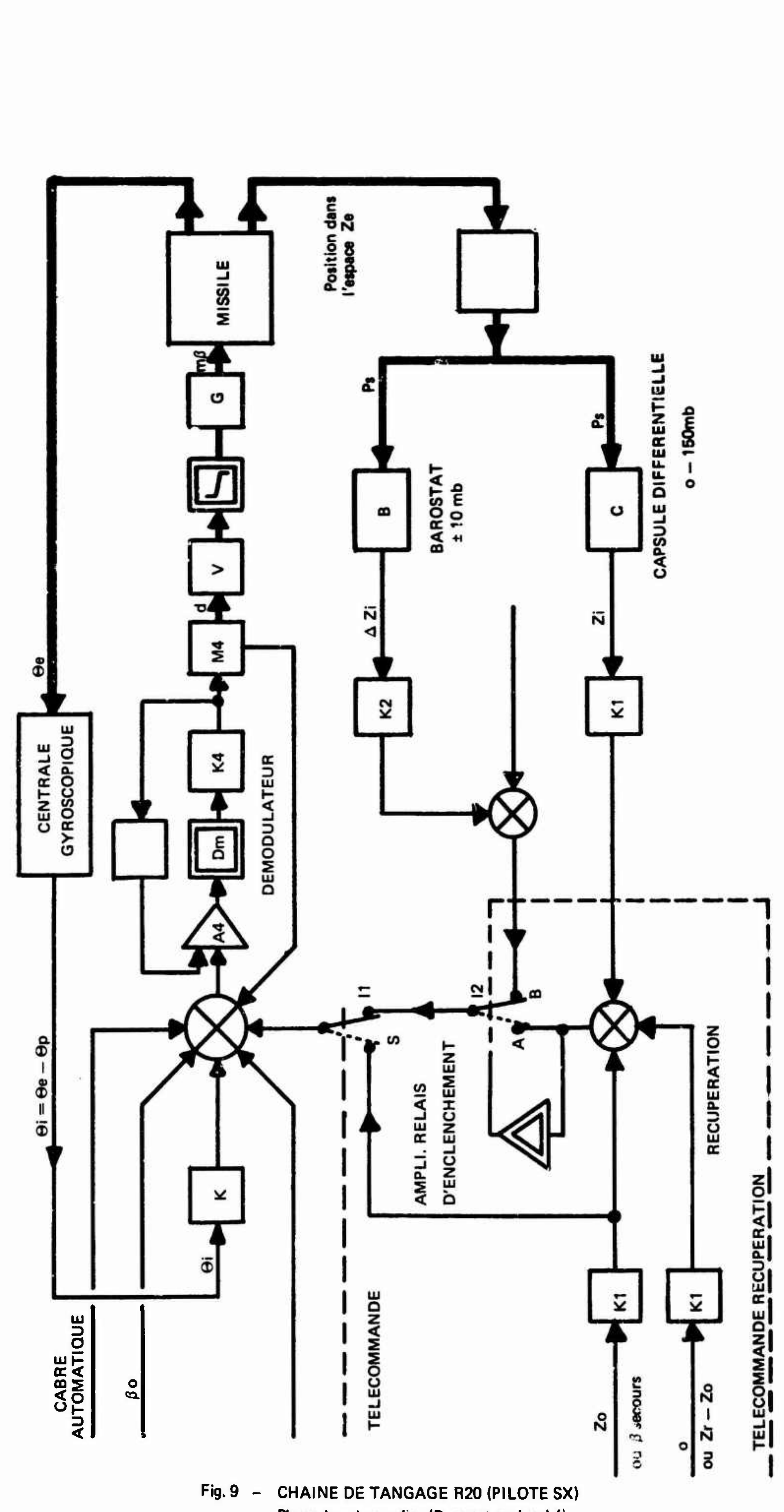


Fig. 9 – CHAINE DE TANGAGE R20 (PILOTE SX)
Phase de vol en palier (Barostat enclenché)

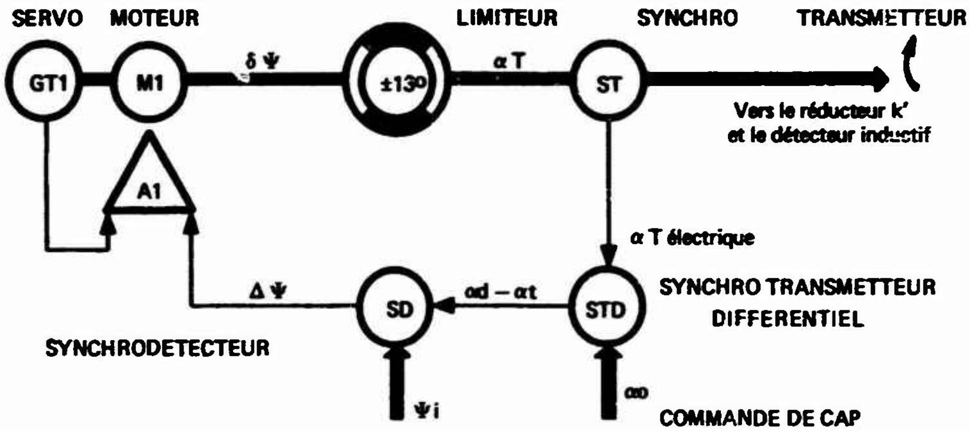


Fig. 10 – CHAINE DE SYNCHRODÉTECTION

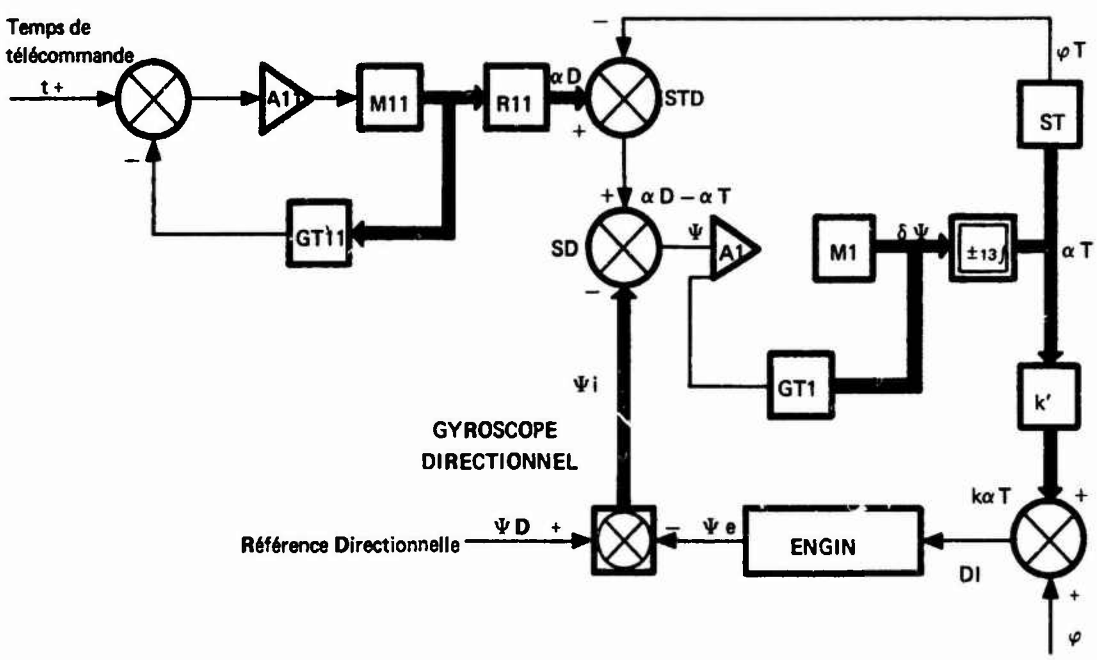


Fig. 11 – DESCRIPTION DE LA CHAINE D'ASSERVISSEMENT EN CAP

SONDE BAROSTATIQUE

La sonde barostatique fournit au pilote automatique le signal de référence permettant d'assurer le vol à altitude stabilisée.

Le schéma électrique et pneumatique de l'ensemble est représenté Figure 12 ci-dessous.

Au départ du missile, l'électrovanne E1 se ferme, l'électrovanne E2 reste ouverte. La pression sol, règnant dans l'enceinte thermostatée fermée au départ sert de référence, pendant tout le vol, à la capsule 0,150 mb.

Pendant la phase de montée du missile, cette capsule mesure la différence entre la pression statique de vol et la pression de référence.

La tension du potentiomètre lié à cette capsule est comparée, dans le pilote automatique à une tension de référence V affichée avant le départ du missile.

Quand ces deux tensions sont égales, l'électrovanne E2 se ferme.

La pression règnant dans le réservoir R2 fermé sert alors de référence à la capsule $\pm$ 10 mb.

Les indications données par le potentiomètre de cette capsule permettent de stabiliser le missile, autour d'une isobare.

Le réservoir R2 a une contenance de 0,5 litre, il est calorifugé.

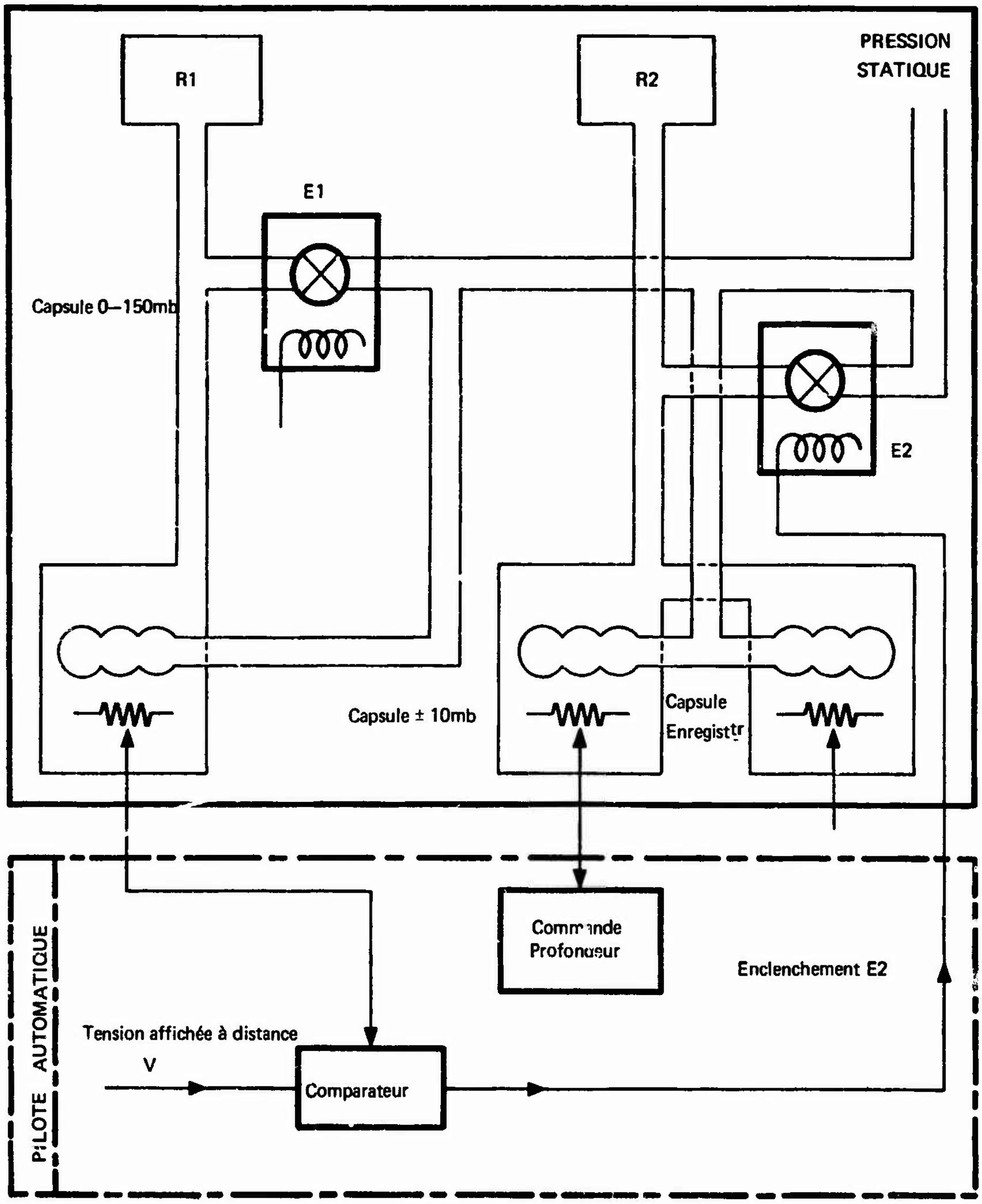


Fig. 12- SCHEMA DE PRINCIPE DE FONCTIONNEMENT DE LA SONDE BAROSTATIQUE

LE PROGRAMMATEUR

Le programmateur est un appareil capable d'emmagasiner, avant le départ du missile, un certain nombre d'informations, de durée déterminée et échelonnées dans le temps. Au cours du vol programmé, il devra les restituer avec précision, aux instants prévus. Les signaux délivrés par le programmateur, agissant sur les divers équipements intéressés, permettent au missile d'exécuter toutes les manoeuvres prescrites, conformément au programme préaffiché avant le décallage. Celles-ci consistent essentiellement en virages, mise en marche et arrêt des équipements de surveillance, éventuellement changement d'altitude, démarrage d'enregistreurs.

CARACTERISTIQUES GENERALES

- Le programmateur est un appareil statique.
- Il permet de passer 56 débuts ou fins d'ordres sur 9 voies différentes.
- Les temps affichables sont compris entre 0 , 1 et 599, 9 sec. Le pas d'affichage est de 0, 1 sec.
- Cet affichage se fait au moyen de fiches placées sur un panneau de programmation (sous forme digitale décimale binaire).
- Précision 10^{-4} de - 30° à + 70° C.

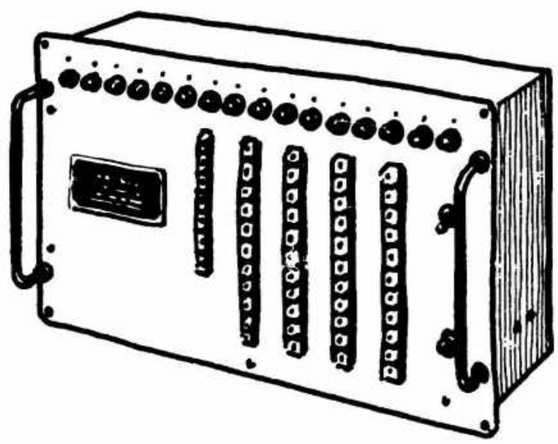


Fig. 13 – CONVERTISSEUR DECIMAL BINAIRE

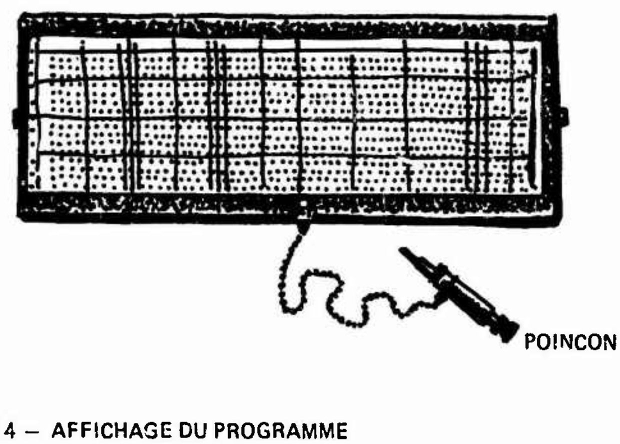


Fig. 14 – AFFICHAGE DU PROGRAMME

FONCTIONNEMENT

Le fonctionnement est entièrement digital, y compris celui de la base de temps, ce qui permet d'utiliser uniquement des circuits statiques "tout ou rien".

L'affichage est réalisé suivant le code binaire décimal. Un compteur cammandé par une horloge, est déclenché à l'instant zéro du programme et faurnit le temps réel sous forme digitale suivant le même code que celui adopté pour l'affichage.

Un circuit logique d'identité campare, paur chaque temps affiché, le temps réel au temps programmé.

Le programmateur permet l'affichage de 56 temps répartis suivant 9 voies.

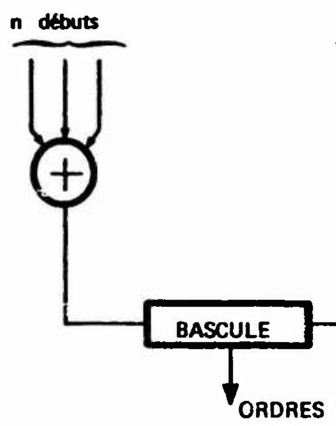


Figure 15

Pour chaque voie les tensions correspondant à un début d'ordre commandent l'état d'une bascule, les tensions correspondant à la fin d'ordre l'autre état de cette même bascule. L'état de la bascule représente alors l'ordre à transmettre.

-=-=-=-=-=-=-=-=-

CONCLUSION

Après 8 années d'expérimentation ayant donné lieu à plus de 80 vols, nous sommes en mesure d'avancer

- la fiabilité du dispositif,
- la reproductibilité des performances,
- la précision satisfaisante de survol.

Nous développerons plus particulièrement ce dernier point.

L'analyse des résultats réels par rapport aux vols programmés a montré qu'il y avait lieu de décomposer l'erreur en deux termes :

- une erreur systématique :
  . imputable aux erreurs d'affichage de l'opérateur humain,
  . aux variations des conditions météorologiques moyennes entre la zone survolée et la zone de départ,
  . aux variations dans le temps entre les conditions théoriques de calcul de préparation du programme et les conditions réelles au moment du survol.

(Voir figure 16, page 18).

Les erreurs aléatoires varient d'un vol à l'autre d'une manière imprévisible dans une certaine plage. Les erreurs aléatoires peuvent être caractérisées par leur ECP moyen qui varie entre 90 et 150 m pour des durées de vol de 15 à 20 minutes.

L'erreur systématique résiduelle peut avoir un caractère croissant en fonction du temps. Dans un certain nombre de cas cependant on a pu arriver à la laisser croître dans un sens puis à la faire décroître, pour l'annuler en un temps précis.

La Figure 17, illustre un vol réussi de ce genre, obtenu en introduisant une correction finale 2 à 3 minutes seulement avant le vol.

La Figure 18, donne une idée de la complexité des vols effectués.

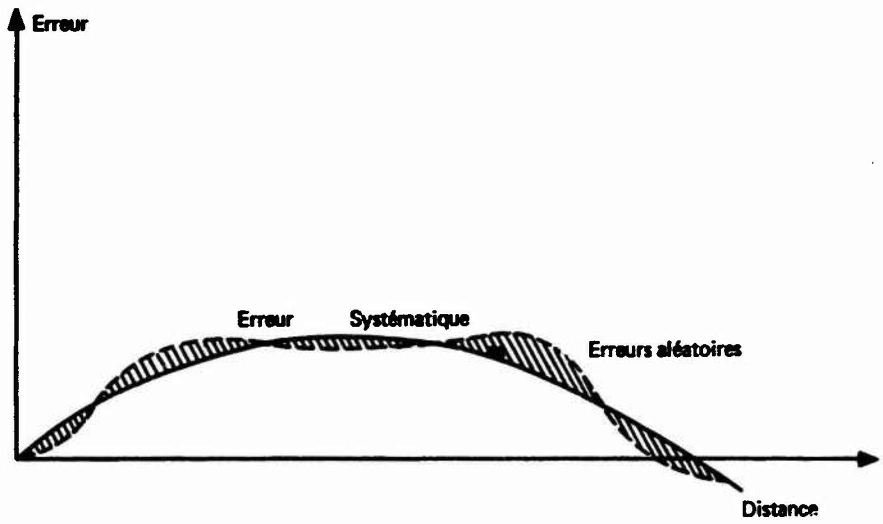


Fig. 16 – REPARTITION DES ERREURS ALEATOIRES ET SYSTEMATIQUES

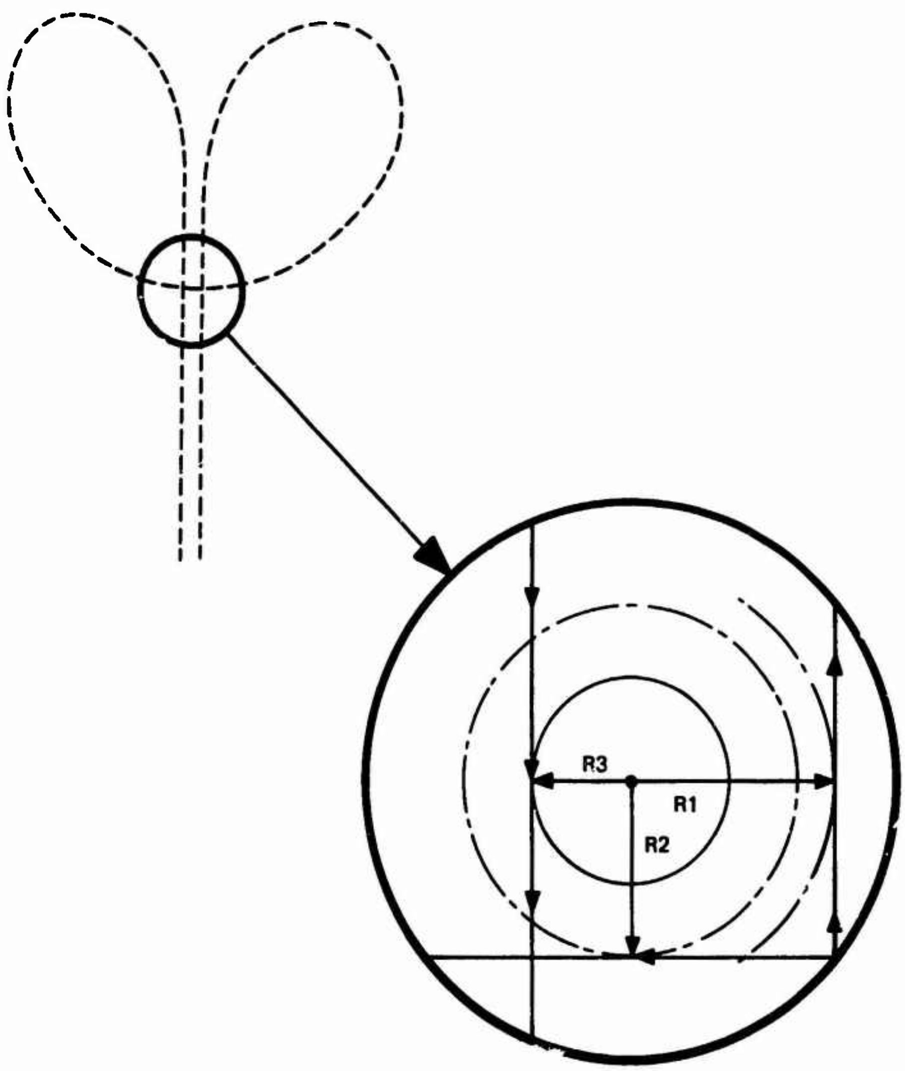


Fig. 17 – EXEMPLES DE SURVOLS SUCCESSIFS DE PLUS EN PLUS PRECIS D'UN POINT DONNE

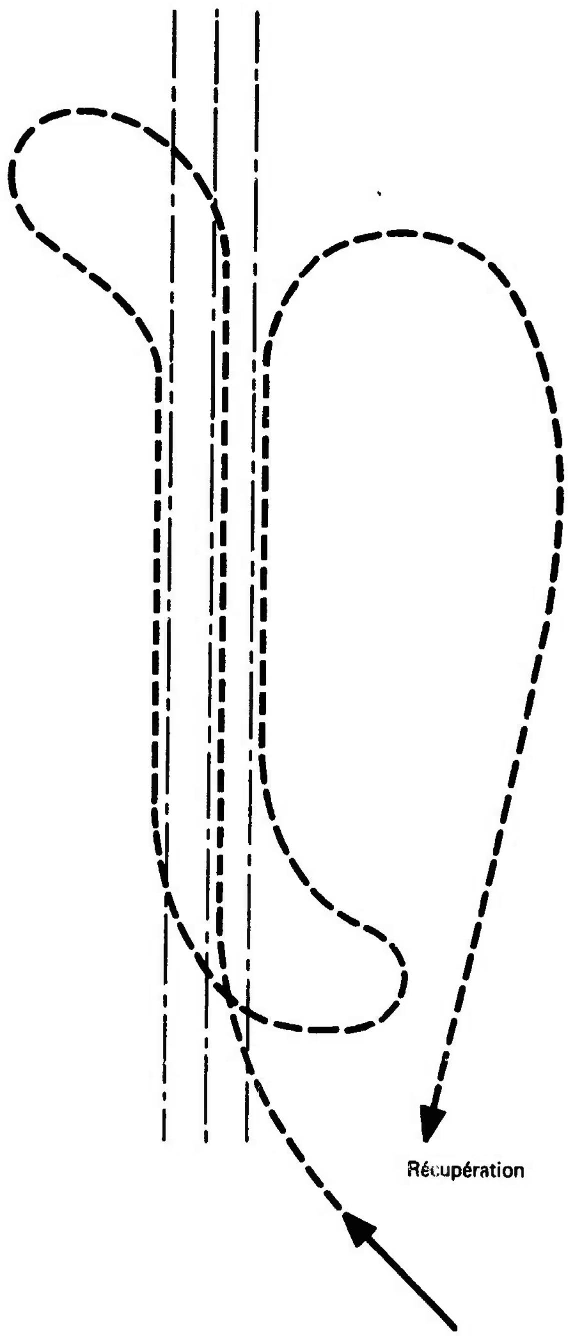


Fig. 18 - EXEMPLE D'UN VOL PROGRAMME COMPLEXE AVEC SURVOL DE TROIS AXES PARALLELES

THE EXPERIMENTAL EVALUATION OF AUTOMATED NAVIGATION SYSTEMS

J.G. Carr
Royal Aircraft Establishment
Farnborough, Hants, England

SUMMARY

Certain aspects of automated avionics systems which are being examined in the RAE Comet exercise are described. The emphasis is on navigation systems and includes the work on digital computers and on-board digital communication techniques, software developments including the use of high level programming languages, and the use of computer controlled electronic displays.

The laboratory work using simulated navigation sensor inputs into an experimental system comprising a digital computer and electronic displays is described. This work has enabled control systems and display formats to be evolved suitable for current and future aircraft projects, and software to be developed and evaluated for the airborne trials.

A Comet 4 aircraft has been re-equipped as a flying laboratory for this work. The installation in the cabin of the aircraft and some of the current experimental investigations are described. The cockpit of the Comet has also been modified by the addition of experimental electronic displays to the second pilot's instrument panel.

1 INTRODUCTION

If we define automation as the automatic performance of operations which were previously manual, then we already have many examples of automation in aircraft that are flying to-day. The evolution towards a fully automated aircraft is taking place gradually as confidence in the reliability and efficiency of the equipment is built up. However, before a given step towards automation is taken there are many alternatives to consider and consequences to weigh and it is hoped that the experiments to be described in this paper will play their part in steering this evaluation along the best path. The field is very wide and with limited resources only certain aspects can be studied. In the first instance the experiments in the Comet are concentrating on navigation sensors and systems although there are parallel programmes of work on flight control and weapon aiming systems in other mor manoeuvrable aircraft.

The Comet is an ex-airline passenger aircraft which has been converted into a flying laboratory mainly for this investigation. The aircraft has just been commissioned and the equipment and facilities it contains will be described later. Whilst waiting for the Comet 4 modifications to be completed laboratory work and preliminary flight trials in an older Comet 2 aircraft have been pursued and have greatly assisted in the planning of the programme for the Comet 4.

The heart of the experimental system is a digital computer and a system is built up by feeding this with appropriate sensor outputs and then supplying the processed information to suitable displays and controls. No real attempt has been made to build up a set of hardware into a complete system suitable for installation in a service aircraft, but rather to combine together versatile units in a flexible manner so that alternative configurations can be readily evaluated as far as principles, software, performance etc are concerned and to leave equipment manufacturers to draw on the results of the experiments in constructing systems to suit particular aircraft. Thus, in practice the equipment we use in a given trial may comprise some fully developed production items with some experimental breadboard type units; the whole system being combined together through a digital computer for which the program has been developed and written inhouse.

The emphasis in this paper will be on navigation systems and will include the work on digital computers and on-board digital communication techniques, software development, including the use of high level programming languages and the use of computer controlled electronic displays. The work will be illustrated by reference to 3 stages (a) laboratory work using simulated inputs and situations, (b) airborne laboratory work in the cabin of a Comet using real inputs and aircrew observers, (c) airborne cockpit work using flight crew and real situations.

2 LABORATORY INVESTIGATIONS

Laboratory work is a prelude to and is complementary to the flight trials. For economic reasons as much of the investigation as possible is done in the laboratory in rigs which as far as possible duplicate the aircraft installation. This has enabled software development to proceed continuously, and the flight programs to be checked in the laboratory by using firstly simulated inputs and secondly inflight recorded data as inputs. This latter facility has been particularly useful in the development of a Kalman filter for an integrated navigation system.

2.1 The Computer

The computer used for most of the work is a Ferranti Argus 400. These computers with other members of the Argus series are widely used in many process control types of applications but have not yet been adopted for any airborne use other than at the RAE.

The Argus 400 is a 24 bit serial arithmetic machine with 8 accumulators, 3 of which can be used as modifiers or index registers. Each of the computers has an 8000 word 2 μs core store and a comprehensive input/output unit so as to cope with a range of types and quantities of peripheral equipment. For example each computer input/output unit has 48 digital inputs, 10 digital outputs, 72 analogue voltage inputs, 2 synchro inputs etc, there is also a 24 bit parallel direct store access system via a multiplexer for priority scheduling.

2.2 Computer Languages

Until recently all program coding for this work has been done by RAE staff using a very basic assembler language 'April' this is about one step up from the basic machine code in that relative addressing can be used and there is a cram facility for specifying independent bit fields within a single word which is useful for packing in display formats. A symbolic assembler language 'Astral' eventually became available for Argus but it was not made extensive use of because it was not considered a significant step forward. After some experience of coding in basic assembler it became clear that a high level language of a particular type was required, not only for ease of programming but also for better communication and improved documentation. Implementation of the well known languages such as FORTRAN or ALGOL tend to be excluded from the airborne real time small computer application because in general they take up much more store and have significantly longer run times than the equivalent code version. This is largely because the programs have to run with parts of the compiler in the store and a program for any particular operation is likely to be far from optimum.

A fairly recent development at RRE is the specification of a suitable high level language called CORAL 66. A principal feature of CORAL is its ability to analyse the source code program using one track syntax analysis techniques and to generate code to perform the operations required. The language itself bears a superficial resemblance to ALGOL 60 but the facilities are tailored to meet the special needs of real time programming. A complete description of the language is published by HM Stationery Office under the title 'Official Definition of CORAL 66' and a user's guide appears in RAE Report TR70102 'A Guide to CORAL Programming'.

CORAL compilers now exist for a number of computers in UK, others are being written and the efficiency of these compilers in real time avionics applications is being assessed. To facilitate this evaluation and also to study the techniques for producing very efficient code an Argus 500 computer with a full range of peripherals such as a line printer, CRT display and drum backing store is being used.

2.3 Displays and Controls

In an integrated system a versatile display of some kind is an essential link between crew and computer. The most suitable device at the present time is a cathode ray tube or CRT, these can now be obtained in a compact and rugged form and with sufficient brightness for use on the pilot's instrument panel. Other types of electronic display are being developed, such as plasma or solid state matrices, and liquid crystals, and these may eventually take the place of the CRT which is rather bulky and power consuming.

When a CRT is driven from a digital computer it is capable of showing (by software changes only) an almost infinite variety of formats both alpha-numeric and diagrammatic and some tubes can also show TV type pictures either separately or superimposed on the alpha-numeric information. Thus it is possible to replace many conventional cockpit instruments by a few CRT displays and to present the crew with suitably processed data for immediate use. There is also the opportunity to avoid a multitude of separate selector switches by time-sharing a few switches through the computer in some form of multi-function keyboard. Essentially a multi-function keyboard is linked to the computer in such a way that the function of each key and the label associated with it changes according to the mode selected. The labels for the switches could be of the optical projection type, or small addressable 6-10 letter matrix labels either using electro-luminescent, plasma or liquid crystal techniques, but at the present time the most convenient method seems to be to write the labels on the face of the CRT in association with an adjacent row of push buttons.

There has been a good deal of laboratory work at RAE using this technique and Fig.1 shows the laboratory rig used for this work. In addition to the cursive CRT display and keyboard the rig carries a TV raster CRT, a projection type topographical display, and a number of controls such as multi-function keyboards, and joysticks and rolling balls for controlled movement of cursors and markers on the display. The displays are driven from an Argus 400 computer and in order to provide a dynamic display the computer also carries a simple flight simulation in which track and speed, for example, can be varied and various navigation and steering calculations performed.

In an aircraft equipped with conventional displays the aircrew spend considerable time and effort scanning an array of instruments, and this is mostly to determine either whether a parameter is between certain limits or whether its rate of change is between certain limits. Both these decisions are ideally suited to the logic of a digital computer. Only if values are outside these limits do the crew need to know their actual values, they could then be displayed automatically and the crew alerted. However, there is danger in this philosophy if carried to extremes. One problem is to be sure that the right information is presented to the man in a form to which he can react quickly in an emergency. A second problem is to retain the man's confidence that all is well when it is, it is no good saying nothing until something goes wrong, the man will probably be asleep by then anyway. The man's attention must be retained and this is difficult if he normally has nothing to do. This brings us to the idea of display modes. These are display formats which keep the crew appraised of the total flight situation at each phase of the flight and their use can be illustrated with reference to the display of navigation information. We divide a typical flight into a number of phases for each of which a special format is devised which displays parameters which experience has shown are the ones most often required during that phase, these modes are selected from the keyboard and having been selected the keyboard function changes so that a whole set of subsidiary facilities are available appropriate to the mode selected. Fig.2-5 show an experimental keyboard and a selection of display formats chosen to illustrate the versatility of the display and keyboard. Fig.2 illustrates the horizontal situation display which can be obtained at any time by pressing the NAV key. It gives on a dynamic display not only the usual HSI information such as across track error and track angle error both numerically and diagrammatically, but also aircraft position, distance and ETA to next way point, B, track required at B etc. Some of these quantities are optional and can be displayed or removed by the operation of the white key which is associated with the appropriate legend along the bottom of the CRT. When the PLAN format is selected the.e legends automatically change to a more appropriate set as shown in Fig.3. This PLAN format shows on a coarse lat-long grid the relative positions of all the way points (letters) and fix points (numbers) that have been stoed in the computer as the flight plan (this particular flight plan is used in the current flight trials, A is Farnborough, K is North Scotland and J is off the Danish coast). The stored information concerning way point B is displayed, the others can be displayed in succession by a repeated pressing of the white toggle switch at the left of the second row of switches.

These geographic coordinates would normally be entered pre-flight by some convenient means such as card or tape, but if it is required to make a change in flight, pressing the 0-9 key converts the 10 multi-function keys to a numerical typewriter, Fig.4. A steerable flashing marker indicates which character can be changed, and when the amendment has been completed satisfactorily on the display, pressing the enter key substitutes the new data for the old in the computer memory and the route is automatically redrawn. If on the other hand the same way points can be used, but in a different order, the route list can be called up, Fig.5, and amended using the multi-function keys to type letters or numbers where appropriate. In flight, a cross represents the geographic position of the aircraft and should move along the line representing the route, the computer having made the necessary calculations to drive the NAV display and steer the pilot along the great circle route between each pair of way points. There are many other formats that have been devised and flown but the ones described serve to illustrate the versatility and flexibility of the CRT when used with a computer and a multi-function keyboard.

For the above experiments the CRT lines were cursively written. This type of tube has advantages as regards brightness and ease of writing characters and dynamic displays, but the raster driven CRT is more suitable if it is important to combine TV pictures, use shading, or shades of grey etc. Our flight experiments will include work on both types also on colour CRTs as they become available. Only by using colour will the CRT successfully replace the film projection type of topographical map display for military purposes.

The display formats keyboards and controls have evolved as a result of experience and also as a result of frequent discussion with experienced aircrew. Some experiments have been performed by ergonomics and human factors specialists, and as a result such things as an optimum character font and optimum cursor controls have been chosen. Now that this type of display is being adopted for military aircraft the flexibility of the laboratory installation is proving valuable in evaluating the relative merits of alternative configurations

3 FLIGHT TRIALS

A Comet 4 aircraft has been specially re-equipped for this experimental work. Fig.6 shows the aircraft accompanied by a Canberra observation aircraft coming into land at the RAE during the test flying after modification. Most of the passenger seats have been removed from the cabin and replaced by rigs or racks of experimental equipment. The rigs are designed to be secured to the seat rails and there is ready access to signal ducting and power supplies so that the rigs and seats can be easily arranged in the most convenient positions. Figs.7 and 8 show views in the cabin and two basic types of rig can be seen. There is a general purpose rig which can carry 500 lb of equipment 200 lb on each of the bottom and top shelves, and 100 lb on the centre shelf which can be removed or its height adjusted as required. These can be seen as the light-coloured rigs with various superstructures and sets of equipment in Figs.7, 9 and 10. The other basic rig carries the Argus 400 computer installation on ATR short beams and can be seen as the black framework in Fig.8. Fig.9 shows the displays rig which, at the time carried two CRT displays and experimental keyboard. This structure is bolted together and can readily be adjusted to accommodate different arrangements and sizes of display head. Fig.10 shows the rig where most of the navigation datum displays such as Decca Navigator, Loran C and Doppler are situated.

In addition to this equipment in the cabin of the aircraft 2 CRT displays have been installed in the cockpit on the second pilot's instrument panel and this is shown in Fig.11. The aircraft installation for a recent series of flight trials is illustrated in block diagram form in Fig.12 and examples of some of the investigations are given in later sections of this paper.

3.1 Navigation Datum

In order to evaluate the various navigation systems in flight some means of determining the aircraft position must be found, which is more accurate than the system being evaluated. However, this is difficult because our most accurate on-board navigation sensors naturally form part of the experimental system. Photographic fixing, and monitoring by ground based radars such as FPS 16, provide very good accuracy with limitations set respectively by weather and range, and these are used as absolute checks when necessary. For most of the trials a sufficiently accurately datum can be derived by careful post-flight processing and analysis of all the navigation information obtained during the flight using the resources of a large ground based computer installation.

The aircraft is equipped with Tacan, VOR/DME, Decca Navigator, Loran C, Doppler and an inertial navigator and the datum is derived from appropriste combinations of these sensors.

3.2 Data Processing

The on-board computations can be divided roughly into 2 groups, those needed to drive the systems being evaluated and those concerned with the recording and quick-look display requirements. All the calculations are performed in the two Argus computers which are linked together. The information is received by the computer in various forms, for example, the Doppler is a Decca Type 72 and the information arrives as trains of pulses representing along and across track velocities; the inertial information arrives as synchro signals for the 3 angles, analogue voltages for the velocities, and digital whole numbers for lattitude and longitude. This information is converted into a common digital form in the computer input-output unit before being processed into the form required for the experiment. As one of the checks that all is going well a selection of the parameters is printed out on a teleprinter at regular intervals, say, every 5 minutes. This is known as the 'quick-look' facility and enables the observers to check the progress of the experiment. An example of a typical printout is given in Fig.13. It can be seen that this not only includes the geographic position as derived from the various sensors but also, a table of differences between pairs of sensor outputs.

In addition to the quick-look printout the values of all the important parameters held in the computer are recorded on punched paper tape and/or magnetic tape every 30 seconds for further processing and analysis on the ground. The facilities, software etc have been so organised that fully processed data and plots of the results are normally available next day. A typical line-printer tabulation of the data is shown in Fig.14. In this case we were flying in good Decca cover and the computer has converted the Decca coordinates to lat-long, chosen the pair of lines with the best cut and indicated the apparent errors between each pair of navigation systems. The line printer is also used to produce graphs of selected parameters.

All these recording and display devices are very good for experimental work but are not suitable for displaying the processed information to the aircrew. The cockpit displays form a vital link between the computer and aircrew and an important section of the work in the Comet is concerned with displays and their control.

3.3 Displays and Keyboards in Flight

Fig 9 shows the display rig in the cabin of the Comet where the display formats and keyboard computer control systems that have been evolved as a result of laboratory simulation experiments are evaluated with real inputs. The aircraft has a number of spare seats so that a group of aircrew can each have an opportunity of using the displays and controls in each flight. In this way it is hoped to obtain a consensus of opinion as to the suitability or otherwise of the displays so that they can be evolved towards the generally acceptable format. The computer programmers are also the supervisors of the experiments on-board the aircraft, so they can readily comment on the feasibility of any suggested display changes and even amend the program to make slight modifications to the formats during the flight.

The CRT displays will be introduced into the cockpit in the second pilot's position in two stages. In the first stage the pictures will be transmitted from the cabin and he will have no direct control over what he sees but he will be able to judge the pros and cons of the displays in relation to cockpit conditions, conventional displays etc. Secondly, when a suitable control panel and computer program has been evolved the panel will be installed for the pilot's operation and the whole system evaluated. To help in this stage of the work and to avoid conjestion in the cockpit the aircraft will be equipped with closed circuit TV and video recording facilities. This will enable the ergonomics aspects of the system to be studied and evaluated in flight in real situations, without hindering the pilot in his normal duties.

3.4 Kalman Filtering

No discussion of an automatic navigation system would be complete without mention of Kalman filtering, which will almost certainly figure in future systems. This method of combining the outputs of a number of navigation sensors should optimise the overall performance of the system so that either the aircraft geographic position will be known with greater confidence; or simpler and cheaper sensors could be used while still maintaining adequate performance. In the limited experience at RAE so far, there seem to be 2 basic problems. Firstly to ensure that as far as possible spurious or unreliable sensor information is excluded from the calculations, and secondly to ensure that the error models of the sensors carried in the computer are adequate, whilst at the same time avoiding the need for excessive computer capacity.

The experiments, which are at an early stage because the Comet has only just become available, have so far been with a doppler/inertia navigation system and with a simple error model, 7 state variables for the inertial navigator and 2 for the doppler. Improved navigation performance in flight was obtained, but the experiments were only partially successful because of electrical faults in the transmission of both the inertial navigator data and the doppler signals. For example the computer did not receive adequate warning that the doppler had switched to memory with consequent injection of spurious velocities into the system. This fault has now been rectified, and in addition input data is filtered and checked for validity before use in the filter.

In the flight trials, in addition to the processed data, complete records of the raw outputs of the doppler and IN systems are obtained, and these recordings are used subsequently in the laboratory as inputs to a computer simulation of the Kalman filtered system. In this simulation, the effect of varying the number and size of the state variables, variances, plant noise etc has been studied and an attempt made to optimise these parameters for subsequent flight trials.

An alternative program has also been written in which the number of state variables can easily be changed up to a maximum of 12. Successful laboratory simulations have been run with as few as 5 state variables and this is one of the models at present being evaluated in flight. After obtaining a satisfactory performance over land, where there should be a good doppler return, the effect of land sea transitions will be studied. The next stage of the Kalman filtering work will be to include position fixes as inputs in addition to the IN and Doppler.

3.5 Digital Data Transmission

In an automated system using digital computers and digital techniques the transmission of information from point to point within the aircraft presents a major task. Digital transmission has both advantages and disadvantages; for example, there are good techniques for verifying digital information but on the other hand spurious pulses can corrupt the data. A comprehensive program on a number of aspects of data transmission has been started and the hardware resulting from this will eventually be tested in the Comet.

Very little information is available about the type and intensity of impulsive electrical noise to which a digital data transmission system is likely to be exposed in a typical aircraft installation. A start has been made on a survey of aircraft using a small recorder which counts the number and records the size and area of any pulses received on a length of typical aircraft cable which is correctly terminated at each end and which follows the likely path of a signal cable. The thresholds are adjustable but in the initial stages are set at 2, 5, 10 and 20 volts and 5, 10, 20 and 40 microvolt seconds. The coded description of the pulse together with its time of occurrence are automatically recorded for post flight analysis and possible correlation with any known electrical disturbance.

Various types of cable and methods of modulation are used for the tranmission of digital data and as always a compromise has to be made to obtain adequate integrity and isolation whilst not paying too high a penalty in cost and weight. Several programmes of work are underway which should help in arriving at a satisfactory compromise.

(a) Work to determine the effectiveness of various cableforms such as coaxial cables and twisted screened pairs for both tranmitting pulses at a high repetition rate without inducing cross-talk in adjacent cables, and also for rejecting impulsive electrical interference.

(b) The design of line drivers capable of high pulse repetition rates, but forming pulse shapes containing the minimum of the high harmonics which would tend to penetrate the screen.

(c) The design of tranmission systems using modulation methods which are as far as possible immune to interference -- this includes the use of fibre optic links in place of conventional cables.

(d) The study of multiplexing methods to take full advantage of the possibility of time sharing of cables to reduce aircraft cable weights.

(e) Exploring the possibility of inflight automatic self testing of digital systems using a computer based aircraft integrated data system (AIDS) to complement existing tools such as built-in self-test.

4 CONCLUSIONS

The facilities provided by the Comet 4 flying laboratory and some of the programs being carried out in it have been described. The installation has been planned to be very flexible and capable of being modified or extended with a minimum of effort. In this way it is hoped to keep pace with the developing techniques and to provide a satisfactory test vehicle for experimental navigation systems for some years to come.

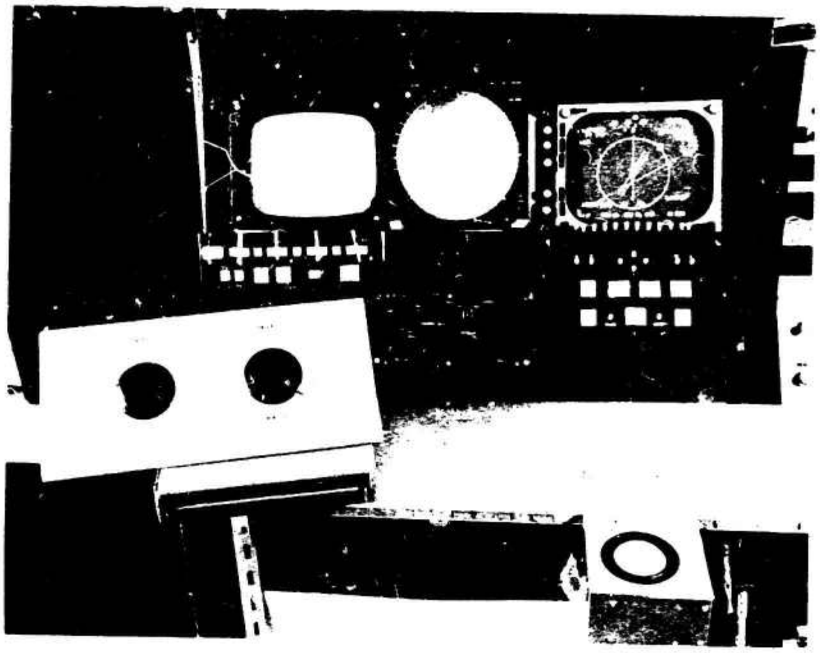


Fig.1 Laboratory rig for displays investigations

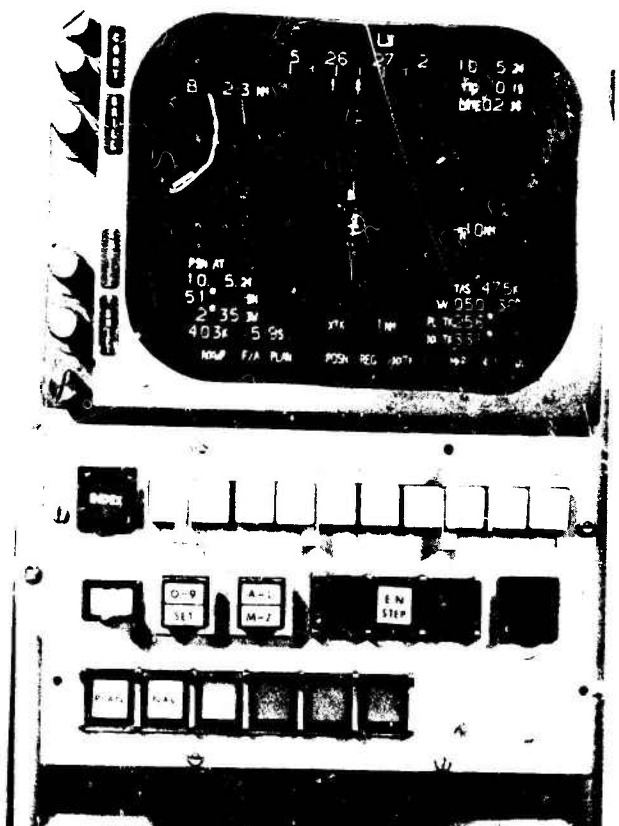


Fig.2 CRT display and multifunction keyboard – NAV format

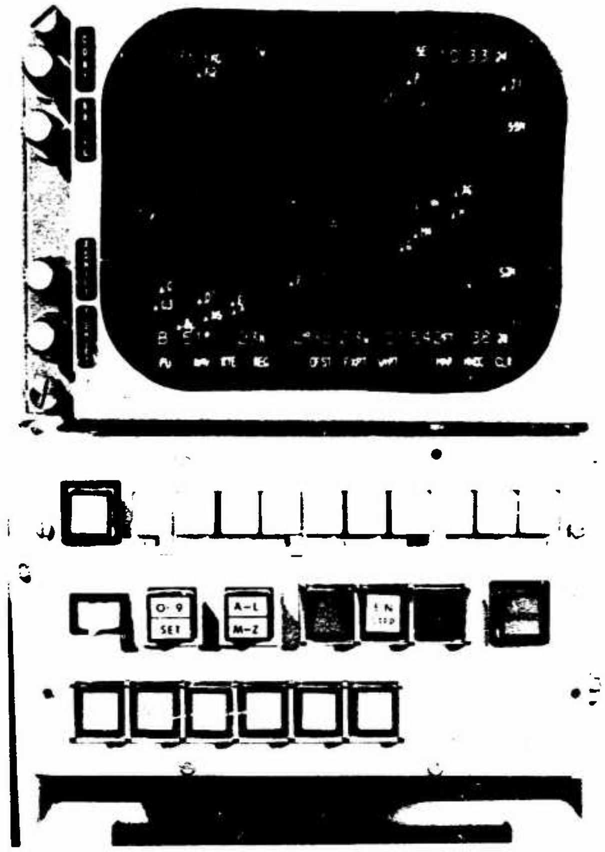


Fig.3 CRT display and MFK – PLAN format

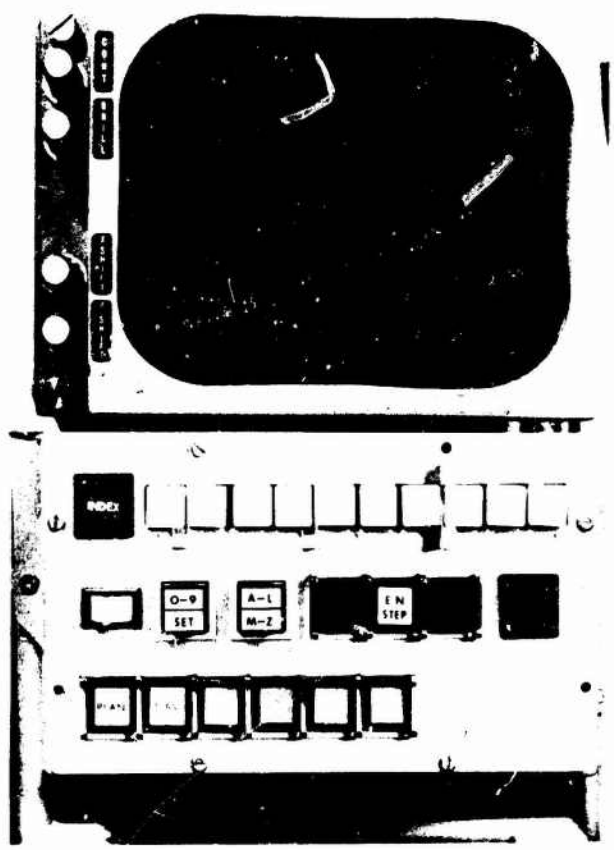


Fig.4 CRT display and MFK – PLAN with numerical keyboard

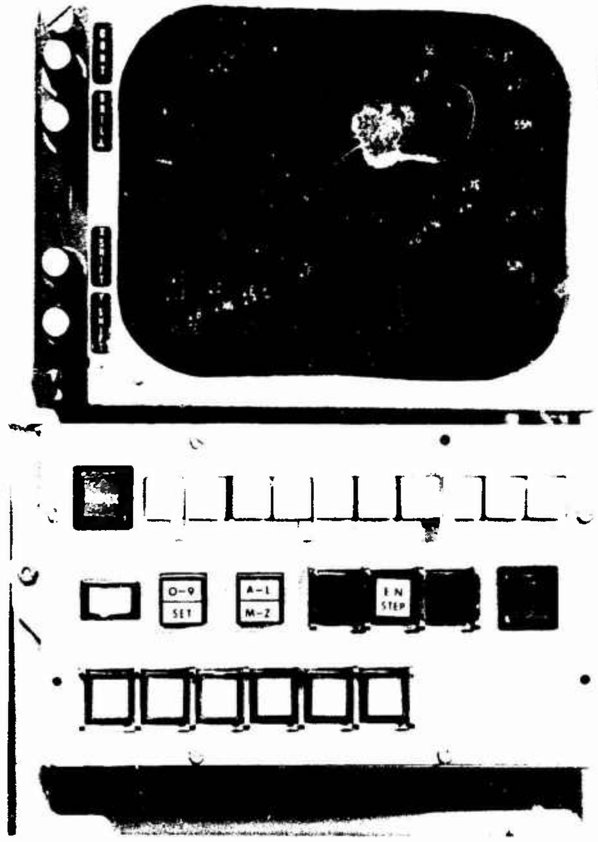


Fig.5 CRT display and MFK – PLAN with alphabetical keyboard

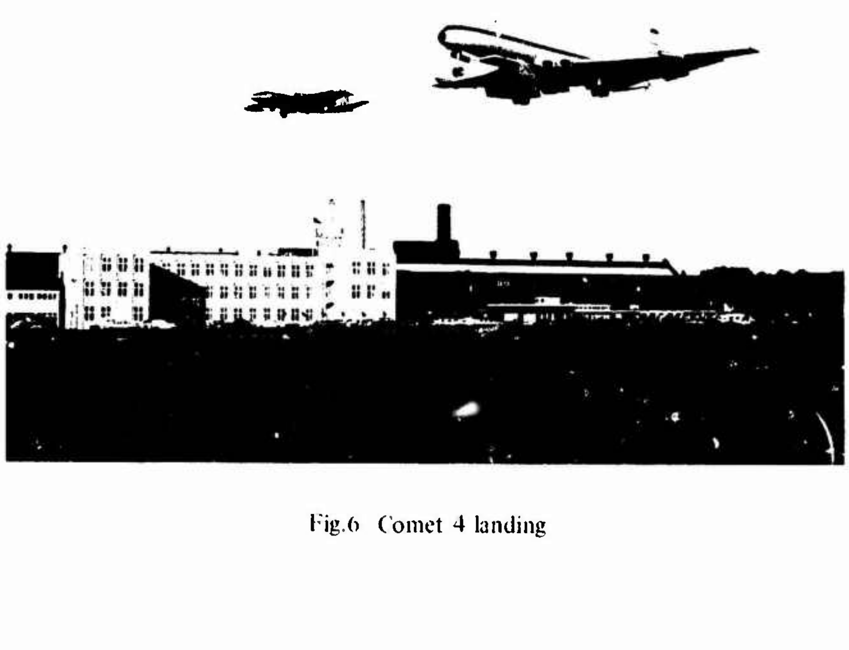


Fig.6 Comet 4 landing

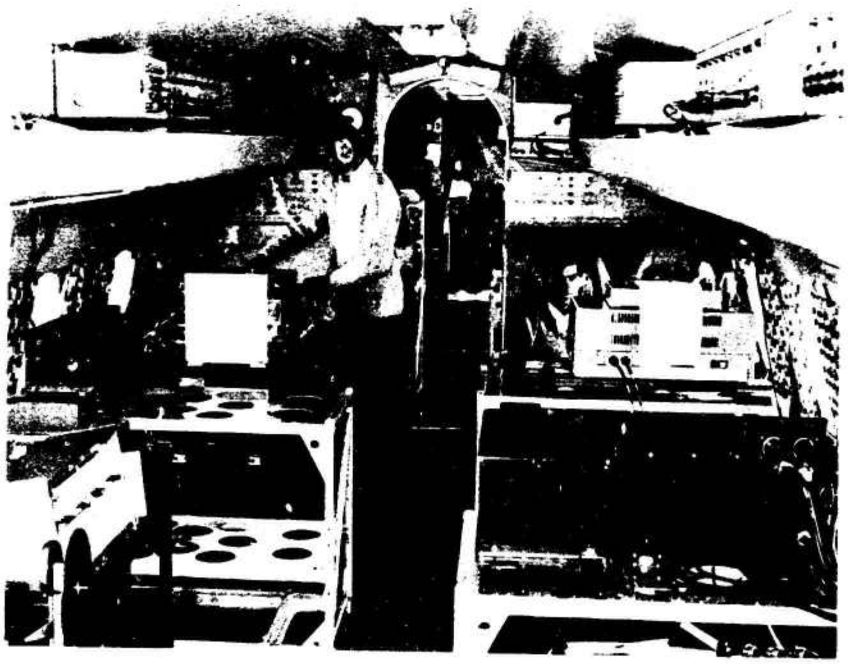


Fig.7 Comet 4 cabin installation

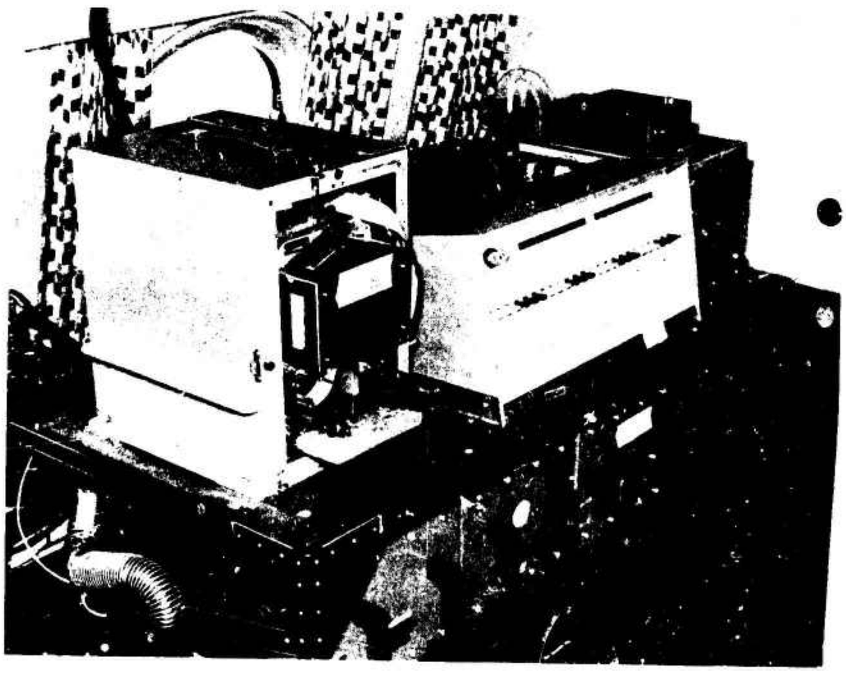


Fig.8 Argus 400 computer rig

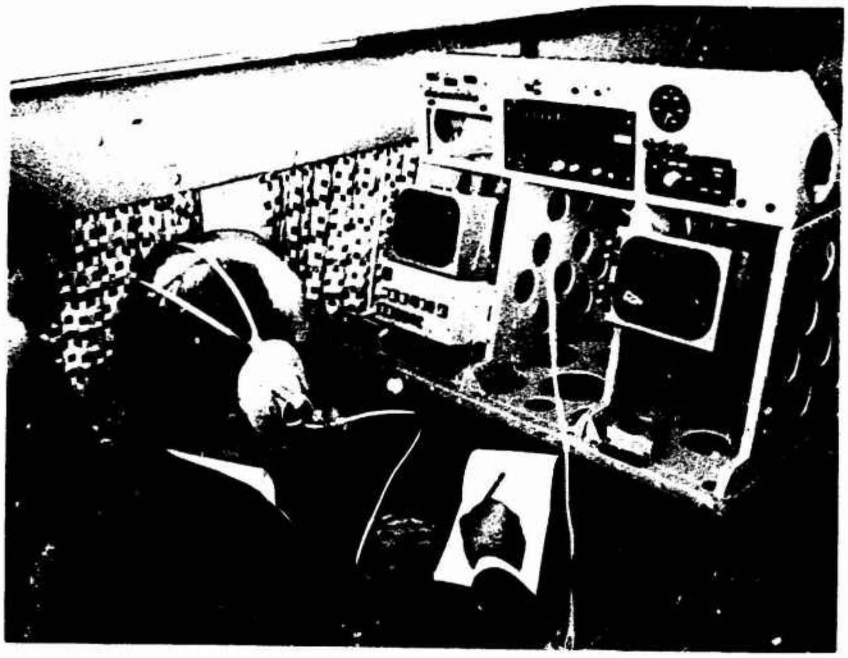


Fig.9 Comet displays rig

Fig.10 Navigation datum displays

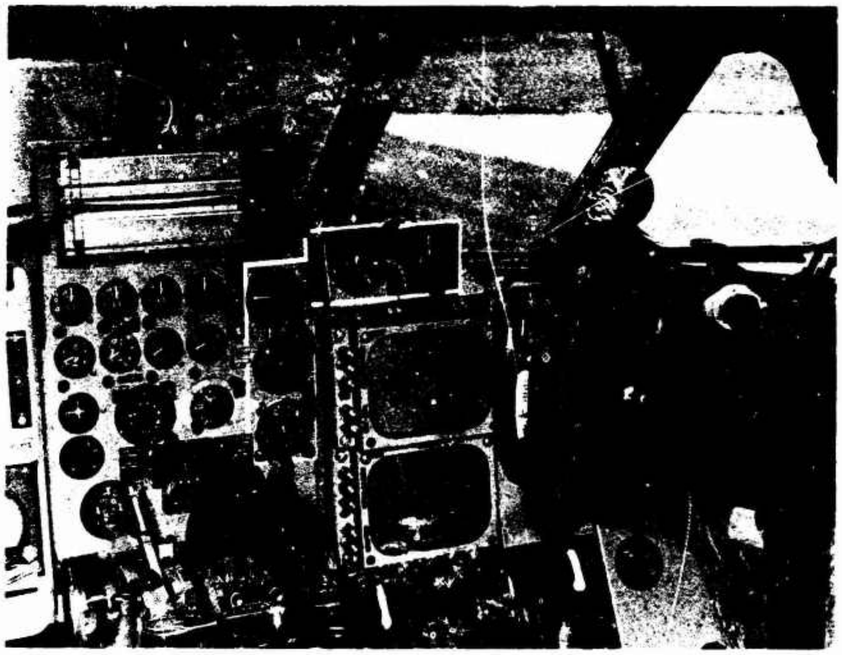


Fig.11 CRT displays in Comet cockpit

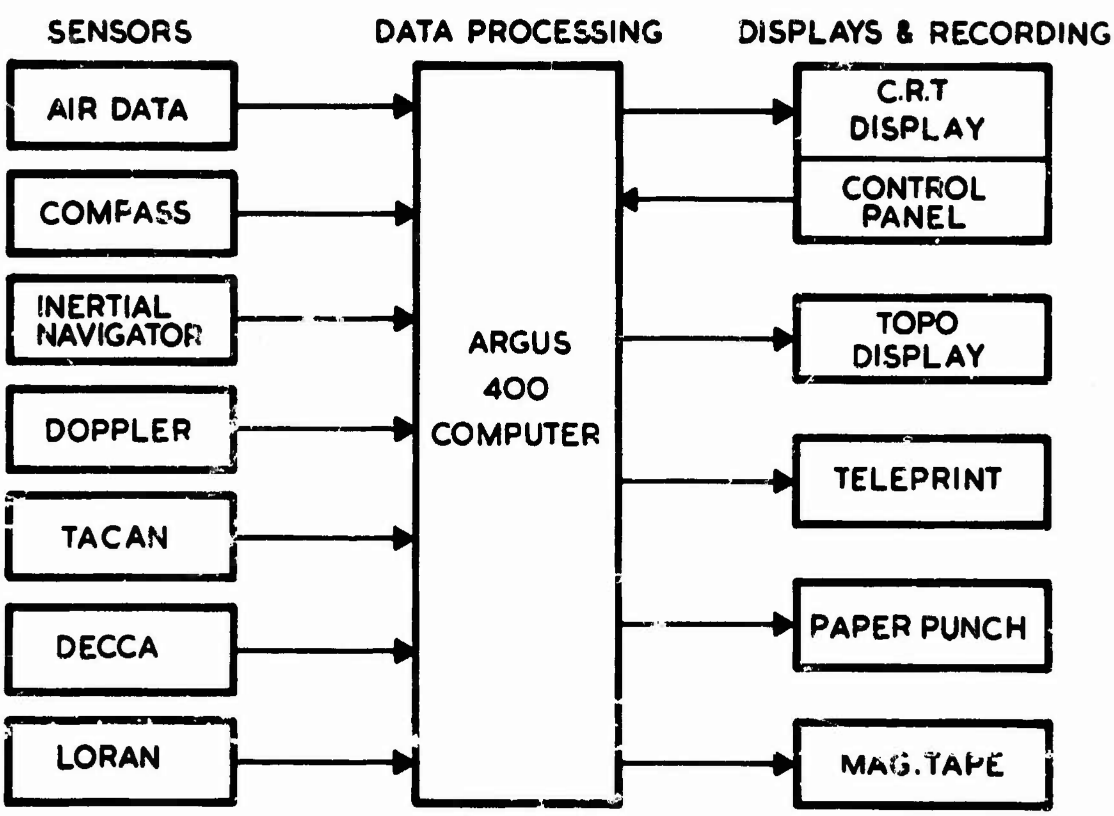


Fig.12 Integrated navigation system – current flight trials

```
00 10 30
  051 12.9   -001 05.7    262 47.5    264 35.1
  051 13.0   -001 05.5    274 14.2    001 47.5
  051 18.1   -001 06.9    256 40.2    025.8
  051 12.5   -001 05.2
  0,407.4   -013.9    0,048.9    298.1    24,408   000.0
  00    05    04    00.58    46.48    67.78
 -000 00.1   -000 00.2
 -000 05.2    000 01.1
  000 00.3   -000 00.5
```

KEY

TIME					
IN LAT	IN LONG	IN HDG	TRACK		
DOPPLER LAT	DOP LONG	COMPASS HDG	DRIFT		
TACAN LAT	TAC LONG	TAC BRG	RANGE		
KALMAN LAT	KAL LONG				
ALG TK ER	AC TK ER	DOP DIST	DOP SPEED	HEIGHT	DOP STATUS
DECCA ZONES			DECCA LANES		
RED	GRN	PURP	RED	GREEN	PURPLE
DOP-IN LAT	DOP-IN LON	DIFFERENCES			
TAC-IN LAT	TAC-IN LON				
KAL-IN LAT	KAL-IN LON				

Fig.13 In flight quick-look – print-out

```
                                                                                                          PAGE 32
                                                                                                        DIFFERENCES
TIME:  13      1      30  =  46890   SECONDS                                                             LAT     LONG
I.N.   54    23.03   -3     0.50      HEADING: 329    16.90  TRACK: 328    34.10
DOP    54    29.90   -3     2.18      COMPASS: 337     9.00  DRIFT:   0  -37.47               IN-DOPPLER  -6.87     1.68
DOP.K  54    29.90   -3     2.18      TRK.KAL: 328    34.10  TRK.ERROR:  0     0.00           DOP-DOPKAL   0.00     0.00
TACAN  54    29.30   -3     2.74      BEARING: 190    54.40  RANGE:      64                   IN-TACAN  -6.27    2.24
KAL    54    24.39   -3     3.21      VEL ERR N.S:     8     E.W:    3  FT/SEC   MARK:  0       IN-KALMAN   -1.36     2.71
  DOP.DIST:    226   N.MILES   GRD.SP:    296   KN   HEIGHT: 23000    23000   FT. DOP UNLOCK:   0.00
      MULCOS:     4       0
        RED ZONE LANE   GREEN ZONE LANE   PURPLE ZONE LANE   CHAIN   IGNORE
READING:  0     4.66         6    38.57          2    73.17     3       7
DOPDEC:      0.01             0.02             0.04
DECCA RED/GREEN: 54    30.51    -3     1.33
        RED/PURPLE: 54    30.83    -3     5.06
      GREEN/PURPLE: 54    30.10    -3     2.29      BEST CUT                                DOP-DECCA:  -0.20         0.11

BLOCK NO.  94


TIME:  13      2       0  =  46920   SECONDS
I.N.   54    25.13   -3     2.63      HEADING: 330     4.30  TRACK: 329    18.70
DOP    54    31.99   -3     4.33      COMPASS: 338    42.80  DRIFT:   0  -40.33               IN-DOPPLER  -6.86     1.70
DOP.K  54    31.99   -3     4.33      TRK.KAL: 329    18.70  TRK.ERROR:  0     0.00           DOP-DOPKAL   0.00     0.00
TACAN  54    30.98   -3     1.73      BEARING: 191    22.70  RANGE:      66                   IN-TACAN  -5.85    -0.90
KAL    54    31.53   -3     4.33      VEL ERR N.S:    -1     E.W:   -6  FT/SEC   MARK:  0       IN-KALMAN   -6.40     1.70
  DOP.DIST:    229   N.MILES   GRD.SP:    294   KN   HEIGHT: 23000    23000   FT. DOP UNLOCK:   0.00
      MULCOS:     4       0
        RED ZONE LANE   GREEN ZONE LANE   PURPLE ZONE LANE   CHAIN   IGNORE
READING:  0     3.77         6    31.56          2    74.99     3       7
DOPDEC:      0.01             0.08             0.04
DECCA RED/GREEN: 54    32.70    -3     4.48
        RED/PURPLE: 54    32.90    -3     7.[illegible]
      GREEN/PURPLE: 54    32.48    -3     4.99      BEST CUT                                DOP-DECCA:  -0.49         0.66

BLOCK NO.  95


TIME:  13      2      30  =  46950   SECONDS
I.N.   54    27.25   -3     4.78      HEADING: 328    47.90  TRACK: 328    49.20
DOP    54    34.08   -3     6.52      COMPASS: 339    13.10  DRIFT:   0   -1.31               IN-DOPPLER  -6.83     1.74
DOP.K  54    34.08   -3     6.52      TRK.KAL: 328    49.20  TRK.ERROR:  0     0.00           DOP-DOPKAL   0.00     0.00
TACAN  54    33.18   -3     1.95      BEARING: 191    11.00  RANGE:      68                   IN-TACAN  -5.93    -2.83
KAL    54    28.72   -3     7.79      VEL ERR N.S:     5     E.W:    6  FT/SEC   MARK:  0       IN-KALMAN   -1.47     3.01
  DOP.DIST:    231   N.MILES   GRD.SP:    296   KN   HEIGHT: 23000    23000   FT. DOP UNLOCK:   0.00
      MULCOS:     4       0
        RED ZONE LANE   GREEN ZONE LANE   PURPLE ZONE LANE   CHAIN   IGNORE
READING:  0     3.02         5    43.67          2    77.09     3       7
DOPDEC:      0.01             0.08             0.03
DECCA RED/GREEN: 54    34.64    -3     6.97
        RED/PURPLE: 54    34.74    -3     8.71
      GREEN/PURPLE: 54    34.60    -3     7.10      BEST CUT                                DOP-DECCA:  -0.52         0.58

BLOCK NO.  96
```

Fig.14 Processed data print-out

CURRENT STATUS OF MODELS FOR THE HUMAN OPERATOR AS A CONTROLLER AND DECISION MAKER IN MANNED AEROSPACE SYSTEMS

A. V. Phatak and D. L. Kleinman
Systems Control, Inc.
Palo Alto, California and Cambridge, Massachusetts
U. S. A.

SUMMARY

The increased complexity of aerospace vehicles, with emphasis on automatic controls, has considerably altered the role of the human operator in such systems. Routine, burdensome tasks previously performed by man are now automated. As a result, the human task requires a greater emphasis on the monitoring and decision making aspects than on the control problem. Any attempt to automate these decision processes or to augment them for semi-automatic procedures must necessarily be based on a quantitative understanding of human capabilities in complex decision an control tasks. Considerable data are available on the information gathering and processing aspects of human behavior. From these results mathematical models of human decision processes and adaptive behavior have been proposed for specific control situations. In this paper we survey accepted techniques and models for analyzing and predicting human performance in complex multi-control and multi-display situations commonly found in aerospace systems. The models that we consider have been developed or proposed for the related human functions of information processing, decision making and control. This paper discusses the relative advantages, disadvantages and limitations of each of the modeling schemes. Prospects for mechanizing all or part of the decision functions performed by human operators are considered -- specific examples being in the automation of human failure detection and adaptation to sudden changes in the system operating conditions.

1. INTRODUCTION

The human's role in manned vehicle systems is changing. The coming decades will see increased reliance on automatic and semi-automatic aircraft landing systems, digital controllers and adaptive augmentation. In these vehicles the human will operate in parallel with an automatic system; consequently, his role as a system monitor, fault detector and decision maker will become paramount. He must be capable of rapidly detecting system anomalies and failures, diagnosing the malfunction correctly, and taking the proper corrective control action.

It is clear that any analytic techniques proposed to study these complex man-machine problems must be based on an appropriate blending of human response theory, control theory and decision and information theory. Considerable expertise already exists in these fields. Mathematical models of the human as a controller have been developed and refined over the past two decades and decision and control theory are both well-developed, albeit diverse sciences. Lacking, however, is the methodology by which human operator models can interface with modern statistical decision theory to develop a comprehensive quantitative representation of the human as a controller <u>and</u> decision maker in complex man-machine tasks.

Since it is possible to regard virtually all aspects of human behavior as involving decision making in some way(Edwards, W., 1969), it is necessary to define the class of problem being considered. We limit this study to the adaptive abilities of the human for detecting and diagnosing <u>unexpected</u> failures and taking corrective action in a closed-loop dynamic system. Typical examples are observed in pilot response to stability augmentation system failures, engine-out or instrument malfunctions. At present, however, there is no validated theory for predicting human response in these critical situations. Several preliminary contributions to manual adaptive control have been made by Sheridan (1960, 1962), Yound <u>et al</u>.(1964), Knoop, Gould and Fu (1964, 1966), Elkind and Miller (1967) and Phatak (1968, 1969). These and other works are summarized in the excellent survey paper by Young (1969).

This paper examines the potential for developing a theory for manual adaptive control that would be applicable to existing and future manned aerospace systems. We examine the major human operator models proposed to date to determine their flexibility, feasibility and potential to interface with the large body of knowledge regarding decision making models. The status of the interface is discussed and candidate models for human fault detection and diagnosis are presented.

2. HUMAN OPERATOR MODELS

Modern research on the manual control of dynamic systems had its origin in the servomechanism and filtering theories of the 1940's. In subsequent years, models for the human as a control element were advanced by numerous researchers using diverse techniques. The human was characterized as a network of nonlinear threshold elements, a sampled-data system, a finite state adaptive machine, a servo element and optimal feedback controller to name but a few. However, in the quantitative description of human response, two main approaches have emerged - the quasi-linear model of McRuer, <u>et al</u>. (1957, 1965), and the optimal control model of Kleinman, Baron and Levison (1971). Both are summarized below.

2.1 Quasi-Linear Pilot Models

From the work of Elkind (1956) and McRuer, et al. (1957) there resulted a quasi-linear model of the human as a controller of single input, single output, linear time-invariant systems. The linear portion of this model consists of a "human" transfer function, the most common form of which is

$$Y_p(s) = K_p\frac{(\tau_L s + 1)}{(\tau_1 s + 1)} \cdot \left\{\frac{e^{-\tau_e s}}{\tau_N s + 1}\right\} \tag{1}$$

The bracketed term models inherent human limitations of reaction time and process delays, and lags attributed to the neuomotor system. The remaining terms represent the human's equalization characteristics and are adjustable in accordance with a set of verbal "rules". Basically, K_p, τ_L, and τ_1 are chosen such that the ensuing closed-loop characteristics approximate those of a good feedback control system.

The portion of the human's response that cannot be related to his observed input by a transfer function $Y_p(s)$ is called "remnant". A multiparameter model for human remnant, based on random sampling considerations, has only recently been added in the quasi-linear context (Clement, W.E., 1969), and is partially validated. While the neglect of remnant was not a serious drawback when studying simple systems, the need to model remnant and human uncertainties is greatly accentuated in complex tasks where $Y_p(s)$ often accounts for only a small portion of the controllers output power.

Attempts to extend the quasilinear models to multi-varible, multi-display situations began about 1960. The approach has been based on classical multi-loop control theory and relies heavily on judgements concerning the closed-loop system structure. As such it is necessary to solve iteratively a sequence of inner and outer loop closure control problems, assuming a fixed form model for the pilot's equalization in each loop. The process is often cumbersome. Nevertheless, considerable application of the quasi-linear models to study multi-loop aircraft problems has been made over the past several years. The most notable applications have been in anlyzing aircraft stability and in correlating pilot response parameters with aircraft handling qualities. Because of the lack of remnant models, predictions of closed-loop performance (i.e., mean squared values of system quantities) were generally not available.

By relying on classical frequency domain techniques, the quasi-linear models are not readily applicable to study time-varying, or transient control phenomena. These models do not contain the means to characterize the information processing behavior of the human, but focus instead on the human's steady-state control behavior. In studying control system failures, the quasi-linear models can be used to predict human response in the steady-state modes both before and after a failure. However, failure detection and identification plus the means by which control re-adaption occurs is not readily handled.

2.2 The Optimal Control Model

An alternative to the quasilinear model has recently been developed and validated by Kleinman, Baron and Levison (1971). Their approach is rooted in modern optimal control and estimation theory and is based on the assumption that the well-trained human controller behaves in an optimal manner subject to his inherent limitations and constraints and the task requirements. The approach is capable of treating single-axis, multivariable and even time-varying systems within a single conceptual framework using state-space and time-domain techniques which are well-suited to the analysis of complex man-machine systems.

The structure of the overall model is shown in Fig. 1. The vehicle's equations of motion are of the general form

$$\dot{x}(t) = Ax(t) + Bu(t) + w(t) \tag{2}$$

where x is the vehicle state, u is the human's control input and w represents the external disturbances. The human observes a set of displayed outputs

$$y(t) = Cx(t) \tag{3}$$

The human limitations that are represented in the model include time delay τ, and a first-order representation for "neuromotor" dynamics (included by limiting the human's control rates) as in the quasi-linear model (1). Human remnant, or randomness, is modeled directly by associating a white "observation" noise with each displayed output, and a white "motor" noise with each control input.†

The human's equalization is modeled by the cascade combination of a Kalman estimator that deduces the vehicle states from displayed information, a predictor that compensates optimally for the inherent delay τ, and a set of gains ℓ^*, that operates on the predicted state estimate to produce the desired control response. The gains are chosen to minimize an appropriate cost functional that relates the human's

† "Average values for human response parameters in the optimal control model generally are available from human response theory and existing data.

control objectives to the task being performed. The selection of a cost functional is a nontrivial matter and requires an understanding of the overall man-machine system requirements. Thus the lead-lag equalization of equation (1) has been replaced by a more complex feedback control characterization in the optimal control model.

The heart of the human equalization model is the Kalman filter, or information processor, F, that constructs a best estimate $p(t) = \hat{x}(t - \tau)$ of the delayed state $x(t - \tau)$ from the delayed, noisy perception

$$y_p(t) = Cx(t - \tau) + v_y(t - \tau) \tag{4}$$

according to the equation

$$\dot{p}(t) = Ap(t) + B\hat{u}(t - \tau) + G[y_p(t) - Cp(t)] \; : \text{Filter } F \tag{5}$$

The filter gain matrix G depends on the magnitude of the external disturbances and the human's remnant. The quantity $\hat{u}(t - \tau)$ is the "humans" estimate of the control input. Because of the motor noise, the actual control input $u(t)$ is not precisely known. The filter output $p(t)$ is processed by the optimal predictor to generate the state estimate $\hat{x}(t)$.

The optimal control model has been applied across a variety of manual control tasks, although fewer in number than the quasi-linear models. However, in the few years since its development the model has shown considerable success in predicting man-vehicle performance and has been extended to treat visual scanning, attention allocation and workload. The modeling of the human cognitive and decision processes within the control context appear to be logical directions in which to extend the existing model. The potential for such extension is discussed in the next section.

3. ADAPTIVE MODELING ASPECTS OF THE OPTIMAL CONTROL MODEL

The primary potential of the optimal control model for studying human decision making lies in the characteristics associated with the Kalman filter-predictor submodels. The combination of these elements provides the framework for modeling the information processing behavior of the human, and consequently, his decision abilities.

3.1 Internal Model

It is interesting to note that the description of the Kalman Filter F given by (5) contains an explicit model of the plant dynamics given by (2) and (3) via the parameter matrices A, B and C. Put another way, the filter includes an internal model of the environment. This concept is appealing. The use of internal models of the dynamical process being controlled has been advocated by several manual control researchers. Broadly, an internal model characterizes, either implicitly or explicitly, the human's knowledge of the controlled vehicle dynamics, a process arrived at and refined through training and experience.

Preliminary experiments in adaptive manual control support some type of model of the environment that is internal to the human (Young, L.R., 1969). Virtually all efforts to model fault detection in manual control have postulated an internal model. Indeed, the concept of expected versus unexpected response associated with detection implies some type of internal model of the controlled element dynamics. The optimal control model satisfies this prerequisite. What is needed in the modeling framework is the ability to detect any discrepancy between the actual system and its internal representation, as would be the case following a change or failure in system dynamics.

3.2 State Estimation

The output of the Kalman filter/predictor $\hat{x}(t)$, is the models' best (i.e. minimum mean-square error) estimate of the vehicle state $x(t)$, generated on the basis of the available information $y(t)$. The human's instantaneous control action and his estimate of vehicle attitude and position are determined directly from $\hat{x}(t)$. This "internal" estimate of vehicle status is updated continuously and provides a mechanism for studying decision/detection phenomena that are wholly dependent on the vehicle state. Examples of such problems are in deciding at time t, whether or not certain variables lie within desired limits, as being within a nominal approach-to-landing "window". Levison (1971) used the generated state estimate $\hat{x}(t)$ to develop a continuous time, monitoring and decision model. His basic assumption was that a human's decision involving $x(t)$ is made on the basis of the estimate $\hat{x}(t)$ and its error covariance $K_{(x-\hat{x})}(t)$. The model has been partially validated by an experiment in which subjects continuously decided whether a signal was within certain bounds on the basis of observing signal plus noise.

The successful modeling of a human's continuous-time state detection process is an important application of the optimal control model. However, this model needs to be appended before it can be applied to the problem of fault detection and identification. Nevertheless, it can be expected that the internal estimate $\hat{x}(t)$ will be of paramount importance for control generation in any adaptive response model of the human that springs from the optimal control model.

3.3 The Innovations Process

Consider the method by which the Kalman filter output $p(t) = \hat{x}(t-\tau)$ is updated as a function of time.

The driving term

$$r(t) = [y_p(t) - C\hat{x}(t-\tau)]$$

represents the difference between the human's perceived information $y_p(t)$ and the filter's internal estimate $C\hat{x}(t-\tau)$. Thus, $r(t)$ is the difference between actual and expected observations, and is called the residual, or innovations process. Basically, $r(t)$ is the new information that is brought to the filter by $y_p(t)$.

In the nominal case, when the internal model in the Kalman filter adequately representa the controlled element dynamics, the process $r(t)$ is a zero mean white Gaussian noise (Mehra, R.K., 1971). In other words, $y_p(t)$ and $C\hat{x}(t-\tau)$ are statistically equivalent and their difference has no information content. However, if there are failures within the actual system, model and actual parameter matrices will differ and the human's estimate of system behavior would deviate in a mean sense from observed dynamic behavior. These differences will produce a non-zero mean, correlated, innovations process. This fact can be used to represent the human's detection of failure by designing a model that is sensitive to the non-whiteneas of $r(t)$. We shall have more to say on the role of the innovations process and its whiteness properties in the sequel.

It is appropriate to note that the model for failure detection as proposed by Miller and Elkind (1967) was based upon monitoring the deviation between an expected error quanity and the actual quantity observed on the display. "Important" deviations between what happened and what was expected were used to develop a detection criterion. The technique seemed to work well for simple systems. Its use in complex systems was never attempted. However, the similarity between this technique and the generalized concept of residual monitoring is striking.

In summary, the three attributes of the optimal control model consisting of (1) the internal model (2) the state estimate, and (3) the properties of the innovations process indicate a potential for using this model as a basis for developing a theory of adaptive manual control. The manner in which this might be accomplished, and the interfacing of the model with decision theoretic concepts are discussed in the following aections.

4. HUMAN ADAPATATION TO UNEXPECTED FAILURES

The human operators task in any complex closed-loop man-machine system is to maintain system performance at or above accepted levels in spite of external distrubances and in the event of unexpected failures. Since steady-state models for the human operator reaponse in the presence of random disturbances are well understood, we will concentrate on human adaptivity to unexpected system failures. The failures considered here ultimately result in an unexpected change in the plant dynamica deacribed by equations (2) and (3). Put another way, an unexpected failure results in a change in the pre-failure plant dynamics characterized by the matrix triplet

$$S_o = (A_o, B_o, C_o)$$

to a new set of dynamics characterized by a matrix triplet S_k.

There are two general types of transitions in plant dynamics. One is relatively slow, with the parameters changing gradually over a period of a few seconds or longer. The other involvea a sudden or steplike change. The slow change usually results in a gradual deterioration of system performance and can be adapted to by the human with some kind of performance adaptive algorithm. We are primarily interested in atep failures or changes in S_o because of their sudden effects on overall system performance and closed-loop stability. This sudden degradation of the syatem compels the human operator to make a quick diagnosis about the state of the system and to restructure the control strategy via some decision-directed algorithm.

A number of investigators in the past have studied human adaptive behavior following a change in the controlled element dynamics. The basic structure of most all of the proposed adaptive modela is very similar and involvea the partitioning of the adaptive process into the following four phaaes

1. Optimal control of pre-failure plant dynamics
2. Detection of a change or failure in plant dynamica
3. Identification or diagnosis of the fault
4. Corrective action or modification of control atrategy appropriate to the post-failure plant structure

The combination of appropriate models for the four phases of adaptation, therefore, constitutes the adaptive human controller model. Modela for the firat and last phaaes are available in the describing function or optimal control form as described earlier in Section 2. Hence the contribution to adaptive modeling lies in the mathematical description of human decision processea represented by the second and third phases. Unfortunately, the direct modeling approach based on processing input-output experimental data has severe limitations because of the lack of adequate techniques for the identification of rapidly changing dynamic systema. Aa a result, the development of models for human adaptive control has remained inductive in approach and limited to general schemes. On the other hand, a considerable amount of research in statistical decision theory has been devoted to fault detection and hypothesis testing problems and can be made to bear upon human decision making and adaptive procesaes. The properties of the optimal control model described in Section 3 render such an interface feasible, as discussed in the following aection.

5. FORMULATION OF THE ADAPTIVE CONTROL PROBLEM

The human operator is involved in controlling or monitoring a vehicle whose dynamics may be described by the linear differential equation

$$\dot{x}(t) = A(t)\, x(t) + B(t)u(t) + w(t) \tag{6}$$

where x(t) is the vehicle n-state vector consisting of quantities like pitch rate, angle of attack, longitudinal velocity, etc., u(t) is the r-vector of control inputs (e.g. elevetor, throttle, etc.), and w(t) is an u-vector of stationary gaussian white process noise with zero mean and covariance w.

As discussed in section 2.2 the human operator perceives the s-vector (the time-delay is omitted for ease of exposition)

$$y_p(t) = C(t)\, x(t) + v_y(t) \tag{7}$$

where $v_y(t)$ is the stationary gaussian white observation noise with zero mean and covariance V.

The plant dynamics defined by the matrix triplet S(t) = [A(t), B(t), C(t)] <u>may</u> unexpectedly change from a known prefailure configuration $S_o \stackrel{\Delta}{=} [A_o, B_o, C_o]$ into one of N possible post-failure configurations $S_i \stackrel{\Delta}{=} [A_i, B_i, C_i]$ i = 1,2,---, N at some unknown instant of time $t \stackrel{\Delta}{=} t_f$.

In other words,

$$S(t) = \left\{ S_o[1 - \gamma(t,t_f)] + S_i\, \gamma(t,t_f) \right\} \quad , i = 1,2,---,N \tag{8}$$

where

$$\gamma(t,t_f) \stackrel{\Delta}{=} \begin{cases} 0 : t < t_f \\ 1 : t \geq t_f \end{cases}$$

[Note that the human operator time delay is not included in the observation vector described by (7). However, it can be easily incorporated without changing the structure of equations (6-8) by shifting the time scale by τ seconds].

The human operators' task is to continuously determine which one of the N systems S_i, i = 0,1,---N is the true system S based on the continuous observation record $Y_t = \{y_p(\tau);\ \tau \epsilon (t_o,t)\}$.

There are two approaches by which the human operator may carry out his on-line fault monitoring task.

1. He may partition the task into one of failure detection followed by fault diagnosis, or

2. He may directly perform simultaneous detection and identification of plant failures.

Partitioning of the decision task into the sequence of tasks of failure detection followed by identification has the advantage of not requiring internal models for the alternative failure hypothesis $S = S_i$, i = 1,2,---N, during the process of detection. The sub-process of fault identification, however, does need these internal models, and involves identification procedures (similar to the second approach) for accepting one of N post-failure hypotheses.

We will first consider failure detection based on the properties of the innovations process, and then follow it with a general discussion of multihypotheses procedures for fault identification.

6. FAILURE DETECTION AND IDENTIFICATION

6.1 <u>Innovations Testing</u>

As long as the system is operating normally, the human operator's internal model representation of the plant is consistent with the actual control situation described by the matrix triplet S_o. Consequently the expected observations $\hat{y}(t/S_o)$ should be close to the actual observations y(t) in a minimum mean square sense, and the innovetions or residual process, $r(t/S_o) \stackrel{\Delta}{=} y(t) - C_o\hat{x}(t/S_o)$, must be a zero mean, white guassian process with covariance V, equal to that of the observation noise. If a failure occurs in the plant dynamics to configuration $\bar{S}_o$ ($\bar{S}_o \stackrel{\Delta}{=}$ not S_o), then there results an inconsistency between the internal model corresponding to S_o and the actual representation $\bar{S}_o$. As a result the model response would no longer match the actual system response thereby causing a change in the properties of the innovations process. Hence, detection of this change in the properties of the innovations process would provide the necessary test for fault detection. In effect, we must test the null hypothesis that the innovation process $r(t/S_o)$ is zero mean, white and has a given covariance, at a certain specified level of significance. Notice that no knowledge of internal models for the post-failure situation is required for fault detection.

At any given time t the innovations process $r(t/S_o)$ for the past T seconds can be tested for whiteness, zero mean and given covariance V, in that order. If the null hypothesis is satisfied, the test may be conducted again, after some interval of time, Δt. In general, fault detection using this procedure, requires that record segments of the innovations process $\{r(t/S_o); \tau \varepsilon (t-T, t)\}$ be tested continuously every Δt seconds using a sliding observation window of T seconds duration. The choice of Δt depends on the particular case under consideration.

There exists several tests in the literature that may be applied to the innovations in order to test for whiteness, zero mean and covariance characteristics. An excellent paper by Mehra and Peschon (1971) on the innovations approach to fault detection gives the necessary details on applying some of these tests to practical problems.

Unfortunately, the method of residual testing gives little information about the nature of the fault after it is detected. Indeed, there are ways to utilize the correlation in the innovations process following failure to identify the new parameters in the plant dynamics. However, these schemes are not rapid enough for use in the adaptive control context. The methods of alternate hypothesis testing discussed below are more suitable to the problem.

6.2 Alternate Hypothesis Testing

The principal assumption made here is that the human operator, in addition to having an internal model for the normal operating plant dynamics, can have internal models corresponding to each of the N possible post-failure plant dynamics.

Let F_i, $i = 0,1,2,$---N be a bank of (N+1) internal Kalman filter models associated with the corresponding (N+1) alternative hypothesis, $S = S_i$ $i = 0,1$---,N. As described in equation (5), the inputs to each filter are the observation vector $\bar{y}_p(t)$ and the control signal u(t). Two vector outputs are obtained from each filter. They are:

$$\begin{aligned}
&\text{i)} \quad \hat{x}_i(t) \equiv \hat{x}(t / S_i) \equiv \text{state estimate vectors of } F_i \\
&\quad \text{and} \\
&\text{ii)} \quad r_i(t) \equiv r(t / S_i) = \text{the innovations vector of } F_i
\end{aligned} \tag{9}$$

$$\text{where} \quad r_i(t) = y_p(t) - C_i\hat{x}_i(t)$$

We postulate here, that the human's decision making must be based on recursively computing the probabilities of each of the individual (N+1) hypotheses on the basis of incoming data y(t), and then comparing them in some way to decide the true state of the plant dynamics. Let $\lambda_i(t)$ be this aposteriori probability that $S = S_i$; that is, let

$$\lambda_i(t) \triangleq \Pr\{ S = S_i \mid Y_t \} \tag{10}$$

Then at any time, t, $\lambda_i(t)$ reflects the confidence with which the internal model S_i can be accepted as the correct representation of the true system. The λ_i's may also be computed in terms of the individual probability ratios (also called the odds ratio).

$$\lambda_{jo}(t) = \frac{\lambda_j(t)}{\lambda_o(t)} \tag{11}$$

by using the expression (obtained by Baye's rules)

$$\lambda_i(t) = \frac{\lambda_{io}(t)}{\sum_{j=1}^{N} \lambda_{jo}(t)} \tag{12}$$

It is also easy to show (Kailath, T., (1970); Lainiotis, D.G., (1971) that the individual odds ratios λ_{jo}'s can be obtained on-line by integrating the scalar differential equations

$$\dot{\lambda}_{jo}(t) = -\lambda_{jo}(t)\,[q_j(t) - q_o(t)], \; j = 1,2, ---, N \tag{13}$$

$$\lambda_{jo}(o) = \lambda_{jo}^{*} = \text{a priori odds ratio}$$

where

$$q_i(t) = \frac{1}{2} r_i^T (t) V^{-1} r_i(t), \; i = 0,1, ---, N \qquad (14)$$

Note that

$$0 \leq \lambda_i(t) \leq 1 \qquad (15)$$

and

$$\sum_{i=0}^{N} \lambda_i(t) = 1$$

Equations (12), (13) and (14), together, give the human all the information that is needed to proceed with the task of system identification. Two plausible hypotheses tests are discussed next.

a. Recursive Multihypotheses Test

Let us assume that the test starts at $t = t_o$. Then at any particular time, the λ_i's give the current probability estimates of hypotheses S_i being true.

The system is assumed to be operating normally if

$$\lambda_o(t) > \lambda_i(t), \; i = 1,2, ---, N \qquad (16)$$

A failure is said to have occured, if at any time

$$\lambda_o(t) < \lambda_j(t), \text{ for some } j = 1,2, ---, N \qquad (17)$$

The post-failure plant is identified to be $S = S_k$ if

$$\lambda_k(t) > \lambda_o(t), \text{ and}$$
$$\lambda_k(t) = \max_j \; \lambda_j(t) \quad , j = 1,2, ---, N \qquad (18)$$

Equations (16), (17) and (18) can be tested on-line in a recursive fashion to give simultaneous results on detection and identification of failures in plant dynamics.

b. Pairwise Sequential Probability Ratio Test

The approach taken here is to apply N parallel sequential probability ratio tests (SPRT's) to the pairs of hypotheses$[(S_o,S_j), j=1,--,N]$. It is hypothesized that the human operator can form the N probability ratios, $\lambda_{jo}(t)$, $j=1,2,---,N$, and perform N sequential tests as follows:

If $\lambda_{jo}(t) \geq A_j$: accept hypothesis $S = S_j$ (19)

$\lambda_{jo}(t) \leq B_j$: accept hypothesis $S = S_o$

$B_j < \lambda_{jo}(t) < A_j$: take another sample

where $A_j = \dfrac{1 - \beta_j}{\alpha_j}$, $B_j = \dfrac{\beta_j}{1 - \alpha_j}$; and $\lambda_{jo}(t)$ are obtained from equations (13 and (14) with $\lambda_{jo} = 1.0$.

α_j and β_j are pre-specified false alarm and miss probabilities for the individual binary hypothesis tests. If any $\lambda_{jo}(t) \leq B_j$, then reinitialize equations (13) and (14) and start testing again according to (19). The first $\lambda_{jo}(t)$, say $\lambda_{mo}(t)$, to cross the upper boundary A_j indicates that the plant has failed to the new post-failure configuration of $S = S_m$.

Once the failure has been detected and identified, the next step in the human adaptive process is to restructure the optimal control strategy so as to make it compatible with the post-failure vehicle dynamics. However, if large errors have resulted prior to the time that identification is completed, then an intermediate control strategy appropriate to reduction of transient errors may have to be adopted before switching to the control technique suitable for the identified post-failure plant. Any transient tracking strategy selected by the human operator must be compatible with the optimal control of the post-failure plant in steady state. Weir and Phatak (1966), for example, proposed reduction of large accumulated errors by using a time-optimal non-linear control strategy. The optimal steady-state control stage of the adaptive process is already well understood since it involves an application of the optimal control model of section 2.2.

7. CONCLUSIONS

In this survey paper, we have tried to expound upon some of the critically important aspects of human operator modeling - namely in the role of the human as a controller and a decision-maker. The models that we propose for joint human control and decision-making behavior are inductive extrapolations of results from simple laboratory experiments on human hypothesis testing and some recent work in statistical decision theory. The adaptive control models proposed here could also be used as aids to human operator decision making. For example, the probabilities of the various alternative failure hypotheses may be computed automatically and displayed to the human operator for appropriate action. Such an approach may become mandatory in future aerospace systems where the human operator's role would be primarily that of a status monitor. In this role, the human would no longer be an active controller and hence his ability to store adequate internal models may be considerably degraded. However, an on-board computer could take over this job of storing internal models and display to the human monitor, a continuous record of system status information. Much more experimental work needs to be carried out before we can intelligently allocate control and decision making roles between the human operator and the on-board computer. It is our hope that this paper will serve the purpose of stimulating further work in this fascinating and practical problem.

REFERENCES

CLEMENT, W.F., 1969, "Random Sampling Remnant Theory Applied to Manual Control," Systems Technology, Inc., TM-183-A.

EDWARDS, W., 1969, "A Bibliography of Research on Behavioral Decision Processes to 1969", U. of Michigan, Human Performance Center, Memo Rep. No. 7.

ELKIND, J. I., 1956, "Characteristics of Simple Manual Control Systems, TR111, MIT Lincoln Laboratory, Lexington, Mass.

GOULD, E.E., and FU, K.S., 1966, "Adaptive Model of the Human Operator in a Time-Varying Control Task," Proc. 2nd Ann. NASA-University Conf. on Manual Control (M.I.T., Cambridge, Mass.) pp. 85-97

KAILATH, THOMAS, May 1970, "Likelihood Ratios for Gaussian Processes", IEEE Trans. on Info. Theory, Vol. IT-6-No. 3.

KLEINMAN, D.L., and BARON, S., 1971, "Manned Vehicle Systems Analysis by Means of Modern Control Theory", NASA-CR-1792

KLEINMAN, D. L., BARON, S., and LEVISON, W. L., 1971, "A Control Theoretic Approach to Manned Vehicle Systems Analysis", IEEE Trans. Autom. Control, Vol. AC-16, No. 6

KLEINMAN, D. L. and BARON, S., 1971, "Analytic Evaluation of Display Requirements for Approach to Landing", NASA-CR-1952, Oct.

KNOOP, D. E. , and FU, K. F., 1964, "An Adaptive Model of the Human Operator in a Control System", 5th National Symposium on HFE, San Diego, Calif.

LAINIOTIS, D. G., April 1971, "Optimal Adaptive Estimation: Structure and Parameter Adaptation", IEEE Trans. on Auto. Control. Vol. AC-16, No. 2.

LEVISON, W. H., 1971, "A Control Theory Model for Human Decision Making", 7th Annual NASA Conference on Manual Control, USC.

McRUER, D.T., KRENDEL, E. S., 1957, "Dynamic Response of Human Operators", Wright Air Dev. Center, WADC TR56-524, Wright-Patterson AFB.

McRUER, D. T., GRAHAM, D., KRENDEL, E. S., and REISENER, W., 1965, "Human Pilot Dynamics in Compensatory Systems", AFFDL-TR-65-15.

MEHRA, R. K., and PESCHON, J., 1971, "An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems", International IEEE Conference on Systems, Networks, and Computers, Mexico.

MILLER, D. C. and ELKIND, J. I., 1967, "The Adaptive Response of the Human Controller to Sudden Changes in Controlled Process Dynamics", IEEE Trans. HFE, Vol. 8, No. 3, 1967.

PHATAK, A. V., 1968, "On the Adaptive Behavior of the Human Operator in Response to a Sudden Change in the Control Situation", Ph.D. Thesis , University of Southern California, USCEE Report 277.

PHATAK, A. V., and BEKEY, G. A., 1969, "Decision Processes in the Adaptive Behavior of Human Controllers", IEEE Trans. on S.S.C. Vol. 5, No. 4, Oct.

SHERIDAN, T. B., 1960, "Time-Variables Dynamics of Human Operator Systems," Air Force Cambridge Research Center, Bedford, Mass., Tech. Rept. AFCRC-TN-60-169 (ASTIA Doc. AD-237045), March

SHERIDAN, T. B., 1962, "Studies of Adaptive Characteristics of the Human Controller", USAF Electronic Systems Division, Rept. ESD-TDR-62-351, December 1962.

SMALLWOOD, R. D., 1967, "Internal Models and the Human Instrument Monitor", IEEE Trans. HFE, Vol. 8, No. 3.

WEIR, D. H., and PHATAK, A. V., 1966, Model of Human Operator Response to Step Transitions in Controlled Element Dynamics, 2nd Annual NASA-University Conf. on Manual Control, NASA SP-128.

YOUNG, L. R., GREEN, D. M., ELKIND, J. I., and KELLEY, J.A., 1964, "Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control", IEEE Trans. on HFE, Vol. 5, No. 1.

YOUNG, L. R., 1969, "On Adaptive Manual Control", Ergonomics, Vol. 12, No. 4.

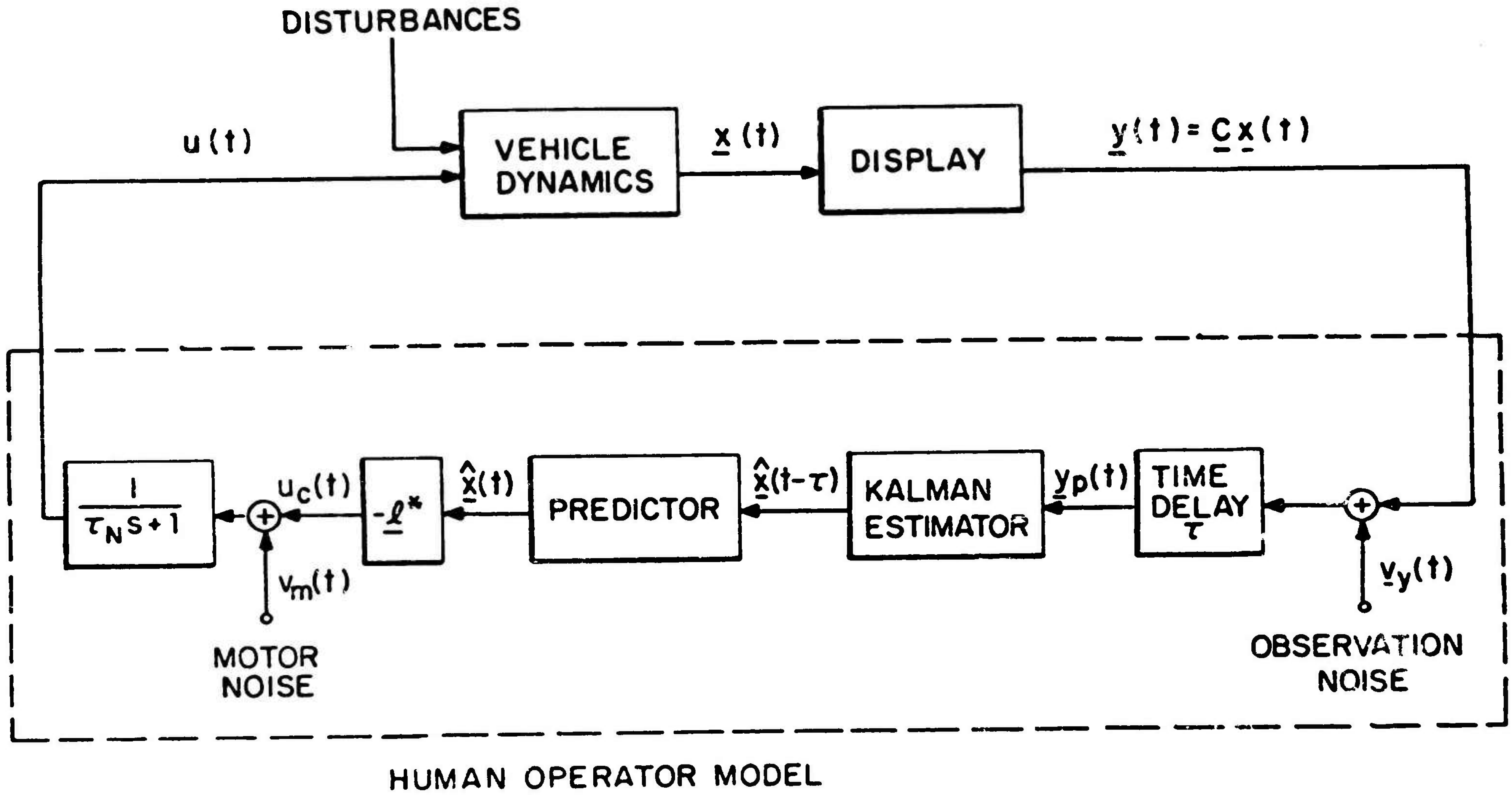


Fig.1 Control Theoretic Model of Optimal Human Behavior.

MANUAL LANDINGS IN FOG

R.R. Newbery
BLEU
RAE BEDFORD
U.K.

SUMMARY

The results of 18 fog-flying sorties using a Category II operation terminated by a manual landing have been analysed in an attempt to learn more about the pilot's capabilities in this environment. Measurements were made to correlate the pilot's decision making process with actual fog structures in real operation. A wide variety of fog structure and visual sequences are illustrated which demonstrate the lack of relationship between the visual segment at high decision heights, the height at which visual contact is first made and the Runway Visual Range measurement. This variation and particularly the potential hazards of shallow fog should be stressed in pilot training. If this sample is representative, go-around rates in Category II operation are likely to be greater than anticipated and will fluctuate violently depending on the contact height. The pilots felt that Category II operation was straightforward provided that good quality approach performance, strict crew drills and accurate RVR reporting to give warning of shallow or changing fog conditions along the runway, were maintained.

1. INTRODUCTION

A wide spectrum of aircraft, equipment and operational procedures have been developed to enable aircraft to land in poor visibility. The broad system requirements are classified into the ICAO Category I, II and III operational objectives. By defining a specific height above which the head-up pilot must make his decision, Categories I and II allow for the landing to be completed manually. The problem for the approving authorities is to assess a safe operational boundary for each application, bearing in mind that while the hardware components of the system may be well tested little is known of the pilot's capability and judgement in the new environment. The question "How safe is Category II operation?" is an easy one to pose but difficult to answer because of the large and important human factor element. Little measured information is available to determine how decisions to land are made and how a pilot's judgement and performance are affected in the variety of visual sequences encountered in Category II operations.

Simulator programmes at the Blind Landing Experimental Unit (BROWN, A.D., 1968 and 1971) studied these questions in detail and a parallel fog flying programme was started to gather more data on visibility sequences and to attempt a measurement of the factors affecting the pilot's decision to land and his subsequent performance. The analysis of the flight results presented in this paper is inevitably an analysis of 'what we could obtain' rather than a balanced scientific overall study but the results have the virtue of being real measurements obtained under real conditions with real pressures on the crew.

2. BLEU OPERATIONAL PROCEDURES

Operation of the BLEU experimental aircraft in fog was broadly along the lines of the "United Kingdom provisional operating requirements for all weather operations in Category II". (HMSO CAP 321, 1969). Landings in fog could be made at sites having the minimum requirements of Category II quality, ILS, approach and runway lighting, and RVR reporting based on visual or transmissometer measurements from near the touchdown zone with a radio link to the aircraft.

The basic crew procedures used during the programme have been in use at BLEU for many years and have a similarity to the crew drills used by British European Airways. A coupled approach is made under the control of P2, the co-pilot, head-down and normally in the right hand seat. P1, who is the Captain and subject pilot, monitors the approach transferring his concentration from head-down to head-up as the approach progresses. Radio heights at 100ft intervals are called from 500ft downwards either by P2 or a third member of the crew and 100ft above the decision height is called as '100 above'. For the purposes of this flight programme, the subject pilot, P1, was instructed to take over manual control and land the aircraft as soon as he considered that sufficient visual information is available to do so safely. If P1 has not assumed control by decision height, P2 carries out the standard go-around procedure by disconnecting autopilot and autothrottles and applying full power. Care is taken to avoid a conflict of decision at this critical time.

3. EXPERIMENTAL EQUIPMENT, PILOTS AND SITES

3.1 Aircraft

A Comet and a Varsity were used during the flying programme. Features relevant to this paper are:-

	Comet	Varsity
Downward View Cut off Angle Approach Speed used	14 degrees 120 knots	20 degrees 110 knots
Experimental Clearance later amended to	Zero Decision Height 200m RVR 150m RVR	100ft Decision Height 400m RVR 80ft Decision Height 200m RVR
Approach/Autoland system	Development model of Smiths Mk 29 SEP 5 Triplex Autoland system	Smiths Mk 10B Military Flight Instrument and Simplex auto-land system.

3.2 Pilots

Five subject pilots took part in the programme, all were drawn from the Royal Air Force but had a variety of experience as follows:-

<table>
<tr><th></th><th>Fog Flying</th><th>Simulator</th></tr>
<tr><td>A</td><td colspan="2">Very experienced. 6 years participation in BLEU fog flying and Simulator experiments both manual and automatic landing.</td></tr>
<tr><td>B</td><td>1½ years with BLEU; largely automatic landing experience.</td><td rowspan="3">Each pilot had at least 72 approaches in simulator night conditions prior to the flight programme.</td></tr>
<tr><td>C</td><td>Some experience of automatic landing during the previous Winter.</td></tr>
<tr><td>D</td><td>No previous automatic or manual landings below Cat I limits</td></tr>
<tr><td>E</td><td colspan="2">New subject pilot during the programme. No fog-flying or low visibility simulator experience.</td></tr>
</table>

The simulator experiments had highlighted the possibility of conflict between the decisions of P1 and P2 at the decision height. Thereafter the crews were very aware of the need to avoid this conflict in flight.

3.3 Sites

Flying commenced at Bedford where 14 out of the 18 sorties were flown. Operations were later extended to Liverpool and Schiphol with the kind co-operation of the local airfield authorities and the assistance of the Netherlands National Aerospace Laboratory (NLR). Bedford and Schiphol lighting patterns are shown in Figures 1 and 2 this being the main difference between the sites. Bedford had a 90 metre wide runway but the lighting pattern conformed to a gauge size of 60 metres. The other two sites had runways 45 metres wide. Liverpool conforms to the standard ICAO-Calvert pattern having a thicker centre-line lighting than Bedford and omitting the elongated bars at Glide Path Origin.

3.4 Recording Equipment

Lock follow radar facilities were available at both Bedford and Schiphol to measure the flight path trajectories and both aircraft carried comprehensive flight recorders. Although film cameras were used during the flying and provided useful training film they were not used for measuring purposes since the camera film does not record the visual sequences in the same way as the human eye. A special observer was carried in the aircraft to record the visual sequences more accurately by what became known as the 'click box'. This was simply a multi-position switch and an identifying button which enabled the observer to mark when first visual contact was made or when the lighting bars, glide path origin and the threshold came into view. Subsequent analysis of the flight records, particularly the lock follow radar and the altimeter carrier return signals, determined both the aircraft position and the object coming into view. Figures 3 to 7 are examples of the visual sequences measured in this way. The run numbers denote the sequence during the sortie. The time between runs was usually about 10 minutes so accounts of the fogs over periods up to several hours were recorded. The horizontal visual range is shown plotted against the aircraft wheel height above a datum runway plane, so that as the aircraft descends, a visual sequence curve is traced out. The steeply sloping faint line through the origin represents the cockpit cut-off angle; hence the horizontal distance between points on this line, and the visual sequence curve indicate the visual segment available to the pilot at each particular height. When the aircraft is close to the glide path it is possible to add a further grid of lines such that the intersection of these grid lines with the horizontal visual segment line, defines the features which the observer was able to see within his visual segment.

This method of measurement contains many of the important practical elements which are inherent in the Category II operation. The observer, like the pilot, has a realisation and an action time delay. Changes in lighting intensity and alignment affect what is seen and the view itself is not sharp and clear. Ground measurements of Slant Visual Range (SVR) where time is not pressing and the observation lights on the pole are of uniform intensity and alignment may give a correct measurement of the actual fog structure in that place but may not be representative of the dynamic situation viewed by a pilot landing on the runway. To consolidate the measurements a second observer was carried to record impressions of the landing manoeuvre and provide a description of the visual sequence and pilot reactions.

RVR reporting at Bedford was initially by direct observation of a calibrated line of lights alongside the runway so that an equivalent RVR in terms of centreline lights was relayed to the aircraft by direct radio link. Later in the programme transmissometer readings became available. RVR reporting at Schiphol was by means of two transmissometers. The readings were taken in a central observation room, converted to RVR and relayed by Air Traffic Control to the approaching aircraft. At the time of the visits to Liverpool, visual observations were relayed via ATC.

4. THE FOG FLYING

At the start of the programme it was decided that the two pilots with least fog flying experience, c and d of para 3.2, should take the bulk of the flying and that one pilot would remain as subject pilot for each complete sortie. This was an attempt to see if visual-clue requirements changed with experience and to study any short term learning process or effect of familiarity. It soon became obvious that familiarity with a particular fog structure during a sortie was rapidly obtained and the procedure was modified so that the eubject pilote were interchanged after they had each achieved a landing. By this procedure it was hoped to present as many fresh situations as possible.

In order to stimulate an operational rather than an experimental frame of mind and preserve a high margin of safety, the Varsity aircraft was only cleared for operation during March and September 1969, in RVR conditions of 400 metres or better and with a decision height of 100ft. When flights had been made to exercise approach and go-around crew drills to these limits, the clearance was lowered to 80ft decision height and 200 metres RVR for the remaining sorties. A number of compulsory go-arounds and automatic landings occurred during the programme but on these occasione the pilot gave his opinion on whether he thought that the visual segment was enough to be able to complete the landing manually.

The subject pilots were instructed to take over manual control as soon as they felt they could do so to complete the landing. This is <u>not</u> recommended practice for normal Category II operation, or course, but was an attempt to confirm the point at which P1 took control, including his reasone for doing so, and to accentuate any possible performance variations during the landing. For airline operation with an automatic approach it is clearly preferable to maintain the automatic approach down to decision height but under the supervision of P1 after he has announced that he has taken control.

In 1969 and 1970, eighteen sorties consisting of 150 approaches were achieved in RVR conditions below 800 metres. All the following well known features were encountered:-

(a) Large but reducing visual segments due to ehallow fog.

(b) Low contact height with relatively large RVR values.

(c) High contact height but small visual segments with relatively small RVR values.

(d) Fogs with rapidly changing RVR values.

The approaches were divided between sites and aircraft as follows:-

	Comet	Varsity
Bedford	3 sorties 23 Approaches	11 sorties 100 Approachee
Liverpool	-	3 eorties 22 Approaches
Schiphol	-	1 eortie 5 Approaches

5. THE DECISION TO LAND

5.1 Factors Affecting the Decision to Land

5.1.1 Decision Time and Visual Segment

The head-up pilot has to decide whether he can identify the flight path of the aircraft in relation to the approach lighting pattern and the runway, and whether he will have enough visual information to complete the landing. Figure 8 describes the situation; there is a period of time (decision time) between first contact and the decision height in which he must assess the information within his field of view. This field of view lies between the cockpit cut-off and the furthest extent of his vision (the slant visual range). A convenient way of describing this field of view is the 'vieual segment'. This is the horizontal distance along the ground within this field of view and will normally be different at different heights depending upon the aircraft's downward view and the fog characteristics, etc.

In figuree 9 and 10 each approach is plotted by a character on axes which describe the decision time available and the extent of the pilot'e view just before decision height. This enables us to investigate the consistency of the decisions and the rules on which they were made. Decision time ie represented on the vertical axis by the difference between contact height and decision height. The field of view on the horizontal axis is expressed as the visual segment at 20 feet above the decision height. Twenty feet above decision height is choeen in preference to decision height since it ie closer to the point at which the pilot is actually making his decision.

Results are plotted on figure 9 for those approaches in which the pilot could be coneidered to be making the approach for the first time after a transit flight. They include approachee being made for the firet time in a eortie or after a break in a sortie during which the subject pilot was either acting as P2 or being carried as a passenger. Each approach is denoted by a symbol which indicates whether the visual sequence down to decision height was considered 'adequate' or 'inadequate' for manual completion of the landing. These approaches were made in widely differing visual conditions and unfortunately only a few resulte were obtained when the limiting conditions of contact time and visual segment were preeent. Nevertheless theee reeults are in good agreement with the more extensive simulator results (Brown A D 1968 and 1971) in that runs with visual segments in excess of 125 metres were considered to have adequate visual reference.

The following comments on some of the more interesting runs help to give more depth to the picture. The visual sequence with a segment of 125 metres was considered by the most experienced pilot, (a), see para 3.2, to be sufficient to land the Varsity but not the Comet. This is a good indication of where the borderline is likely to exist between acceptable and unacceptable visual segments. For the approach in which first contact was made 30 feet above the decision height pilot (d) comments, "It all happened too quickly". This approach in the Comet was the first Category II approach ever made by this particular pilot. The go-around and the successful landing plotted close together with contact 50 feet above decision height were made in one sortie by pilot (c). The go-around was his first approach in the sortie and the successful landing followed a period during which he was acting as P2. This again indicates a borderline condition.

The results plotted in figure 10 are for those approaches which were made by the same pilot within fifteen minutes of a previous approach. Even in these approaches, the lowest contact above the decision height which was considered to be acceptable was 30 feet, or equivalent to about 3 seconds of time. For first time approaches a criterion of more than 40 feet (or 4 seconds) is thought more likely to apply. The low contact heights invariably occurred with small visual segments so no really independent measurement of this criterion was possible from these results. Eight approaches in shallow fog with high contact height, low RVR, and large but reducing visual segments, are plotted as question marks. These were compulsory go-arounds since the RVR was below aircraft limits. Most of them would probably have been go-arounds had there been no aircraft RVR limit, but it is not clear how much the pilot was influenced by the RVR value and how much by the visual sequence. The effect of shallow fogs is dealt with separately in paragraph 5.2.

5.1.2 Short Term Learning

The repeated approaches plotted in Fig 10 contain at least ten successful landings with less than the 'decision time' and 'visual segment' criteria for first time approaches suggested in the previous paragraph. This was almost certainly due to recent familiarity and experience of the conditions. During an automatic landing or go-around the pilot becomes familiar with the appearance of the visual clues, becomes sure of their identification and has the opportunity to confirm that the visual sequence opens. A smaller segment and a shorter decision time then appear to be sufficient to assess the situation on the next approach, if the time interval between the approaches is small. During some of the go-arounds remarks were made such as "It all happened too quickly" or "In daytime it doesn't look like the simulator" whereas an approach or two later in similar conditions the pilot said that he felt that he had "mind to spare". Three of the landings made with the lowest contact above the decision height were made by three different pilots (a, b and d of Para 3.2), one followed two automatic landings, another followed four go-arounds with slightly improving contact height and the third followed two go-arounds in a stable fog. Once a landing had been achieved there was progressively less difficulty in deciding to land in the succeeding runs even with worse visual segments. The pilots suggested that this came from an increasing familiarity with details or recognisable features even such as a misaligned light, enabling them to assess their position and flight path earlier.

5.1.3 Long Term Experience

While there was a great difference in the previous experience of the five pilots, this was not obvious when comparing their decisions to land. All five pilots were subject pilots for go-arounds plotted in Figure 9 with contact heights between 25 and 50 feet above decision height, but four of the pilots made repeated landings later in the sortie in the same conditions, the fifth (e) not having the opportunity to do so. Three of the pilots (b, c and d) experienced these limiting conditions in their first fog sortie of this manual landing programme and pilot (a) in his second fog sortie of the programme. Pilot (d) said that he felt pre-disposed to a go-around on his first approach in 400 metres RVR. It must be remembered that all five pilots were well informed and well aware of the danger of shallow or patchy fog and the possibility of P1 and P2 actions at decision height conflicting.

5.2 The Special Case of Very Shallow Fog

Until the night sortie described by Figure 7, the decision to land had been the answer to the straightforward question 'Can the pilot see enough before decision height?', and the results gave straightforward conclusions as to pilot requirements (para 5.1). However, the shallow fog conditions of this flight, underlined one of the well-known, but insidious, dangers of the decision to land. On the first run with an RVR of 210 metres the glow of the lighting could be seen when turning on to the final approach at seven miles range and good contact with the approach lighting was established before joining the glide path. Manual control was taken at 155 feet and the landing completed. No difficulty was found with azimuth control but pitch plane control was described as "not easy" with a marked drop in the visual segment to 160 metres in the flare. It seems that while 125 metres is considered acceptable to make a decision to land the pilot is anticipating an opening sequence. Difficulty was found in recording this type of fog structure and the visual sequence for run 1 had to be discarded. The visual sequences of the following runs record the height at which individual lights became visible rather than their indefinable glow. Runs 2 to 9 were compulsory go-arounds from an 80ft decision height because the RVR had fallen below the aircraft limit of 200 metres. The visual segment above decision height was still quite large and on four of these runs "there would have been a strong temptation to complete the landing had not a go-around been obligatory". The pilots were however well aware of the inherent dangers associated with low reported RVR's and on only one of these runs did the pilot feel sure that he would have been able to complete the landing. When the RVR rose above 200 metres, landings were again made (runs 10 to 13) with the minimum visual segments occurring during the flare. Pitch control was again described as not easy.

5.3 The Point of Manual Take-Over

The subject pilots were instructed to take over manual control when they had enough information to complete the landing. There appears to be little relationship between the point at which manual control was taken and either the contact height or the visual segment. However, there does appear to be a strong relationship between <u>what</u> the pilot saw and the point of takeover. If we observe what object (first, second, third crossbar, threshold, glide path origin etc) is just within his furthest field of view at the moment of takeover, a distribution may be drawn as shown in Figure 11. Most takeovers are seen to occur before the pilot had seen threshold, let alone the aiming point or glide path origin. The visual sequences indicate that there is little chance of seeing glide path origin from the decision height in Category II operations. Most takeovers occurred as the latter half of the red barrettes between threshold and the innermost lighting bar, came into view. On some occasions the decision to land was announced well before the pilot assumed manual control. The two obvious cases are marked D. While this is understandable and good practice to maintain automatic control, it tends to obscure the relationship between the decision to take over and what was seen. There appears to be no strong trend in shallow fog conditions although some pilots expressed the view that with such a low reported RVR they preferred to wait for threshold to come into view. When deliberate offsets had been injected into the ILS flight path control there appeared to be a tendency to recognise this and take over early provided that a suitable visual segment existed.

5.4 The Go-Around Rate

5.4.1 Flight Results

An important aspect of operation in low weather minima is that the go-around rate should be kept small. Not only is there some risk in the go-around procedure but a large go-around rate would present operational problems. To show the relationship between go-around rate and reported RVR, the flight results are given diagrammatically in histogram form in figure 12. In the band 400 to 800 metres RVR, the average go-around rate was 10% and for the band 200-400 metres RVR the average go-around rate exceeded 40%. If account is only taken of the fresh situations which were presented to the pilots and the repeated runs ignored, the go-around rate for 400-800 metres exceeded 15%. The approaches were grouped in bands of RVR and the relationship between go-around rate and RVR is shown by the curved line in Figure 12. This indicates a go-around rate of approximately 20% at 400 metres RVR.

These flight trials' results represent a mixture of approaches made in Varsity (127) and Comet (23) at decision heights of 100 ft (68) and 80 ft (82) with 74 of them being repeated runs by the pilot who had just made the previous approach. These go-around rates are likely therefore, to be less than those expected in airline service, when the approach is being made after a transit flight in present day types of aircraft with decision heights of 100 ft.

5.4.2 The Effect of Decision Height and Aircraft Geometry on Go-Around Rates

In order to gain a more accurate impression of the go-around rates likely in service, the sample of 150 measured sequences and observers reports were reviewed in the light of the criteria emerging from paragraph 5.1 and figures 9 and 10. For this analysis 'adequate visual reference' was taken to be at least 45 ft between contact height and decision height and a minimum visual segment of 130 metres at 20 ft above the decision height. This review was made for decision heights of 80 ft, 100 ft and 150 ft, also for two cockpit cut-off angles of 0.24 radians (14 degrees) and 0.35 radians (20 degrees). The resulting go-around rates are shown plotted against reported RVR in figure 13. Two effects are worth noting. In the band 300 to 500 metres RVR, go-around rates increase sharply as the decision height is increased and the points exhibit a scatter defying accurate fitting to a curve. This effect is due to the particular fog structures making up each sample of RVR.

For a 150 ft decision height and reported RVR's below 500 metres the go-around rate is high and bears little relationship to the RVR value. This is probably an indication of the lack of correlation between these low RVR values and the visual sequence at 150 to 200 ft on the approach.

5.4.2.1 Go-Around Rates in the Band 400-800 Metres RVR

The sample of 58 approaches in the RVR band 400 to 800 metres is small, but it consists of approaches made in fifteen of the eighteen sorties and the trends are worth noting. For the most favourable combination of 80 ft decision height and cockpit cut-off of 0.35 radians (20 degrees) an average go-around rate of 7% results.

The more representative case of 100 ft decision height and 0.24 radians (14 degrees) cockpit cut-off results in a greatly increased average go-around rate of over 20%. For the higher decision height of 150 ft the average go-around rate rises to 45%.

Only 30 approaches were made in reported RVR values between 500 and 800 metres but these are drawn from eleven different sorties and should contain a representative variety of fog structure. The average go-around rates for this band of RVR now become 13% for a decision height of 150 ft and less than 7% for the 80 ft and 100 ft decision heights.

5.4.2.2 The effect of Cockpit Cut-off

At low decision heights the cockpit cut-off appears to have considerable effect. The average go-around rate for 400 to 800 metres with a decision height of 80 ft increases by a factor of 2:1 when the cut-off angle is changed from 0.35 radians (20 degrees) to 0.24 radians (14 degrees). The effect on the go-around rate for 150 ft decision height however, is negligible. At the higher decision height the decision to land appears to be governed almost entirely by the contact height criterion, the predominant feature of the 'successful landings' being that of descending below a cloudbase with a wide visual segment persisting down to the runway RVR. The cut-off angle had little effect on these conditions, whereas in the visual sequences with lower contact height, the visual segments were frequently small and opened gradually.

5.4.2.3 Variation in Go-Around Rate

The go-around rates calculated in the previous paragraphs for the presently accepted pairings of decision height and minimum RVR are very much larger than most authorities expect or are willing to accept. The effect of increasing the eye to wheel height is similar to an increase in decision height as far as the pilot's visual sequence is concerned and approach speed will affect the time available for his decision. We must therefore expect go-around rates to be influenced not only by fog structure but also by decision height, cockpit cut-off angle, eye to wheel height and approach speed.

6. ANALYSIS OF PERFORMANCE

Having studied in some detail the pilot's decisions and arrived at the criteria used in reaching them, we should now consider if these decisions were correct in leading to landings with a sufficiently low risk of an accident. All the landings were safely completed, evoking little observer comment, and on no occasion did the pilot think that in retrospect the decision to land was to 'risky'. In all except the shallow fog situations the landing became easier as it progressed. The subjective opinion of the crews was that operation to Category II limits was straightforward.

The traditional approach to the certification of an automatic system is to investigate statistical distributions of critical parameters such as rate of descent, lateral error, height over threshold, etc and extrapolate them to cover the extreme envelope of operation. While this approach is not feasible with the limited data at our disposal the comparison between the performance actually achieved in fog and that achieved in clear weather is worth recording.

6.1 Comet Performance

Clear weather landings were made in the Comet with the same pilots who took part in the fog flying, but in winds between 10 and 15 knots. Only two items in the performance comparison are of interest; in fog the height scatter over threshold was smaller and the lateral scatter at touchdown was bigger.

6.2 Varsity Performance

Prior to this fog flying programme an experiment had been conducted in clear weather during day and night to measure piloted landing performance (Armstrong B D 1970) and these results are used as a comparison. Only two pilots were common to both this experiment and the fog flying but the comparison is still thought worthwhile. In the tables below performance is compared over the runway threshold and at touchdown, and the samples are divided into day and night.

OVER RUNWAY THRESHOLD

Time of Day	Weather Condition	Sample	Standard Deviations			Mean Height ft
			Lateral Error m	Rate of Descent ft/sec	Height ft	
Day	Fog	26	3.3	2.0	6.7	58.5
	Clear	32	* 1.6	2.2	* 11	* 48
Night	Fog	16	1.5	1.3	5.0	49
	Clear	24	* 3.0	1.5	* 13.3	55

Reduced height scatter over threshold is highly significant in both day and night. The lateral performance is shown to be better during the night and worse during the day both being highly significant. This is thought to be due to the poor contrast or "washed out" appearance of the approach and runway lights during daytime fogs. The tendency is noticed for the average height above threshold to be greater in fog during the day. On several occasions the pilot was under the impression that the rate of descent was too high after take over and he deliberately flew a shallower flight path.

* Statistical significance at the 1% level is denoted by an asterisk. This means that had the two processes been identical our method of sampling would have a chance of producing these results accidentally once in a hundred equivalent samples. 1% is regarded as highly significant.

AT TOUCHDOWN

Time of Day	Weather Condition	Sample	Standard Deviations				Mean Touchdown Range From GPO m
			Lateral Error m	Lateral Rate m/sec	Rate of Descent ft/sec	Touchdown Range m	
Day	Fog	Between 10-29	3.4 *	0.91 *	0.8 *	140 *	320 *
	Clear	32	1.4	0.3	0.4	76	152
Night	Fog	Between 10-16	2.8	0.85	1.1 *	104	232
	Clear	24	2.2	0.52	0.45	110	168

Fog flying performance at touchdown during the day is significantly worse in all respects than in clear weather. At night only the rate of descent is significantly worse. The day fog results themselves are not very much worse than the night fog results but what is outstanding is the improvement in the day clear weather situations.

The performance results were divided into three sample bands of reported RVR. The samples were small and there was little of significance except the indication that the denser fogs exhibited the greater lateral scatter and the smaller height scatter over threshold. Examination of the probability distributions of these performance parameters gave no more information other than to confirm the reluctance of the pilot to be low over threshold. Separate analysis of the thirteen landings made on 9 December 1969 and 12 December 1969 in shallow fog indicates no degradation in performance resulting from these conditions. A comparison of performance between pilots gave no indication that the fog environment affected the performance of one more than another. A slight difference in emphasis between pitch plane and azimuth could be detected at the just significant level, but this was probably just as much due to variations of fog, site or sortie.

There is thus evidence to show that landing performance in day fog is not as good as in daytime clear weather but little to show that it is unacceptable and no reason at all to question the validity of the decision to land made by the pilots. There is no alarming worsening in performance as the RVR reduces, only a concentration on the pitch plane at the expense of azimuth. Perhaps this effect should be taken into account in certification studies. The sample sizes are, however, too small to extrapolate these results into airline service without confirmation among a wider group of pilots, aircraft and sites.

The automatic approach system performance in the Varsity gave standard deviations at the manual take over point of:-

Rate of Descent	1.7 ft/sec
Glide Path Signal	20 microamps
Localizer Signal	4.3 microamps

This good standard of performance was thought to play a large part in making the tasks of decision making and subsequent control more straightforward for the pilot.

7. VISUAL SEGMENT, SLANT VISUAL RANGE (SVR), RUNWAY VISUAL RANGE (RVR) AND CONTACT HEIGHT

Figure 8 already referred to in para 5.1, illustrates the definitions of the terms slant visual range, visual segment and contact height. The relationship between SVR and visual segment is strictly non-linear, but cut-off angles of 14 and 20 degrees are sufficiently small to make a linear approximation accurate to within 2 metres of visual segment. Diagrams may therefore be drawn with both these quantities on the same axis but with a datum shift and a very small difference in scaling. On a particular approach there is only one contact height and only one latest reported RVR although this may differ, due to tolerances and space and time variations, from what is actually seen along the runway. The SVR will, of course, be different at different heights depending on the fog structure. If the fog and lighting brilliance are uniform there should be little difference between the SVR and the RVR.

Figures 14, 15 and 16 compiled from the 'click box' observations illustrate the variation between the reported RVR and the encountered SVR at three heights, 50ft, 100ft and 150ft. There is clearly no simple relationship between SVR and reported RVR, for even at 50 ft wheel height there is a large scatter and the scatter increases with height. Operational crews must be made aware of the fact that a reported RVR does not guarantee a particular visual segment even at 50ft height and much less so at increased heights.

Figure 14 does not show a correct balance at RVR values below 300 metres since the thickest fogs resulted in many go-arounds with negligible contact so that the segment at 50 ft could not be assessed.

7.1 The Probability of a Given Visual Segment

The flight results shown in figures 14, 15 and 16 are expressed in a different way in figures 17, 18 and 19 where curves are drawn of the probability of a given visual segment at each of the three wheel heights 50 ft, 100 ft and 150 ft. In each figure the results are divided into 4 RVR groups and the sample size is quoted against each curve. These curves are all drawn for a cockpit cut-off angle of 0.24 radians (14 degrees). The samples are small and care should be taken in extrapolation or using them in a wider context. The high probability of small visual segments at 100ft and 150ft wheel height confirms the high go-around rates already discussed in paragraph 5.

7.2 Variation of Contact Height with Reported RVR

Figure 20 illustrates the variation of contact height with RVR which was measured during the trials. Such a large scatter serves, yet again, to emphasise the wide variety of visual sequences which are possible.

8. SPECIAL EFFECTS AND CREW COMMENTS

This section contains a collection of miscellaneous but interesting features which arose during the fog flying. Many of the points have been made by previous investigators but a repetition may help to complete the picture for operational crews and other readers.

8.1 The Lighting and Runway Markings

Having been engaged on experimental work on the BLEU simulator (72 approaches each in simulated night conditions), the newer pilots were surprised by the appearance of the lights in real daytime fog. They appeared 'washed-out' and not at all sharp in intensity or colour. The observer was sometimes not sure of the exact moment of becoming conscious of a light as it passed through the stages from nothing to a vague glow to becoming a recognisable light. In shallow fog he noted that he tended to look in the wrong place since the glow appeared first on the fog top, directly above the lights. The colour of the red barrettes, and more frequently the green threshold bar, were often not consciously observed in thick daytime fog. The painted white threshold bars and the runway marking at Bedford stood out in the daytime, invariably giving better guidance than the touchdown zone lighting and the threshold bar.

The so called 'black hole' phenomenon where no visual information seems to appear beyond threshold can be recognised in many of the visual sequences accentuating the effect of shallow fog. The way many of the lines bulge to the left below the threshold line, strengthens the impression that shortening sequences are not only the result of fog structure but that a large contribution comes from the lighting intensity ratio between approach and runway lighting. The alignment of the lights also has a large effect, on some occasions the VASI lights were seen before any of the runway lights, thus providing a most useful reference.

During the day, the runway texture at Bedford - concrete with many tyre marks - gave useful knowledge of the ground plane. The Liverpool runway appearing uniformly black, lacked this feature. During the second sortie at Liverpool a light covering of snow on the runway masked the lights, reduced the contrast and hence the effectiveness of the touchdown zone lighting.

At Liverpool and Schiphol the crews commented on the lack of the extended lighting bars marking glide path origin. These bars give useful range information during the flare and help to identify the ground plane.

8.2 RVR Fluctuations

Most operational crews are well aware that patchy and rapidly changing fogs occur, but they may still be somewhat critical when the reported RVR does not agree with what they see during landing. Even with the direct-contact reporting used during the flying at Bedford, changes of RVR from 330 to 500 metres and 430 to 300 metres occurred without time to update the pilot with the new value. The latter occasion resulted in a go-around due to lack of visual reference at 100 ft when the pilot was expecting a Category II RVR. The previous approaches in the sortie had been made with increasing RVR.

On the sortie at Schiphol the reported RVR values were noticeably pessimistic. Some of this was probably due to a time lag in reading and relaying the RVR and some probably due to a natural caution in announcing an improved situation. It is important that procedures of this nature, however reasonable, do not discredit the reported RVR in the pilot's mind. A situation can easily be imagined where on one day a pilot is passed an RVR value which he finds is pessimistic. If on the next occasion he encounters a shallow fog in which there is a large visual segment above decision height he may well be tempted to disbelieve the low RVR value and continue the landing.

9. CONCLUSION

The pilots felt that operation in Category II weather conditions with good quality approach performance, strict crew drills and accurate RVR reports, was straightforward and relatively easy. Only when the results were analysed did the difficulties appear in their true perspective. What began as a subsidiary objective, the collection of more data on visual sequences, became one of the more important aspects of the programme because it demonstrated the wide variations which can be met. Figures 3 to 7 and 14 to 20 speak for themselves on the lack of relationship between RVR, contact height and visual segment at decision height. There are important lessons for operational crews in these diagrams. Because there is no magically sharp change in conditions at 400 metres RVR, or indeed, at any other value, on many occasions when the RVR is just below limits a successful landing could be feasible and safe. To avoid RVR reports getting discredited these variations in visual sequences must be appreciated by the crews.

A feature which will probably determine the minimum limits for a manual type of Category II operation is the go-around rate. The pilots appeared to require approximately 4 seconds to assess a minimum visual segment of 130 metres in order to make a decision to land, low contact heights which occurred in several sorties result in a go-around rate higher than expected in Category II RVRs. If this sample is representative, the effect of these low contact heights is accentuated by the choice of decision height, the cockpit cut-off angle, the eye to wheel height and the approach speed. The expected go-around rates for this sample of fogs range from 7% for a decision height of 80 ft in a Varsity which has a cockpit cut-off of 20 degrees, to 20% for a 100 ft decision height in a Comet with a 14 degree cockpit cut-off. Importance has rightly been attached to the go-around rate tolerable for Air Traffic Control purposes. However, the average rates of go-around really have little relevance to Air Traffic workload since the go-around rate may be very high in particularly adverse fog structures, and zero in others with the same RVR. When consideration is being given to operating limits the particular go-around performance and geometrical characteristics of the aircraft type must also be taken into account.

The criteria established for adequate visual reference in fogs with low contact height do not apply to shallow fog conditions, and there is potentially a risk of the pilot being led into a dangerously shortening sequence. The pilot must temper what he sees with a knowledge of the reported RVR. Strict adherence to the RVR limits, accurate RVR reporting and decision heights as low as consistent with the approach and go-around performance, would appear to solve this problem. If the decision height is low there is little chance of a dangerously short segment occurring after decision height before the visibility improves to the RVR on the runway. Descriptive assistance to the approaching pilot such as 'low contact height', 'shallow fog with minimum segment at 'Z' ft', 'patchy fog' etc would undoubtedly be useful.

The flight path and touchdown performance achieved in fog is not easy to assess. The touchdown performance was worse in many respects than in an equivalent clear weather sample, but it was by no means unacceptable. Performance in fog compared with clear weather was worse during the day than at night. The pilots commented that guidance from the lighting, and in particular the touchdown zone lighting, was poor during the day due to the 'washed out' appearance and poor contrast. Azimuth performance appeared to be degraded more than pitch performance due, perhaps, to the concentration on the pitch plane. A tendency was noticed to keep high over threshold.

Category II operations with manual landing down to limits of 100 ft decision height and 500 metres RVR appear to present no problems provided that high standards of equipment and crew drills are maintained. As the RVR is reduced to 400 metres the effect of fogs with low contact height is to increase the go-around rate. Shallow fogs are potentially deceptive in tempting the pilot to land in very low RVR's. Decision heights should not be set arbitrarily high, but should be appropriate to the approach and go-around performance. The flying described in this paper is necessarily a small sample and the author sees the logical continuation being an analysis of instrumented Category II landings in service.

REFERENCES

1	H M Stationery Office CAP 321	United Kingdom Provisional Operating Requirements for All Weather Operations Category II. Board of Trade London 1969
2	A D Brown	A Preliminary Simulator Investigation of the Problems of Category II Operation. Unpublished MOD(PE) Report (1968)
3	A D Brown	Category II. A Simulation Study of Approaches and Landings at Night. RAE Technical Report 71044 (1971)
4	B D Armstrong	Flight Trials to Discover Whether Peripheral Vision is needed for Landing. RAE Technical Report 70205 (1970)

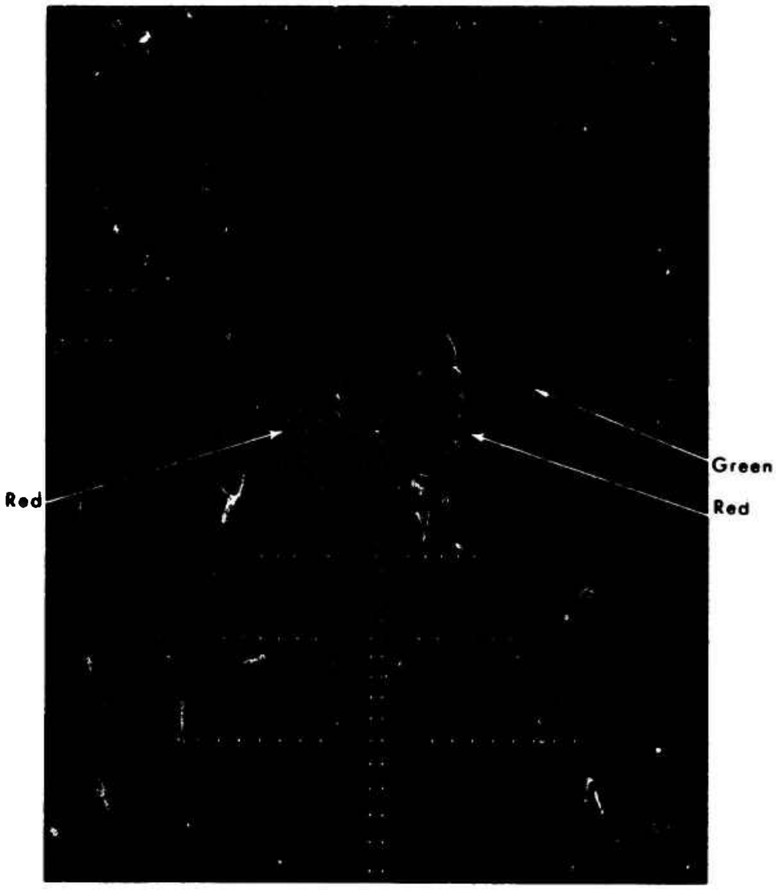


Fig.1 Cat II lighting Bedford airfield

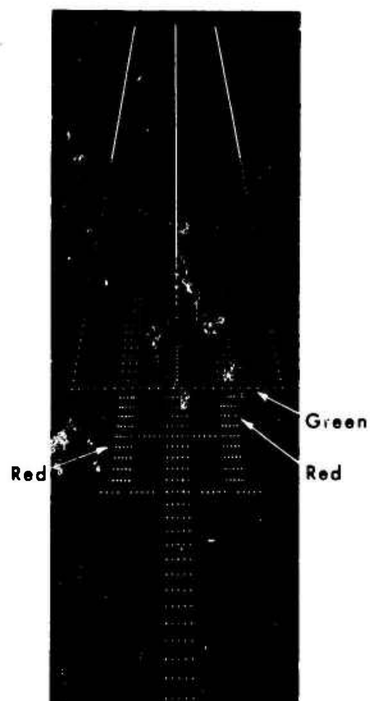


Fig.2 Cat II lighting Schiphol airport

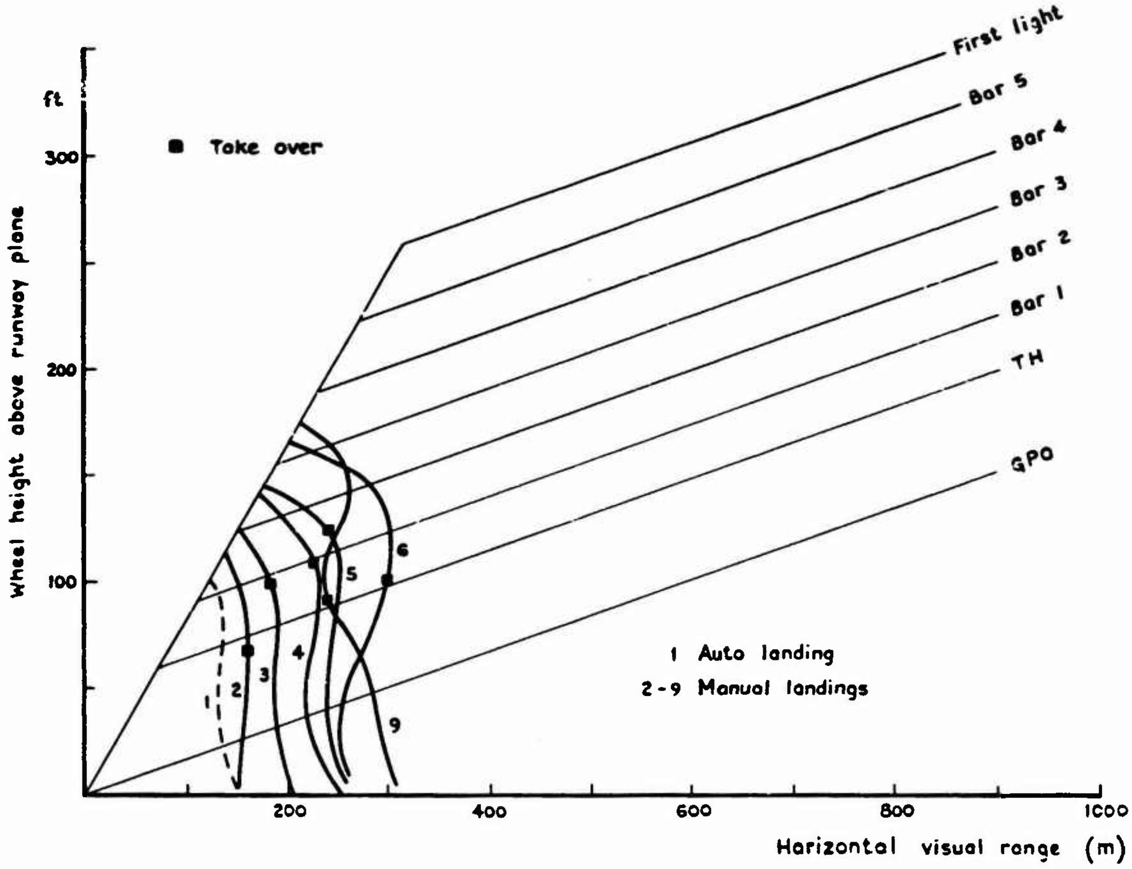


Fig.3 Visual sequences. Bedford day fog 19 March 1969

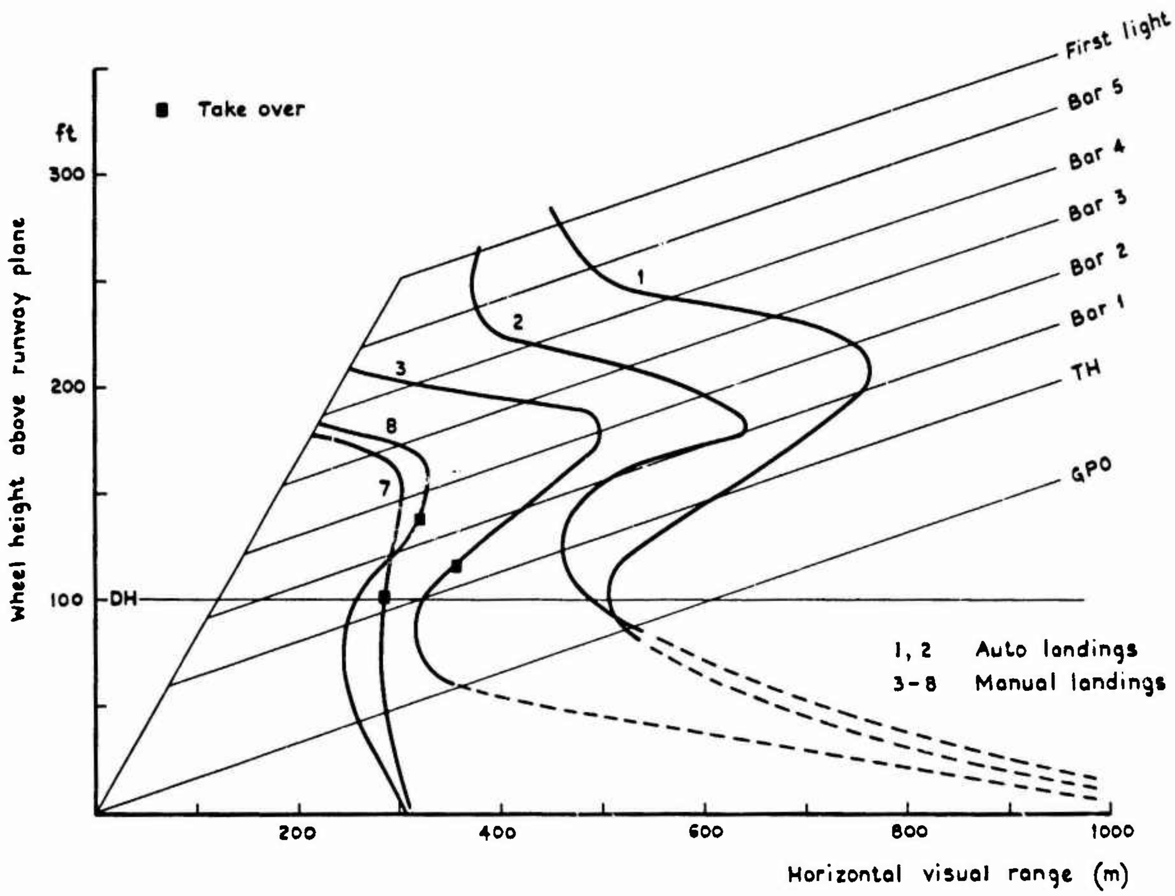


Fig.4 Visual sequences. Bedford night fog 19 March 1969

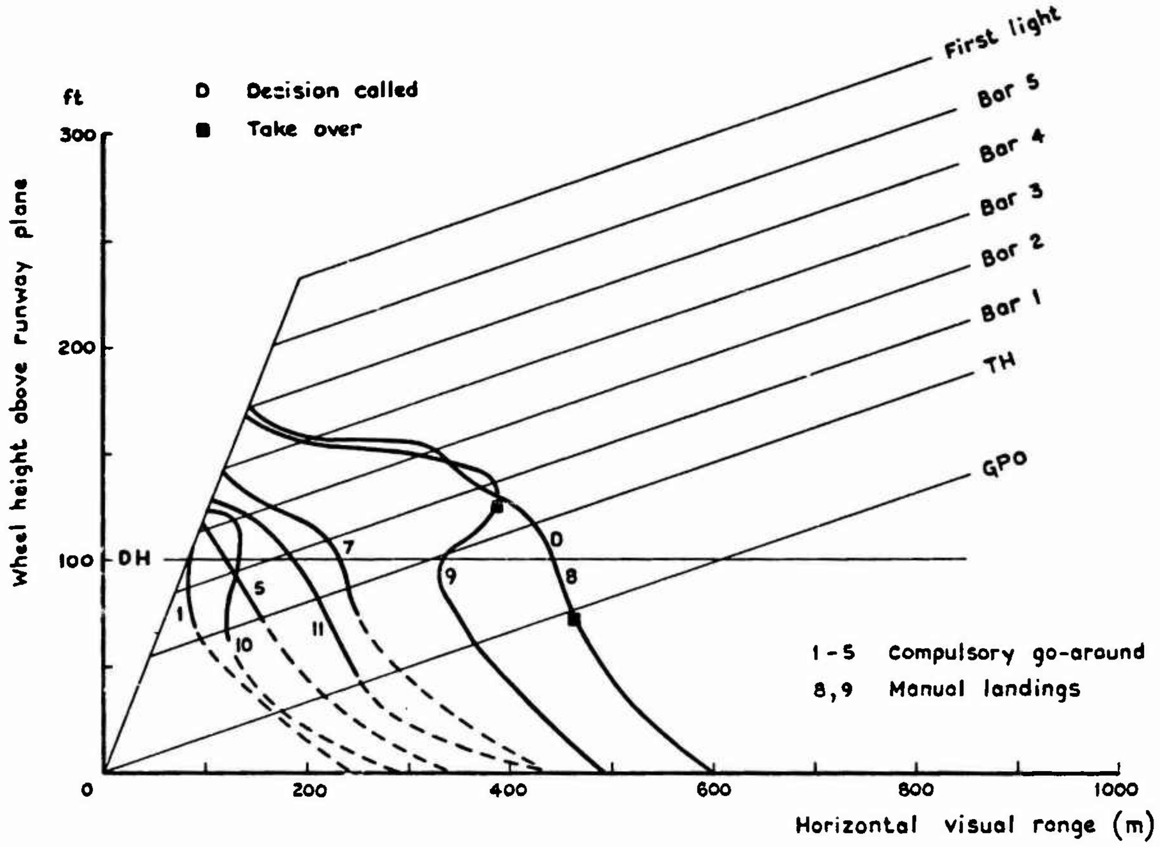


Fig.5 Visual sequences. Bedford day fog 18 September 1969

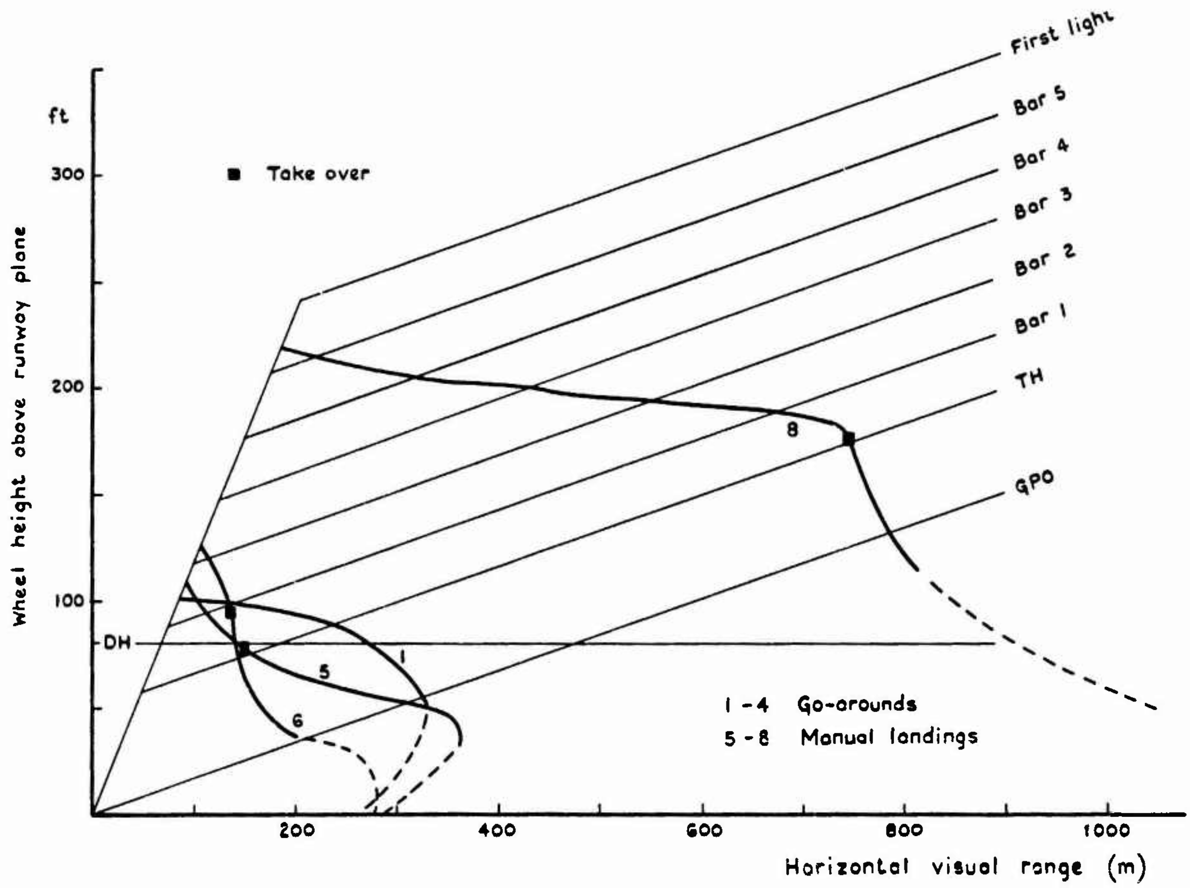


Fig.6 Visual sequences. Bedford day fog 5 May 1969

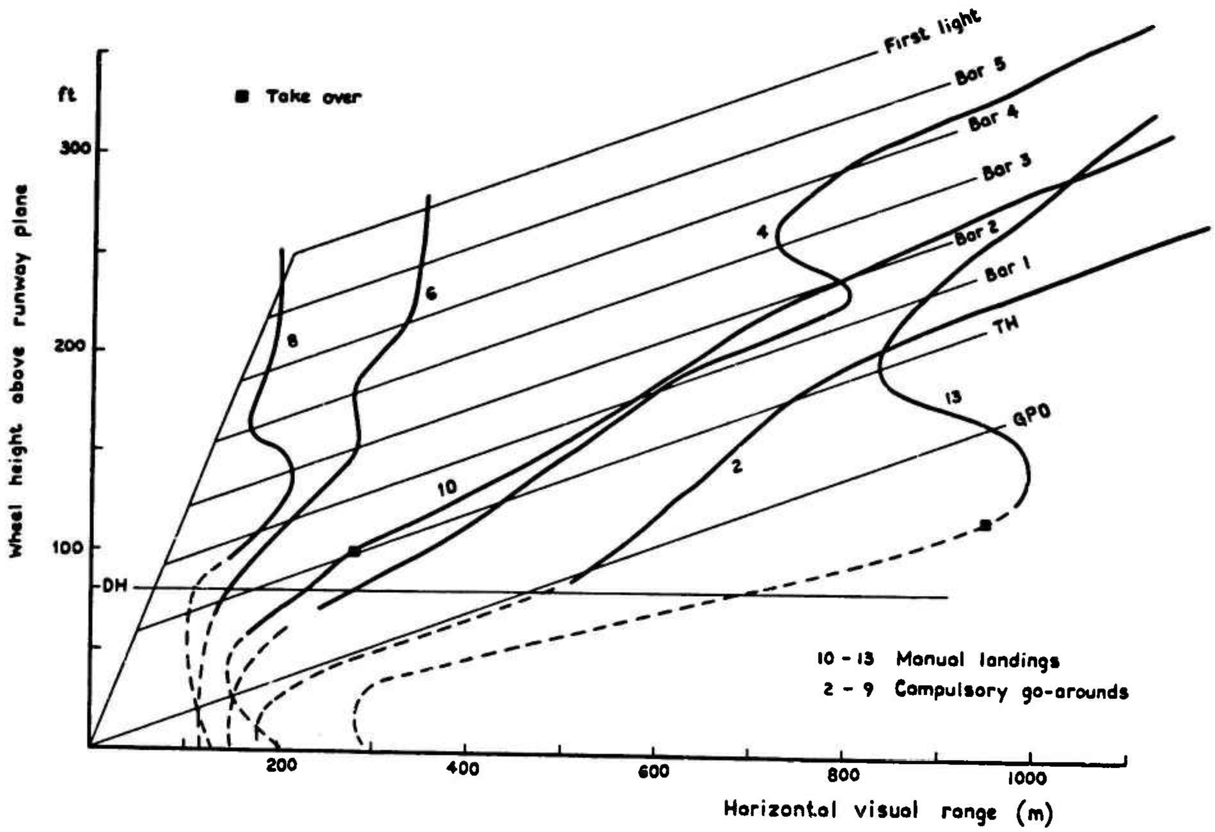


Fig.7 Visual sequences. Bedford night shallow fog 9 December 1969

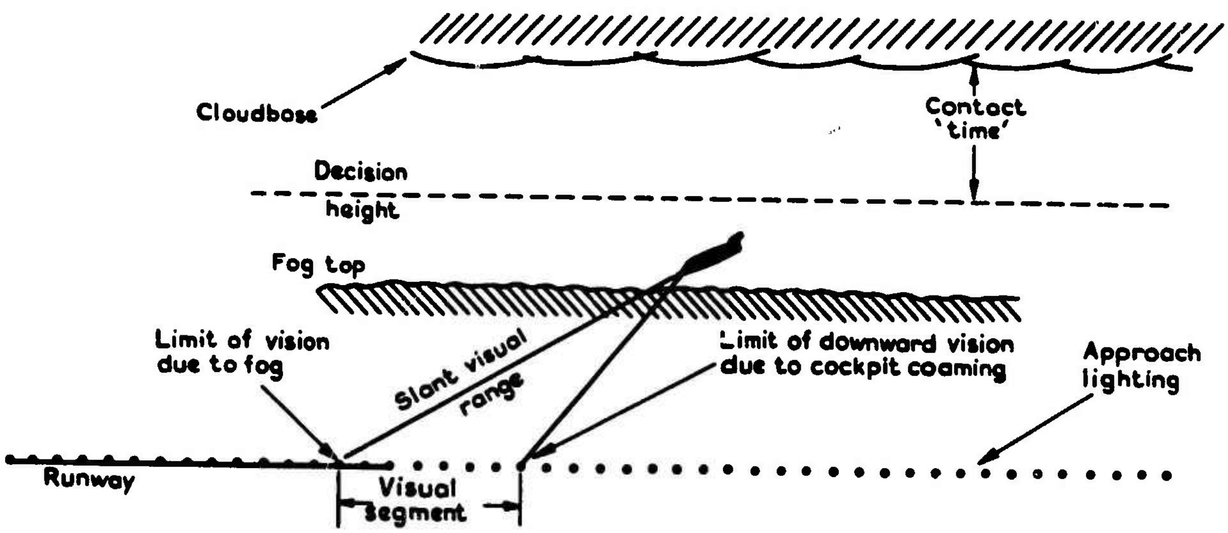


Fig.8 The category II operation – relevant terminology

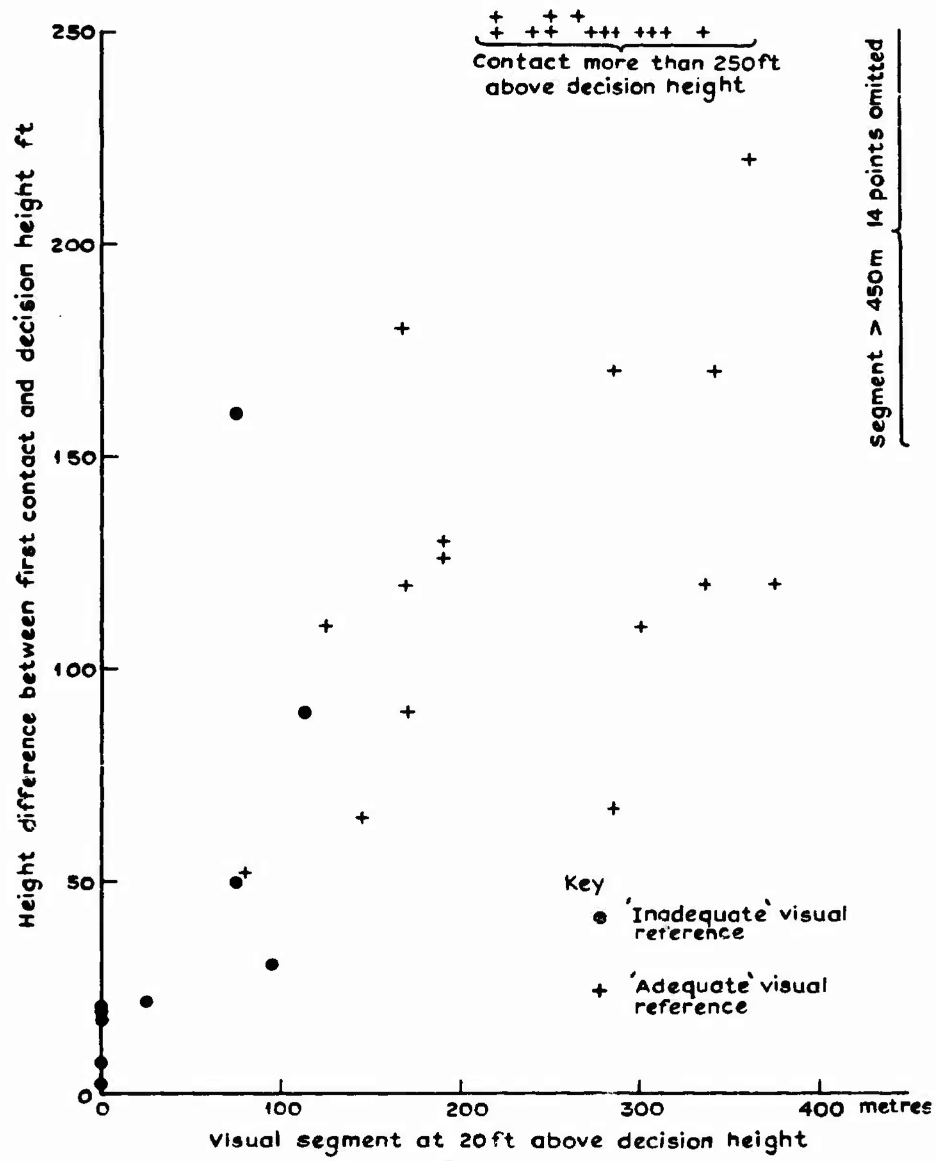


Fig.9 The pilot's decision to land – first time runs

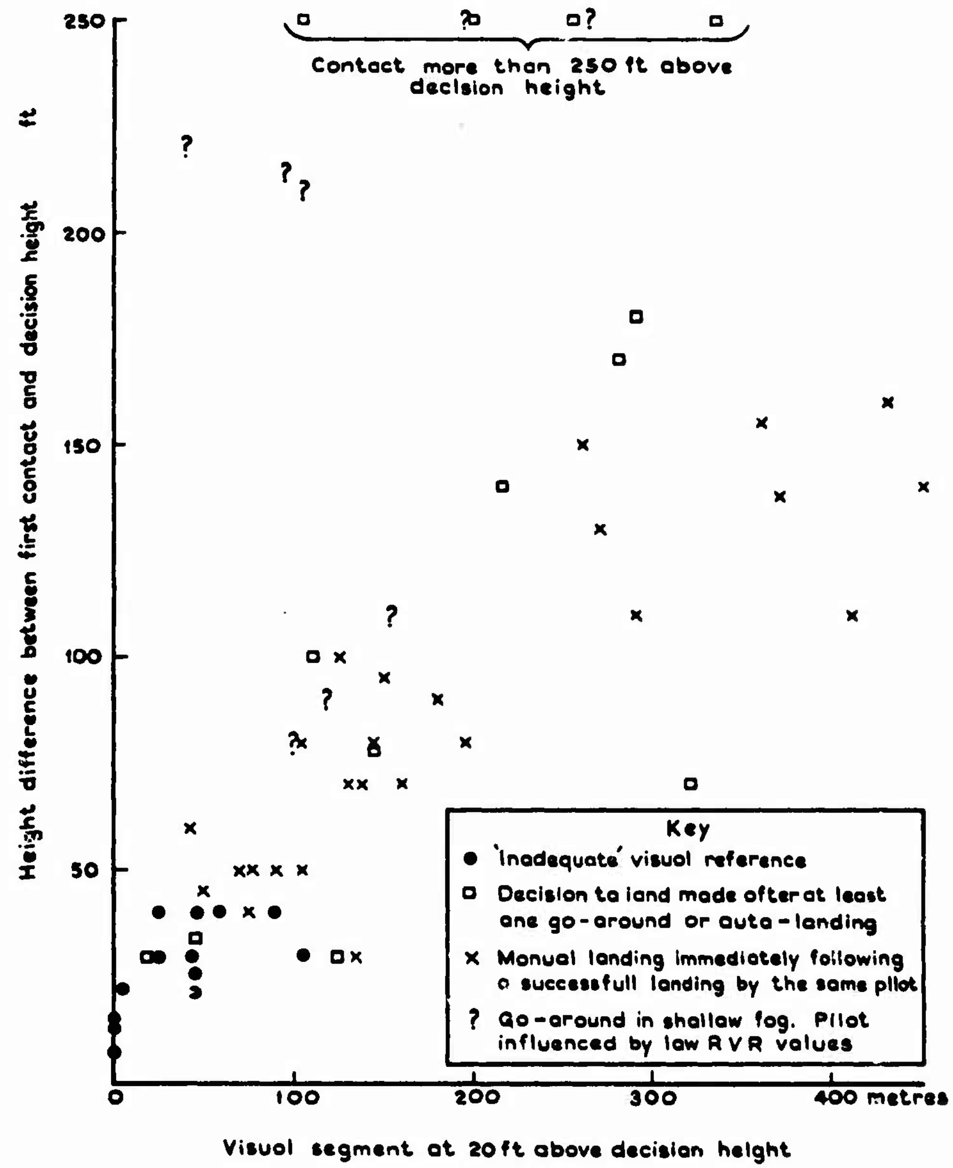


Fig.10 The pilot's decision to land – repeated runs

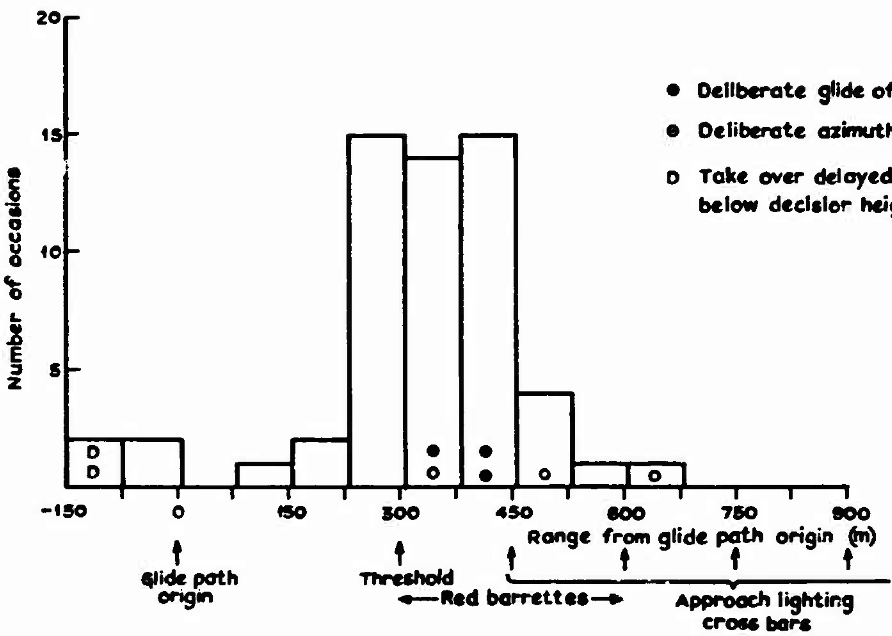


Fig.11 Far point of vision at the moment of take over

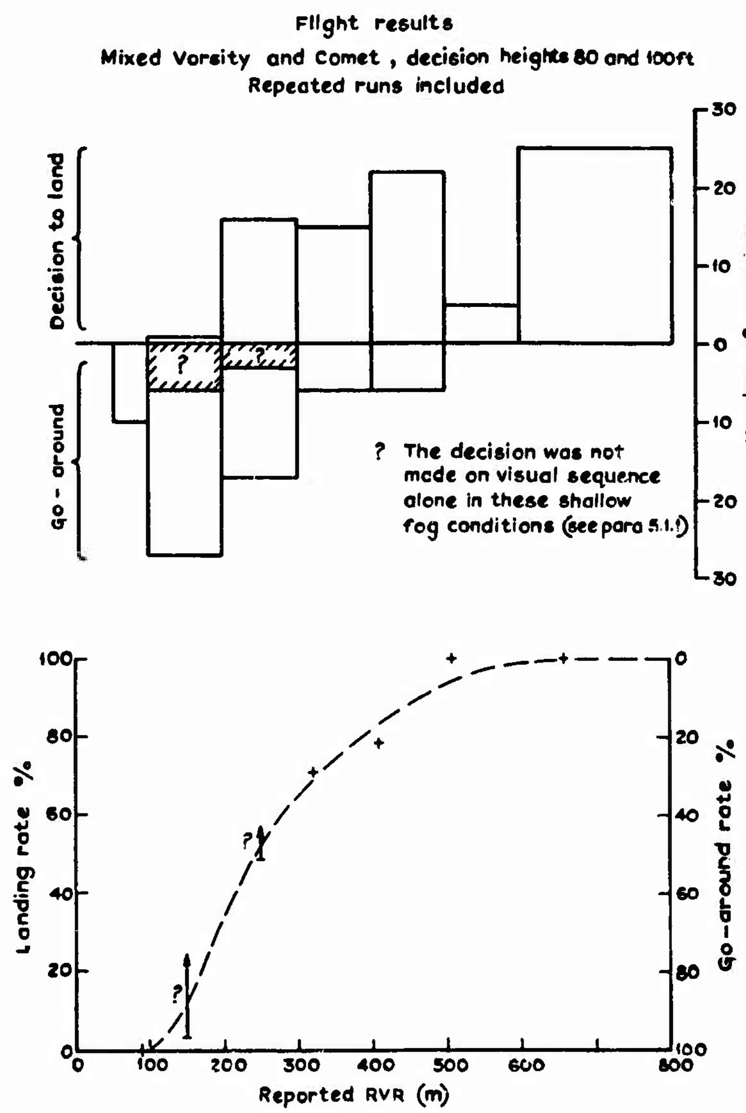


Fig.12 The variation of go-around rate with RVR

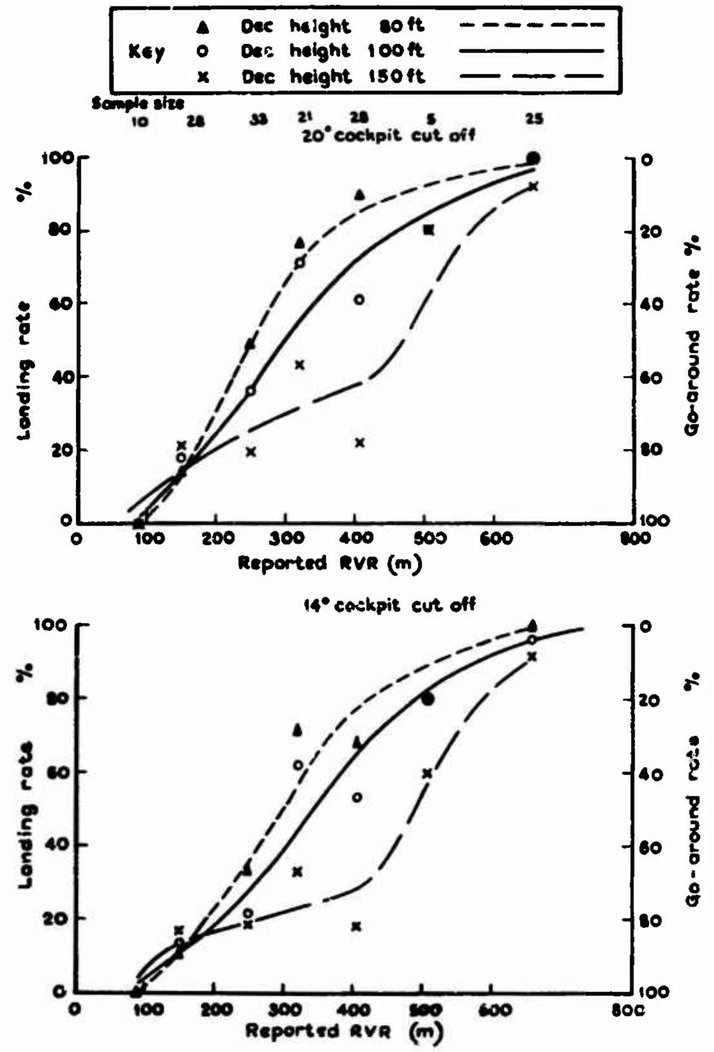


Fig.13 Probable effect of cockpit cut-off and decision height on the go-around rate for 'first time' approaches

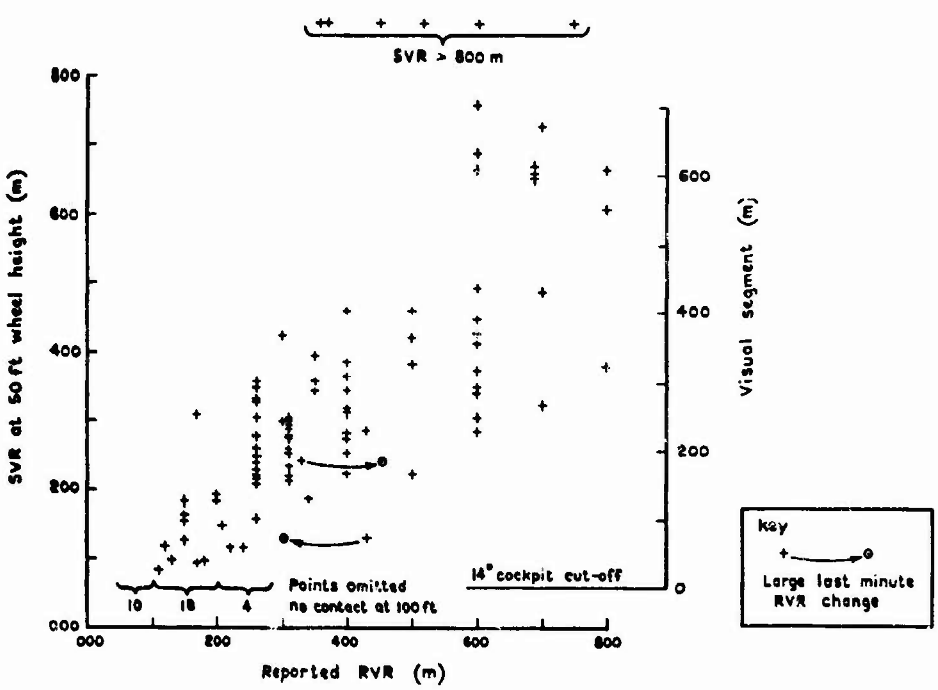


Fig.14 The variation of SVR at 50ft with reported RVR

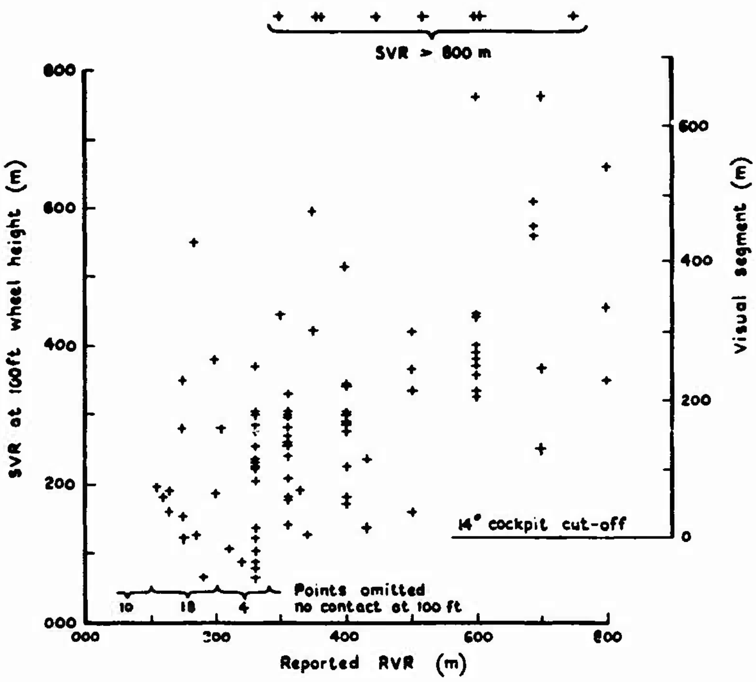


Fig.15 The variation of SVR at 100 ft with reported RVR

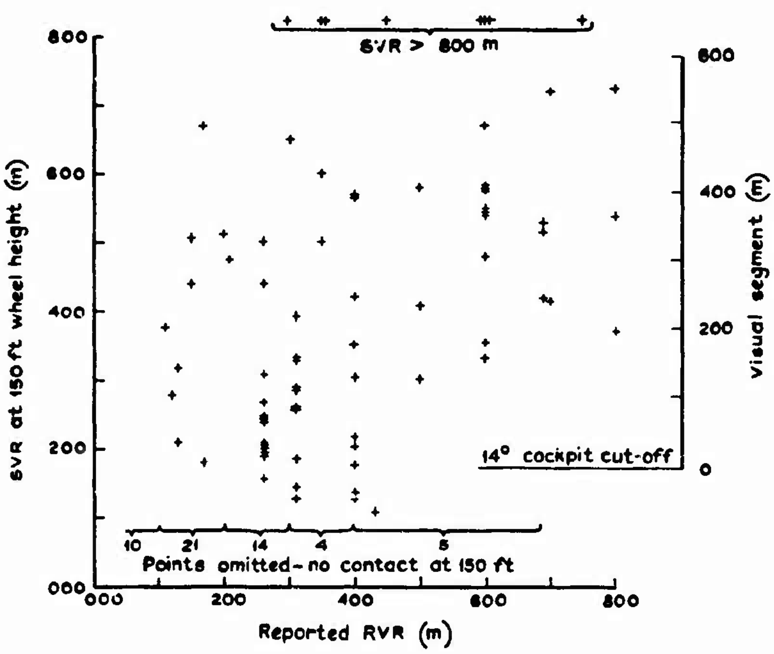


Fig.16 The variation of SVR at 150ft with reported RVR

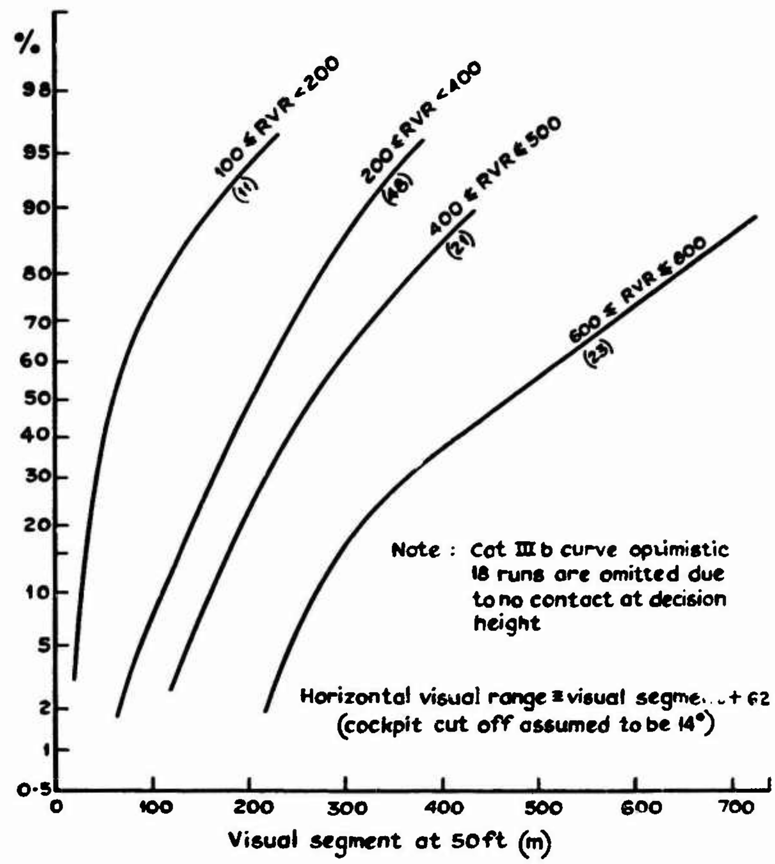


Fig.17 Probability of less than a given visual segment at 50 ft

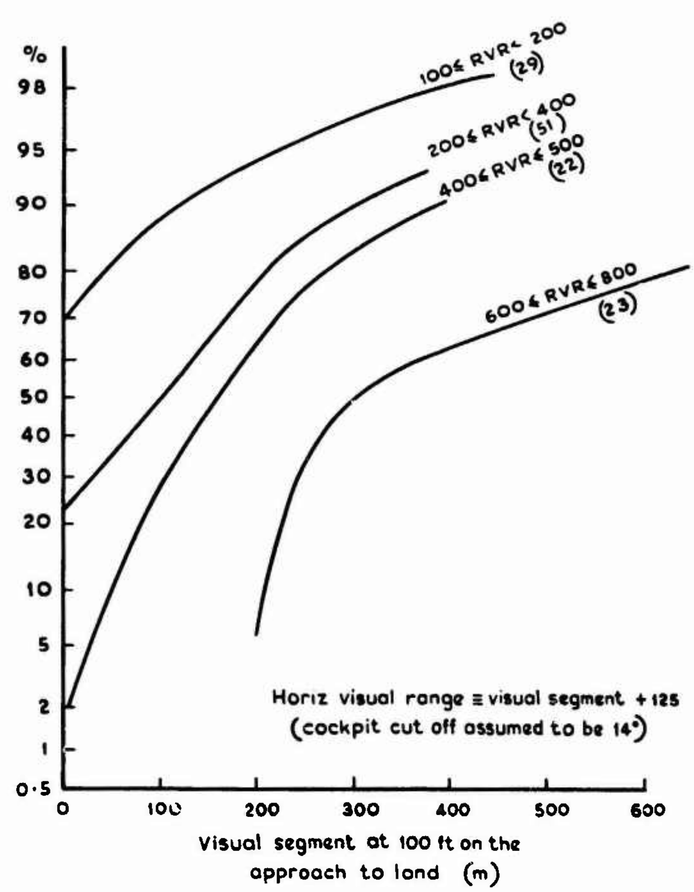


Fig.18 Probability of less than a given visual segment at 100 ft

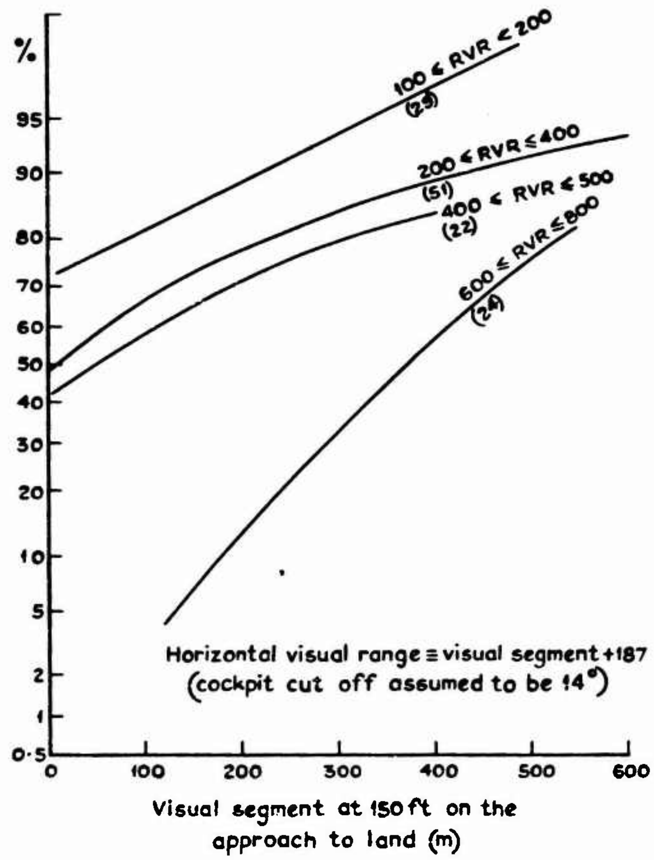


Fig.19 Probability of less than a given visual segment at 150 ft

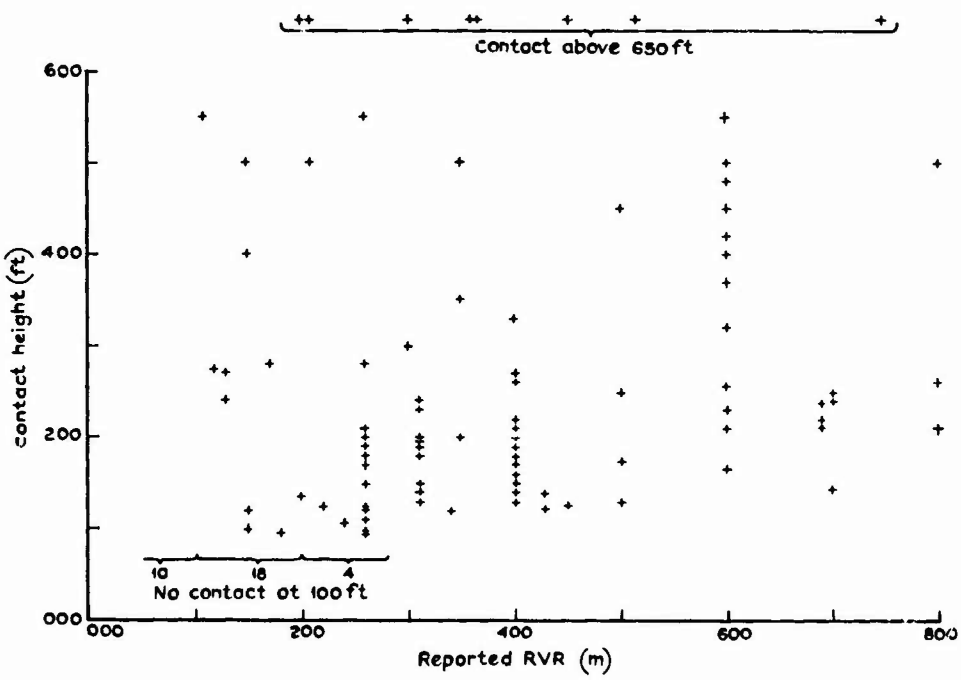


Fig.20 The variation of contact height with reported RVR

MONTE CARLO SIMULATION OF DEGRADED MAN-MACHINE PERFORMANCE

Gerald P. Chubb
Aerospace Medical Research Laboratory
Aerospace Medical Division
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio

SUMMARY

System vulnerability is a function of both human and hardware vulnerabilities to anticipated threat environments. The feasibility of considering the interaction of man and machine degradation under nuclear attack conditions has recently been demonstrated. It appears that the technique may be useful in identifying certain situations where automation may be particularly useful under these attack conditions, although the requirement is not obvious from (or justified by) analyses of system performance under nominal operating conditions. The approach taken appears generalizable to other degradation conditions, such as inflight malfunctions and conventional weapons battle damage. Although the model used does not directly imply what design changes would facilitate system performance, it should aid the designer in such analyses. Moreover, given suggested changes in man-machine task sequencing, the model can aid in assessing how these changes may affect selected systems effectiveness measures. A number of refinements and extensions to the current capabilities of this model are envisioned and briefly discussed.

1.0 INTRODUCTION

By developing techniques to assess the vulnerability and survivability of existing systems, methods will be available for balancing the hardness criteria for future systems. To achieve this goal of balanced hardness to weapons threats, one must consider not only the viability of each subsystem but its interaction with other subsystems. In manned systems, it is necessary to know not only how the man and the machine degrade following exposure to weapons effects (e.g., ionizing radiation, conventional weapons, battle damage, etc.) but also how the machine degradation affects man when he too may be degraded in his performance capabilities. Vulnerability of man or machine does not directly imply system vulnerability, and certainly vulnerability does not directly imply a lack of survivability. It is most difficult, if not impossible, to contrive an analytic technique which adequately captures the structure and dynamics of the interaction between man and machine performance. Monte Carlo simulation is one expedient means for trying to capture a more realistic (though perhaps currently crude) description of system performance. By describing how the performance of human and hardware change when degraded by exposure to attack conditions, the model can help postulate what happens to system performance and perhaps to mission success. While this may not answer all of the concerns about survivability, it is believed that such techniques are especially useful to designers whose primary objectives may be simply to evaluate the preferability of proposed alternative designs.

This sort of approach seems particularly useful in the design of systems with automated features where man serves as a monitor and back-up in case of equipment malfunction, due to whatever cause. The status of the man, his current and projected workload can all affect how well he executes his role of back-up to the automated features; and in turn, knowledge of such inabilities to manage this workload if his own capabilities are degraded will be important in specifying the required reliability and hardness of the equipment. While most of the efforts to date have focused on man-machine modeling under nuclear attack conditions, the potential exists for expanding the technology to treat other kinds of induced man-machine degradation.

1.1 Background

In approaching the problem of survivability/vulnerability (S/V), prior efforts have focused on determining equipment vulnerabilities. From these vulnerability assessments, estimates have been derived of subsystem and system operability. Component vulnerabilities were typically one of two sorts: a transient disruption of function, or the total cessation of function. Since different components make up a system and alternate structural arrangements of components constitute the alternative designs for a system, each new system requires a wholly new vulnerability analysis. But, it should be pointed out, survivability depends not only on the inherent vulnerabilities of the components but to a large extent on the use, strategic and tactical deployment procedures, mission contingencies, threat environments, scenario characteristics, etc. However, these latter considerations are probably best treated in what is referred to as "war gaming" analyses.

Human vulnerabilities are somewhat different than equipment vulnerabilities. Man comes with a fixed complement of components in a virtually unalterable design configuration. Human vulnerability is variable only to the extent of intra- and inter-person variability. While this is true, whatever the variability in performance capability, it can have a disproportionately large (or small) impact on the mission, depending on what tasks are to be performed, when and under what workload conditions. It appears that the physiological vulnerabilities, if completely and accurately described, would remain static, as "defined". However, performance vulnerability would be expected to be dynamic in character. While man himself cannot be redesigned, his tasks, procedures, and tactics can be. Thus, the human's contribution to system vulnerability can be viewed as having both static and dynamic components.

Because of the dynamic component, meaningful ad hoc analyses of human vulnerability per se are difficult, if not in fact impossible. However, it does appear possible to examine man's performance in the specific context of mission relevant events. To do this, it has been proposed that an existing

Monte Carlo simulation model of operator performance (Siegel and Wolf, 1969) be modified and adapted to treating the problem of human vulnerability. This development forms the basis of this paper.

1.2 Rationale

It was believed that a demonstration of feasibility was required in determining whether a suitable approach could be devised for treating the S/V problem in a more systematic manner than had been done to date. As part of an S/V study of an existing interceptor, an attempt was made to develop the techniques for treating human vulnerability while hardware vulnerabilities were being separately analyzed.

Since efforts were conducted in parallel, the hardware studies essentially assumed the human component could be treated independently, and the human studies assumed the hardware was wholly operational. In a subsequent effort, the problem of man-machine interaction was examined in an attempt to consider the impact hardware vulnerabilities have on the dynamic human vulnerability.

In implementing these considerations into the resulting computer programs, a conceptual scheme was adopted that would permit future developments that were not necessarily directed to nuclear S/V. The treatment of static human vulnerabilities were handled in a "preprocessing" routine that adjusts selected task parameters (input data) to reflect the expected impact of the radiation abosrbed dose. A more complete description may be found in Chubb (1971).

The impact that radiation induced hardware degradation has on man's performance was treated in a separate preprocessing routine. This separation of hardware and human vulnerability assessments necessitated the following sequence of processing. Given the normal sequence of tasks executed by an operator, the equipment degradation preprocessor was entered first to examine the feasibility of accomplishing this a priori sequence under the specified exposure conditions. If an equipment required for a task was unavailable due to radiation induced degradation, an alternate task or family of tasks was substituted. Thus, a new sequence of tasks was generated by the equipment preprocessor. This new task sequence, representing the operator's response to equipment degradation,was then passed through the human degradation preprocessor. Here, the task parameters (performance time and probability of success) were adjusted to reflect the impact the specified radiation exposure would have on man.

1.3 Utility

Because of this separation of human and hardware preprocessing, it is possible to use the routines in different combinations to explore distinctly different mission/threat contexts, as illustrated in Table I. While the threat environment depicted is for a nuclear attack, the scheme is equally applicable to other contexts. A similar matrix could be contrived for any of the following: (1) in-flight degradation under nominal conditions (e.g., in-flight malfunction); (2) conventional weapons (e.g., battle damage and personal injury); (3) chemical agents; and (4) bacteriological agents. Although programs do not now exist to treat each of these, it appears possible to develop them. They then might be used individually or in some combinatorial manner with other routines. Perhaps the greatest utility of this approach is the ability to examine what hardening and protection may afford in enhancing mission effectiveness. Relative comparisons can be made between the outcomes of model runs under each of the conditions depicted in Table I, for any of the threats or threat combinations deemed to be of interest. Such comparisons could indirectly imply the value associated with different approaches to enhancing system vulnerability: protecting man, protecting the machine, or protecting both.

2. METHODOLOGY

It is apparent that the usefulness of the simulation of man-machine performance is not solely grounded in the assessment of nuclear vulnerability/survivability. An attempt will be made to describe the nature of the general capability to treat equipment degradation and its impact on man-machine effectiveness. Particular emphasis is given to progress and plans for future developments rather than simply citing past accomplishments.

2.1 Concepts

From an operator's point of view. it makes little difference what causes a deficiency in the equipment he uses. The realities are that he must compensate for that operating deficiency, either by changing the nature of his performance or by electing to execute some alternate course of action. In a gross attempt to identify categorically the kinds of problems faced by an operator, the matrix presented as Table II was prepared. It recognizes a distinction between malfunction and symptom and further allows a graduation in the severity of the malfunction. The text in each cell is a terse commentary or the potential significance of the associated symptom/equipment-status combination.

From even this gross comparative analysis, it is apparent that rather dramatic differences exist which could dominantly affect system performance as a function of the manner in which man must respond to the equipment condition.

The immediately detectable malfunctions can lead the operator to perform a series of operational checks to determine systems status so a tactical decision can be made to abort some or all of the mission objectives or to continue under some appropriate but alternate attack mode against the same or some secondary set of mission objectives.

During system design, one would like to explore the impact not only of automated functional systems but automated support systems, such as in-flight malfunction detection and fault isolation systems. In each case, however, one would also like to explore the impact operating deficiencies in these systems might have on mission success. Fault trees are often generated during design to depict the interdependencies among signal events. From such data, it appears feasible to determine the nature of the display(s) presented to the operator for any set of conditions of interest. To close the loop and determine systems performance, it is necessary then to examine the operator's response to these conditions. Table III exemplifies a possible starting point for such analyses—a tabular examination of control/display combinations.

2.2 Approach

Several approaches can be taken in attempting to define the operator's response to the specified set of equipment contingencies. First, one can resort to expert opinion and contrive an assumed response or prepare a statement of a "preferred" response to the given contingency condition. Second, one could expand this approach to consider classes of experts: (1) the designer, (2) "naive" subjects with requisite technical backgrounds and experience, or (3) representatives of the user population—either staff representatives or operational personnel. Third, one could attempt to simulate the contingency events and observe the responses elicited in trained or untrained subjects. Fourth, one could provide extensive training to operators on each of a set of alternative responses and measure their performance under the contingency situation. None of these approaches are novel or new, and neither are they solutions in and of themselves, which is the essential point of this digression.

The difficulty is in extrapolating from the particular to the general. While such analyses are necessary in defining possible or representative responses to an event or set of events, they do not directly imply what happens to the mission as a function of their occurrence. This will depend not only on that response but when the events occur that precipitate that response (within the mission profile) and what consequent actions may be necessary due to that response.

In turn, this set of system dynamics requires a set of techniques which will capture the structural relationships and interdependencies among the event contingencies antecedent or subsequent to any event set of particular interest.

2.3 Techniques

Assuming that one can, by whatever approach, define the precedence relationships among events and assign to each event a set of attributes which aptly depict the character (time and duration of occurrence, etc.) of that event, it should be possible to explore the dynamic interactions which occur among these events in the course of executing a mission. Tractable analytic techniques have been proposed for linear networks of events (Pritsker, 1966), but it is necessary to resort to simulation as multiple contingencies and nonlinearities are added to the network of events that make up the description of the man-machine system.

The Siegel-Wolf model was programmed in FORTRAN. Among its nonlinearities are time dependent changes in event attributes (performance time and probability of task success) which are conditional upon work-load, time stress, and other factors (such as operator proficiency). In exploring alternative program languages, an application of the Siegel-Wolf model was found which used GPSS/360 (Kochhar, 1971). The contrasts between these two computer languages as modeling techniques highlight some of the advantages/disadvantages of each.

In the FORTRAN simulation, task attributes and the dependency relationships in the network depicting the task sequence are all input data. The model uses the time dependent relationships of time stress and crew cohesiveness to process these input data. In the GPSS/360 version, the relationships are treated as the input, and the structural relationship among the tasks constitutes the model. This latter representation is particularly awkward for design analyses, since it is the task sequencing and task attributes which will be affected by changes in panel layout or control/display design features or characteristics.

However, there is an inherent advantage in selecting a "simulation language" for executing the Siegel-Wolf model. If the model is to be used, it must be cast within the vernacular of the intended user. If that user is to be a systems analyst, the model should be easily programmed in an "available" computer language currently used for systems simulation.

Consequently, other alternatives were explored but no other applications of the Siegel-Wolf model were found. The Simscript language, while originally developed under the Air Force's Project Rand, contains concepts which appeared analogous to those used in the Siegel-Wolf model. However, the use of Simscript (Kiviat, Villanueva, and Markowitz, 1968) depends on the availability of a Simscript compiler and that limits one's ability to adapt, modify and extend the capabilities of the language.

Another and more promising alternative was GASP II (Pritsker and Kiviat, 1969) in that this language was FORTRAN based so that no special compiler is required; and therefore, the routines can be readily adapted or modified. GASP is conceptually analogous to Simscript. In exploring the suitability of GASP as a vehicle for programming the Siegel-Wolf model, it was learned that a specially tailored modification of selected GASP routines was more suited to the purpose. This package of routines is titled P-GERTS (Pritsker, 1971) and is an outgrowth of the difficulties encountered in attempting to study network relationships analytically. Even so, the P-GERTS programs require some modification to permit representation of the time dependent workload management dynamics of the Siegel-Wolf model.

Another potential advantage to exploring the use of GASP as a programming technique for man-machine simulation models is that attempts are being made to "hybridize" the model. Methods are being devised to treat continuous time variables within the context of discrete event simulation. This promises an opportunity to overcome some acknowledged deficiences in the way the Siegel-Wolf attempts to treat continuous types of tasks, such as compensatory or pursuit tracking. Again, this goal has yet to be attained.

2.4 Constructs and Data

While some areas of model refinement and enrichment are being retarded because of a need to develop suitable techniques, other problem areas suffer from a lack of suitable constructs or data. An example of the construct problem is treating continuous time variables in discrete event simulations, and an example of the data problem is describing the impact changes in information uncertainties have on performance.

2.4.1 Constructs

The expenditure of consumables such as fuel depend on the geometry of the flight profile, which is a function of operator actions and vehicle dynamics, but the status of such consumables has a bearing on time stress imposed on an operator/pilot and should be expected to affect his performance—for example, forcing certain decisions to be made. Also, there are a number of tasks which are only awkwardly handled on a discrete basis. Compensatory and pursuit tracking tasks can be represented as a series of discrete tasks, but the manner in which this is currently being accomplished is highly artificial and could stand substantial improvement. What appears to be necessary is a set of constructs for connecting the discrete task workload paradigms with the servo theoretic representations of the dynamics of the human operator. For example, in some applications, one is interested in how much time is required to achieve a specified degree of control, like "the cursors are aligned ±X for at least a period Y". Alternately, one might need to determine whether the cursors were within a prescribed tolerance at time Z, ±X, when the initial conditions were Y and W amount of time exists between "now" and time Z. The first case might be exemplified as target acquisition, the second as target tracking until time of weapons launch.

2.4.2 Data

The data difficulty became apparent in exploring what literature might be applicable to estimating how performance would degrade with marginally operable equipment. It was proposed that a marginally operable instrument could be depicted as an alteration in the information content or uncertainty of the display output. Assuming this uncertainty could be defined, it is still necessary to estimate how changes in uncertainty affect performance time and error. Very little empirical data were found that addressed this issue, despite numerous information theoretic oriented studies of human behavior. Much work remains to be done.

3. RESULTS

The purposes of simulation can vary markedly with the objectives of different users. One of the more instrumental uses of modeling is to assess the impact of design changes or evaluate design alternatives. It is desirable to have a model which in some sense provides details which help explain the results obtained. From an analysis of the available output data, one hopes to at least isolate areas of potential design weaknesses. An attempt will be made to illustrate how this might be done using the modified Siegel-Wolf model.

3.1 Principal Output Data

The Siegel-Wolf model views the performance of the man-machine system as a network of activities which consume time and have varied probabilities of success. The primary issue addressed is how long does it take to get to the last task of the mission and will this task be accomplished within a prescribed time frame. Because of this time orientation, the model is applicable to that limited class of systems and missions where a physical end point is operationally definable relative to some initial point in time--for example, time to weapons release from time of order to "scramble". The model attempts to represent the dynamics of the human as a task manager who can adjust his sequence of activities and adapt his speed and accuracy of performance in accordance with the time stress imposed by any discrepancy between time allocated and workload assigned. A more complete description of the model, its constructs, and various applications can be found in Siegel and Wolf (1969).

For a given set of run conditions, the model is iterated a specified number of times. The time used by the simulated operator is accumulated and a count is kept of the number of iterations which required (i.e., used) less time than allotted (and therefore "succeeded") versus those which required (i.e., used) more time than had been allotted (and therefore "failed"). A record is also kept of the number of times a task was failed, how often nonessential tasks were skipped due to time stress, how much time was consumed by failed tasks, the amount of time lost in waiting for one's partner or for other causes, and the number of times a task ended up being the last task completed before the allotted time ran out. Examples of these outputs are shown in figures 1 through 4.

The time used and percent success outputs are typically the first data of interest, but when run conditions exhibit some "surprising" increase or decrease in percent success, the other output data become useful in trying to explain possible causes for the observed results.

3.2 Use of These Output Data

Knowing when a task was completed, one can plot the completion time experienced under the run conditions with the a priori expectations for the task sequence. Figure 5 shows a plot of this sort. Each of the curves shows a different set of run conditions--in this case, exposure to increasing levels of supralethal doses of ionizing radiation.

The pattern of these curves is of particular interest. A downward (negative) slope is desired, since it indicates work is progressing "ahead of schedule"; but an upward (positive) slope indicates a block or series of tasks where the operator is losing ground, so to speak; i.e., he is slipping behind his schedule for task completion.

With this information, one can examine areas of concern. Two are of interest in figure 5. The block of tasks from 150 to 180 consistently leads to schedule slippage under all conditions, but it is only at the more severe exposure levels that mission success is grossly compromised. The block of tasks between 30 and 60, though, tend to aggravate this condition and appear differentially sensitive to the increasing exposure levels.

At this point, one can examine the block diagram of the mission, relevant parts of which are presented in figure 6. From the diagram, one gets a general impression of the nature of the tasks and can begin to postulate the kinds of operations which might be implicated as the contributive factors to an undesirable situation. As shown in figure 6, the climb phase (tasks 30-60) consists of a number of checks and some instrument adjustments. The lead collision attack phase (tasks 150-180 in figure 5 correspond to tasks 181-206 in figure 6) is the other area of concern.

At this point, one can examine in detail the input data associated with these tasks, as shown in Tables IV and V. Proceding left to right, there is a description of the task, the operator and task numbers, identifiers E and N denoting task type and task essentiality, the task branching directions upon success or failure of this task, the average and standard deviation of task times, and the estimated probability that this task will succeed.

3.3 Detailed Diagnostic Examination

One of the first steps in examining these data is to look for tasks with low probabilities of success, such as tasks 43, 44, and 45. In each of these cases, the given task is simply repeated. Tasks 57 and 58 present a different situation, however, since if task 58 is performed, the operator is obliged to loop back to task 55. The low probability associated with task 58 is not a controlling mechanism in this case but rather serves only as a means of counting task successes versus failures. The controlling mechanism here is task 57. If the airspeed and vertical velocities are not what they should be, which nominally is expected to happen only twice in a hundred occasions, then corrective action is necessitated and the series of checks must be made to assure the correction sufficed. Any serious degradation of task 57 acts as a forced repetition of four tasks.

The implications then are that demands placed on the pilot to make manual updates can be tolerable under nominal conditions but may degrade mission success under exposure to certain levels of ionizing radiation. In this instance, one might propose automated updates to circumvent this potential difficulty. As for the airspeed altitude problem, there appear to be two factors to consider. One certainly is automation—simply assuring autopilot functions are adequate and do not degrade such that manual intervention is required. The other is to examine the vehicle dynamics to discover how sensitive the vehicle's responses are to operator inputs and how well (in what time frame and for how long) the vehicle will maintain a steady state condition once stabilization is achieved.

Examining tasks 181-206, one finds low probabilities associated with tasks 187, 190, 193 and 195, in particular, as shown in Table VI. Each of these is a form of "continuous" task, beginning with task 187 which consists of monitoring the radar scope for an assigned target; task 190 is a manual tracking task using a thumbwheel; task 193 requires a lever positioning, and again is a tracking task; and task 195 requires using the same lever in a second axis, which is also a tracking task. At this point, it is apparent that the problem here is a classic example of the control/display interface design issue which has long been of fundamental interest in human factors engineering. One would thus recommend a thorough examination of control/display ratios, display brightness and contrast, stimulus-response compatibility, response-response compatibility, etc.

3.4 Exploring the Value of Corrective Action

By now it should be apparent that the model is not an end point in the design process but instead may be the beginning of detailed design analyses. However, it does serve as an aid in focusing attention on potential problem areas and in evaluating the impact a change may have not just on man's performance but on system performance.

For example, automatic target lock-on could eliminate the need for perhaps tasks 193 and 195 and maybe alter the characteristics or needs for other steps in this procedure. One might be interested in altering the values of these tasks (or eliminating them) to see what impact such changes might have on the output results. If these changes demonstrate that the task sequence is more often completed "on time", one then has some basis for justifying the resource expenditure required for the design analyses which will provide this revised capability in the system.

Alternately, one could postulate other kinds of alternatives and use the model as a means of evaluating their relative effectiveness so a trade-off can be made between them. Rarely does one get a large change in effectiveness for a small change in cost, so such modeling analyses could often be valuable methods for obtaining data relevant to design decisions.

4. DISCUSSION

The Siegel-Wolf model appears to be useful in a wide variety of applications, has been used by several different firms, and recently was modified to treat nuclear survivability/vulnerability. Attempts are already underway to extend and refine the model and integrate it within the framework of available systems simulation routines. This is expected to result in a general man-machine modelling capability that can be used throughout the design cycle.

The model does not offer "automated design" but is potentially useful in examining areas where automation could prove beneficial in manned systems, particularly under adverse attack environments. While the model currently focuses on discrete event simulation of a time-based network, development plans include the treatment of continuous time variables.

Interest has also been expressed in expanding the model to treat nonnuclear vulnerabilities and the impact of those nonlethal multiple stresses common to the flight regime: noise, vibration, heat, fatigue, etc. Hopefully, these can all find their way into the model so that systems performance capability can be faithfully represented as a composite of both human and hardware factors.

5. ACKNOWLEDGEMENT

The research reported in this paper was sponsored by the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio. This paper has been identified by the Aerospace Medical Research Laboratory as AMRL-TR-72-49. Further reproduction is authorized to satisfy needs of the U.S. Government.

6. REFERENCES

Chubb, Gerald P., April 1971, "Using Monte Carlo Simulation to Assess the Impact Radiation Induced Performance Degradation Can Have on Mission Success", Symposium on Computer Simulation as Related to Manpower and Personnel Planning, Naval Personnel Research and Development Laboratory, pp. 237-263.

Kiviat, P. J., Villanueva, R., and Markowitz, H. M., 1968, The Simscript II Programming Language, Prentice-Hall, Inc., Englewood Cliffs.

Kochhar, D., and Wills, B., 8-10 December 1971, "Simulation of a Two-Man Interaction System", 1971 Winter Simulation Conference, Fifth Conference on the Applications of Simulation, held at the Waldorf-Astoria, New York, pp. 56-62.

Pritsker, Alan B., April 1966, GERT: Graphical Evaluation and Review Technique, Rand Corporation, RM-4973-NASA, Santa Monica, California.

Pritsker, Alan B., and Kiviat, P. J., 1969, Simulation with GASP II, A FORTRAN Based Simulation Language, Prentice-Hall, Inc.

Pritsker, Alan B., December 1971, Precedence GERT, Research Memorandum No. 71-14, Center for Large Scale Systems, Purdue University, Lafayette, Indiana.

Siegel, Arthur I., and Wolf, J. Jay, 1969, Man-Machine Simulation Models, John Wiley and Sons, Inc., New York.

		MAN	
		UNDEGRADED	DEGRADED
HARDWARE	UNDEGRADED	NOMINAL, GENUINE, PEACETIME OPERATION.	HARDENED SYSTEM. MAN LEFT UNPROTECTED OR INADVERTENTLY EXPOSED.
	DEGRADED	MAN PROTECTED UNTIL ASSIGNED TO SYSTEM/ MISSION; SYSTEM PARTIALLY PROTECTED.	BALANCED HARDENING, WHERE EQUIPMENT DEGRADES AT LEVELS WHERE MAN ALSO SUFFERS DEGRADATION.

Table I. Example of Different Mission/Threat Contexts That Can Be Treated.

The Implications of Detectability Versus Severity of System Malfunction

	Actual System Status		
Deficiency Symptoms	Wholly Operable	Marginally Operable	Completely Inoperable
Immediately detectable and permanent.	False alarm should be ignored; may not be.	Must be evaluated and acted upon as appropriate.	Contingencies must be evaluated and action taken accordingly.
Immediately detectable but transient.	False alarm and should be ignored; if not, true nature will eventually manifest and be detectable.	If evaluated immediately, will result in lost time if symptom goes away; implies recovery from inoperability.	Any premature initial reaction loses time and imposes an additional evaluation workload on pilot.
Eventually detectable (with use) and permanent.	This false alarm may be more disruptive if stress is building up, especially if above threshold.	Symptom may prove worse than problem if continued use is tried; stress at time of first use may stiffle discovery of true status versus indicated status.	Impact depends on howlong* true status remains undetected; can lead to late abort unnecessarily complicating recovery.
Eventually detectable (with use) and transient.	Depending on when use was attempted, condition might not be encountered; if encountered, will be minimally disruptive.	May he of no consequence if recovery occurs before use; disruptive if recovery occurs during the evaluation following attempted use.	If encountered, could lead to premature abort; if no abort, stress may still complicate the response.

* Inappropriate "false confidence" can provoke greater stress when contingencies are recognized. Appears to be the worst case condition.

Table II. The Implications of Symptom Indication Versus Malfunction Severity.

	CONTROLS	DISPLAYS
CONTINUOUS	CHANGE IN PLANT DYNAMICS, HANDLING QUALITY, STABILITY, FORCE EXERTION REQUIREMENTS, ETC.	INTERMITTENCIES, BLURRING, CHANGE IN CONTRAST, LOSS OF SYMBOLOGY, FALSE SYMBOLOGY OR LOWER SIGNAL TO NOISE RATIO, SCALE CHANGES.
DISCRETE	LOCKED OUT/IN. SELECTED POSITIONS INOPERATIVE. UNANTICIPATED CONSEQUENCES ENSUE.	PRESENT NO INFORMATION. --ERRATIC, RANDOM --CONSTANT BIAS --REMAIN ON/OFF AT "WRONG" TIMES PROVIDE FALSE INFORMATION.

Table III. Some Manifestations of Control or Display Operating Deficiencies.

Read EGT.	1	38	N	39	38	1.26	0.74	0.99
Read ADI.	1	39	N	40	39	1.26	0.74	0.99
Read AMI and AVVI.	1	40	N	41	40	1.34	0.48	0.98
Check standby ADI.	1	41	N	42	41	1.94	0.88	0.98
Check standby altimeter, and airspeed	1	42	N	43	42	3.24	1.08	0.96
Reset altimeter	1	43		44	43	8.60	3.00	0.50
Read TSD.	1	44		45	44	1.26	0.74	0.40
Read HSI.	1	45		46	45	0.66	0.34	0.40

Table IV. Selected Input Data for Tasks Between 31 and 45.

Release mike button	1	53	N	54	53	1.10	0.76	0.99
Assume manual control	1	54		55	54	4.20	1.02	0.99
Read TSD.	1	55		56	55	1.26	0.74	0.99
Read HSI.	1	56		57	56	0.66	0.34	0.99
Read AVVI and AMI.	1	57		59	58	1.34	0.48	0.98
Correct course to reflect instructions	1	58		55	55	3.80	0.48	0.50
Select auto flight mode	1	59		60	59	8.60	3.00	0.95
Release manual control	1	60		61	60	1.10	0.76	0.99

Table V. Selected Input Data for Tasks Between 46 and 61.

Read radar	1	187	188	187	1.20	0.40	0.50
Locate target on scope	1	188	189	188	1.20	0.40	0.90
Press action switch to super search	1	189	190	189	1.10	0.76	0.99
Adjust ANT ELEV for maximum target return	1	190	191	190	3.80	0.48	0.40
Is target within range?	1	191	192	191	1.20	0.40	0.90
Press action switch to manual radar search	1	192	193	192	1.10	0.76	0.99
Spotlight target	1	193	194	193	3.80	0.48	0.30
Set action switch to 1st detent	1	194	195	194	1.10	0.76	0.99
Adjust range gate	1	195	196	195	3.80	0.48	0.40

Table VI. Selected Input Data for Tasks Between 182 and 196.

SUMMARY OUTPUT OF ITERATION 5 RUN 1 TRIAL 1
UNDE RUN TOTAL TIME USED 711.89

OPR NO	THRES HOLD	SPEED	AVAIL	USED	DIFF	WAIT
			----- T I M E S -----			
1	2.30	0.90	750.00	516.93	-233.1	0.0
2	2.30	1.00	750.00	711.89	-38.1	365.0

****

PEAK TASK	STRESS VALUE	FINAL STRESS	LAST COHES	CYCLIC WAIT
1	1.17	1.00	0.0	0.0
1	1.00	1.00	0.0	0.0

****

Figure 1. Iteration Summary, Showing Time Used, Peak and Final Time Stress, and Time Spent Waiting.

SUMMARY OUTPUT OF RUN 1
NUMBER OF ITERATIONS 20
NUMBER OF SUCCESSES 20
PER CENT SUCCESSES 100.0 FOR TIME DURATION 750.00

OPR NO	THRES HOLD	SPEED	AVAIL	USED	DIFF	WAIT
			------ AVERAGE TIMES ------			
1	2.30	0.90	750.00	445.26	-304.74	7.57
2	2.30	1.00	750.00	521.08	-228.92	163.93

PEAK STRESS	FINAL STRESS	CYCLIC WAIT
1.20	1.00	0.0
1.00	1.00	-734.27

Figure 2. Run Summary, Showing the Number and Percent of "Successful" Iterations, Average Time Used, Average Peak and Final Time Stress.

FREQUENCY DATA

TASK NO	LAST TASK COMPLETED OP 1	OP 2	TASK FAILED OP 1	OP 2	TASK IGNORED OP 1	NE	OP 2	NE
1	0	0	3	1	0		0	
2	0	0	1	1	0		0	
3	0	0	0	2	1	N	1	N
4	0	0	0	1	0		0	
5	0	0	2	0	1	N	0	
6	0	0	0	0	0		0	
7	0	0	0	0	0		0	
8	0	0	0	0	0		0	
9	0	20	0	1	0		0	N
10	0	0	0	0	0		0	
11	0	0	0	0	0		0	
12	0	0	0	0	0		0	
13	0	0	9	0	0		0	
14	0	0	0	0	0		0	
15	20	0	2	0	0		0	
		TOTAL	17	6	2		1	
		AVERAGE	0.8	0.3	0.1		0.0	

Figure 3. Extended Run Summary, Showing Number of Times Each Task Was the Last One Completed Within the Allotted Mission Time, the Number of Times A Task Failed, and the Number of Times a Task Was Ignored.

SUMMARY OUTPUT OF RUN 1

TASK NO	FAILURE TIMES OP 1	OP 2	PEAK STRESS OP 1	OP 2	AVG TIME CMPLTD OP 1	OP 2
1	397.42	4.73	19	20	137.48	5.42
2	122.83	6.39	0	0	135.95	14.98
3	0.0	203.25	0	0	271.46	213.57
4	0.0	0.75	0	0	251.48	0.0
5	20.30	0.0	0	0	157.91	0.0
6	0.0	0.0	1	0	242.42	270.02
7	0.0	0.0	0	0	243.07	367.14
8	0.0	0.0	0	0	241.61	49.17
9	0.0	145.02	0	0	291.39	521.08
10	0.0	0.0	0	0	0.0	0.0
11	0.0	0.0	0	0	0.0	0.0
12	0.0	0.0	0	0	0.0	0.0
13	827.53	0.0	0	0	386.87	0.0
14	0.0	0.0	0	0	395.02	0.0
15	90.44	0.0	0	0	445.26	0.0
	1458.52	360.13	TOTAL			
	72.93	18.01	AVERAGE			

Figure 4. Additional Run Outputs, Time Spent on Failed Tasks, Times a Task Had the Peak Stress and Average Time for Task Completion.

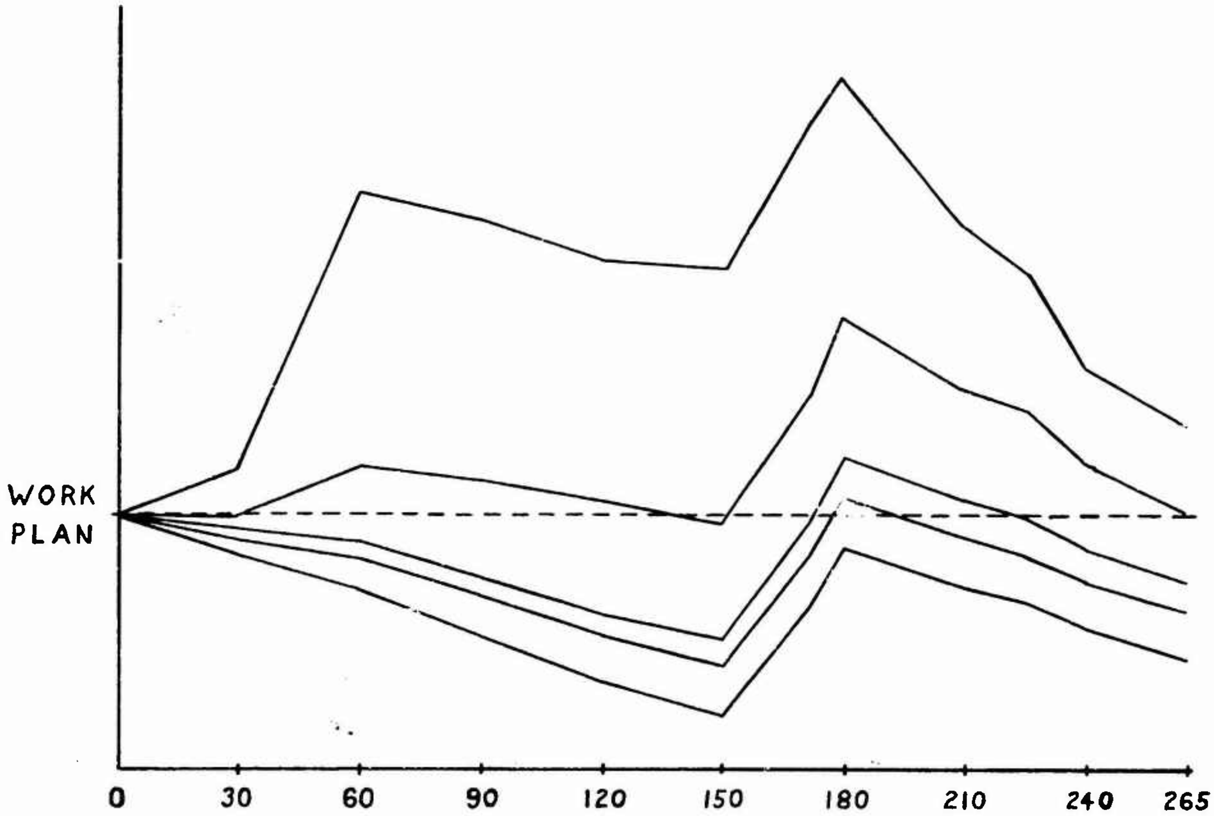


Figure 5. Simulated Task Completion Versus Scheduled or A Priori Expectation of Task Plan.

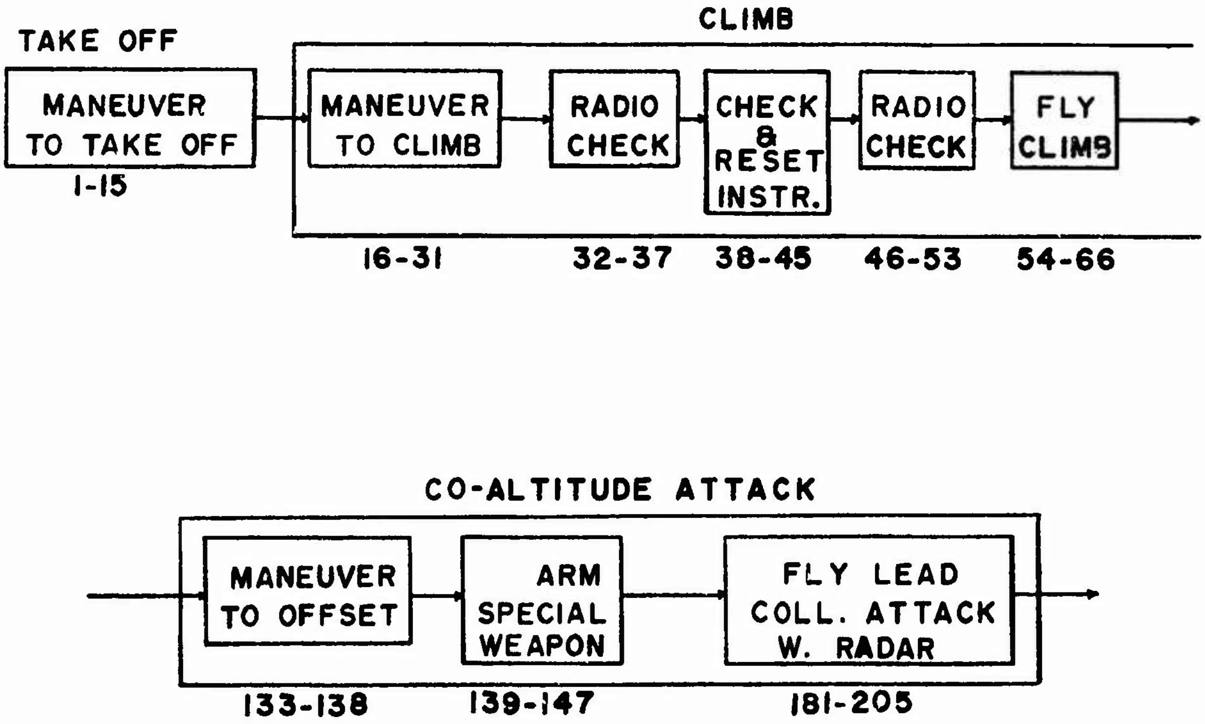


Figure 6. Portions of the Mission Profile Block Diagram.

DEVELOPMENTS IN AIRCRAFT DIGITAL SYSTEMS

R. Ruggles
Flight Controls Division

E. M. Scott
Flight Automation Research Laboratory

Marconi-Elliott Avionics Systems Ltd
Rochester, Kent, England.

SUMMARY

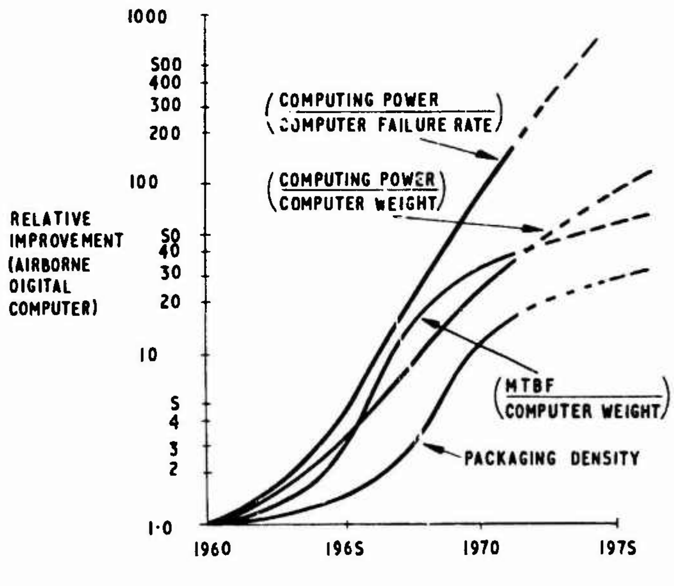
The effects of the relationship between user need and technological capability are considered for flight control as opposed to navigation and some physical characteristics of current digital autopilot are given. The functional division and integration of avionic subsystems are considered and it is concluded that integration in the form of loosely federated groups of related systems is preferred to the centralised computer complex in spite of its apparent conceptual simplicity.

The concept of task oriented computers is discussed and the main parameters of some existing examples are given. Some details of the architecture, software and hardware for this type of computer are given. An example of the application to automatic flight control with a requirement for a fail operative capability is given and the problem of dealing with tolerances between operating lanes is briefly discussed.

1. RECENT PROGRESS AND CURRENT CAPABILITY

Technological progress arises from the interaction of a developing capability by designers on the one hand and the evolution of specific user needs on the other. There is a tendency for the specific user needs to lag behind the technological capability because it is usually necessary to build up confidence in new technology before general user acceptance is achieved. How much this applies is affected greatly by the user's experience of existing systems and his degree of satisfaction with and confidence in them.

It is interesting to consider this interaction between capability and user need in relation to some avionic systems requirements and the availability of airborne digital computers.



NOTES:
1. COMPUTING POWER — OUTPUT BITS/SEC (PROCESSOR)
2. PACKAGING DENSITY — DISCRETE COMPONENTS AND EQUIVALENT PER LB
3. IMPROVEMENTS IN ANALOGUE COMPUTING HAVE BEEN SUBSTANTIALLY SIMILAR IN RELATION TO MTBF/LB AND PACKAGING DENSITY

Figure 1 Improvements in airborne digital computers

The last decade has seen a very rapid development in the digital computer for airborne use and Figure 1 gives some indication of this. This has caused a steady changeover from the primarily analogue subsystems in service in the early sixties to the extent that currently very few new analogue avionic subsystems are being contemplated. There are two very different aspects to this changeover. For instance, in the case of navigation systems the performance of the old analogue systems was severely resistricted by achievable computational accuracies so the advent of the airborne digital computer completely changed the computational capability. Since the analogue systems were complex electro mechanical devices which were being operated at the limits of their capability they had obviously little flexibility and severe maintenance and reliability problems. Thus in this case acceptability of a digital implementation was rapidly achieved.

In the case of flight control systems, where safety and integrity are major considerations and bandwidth rather than accuracy is important, many sensors are still analogue and actuation systems, which are a major interface, are likely to remain analogue for sometime, so that the advantages of a digital system are reduced by the additional interface requirements. As a result the advent of the airborne digital computer did not immediately offer any advantage to the flight control system user and it is only in the most recent applications where a digital autopilot has been shown to be competitive in cost, weight and reliability for the same performance as its analogue competitor. Figure 2 shows the approximate breakdown of some digital autopilot computers into processor (including store), interface and power supplies. The first is an R and D application in an existing in service airframe with its current interfacing systems all of which are analogue. The second is a current program and the interfaces are largely digital. A particular point to note is that although the rapid pace of development in component technology has enabled a big reduction in processor hardware to be made, this is not matched in the other areas. The third example is a projection of the second taking into account likely improvements in technology in the next few years.

APPLICATION	IN SERVICE AIRCRAFT (R&D)	NEW AIRCRAFT DESIGN	FUTURE AIRCRAFT DESIGN
SIZE	1 ATR	¾ ATR	½ ATR
PROCESSOR	25%	15%	17%
INTERFACE	55%	55%	50%
POWER SUPPLY	20%	30%	33%

Figure 2 Typical digital autopilot computers

2. FUNCTIONAL DIVISION AND INTEGRATION

The intrinsic time sharing nature of digital computers, displays and data transmission has made digital mechanisations of avionic systems competitive in cost weight and reliability with their analogue counterparts. The benefits of time sharing are best realised if the tasks are grouped together in a manner consistent with the overall system requirements. The continuing trend towards larger task groupings has been extrapolated in certain quarters to the conclusion that all the avionic systems will eventually be grouped together into sensor groups, centralised computing and time shared displays interconnected by digital data transmission highways. Such extrapolations do not take account of the constraints which have to be placed on the primary functions of any avionic system. Neither do they take account of the incompatibility and inefficiency which result when functions with dissimilar computing

and interface characteristics are grouped together. This centralised approach leads to systems designs of the form shown in Figure 3. The conceptual simplicity of such designs, which is their main attraction, belies their inherent complexity and inefficiency.

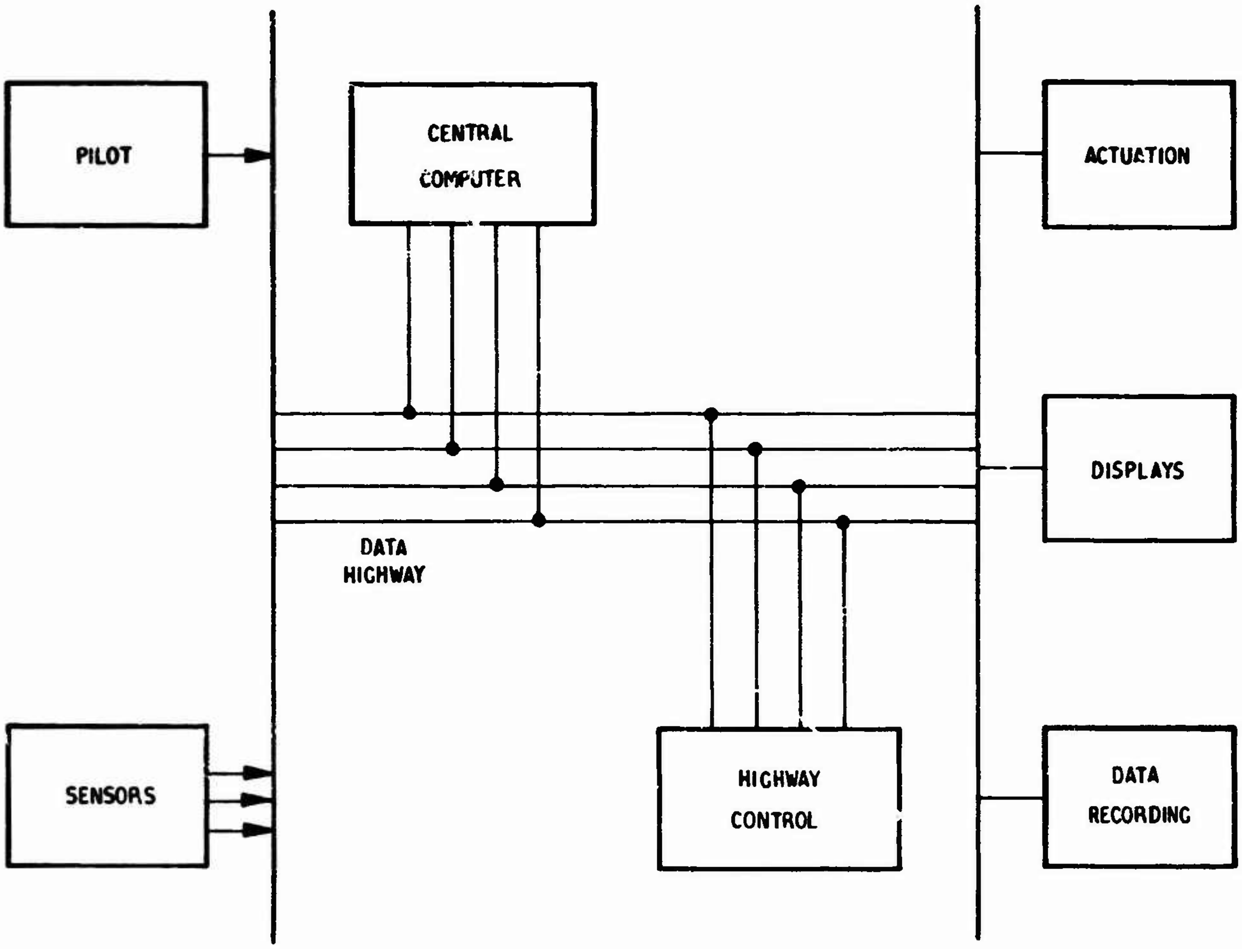


Figure 3 A centralised system

While such approaches to avionic systems design are unacceptable, it is equally unacceptable to permit avionic systems to continue to develop in their current piecemeal fashion. The functions of current and projected avionic systems must be rationalised into more optimum system arrangements so as to yield the following improvements:

- The increased availability of the aircraft as a total system.
- The provision of flexibility in the capability of the total system for a range of requirements and the facilitation of the incorporation of new requirements.
- The reduction of the ever increasing pilot and/or crew work loads.
- Reductions in the cost, weight and power requirements of the total system.

Systems integration must be defined as the attainment of the above objectives and should not be constrained to any preconceived system organisations. Any such systems integration must conform to the primary constraints of the individual functions being integrated. These are:

- Flight safety reliability
- Mission or dispatch reliability
- Maintenance reliability

The requirements for reversion operation and graceful degradation of the system performance must also be given due consideration in any integrated system, and will determine the extent to which the avionics can be regarded as a total system.

The primary functions of a typical short haul civil aircraft are shown in Figure 4. These are grouped into sets of differing integrity levels. Any integration of these functions must either conform to these boundaries or accept the inefficiencies associated with the integrated system having an integrity level equal to the highest integrity requirement of the constituent functions.

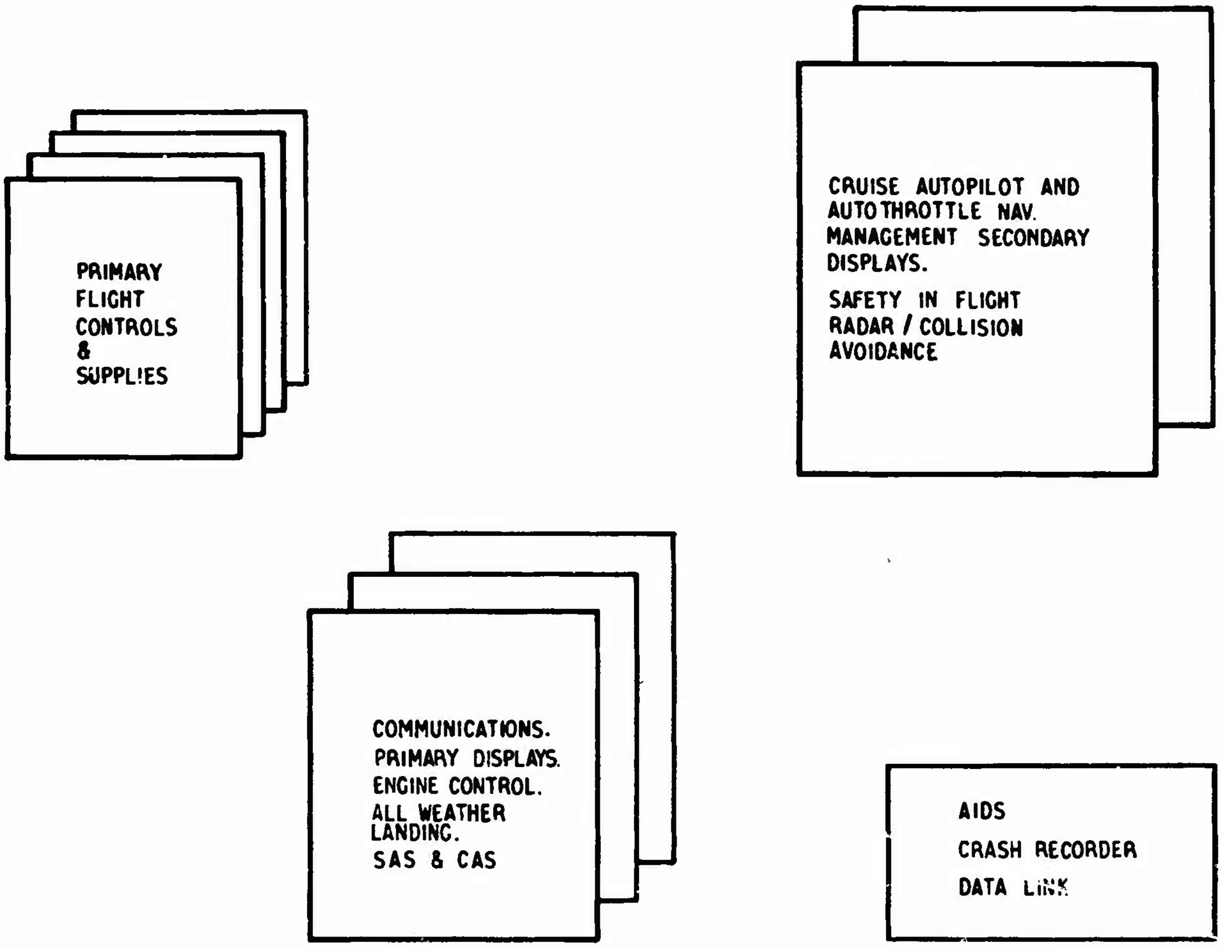


Figure 4 Primary functions of a typical short haul civil aircraft

In any avionic system the integrity and availability requirements, which demand high levels of redundancy, are in direct conflict with the maintenance reliability requirement, which demands minimum component content and this conflict increases with the level of systems integration. Full account has also to be taken of the impact of the system complexity on the maintenance strategy, the cost of spares holding and the ability of the system to detect and locate failures.

Particular consideration must also be given to the man/machine interface in order to minimise pilot and crew work loads. Increased automation and integration changes the pilots role from one of direct involvement in the operation and control of the individual avionic functions to one of overall system monitoring. This monitoring becomes increasingly difficult as the level of integration increases and the pilot becomes further removed from the constituent functions of the system. A compromise has to be reached between the increasing complexity of the systems and the diagnostic capabilities of the pilot. Reversionary operation must also be considered in the context of the man/machine interface. The pilot, having been concerned with monitoring for the majority of the flight must be able to intervene manually when necessary. Manual operation will also continue to be mandatory for take off, departure, approach and landing.

The compatibility of the constituent functions must also be considered in any systems integration. Incompatibility in any of the following areas will inevitably lead to inefficiencies in the integrated system:

- Interfacing and interconnections
- Iteration rate
- Data Transmission rate
- Accuracy and resolution
- Amount and type of computing
- Storage media and organisation

A further factor to be considered is the mixture of contributory disciplines involved in an integrated system. These can place constraints on the systems design and can lead to problems of project management and control.

If the centralised system shown in Figure 3 is considered in the light of the above discussion it can be seen that it fails to meet most of the above criteria. This can be illustrated from a consideration of the data highway and the centralised computing complex.

The data highway must be capable of transferring the longest digital word required by any of the constituent functions and be capable of operating at data rates compatible with the highest and lowest bandwidth of the individual functions. The highway must also have a higher integrity than that of the highest integrity function in the system.

The incompatibility of these parameters leads to inefficiencies in the systems design and increases hardware content well above that which is essential for systems operation. This inefficiency can only be justified if it is more than offset by the difference between the reduction in interconnections required in a non-integrated system and the additional hardware associated with the highway control. As the trend towards integration in the separate subsystems proceeds, inter subsystem interconnections will decrease and the use of a data highway will be even less justifiable than it is at present.

The failure modes of the data highway are also more complex than the star interconnections of a more conventional system. Difficulties also arise in the provision of adequate electrical isolation between the redundant lanes of the data highway and a high probability of secondary failures exists.

A block schematic of a typical centralised multiprocessor is shown in Figure 5. The use of multiprocessing is a recognition of the integrity requirements of the individual functions and of the lack of adequate computational speed in a single computer for the integrated tasks. The centralised computer must also be capable of simultaneously computing the highest accuracy and highest bandwidth functions in the system. The integrity of the central computer must also be higher than that of the highest integrity function. Such a centralised computer must also cater for the different types and amounts of storage required by each function.

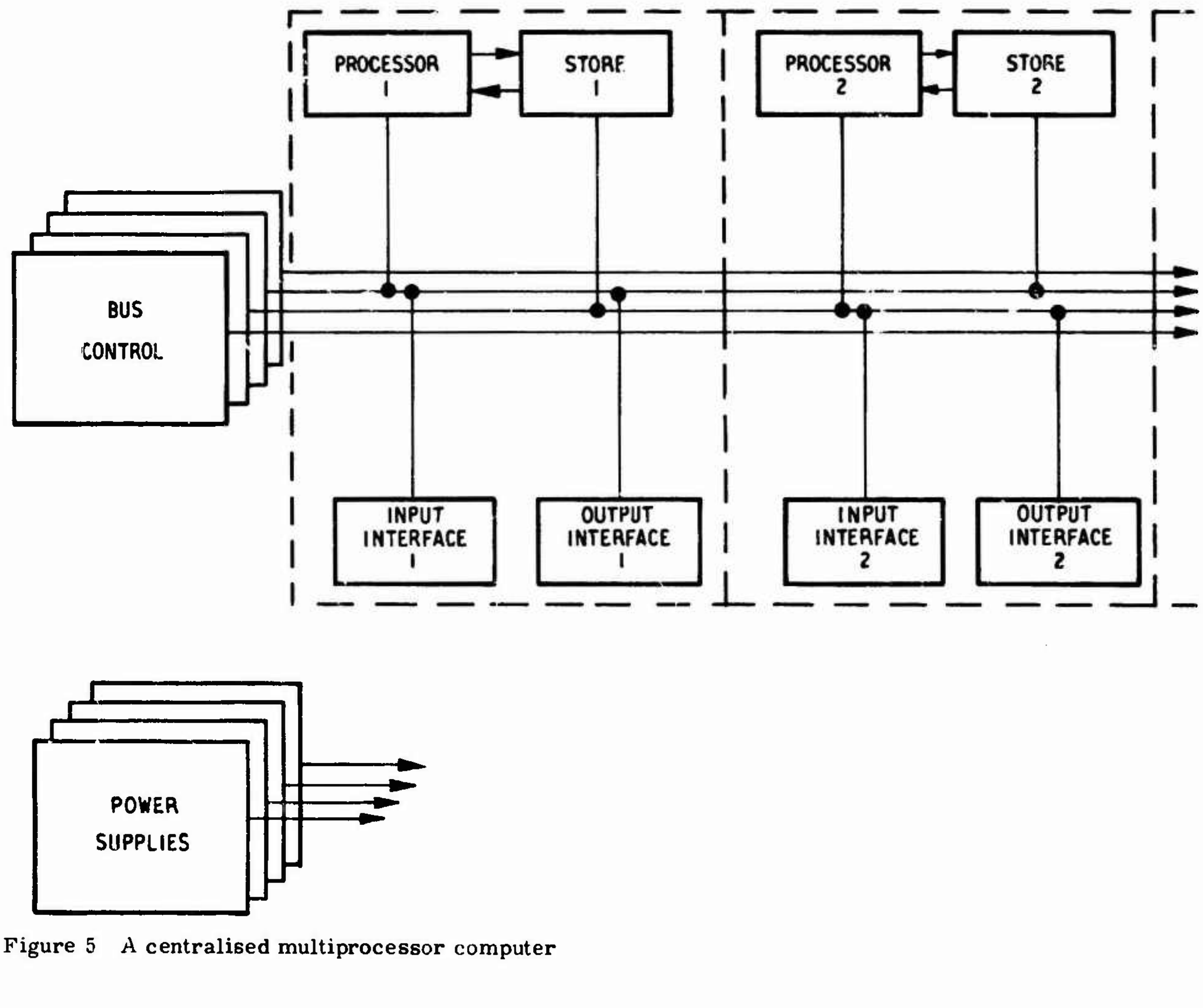


Figure 5 A centralised multiprocessor computer

The complexity of the central computer also makes failure modes and effects analyses a prohibitively expensive and prolonged exercise. The total software package for the diverse range of functions in the integrated system would also be prohibitively difficult to control, debug and modify, and would lead to extended development timescales.

There is little justification for the centralised integrated system because no advantages can be claimed which might outweigh the objections noted above. Providing care is taken to ensure that the piecemeal integration of individual systems develops in a rational manner which is consistent with the total system requirements, a loosely federated group of subsystems will provide a more acceptable solution to systems integration than the centralised multiprocessing schemes.

3. TASK ORIENTED COMPUTERS

The digital computer can only be competitive with its analogue counterpart if it is designed to meet the specific system requirement. General purpose computers, while being ideally suited to ground borne and certain avionic management tasks, have no place in on-line avionic control applications such as autopilot and engine control systems.

While recognising the requirement for large general purpose computers in some avionic applications, it is considered that optimum mechanisations of on-line control in avionics can be achieved only by a 'task oriented' approach to the processor design. (It should be noted that the term processor is used to refer to the combined central processor unit and store in order that it can be distinguished from the avionic use of the term 'computer' meaning a complete LRU; in the case of an autopilot computer for example).

The first generation of these task oriented processors was designed to minimise their hardware content. However, the rapid advances which have been made in component technology now enable the processors to be designed to improve the interface design and to optimise the performance of other components in the system without incurring a significant increase in the processor size.

The essential features of task oriented processors are as follows:

3.1 Architecture and Operational Mode

The register and store organisation is chosen to suit the system requirement, provision being made for error detection, multiple length arithmetic, overflow protection etc., as appropriate. The operation of the processor and the interface may be serial, parallel or byte serial.

3.2 Instruction Sets

The instruction sets are constrained to be subsets of a comprehensive master instruction set which has been evolved over years of on-line control experience. In this way, good assembling, simulating and diagnostic facilities can be provided without incurring a substantial software investment for each new application.

The instruction sets are optimised for the particular tasks and consideration is given to the inclusion of special functions not included in the master instruction set. Potential benefits accruing from the inclusions of these functions are traded-off against the cost of extending the master instruction set and hence modifying the standard software.

3.3 Software

Program lengths for these applications are typically one to four thousand words. The total program bit storage is of the same order as that required in the microprograms of current general purpose computers. The applications programs are assembled and debugged off-line on assemblers and simulators on a large ground based computer. Software compatibility has ensured that good diagnostic software and monitoring facilities are available to enable programs to be debugged in parallel with the commissioning of the hardware.

3.4 Storage Media and Organisation

The use of semiconductor storage media has enabled storage to be readily separated by function. This has permitted the use of different data and program wordlengths and the use of the most appropriate storage media in each area. The data wordlength is chosen to suit the accuracy and resolution required in the application, while the program wordlength is chosen according to the required processor facilities and to the required direct address field.

Program storage for on-line control applications is normally invariant and is implemented with read only memories for maximum program integrity. The scratchpad or workspace is implemented with read/write random access memory and constants are implemented with programmable read only memories. During development the storage of both program and constants is implemented with electrically alterable read only memory for ease of modification.

3.5 Input/Output

The input/output facilities of the processor are designed to suit the particular system requirement and to minimise the total hardware content of the processor and interface. The input/output may be via the accumulator, by autonomous data transfer to the data store, by direct data store access under program control, or any combinations of these, depending on the particular system requirement.

3.6 Component Technology

As the processors are designed for each specific application, changes in the component technology can be made best use of in each application. By comparison, general purpose computers are designed for implementation with obsolescent components and can only partially benefit from advances in the component technology.

3.7 Integrity

Due consideration is given in the design of the processors to the requirements for failure mode and effects analyses, which are essential for the certification of on-line control systems. Facilities are also provided for rapid initialisation of programs following power interrupts and other transient disturbances.

Figure 6 summarises a few of the main parameters of some of the task oriented processors which are currently in production for avionic applications. The diversity in design can be clearly seen from the parameters presented. A Marconi-Elliott Avionics general purpose computer is also included in the list of comparison purposes. The computer is currently in production for Navigational Attack applications.

TYPE	APPLICATION	DATA WORDLENGTH	ITERATION RATE	ADD TIME
MANAGEMENT	NAV ATTACK	18	50Hz max	8µS
DISPLAY	HUD	12	50 Hz	2µS
	HDD	12 / 20	50 Hz	250nS
CONTROL	MILITARY AUTOPILOT	12	20Hz	3µS
SENSOR	AIR DATA	24	16Hz	10µS

Figure 6 Some parameters of task oriented processors.

4. SUBSYSTEM ASPECTS

From the consideration of how functional division and integration should be applied in the overall system it seems clear that systems directly affecting flight safety such as automatic flight control, engine and intake controls should be looked at as a separate set and integration between these subsystems considered. Any such resulting integrated subsystem must obviously be capable of meeting all of the separate subsystem integrity requirements and must not introduce additional cross combinations of failures.

Since fully digital automatic flight control in terms of in-service experience is new and as yet unproven it is appropriate to give some detailed attention to the performance advantages that can be gained and the areas which will need to be carefully considered in achieving the necessary acceptability of their integrity. For instance a rigorous failure modes and effects analysis is a prerequisite for the acceptance of current flight control systems and a large background of hard won experience has been acquired by those involved in this field. Bringing digital flight control to the same level of confidence will be no easy task.

As an example consider a requirement for a single fail operative system. There are two types of candidate systems.

- Dual redundant, with both lanes independently monitored
- Triplex

The dual redundant system can be configured in various ways examples of which are shown in Figures 7 and 8. In each case the system is shown with one lane active and on-line and the other active but on standby. This form of output consolidation is generally preferred from redundancy management considerations where the next element of the system such as actuation may have a different level of redundancy.

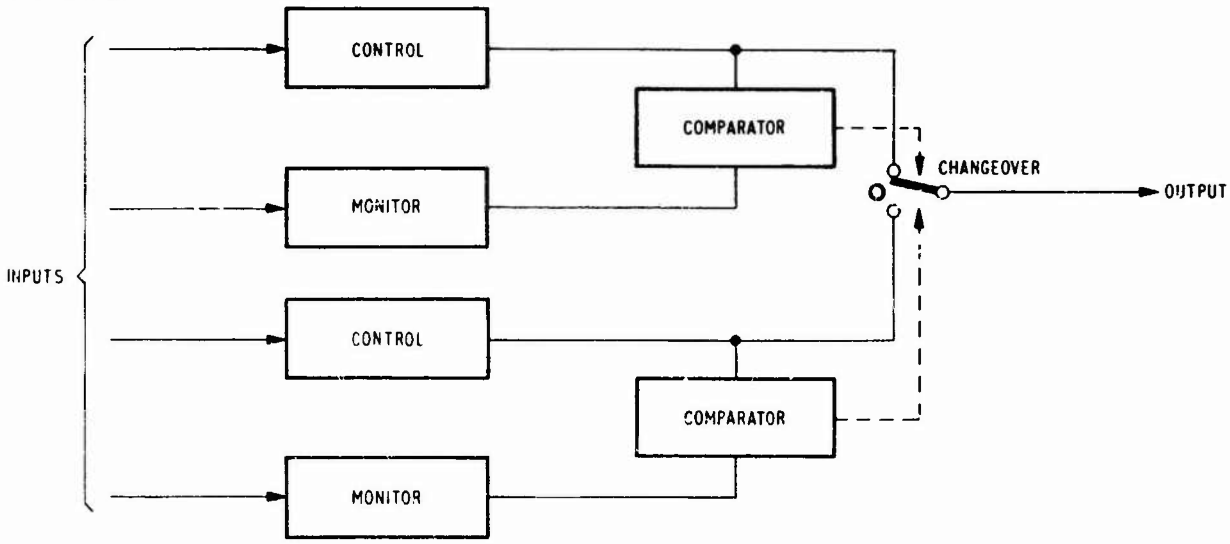


Figure 7 Duplicate monitored system

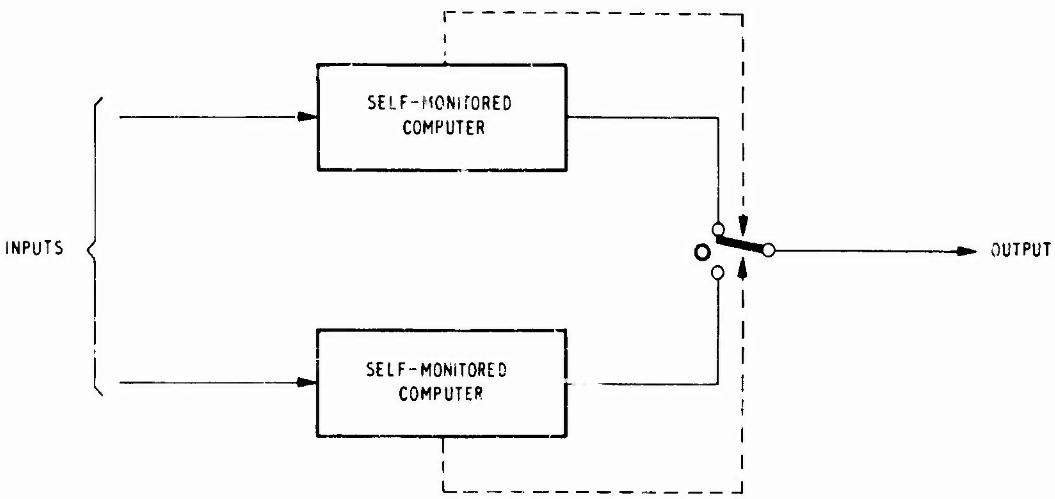


Figure 8 Duplicate self-monitored system

The system in Figure 7 has much to commend it on the score of integrity since the monitor computer and the control computer can be dissimilar in hardware and software but a price in total system hardware may be paid. The system in Figure 8 is at the present time an as yet unrealised ideal with digital systems of 100% self monitoring. In practice this system has fail safe cross monitoring between the two outputs to provide safety; failure identification and isolation is carried out by the self monitoring at less than 100% probability and the system is not truly single fail operative.

Figure 9 shows a triplex system with output voting and this system will be considered in more detail.

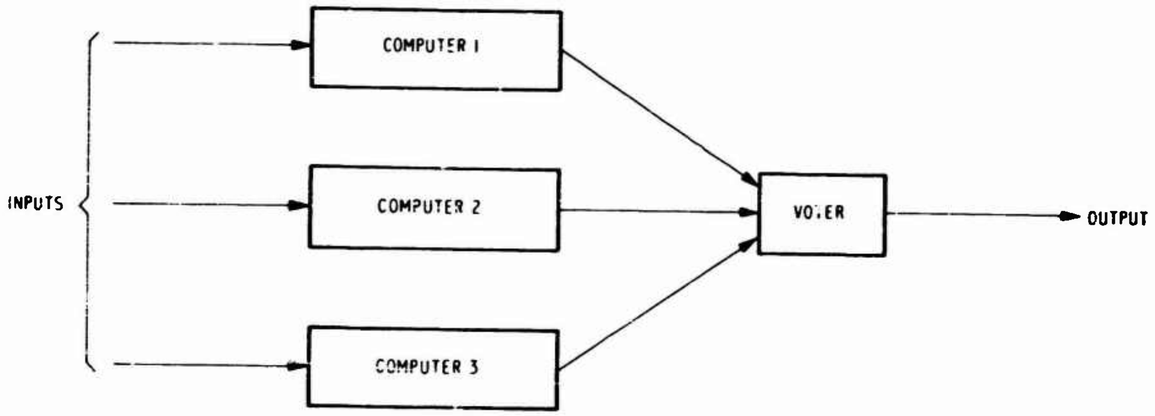


Figure 9 Triplex system

The main problem in a failure survival system of this kind arises from the existence of differences between the information in each lane of the system and in the case of flight control systems the output transient of the system when failures, real or apparent, occur. Meeting performance requirements and achieving safe operation after failures with realistic levels of inter lane tolerances is a major design parameter in this type of system.

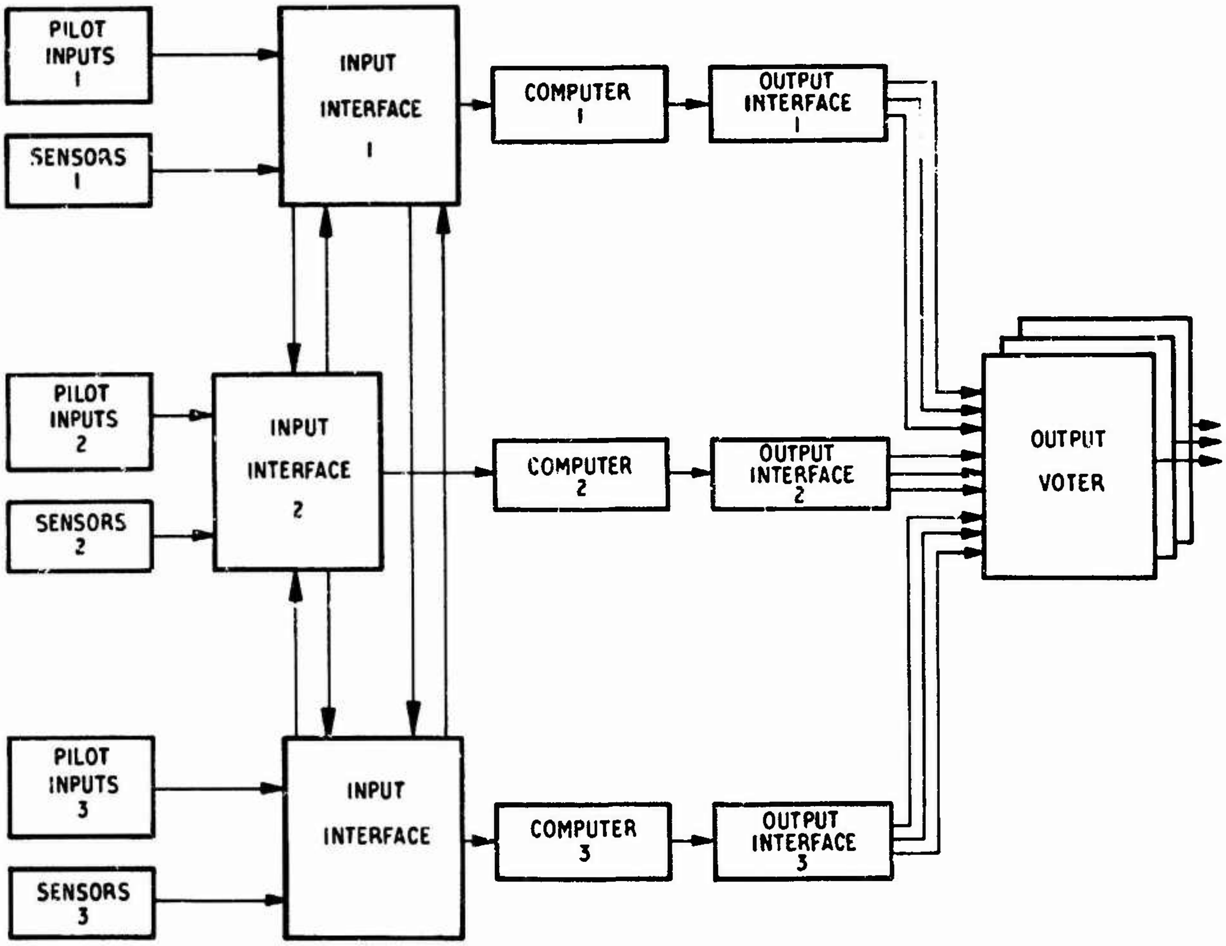


Figure 10 Simple triplex system

These inter lane tolerances arise in three areas, the sensors and other inputs to the system, the computing and the actuation. One of the benefits of the digital system is the accuracy to which the individual computations in each lane can be matched (the absolute accuracy requirements for the system can be met with 12 bits). However this does not help the other areas and although the actuation problem can be overcome if it is isolated from tolerances upstream the sensors are normally a severe problem requiring the incorporation of voters.

Figure 10 shows a representative system including the cross connections between the three lanes required for voting on the sensor inputs. For a system of any complexity with several sensor inputs requiring voting these interconnections are a severe aircraft wiring penalty and very much degrade the integrity and vulnerability of the system. In a relatively simple system and in cases where the forward gains are not high the penalty may be acceptable but for high gain stability augmentation and autopilots where numerical integrations are normally required an alternative solution must be found.

Figure 11 shows a preferred system where the cross connections are made with optical links between the digital processors. In this system sensor 1 feeds only computer 1 and similarly sensors 2 and 3 feed only computers 2 and 3 respectively. Any necessary A-D or D-D conversion is carried out and then the signals are cross fed to the other two computers. The optical links offer big advantages in integrity and vulnerability. Only three cables are required for all cross connections and each is just over 0.1 inch diameter. There are no earthing problems associated with screens and no problems of EMI generation or reception and the propagation of electrical faults or fires is prevented.

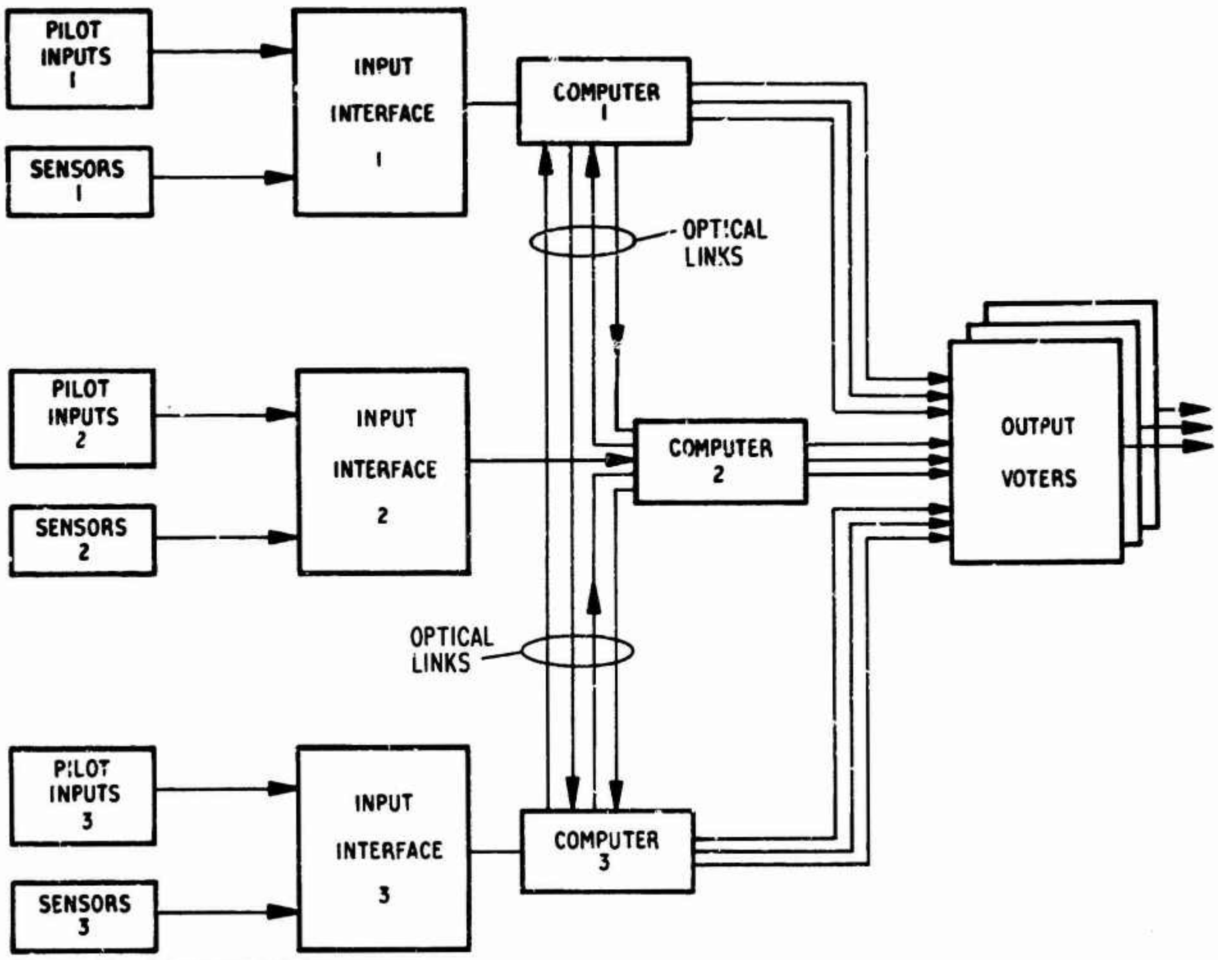


Figure 11 Advanced triplex system

5. FUTURE TRENDS

The rate of integration of avionic systems will be determined by the acquisition of experience with them and the establishment of user confidence in them, because current technology is well able to cater for increased integration in a competitive manner. Nevertheless, the trends in the technological development will further aid systems integration and every advantage should be taken of these developments to further improve flight safety and to make the integrity of systems more visible.

The smallest digital processors which are currently employed in avionic systems comprise 50 to 70 standard MSI and LSI components (including storage). This number will undoubtedly decrease with the advent of the single chip central processor and the size of the total processor will be reduced to limiting proportions. The computing power of these processors will be best utilised by distributing them throughout the system to supplement the operation of those parts of the system which have not been subjected to the same rate of development as the semiconductor components, such as sensors and actuators. These processors will also supplant the more dedicated logic associated with interfacing to sensors, displays, actuators and controllers, and data transmission control. This will lead to systems which can be readily modified by re-programming the appropriate processor without either causing interactions with other parts of the system or degrading the system integrity.

Time shared display surfaces are already available in the form of raster and stroke written CRTs, which are being applied in integrated systems. However, the developments in solid state displays indicate that they will soon supplant the CRT because of their greater reliability and smaller volume.

Optical data transmission can be used to increase the integrity and the availability of integrated systems. It is already being employed in avionic applications and, as the losses in the fibre materials are further reduced, so the distances over which optical data transmission is employed can be increased and the transmitting and receiving components can be reduced. The use of optical data transmission also frees the system designer from many of the constraints which were placed on him in analogue systems.

HUMAN FACTORS IN LOW WEATHER OPERATION OF TRANSPORT AIRCRAFT

J.W. WILSON
CHIEF OPERATIONS ENGINEER
HAWKER SIDDELEY AVIATION LIMITED
HATFIELD - HERTFORDSHIRE
ENGLAND

SUMMARY

Practical experience gained during the manufacturer's flight development testing and airline in-service operation of a failure-survival Category 3 automatic landing system is reviewed for indications of the extent to which human factors have affected the design of the system and the techniques used by the airline in order to reach the very high safety levels that are necessary.

It is concluded that the stressful environment of low weather operation further restricts the capacity of the crew to respond reliably to instrumental information or visual warnings and that great care must be taken to simplify tasks allocated to each crew member to ensure that the work load remains at a level low enough to provide spare capacity to cope with unusual variations in human and system performance.

The important factors influencing the complexity of the task are:-

1. Provision of adequate monitoring devices located in the optimum area of each crew member's primary visual scan, to enable the pilot to keep ahead of the operation of the automatic control systems.

2. Application of identical procedures for use in Category 1, 2 or 3 weather which lead to familiarity with and confidence in the integrity and performance of the system in a wide range of meteorological and visual conditions.

3. Design of the system and development of procedures such that the maximum number of manual and automatic functions that require action, checking or monitoring can be completed before the final stage of the approach to land.

4. The decision to land should be made as low as possible, compatible with a go-around performance which will not normally result in touchdown.

1. INTRODUCTION

In 1960 the author became involved with the Autoland programme for the (then de Havilland) Trident, which had the aim of completing by 1970 a clearance for British European Airways to operate in 'zero-zero'. Although in the years since Trident low weather flight testing began there has been a steady reduction in certificated limits from 300 ft DH/1,200 metres RVR in 1962 to the present 12 ft/270 metres, and in the near future to 12 ft/120 metres, routine airline operation below 60 metres RVR seems most unlikely to be achieved this century.

During the development programme more than 2,000 automatic landings were carried out in a wide variety of wind, turbulence, temperature and visibility conditions and in the winter months of 1964, 1966 and 1971 opportunities occurred to operate Tridents, crewed by company pilots, in runway visibilities at London Airport between 50 - 150 metres. In order to assess human and system capability and safety in the presence of faults or unusual conditions, more than 300 landings and take-offs were made in various Trident aircraft fitted with fog screens, graded and randomly adjusted during the ground roll to give a realistic RVR varying between 0 - 35 metres.

This paper is based on assessments made during the programme and the opinions expressed are largely subjective, gathered from participation in flight crew activities and discussions with company and airline pilots of many organisations. Many decisions were significantly affected by the need to achieve operational solutions which were acceptable in terms of cost, timescale and airline requirements: to this extent the conclusions are specific and should not be interpreted as if the outcome of a controlled human factors investigation.

British European Airways Tridents are certificated for operation in Category 3 conditions with a triplex level of automatic control in pitch and roll: degradation to a duplex level below the equipment decision height (300 feet) has been demonstrated to be safe to the 10^{-7} rules. The minimum decision height is 12 ft, down to which height automatic go-around without touching the ground can be initiated, allowing for engine failure occurring at any point and critical WAT conditions.

In-service recording has produced over 1,400 automatic landings to date and shows a mean touchdown lateral error of less than 1 ft from the centre line and a mean vertical speed of 1.70 ft/sec. BEA are operating 60+ Tridents, the majority equipped to this standard, and have some 800 pilots allocated to the Trident fleet most of whom are trained for Category 3 operations.

BEA have for many years used a precise method for operating their aircraft known as the Monitored Approach Technique, and it was essential that this drill was maintained as automatic landing was introduced and weather limits lowered. The technique had shown itself to be effective in reducing workload in poor weather conditions and is readily applicable to the three pilot crew complement. Summarising, the pilot in the right hand seat carries out the approach (manual or automatic) and passes control to the Captain only after the Captain has called 'I have control'. The co-pilot will execute a go-around (manual or automatic) if the Captain has not called 'I have control' before the decision height is reached: the third pilot crosschecks all actions made by the Captain and co-pilot and will call for a go-around if they both fail to take appropriate action before reaching the decision height. The go-around mode is automatically initiated by opening the throttles fully, there being no restriction on engine operation.

To achieve the lowest level of workload, approaches are normally automatic and followed by an automatic landing where ground limitations permit; autothrottle is used whenever available. As the autopilot is a failure-survival system in the approach, landing and go-around modes, it is usual for these modes to be used at suitable airfields in any weather conditions within the limitations of the airborne system.

2. THE HUMAN FACTORS INVOLVED

Human factors in low weather operation are, as in other flight regimes, primarily concerned with problems which arise from the existence of a high workload and the occurrence of dangerous errors. The critical nature of the landing and take-off, especially in low visibility, generates additional stress which can lead to a reduction in crew capacity and increase in the probability of making a mistake, which in its turn can have more serious consequences due to the small correction time available and the path accuracy required. The major factors can be identified as:-

(a) The task is too difficult for the crew to execute, probably due to lack of useable information, system complexity or equipment malfunction.
(b) Crew capacity, training or procedures are not adequate, probably due to illogical drills, incorrect allocation of tasks or lack of experience.
(c) Crew actions or systems performance are not properly monitored.
(d) Training and procedures are not designed to ensure involvement of crew members in monitoring and crosschecking.

Whilst keeping (b) and (d) very much in mind (the Monitored Approach Technique), the aircraft manufacturer is perhaps more concerned with resolving (a) and (c), which points to:

(e) Making the tasks simple and unhurried.
(f) Using internally-monitored and reliable automatic equipment for routine tasks.

If the tasks are to be simplified and rationalised they must also be positively allocated to the appropriate crew members. Of the qualms and critical decisions affecting the pilot in low weather landing conditions the more important are:

(g) Will the visual cues be recognised before reaching decision height ?
(h) Will the correct decisions be taken at and below the decision height ?
(i) Is the approach and landing flight path performance satisfactory ?
(j) Is the performance being monitored frequently enough ?
(k) Is the ILS trustworthy and free from interference ?
(l) Is a go-around and diversion likely ?
(m) Will the automatics remain engaged ?
(n) Will there be adequate visibility for the runway ground roll?

Of these, it is the first four that primarily concern the Captain and the first two that he cannot delegate. If he can be freed of frequent instrument monitoring the Captain is that much more able to concentrate on assessing the visual picture, while the co-pilot is committed to panel scanning to monitor the airborne and ground equipment performance and it follows that the latter should be better prepared to carry out a go-around if it is required: this is the essence of the Monitored Approach Technique.

The last two questions can only be resolved by confidence in the reliability and integrity of the airborne and ground systems, requiring:
(o) Encouraging in-service experience and routine use of the equipment over a considerable period.

3. EXPERIENCE AND APPLICATION

3.1. CREW CAPABILITY

A strong tendency to vision tunnelling or visual load shedding is very apparent in low weather landing - and it was often found that accurate monitoring of speed, height and altitude became so demanding for a crew member who was also carrying out a manual task such as speed or flight path control that he absorbed only the information contained in the centre of his primary visual target.

Speed error, beam error, failure flags, progress lights, all were sometimes outside his scan or passed unnoticed when the stress level was high, and this leads to some doubts as to the value of displaying rising runway symbols, expanded localiser or runway alignment commands in the already cluttered ADI. The Trident primary instruments have moved progressively towards stark simplicity; the ADI shows only attitude information and even the flight director is biassed from view when low weather automatic approach and landing is scheduled.

Although red has traditionally been used to indicate a need for immediate action, if the workload is high enough warning lights must also flash, and amber was found to be a more noticeable colour.

3.2. DECISION HEIGHT

Since the visual cues that the Captain needs to make his judgment improve in quality and reliability as the aircraft descends, the later (or lower) the decision is made the more likely it is to be correct. Very low minimum decision heights (12 ft for the Trident) allow a less hurried assessment and lead to lower stress when it is appreciated that the landing and go-around capability is compatible with a late decision. Another advantage is that at 12 ft. (wheel height) where the pilot's head height is 25 feet, the eye is close enough to the height of the transmissometer that RVR can be taken as equal to slant visual range.

It was noticed on a significant number of occasions that the Captain found it extremely difficult to reverse a correctly-taken initial decision, for example, due to the shortening visual segment that occurs with some types of fog structure. Such a change of initial intentions is much less likely in the few seconds between 12 feet and touchdown; it is obvious that the lowest possible decision height allows the least hurried and most accurate decision without committing the Captain to accepting a touchdown.

3.3. MONITORING

Removal of the Flight Director bars also prevents any instinctive attempt by the crew to monitor the triplex autopilot with a simplex director, and avoids the technical difficulties of certificating directed go-around from very low altitude when adequate performance can be achieved using monitored attitude information. Furthermore, in many systems complete disengagement of a redundant autopilot can be caused by multiple failures which are quite likely to affect director computation and unless it can be shown that this will 'never' happen, reversion to raw attitude displayed on either a valid ADI or standby horizon is the safer course.

3.4. WARNING

When utilising the monitored approach technique and automatic flight path control for approach and landing, the Captain must be advised of disengagement of the automatic system or significant loss of performance as rapidly as possible, particularly during the period immediately before reaching the decision height while he is searching for visual cues and barely aware of his instruments. As it was found that red lights, whatever their instrument panel position, did not reliably warn him, distinctive audio warning of autopilot disconnect, and amber flashing warnings of beam deviation were added; for both of these, as an additional safeguard, the co-pilot must initiate a go-around below the equipment decision height whether the Captain recognises the indication or not, up to the point that the Captain assumes control.

3.5. DISPLAYS

In non-visual conditions below 200 feet the Captain becomes increasingly anxious to know his whereabouts in the landing sequence; this leads to a tendency to glance down to the essential instruments (radio altimeter, mode indicator, attitude indicator) with consequential effect on workload and eye accommodation time.

Since visual cues appear and are first identified immediately above the windscreen vision cut-off line, it was found to be essential to locate progress information in this area, the particular arrangement being by a light cluster switched to go on at 130 feet, off at 65 feet, on at 12 feet and off at Autopilot disconnect at touchdown. These heights correspond to switching out the glide signal, start of flare, and runway alignment; with this arrangement it has been found to be practicable for the Captain to search for visual information and still be aware of aircraft height and autopilot mode switching in his peripheral head-up vision.

Attempts to increase the Captain's monitoring load any further (speed error, flare shape, beam error, director commands), resulted in a significant degradation of his performance in assessing the landing from external references particularly when the decision height was reduced to below the height at which the flare manoeuvre is started.

The combination of a vision line running through the bottom of the windshield and the pilot's inability to accept any but the simplest instrumental information is not compatible with low-weather use of the collimated type of head-up display, at least whilst its visual window is so much less than the aircraft's and its useability and integrity so unproven. Experience gained during the programme with several different displays also confirmed the pilot's difficulty in 'seeing' the collimated data even in fine weather conditions once the aircraft was below about 200 feet; no experience was gained in low weather conditions. It certainly seems that once the Captain is involved in assessing the landing picture there is very little spare capacity that can accept displayed data.

Early attempts to clear the paravisual displays (recessed in the glareshield) giving pitch and roll commands during landing met with problems similar to those described for head-up displays; also, the consequence of human lag in response to a pitch command was often a substantial over-correction, resulting in a very variable flare shape and lack of confidence in achieving even a reasonable touchdown in poor visual conditions. In addition, a significant number of pilots 'reversed' their initial pitch response and in the end pitch PVD was discarded. Fortunately, roll PVD was quite easily cleared as an en-route monitor of the autopilot roll channel and was subsequently found to be suitable for use by the pilot along the runway in zero visibility with either rudder control or nosewheel steering. Again the implication is that this steering task is simple enough and the display sufficiently large and compulsive that the pilot can perform reliably when under stress.

Recent simulation work has, however, confirmed that visual load shedding can easily be induced during the ground roll phase in visibilities below 100 metres if the workload is increased by, for example, requiring the pilot to carry out a programmed deceleration at the same time as steering from an offset position back to the centre of the runway - attention has to be paid to the detailed design of a deceleration director that is simple to use and need be followed only when the steering task is undemanding. Since the steering command should be seen paravisually, it is desirable to combine the deceleration director with steering guidance in the glareshield along with other information required during the landing run such as symmetrical wheel braking, runway distance-to-go and ground speed; in the Trident system the latter are located in the centre instrument panel and passed to the Captain by voice by the third crew member. This arrangement ensures that the Captain's primary effort is concentrated on steering control and relies on the third crew member to avoid overloading him by too frequent reporting of distance, speed and braking data. Its disadvantages are the use of an audio channel subject to both interference and saturation at critical moments, and the inevitable temptation on the part of the Captain to look down at the panel rather than out at the runway.

3.6. STATUS INFORMATION

Apart from peripheral progress information and ground roll steering commands, system information is restricted and can be disregarded once the aircraft is below the equipment decision height. This height is set at 300 ft. to ensure commonality of crew drills for any decision height between 0 - 250 ft. and allows time for the crew to confirm or correct the selected decision height as a function of the system status.

Status information is located in the glareshield in a position ahead of the third crew member which does not distract the Captain or First Officer. After much trial and error the display has been reduced to two magnetic indicators - one giving integrity status as a red, yellow or green indication, the other redundancy status as to whether the Triplex pitch and Triplex roll channels are all engaged or not.

3.7. THE INDEPENDENT LANDING MONITOR

Although the Trident system does not contain an ILM as such, thought was given to this subject some years ago and subsequent experience has not much modified our initial views.

For the ILM to be valid in non visual conditions its accuracy and integrity must be as good as, or better than, the autopilot's; this is a daunting thought when the cost of the automatic system is borne in mind.

There is evidence that the Captain is already loaded to a practical limit and would find it difficult to use another display particularly if it is an additional panel instrument or part of a head-up display. The co-pilot might be able to use it in instrumental form, but is it really necessary? In the pitch plane, the use made of the glideslope information below 100 feet is rapidly decreasing and multiple radio altimeters provide independent monitoring and control of the vertical flight path in relation to the ground. In the horizontal plane, the use of a very low decision height allows the Captain to assess the aircraft position in relation to the runway centre line in RVR's down to 100 metres without committing himself to touchdown - a landing picture which is a lot more convincing than that produced by an electronic ILM.

It seems that it would be more effective to put effort into making the radio guidance system less susceptible to significant faults than to develop independent devices that will be difficult to use and add further complication and cost to the total system. A few failures of the ILM in good weather conditions would soon result in it being discredited by the flight crew and instead of being used as a Category 3 monitor it might, like the Flight Director, be switched off whenever low weather approaches were made.

However, if it happens that ATC cannot guarantee a clear runway on which to land or take-off, or taxying separation cannot be safely and expeditiously provided, a form of ILM that allows the Captain to 'see' further ahead than his eyes allow would be of some value in reducing crew stress.

3.8. TAKE OFF

The Trident is certificated and British European Airways are approved for take-off in RVR conditions down to 90 metres, the limiting factor being the safe taxying of the aircraft from the terminal to the runway. Once on the runway the Captain is provided, by the PVD, with guidance that will ensure that he aligns the aircraft with the far-end localiser and continues to track down its centre line during take-off.

It was found that with this aid there was no bottom limit to the take-off RVR other than that set by the need to recognise a failure of the guidance system, and this was achieved with 30m RVR even when interspersed with sizeable blobs of fog giving zero visibility. It is obvious that an optimum future design should be duplex so that failures could be restricted to lack of guidance by shuttering the display.

In very low visibility the start of the take-off run can cause some difficulty since runway centre line lights may only be in view singly and the repetition rate too low to give a dynamic picture that the pilot can use to obtain tracking information. The same applies to runway texture given by tyre marks or slab joints; above about 60 knots, the lights and texture provide virtually continuous guidance and the task is relatively simple - not unlike a routine night take-off from a poorly equipped runway. No need was found for a rotation or climb out director, a Category 3 approved airfield being non-critical for a short haul aircraft in low visibility conditions.

Because of the texture and repetition aspects described above a significant contribution to the reduction of pilot stress can be made by appropriate painting of the runway, particularly a continuous centre line, and by insistence on a high standard of maintenance of all centre line and taxiway lights and markings.

4. FURTHER WORK

The development programmes and initial in-service operation have revealed how large are the gaps in our knowledge of the quantitative effects of those Human Factors that dominate in low weather landing and take-off, although a much better qualitative understanding of the limitations and requirements of the pilot has been obtained. It would be encouraging to think that with this knowledge it would be possible to proceed with more analytical simulator and flight research programmes, but the author is of the opinion that without inclusion of real-life stress the results would be of doubtful value or validity.

It seems clear that the effect of stress associated with a non visual landing significantly reduces the capacity of the pilot to respond in a time-shared or sequential fashion to several different indications, and therefore that concurrent or difficult tasks should be eliminated, either by design or technique. More work could be undertaken in representative transport aircraft to determine the optimum task sequence and which of them should be automated or suppressed during the landing manoeuvre.

Investigation is also needed into the extent to which the characteristics of the machine should be biassed to imitate the pilot's normal performance, rather than to achieve the highest obtainable performance.

In a contrary way the need exists to compensate for human lag by, for example, initiating a manoeuvre a little 'early' or using phase-advanced warnings.

Finally, any further work must reflect actual airline operations and piloting standards and should therefore be organised in close co-operation with airline pilots and those familiar with human factor features of current transport aircraft.

5. ACKNOWLEDGEMENT

The author wishes to thank the directors of Hawker Siddeley Aviation Limited for permission to present this paper. The opinions expressed are, however, those of the author and do not necessarily represent the views of the company.

AVIONIC SYSTEMS INTEGRATION USING DIGITAL COMPUTERS

Erwin C. Gangi
Navigation and Guidance Division
Airborne Computer Engineering Branch
Wright-Patterson AFB, Ohio

ABSTRACT

Present weapon systems use a multiplicity of signal formats and transmission techniques for information transfer within an integrated avionics system. Since the number of sensors and data processing devices vary as a function of the mission requirements of each weapon system, the interface equipment can become as complex as the equipment being integrated. One approach toward minimizing this complexity is the implementation of a serial digital data bus as the primary means of functionally communicating and interconnecting the various equipments.

If a system is logically partitioned to the data it supplies, requires or processes, then with a flexibly designed digital data bus and standard interfaces, it can easily be integrated through the computer software.

Modifications or redesign of the multiplexed data bus concept is a matter of reconfiguration of the building blocks, adding and deleting as required and then changing the software to reintegrate the new configuration, saving the costly rewiring and redesigning of the computer converter box. The computer, no longer an integral part with the converter, is now a separate line replaceable unit, not subject to obsolescence due to systems modifications.

I. INTRODUCTION

With today's rapid advancement in technology and systems architecture, it is difficult to define in detail what our future avionics systems will look like. Without stifling progress, it is desirable to guide future developments in such a way that when the hardware emerges it will fit into a workable pattern. This can be done by partitioning the avionics on a large scale and standardizing the interfaces. Total standardization is undesirable because technology changes are imminent and future generation avionics will assume new configurations and architectures. This problem, however, can be solved using a programmable digital computer and a flexible digital multiplexed data bus with standard interfaces.

The past decade has seen airborne computers grow up -- from the drum computers in the early sixties to the LSI versions today. Their capability has increased beyond expectations and their applications are many. Speed, capacity, and reliability have increased; weight volume and power have gone down. All this is due to the rapid advancement of integrated circuit and memory technologies. We are entering an era where more and more avionic processing capability is required and the single CPU computer has reached its practical limit.

II. CENTRAL vs. FEDERATED PROCESSING

Whenever there is more processing in an avionic system than a single computer can handle, then there is no choice but to partition the workload among several computers. There are logical partitions identifiable with distinct functions: Navigation, Weapon Delivery, Communication, Reconnaissance, and Electronic Warfare; others are not so evident: Control and Display, Signal Processing, Built-In-Test and Management functions.

If a single computer's hardware is fast enough that it can handle the real time workload of the total avionic system, when the partitioning must be done internally in the software. This results in functional software modules which are serviced by a real time executive and must be responsive to the system requirements and real time interrupts.

Should one computer prove insufficient, then logically more than one or a multiprocessor with multiple arithmetic sections is required. Now partitioning comes into the picture again. The computer workload should be divided into approximately equal parts for each CPU. For programming ease, it is desirable to keep interrelated software modules in the same computer and avoid splitting any up unless a definite gain can be shown by doing so. There appears to be two ways of integrating multiple computers into an avionic system.

If the workload is such that the processing can be easily identified with particular sensors such as in Navigation, Radar or Electronic Warfare, then it appears logical to assign federated computers, one for each sensor with digital communication between them. However, if in addition a central management computer is used, then these federated computers will become preprocessors with central control.

When a single computation center performs all the avionic processing in a simplex or multiprocessor mode, then the system is central. In this approach one or more computers stand side by side to carry the load with a central A/D converter doing the signal conditioning. Dual computers are frequently used for redundancy to give the system higher reliability through its graceful degradation feature and higher throughput without having any idle computers standing by.

This integration approach can be used on the whole aircraft avionics whatever it may entail or just on a particular subsystem such as the navigation. Figure 1 shows a set of avionic sensors that could be integrated through a central converter using the present state-of-the-art approach. Figure 2 shows these same sensors integrated through the multiplexed digital data bus.

III. DATA TRANSMISSION TECHNIQUES

Once the avionic system has been properly partitioned, the data transmission technique has to be investigated. Real time communication between sensors, computer, controls and displays is essential for an effective and accurate system. Distances between the sensor and the computer are often in excess of 50 feet. Electromagnetic interference, signal degradation over distance and high central conversion rate requirements are a major problem.

The majority of avionic sensors are designed to provide data in an analog form most convenient to the sensor manufacturer or peculiar to a one-time application. Therefore, the computer contractor, to properly communicate with the subsystem, must build a special purpose box to interface the computer with the other subsystems. Each analog or discrete parameter is routed separately from the subsystem to the central converter. The converter unit provides such things as signal terminations, conditioning, sampling, scaling and conversion to the proper digital format. The analog to digital and digital to analog converters must be extremely high speed because they are time shared among all the input/output signals. Often this converter box is twice as complex as the computer and highly special purpose. That is, if any signal format is changed, there is a definite impact on the converter hardware. Changes are often costly and major avionic modifications impossible without starting from scratch.

A digital avionic multiplex data bus has a number of definite advantages. Substantial wire savings can be realized by using the same data bus in a time sharing mode. Wire and connector savings show up as weight savings. Ease of changing sensors due to standard interfaces is another extremely desirable feature. The digital transmission technique used is less susceptible to EMI and it is easier to detect and correct errors. System configuration modifications can be performed by merely changing the computer software. Computers will no longer become obsolete because they will be truly general purpose (i.e., no special purpose converter hardware will be an integral part of it). In this digital bus concept, sometimes call MUX for short, the sensors that provide real time data to the computer or receive data from it will all be tied in parallel to a common data bus. Transmission will be digital under central computer control. This requires federated conversion and federated data storage in a sensor scratch pad memory. The sensor will generate the data, convert it and store it in its own scratch pad (remote data memory) at its own iteration rate. The central computer, under software control, will sample these data memories as required by the operational program and treat them as if they were an extension of the central computer's main memory. That is, the software may address a sensor's scratch pad memory to get data directly without having it go through temporary storage in the computer's main memory. (See Figure 3.) Data gathering is done by use of the computer's I/O controller that detects, through address look-ahead, that an external word is needed and fetches it.

IV. ADVANTAGES OF MULTIPLEXING

An interesting advantage to this technique is that the data transfer rates on the bus are much lower, requiring a much lower bandwidth transmission line. A one megahertz data transmission rate is usually sufficient. A data word is transferred only when needed rather than continuously as in the analog. Continuous transmission is required in the analog transmission technique because the sensor and converter hardware never know when and even if the software requires the newly generated information. Consequently, the computer memory must always be updated with the most current data. (See Figure 4.) In the multiplex data bus technique, the sensor scratch pad memories are treated as an extention of the computer's main memory and its use is under programmer control. Therefore, additions and deletions of subsystems are possible with minimal integration problems. Sensor modifications that change the sensor's analog input/output format will require the sensor contractor to modify his converter and it will be totally his responsibility to comply. These modifications will be performed by the subcontractor in the subsystem's own lower speed converter box requiring only software modifications within the central computer.

The integration of a digital computer into an avionic system is no easy task. A lot of study is required in the interface area to make a subsystem more electronically compatible. If the system is logically partitioned to the data it supplies, requires or processes, then with a flexibly designed digital data bus and standard interfaces it can easily be integrated through the computer software. Since technology will continue to progress and configurations will keep on changing, it is desirable to be independent of both. That is, let the data bus be two electrical wires connecting the sensor electronics to the computer electronics with no electronics of its own. Let us then specify a flexible format for the data, addressing, tests, etc. that flow on this line. We will specify what the bus looks like to the computer when it looks into the bare wire and similarly for the sensors. It is not desirable to standardize word lengths since due to the low bus transmission rates, communication can be serial and therefore the word length varied. To design a flexible standard that does not age, let's look at a somewhat similar system for analogy, the A/C power in your lab. It supplies the 110 volt, 400 hertz, with a generator (computer) on one end and its output is used by all subsystems' power supplies (sensor scratch pad memories) at the other end. The 110 volt, 400 hz (format), is fixed, but the amount of current transmitted (transmission rate) is variable. The power supplies are federated as are our converters, one per system. Tying one more system into the circuit is dependent on the generator capacity (computer) not the tranmission line as long as you don't exceed the current capacity (bandwidth). If more capacity (bandwidth) is required, another line (bus) is installed to handle the extra load.

V. CONCLUSIONS

In the future, there will be more and more digital subsystems available and it would definitely be advantageous if all had the same flexible standard interface. With all subsystems and computers having identical digital interfaces (electrical and format), the aircraft avionics could be configured in a building block fashion using the central computer software to integrate it. Systems could be simulated via ground computers to show data requirements, system responses and accuracy, and system flexibility. AGE could be made programmable, simulating the remainder of the avionic system via data bus. Subsystems would be interchangeable between aircraft and computers would no longer become obsolete.

Some of the prime advantages of the digital TDM multiplex data bus are:

1. Substantial weight savings (wire, connectors)

2. Simpler wire routing in aircraft (fewer bulkhead holes, clamps, connectors)

3. Fewer data transfers (sensor information on demand only)

4. Ease of changing sensors (because of standard digital interface)

5. Less tendency to obsolescence

6. Less noise from EMI (because of digital transmission)

7. Higher reliability

This transmission technique is being recognized within the Government as advantageous in complex systems integration and research and development is in process in these areas: avionic system integration, automatic built-in-test, electrical power switching, fly by wire, digital voice communications, flight instrumentation and digital coded firing systems.

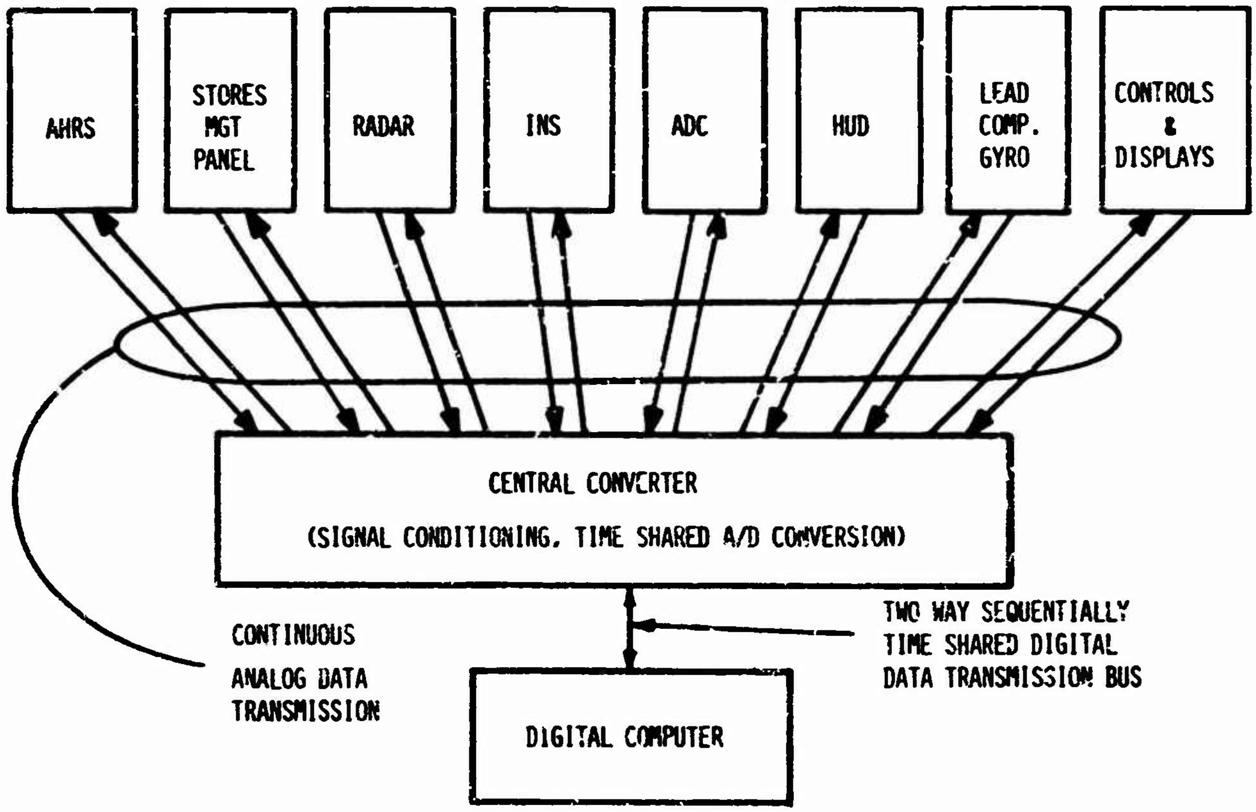


Fig.1 Conventional analog integration

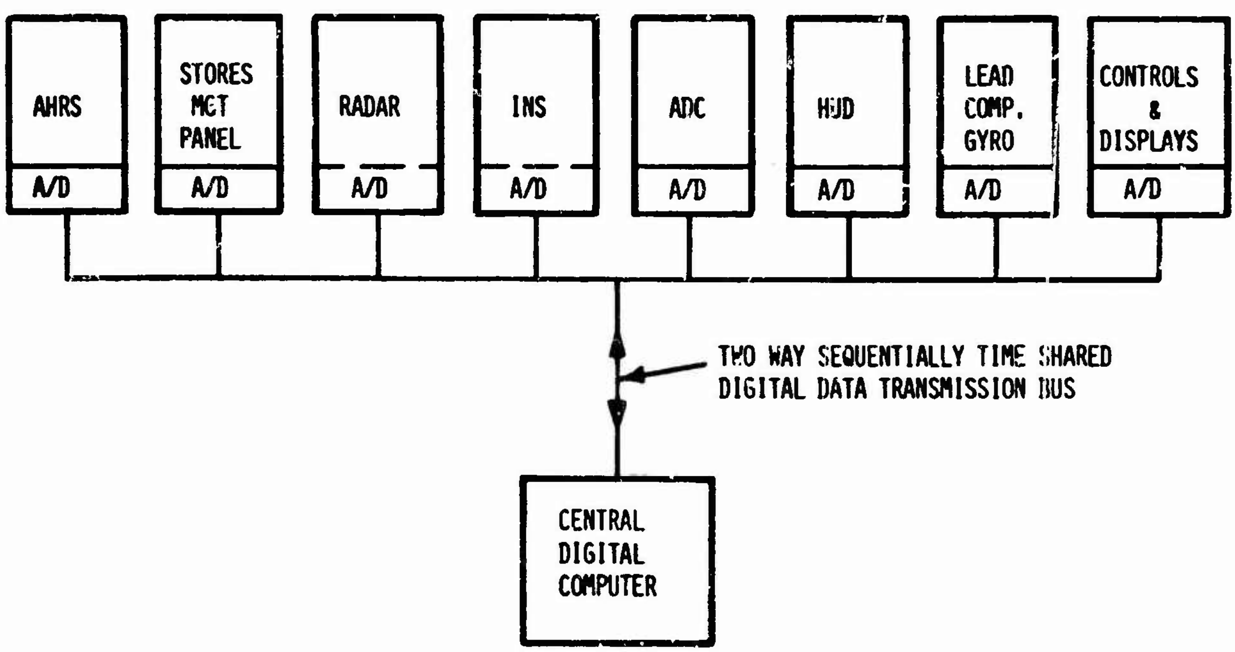


Fig.2 Multiplexed data bus

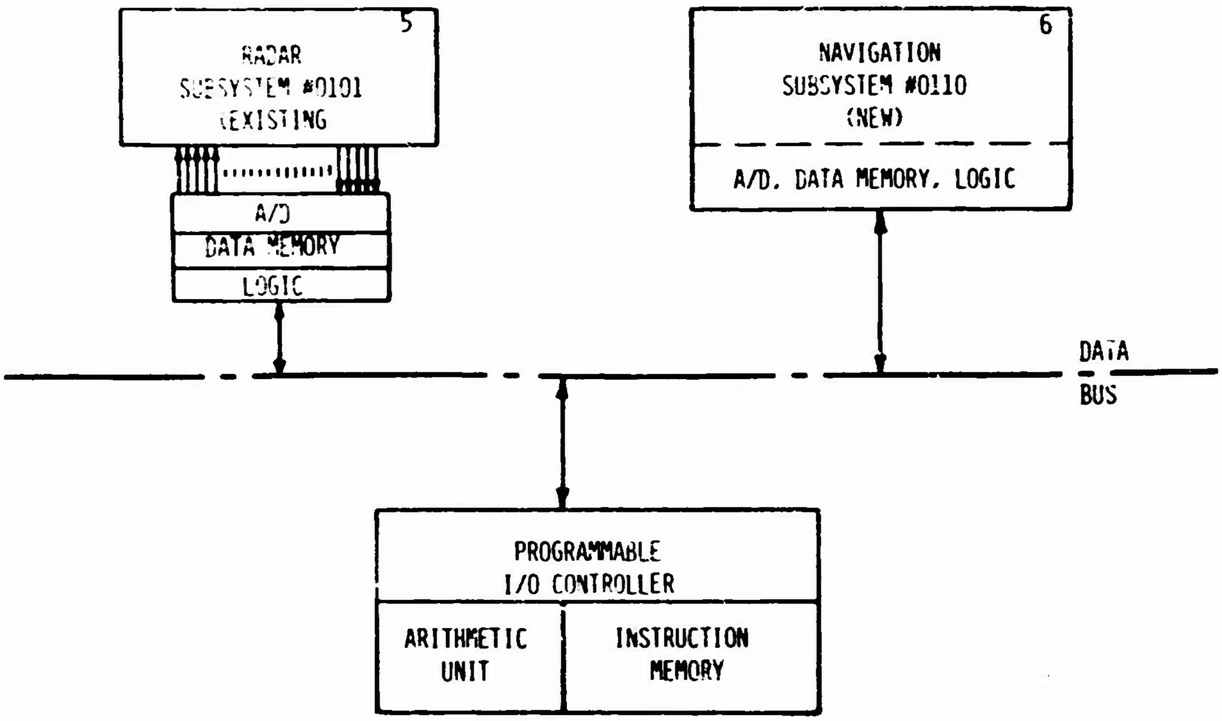


Fig.3 Data bus terminals

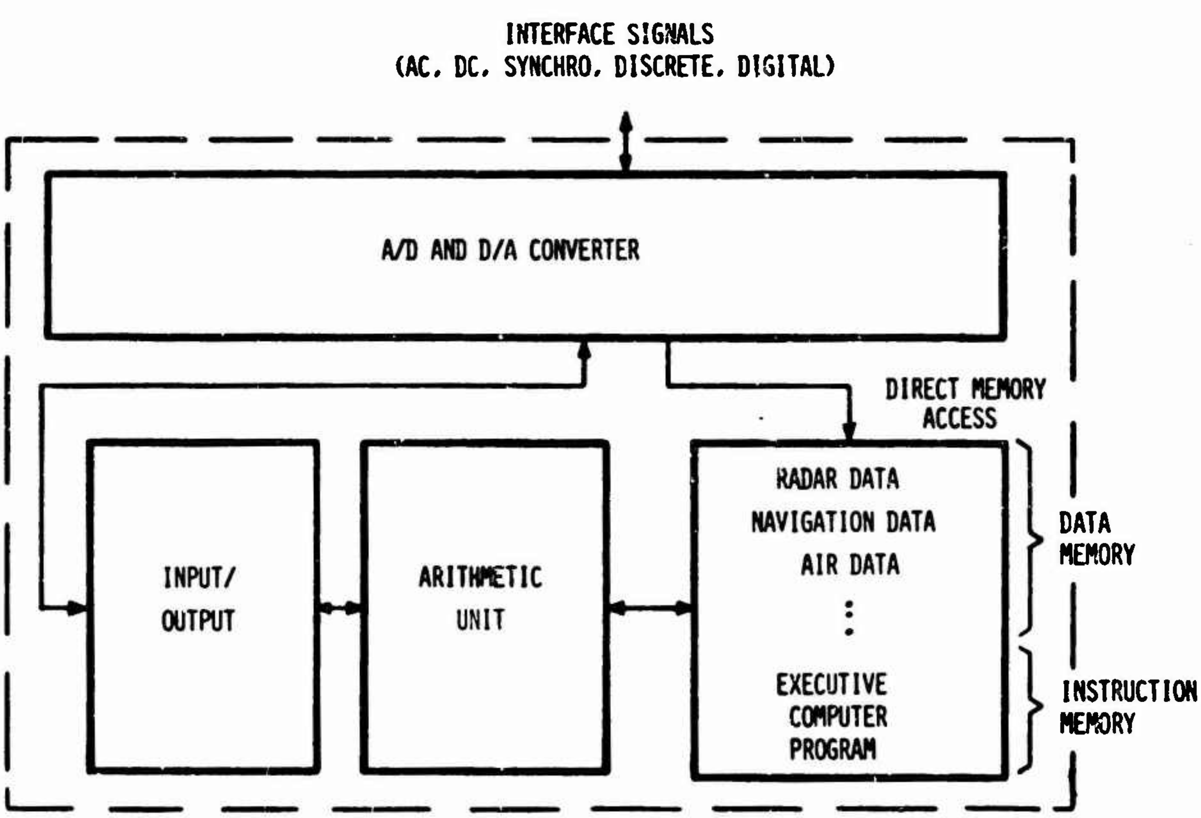


Fig.4 Analog interfaced digital computer

THE EXPONENTIAL PROBABILITY DISTRIBUTION AND ITS USE IN ASSESSING THE PERFORMANCE STATISTICS OF AEROSPACE SYSTEMS

D.A. LLOYD
Smiths Industries Limited, Aviation Division,
Cheltenham, England.

SUMMARY

The statistics of the output state variables of automatic aerospace systems are of considerable interest and of wide application, particularly in the case of manned systems. The paper shows that the exponential probability distribution can be used as an approximation to the distributions of the output state variables of practical aerospace systems for a wide range of practical situations.

The use of the exponential distribution as a practical mathematical tool is suggested, in the assessment of some of the performance statistics of aerospace systems, both for preliminary calculations and for final calculations involving the extrapolation of test results.

Symbols and Definitions

The symbols used in this paper are those defined in:

Higher Transcendental Functions Volumnes 1 and 2, Bateman Manuscript Project, California Institute of Technology, Edited by Arthur Erdélyi, Published McGraw-Hill, 1953.

Standard definitons are used in almost all cases. Possible exceptions are the error function definitons which are defined here as:

$$\mathrm{Erf}(x) = \int_0^x \exp\left\{-t^2\right\} dt, \quad \text{Error function,}$$

and

$$\mathrm{Erfc}(x) = \int_x^\infty \exp\left\{-t^2\right\} dt, \quad \text{Complementary error function.}$$

These definitions differ by a factor of $\frac{2}{\sqrt{\pi}}$ from those used in some text books.

A list of some of the less commonly used symbols is given below:

- $\Gamma(a,x)$ - Incomplete gamma function.
- $J_v(x)$ - Bessel function of the first kind.
- $I_v(x)$ - Modified Bessel function of the first kind.
- $K_v(x)$ - Modified Bessel function of the third kind.
- $P_v(x)$ - Legendre function. (Except when $P_1(x)$ is used as a cumulative probability function - the meanings of these terms are apparent from the context).
- $F(a,b;\gamma;x)$ - Hypergeometric function.
- $L_v(x)$ - Modified Struve function.

The following definitions are used for the purpose of this paper:

Random variable (or random process): a variable (or process) defined by its statistical properties.

Gaussian random variable: a random variable for which for every finite set of instants t_n, the random variables $x_n = x_{t_n}$ have a normal joint probability distribution.

Stationary random variable: a random variable having constant statistical properties over all time intervals, that is, constant mean, standard deviation.

Gaussian noise: a noise signal having the properties of a Gaussian random variable.

1. INTRODUCTION

The probability distributions of the output state variables of aerospace systems are of interest from many points of view, particularly in the case of manned systems where these statistics have a considerable influence on the safety statistics of the systems. The aspects considered in this paper are those concerned with the assessment of some of the performance statistics of the (automatically) controlled vehicle. Two types of effects are examined; these are the effects due to some particular noise inputs and the effects due to system tolerances.

Except in the practical results given in section 4, noise inputs with zero mean values are assumed throughout the paper. This assumption is not considered to affect the usefulness of the results, because with the linear or almost linear systems considered here conventional methods can be used to combine the effects described here with those due to non-zero mean values of the inputs.

It is well known that the output state variables of linear systems with Gaussian noise inputs will have normal probability distributions, and that (by the Central Limit Theorem) output variables which are formed from the sums of many small random variables having a wide range of possible probability distributions tend to have normal distributions. For these reasons, and also because many mathematical advantages accrue if normal probability distribution can be assumed, wide use is made of the normal probability distribution in the design and assessment of automatic control systems.

It is the object of this paper to show that although the normal probability distribution is a useful and valid tool for the assessment of the performance statistics of aerospace systems in particular cases, there are advantages to be gained by using the exponential distribution in many practical cases.

Many workers in the field of flight control systems and navigation have noted that the probability distribution of many flight test and simulator results exhibit considerable departures from the normal distribution, and some have noticed that many of these results approximate closely to the exponential distribution; see, for example, reference 1. It is shown in reference 2 how the apparent variation of results that arise in practice may be unified into a simple inclusive pattern. A different approach (from that of reference 2) is adopted in this paper.

2. THEORY

2.1 Noise Inputs

To illustrate points made in later sections, a model of a practical situation is used here. In this model a large number of experiments is assumed to be carried out with a linear system with Gaussian noise inputs. It is assumed that the noise inputs are defined by their power spectral densities $S_n(\omega)$, where

$$S_n(\omega) = \sigma_n^2 \, \phi_n(\omega), \qquad \text{------- (1)}$$

where $\phi_n(\omega)$ is a constant function such that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_n(\omega) d\omega = 1 \qquad \text{------- (2)}$$

and the variance of the noise input, which is assumed to have zero mean value, is given by

$$\text{Var.}(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_n(\omega) d\omega = \sigma_n^2 \qquad \text{------- (3)}$$

During each experiment σ_n is constant so that the noise input is a stationary random process, but σ_n is varied between experiments so that, considered over all experiments, σ_n can be considered to be a random variable with a probability density function $p(\sigma_n)$.

It can be seen that during any one experiment an output state variable x will have a normal probability distribution with a standard deviation σ_x that is proportional to the value of σ_n used in the experiment. Thus, over all experiments, σ_x will have a probability distribution of the same type as σ_n. In mathematical terms, the conditional probability density function of x given σ_x, (or σ_n) is given by

$$p(x|\sigma_x) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \exp\left\{-\frac{x^2}{2\sigma_x^2}\right\} \qquad \text{------- (4)}$$

The marginal p.d.f. of x is given by

$$p(x) = \int_0^{\infty} p(x|\sigma_x)\, p(\sigma_x)\, d\sigma_x \qquad \text{------- (5)}$$

Equation 5 can be used to determine the function p(x) for any given functions $p(x|\sigma_x)$ and $p(\sigma_x)$. Some functions p(x), for $p(x/\sigma_x)$ a normal distribution and for various forms of $p(\sigma_x)$, are given in Table 1.

For conditions defined by equations 4 and 5

$$\overline{x^2} = \sigma^2 = 2\int_0^{\infty} x^2 p(x)dx, \text{ where } \sigma^2 \text{ is the variance of x over all experiments}$$

$$= 2\int_0^{\infty} \left\{ x^2 \int_0^{\infty} \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{x^2}{2\sigma_x^2}\right] . \, p(\sigma_x)\, d\sigma_x \right\} dx \qquad \text{------- (6)}$$

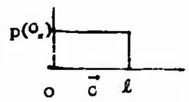
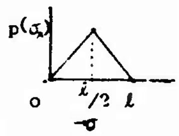
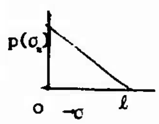
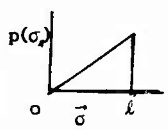
Re-arranging and changing the order of integration gives

$$\sigma^2 = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \left\{ p(\sigma_x) \int_0^{\infty} \frac{x^2}{\sigma_x} \exp\left[-\frac{x^2}{2\sigma_x^2}\right] dx \right\} d\sigma \qquad \text{------- (7)}$$

$$= \int_0^{\infty} \sigma_x^2\, p(\sigma_x)\, d\sigma_x = \overline{\sigma_x^2}\,, \qquad \text{------- (8)}$$

or, the variance of x is equal to the mean square value of σ_x over all experiments.

TABLE 1

ITEM	$p(\sigma_x)$	$\bar{\sigma}_x$	VARIANCE of σ_x	$p(x)$	$\sigma^2 = \overline{\sigma_x^2}$
(i)	$\frac{\sigma_x}{\mu^2} \exp\left\{-\frac{\sigma_x^2}{2\mu^2}\right\}$ (Rayleigh)	$\sqrt{\frac{\pi}{2}}\cdot\mu$	$(4-\pi)\frac{\mu^2}{2}$	$\frac{1}{2\mu} \exp\left\{-\frac{\lvert x\rvert}{\mu}\right\}$ (Exponential)	$2\mu^2$
(ii)		$\frac{\ell}{2}$	$\frac{\ell^2}{12}$	$\frac{1}{2\sqrt{2\pi}}\cdot\frac{1}{\ell}\cdot\left[-E_1\left(\frac{-x^2}{2\ell^2}\right)\right]$	$\frac{\ell^2}{3}$
(iii)		$\frac{\ell}{2}$	$\frac{\ell^2}{24}$	$\frac{4}{\sqrt{2\pi}}\cdot\frac{1}{\ell}\left[\exp\left\{-\frac{2x^2}{\ell^2}\right\}-\exp\left\{-\frac{x^2}{2\ell^2}\right\}\right] - \frac{\lvert x\rvert}{\sqrt{\pi}}\cdot\frac{4}{\ell^2}\left[2\,\mathrm{Erfc}\left(\frac{\sqrt{2}\lvert x\rvert}{\ell}\right) - \mathrm{Erfc}\left(\frac{\lvert x\rvert}{\sqrt{2}.\ell}\right)\right] + \frac{2}{\sqrt{2\pi}}\cdot\frac{1}{\ell}\cdot\left[E_1\left(\frac{-2x^2}{\ell^2}\right) - E_1\left(\frac{-x^2}{2\ell^2}\right)\right]$	$\frac{7\ell^2}{24}$
(iv)		$\frac{\ell}{3}$	$\frac{\ell^3}{18}$	$\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\ell}\left[-E_1\left(\frac{-x^2}{2\ell^2}\right) - 2\exp\left\{\frac{-x^2}{2\ell^2}\right\} + \frac{2\sqrt{2}}{\ell}\cdot\lvert x\rvert\cdot\mathrm{Erfc}\left(\frac{\lvert x\rvert}{\sqrt{2}\ell}\right)\right]$	$\frac{\ell^2}{6}$
(v)		$\frac{2\ell}{3}$	$\frac{\ell^2}{18}$	$\frac{2}{\sqrt{2\pi}}\cdot\frac{1}{\ell}\cdot\exp\left\{\frac{-x^2}{2\ell^2}\right\} - \frac{2\lvert x\rvert}{\sqrt{\pi}}\cdot\frac{1}{\ell^2}\cdot\mathrm{Erfc}\left(\frac{\lvert x\rvert}{\sqrt{2}\ell}\right)$	$\frac{\ell^2}{2}$
(vi)	$\frac{2}{\sqrt{2\pi}\mu}\exp\frac{-\sigma^2}{2\mu^2}$	$\sqrt{\frac{2}{\pi}}\cdot\mu$	$\left(1-\frac{2}{\pi}\right)\mu^2$	$\frac{1}{\pi\mu} K_0\left(\frac{\lvert x\rvert}{\mu}\right)$ *	μ^2

* The distribution $\frac{1}{\pi\mu} K_0\left(\frac{\lvert x\rvert}{\mu}\right)$ is also given by the system defined by $x = ab$, where a and b are independent normally distributed random variables with zero mean values. In this case $\mu = \sigma_a.\sigma_b$.

The cumulative distribution of x is defined by

$$P_{(x)} = \int_{-\infty}^{x} p_x(t)dt, \qquad \text{------- (9)}$$

but for the present purposes the function $P_1(x)$ defined by

$$P_1(x) = 1 - P_{(x)} = \int_x^{\infty} p_x(t)dt \qquad \text{------- (10)}$$

is probably more useful than $P_{(x)}$.

For the forms of $p_{(x)}$ given by items (i), (ii), and (vi) of Table 1 the functions $P_1(x)$ are given by:

$$\text{(i)}\quad P_{1(x)} = \frac{1}{2\mu}\int_x^{\infty} e^{-(t)/\mu}\,dt = \frac{1}{2}e^{-(x)/\mu}, \quad x \geq 0 \qquad \text{------- (11)}$$

$$\text{(ii)}\quad P_{1(x)} = \frac{1}{2\sqrt{2\pi}\,\ell}\int_x^{\infty} -E_1\left(-\frac{t^2}{2\ell^2}\right)dt$$

$$= \frac{1}{2\sqrt{\pi}}\left[\Gamma\left(\frac{1}{2}, \frac{x^2}{2\ell^2}\right) + \frac{x}{\sqrt{2}\ell}\,E_1\left(-\frac{x^2}{2\ell^2}\right)\right], \quad x \geq 0 \qquad \text{------- (12)}$$

$$\text{(vi)}\quad P_{1(x)} = \frac{1}{\pi\mu}\int_x^{\infty} K_0\left(\frac{t}{\mu}\right)dt$$

$$= \frac{1}{2}\left\{1 - \frac{x}{\mu}\left[K_0\left(\frac{x}{\mu}\right)\cdot L_{-1}\left(\frac{x}{\mu}\right) + K_{-1}\left(\frac{x}{\mu}\right)\cdot L_0\left(\frac{x}{\mu}\right)\right]\right., \quad x \geq 0 \qquad \text{------- (13)}$$

Some of the probability distributions of Table 1 and equations 11, 12 and 13 are plotted (for $\sigma^2 = 1$) in figures 1, 2 and 2a. It can be seen from Table 1 that if $p(\sigma_x)$ is a Rayleigh distribution then $p(x)$ is (exactly) an exponential distribution. An examination of Table 1 and figures 1 and 2 shows that in all cases the exponential distribution gives a better approximation of $p(x)$ than the normal distribution for values of x greater than about 2.5 for the ranges of x considered in the figures

2.2 The Rayleigh Distribution, (and Similar Distributions)

In section 2.1 it was shown that an exponential distribution is generated if the conditional probability $p(x \mid \sigma_x)$ is normal and $p(\sigma_x)$ is a Rayleigh distribution. It may also be inferred from the results given in Table 1 that forms of $p(\sigma_x)$ of the general shape of the Rayleigh distribution lead to distributions of $p(x)$ approximating closely to the exponential distribution. It is therefore of some interest to examine the Rayleigh distribution and approximations to distribution in more detail.

It is well known that if

$$z^2 = a^2 + b^2 \qquad \text{------- (14)}$$

and a and b are normally distributed random variables, and if $\sigma_a^2 = \sigma_b^2 = \sigma^2$, and $\bar{a} = \bar{b} = 0$, then

$$p(z) = \frac{z}{\sigma^2} \exp\left\{-\frac{z^2}{2\sigma^2}\right\}, \qquad z = |z|. \qquad \text{------- (15)}$$

That is, z has a Rayleigh distribution.

Equation 14 arises naturally in many practical situations where z is the vector sum of two orthogonal vectors a and b with statistically independent normally distributed amplitudes . In some cases the relationship $\sigma_a^2 = \sigma_b^2$ will also arise naturally from the physical properties of the process under consideration, but it is of some interest to examine the nature of p(z) when $\sigma_a^2 \neq \sigma_b^2$. Under these conditions equation 5 can be used to show that

$$p(z) = \frac{z}{\sigma_a\sigma_b} \exp\left\{-\frac{z^2}{2}\left(\frac{1}{2\sigma_a^2} + \frac{1}{2\sigma_b^2}\right)\right\} \cdot I_o\left\{\frac{z^2}{2}\left(\frac{1}{2\sigma_a^2} - \frac{1}{2\sigma_b^2}\right)\right\}, \qquad \text{------- (16)}$$

for $z = |z|$.

$$\bar{z} = \sqrt{\frac{\pi}{2}}\,\sqrt{\sigma_a\sigma_b}\; P_{\frac{1}{2}}\left\{\frac{\sigma_a^2 + \sigma_b^2}{2\sigma_a\sigma_b}\right\} \qquad \text{------- (17)}$$

$$\overline{z^2} = \left(\sigma_a^2 + \sigma_b^2\right) \qquad \text{------- (18)}$$

Consider the case where $\sigma_b = n\sigma_a$, $n \geq 1$, and $\overline{z^2} = 1$

Equation 16 can be written in the forms:

$$p(z) = \frac{(n^2+1)}{n}\, z \exp\left\{-\frac{z^2}{2}\left(\frac{n^2+1}{n^2}\right)\right\}\left[\exp\left\{-\frac{z^2}{4}\left(\frac{n^4-1}{n^2}\right)\right\} I_o\left\{\frac{z^2}{4}\left(\frac{n^4-1}{n^2}\right)\right\}\right] \qquad \text{----(19a)}$$

$$p(z) = 2\frac{(n^2+1)^{\frac{1}{2}}}{(n^2-1)^{\frac{1}{2}}} \exp\left\{-\frac{z^2}{2}\left(\frac{n^2+1}{n^2}\right)\right\}\left[\frac{z}{2}\,\frac{(n^4-1)^{\frac{1}{2}}}{n} \cdot \exp\left\{-\frac{z^2}{4}\left(\frac{n^4-1}{n^2}\right)\right\} \cdot I_o\left\{\frac{z^2}{4}\left(\frac{n^4-1}{n^2}\right)\right\}\right] \qquad \text{------- (19b)}$$

$$\left.\begin{aligned} &\lim_{q\to 0} e^{-q} = 1, \text{ and } \lim_{q\to 0} I_o(q) = 1 \\ &\lim_{q\to\infty}\left[q^{\frac{1}{2}}e^{-q}\, I_o(q)\right] = \frac{1}{\sqrt{2\pi}} \end{aligned}\right\} \qquad \text{------- (20)}$$

From equations 19a, 19b and 20 it can be seen that

$$\lim_{n\to 1} p(z) = 2ze^{-z^2} \qquad \text{------- (21a)}$$

$$\lim_{n\to\infty} p(z) = \frac{2}{\sqrt{2\pi}}\, e^{-z^2/2} \qquad \text{------- (21b)}$$

That is, the limiting forms of p(z) are a Rayleigh distribution for $n\to 1$, and a (one sided) normal distribution for $n\to\infty$.

If $z = \sigma_x$, and the conditional probability $p(x \mid \sigma_x)$ is normal, then for the unit variance case ($\overline{x^2} = 1$, $\bar{x} = 0$), the results of section 2.1 show that, for all practical purposes for p(z) given by equation 16, p(x) and $P_1(x)$ will lie in the shaded regions of figures 1 and 2

Figure 1 includes some points from the distributions p(x) derived from $p(\sigma_x)$ of the form of equation 16 with values of n (where $n = \sigma_b/\sigma_a$ in equation 16) of 2 and 5. These points indicate that for n ≤ about 2, the exponential distribution provides a reasonable approximation to p(x).

2.3 Tolerances

The effect of tolerances within a system on the statistics of the output state variables of the system is usually small, (in most practical systems), in comparison with the effects due to the noise inputs. An important exception to this statement is in the case of multiplexed systems where the effects of tolerances have a major influence on the statistics of nuisance cut-outs (see section 2 3.1).

To illustrate the effects of tolerances a simple model will be used. The model consists of a duplex system with two nominally identical sub-channels.

Let the nominal transfer function between the (noise) input and the output state variable x be G(D), where

$$G(D) = \frac{B_o + B_1D + \text{----} + B_mD^m}{A_o + A_1D + \text{----} + A_nD^n} \qquad \text{-------} (22)$$

where $n > m$.

Let the actual transfer functions of the two sub-channels be $G_1(D)$ and $G_2(D)$, where

$$G_1(D) = \frac{(B_o+\beta_o) + (B_1+\beta_1)D + \text{--} + (B_m+\beta_m)D^m}{(A_o+\alpha_o) + (A_1+\alpha_1)D + \text{--} + (A_n+\alpha_n)D^n}, \qquad \text{-------} (23a)$$

$$G_2(D) = \frac{(B_o+\beta_o^1) + (B_1 + \beta_1^1)D + \text{----} + (B_m+\beta_m^1)D^m}{(A_o+\alpha_o^1) + (A_1 + \alpha_1^1)D + \text{----} + (A_n+\alpha_n^1)D^n}, \qquad \text{-------} (23b)$$

where β_r, β_r^1, α_r, α_r^1 are terms due to the tolerances in the two sub-channels.

The state variable of interest for the present purpose is the difference between the outputs of the two sub-channels. Let this difference be denoted by e, where

$$e = x_1 - x_2 \qquad \text{-------} (24)$$

where x_1 and x_2 are the outputs of the two sub-channels due to the common input.

It can be seen that the transfer function between the input and the state variable e is given by $G_e(D)$, where

$$G_e(D) = G_1(D) - G_2(D) \qquad \text{-------} (25)$$

If the individual tolerances are small, so that $\beta_r \ll B_r$, $\alpha_r \ll A_r$, and similarly for β_r^1 and α_r^1 so that second order terms can be neglected in the numerator of $G_c(D)$ and first order terms can be neglected in the denominator, $G_e(D)$ can be approximated by

$$G_e(D) = \frac{E_o+E_1D + E_2D^2 + \text{----} + E_{m+n}D^{m+n}}{(A_o+A_1D + \quad \text{----} + A_nD^n)^2} \qquad \text{-------} (26)$$

$$\left.\begin{aligned} \text{where } E_o &= A_o(\beta_o-\beta_o^1) + B_o(\alpha_o^1 - \alpha_o) \\ E_1 &= A_o(\beta_1-\beta_1^1) + B_o(\alpha_1^1 - \alpha_1) + A_1(\beta_o-\beta_o^1) + B_1(\alpha_o^1-\alpha_o) \\ &\vdots \\ &\text{etc.} \end{aligned}\right\} \qquad \text{-------} (27)$$

$$\text{Let } y = \frac{1}{(A_o+A_1D + \text{---} + A_nD^n)^2} \cdot n \qquad \text{-------} (28)$$

where n is the noise input. Then

$$e = E_oy_o + E_1 y_1 + E_2y_2 + \text{---} + E_{m+n} y_{m+n} \qquad \text{-------} (29)$$

where $y_r = D^r y$.

Two particular cases will now be considered.

Case 1, m large

For a large number of systems with different tolerances, if it is assumed that the terms E_r are statistically independent, from equation 29 it can be seen that e is given by the sum of many small independent random quantities of the form E_ry_r, ($r = 1,2, \ldots, m+n$). Assuming zero mean values for input and tolerances the Central Limit Theorem can be applied to equation 29 to indicate that the function p(e) will tend to a normal distribution function. That is, for a stationary noise input, (which need not be Gaussian),

$$p(e) \approx \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left\{-\frac{e^2}{2\sigma_e^2}\right\} \qquad \text{-------} (30)$$

Case 2, n+m small (about 5 < (n+m) < about 10)

In equation 29, with a Gaussian input the quantities y_r will be normally distributed, and the terms E_ry_r will have probability distributions approximating to the form

$$p(q_r) = \frac{1}{\pi\mu} K_o\left(\frac{(q_r)}{\mu}\right), \quad q_r = E_ry_r$$

These distributions are considerably different from the normal distribution and the application of the Central Limit Theorem to a limited number of terms of this form is not valid.

Consider the case (n+m) = 5. From equation 29 $e = e_1+e_2$ -------- (31)

$$\left.\begin{aligned} \text{where } e_1 &= (E_oy_o + E_2y_2 + E_4y_4) \\ e_2 &= (E_1y_1 + E_3y_3 + E_5y_5) \end{aligned}\right\} \qquad \text{-------} (32)$$

It can be shown that if the noise input is a stationary Gaussian process e_1 and e_2 are statistically independent, and

$$\sigma_e^2 = \sigma_{e1}^2 + \sigma_{e2}^2 \qquad \text{-------} (33)$$

For a particular set of tolerances E_r,

$$\begin{aligned} \sigma_{e1}^2 &= \overline{e_1^2} = E_o^2\overline{y_o^2} + E_2^2\overline{y_2^2} + E_4^2\overline{y_4^2} \\ &\quad + 2E_oE_2\overline{y_oy_2} + 2E_oE_4\overline{y_oy_4} + 2E_2E_4\overline{y_2y_4}, \end{aligned} \qquad \text{-------} (34)$$

$$\left.\begin{array}{l} \text{but } \overline{y_r^2} = \sigma_{yr}^{\;2} = \sigma_r^{\;2} \\ \text{and it can be shown that } \overline{y_r y_s} = (-1)^{\frac{r+s}{2}} \sigma^2_{y\left(\frac{r+s}{2}\right)} \end{array}\right\} \qquad \text{------- (35)}$$

Thus equation 34 can be written in the form

$$\sigma_{e1}^{\;2} = \left[E_o\sigma_o - E_2\sigma_2 + E_4\sigma_4\right]^2 + 2E_oE_2\left(\sigma_o\sigma_2 - \sigma_1^{\;2}\right)$$

$$-2E_oE_4(\sigma_o\sigma_4 - \sigma_2^{\;2}) + 2E_2E_4(\sigma_2\sigma_4 - \sigma_3^{\;2}) \qquad \text{------- (36)}$$

It can be shown that $\sigma_r\sigma_s - \sigma^2_{\frac{r+s}{2}}$ is always positive, and must be less than $\sigma_r\sigma_s$.

In the case considered here ($n+m < 10$) it is apparent that the magnitude of the sum of the terms outside the squared bracket in equation 36 is considerably less than the square of the sum of the terms inside the squared bracket. (This is not necessarily true for $(n+m) >> 10$ due to the fac[illegible] that in this case the number of terms of the form $2E_rE_s(\sigma_r\sigma_s - \sigma^2_{\frac{r+s}{2}})$ outside the squared bracket in equation 36 is much larger than the number of terms of the form $E_r\sigma_r$ inside the squared bracket).

Thus, in the case being considered here, equation 36 can be approximated by

$$\sigma_{e1}^{\;2} = \left[E_o\sigma_o - E_2\sigma_2 + E_4\sigma_4\right]^2 \qquad \text{------- (37a)}$$

Similarly $\sigma_{e1}^{\;2}$ can be approximated by

$$\sigma_{e2}^{\;2} = \left[E_1\sigma_1 - E_3\sigma_e + E_5\sigma_5\right]^2 \qquad \text{------- (37b)}$$

Thus, from equation 33, $\sigma_e^{\;2}$ can be approximated by

$$\sigma_e^{\;2} = \left[E_o\sigma_o - E_2\sigma_2 + E_4\sigma_4\right]^2 + \left[E_1\sigma_1 - E_3\sigma_3 + E_5\sigma_5\right]^2 \qquad \text{------- (38)}$$

From equation 27 it can be seen that each of the terms E_r is composed of the sum of a number of small quantities. Considering a large number of systems with different tolerance values, each term E_r will tend to have a bell-shaped distribution of approximately normal form. Thus the terms $[E_o\sigma_o - E_2\sigma_2 + E_4\sigma_4]$ and $[E_1\sigma_1 - E_3\sigma_3 + E_5\sigma_5]$ are of the ideal forms for the application of the Central Limit theorem, in spite of the limited number of terms in each bracketed term. The application of the Central Limit Theorem to the terms in the brackets will lead to the conclusion that each of the squared terms in equation 38 will have a probability distribution approximating to that of the square of a normally distributed random variable, except in the tail regions which will be restricted by the worst tolerance case so that $p(\sigma_e)$ will be zero for values of $\sigma_e \geq$ the value corresponding to the worst tolerance case.

If the terms $[E_o\sigma_o - E_2\sigma_2 + E_4\sigma_4]$ and $[E_1\sigma_1 - E_3\sigma_3 + E_5\sigma_5]$ were normally distributed, then from section 2.2 it can be seen that $p(\sigma_e)$ would be of the form given by equation 16. The effects due to the facts that σ_{e1} and σ_{e2} are unlikely to be equal and the restriction on the tail regions mentioned above will tend to cancel, ($\sigma_e \neq \sigma_{e2}$ will tend to raise the tail region of $p(\sigma_e)$ above that of the Rayleigh distribution and the restriction due to the restricted value of σ_e maximum will tend to lower the tail region. Thus continuing this argument to the probability distribution p(e), it appears likely that in most practical cases the probability distribution of e can be approximated by an exponential distribution, for Case 2 with a stationary Gaussian noise input. That is

$$p(e) \simeq \frac{1}{2\mu} \exp\left\{-\frac{|e|}{\mu}\right\} \qquad \text{------- (39)}$$

It should be noted that the equations of motion of an aircraft can be approximated by a [illegible] order transfer function ($m = 3$, $n = 4$, $n + m = 7$) to sufficient accuracy for many pra[illegible] purposes, and such an approximation is frequently used for simulation purposes. It c[illegible]een from the discussion relating to Case 1 that the value of p(e) derived for this case [illegible]ower values for large values of e than the corresponding value of p(e) given by equation [illegible] It is therefore suggested that for aircraft calculations the value of p(e) given by equatio[illegible]9 can provide a useful practical upper limit for p(e) for e large.

2.3.1 Nuisance cut-outs

In multiplexed systems it is usual to cut out faulty sub-channels using a discriminant depending on the magnitude of the difference between the outputs of nominally identical sub-channels. In the presence of tolerances, the level of the difference signal causing cut-out is usually a compromise between the conflicting requirements to detect small faults and to avoid cut-outs caused by the effects of tolerances when no faults are present. The latter type of cut-out is known as a nuisance cut-out. Methods of cut-out depend on the particular system requirements and the degree of multiplicity of the system. The principles involved will be illustrated by means of the simple duplex arrangement used in section 2.3.

It is assumed that a nuisance cut-out occurs with no fault present when the difference signal (denoted by x in this section) passes through the levels $\pm$ c in the direction $|x|$ increasing. For a stationary Gaussian noise input the output x will be normally distributed. It is well known (see, for example, reference 3) that under these conditions the mean rate of crossing the level $\pm$ c in the direction $|x|$ increasing is given by N, where

$$N = \frac{1}{\pi}\frac{\sigma_1}{\sigma_o} \exp\left\{-\frac{c^2}{2\sigma_o^{\;2}}\right\}, \text{crossings per unit time}, \qquad \text{------- (40)}$$

where

σ_o^2 = variance of x,

σ_1^2 = variance of $\dot{x}$.

Equation 40 can be applied to any linear system with any particular set of tolerances and with a stationary Gaussian noise input.

The results of Section 2.3 for Case 1 suggest that this result can also be applied to a large number of systems (of with high order transfer functions) with a wide range of tolerance values and a stationary noise input.

The case of more interest here is that of Case 2 which has been shown to provide a reasonable approximation in the case of aircraft calculations. In this case, for a stationary Gaussian noise input and assuming that p(x) and p($\dot{x}$) are approximately exponential, the methods of reference 3 can be used to show that the mean crossings the levels $\pm$ c in the direction $|x|$ increasing, in unit time, can be approximated by N_1, where

$$N_1 = \frac{\mu_1}{2\mu_o} \exp\left\{-\frac{|c|}{\mu_o}\right\} \qquad \text{------- (41)}$$

where μ_o and μ_1 are defined by the probability distributions of x and $\dot{x}$:

$$p(x) = \frac{1}{2\mu_o} \exp\left\{-\frac{|x|}{\mu_o}\right\}$$

$$p(\dot{x}) = \frac{1}{2\mu_1} \exp\left\{-\frac{|\dot{x}|}{\mu_1}\right\}.$$

In practical systems the mean nuisance cut-out rate must be small, ($^c/\mu_o$ must be large). The probability of cut-out in any very short interval of time δt is given by $N_1\delta t$. Under these circumstances the following probabilities can be derived:

$$\left.\begin{aligned}
&\text{probability of n cut-outs in interval (O,T)} = \frac{1}{n!}(N_1T)^n, \ \exp\left\{-N_1T\right\}\\
&\text{probability of no cut-outs in (O,T)} = \exp\left\{-N_1T\right\}\\
&\text{probability of one or more cut-outs in (O,T)} = 1 - \exp\left\{-N_1T\right\}
\end{aligned}\right\} \qquad \text{------- (42)}$$

In both cases discussed in Sections 2.3 and 2.3.1 it is assumed that many tolerances are present in in the system. This assumption is necessary in both cases in order that the terms E_r in equations 26-38 shall exist and can be assumed to be statistically independent. If this assumption is not true then the application of the Central Limit Theorem to equations 29 and 38 is not valid.

In cases where small numbers of tolerances are involved, departures from the exponential distribution for Case 2 systems with stationary Gaussian noise inputs can be expected. In general, the resulting distributions for these cases can be expected to give smaller values of p(x), and smaller values of the mean rates of cut-out for large values of x and c than those given by the results of sections 2.3 and 2.3.1.

It is apparent from the results given here that the mean cut-out rates for low order systems (such as aircraft) will generally be much larger than the corresponding values given by the assumption of normal distributions for the state variable x.

2.4 Duration of Level Exceedances

A variable of interest in many aerospace system calculations, from the points of view of safety and stressing, is the duration of the intervals for which a state variable x exceeds some level x = c. This topic is obviously connected with the subjects discussed in Section 2.3 and 2.3.1, but in the present context it is assumed that the system transfer function is constant and the noise input can be represented by a power spectral density of the form shown in equation 1 with the restriction given in section 2.1, (that is, $\phi_n(\omega)$ is assumed to be a constant function). Under these conditions the ratio σ_1/σ_o, where $\sigma_o = \sigma_x$ and $\sigma_1 = \sigma_{\dot{x}}$, is a constant, (which is proportional to the number of zero axis crossings in unit time of the noise input and also to the number of zero axis crossings in unit time of the variable x).

$$\text{Let } \frac{\sigma_1}{\sigma_o} = R \qquad \text{------- (43)}$$

Then for a stationary input the mean duration of the periods for which $|x| \geq c$ is given by $\bar{\tau}$, where

$$\bar{\tau} = \frac{2}{N}\int_c^{\infty} p(x)\,dx \qquad \text{------- (44)}$$

where N = average number of crossings of x = c in the direction $|x|$ increasing in unit time.

Thus, from equation 40, for a Gaussian noise input

$$\bar{\tau} = \frac{\frac{2}{\sqrt{2\pi}}\frac{1}{\sigma_o}\int_c^{\infty}\exp\left\{-\frac{x^2}{2\sigma_o^2}\right\}dx}{\frac{1}{\pi}\,R\,\exp\left\{-\frac{c^2}{2\sigma_o^2}\right\}} \qquad \text{------- (45)}$$

It can be shown that for $^c/\sigma_o$ large

$$\frac{2}{\sqrt{2\pi}\sigma_o}\int_c^{\infty}\exp\left\{-\frac{x^2}{2\sigma_o^2}\right\}dx \rightarrow \frac{2\sigma_o}{\sqrt{2\pi}}\cdot\frac{1}{c}\exp\left\{-\frac{c^2}{2\sigma_o^2}\right\} \qquad \text{------- (46)}$$

Thus for $\frac{c}{\sigma_o}$ large

$$\bar{\tau} = \frac{\sqrt{2\pi}\sigma_o}{Rc} \qquad \text{------- (47)}$$

For an exponential distribution of x generated by the system

$$\left.\begin{aligned} p(x|\sigma_o) &= \frac{1}{\sqrt{2\pi}\sigma_o} \exp\left\{-\frac{x^2}{2\sigma_o^2}\right\} \\ p(\sigma_o) &= \frac{\sigma_o}{\mu_o^2} \exp\left\{-\frac{\sigma_o^2}{2\mu_o^2}\right\} \end{aligned}\right\}, \qquad \text{------- (48)}$$

the quantity N is given by

$$N = \frac{R}{\pi} \cdot \frac{C}{\mu_o} \cdot K_1\left\{\frac{C}{\mu_o}\right\}, \qquad \text{------- (49)}$$

and

$$\bar{\tau} = \frac{\exp\left\{-\frac{C}{\mu_o}\right\}}{N} \qquad \text{------- (50)}$$

For $\frac{C}{\mu_o}$ large, $\frac{C}{\mu_o} K_1\left\{\frac{C}{\mu_o}\right\} \rightarrow \sqrt{\left(\frac{C.\pi}{\mu_o.2}\right)} \exp\left\{-\frac{C}{\mu_o}\right\}$ ------- (51)

and for $\frac{C}{\mu_o}$ large,

$$N \rightarrow \frac{1}{\sqrt{2\pi}} . R . \left(\frac{C}{\mu_o}\right)^{\frac{1}{2}} . \exp\left\{-\frac{C}{\mu_o}\right\}, \qquad \text{------- (52)}$$

and

$$\bar{\tau} \rightarrow \frac{\sqrt{2\pi}}{R} \left(\frac{\mu_o}{C}\right)^{\frac{1}{2}} \qquad \text{------- (53)}$$

The probability distribution of τ for a normal distribution of x is discussed in reference 4, where it is shown (pages 604 and 605) that for $\frac{C}{\sigma_o}$ large, $p(\tau)$ can be approximated by

$$p(\tau) = \frac{\sigma_1^2 C^2 \tau}{4\sigma_o^4} \exp\left\{-\frac{\sigma_1^2 C^2 \tau^2}{8\sigma_o^4}\right\}, \qquad \text{------- (54)}$$

(a Rayleigh distribution).

Using this result and R constant it can be shown that for the exponential distribution of the type used here $p(\tau)$ can be approximated for C/μ_o large, by

$$p(\tau) \approx \frac{R^2 C\tau}{2\mu_o(4+R^2\tau^2)^{\frac{1}{2}}} \exp\left\{\frac{C}{\mu_o}\right\} . \exp\left\{-\frac{C}{2\mu_o}(4+R^2\tau^2)^{\frac{1}{2}}\right\} \qquad \text{------- (55)}$$

or, as τ will usually be small for $\frac{C}{\mu_o}$ large, this expression can be further approximated by

$$p(\tau) \approx \frac{R^2 C\tau}{4\mu_o} \exp\left\{-\frac{R^2 C\tau^2}{8\mu_o}\right\} \qquad \text{------- (56)}$$

(a Rayleigh distribution).

It should be noted that if a variable x has a Rayleigh distribution of the form

$$p(x) = \frac{x}{\mu^2} \exp\left\{-\frac{x^2}{2\mu^2}\right\} \qquad \text{------- (57)}$$

the variable q, where $q = x^2$, has a distribution of the form

$$p(q) = \frac{1}{2\mu^2} \exp\left\{-\frac{q}{2\mu^2}\right\} \qquad \text{------- (58)}$$

(a one sided exponential distribution).

As discussed in reference 4 (page 606) it is to be expected for any distribution of τ for τ small, that $p(\tau)$ will decrease exponentially for large values of τ.

Practical distributions for τ can thus be obtained empirically by joining the distributions for τ small of the type given above to an exponentially decreasing function of τ for τ large by means of a smooth curve asymptotic to both $p(\tau)$, τ small and $p(\tau)$, τ large, in such a manner that the mean value of τ given by equation 44 is satisfied.

The derivation of $p(\tau)$ in the general case is rather difficult. Some of the difficulties are discussed in references 3 and 4.

2.5 The Probability Distribution of the sum of two exponentially distributed random variables

The main thesis of this paper is that for some particular noise inputs the output state variables of linear systems have probability distributions approximating to exponential distributions. In many cases, aerospace systems will be subjected to more than one noise input. However, in many practical cases one or two noise inputs predominate in their effects on the statistics of the output state variables; for example, atmospheric turbulence and radio noise in the case of automatic landing of aircraft. It is of some interest, therefore, to examine the probability distribution of the sum of two independent exponentially distributed variables. Statistical independence is assumed because it is assumed that the two noise sources are statistically independent.

Consider the system defined by

$$z = x + y \qquad \text{------- (59)}$$

where x and y are independent exponentially distributed variables, that is,

$$\left.\begin{aligned} p(x) &= \frac{1}{2\mu_1} \exp\left\{-\frac{|x|}{\mu_1}\right\} \\ p(y) &= \frac{1}{2\mu_2} \exp\left\{-\frac{|y|}{\mu_2}\right\} \end{aligned}\right\} \qquad \text{------- (60)}$$

Using equation 5, (or the method of characteristic functions), it can be shown that

$$p(z) = \frac{1}{2(\mu_1^2-\mu_2^2)}\left[\mu_1 \exp\left\{-\frac{|z|}{\mu_1}\right\} - \mu_2 \exp\left\{-\frac{|z|}{\mu_2}\right\}\right] \qquad (61)$$

Let $\mu_2 = n\mu_1$, $0 \leq n \leq 1$, and consider the case $\overline{z^2} = \sigma_z^2 = 1$

For these conditions p(z) is given by

$$p(z) = \frac{(1+n^2)^{\frac{1}{2}}}{\sqrt{2}(1-n^2)}\left[\exp\left\{-(1+n)^{\frac{1}{2}}\cdot\sqrt{2}.|z|\right\} - n\exp\left\{-\frac{(1+n^2)^{\frac{1}{2}}}{n}\cdot\sqrt{2}\,|z|\right\}\right] \qquad (62)$$

For the limiting cases n = 0, and n = 1,

$$n = 0,\ p(z) = \frac{1}{\sqrt{2}}\exp\left\{-\sqrt{2}.|z|\right\} \qquad (63)$$

$$n = 1,\quad p(z) = z\exp\left\{-2|z|\right\} + \tfrac{1}{2}\left\{\exp -2|z|\right\} \qquad (64)$$

The corresponding cumulative distributions $P_1(z)$ are

$$n = 0,\ P_1(z) = \tfrac{1}{2}\exp\left\{-\sqrt{2}\,z\right\},\quad z\geq 0 \qquad (65)$$

$$n = 1,\ P_1(z) = \tfrac{1}{2}(z+1).\ \exp\left\{-2z\right\},\quad z\geq 0 \qquad (66)$$

and for the general case

$$P_1(z) = \frac{1}{2(1-n^2)}\left[\exp\left\{-(1+n^2)^{\frac{1}{2}}\sqrt{2}\,z\right\} - n^2\exp\left\{-\frac{(1+n^2)^{\frac{1}{2}}}{n}\cdot\sqrt{2}z\right\}\right], z\geq 0 \qquad (67)$$

The distributions p(z) are shown in Figure 3 for n = 0,½,1, and $P_1(z)$ is shown in Figure 4 for the same values of n. The normal distribution is also included in these figures for comparison with the other results.

For most practical purposes the functions p(z) and $P_1(z)$ can be assumed to lie in the shaded regions of Figures 3 and 4. It can be seen from these figures that for all values of n the exponential distribution gives a much better approximation to p(z) than is given by the normal distribution, and for values of n less than about 0.5 the assumption of an exponential distribution for z gives results that are not seriously in error and are pessimistic, (that is, the values of p(z) given by the exponential distribution are greater than the true values of p(z) for large values of z).

2.5.1 <u>The probability distribution of the sum of two random variables with zero order Bessel function distributions</u>

From the preceding discussion if z = x+y

where x and y are statistically independent and

$$\left.\begin{aligned} p(x) &= \frac{1}{\pi\mu_1} K_o\left(\frac{|x|}{\mu_1}\right) \\ p(y) &= \frac{1}{\pi\mu_2} K_o\left(\frac{|y|}{\mu_2}\right)\end{aligned}\right\} \qquad (68)$$

If $\mu_2 = n\mu_1$, $n\leq 1$, it can be shown that for $\sigma_z^2 = 1$

$$\text{if } n\to 0,\ p(z) \to \frac{1}{\pi} K_o(|z|) \qquad (69)$$

$$\text{if } n\to 1,\ p(z) \to \frac{1}{\sqrt{2}}\exp\left\{-\sqrt{2}\,|z|\right\} \qquad (70)$$

and in the general case

$$p(z) \simeq \frac{1}{\sqrt{2}}\exp\left\{-\sqrt{2}|z|\right\} \text{ for } \tfrac{1}{2} \leq n \leq 1 \qquad (71)$$

2.6 <u>Discussion of Theoretical Results</u>

(i) It has been shown that for many forms of $p(\sigma_x)$, if the conditional probability $p(x|\sigma_x)$ is a normal distribution, the probability distribution p(x) is much closer to the exponential distribution than to the normal distribution. This is illustrated for some particular forms of $p(\sigma_x)$ in Figures 1 and 2.

(ii) It has been shown that for $p(\sigma_x)$ a Rayleigh distribution p(x) is exactly an exponential distribution of the form

$$p(x) = \frac{1}{2\mu}\exp\left\{-\frac{|x|}{\mu}\right\},$$

and for many distributions of $p(\sigma_x)$ of the general shape of the Rayleigh distribution, but considerably different in detail, it may be inferred that p(x) approximates closely to the exponential distribution given above for a wide range of x.

(iii) It has been shown that if σ_x is proportional to a random variable z, and z is given by

$$z^2 = a^2 + b^2,$$

where a and b are statistically independent normally distributed random variables with zero mean values with $\sigma_b = n\sigma_a$, p(z) has limiting forms of a Rayleigh distribution for $n\to 1$ and a one sided normal distribution for $n\to\infty$, and $p(\sigma_x)$ has similar limiting forms.

(iv) Under the conditions of item (iii) it has been shown that p(x) (for the unit variance case) will lie in the shaded regions of Figures 1 and 2, and for n in the range $1 \le n \le 2$, p(x) is approximately an exponentially distribution.

(v) The theoretical results indicated that the distribution $p(x) = \frac{1}{\pi} K_0(|x|)$ can be taken as an upper limit for p(x) in many situations of the type described in sections (iii) and (iv). However, the values of p(x) for large values of $|x|$ will tend to be reduced by the fact that if the noise input is very large the vehicle containing the system being considered will either avoid the regions containing this noise, (for example, aircraft will tend to avoid hurricanes and similar phenomena), or the noise inputs will not be used for automatic control purposes, (for example, on excessively noisy radio beams pilots will tend to use manual control). For these reasons it is suggested that the exponential distribution can be used to provide a limiting distribution for p(x) for large values of $|x|$ in most practical cases for automatically controlled systems.

(vi) In most real situations many noise inputs are present, but in many cases one or two noises will predominate. In these cases it has been shown that the exponential distribution provides a useful approximate limiting distribution for the system output state variables in many practical situations.

(vii) In well designed systems the effects of tolerances on the statistics of the output state variables are usually small compared with the effects due to the noise inputs. However, in the case of multiplexed systems, which are often used to improve the overall safety statistics of manned aerospace systems, the tolerances have an important effect on nuisance cut-outs. It has been shown that the average rate of nuisance cut-outs for simple duplex systems with tolerances can be approximated by an exponential function under certain conditions. The instantaneous probability of cut-out for these systems has been shown to approximate to an exponential distribution. It is suggested that the exponential distribution can provide a useful tool in the assessment of the cut-out statistics for many practical systems.

(viii) The application of the exponential distribution to the rate and duration of the exceedances of a level x = c has been described.

(ix) The restrictions of the noise models used, particularly the restriction imposed by equations 1 and 2, are not considered to be of great significance. For example, although references 7 and 8 show that the form of the power spectral density of atmospheric turbulence is not constant, the variations in form (that is variations in the form of $\phi(\omega)$ in equations 1 and 2) are not great, and it is common practice in simulation exercises to assume a constant form, (for example, the von Karman or Dryden spectra).

The assumptions used here are equivalent to specifying that noise is generated by a constant physical process which defines a constant form of power spectral density, but the gain factor c_n^2 in equation 1 varies as discussed in this paper.

(x) Under the conditions assumed here it is only necessary to know the ratio rms output/rms noise input to apply the methods of this paper to practical situations This ratio may be obtained from test results or may be calculated, using well known methods from the parameters of the system transfer function, (see, for example, Chapter 5 and Appendix E of reference 5).

3. SOME PARTICULAR NOISE MODELS

3.1 Atmospheric Turbulence

Reference 6 gives the following relationship between rms turbulence intensity and reported mean wind speed.

$$\left.\begin{aligned} \sigma_u &= 0.18U \\ \sigma_v &= 0.09U \end{aligned}\right\} \quad \text{------- (72)}$$

where U = reported mean wind speed 10 metres above ground level,

σ_u = rms intensity of the horizontal component of turbulence.

σ_v = rms intensity of the vertical component of turbulence.

Extrapolating the results given in equation 72,

$$\left.\begin{aligned} \sigma_u &= K_1 U \\ \sigma_v &= K_2 U \end{aligned}\right\} \quad \text{------- (73)}$$

where the values of K_1 and K_2 given in equation 72 apply to altitudes from 3 to 100 metres above ground level. K_1 and K_2 will be functions of altitude and terrain conditions, and $K_1 \to K_2$ as homogeneous isotropic conditions are approached at large values of height above ground level.

If it is assumed that the spectral densities of turbulence are as defined by equations 1 and 2, at any particular height or small variation of height, K_1 and K_2 can be assumed to be constants. Under these conditions it is apparent that the probability distributions of σ_u and σ_v will be of the same type as that of U. The probability distribution of U is discussed in the following section.

3.1.1 A Wind Model

The cumulative probability distributions of total wind and cross wind amplitudes during the aircraft landing manoeuvre are given in reference 6. The relevant information is given here in figure 5. Meteorological Office data for airports at London, Manchester and Renfew are also plotted in figure 5, and it can be seen that the data for British airports agrees very well with that of the general data of reference 5, (which is based on world-wide in-service operations of British airlines).

An examination of figure 5 shows that the probability distribution of cross-wind can be closely approximated by a normal distribution with a standard deviation of 7 knots and zero mean value, and that of the magnitude (modulus) of total wind by a Rayleigh distribution with a rms value of

$8\sqrt{2}$ knots, that is,

$$p(U_c) = \frac{2}{\sqrt{2\pi}} \cdot \frac{1}{\sigma_{uc}} \exp\left\{-\frac{U_c^2}{2\sigma_{uc}^2}\right\}, \quad \sigma_{u_c} = 7, \qquad \text{------- (74)}$$

$$p(U) = \frac{U}{\mu^2} \exp\left\{-\frac{U^2}{2\mu^2}\right\}, \quad \mu^2 = 64, \qquad \text{------- (75)}$$

where U_c = modulus of crosswind amplitude

U = modulus of total wind amplitude

(Units: knots).

If U_R = modulus of component of wind along runway, then

$$U^2 = U_R^2 + U_c^2 \qquad \text{------- (76)}$$

The results of section 2.2 suggest that a possible wind model can be derived by making the following assumptions:

(i) $$p(U_c) = \frac{2}{\sqrt{2\pi}\sigma_{uc}} \exp\left\{-\frac{U_c^2}{2\sigma_{uc}^2}\right\}$$

(ii) $$p(U_R) = \frac{2}{\sqrt{2\pi}\sigma_{uR}} \cdot \exp\left\{-\frac{U_R^2}{2\sigma_{uR}^2}\right\}$$

where $\sigma_{U_c} = 7$, $\sigma_{U_R} \approx 9.143$,

so that $\sigma_{u_c}\sigma_{u_R} = 64$, $\frac{\sigma_{uR}}{\sigma_{uc}} = 1.306$.

With the above assumptions, from section 2.2. p(U) is given by

$$p(U) = \frac{U}{\sigma_{u_c}\sigma_{u_R}} \exp\left\{-\frac{U^2}{2}\left(\frac{1}{2\sigma_{u_c}^2} + \frac{1}{2\sigma_{u_R}^2}\right)\right\} \cdot I_o\left\{\frac{U^2}{2}\left(\frac{1}{2\sigma_{u_c}^2} - \frac{1}{2\sigma_{u_R}^2}\right)\right\} \qquad \text{-----(77)}$$

This function is plotted in figure 6 and is shown in this figure to approximate closely to p(U) of figure 5 and equation 75 over the range of wind speeds considered in figure 5.

3.1.2 Turbulence

If $\sigma = KU$

where σ^2 = variance of the component of turbulence in a vertical or horizontal plane, and the probability density function of U is given by p(U), then

$$p(\sigma) = \frac{1}{K} p_u\left(\frac{\sigma}{K}\right) \qquad \text{------- (78)}$$

It has been shown that

$$p(U) \doteq \frac{U}{64} \exp\left\{-\frac{U^2}{128}\right\} \qquad \text{------- (79)}$$

U knots.

Therefore

$$p(\sigma) \doteq \frac{\sigma}{\mu^2} \exp\left\{-\frac{\sigma^2}{2\mu^2}\right\} \qquad \text{------- (80)}$$

where $\mu^2 = 64K^2$

For σ_u in feet per second

$$p(\sigma_u) \approx \frac{\sigma_u^2}{182.4K^2} \exp\left\{-\frac{\sigma_u^2}{364.8K^2}\right\} \qquad \text{------- (81)}$$

For rough terrain and smooth terrain values of K of 0.2 and 0.12 respectively have been suggested by Aero Flight, Royal Aircraft Establishment, Bedford, (for conditions during aircraft landing manoeuvres). $p(\sigma_u)$ of equation 81 for K = 0.12, 0.2 and 0.18 are plotted in figure 7.

3.2 A Radio Noise Model

It is possible to use an argument similar to that used in sections 3.1.1 and 3.1.2 to suggest that the probability distributions of the standard deviations of the radio noise from ILS beams approximate to Rayleigh distributions. However, the available evidence is not sufficient, at this time, to support this line of reasoning.

Figure 8, distributions 1 and 2, shows two distributions which have been used in simulation studies for the distribution of radio noise on localiser beams, and the results have been shown to give reasonable agreement with flight test results.

Figure 8 also shows a Rayleigh distribution (with μ = 1). From the results of section 2.1 it appears to be reasonable to use distribution 3 of figure 8 as a model for the radio noise on localiser beams. The results for p(x) will be slightly more pessimistic than those of distributions 1 and 2 of figure 8 for large values of x, but the overall results should be in reasonable agreement with those obtained in practical situations.

It is suggested that a similar model of the form $p(\sigma_N) = \frac{\sigma_N}{\mu^2} \exp\left\{-\frac{\sigma_N^2}{2\mu^2}\right\}$, σ_N = rms noise level, with suitable values for μ, can be used for many types of radio noise, in particular the radio noise on ILS localiser and glide path beams. This model will, of course, lead to exponential probability distributions for the output state variables.

3.3 A General Noise Model

Particular noise inputs will have probability distributions depending on the type of input being considered. However, the wide occurrence of exponential distributions observed in a variety of practical situations suggests that a noise model on the lines of section 3.1.2 could fit possible practical situations.

In this model the noise power, specified by its mean squared value, is assumed to be proportional to the power in some variable z, where z is formed by the sum of two independent orthogonal vectors with normally distributed amplitudes and zero mean values. In many situations the physical conditions can be expected to be such that the standard deviations of the two orthogonal components of z are approximately equal. This will lead to (approximately) exponential distributions for the system output state variables.

3.4 Discussion of Noise Models

The results of sections 3.1.1, 3.1.2, 3.2 and 3.3 indicate that a Rayleigh distribution of σ_N is approximated by a wide variety of noises.

This conclusion agrees well with the atmospheric turbulence of the type defined by reference 6, and also with the radio noises frequently used in simulation exercises. In addition to the agreement obtained in these cases, the general shape of the Rayleigh distribution is intuitively very attractive, and further work to establish its theoretical justification in other particular noise cases would seem to be desirable.

4. PRACTICAL RESULTS

4.1 Noise Inputs

Some results of flight trials carried out by Hawker Siddeley Aviation Limited are shown in Figures 10 to 14. It can be seen that the distributions of all state variables shown in these figures approximate to exponential distributions of the form

$$p(y) = \frac{1}{2\mu} \exp\left\{-\frac{|y|}{\mu}\right\}$$

where $y = x - \bar{x}$, for the state variable x with mean value $\bar{x}$. The quantity $\bar{x}$ is due to bias, steady wind, or datum axes, the effects of which are not considered in this paper, but which can be calculated or measured by well known methods. All state variables mentioned in earlier sections refer to random variables with zero mean value (such as the variable y in the above expression).

Figure 9 gives a typical result for an output state variable with a stationary Gaussian noise input. This figure shows that with an input with stationary Gaussian properties the output distribution approximates closely to the normal distribution (which gives a straight line on the arithmetical probability graph paper used for the figures of this section). This fact supports the assumption used throughout this paper that the controlled aircraft can be assumed to be a linear system for the purposes of the work described here.

NOTE: The results shown in figures 9-14 were obtained over a long period during development phases, and do not necessarily apply to the current version of the aircraft and control system.

4.2 Tolerances

The results of two experiments carried out to investigate the effects of tolerances in a duplex system are shown in Figures 15 and 16. The experiments consisted of a large numbers of simulator runs (about 2000 for each experiment) with stationary Gaussian noises (atmospheric turbulence and radio noise) applied to a simulated automatically controlled aircraft with selected parameters in the duplex control system varied in a manner corresponding to their expected statistical properties.

In experiment B seven parameters were varied, and in experiment A fourteen parameters were varied. In each case the varied parameters were assumed to have rectangular probability distributions. Particular tolerance variations were not the same in the two experiments.

Figure 15 shows the probability distributions of the standard deviations of error (the difference between the outputs of the two sub-channels). It can be seen that the result for experiment A are not very different from the corresponding result for a Rayleigh distribution with $\mu = 0.35$, but the result for experiment B are considerably different from the corresponding Rayleigh distribution with $\mu = 0.48$. In each case the values of μ were selected from the mean of μ_1 and μ_2 where

$$\mu_1 = \sqrt{\frac{2}{\pi}}\ \bar{\sigma}_e$$

$$\mu_2^2 = \frac{\overline{\sigma_e^2}}{2}$$

In both experiments μ_1 and μ_2 were approximately equal (within about 10%).

The probability distributions of the error signal (x) were calculated from equations 1 and 2 using the experimental results and the results obtained by assuming Rayleigh distributions for σ_e. The instantaneous probabilities of cut-out for cut-out settings l (l = 0 to 6 degrees) were calculated from instantaneous probability of cut-out = p c. = $2\int_c^\infty p(x)\,dx$. The results for the experimental results are shown by curves A and B of Figure 16. The corresponding results assuming Rayleigh distributions for σ_e are shown by curves A_1 and B_1 of this figure.

It can be seen that the function p.c. for experiment A is not very different from the exponential distribution expected for curve A_1 from the results of previous sections. The curve B is not very close to the exponential curve B_1, but this may be expected due to the comparatively small number of tolerances considered in experiment B - as predicted in section 2.3.1.

Although the results given in this section are not very conclusive they do support the general arguments given in sections 2.3 and 2.3.1 for linear systems with a large number of tolerances and a transfer function of low order.

4.3 Wind and Turbulence

References 7 and 8 contain a large quantity of useful information on low altitude wind and turbulence obtained from practical measurements.

The results of these references will not be discussed in detail here, but the following points are noted:

(i) At first sight the results contained in References 7 and 8 do not appear to agree with the results of this paper. However, further investigations show that if the results for low altitudes over high and low mountains are ignored, good qualitative agreement is obtained.

(ii) In spite of the fact that figures 5.66 - 5.73 of Volume 1 of Reference 8 show that the correlation between rms turbulence and windspeed magnitude is small, figure 5.45 of the same reference shows that, except for large rms gust velocity values, the probability distribution of rms gust velocity is approximately Rayleigh (with $\mu \doteq 2$ ft/second). It can be seen from the other data in Reference 8 that the probability distribution of rms gust velocity is distorted, so that $p(\sigma)$ is greater than the corresponding Rayleigh distribution for large σ, by the effects of mountains.

(iii) Figures 5.11 - 5.13 of Volume 1 of Reference 8 show that the probability distributions of gust velocity amplitude are approximately exponential for all cases except those for low altitudes over mountains. This conclusion is also supported by figures V-16 to V-18 of Volume 2 of Reference 8.

(iv) In general, the results of References 7 and 8 support the forms of wind and turbulence models suggested in this paper for all except thecases over mountainous regions, although the parameters of the probability distributions tend to be rather different from those suggested here - as may be expected from the limited geographical regions in which the LO-LOCAT data were obtained.

(v) The disagreement between the results of this paper and the low altitude (750 feet and below) results for mountainous regions obtained in References 7 and 8 is not considered to be significant for most airline operations, because low altitude flying over mountains is not common for airline operations. However, this disagreement should be remembered in the case of military aeroplanes which commonly use low altitude terrain following for radar detection avoidance.

5. CONCLUSIONS

(i) It has been shown that the probability distributions of the output state variables of linear or almost linear systems, with various non-stationary noise inputs likely to be met in practical systems, can be approximated by exponential distributions over wide ranges of the variables.

(ii) It has been shown that the exponential distribution can provide a useful and valid tool in calculations used to assess the performance statistics of systems with noise inputs and with system tolerances.

(iii) The exponential distribution has been shown to apply to some practical results obtained from flight trials and simulation for a modern aircraft.

(iv) It is suggested that the use of the results of this paper in practical calculations can give more realistic results than those obtained from the commonly used assumption of normal distributions for the system output state variables. This is considered to be of major importance in calculations involving safety aspects.

(v) The most obvious application for the use of the results of this paper are for preliminary calculations of the statistics of the output state variables. Some other possible applications are:

- (a) the extrapolation of experimental results,
- (b) calculations involved inthe estimation of accident or incident rates - both in preliminary calculations and in calculations based on experimental results,
- (c) as a design tool to avoid some of the late system changes resulting from the less realistic results obtained from methods based wholly on the normal distribution.

(vi) It is apparent from Reference 1 that the results of this paper also apply to many cases of manual control, but due to the eccentricities and non-linearities associated with human operators, larger departures from the exponential distribution may be expected in these cases.

(vii) The arguments and discussions in this paper are based on work carried out in connection with the automatic landing of aircraft, and are essentially of a practical nature. It is apparent that further work is desirable to provide a more rigorous mathematical foundation for the subject.

(viii) In spite of the lack of mathematical rigour mentioned in (vii) it is suggested that the practical results of section 4 justify the use of the exponential distribution, in the manners described above, for a wide variety of problems in the field of aerospace systems.

(ix) The work leading to this paper was started as the result of Wing Commander E.W. Anderson's famous question, "is the Gaussian distribution normal?" (Reference 1) Although the negative answer to this question has certainly not been proved here, it is hoped that some doubt has been cast upon the affirmative answer, at least as far as linear aerospace systems are concerned, and that the exponential distribution has some right to the title "normal" in these cases.

ACKNOWLEDGEMENTS

The author wishes to thank Hawker Siddeley Aviation Limited for permission to use the test data shown in Figures 9 to 14 inclusive.

REFERENCES

1. Anderson E.W. (1965). Is the Gaussian distribution normal? J. Inst. Nav., Vol 18, p 65.

2. Anderson E.W, and Ellis D.M. (1971). Error distributions in Navigation. J. Inst. Nav., Vol 24, p 429.

3. Rice S.O. Mathematical Analysis of random noise. Selected papers on noise and stochastic processes, Ed. Wax, Dover, 1954.

4. Rice S.O. Distribution of the Duration of Fades in Radio Transmission - Gaussian Noise Model, B.S.T.J., Vol 37, No. 3, May 1958.

5. Lanning J.H., and Battin R.H. Random Processes in Automatic Control, Mc Graw-Hill, 1956.

6. Air Registration Board, Airworthiness requirements for Automatic Landing including automatic landing in restricted visibility down to Category 3 British Civil Airworthiness Requirements, Paper No. 367, Issue 3, 1970, (Working draft).

7. Monson K.R, Jones G.W., Mielke, R.H, et al. Interim Analysis of Low Altitude Atmospheric Turbulence (LO-LOCAT) Data. Tech. Report ASD-TR-69-7, Aeronautical Systems Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, February 1969.

8. Authors and source as for Reference 7. Low Altitude Atmospheric Turbulence LO-LOCAT Phase III Interim Report. Tech Report AFFDL-TR-69-63 Volumes 1 and 2, October 1969.

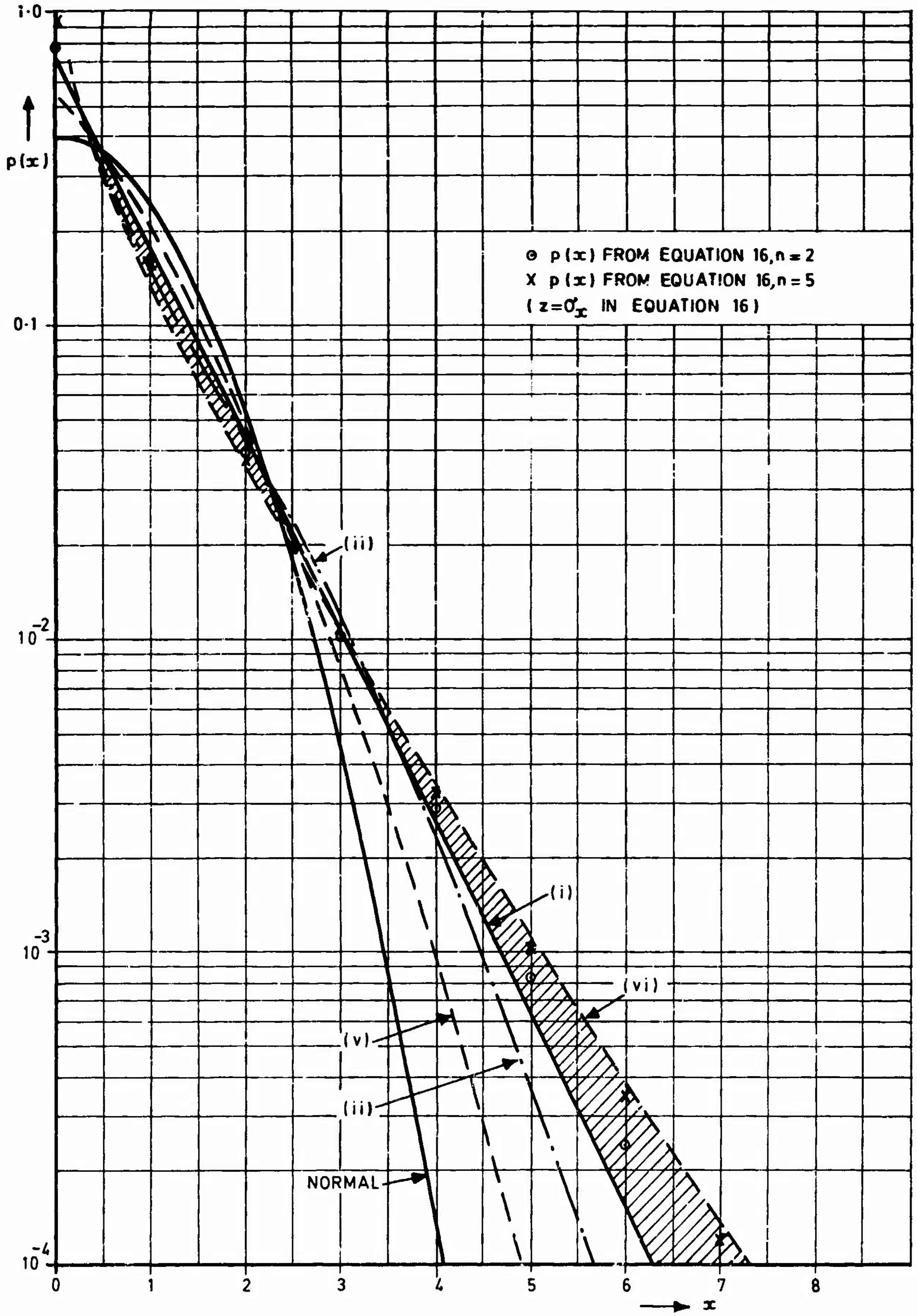


Fig.1 Probability distributions from Table 1 (numbering corresponds to item numbers in Table 1)

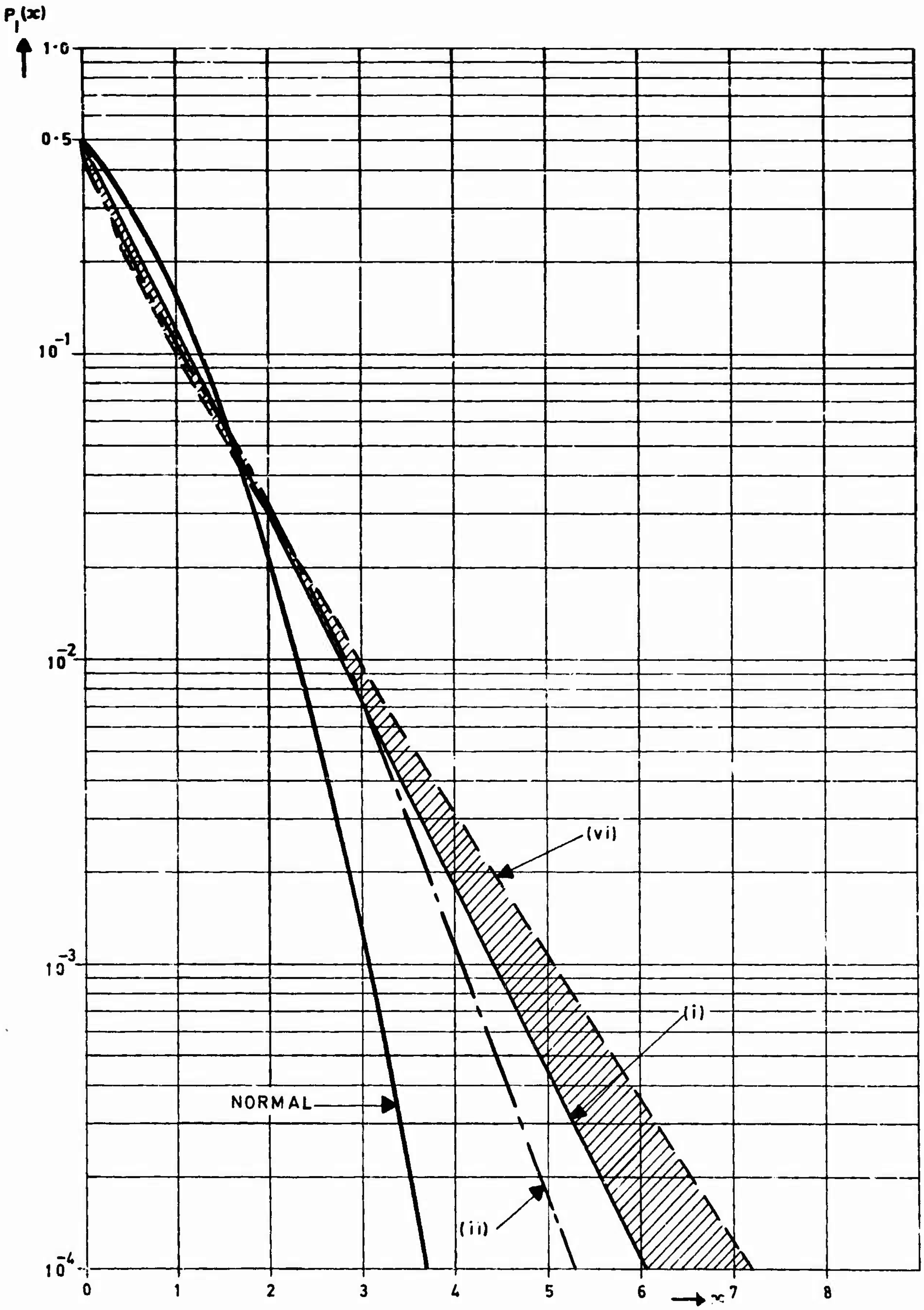


Fig.2 Cumulative probability distributions

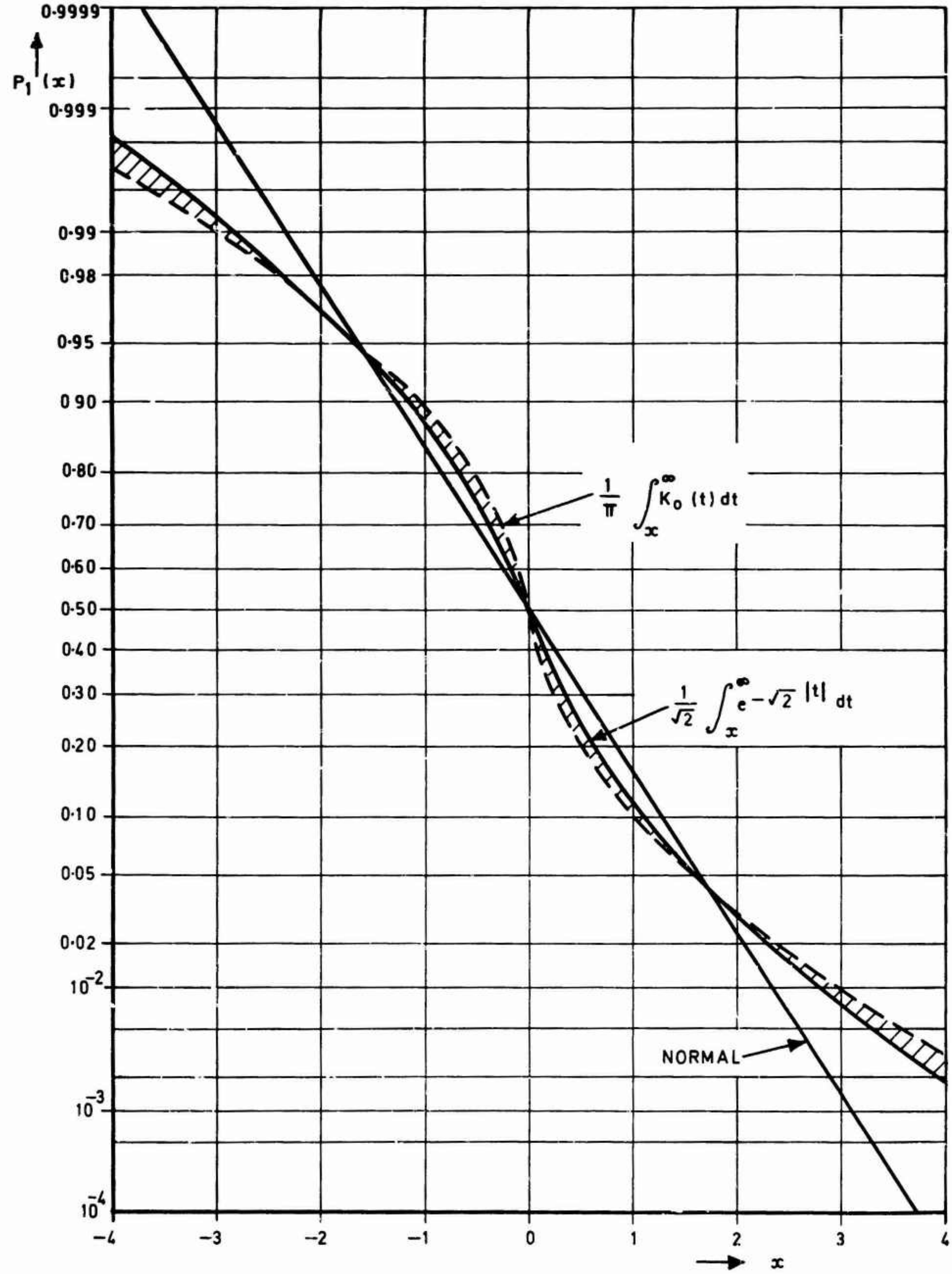


Fig.2(a) Cumulative probability distributions

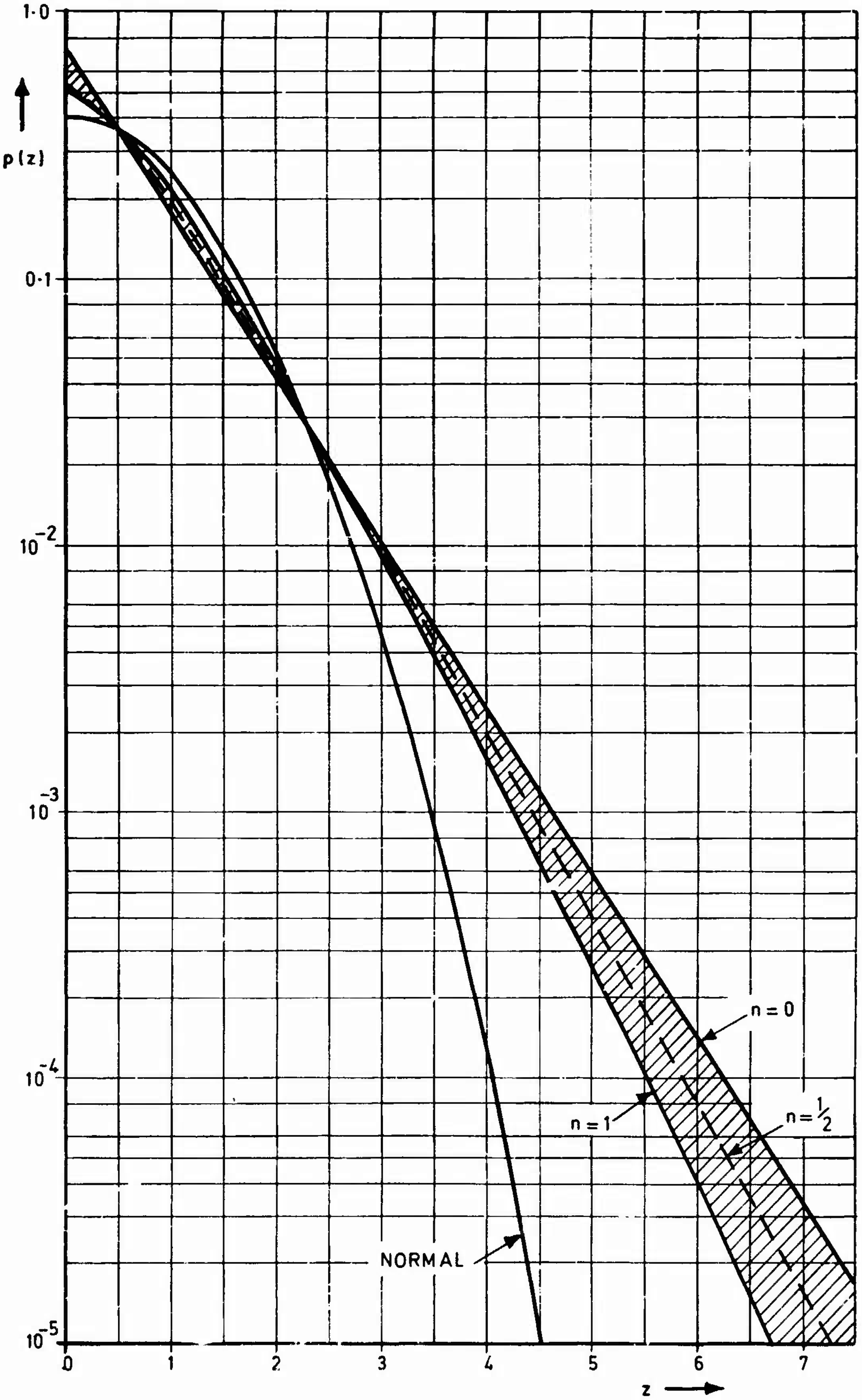


Fig.3 Probability distributions, sum of two exponential distributions

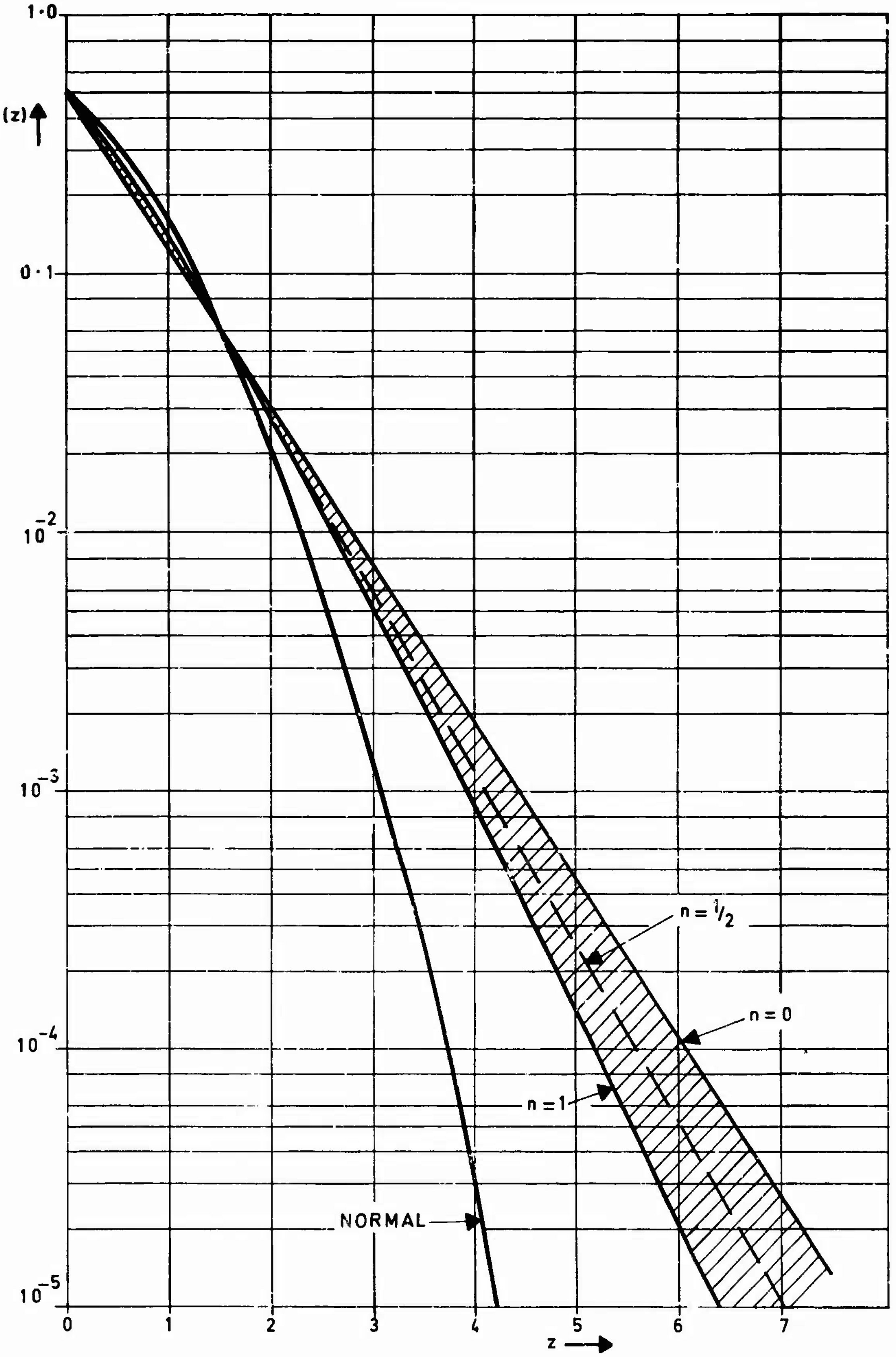


Fig.4 Cumulative distributions sum of two exponential distributions

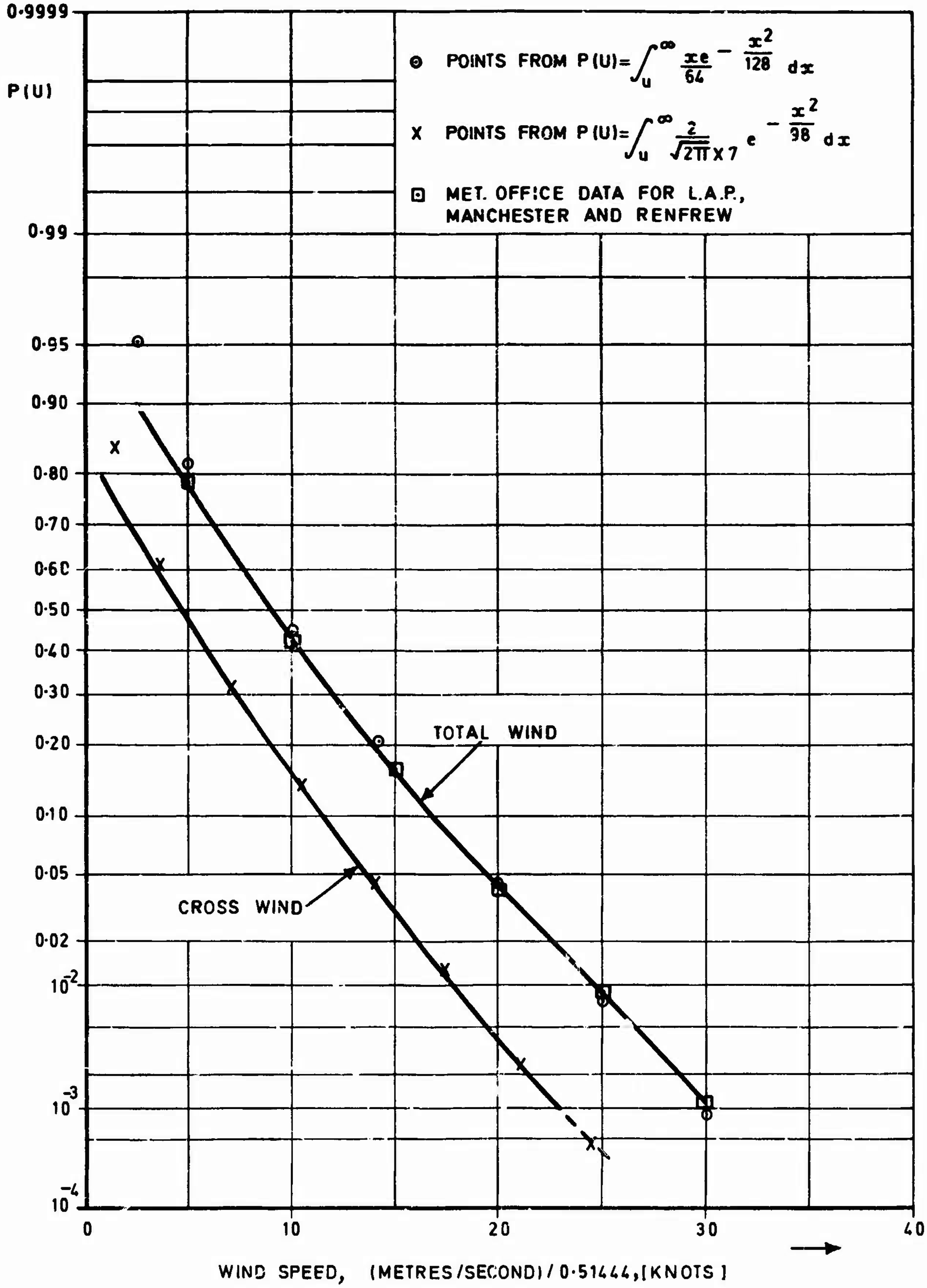


Fig.5 Cumulative probability of reported mean wind and cross wind when landing

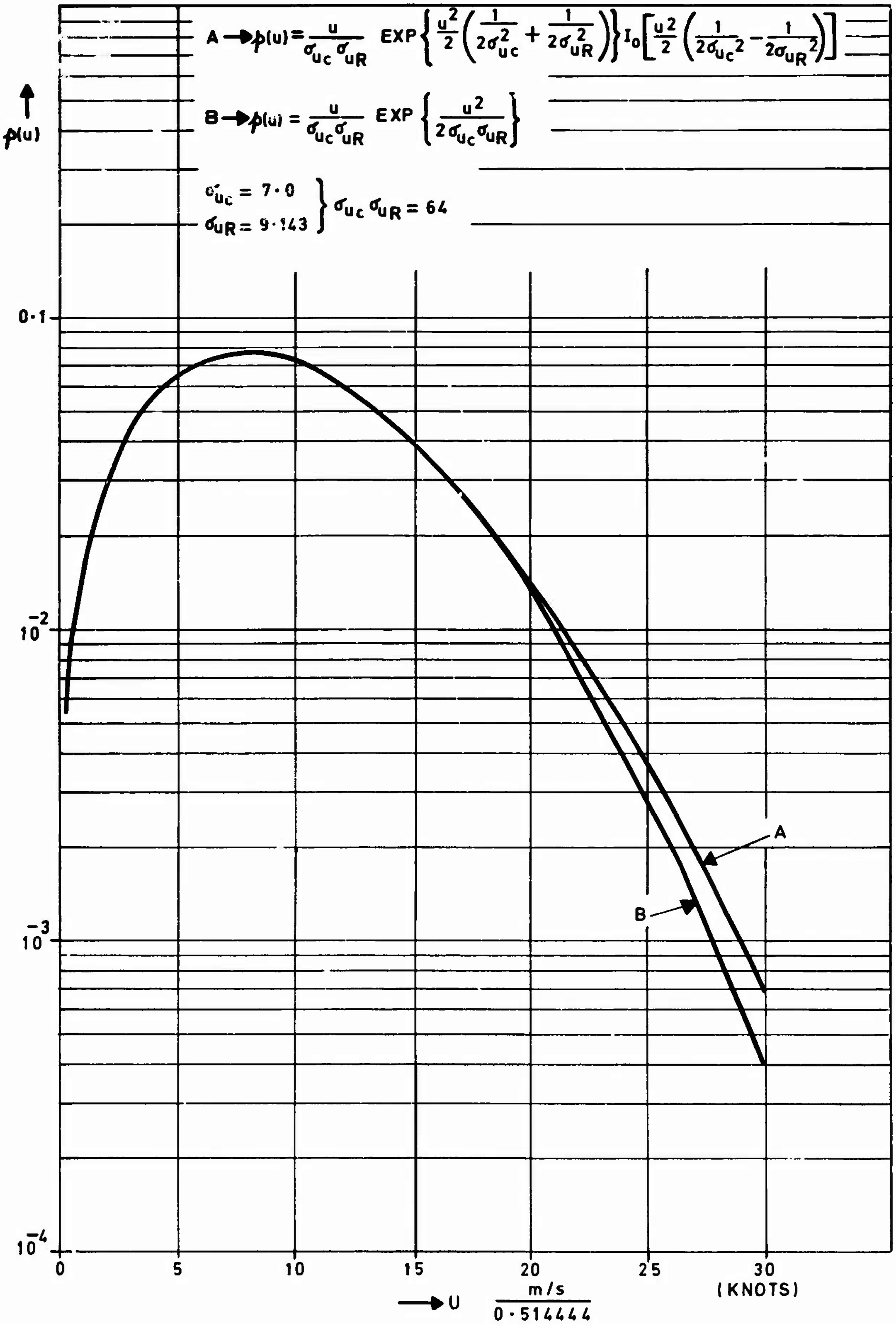


Fig.6 Probability density function of total windspeed

$$p(\sigma_u) = \frac{\sigma_u}{\mu^2} e^{-\frac{\sigma_u^2}{2\mu^2}} = \frac{\sigma_u}{K^2\mu_1^2} e^{-\frac{\sigma^2}{2K^2\mu_1^2}}, \quad \mu_1^2 = 182 \cdot 4$$

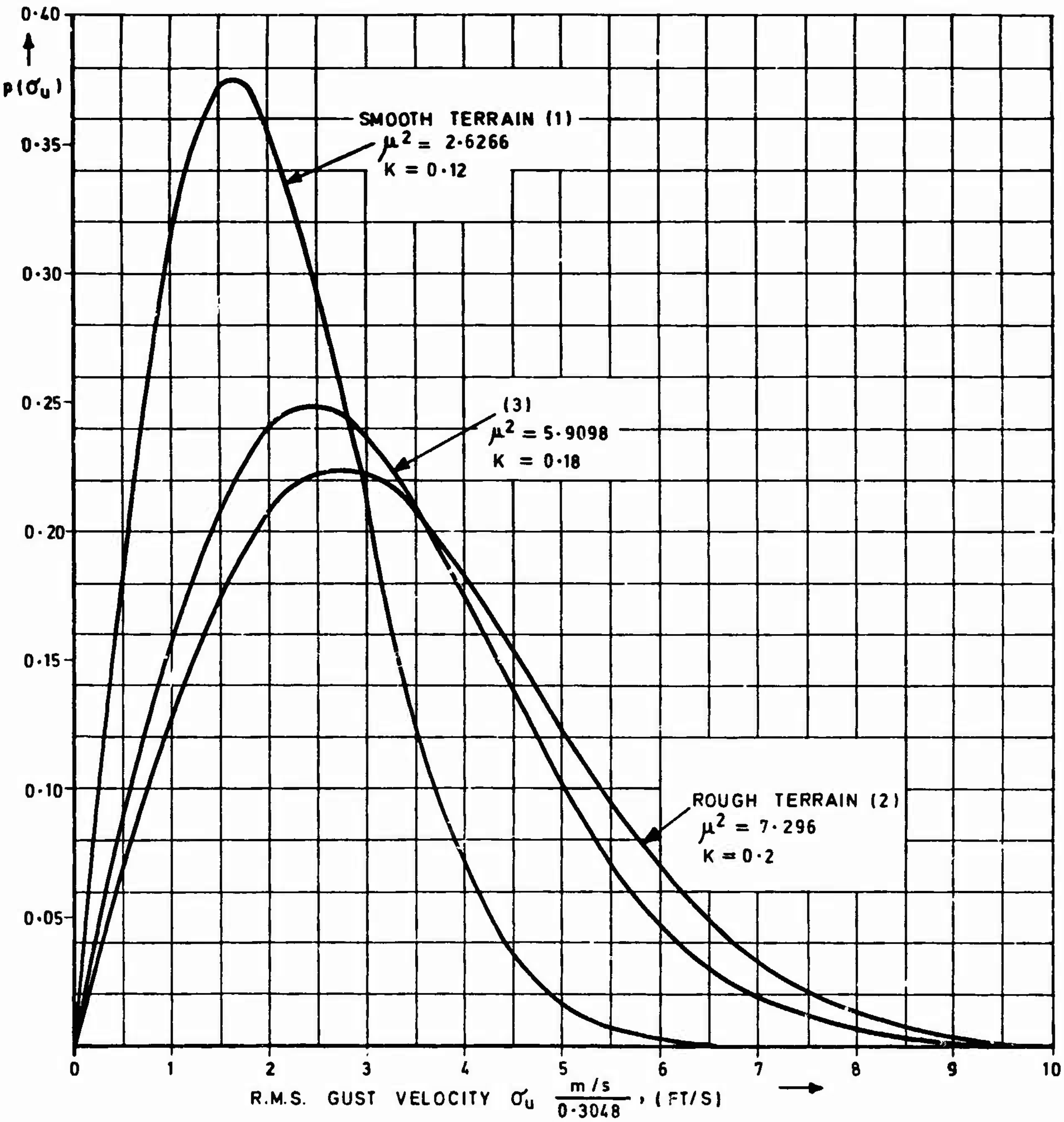


Fig.7 Probability density functions for horizontal gusts

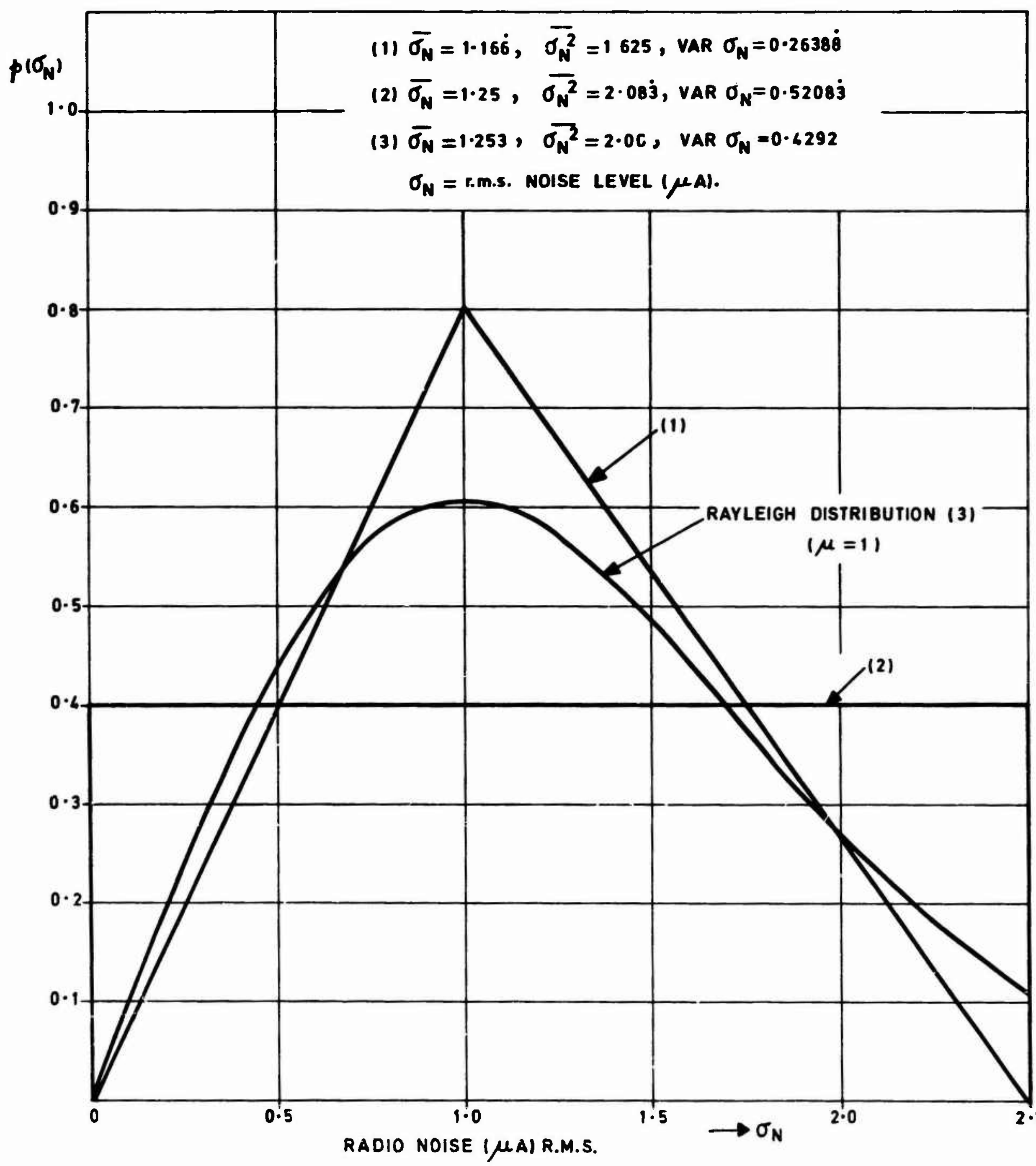


Fig.8 Probability densities for radio noise

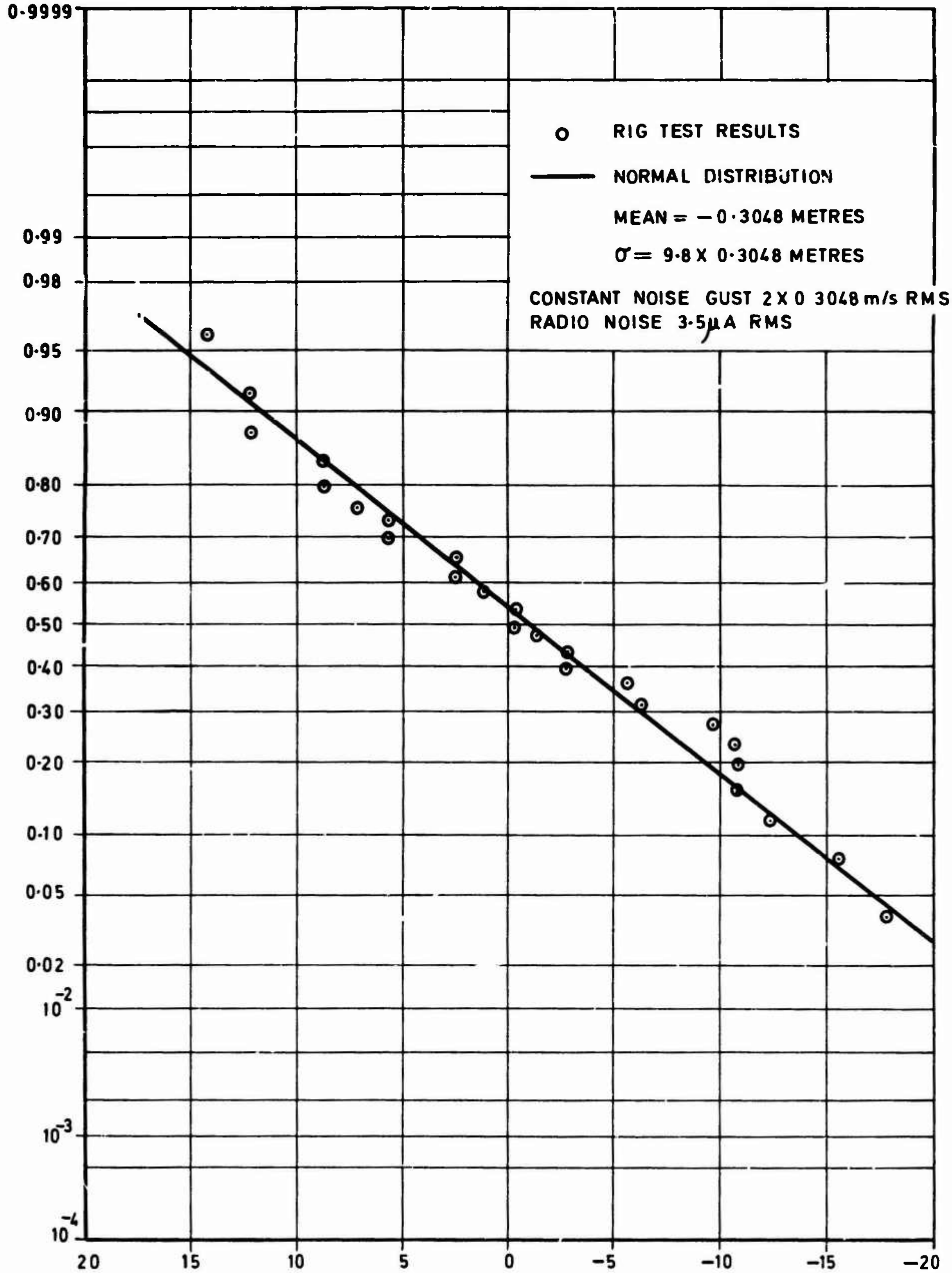


Fig.9 Displacement from runway centre line at 152.4 metres (500 feet) from touchdown, meters, (feet). (Rig test results) 0.3048
0.3048

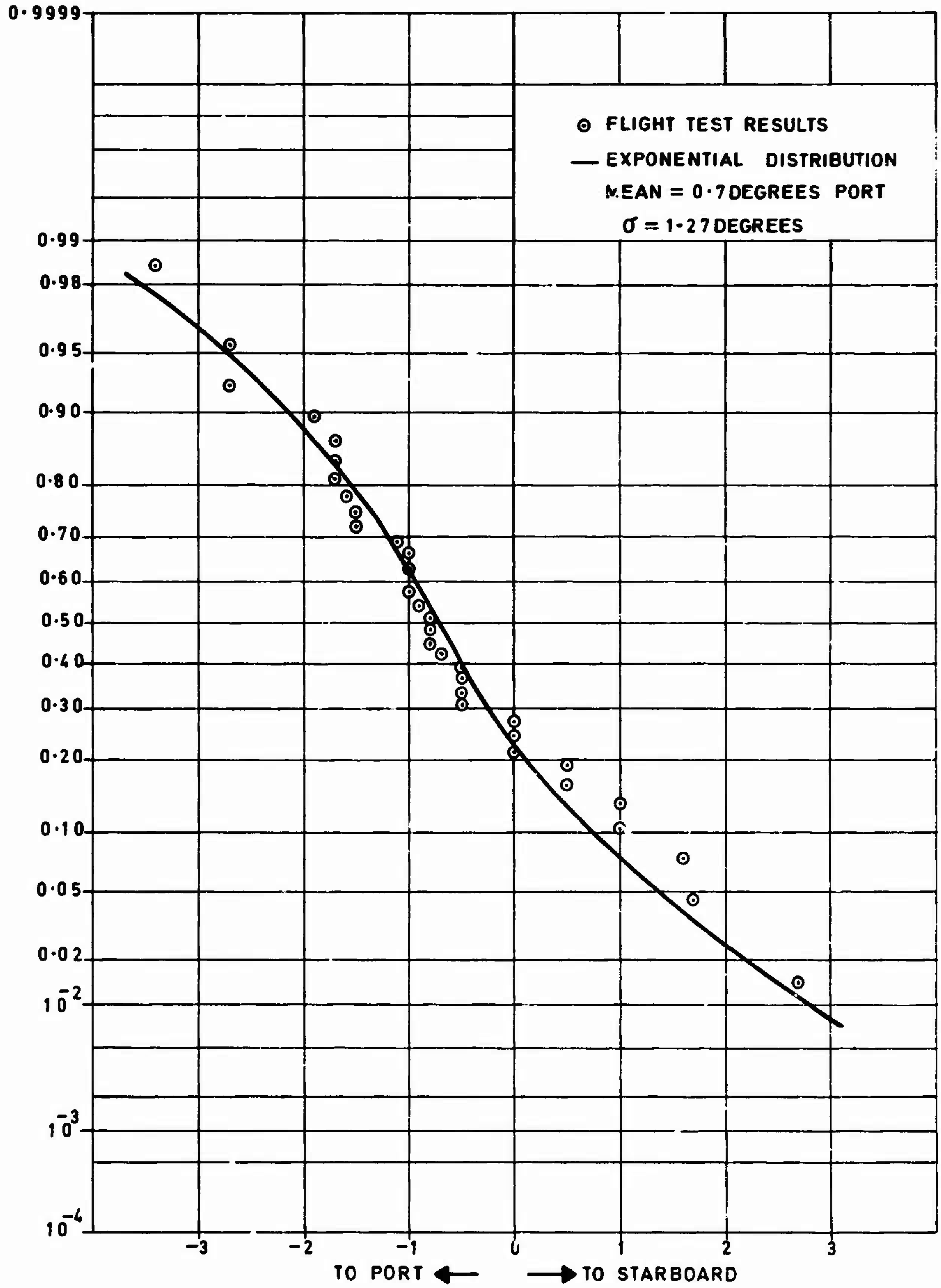


Fig.10 Bank angle at touchdown (degrees)

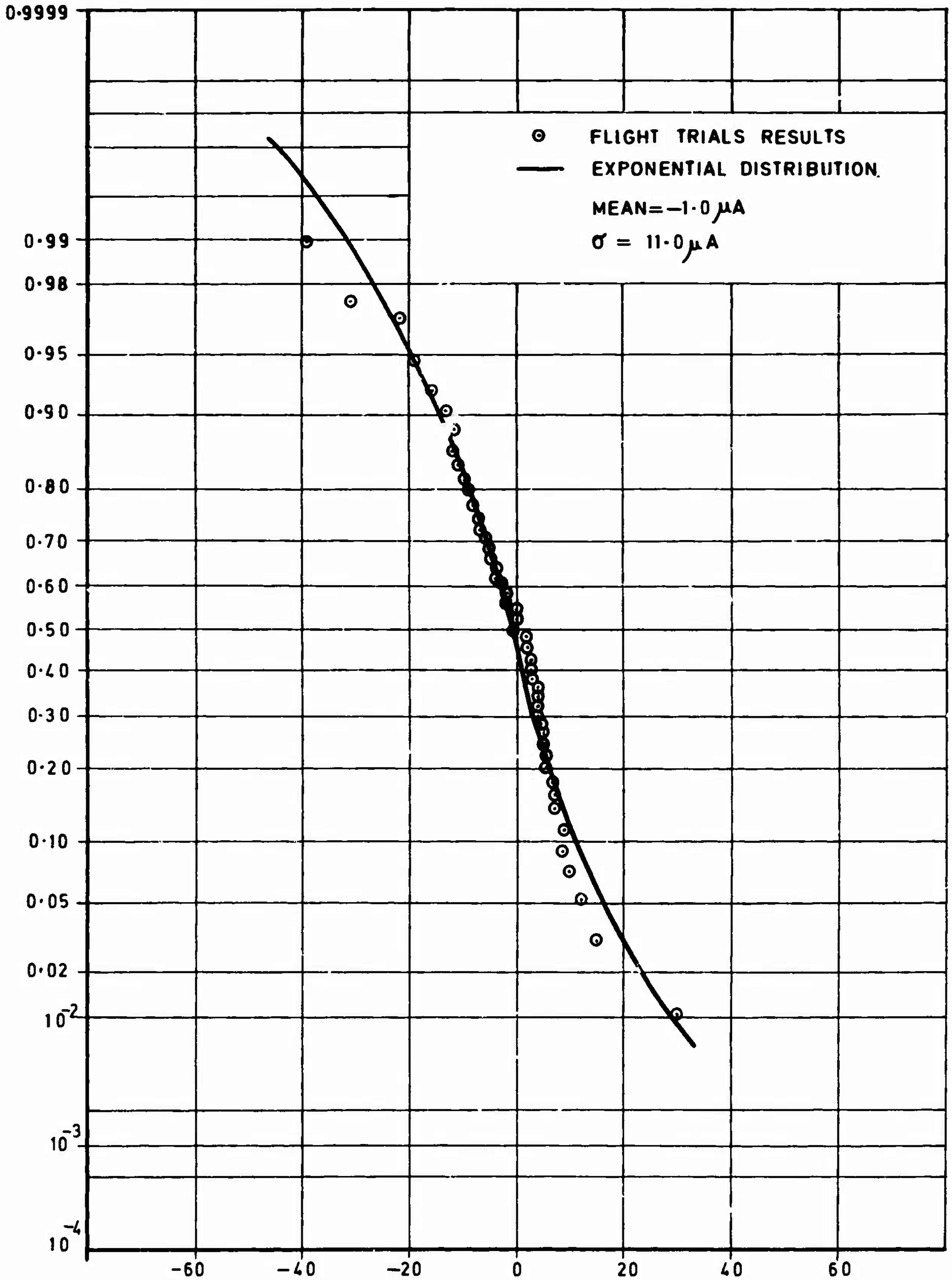


Fig.11 Gude path deviation at 97.536 metres (320 feet), μA

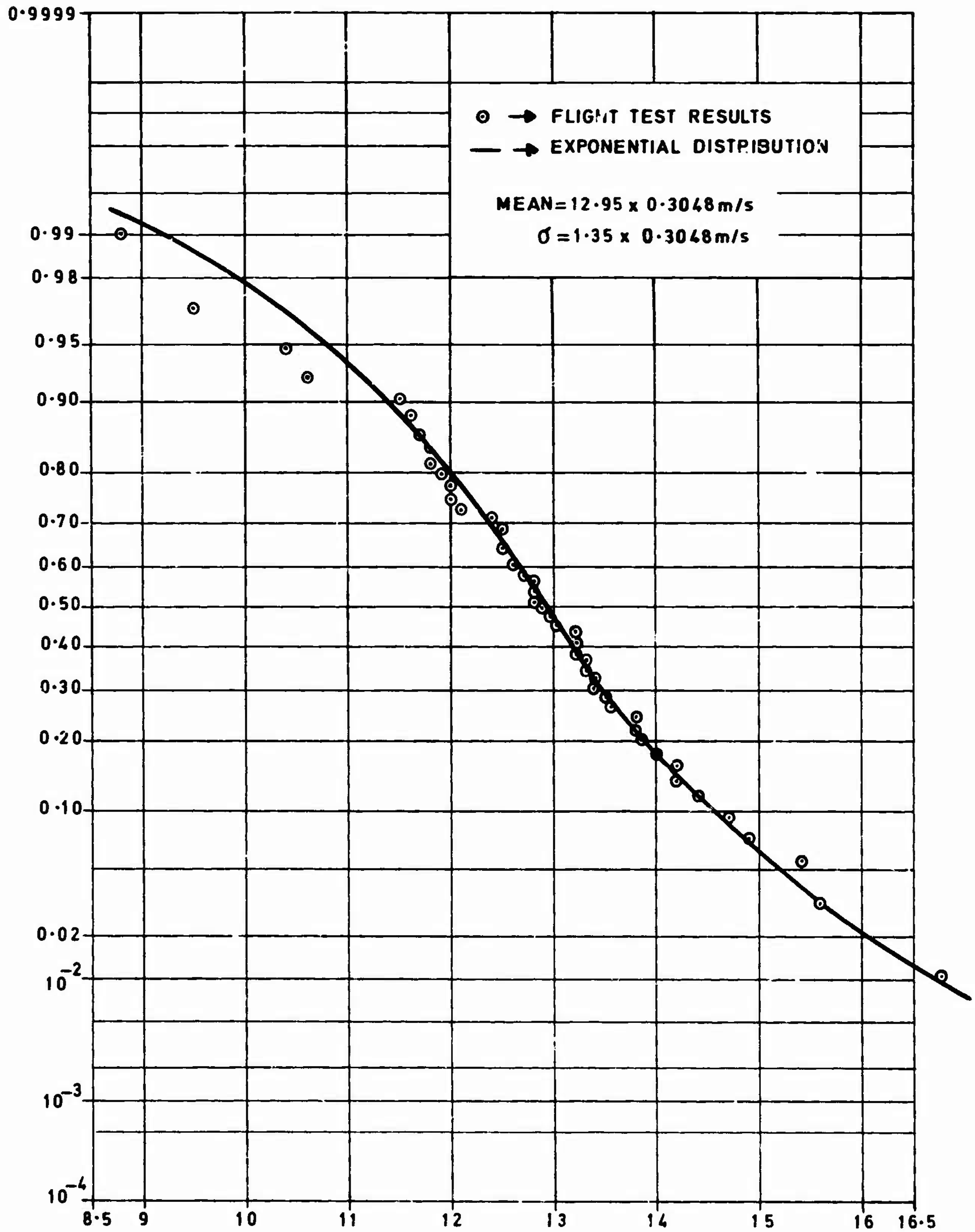


Fig.12 Rate of descent at start of flare, $\frac{m/s}{0.3048}$ (ft/s)

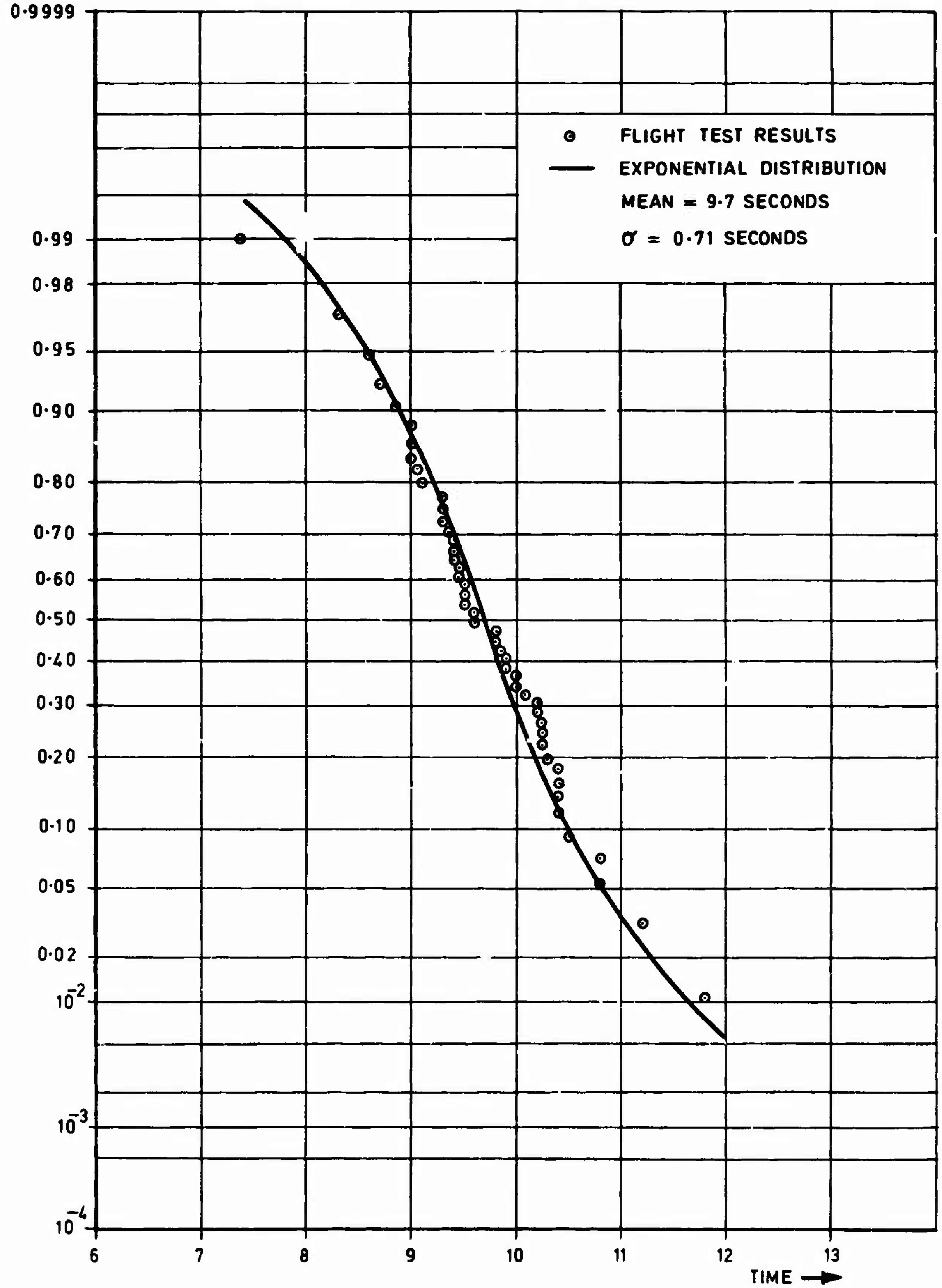


Fig.13 Time 21.336 metre (70 feet) to touchdown, seconds

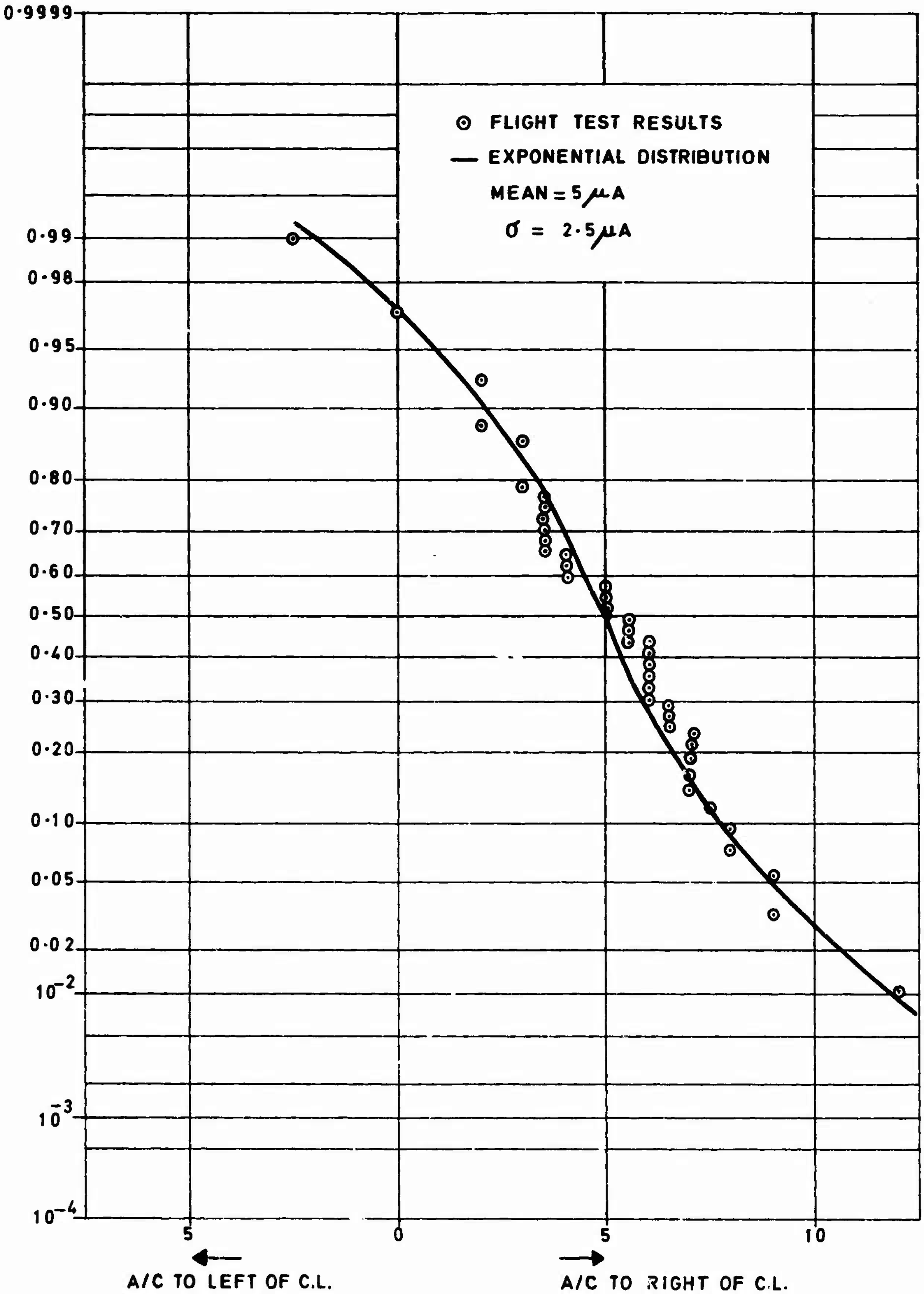


Fig.14 Localizer deviation at 21.336 metres, μA

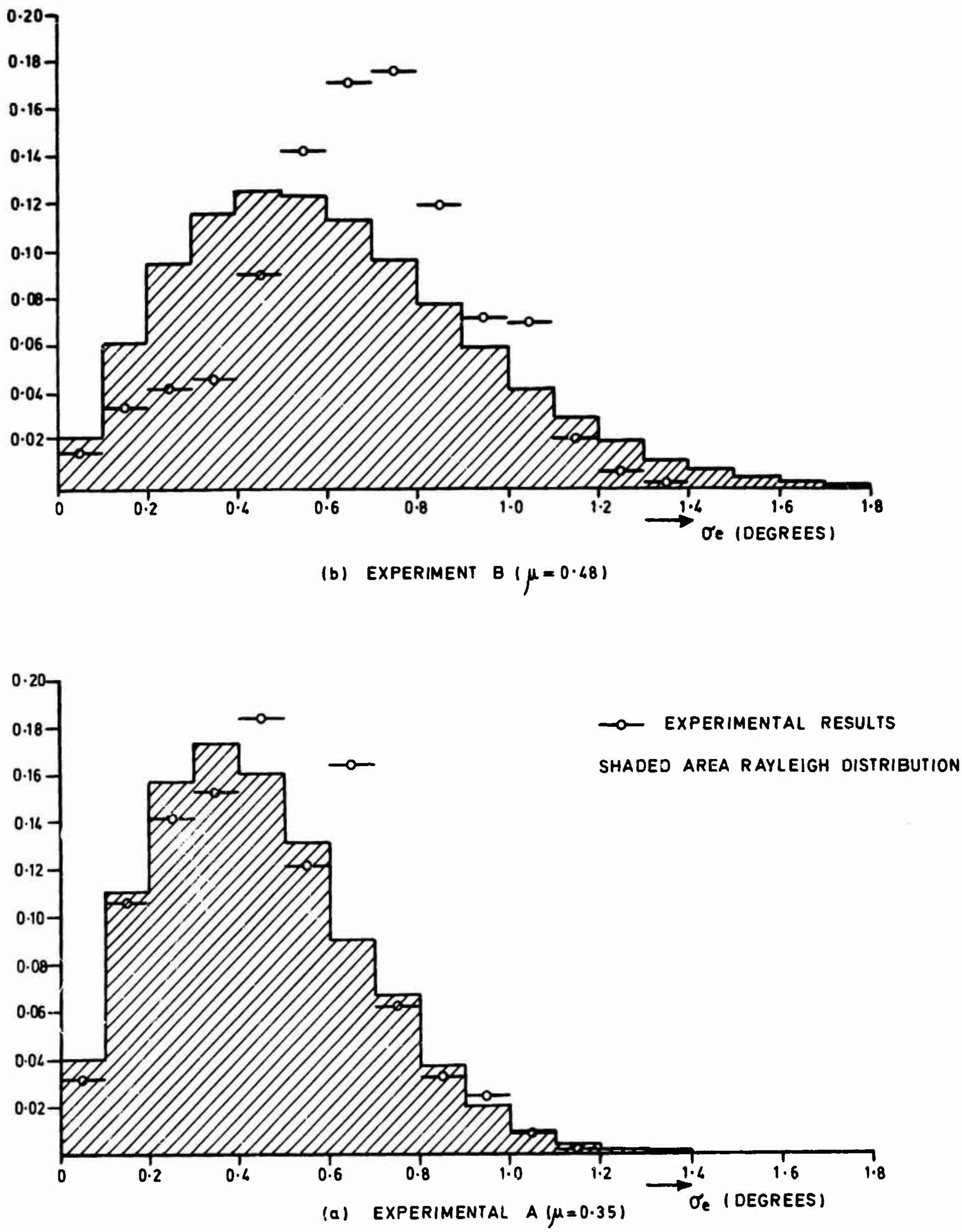


Fig.15 Probability density histograms of standard deviation of error signals in a duplex system with tolerances.

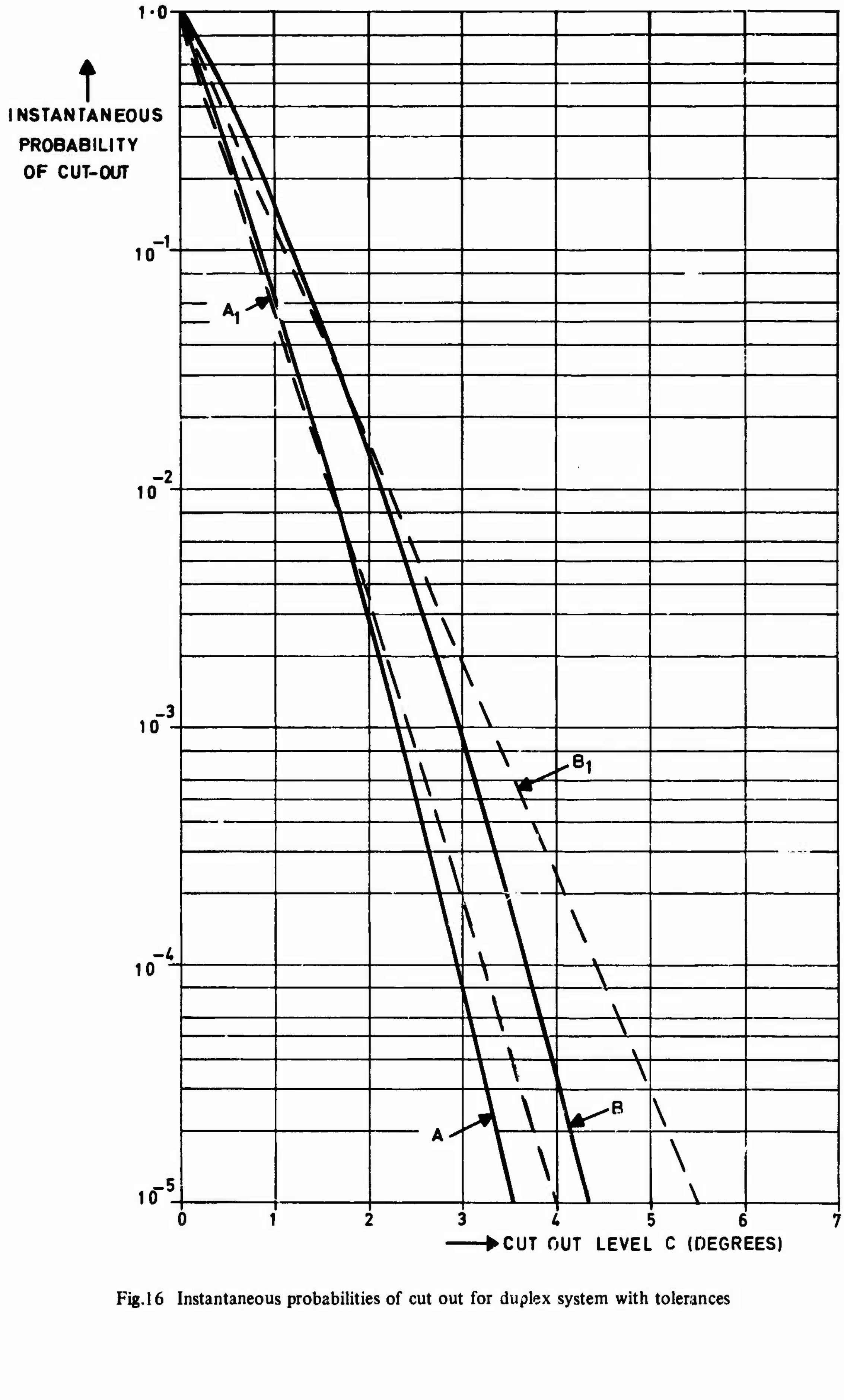


Fig.16 Instantaneous probabilities of cut out for duplex system with tolerances

POTENTIAL TELEOPERATOR APPLICATIONS IN MANNED AEROSPACE SYSTEMS

Edwin G. Johnsen
National Aeronautics and Space Administration
Washington, D. C.
U. S. A.

SUMMARY

Teleoperators are defined as general-purpose dextrous man-machine systems that can augment man, and extend man's sensory and manipulative capabilities across distances and through physical barriers into distant or hostile environments. Teleoperators can bring a new dimension to the automation of manned aerospace systems, particularly where control involves simultaneous multi-channel, multi-level actuation and sensory feedback. The trend of teleoperator development is toward digital computer controlled systems which utilize local sensor-computer-actuator loops to avoid obstacles and to sense manipulator grip-and-slip.

The potential applications of advanced teleoperator technology to manned aerospace systems have only recently begun to be identified. Applications currently being studied and evaluated include long manipulator booms to be mounted on the shuttle. These can transfer cargo from the space shuttle and can acquire and retrieve objects in space. Free-flying teleoperators capable of acquiring, inspecting, repairing or refurbishing satellites in orbit are another space application.

Another potential application of teleoperator technology is the concept of using an anthropomorphous teleoperator in lieu of man to control aircraft or spacecraft normally controlled by a human pilot. Instead of incurring the high cost of modifying the controls of each craft that is to be operated by remote control, the cost will be limited to the development of a small number of teleoperators that can be used to remotely control any spacecraft or aircraft.

1. INTRODUCTION

The development of modern control systems incorporating various kinds of feedback got its initial impetus during World War II when servo systems were developed to meet the fire control requirements of new weapon systems. Closed-loop control technology has come a long way since then and with the availability of minicomputers that can be economically dedicated to specific systems, it appears that a higher level of automation is on the horizon. Teleoperator technology, which occupies a position midway between manual operations and total computer controlled operations, is a part of this higher level automation. Combining the advantages of automatic control and the best features of human control, teleoperators can bring a new dimension to the automation of manned aerospace systems.

Teleoperators is a word invented to describe general-purpose dextrous man-machine systems that can augment a man by amplifying his ability to perceive, to move, and to manipulate. (Johnsen, E. G. and Corliss, W. R. 1971) Teleoperators can extend man by projecting his capabilities across distances and through physical barriers into distant or hostile environments. The unique characteristic of a teleoperator system is that man is always in the control loop. The man and the teleoperator are essentially symbiotic; man needs the machine's strength, endurance, quick response and resistance to hostile environments, while the machine depends upon man's intelligence for unanticipated or complex operations.

2. GENERAL INFORMATION

2.1 Teleoperator Configurations

The configuration of a teleoperator system is usually determined by its intended use. Some system components, particularly the manipulators, may have an anthropomorphous, or man-like, configuration. The anthropomorphous manipulator is usually desirable if the tasks to be performed by the manipulator include the maintenance or operation of equipment that was originally designed to be operated or maintained by manual methods. Anthropomorphous master-slave manipulators are also the most efficient manipulators if the operation must be completed in a very short period, or if it requires unusual skills.

However, the trend is toward non-anthropomorphous teleoperators. Familiar examples are the deep diving research submarines equipped with manipulators. Other examples of non-anthropomorphous teleoperators include head-aimed gun-control systems that have developed for use on some military helicopters and hydraulic excavating and trenching equipment modified to be remotely controlled.

2.2 Teleoperator Technology

Since the purpose of teleoperators is to augment and extend man's sensory and manipulative capabilities, teleoperator technology can be applied to the control of many manned aerospace systems, particularly where control involves simultaneous multi-channel, multi-level actuation and sensory feedback.

2.2.1 Visual Systems

Probably the most important sensory system developed for teleoperators to extend man's sensory capabilities is the visual sensory system. Current research and development work in the remote vision technology is being directed towards visual systems that will give the human operator a "sense-of-presence" or a psychological feeling the he, the operator, is at the remote location and that he is looking directly at the scene or object, instead of looking at a TV monitor. The "sense-of-presence" effect is partially achieved by slaving the motions of the TV camera to the motions of the operator's head. This type of head-aimed television camera control has been demonstrated to be an excellent control technique for remotely controlled vehicles. Other visual technology improvements such as foveal/peripheral visual display formats, (Control Data Corp. 1967), stereo presentations, and color TV can also significantly improve the visual sensor system. In addition to these demonstrated developments, other untested techniques might also improve the visual system. These untested techniques include systems that can provide real-time X-ray images. Combining several techniques, such as visual, sonic or radar imaging would permit superimposing several images on the same display. We anticipate that such visual systems will augment man's visual capabilities far beyond his natural abilities.

2.2.2 Manipulator Systems

The manipulative capabilities of remotely controlled manipulators are being constantly improved. A recent development is a system with two hydraulic manipulators capable of handling armed ordnance with great care. These manipulators have very sensitive controls and a force feedback system that provides a capability to work with small tools and to do delicate tasks. However, these manipulators are also capable of lifting twenty pounds per manipulator. Further, they are able to operate several types of power tools.

2.2.3 Sensors

Although these manipulators are reasonably dextrous, the technology exists to build manipulators that have far more dexterity than this system. For example, manipulators can be equipped with tactile sensing and feedback systems that will permit the operator to feel the location of corners, edges, and gross surface irregularities. (Bliss, J. C. and Hill, J. W. 1967-1968) Other recent improvements include proportional force feedback systems to reflect the forces on every joint, and terminal devices, or "hands" with dexterities approaching those of the human hand. These "hands" can be built with local "imminent slip" sensors that will cause the hand to automatically increase its grip force to prevent slipping once an object has been grasped by the device. We anticipate that within the next few years manipulators can be built that will have nearly all of the flexibility and dexterity of the human arm and hand, and which in addition will be able to handle very heavy loads and to work in any environment. Although normally manually controlled, the manipulators can be computer controlled in order to perform repetitive operations that can be programmed into the computer.

2.2.4 Autonomous Control

Although teleoperators are defined as systems where man is always in the control loop, there is no clear line separating teleoperators from robots. (In this discussion robots are defined as systems that are completely autonomous, possessing a self-contained artificial intelligence system capable of sensing and adjusting to its immediate environment). With the anticipated development of several types of localized control loops such as obstacle avoidance systems and grip-and-slip sensors, teleoperators will progress further toward a robot-type of operation. It is anticipated that when practical developments occur in the technology of artificial intelligence, these developments will be incorporated into teleoperator systems in order to sense and predict the reliability of critical mechanical systems. For example, critical operating data such as temperatures, vibrations, rate of wear, and presence of impurities can be the input into a minicomputer which would compare the data with information in its memory. If predetermined thresholds predicting failures are approached, a warning system could be activated. However, instead of allowing the sensing system to automatically stop the system being monitored, or to switch the operation to a predetermined series of events, a human operator would be alerted and would have the opportunity to observe, adjust or repair the system if these actions are possible or advisable. It thus appears that when artificial intelligence technology for teleoperators is developed, future manned aerospace systems could apply the technology in several ways in order to realize a number of benefits, such as decreasing turn around times or optimizing flight paths.

2.3 Rationale for Human Augmentation

In evaluating the technology of teleoperators, it appears that many of the advanced concepts of teleoperators, including artificial intelligence, are applicable to the automating of manned aerospace systems.

Although the concept of automating a manned aerospace system appears to be a contradiction of terms, there are sound reasons for automating manned systems wherever feasible. For example, in most control situations, the human operator behaves somewhat as an intermittent correction servo. His ability to track has a wave like form, and his correcting action appears to be a series of ballistic movements. All of these control movements are relatively slow since man appears to be physiologically incapable of any repetitive motions faster than 10 cycles per second, and more nearly 5 cycles per second. These times will vary widely between different individuals. For each individual the times will vary as function of fatigue, training, age, and general well-being. Because of these widely varying physiological conditions, man cannot be plugged into precise control equations. To describe man's control capabilities as a series of variable low speed jerks is accurate as well as graphic.

However, man's poorly coordinated or inadequate control movements are only a part of the picture. Fortunately, man also has a very favorable control feedback system in his sensory-neuromotor system which not only maintains closed loop control of responses and receptor input but can adapt its mode of control in a learning situation.

When these physiological facts are factored into man-teleoperator systems, we see that man's ability to change his mode of control adaptively in response to varying situations is far beyond anything in electronic or mechanical technology, but his physical movements will always be slow and erratic.

Once these parameters of man's abilities and man's limitations are recognized and accepted, it appears that since high speed and repetitive operations can be handled by a computer with far greater speed and precision than is possible with manual control, an on-board digital computer could replace many of the repetitive motions of a man and be a part of advanced manned aerospace systems.

Computers, however, have many drawbacks. Most current state-of-the-art computers cannot change modes of control adaptively. They are completely incapable of anticipating events or situations. They cannot determine or seek goals, and they are completely helpless if confronted with tasks not stored in their memories. In view of these limitations, it may be that in manned aerospace systems, man must be able to supervise the computer, to direct it, as a mindless slave in coping with unanticipated situations.

2.4 Potential Applications

Just as teleoperators need not conform to traditional configurations, teleoperator systems are not limited to a narrow band of applications. On the contrary, the list of possible applications is constantly growing, and new configurations are being created to satisfy new functional requirements.

2.4.1 Shuttle Applications

The potential applications of advanced teleoperator technology to manned space systems have only recently begun to be identified. A possible space application receiving considerable attention at the present time is the concept of using long manipulator booms to transfer cargo from the space shuttle or to acquire and retrieve objects in space and bring these objects on board the shuttle.

2.4.2 Free-Flying Space Teleoperator

Free-flying teleoperators are also being studied. The mission of a free-flying teleoperator would be to acquire, inspect, repair, or refurbish satellites that are in orbit. Since these studies and concepts of shuttle and free-flying teleoperators are the subject of numerous other reports and papers, this paper will concentrate on discussing other possible applications that have not been studied or discussed in any significant detail.

2.4.3 Teleoperator Pilots for Aircraft

Perhaps the most interesting concept is the idea of developing a teleoperator with a semi-human configuration. This teleoperator would be used in lieu of man to control aircraft and spacecraft designed for manual control. This system would be provided with anthropomorphous manipulators, anthropomorphous "legs" and a head controlled TV system located between the two manipulators in a position corresponding to the location of a man's eyes. Equipped with a high fidelity visual system and with force and tactile sensors, this system should be able to control any aircraft or spacecraft that has been built to be manually controlled, but without requiring any modification to the control system of the craft. This concept differs from the standard approach to remotely piloted vehicles in that instead of bringing the controls of an aircraft back to a control console, the ability to manually control a craft is extended from the control point to the craft itself. In other words, by means of the teleoperator, the sensory and manipulative capabilities of a human operator are extended out to the space or aircraft.

One of the principal advantages of using this technique is the significant economy that can be realized if it is desirable to fly aircraft or spacecraft by remote control. Instead of incurring the high cost and loss of time required to modify controls on each craft that is to be operated by remote control, the total cost will be limited to the development of a small number of teleoperators. Just a few anthropomorphous teleoperators can provide capability to remotely control virtually all types of spacecraft or aircraft. Another less obvious benefit would be the ability to use the same spacecraft or aircraft controls that a pilot would use in operating the craft. In instances where it would be desirable to flight test the craft at extreme operating conditions where manned control is impracticable, such as 12g accelerations, the ability to test the controls and instruments could be an important advantage.

2.4.4 "Spare Spacecraft" Concept

Another possible application, relevant particularly to space operations, would be the capability to send an unmanned spacecraft along with the manned spacecraft on long space journeys. This redundancy of vehicles and life support systems could provide a margin of safety of significant value in assuring the completion of important exploration trips. In this concept, the unmanned craft would be controlled from either the ground or the manned spacecraft, and its life support supplies would be used only in an emergency.

A different application of teleoperator systems which can be achieved using the current state-of-the-art is the possibility to guide parachute drops to land cargo in a predetermined area with the same or perhaps better accuracy than could be achieved by a manned parachute drop. This concept would permit parachute drops of unmanned systems from aircraft flying at any height and at any speed.

2.4.5 Kamikaze Application

A third possible application of teleoperator systems is so obvious that it will only be mentioned. This, of course, is the use of low cost expendable teleoperators to fly worn out or otherwise expendable aircraft on "kamikaze" missions.

2.4.6 Applications of Component Technology

In addition to the possible application of complete teleoperator systems, it is possible that many of the components and subsystems currently being developed for teleoperator use would find useful application in manned aerospace systems. For example, it appears that advanced systems of the head controlled TV system might be of value. In this application, miniature, gimbal-mounted TV cameras located in the cargo bays in the engine nacelles, on the landing gear and at other critical locations could give the crew visual confirmation of indicated malfunctions.

It is possible that one of the most important contributions that the teleoperator technology might make to the automation of manned aerospace systems is the development of actuators and feedback systems that will be applicable to fly-by-wire systems. Since all teleoperators are essentially fly-by-wire systems that use electrical, hydraulic, or pneumatic actuators, and combinations of these types of actuators, it is possible that the continuing development of advanced teleoperator technology will result in the invention of new powerful reliable and sensitive actuators and feedback systems that will be of value to the aerospace applications.

3. CONCLUSION

In summary, teleoperators have been defined as extenders and augmenters of man. Man is an intelligent, general purpose versatile system, but with physiological limitations that are incompatible with some of the demanding requirement of high performance aircraft and spacecraft. It therefore follows that a well designed combination of man, computer, teleoperator, and aircraft or spacecraft would provide a total system that will have higher performance, better reliability, shorter turn-around times, and more versatile uses than can be realized with systems that depend solely upon human control.

4. REFERENCES

Bliss, J. C. and Hill, J. W., 1967-1968, "Tactile Perception Studies Related to Control Tasks," Progress Reports on NASA Contracts.

Control Data Corporation, 1967, "Foveal HAT - A Head-Aimed Television System With a Dual Field of View." Report TM-74-98 Rosemont, Pa.

Johnsen, E. G. and Corliss, W. R., 1971, Human Factors Applications in Teleoperator Design and Operation, John Wiley and Sons, New York.

MAN-MACHINE CONSIDERATIONS IN THE DEVELOPMENT OF A COCKPIT FOR AN ADVANCED TACTICAL FIGHTER

S. Joel Premselaar
The Boeing Company
Seattle, Washington

D. E. Frearson
USAF Flight Dynamics Laboratory
Wright-Patterson Air Force Base, Ohio

SUMMARY

A revolutionary cockpit concept for a 1975-85 one-man, multi-mission fighter aircraft completed an initial simulation phase recently. The design goal of this concept is to achieve a one-man workload level by presenting the pilot only the information necessary for the particular mission segment he is performing, and yet provide maximum flexibility in terms of pilot options. Key elements of the cockpit design are: multiple, time-shared electronic displays; keyboard and voice command computer input devices; "wrap-around" cockpit arrangement for ease of access to the control-display devices; an integrated total energy command; and a system of dependent automation that permits reduced pilot workload during anomalies. The simulator provides a one-of-a-kind capability for examination of the flight deck design issues involved in tailoring the power and flexibility of the computer to the capabilities and limitations of the human pilot in the performance of his mission.

1. INTRODUCTION

Developmental trends in high-performance aircraft and avionics are expected to result in a proliferation of controls and displays that demand an unprecedented level of occupation by the pilot -- particularly those pilots flying advanced tactical fighters. Mission requirements for future tactical fighters will dictate system capabilities that will cope with not only the diverse environs of geography and weather both day and night, but also with the various gradations of cold, hot, and limited wars. The tactical fighter's arsenal must then range from guns to nuclear weapons with the gamut of conventional weapons between.

To be effective in its role, the tactical fighter weapon system must be capable of pinpoint navigation and high-accuracy weapon delivery. Additional complications include weapon deliveries at low altitude and supersonic speeds requiring a highly discriminatory and rapid target acquisition capability.

Technology in the 1980 time frame will enable production of tools necessary to complete a complex mission; however, the pilot faces the overwhelming task of collecting, collating, integrating, and interpreting a mass of information during normal and contingency operations.

Recognizing the need for revolutionary control-display concepts to realize the goal of a one-man, multi-role future tactical fighter aircraft, the U.S. Air Force Flight Dynamics and Avionics Laboratories initiated a program with the Boeing Company to address this challenging design problem. Accepting the implicit assumption that the computer will be the basic device within the cockpit by which the operational potential of this future multi-role aircraft is realized, the program has emphasized the flight deck design issues involved in tailoring the power and flexibility of the computer to the capabilities and limitations of the human pilot in the performance of his mission.

Experience gained from previous developmental programs offers a guideline for cockpit and avionics integration requirements. Some precepts may be derived from the lessons learned from the history of cockpit and avionics integration programs. Foremost among the precepts are:

- Cockpit and avionics concept definition must mate the machine to the man and the machine to the machine by detailed systems analysis of the mission(s) to be accomplished.
- The whole weapon system must be considered in develooing the cockpit and avionics subsystem configuration concepts.
- Flight simulation of the system is required to validate the functional capability of the concept as a whole.
- Flight test of prototype system hardware is necessary to demonstrate that the concept is compatible with the operational environment.

The basic objective of the system approach, as employed in the development of the integrated fighter cockpit, is to ensure the inclusion of all systems requirements through logical and sequential analyses. The tools developed to perform the analyses lend themselves readily to the iterative process. This facilitates reaction to test and evaluation results and provides a means for injecting technological developments into design or retrofit requirements.

2. DISCUSSION

Three separate studies were conducted to develop the advanced integrated fighter cockpit. The salient objective of each study is quoted below.

- First program: "...to develop an overall concept for a tactical fighter cockpit that will significantly reduce pilot workload."
- Second program: "...to finalize the advanced tactical fighter control-display system concept... (concentrating on) operating in degraded modes."
- Third program: "...to conduct a mission simulation program...to evaluate the 'wrap-around' cockpit."

Using a systems approach, the initial cockpit concept reflected the mission requirements for normal operations. The cockpit concept was exposed to anomalies and again subjected to analysis then reconfigured to reflect mission requirements for degraded mode operations. Figures 1 and 2 depict the activities of each phase of the analyses. Each activity is discussed in the following paragraphs.

2.1. Methodology

The analytical approach used to develop the cockpit concept for the next generation of tactical fighters required an examination of the 1980's tactical fighter pilot's operational needs. To this end, a composite mission profile and scenario was generated.

Considerable background information was used in determining the missions the tactical fighter would be expected to perform. Various sources were exploited to obtain this information. These sources included: a literature study; fact-finding trips; Boeing consultants; and interviews with military personnel.

The objective of this phase of the analysis was to fabricate a composite mission profile that would represent a major workload impact on the tactical fighter system -- man and machine. As such, the mission profile, Figure 3, though unrealistic in many respects, serves to force consideration of operational contingencies representative of most modes of weapon delivery.

A mission scenario was developed expanding the close air support mission profile that includes several unplanned events to complicate the operational situation. References in the scenario to basic mission functions, e.g., "employ countermeasures," "evade attack," "arm weapon," etc., were made without implying specific allocation of such functions to the pilot or the equipment.

The definition of system functions is achieved by constructing a functional flow block diagram such as Figure 4. The primary objectives of functional flow diagrams are to structure system requirements into functional terms, identify functional interfaces, and categorize the functions into tiers of a relevant hierarchy. The top, first, and second level functional flow block diagram divides the mission into coherent and easily recognized segments.

The information/action requirements activity illustrated in Figure 5 relates directly by number and name with those functions derived from the functional flow block diagram. In this portion of the analysis, the functions were reduced to tasks. Each function was divided into related information requirements at a task level. The action required to perform the task was then defined.

At this point the design process became iterative. Extensive fact-finding visits were conducted to military and industrial facilities to determine the stipulated 1980 technological state-of-the-art. By definition, only equipment successfully demonstrated, at least at a laboratory level, was considered as a candidate component for the advanced tactical fighter. Based upon the information derived from the fact-finding trips, the total weapon system was described in detail.

The total system was then divided into two basic elements for purposes of executing the action necessary to perform the task -- man and equipment. Based upon the equipment's capabilities and accepted human engineering practices, the tasks required to perform the function under consideration was allocated to man or equipment or to both.

The pilot action requirements necessary to accomplish the defined tasks were examined for the purpose of ranking them into one of three levels with respect to reach and vision envelopes. To determine its ranking level, each task was assigned a level of criticality: First, according to safety, then according to mission. Additional criteria upon which the determination of primary, secondary, or tertiary action requirement location was based were with respect to four considerations in pilot task performance: (1) frequency of use, (2) precision requirements, (3) response time required, and (4) dwell time.

Concurrent with the ranking of the action requirements, the cockpit volume was divided into reach and vision envelopes with respect to the combat eye reference point. The reach and vision envelopes were ordered consistent with the ranking of the action requirements -- primary, secondary, and tertiary areas. The limits are defined so that controls and displays may be located in their ranked positions.

Equipment design requirements were then introduced. Each pilot action requirement was examined and methods for implementing the action were defined. Each method was, in turn, examined, and human factor pros and cons relative to the pilot's performance of the task were listed.

The trade study options arrived at by defining methods for accomplishing the tasks assigned to the pilot were evaluated. Based on the requirements previously stipulated, the most promising option in terms of pilot performance was selected for inclusion in the cockpit as illustrated in Figure 6. The option selected was based on pilot performance requirements without respect to cost or engineering considerations.

After selecting the best method for performing a specific task, specifications for the equipment were defined. Although all aspects of control and display specifications were considered, the salient considerations, in order to establish the design concept, were symbology, sizes, and shapes.

The controls and displays were defined and allocated to specific reach or vision envelopes and were grouped functionally rather than on an equipment subsystem level. All pieces of equipment were located in their assigned areas or, as space permitted, to the next higher area but never to the next lower.

After the controls and displays were grouped functionally, the configuration was examined to ensure that conflicting cues were not presented to the operator. If equipment in proximity bore too close a resemblance, trades were made to eliminate the possibility of error.

The detailed control and display analyses were predicated on normal operations. However, no cockpit configuration may be considered complete without a complete degraded mode analysis. As a consequence, after the system was defined for normal operations, a degraded mode analysis was conducted within the constraints of the previously mentioned ground rules.

It was necessary to develop a list of failure modes in order to determine the equipment, controls, and displays required to survive anomalies. The list of systems and subsystems developed was examined in every flight phase for its impact upon safety of flight or mission completion. Critical systems were faulted without regard to failure probabilities since, ultimately, the anomaly could be caused by battle damage.

The analytical process for the degraded mode analysis then followed the methods employed in the previous analysis wherein the system and cockpit concepts were evolved.

A mock-up of the cockpit resulting from the normal and degraded mode analyses was fabricated. A computerized workload evaluation was conducted for both normal and degraded mode operations.

2.2. Study Airplane

The systems, based on 1980-1985 technology, developed in the initial study, then updated as a result of the degraded mode analysis, are presented in the present tense -- as though the system existed in actual hardware form. Equipment identified in contemporary terminology is done so in the generic sense only. The intent in these cases is to describe the function a system performs rather than existing 1972 hardware.

The cockpit of the tactical fighter is the focal point for the aircraft's systems and subsystems. Definition of the aircraft and its avionic system is a necessary step to the development of an integrated control-display complex. The aircraft and avionic systems required to meet the needs of an advanced tactical fighter were developed within the framework of the following ground rules and assumptions:

Ground Rules

- One-man crew
- Air-to-ground combat is the primary operational mission
- Time-shared display techniques
- 1980 avionics state-of-the-art
- Pilot retains executive control over an automatic system
- Systems approach in the analysis

Assumptions

- Twin-engine airplane
- Maximum speeds: Mach 2.5 at altitude and 1.5 at sea level
- Air-to-air defense capability
- Variable sweep wing
- Conventional takeoff and landing capabilities only
- Digital avionics interface
- System cost, weight, and reliability secondary considerations to crew performance

The study airplane's maximum gross weight is approximately 52,000 pounds. Using engines strengthened for operations to Mach 1.5 at low altitude, the thrust-to-weight ratio is 1.4 at a combat weight of 35,000 pounds. At optimum altitude, speed performance is Mach 1.6 at military power, Mach 2.3 sustained, and Mach (dash) 2.5 with a variable inlet.

Although the mission requires effective weapon delivery under instrument flight conditions, certain aspects of the stipulated close-support mission (e.g., troop strafing, etc.) require visual capability for the delivery of ordnance. Accordingly, the airplane concept is configured to provide the pilot with downward vision over-the-nose that will meet all mil lead requirements for the firing of guns and the release of all stores. The over-the-nose vision provided also permits steep landing approaches to forward area bare bases without having the aircraft's nose obscuring direct view of the runway.

The requirement for an automatic variable sweep wing was dictated by the desire to achieve optimum lift-to-drag ratios for attack operations over a broad velocity spectrum. A manual control option enables the pilot to improve airplane ride quality in turbulent air.

The primary flight controls of the aircraft are provided with continuous automatic damping. A self-adaptive gain system is incorporated to optimize aircraft response. The spectrum of conditions to which the self-adaptive gain system is tuned is modified by the mode of flight selected. Control harmonization (responses to the controls with respect to the other controls) is also affected by the mode of flight selected. Adverse yaw compensation provides signals to the rudder to reduce control cross-coupling effects. The survivable flight control system combines fly-by-wire aircraft control with duplex integrated hydraulic servo-actuator packages installed at each flight control location.

The aircraft is powered by two low bypass turbofan engines equipped with afterburners. Interchangeable engines, mounted side by side, may be started by the on-board APU. Bleed air from either operating engine provides the means for starting the other.

The aircraft fuel system consists of self-sealing wing and fuselage tanks filled with polyurethane to reduce sloshing. An inert gas or a fuel spray enrichening blanket is used to minimize fire and explosion hazard during combat. External fuel is carried in inflatable auxiliary tanks or may be carried in drop tanks. Transfer of fuel maintains the aircraft center of gravity within specified limits for any selected flight mode. The aircraft can be refueled in flight from a tanker aircraft equipped with either a drogue or flying boom.

The prime electrical power supply is an engine-driven variable speed generator (alternator), which uses a hybrid arrangement with a solid state DC link converter to provide controlled frequency required for limited on-board equipment. Additional on-board power sources include an APU and a battery.

The retractable, tricycle landing gear includes fairing doors that are sequenced closed in both the retracted and extended positions, since the aircraft will operate from unprepared fields. Brakes with anti-skid provisions are included on all three wheels. Collapsible, high-inflation tires inflate rapidly as the landing gear extends.

The upward firing crew escape module encompasses the pressurized cockpit and provides for safe descent with an escape envelope extending from zero altitude and zero speed through maximum operating altitudes and maximum dynamic loadings. Chaff dispensing and emergency beacon transmission capabilities are provided and may be deselected if desired.

The environmental control system provides the functions of cockpit air conditioning and pressurization, transparent area clearance, avionic equipment cooling, and anti-icing of flight surfaces and of sensors as required. Emergency operation provisions are also included.

Successful advanced tactical aircraft system operation depends heavily on the continued availability of computing power. This dependency exists even when battle damage of component failure occurs. To achieve the high degree of continued computing power availability necessary, the advanced tactical fighter system has a computer system that can:

- Detect more than one failure.
- Automatically establish a circumvention or recovery procedure to eliminate the effect of failures.
- Provide a residual level of computing power that fully meets minimum essential computing requirements for the tactical fighter system.

The computer system can also handle specialized computing associated with the operation of synthetic aperture radars, pattern matching associated with automatic target recognition, and signature analysis of complex wave forms. In addition, the computer system is supplemented with a digital data transmission system to permit communication with the rest of the avionics equipment within the advanced tactical aircraft system. The digital data transmission system can operate in spite of component failure or battle damage.

The computer also performs computation, processing, memory, and time standardization for the total airplane system. Multiplexing greatly simplifies interconnecting hardware and system maintenance by eliminating the need for a multitude of wires, connectors, and junction boxes. System modifications to add new functions or to change functions can be accomplished without extensive rewiring. Multiplexing makes practical the wrap-around cockpit in which the instrument panel is structurally secured to the hinged canopy (Figure 7).

Complementary to the essential system capabilities described above is the software system. This system includes an executive program that is responsible for the operational management of the computing system while simultaneously checking for loss of computing power. If such a loss occurs, the executive controls the circumvention procedures necessary to overcome its effects. Application programs carry out the many functional tasks required by the advanced tactical fighter system. Supporting software is also present in such areas as self-test and diagnostic procedures. This support software checks for failure of any computing element or other avionics systems able to communicate (through the digital data transmission system) failure data to the central computer. Appropriate evaluation and circumvention procedures are initiated when a failure is detected. Lastly, ground support programs exist to assist in the design, construction, maintenance, and documentation of the operational computer programs. These programs include compilers, assemblers, file programs, etc.

Four methods for feeding data into the system are suggested for use in the tactical fighter. These are a tape cassette, a data link, a keyboard, and human voice.

The tape cassette is used to program the mission into the computer. This tape is produced by a ground-based computer and carried on-board by the pilot.

Data link input is used to reprogram the flight from a remote station. When the pilot is physically incapacitated, the aircraft may be controlled remotely.

The keyboard, Figure 8, consists of an 8-key master keyboard select panel and 38 multifunction keys. Dedicated keys are included to perform the functions of clear, modify, display, enter, backspace, and space. A pre-entry readout is provided for verification of keyed instructions before entry into the system.

To allow the pilot to command certain operational changes during critical flight phases, without requiring him to use the hands, the human voice can be used to call out specific tasks. It recognizes the sequential usage of key command words. The voice command entry technique allows the pilot to enter changes into the system without using the keyboard or the cassette.

The Failure Monitoring and Control System (FMACS) for advanced tactical aircraft is more than the traditionally conceived on-board test and checkout system; it is a programmed control system that performs many airborne and ground-oriented functions. While airborne, these functions include system status reporting to the pilot and data linked to a remote monitor, warning of flight safety conditions, automated corrective actions and display of these actions, and notification of impending failures. On the ground, these functions include automatically controlled preflight, flight/mission readiness, and fault isolation tests.

The FMACS primarily reduces and simplifies the pilot's workload if any flight subsystem is in a degraded mode of operation. In many cases, FMACS can completely assume the tasks of restructuring the operational parameters of various systems to eliminate a failed or degraded system's impact on the pilot. In other cases, simple status displays in go, no-go, formats, plus simple corrective directions are displayed automatically to the pilot.

Failures affecting flight safety are presented to the pilot automatically through the Master Warning System. Emergency procedures are displayed on the preselected MPD. Failures that may affect mission performance are presented automatically on the MPD.

Warning and Caution are presented to the pilot in up to four ways depending on criticality of the failure or malfunction detected. The types of warning are tactile, voice, master caution light, and video display.

The navigation system selected for IIPACS is comprised of three basic subsystems -- Doppler, Inertial and Satellite (DIS), augmented with Radar, Data Link and DME position fixing. Two inertial systems with a backup Heading and Attitude Reference System (HARS) provide the redundancy and reliability required for "fail operational" during critical mission segments.

Spherical coverage is needed to operate against threats and targets approaching from any directions. An integrated antenna design concept is desired so that aircraft weight volume, structural shadowing, and interference problems are minimized.

The integrated antenna system consists of multiple phased array systems to satisfy the requirement for spherical coverage. These multifunction arrays use an inertialess, computer-controlled electronic beam steering system to allow coordinated control with minimum reaction time.

Although acceptable designs with either planar or conformal arrays are possible, a conformal array approach with the antennas integrated into the aircraft structure is expected to provide best coverage with minimum aircraft weight and volume penalties.

Communications, navigation, and identification systems are programmed for integration into one spread-spectrum bandwidth. Each function has a separate address code. Knowledge of the code and compatible equipment allows the users to communicate with each other.

The spread-spectrum technique is designed to reduce vulnerability to enemy jamming and to provide simultaneous service to many stations or addressees. Satellites are expected to handle hundreds of small mobile terminals with varying power levels without the formal network discipline normally used in satellite communications between large ground terminals.

A secure or private code must be used to provide anti-jam capability. Since each link is spread across the entire available bandwidth, both CW and broadband jamming must overcome the process gain of the spread-spectrum system. Therefore, a great deal of jamming power is required.

Although the tactical fighter's primary mission is to destroy ground targets, self-defense using penetration aids is also necessary. Self-defense options, which include evade, degrade (not to be confused with degraded modes as implied by the nature of this study) and destroy, must be used selectively to minimize primary mission penalties.

The prime purpose for the tactical fighter is weapons delivery. As such, all systems related to the aircraft are oriented to achieve this end. Consequently, the stores management system (SMS) depends both directly and indirectly on virtually every system involved in the placement of the appropriate store on the specific target.

All-weather operation is required to continuously maintain a tactical advantage. Visual delivery and delivery with options selected from the full complement of on-board sensors are needed for flexibility. Precision data link or navigation-system-controlled automatic blind weapon delivery is desired for preplanned targets. The SMS integrates all equipment affecting weapon delivery.

The automatic target acquisition system correlates data received from the following sensors:

- Forward looking infra-red (FLIR)
- Low light level television (LLLTV)
- Multi-mode radar (MMR)
- Radar homing and warning system (RHAW)
- Infra-red warning system
- Target identification system, electro-optical (TISEO)

Sensor data are processed through the center computer complex and presented to the pilot on demand. Techniques such as the image registration process deal with the structure of the image as a whole, as opposed to the detailed comparison of individual resolution elements.

More than one sensor symbol can be presented simultaneously on the target acquisition display (VSD) for identification by superposition of target data from each sensor. For preplanned targets with known coordinates, the cursor appears at the top of the HSD at the predicted location.

The weapon system can gather battle damage and/or reconnaissance data. On-board sensors sense the ground target area and aid in compiling sufficient data to determine target viability before and after a strike. Such information is transmitted through data link to the Battle Area Commander.

On-board equipment can detect and record sensor data from the MMR, LLLTV/FLIR, RHAW, and EMP measuring devices. In addition, targeting and weapon delivery information, navigation data, and voice recording are continuously available through tape decks for playback to the Command Post via data link. Such data may be requested at any time by the Command Post without assistance from the pilot.

The identification (IFF) functions of the IIPACS aircraft are accomplished with either of two independent systems. In actual operation, however, they will have a common phased or conformal array antenna system. The IFF systems are described below.

The directional communication IFF system is an integral part of the spread-spectrum communication band. Frequency will be in the SHF band or higher for improved directivity. Spherical (360^{o}) coverage is provided. The antennas are directional during transmit and receive. The system can interrogate another aircraft at a specified bearing when directed by on-board systems through the central computer.

The multi-mode radar IFF system is integrated with other MMR functions at the operating radar frequency. Spherical (360^{o}) coverage is provided. Modes and codes are integrated with existing radar pulses with secure message structure contained within each pulse.

2.3. Control-Display Concepts

The study analyses were primarily oriented toward reducing the interpretive and integrative processes imposed on the pilot while he performs the complex mechanics of flying a sophisticated aircraft. The result of these analyses produced the concept of an integrated total energy management system (ITEMS).

Energy management is by no means a new concept. Attempts at energy management first appeared with the issue of pilots' handbooks. The most advanced work in the field is found in space vehicles and may be applied to aircraft. ITEMS dispenses with the requirement to present quantitative information. The ITEMS may be programmed to conserve energy in one of three ways; (1) in the form of fuel-maximum range or maximum endurance, (2) in the form of time-minimum time between any two points, and (3) in the form of energy-maneuverabilityvelocity/altitude.

Displays fall into a head-down or head-up (HUD) category. It is well known that a finite time is required to refocus the eyes from near to far and from far to near objects. Many factors enter into accommodation time. Data on accommodation inertia (focusing time) obtained in the Boeing Visiometrics Laboratory indicate the time to obtain a clear image when the observer shifts from near visual tasks to extra-cockpit vision can be significant, (Figures 9, 10 and 11). Consequently, the requirement for a HUD system to display primary flight instrument information as well as fire control information becomes manifest.

Three display levels are used in configuring these conceptual cockpits. Primary displays are the vertical situation display (VSD), HUD, and horizontal situation display (HSD). A secondary display group, referred to as multi-purpose display (MPD's), performs various display functions for the pilot depending on the flight phase. The third display level is that associated with the controls.

The VSD/HUD is the primary attitude reference. In presenting attitude data to the pilot, other optional information can be added to the VSD/HUD without cluttering it beyond the pilot's capability to note changes. These data are used for terrain following, target acquisition, and target attack. The VSD is depicted on Figure 12.

When a designated target nears the outer release envelope of the selected weapon, a fire control symbol appears on the VSD/HUD. The energy management symbol is retained. The movement of the dot on the fire control symbol represents qualitative range rate. It appears at the 11 o'clock position and moves counterclockwise. The 6 o'clock position represents the maximum effective range for the selected weapon, and the 12 o'clock position is the minimum effective or safe firing range. The computer continually updates weapon release parameters; therefore, weapons are dispensed automatically when the dot is at the 3 o'clock position and manually when the dot is between the 6 and 12 o'clock positions. The fire control symbol is also used as an aiming window or reticle.

The HSD is the main source of heading and position information. The information sources feeding the display are the navigation system, the multi-mode radar, and the passive identification system. A moving map can be displayed in all modes. Radar PPI data is selectable. VSD terrain following information is augmented by a terrain clearance PPI presentation (situation display) on the HSD.

MPD's are grouped around the VSD and HSD. They present more vital information in the available primary space than is possible with conventional instrumentation. Graphic and alphanumeric information in color is tailored to the flight phase and provides the pilot with comparative data for increased confidence in system operation.

Figures 13 through 16 typify subsets of display formats that may be made available by preprogramming them as a function of flight modes. The cockpit layout and other display formats and controls are shown in Figures 16 through 29.

2.4. System Evaluation

A simulation program was conducted to evaluate the advanced integrated fighter control-display concept from the point of view of operational pilot acceptance and performance. The evaluation was structured around the following objectives:

- Investigate "Stand Alone" capability of the energy management display concept.
- Evaluate pilot usability of the multi-purpose display concept in support of multi-mode operations.
- Examine potential of voice command and the general purpose keyboard as computer input devices.
- Evaluate usability of the wrap-around cockpit concept in a multi-mission fighter application.
- Examine integrity of the computer augmented system in degraded mode operation.

Three kinds of missions, graded in complexity, were simulated for the evaluation; (1) ferry missions, (2) single purpose missions, and (3) full missions with degraded mode operations. The missions were all derived from the original composite mission. All but one of the evaluators were seasoned combat pilots with an average of over 3,000 hours of flight time in sophisticated high performance aircraft. Each pilot received about 15 hours of ground school and 10 hours of simulated flight time after which they completed a questionnaire designed to elicit original constructive criticism regarding the IIPACS concept.

The concept was received well with comments ranging from acceptance to enthusiasm. Noteworthy is the fact that, according to the data, the inexperienced pilot (less than 300 hours) did at least as well as the others.

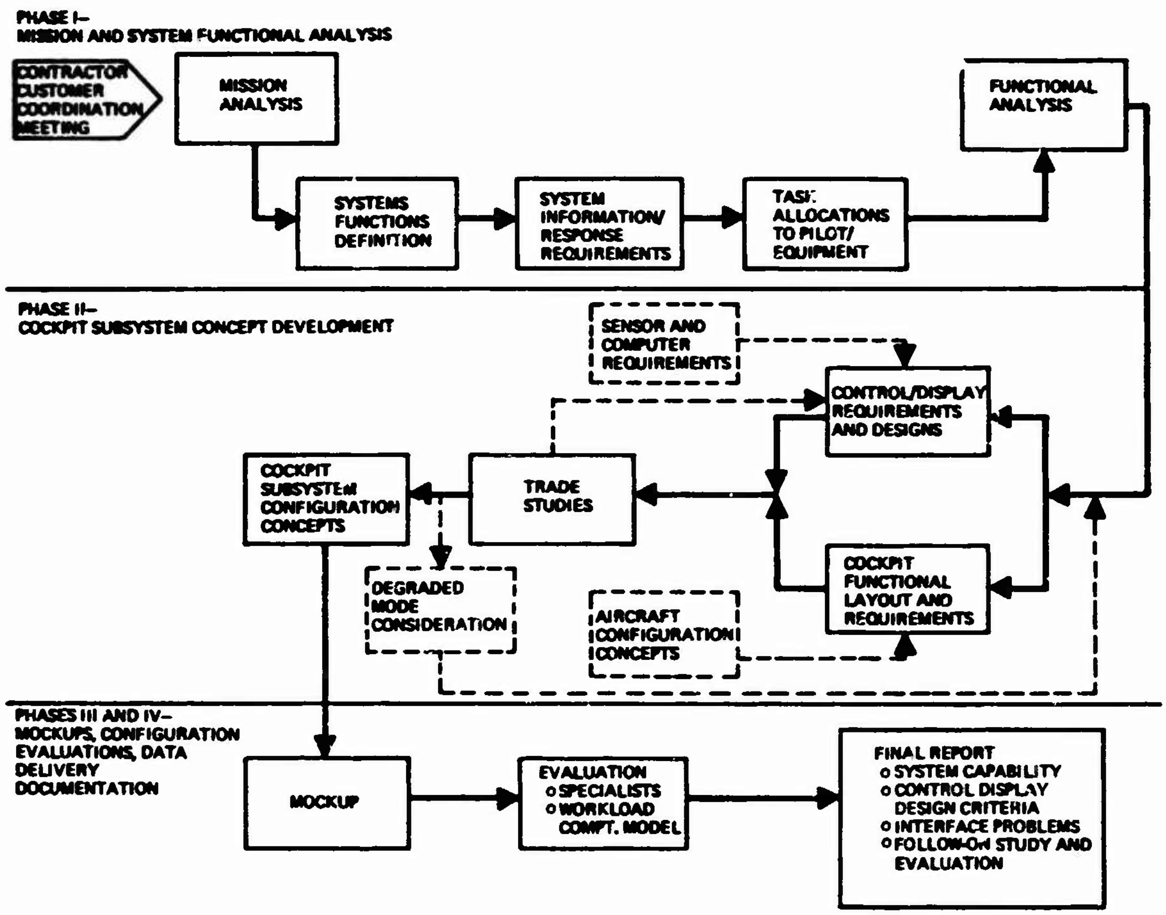


Fig.1 IIPACSS-1 program flow chart

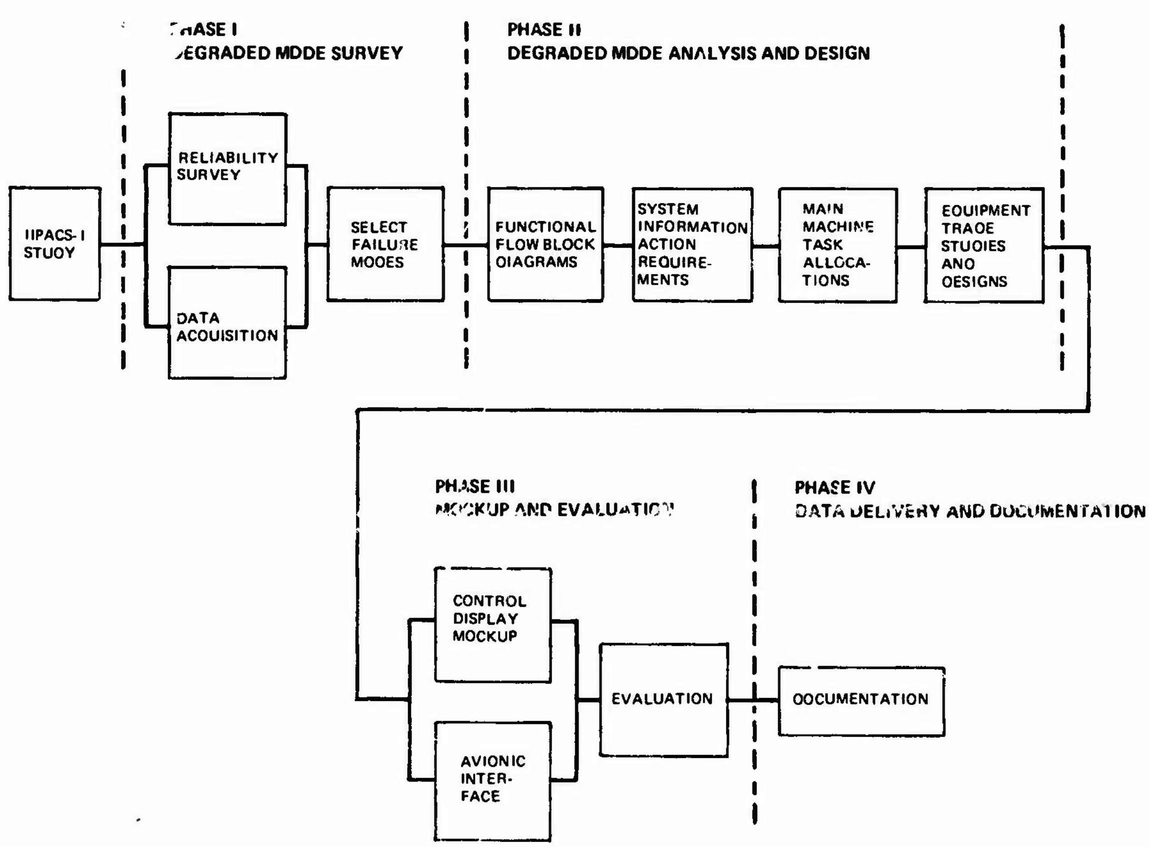


Fig.2 IIPACSS-2 program flow chart

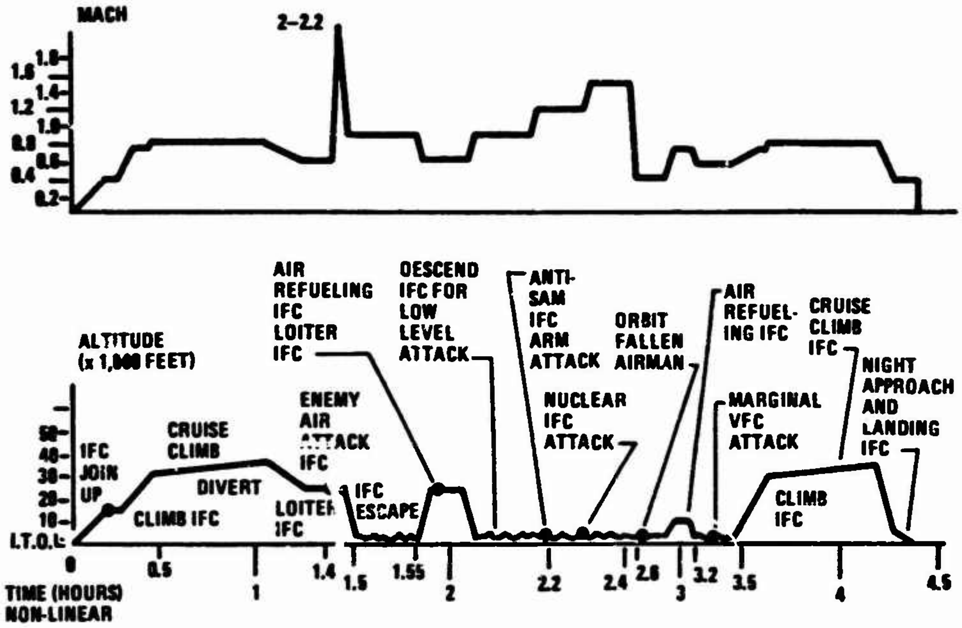


Fig.3 Mission profile

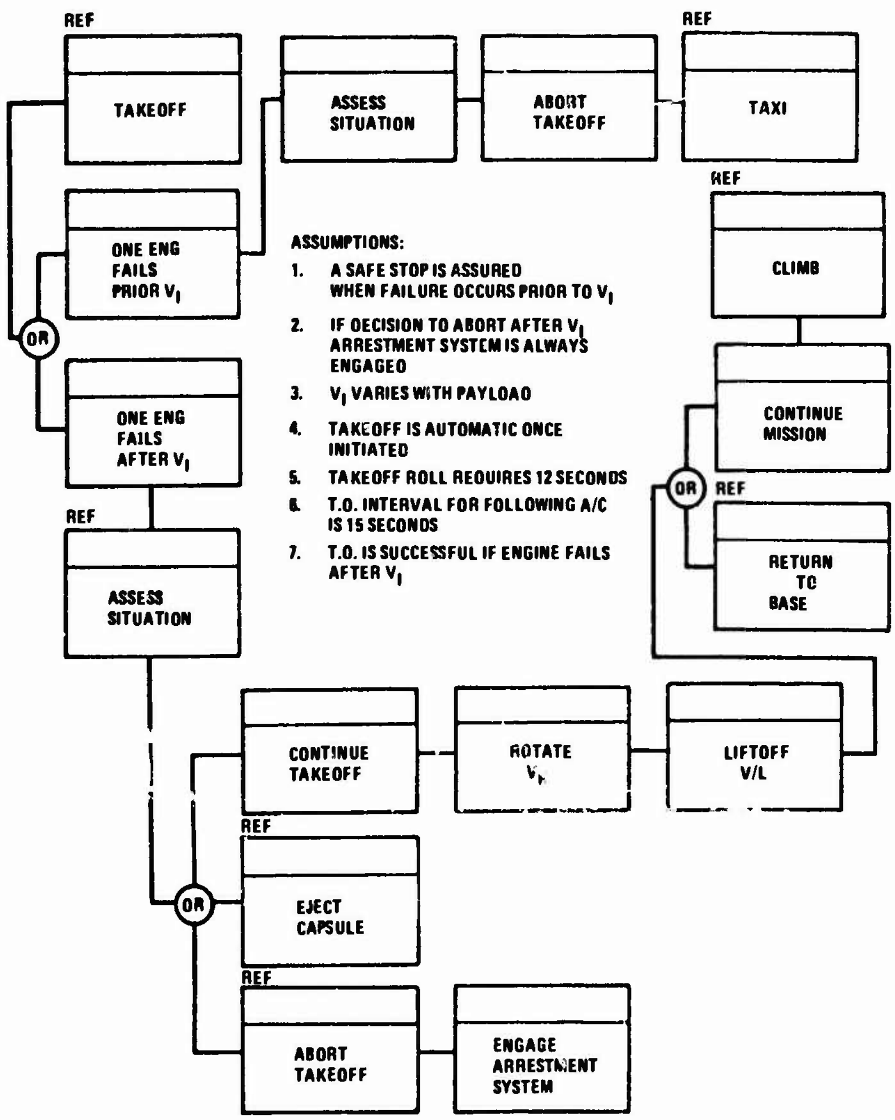


Fig.4 Functional flow block diagram

DEGRADED MODE: ENGINE FAILURE - TAKEOFF

FUNCTION NUMBER CONDITION	ALTERNATIVE ACTIONS	TASK/ACTION REQUIREMENTS	INFORMATION REQUIREMENTS	INFO. AVAIL/ WHERE
REF. 2.1.1.4 TAKEOFF 2.1.1.40.3 ONE ENGINE FAILS PRIOR TO V_1		1. DETECT FAILURE. 2. WARN CREW. 3. MONITOR WARNING AND PROCEDURES.	1. THRUST/TEMP./PRESSURE, SPEED/V_1/STEERING 2. VISUAL, AUDITORY AND TACTILE 3. PREPROGRAMMED MSG. IN STORAGE	 MASTER CAUTION VOICE, VSD/HUD MPD
	2.1.1.40.3.1 ASSESS SITUATION	1. CONSIDER: ● RUNWAY LENGTH REQUIRED TO ABORT ● RUNWAY CONDITIONS ● USABLE RUNWAY REMAINING 2. DECISION - ABORT CAN BE ACCOMPLISHED		
	2.1.1.40.3.2 ABORT TAKEOFF	1. ACTUATE THRUST REVERSER 2. ACTUATE SPOILERS 3. ACTIVATE WHEEL BRAKES 4. ACTUATE ARRESTMENT DEVICE 5. STEER AIRCRAFT 6. COMMUNICATE AND INFORM	1. THRUST REVERSER POSITION, POW... SETTING 2. SPOILER POSITION 3. BRAKING AVAILABLE 4. DEVICE AVAILABLE 5. VISUAL/INST. STEERING CUES 6. RADIO AVAILABLE (VOICE)	 MPD MPD

Fig.5 Information/action requirements

DEGRADED MODE: ENGINE FAILURE - TAKEOFF

DISPLAY/CONTROL REQUIREMENTS ABORT SWITCH	OPTION NO. 1 PLUNGER TYPE ON PANEL	OPTION NO. 2 AUTOMATIC ACTUATION WHEN ENGINE FAILS	SELECTION
• CRITICALITY HIGH • FREQUENCY OF USE INFREQUENT • RESPONSE TIME RAPID • PRECISION REQUIREMENTS HIGH • ENVIRONMENT CONSTRAINTS • LOCATION ALLOCATION • VISION • REACH PRIMARY	PRO: 1. MAY BE ACTUATED AT CREW'S DISCRETION. 2. MAN REACTS WELL IN CONTINGENCIES. 3. SIMPLE. 4. TACTILE CUE ELIMINATES NEED FOR DISPLAY. CON: 1. REQUIRES CREW DECISION. 2. REQUIRES DISCRETE ACTION. 3. MUST BE MANUALLY OPERATED WHEN TIME IS CRITICAL. 4. REQUIRES PANEL SPACE. 5. MUST BE RESET. 6. REQUIRES ILLUMINATION.	PRO: 1. WILL PERFORM FUNCTION WHERE CREW CAPABILITY IS MARGINAL. 2. CAN SENSE SMALL CHANGES IN STIMULI. 3. RESPONDS RAPIDLY TO REQUIREMENT. CON: 1. SUBJECT TO INTERFERENCE. 2. CAN EXECUTE ONLY AS PROGRAMMED. 3. MUST BE MONITORED. 4. COMPLEX AND PROGRAMMING REQUIRED. 5. REQUIRES DISPLAY.	OPTION NO. 1 1. SIMPLICITY. 2. PROVIDES POSITIVE CONTROL. 3. DISCRETIONARY.

Fig.6 Design trade study

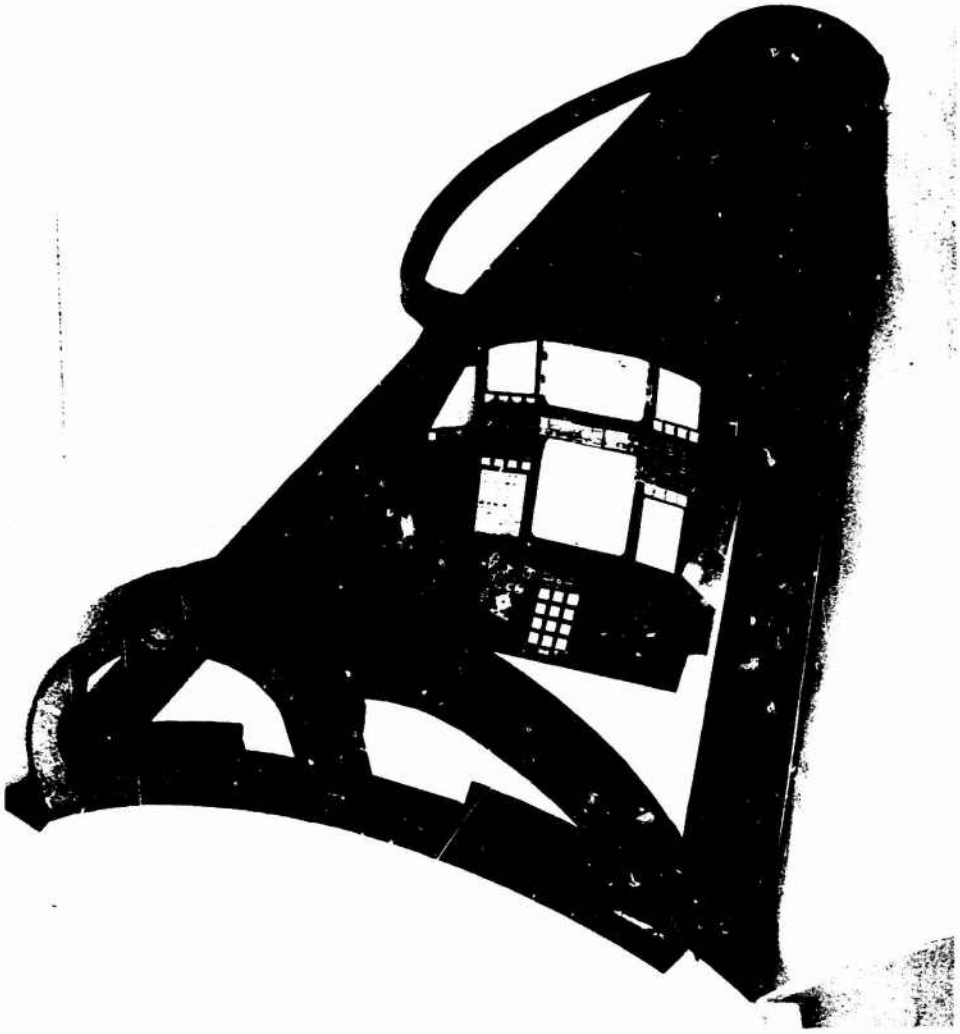


Fig.7 Hinged canopy concept

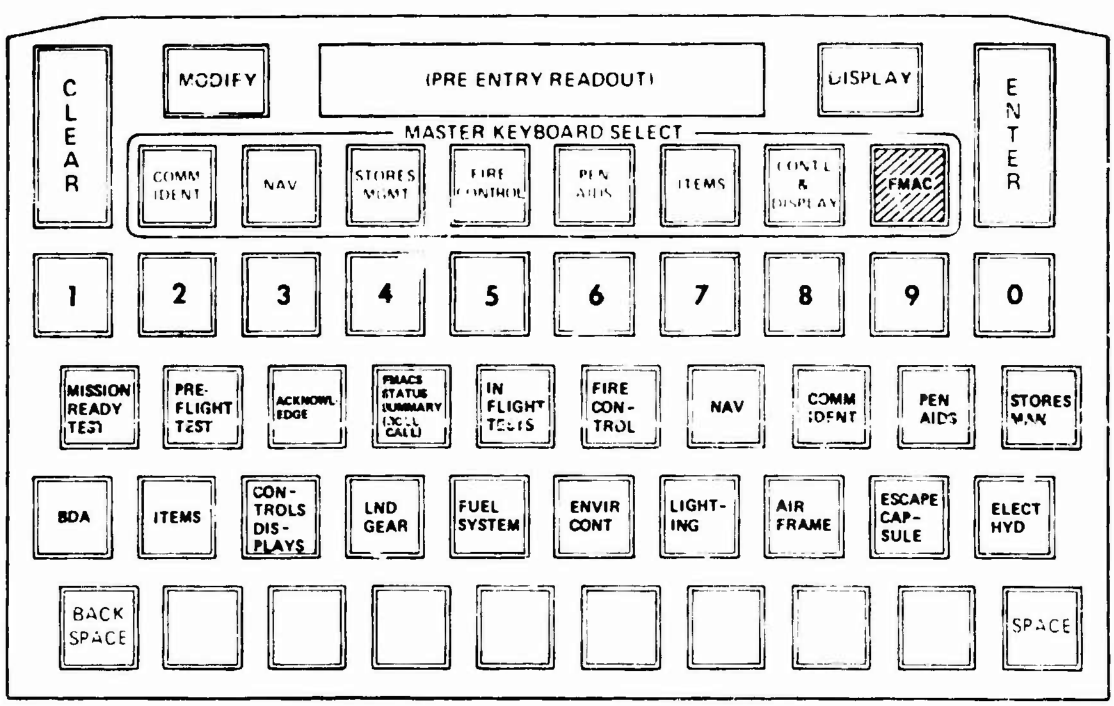


Fig.8 Master keyboard

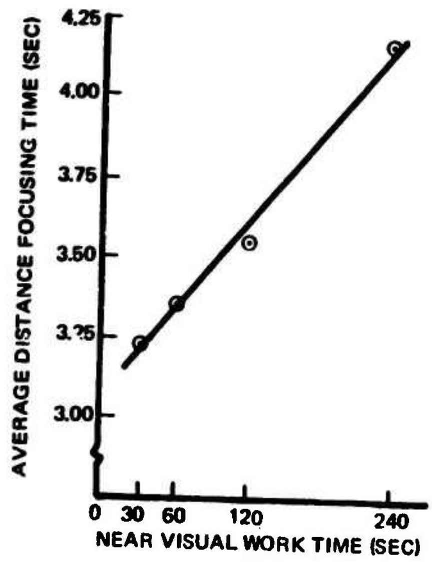


Fig.9 Average time for 35 observers (ages 21 to 55) to change focus from 26 inches to 20 feet (infinity)

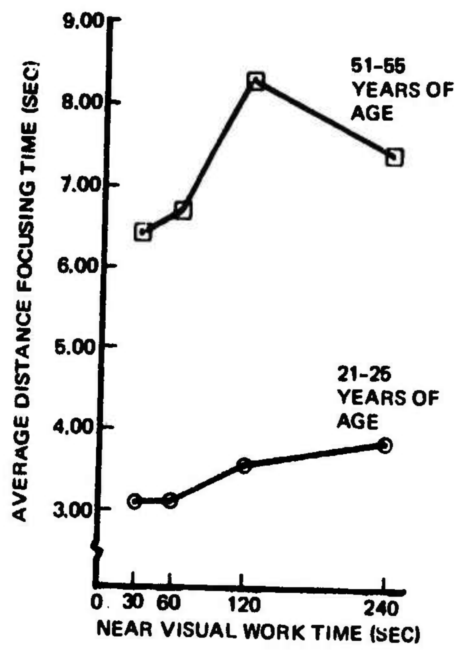


Fig.10 Time required to change focus from near to far for two different age groups

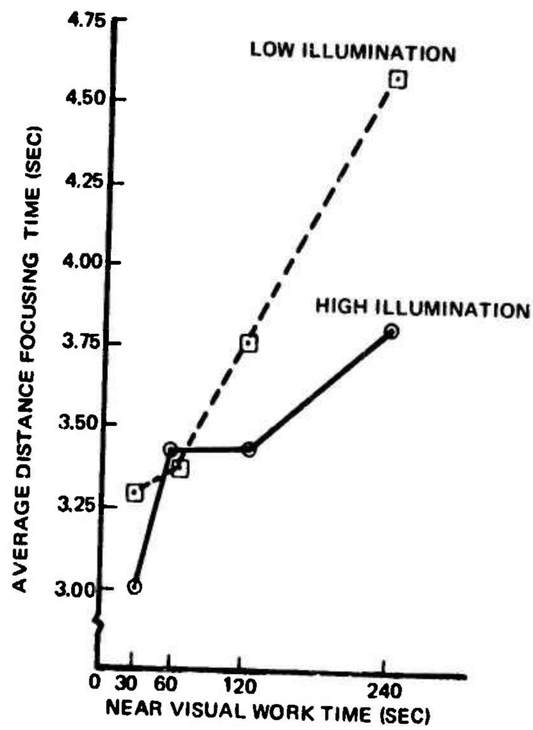


Fig.11 Effect of illumination on focusing time (near to far)

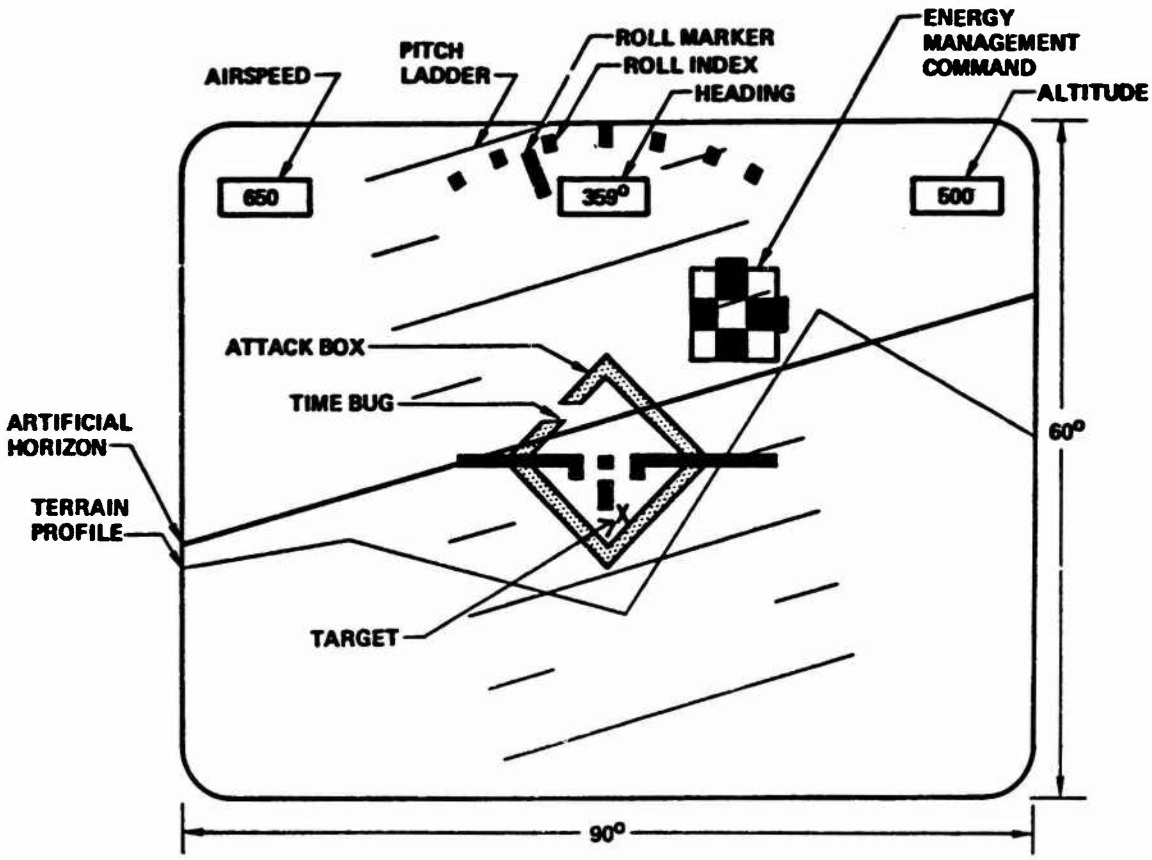


Fig.12 Vertical situation display

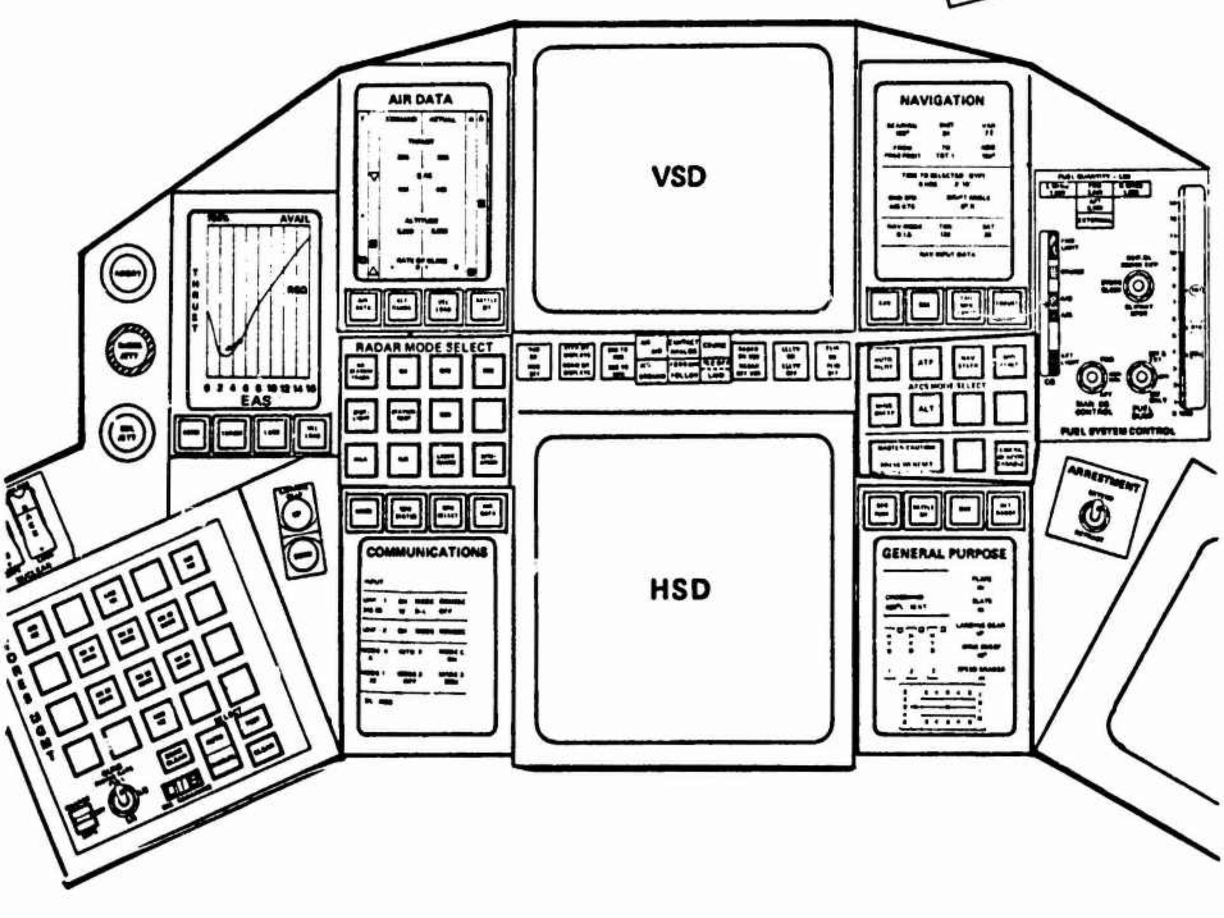


Fig.13 Takeoff mode MPD formats

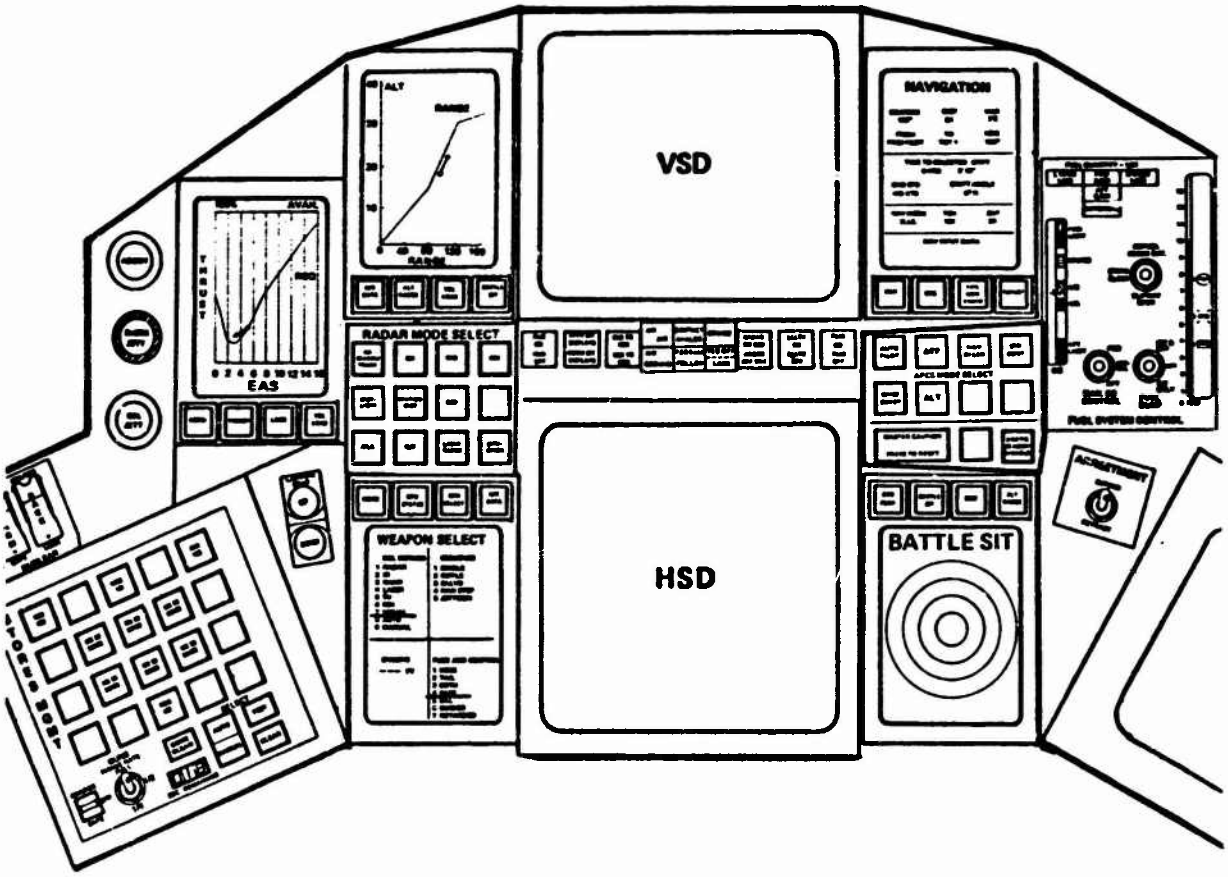


Fig.14 Cruise mode MPD formats

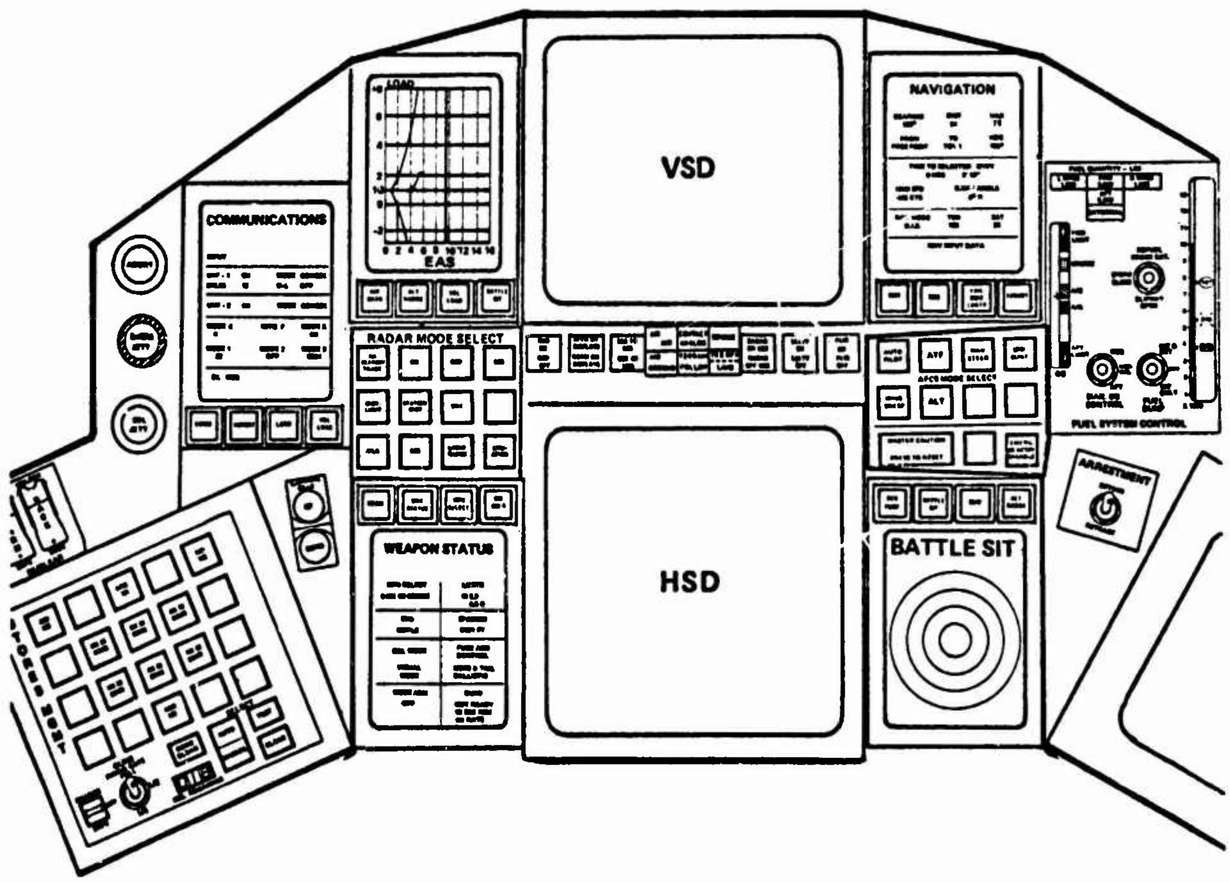


Fig.15 Air to air mode MPD formats

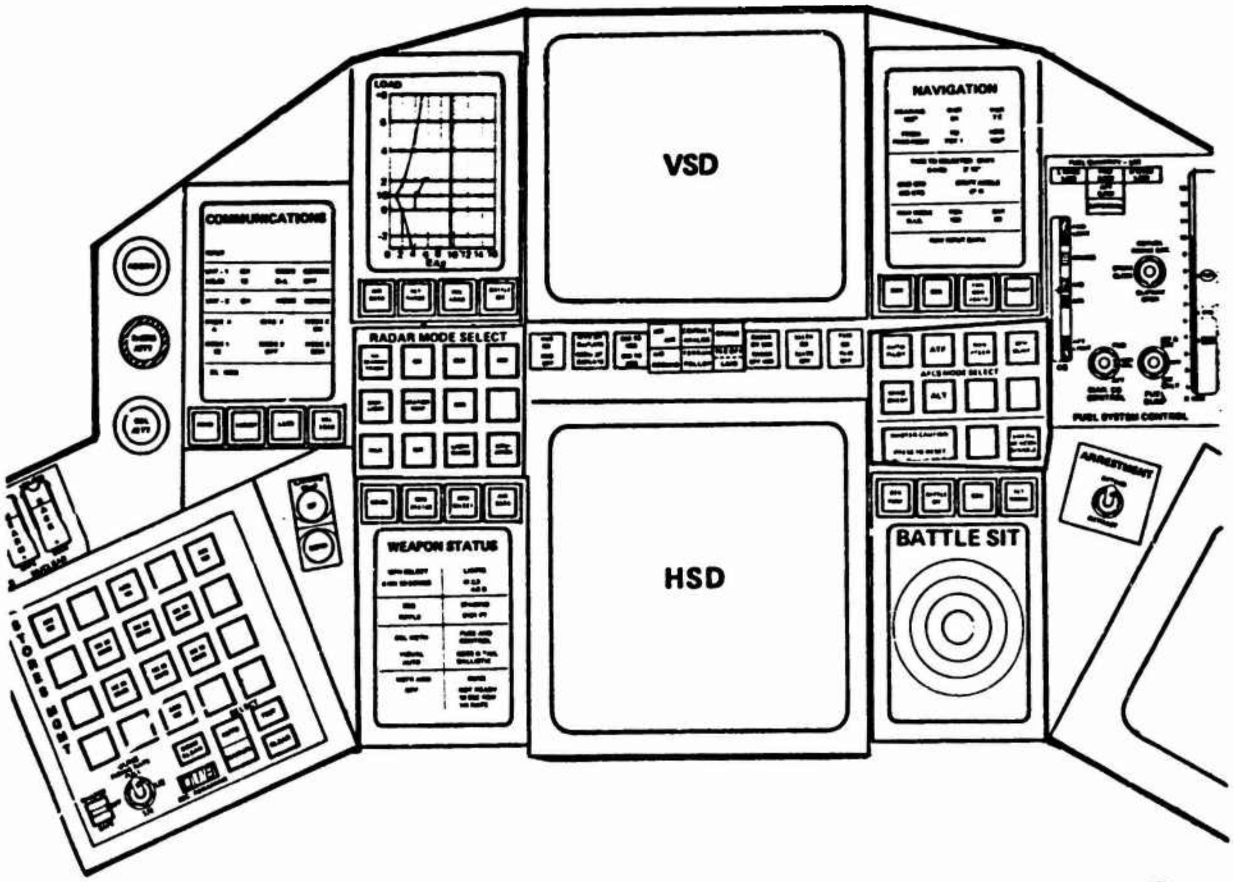


Fig.16 Air to ground mode MPD formats

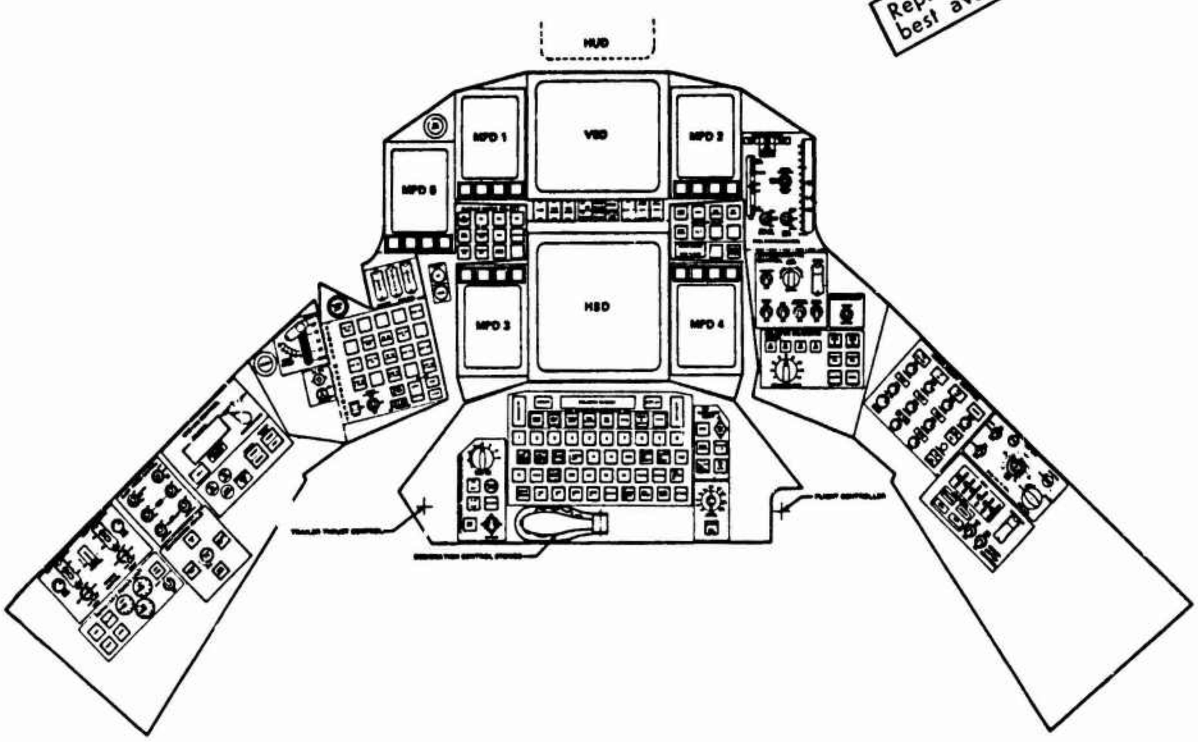


Fig.17 Panel layout

EXAMPLE SHOWS AUTOPILOT PERFORMANCE

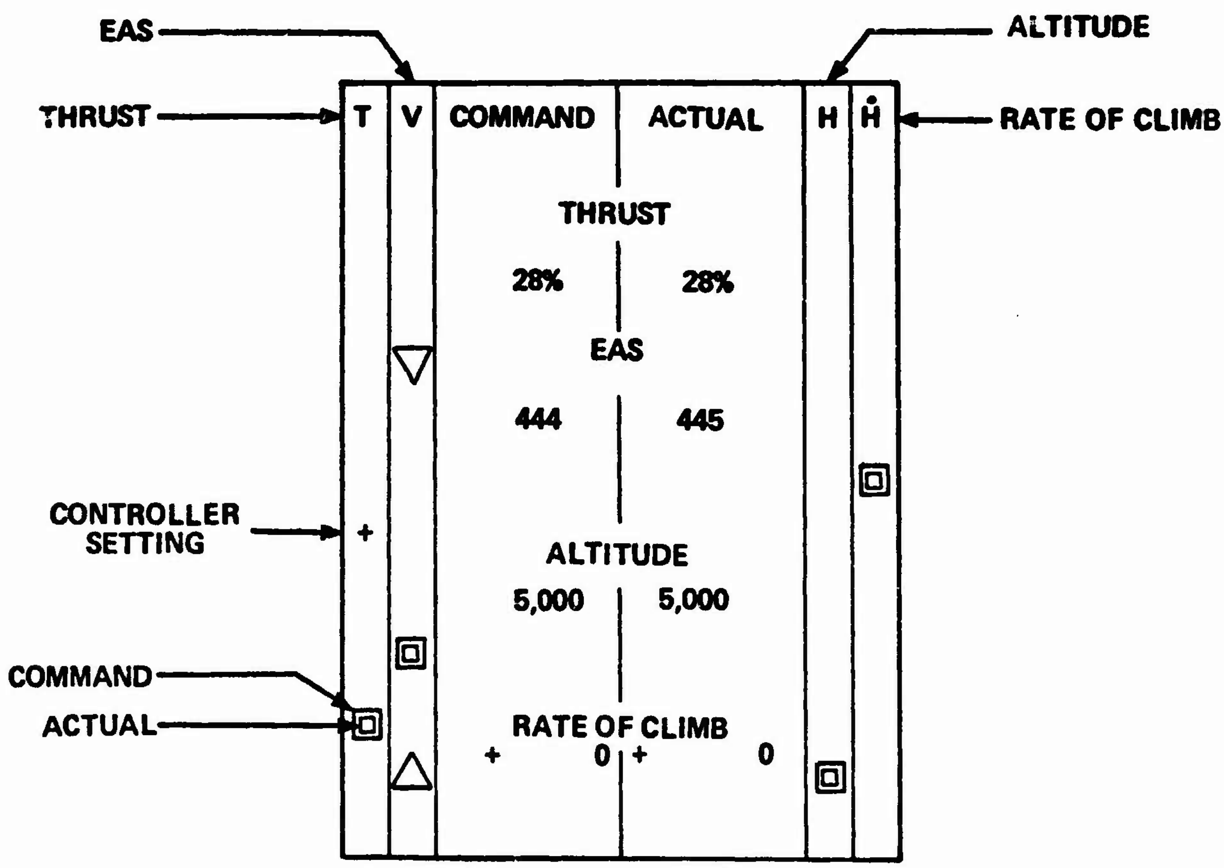


Fig.18 Air data MPD format

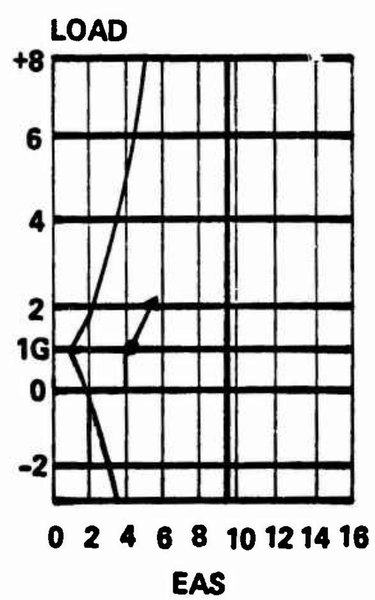


Fig.19 Velocity-load MPD format

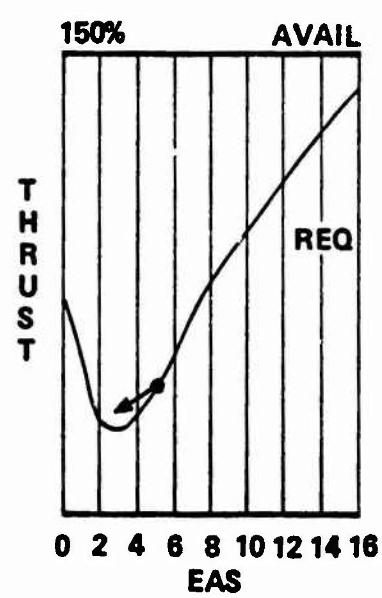


Fig.20 Thrust available - required MPD format

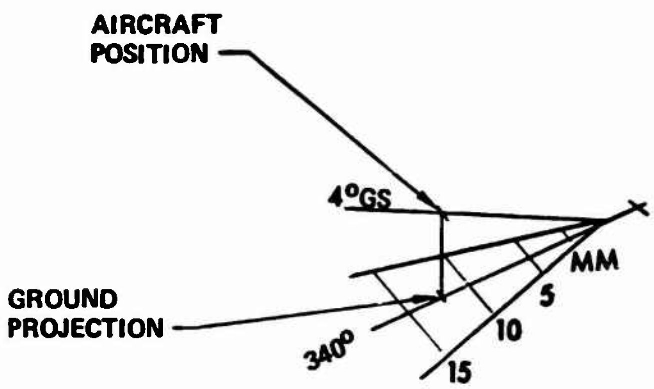


Fig.21 Landing MPD format

BEARING	DIST	VAR
155°	24	7E
FROM	TO	HDG
PRES POSIT	TGT 1	155°

TIME TO SELECTED WYPT
0 HRS 3' 18"

GND SPD	DRIFT ANGLE
445 KTS	0° R

NAV MODE	TCN	SAT
D.I.S.	126	26

NAV INPUT DATA

Fig.22 Navigation MPD format

DEL METHOD
1 RADAR
2 IR
3 RHAW
4 LASER
5 TV
6 RDI
7 VISUAL
8 AUTO
9 MANUAL

SEQUENCE
1 SINGLE
2 RIPPLE
3 SALVO
4 MAN STEP
5 JETTISON

SPACING
--- FT

FUZE AND CONTROL
1 NOSE
2 TAIL
3 BOTH
4 SAFE
5 BAL
6 GUIDED
7 RETARDED

Fig.23 Weapon select MPD format

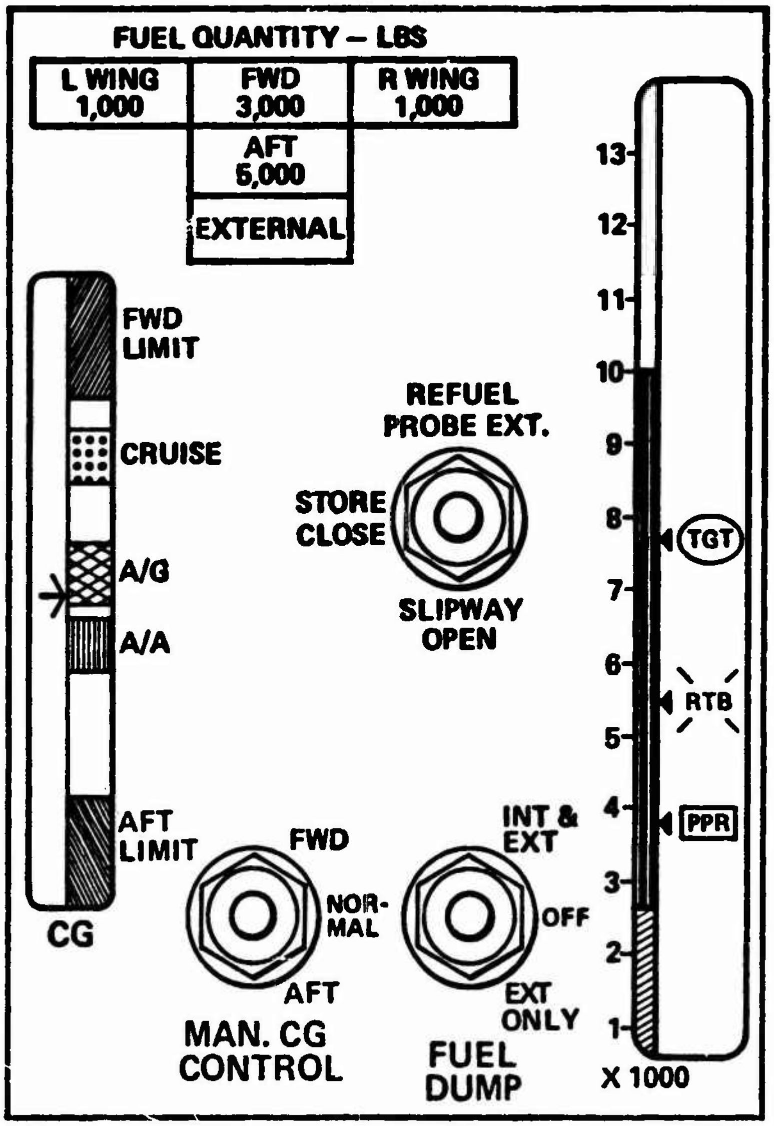


Fig.24 Fuel system control

WPN SELECT 6 MK 99 BOMBS	LIMITS M 2.3 4.5 G
SEQ RIPPLE	SPACING 0181 FT
DEL METH VISUAL AUTO	FUSE AND CONTROL NOSE & TAIL BALLISTIC
MSTR ARM OFF	GUNS NOT READY 18 SEC REM 1/4 RATE

Fig.25 Weapon status MPD format

FAILURE → PITCH AFCS FAILURE

CONSEQUENCES → UNAVAILABLE
- ATF
- ALT

REQUIRED ACTION → CORRECTIVE ACTION
- DISCONNECT AFCS
- PULL UP TO 500 FT TC
- RESET MASTER CAUTION
- FOLLOW TERRAIN MANUALLY

CRITICALITY → FLIGHT SAFETY CRITICAL

Fig.26 Failure monitor MPD format (example)

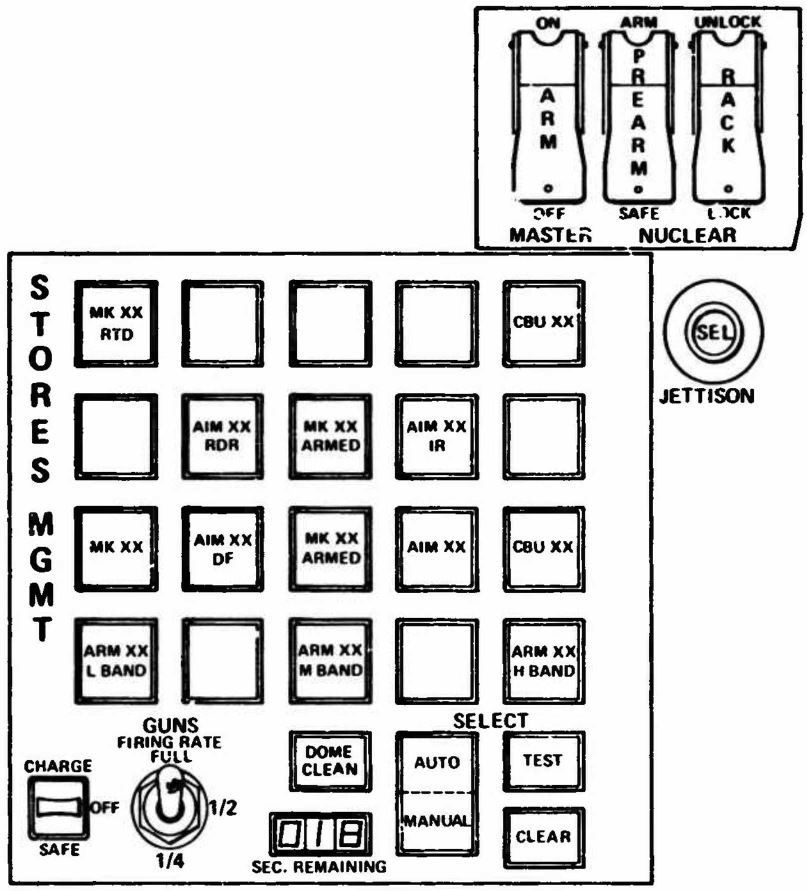


Fig.27 Stores management panel

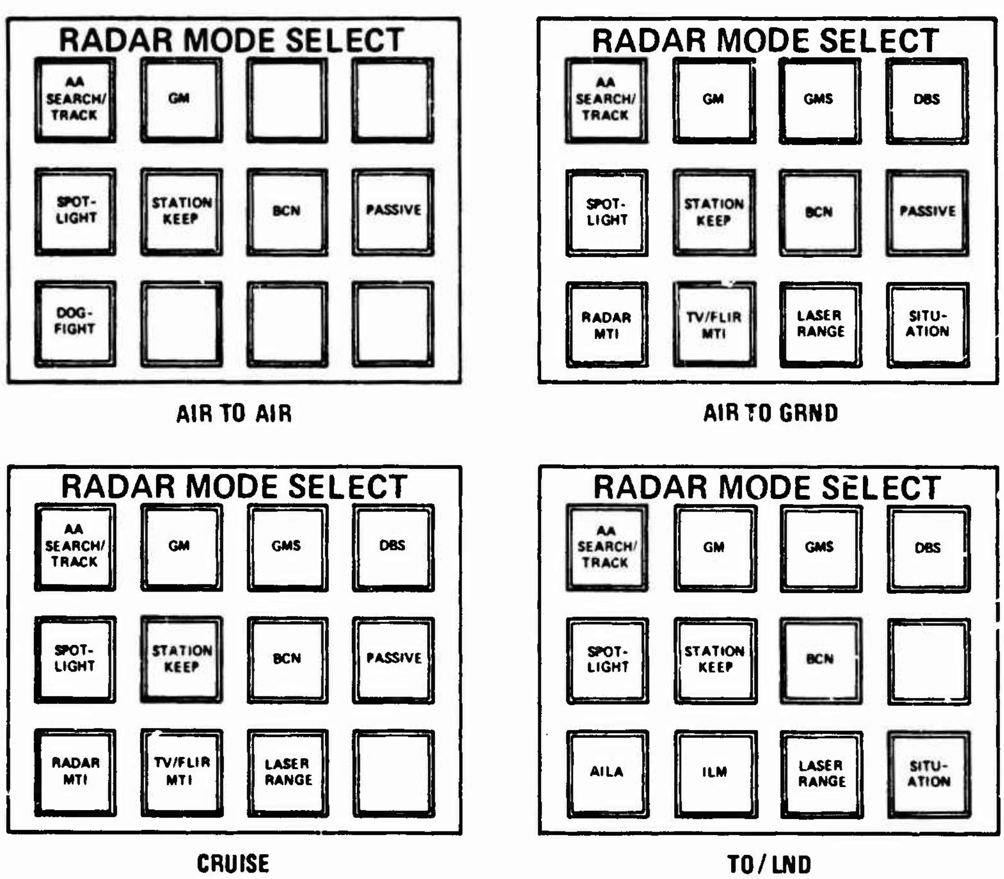


Fig.28 Radar mode select panel formats
for primary mission modes

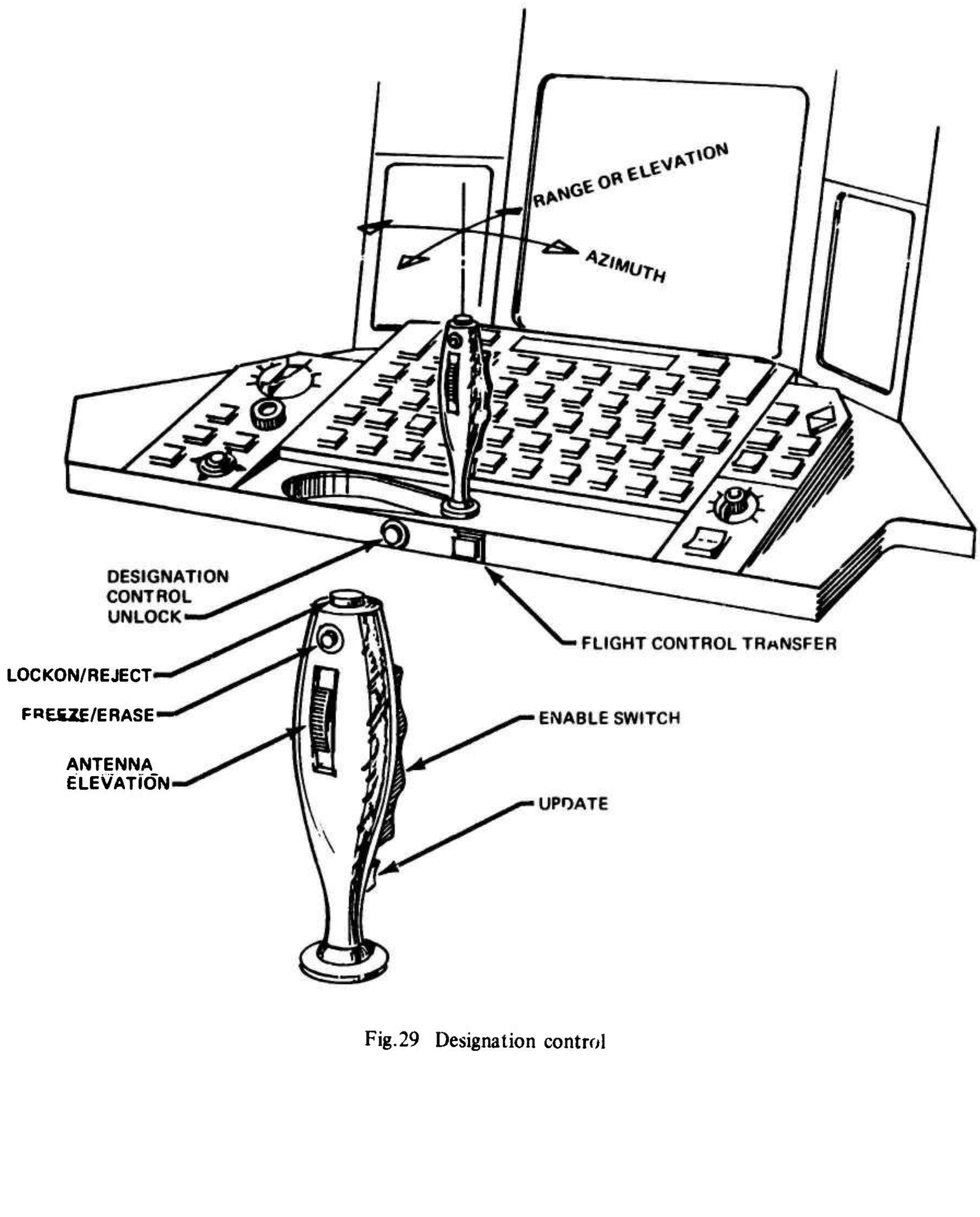


Fig.29 Designation control