

Also published on behalf of the  
INTERNATIONAL ASTRONAUTICAL FEDERATION  
Editor L. G. NAPOLITANO

Applications of Space Developments  
Space Developments for the Future of Mankind  
Astronautics for Peace and Human Progress  
Using Space - Today and Tomorrow  
    Volume 1 - Space Based Industry Symposium  
    Volume 2 - Communications Satellite Symposium  
A New Era in Space Transportation  
Space and Energy  
Space Stations, Present and Future  
Space Activity, Impact on Science and Technology

*Pergamon Journals of Related Interest*  
*Free Specimen Copy Gladly Sent on Request*

Acta Astronautica  
Advances in Space Research  
Earth Oriented Applications of Space Technology  
Planetary and Space Science  
Progress in Aerospace Sciences  
Space Solar Power Review  
Vertica

Dear Reader

If your library is not already a standing/continuation order customer to this series, may we recommend that you place a standing/continuation order to receive immediately upon publication all new volumes. Should you find that these volumes no longer serve your needs, your order can be cancelled at any time without notice.

ROBERT MAXWELL  
*Publisher at Pergamon Press*

# SPACE MANKIND'S FOURTH ENVIRONMENT

SELECTED PAPERS FROM THE XXXII INTERNATIONAL ASTRONAUTICAL CONGRESS  
ROME, 6—12 SEPTEMBER 1981

*Edited by*  
**L. G. NAPOLITANO**  
*University of Naples, Italy*



**PERGAMON PRESS**

OXFORD · NEW YORK · TORONTO · SYDNEY · PARIS · FRANKFURT

U.K.	Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England
U.S.A.	Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.
CANADA	Pergamon Press Canada Ltd., Suite 104, 150 Consumers Rd., Willowdale, Ontario M2J 1P9, Canada
AUSTRALIA	Pergamon Press (Aust.) Pty. Ltd., P.O. Box 544, Potts Point, N.S.W. 2011, Australia
FRANCE	Pergamon Press SARL, 24 rue des Ecoles, 75240 Paris, Cedex 05, France
FEDERAL REPUBLIC OF GERMANY	Pergamon Press GmbH, 6242 Kronberg-Taunus, Hammerweg 6, Federal Republic of Germany

---

Copyright © 1982 Pergamon Press Ltd.

*All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from the publishers.*

First edition 1982

**Library of Congress Cataloging in Publication Data**

International Astronautical Congress (32nd:  
1981: Rome, Italy)  
Space—mankind's fourth environment.  
(Astronautical research)  
Includes bibliographical references.  
1. Astronautics—Congresses. I. Napolitano,  
Luigi G. II. Title. III. Series.  
TL787.I443 1981 629.4 81-22736

**British Library Cataloguing in Publication Data**

International Astronautical Congress (32nd: 1981:  
*Rome*).  
Space: mankind's fourth environment.  
1. Astronautical research—Congresses  
I. Title II. Napolitano, L. G.  
500.5'072 TL565  
ISBN 0-08-028708-5

*In order to make this volume available as economically and as rapidly as possible the authors' typescripts have been reproduced in their original forms. This method unfortunately has its typographical limitations but it is hoped that they in no way distract the reader.*

## PREFACE

The 32nd Congress of the International Astronautical Federation took place in the year marking both the 20th Anniversary of the first manned flight and the beginning of a new space transportation era with the first successful flight of the space shuttle.

The scientific and technological developments and the human achievements achieved during the past two decades, on one hand, and the promises and challenges of the new era ahead, on the other hand, combine to emphasize the fact that space must be considered as mankind's fourth environment, to be further explored, exploited and managed peacefully and with even greater concerns than those for the other three environments: geosphere (land), hydrosphere (oceans), atmosphere (air). These considerations lead to the choice of the congress theme, whose development is faithfully portrayed in the selection of survey, status of the art, assessment and trend papers collected in these proceedings.

The proceedings consist accordingly of four parts. The first part addresses the problem and the challenges of the access to and permanence in the fourth environment. The survey of the development of USSR manned space flight program in the last 20 years, and the overview of the United States manned space crafts from Mercury to the shuttle are followed by papers dealing with different aspects of the space shuttle and of unmanned space transportation systems, including Ariane.

The second part deals with the exploitation of the fourth environment as embodied in the four generations of peaceful utilizations of space technology. Thus, for the first generation (telecommunications) there are papers on the technology of communications satellites, on domestic and regional systems such as Intelsat 5 and the European L-sat and on direct broadcast system. For the second generation, (TEE observations of the three earth environments) there are papers giving the state of the art for active and passive microwave sensors and other advanced sensors, or describing plans for such systems as the first ESA remote sensing satellite (ERS). For the third generation (science and processes in space) only the life sciences part is represented in this book of proceedings with a paper on the near-space radiation situation and another one presenting the main results from medical investigations during the 185-day flight of the Salyut-6-Soyuz. The materials and fluid sciences part is covered in a special issue of Acta Astronautica. Finally, the last papers pertain to the fourth generation: energy from space and for space.

The third part deals with the exploration of the fourth environment and with a

number of aspects related to "space civilization". There are papers on Voyageur mission to Saturn, on Galileo mission to Jupiter, on Intercosmos-Bulgaria 1300 followed by papers discussing aspects of the communication (CETI) and search (SETI) for extra-terrestrial intelligence and by papers dealing with the "history" of space science and technology.

The last part surveys state of the art or advances in the supporting soft and hard technologies and contains papers on astrodynamics and propulsion as well as the prize-winning student paper.

It is perhaps appropriate to mention, for readers convenience, that additional selections of papers presented at the 32nd IAF Congress are to be found in two special issues of *Acta Astronautica*, the archive journal of the International Academy Academy of Astronautics, also published by Pergamon Press. The title of these special issues are: "Space Mankind's Fourth Environment" and "Microgravity Sciences and Processes", respectively.

In closing it is my pleasant duty to thank all those who have contributed to the production of this book of proceedings.

Luigi G. NAPOLITANO  
Editor in Chief  
Proceedings XXXIIInd  
IAF CONGRESS

## AN OVERVIEW OF UNITED STATES MANNED SPACE FLIGHT FROM MERCURY TO THE SHUTTLE

M. A. Faget

*Director of Engineering and Development, NASA Lyndon B. Johnson  
Space Center, Houston, Texas 77058, USA*

### ABSTRACT

This paper is an overview of the U.S. manned-space-flight programs beginning with Project Mercury and continuing to the present status of the Space Shuttle Program. The nature of previous programs is discussed as to design and mission philosophy, characteristics of the launch vehicle and spacecraft, mode of operation, significant flight results, and the manner in which the experience from one program affected the nature of the next. Interesting physical and performance comparisons of spacecraft, launch vehicles, and ground-support systems are discussed.

The Space Shuttle is shown to clearly represent a major departure from the trend established in previous programs. Some of the major consequences of this difference will be revealed in a discussion of the design features of the Shuttle as well as the more significant new advancements in technology that were required during the development program. Finally, some interesting aspects of the first orbital flight are discussed.

### KEYWORDS

Spacecraft design and development; flight control; launch vehicles; propulsion; trajectory analysis; ground-support systems; payloads.

### INTRODUCTION

There are many publications relating to Mercury, Gemini, Apollo, Skylab, Apollo-Soyuz Test Project, and Shuttle. Thus, the idea of condensing the American manned space flights into a short overview paper is somewhat frightening. The purpose of this paper is to present an objective review of some of the technical considerations of the design, development, and operation of these spacecraft. Although there were many triumphs, a number of surprises, and not a few setbacks during the 23-year period covered, such incidences will only be mentioned in the context of technical significance. A little more than 19 years elapsed between February 20, 1962, when Glenn first rode Mercury into orbit, and April 12, 1981, when Young and Crippen inaugurated orbital flight for the Space Transportation System. A comparison of the Mercury capsule and the spaceship Columbia reveals the extent of progress made during these first years of manned space flight.

During the early 1920's after approximately the same number of years of development, the aviation industry was also entering the transportation phase. But there is a significant difference between the development of airplanes and the development of manned spacecraft. The early airplanes were relatively cheap and easy to build. Often, only one man and at most just a handful would be sufficient to do the entire design job. Consequently, a great number of airplanes were built and flown. This led to a rapid evolutionary process for both the design and the operation of aircraft with a number of accidents and fatalities providing a strong corrective influence on any misdirection. The philosophical basis for the design and operation of spacecraft has had to progress without this impartial and unerring guidance of "survival of the fittest." As a substitute, we have had to rely on intensive analysis, extensive testing, imperfect simulations, and the judgment and experiences of the hundreds of key people working on these programs.

American manned spacecraft by type together with a number of weights, dimensions, and features that characterize them are listed in Appendix A. Mercury was a first of a kind design with Gemini and the Apollo command and service module (CSM) (Fig. 1) following in an evolutionary trend. The Apollo lunar module (LM) and the Space Shuttle, on the other hand, represent new and distinctive designs with unique features for which no prior art existed.

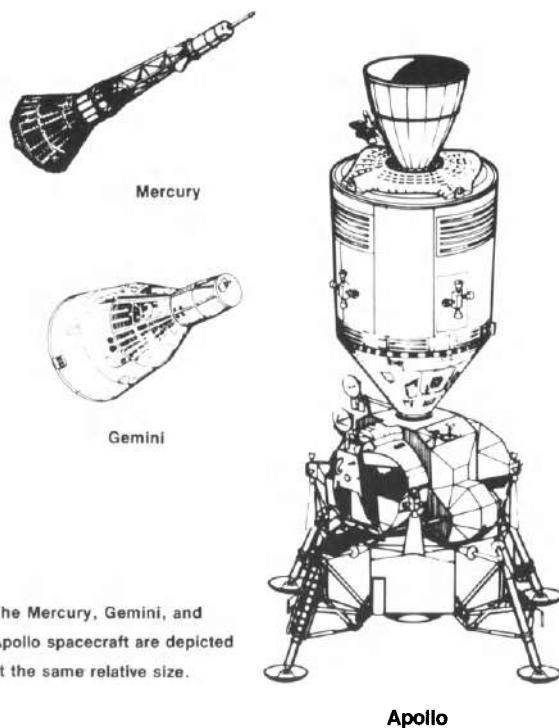


Fig. 1. Mercury, Gemini, and Apollo spacecraft.

Project Mercury

The National Aeronautics and Space Administration was established in October 1958, and within a few weeks, Project Mercury was initiated. In order that the program could be executed as quickly as possible, it was deemed desirable to choose an existing missile for the launch vehicle. The Atlas, considered to be the most powerful missile that would be available during the desired time period, was chosen. Conservative estimates indicated that the Atlas would be capable of orbiting approximately 900 kg (2000 lbm) of payload. Therefore, the spacecraft would necessarily be very small. Actually, the weight of the Mercury spacecraft grew almost 50 percent during its development period. When finally developed, the entry weight of the Mercury spacecraft without the retrorockets was in excess of 1180 kg (2600 lbm). Fortunately, the Atlas performance also grew sufficiently during this period.

The basic purpose of Project Mercury was to expose several test pilots to orbital flight and to have them evaluate the experience in order that future more substantive programs could be planned. To move the program rapidly and because of severe weight constraints, the design of all systems was as simple as possible providing only the basic necessities of launch, a short orbital flight, and safe descent and landing. A simple ballistic configuration designed to minimize reentry heating was chosen. Although kept as simple as possible, each active system had at least one level of redundancy. Descent was initiated by firing three solid retrorockets; however, a safe reentry was assured if only two of these rockets fired. A parachute system that was backed up by a reserve system of identical design provided a safe landing on either water or reasonably level terrain. Redundant sets of hydrogen peroxide monopropellant reaction control jets were used. To guard against electrical failure, one set of jets had its electrically actuated valves backed up with a set of mechanical valves directly linked to the astronaut's rotational hand controller. Life support was provided by a pure-oxygen cabin atmosphere at one-third sea-level pressure. This was backed up by a pressure suit that would automatically inflate if the cabin atmosphere dropped to less than one-fourth sea-level pressure. Redundant voice, telemetry, and command signals could be sent and received over a number of communication channels. The Mercury capsule also carried both C-band and S-band radar beacons to assist the radars at the various ground stations in tracking the flight. Finally, power was supplied by redundant battery sets with independent buses. Heat was dissipated from the Mercury spacecraft by the use of a water evaporator.

The concept for flight control was very simple. Basically, the Mercury spacecraft was inserted into a low Earth orbit using the launch vehicle guidance system. Once in orbit, no further velocity change maneuvers were required until it was time for descent. Return to Earth was accomplished by maneuvering the spacecraft to retrorocket firing attitude and then firing a cluster of three solid rockets. This deflected the flightpath to one which entered the Earth's atmosphere. Since the spacecraft was designed to produce no lift, it followed a highly predictable ballistic entry trajectory. Consequently, the time at which the retrorockets were fired primarily determined the location of splashdown in the ocean. It was recognized that there would be a fairly large dispersion about the planned landing location. However, consideration for emergency descent or aborts during launch which could result in a landing anywhere along the flight track made survival for a period of time on the water after landing and extensive use of location aids a basic design requirement anyway.

Flight control equipment onboard the spacecraft was needed only for attitude control and firing the retrorockets. These functions could be done both manually and automatically. When in automatic flight, an autopilot and two horizon scanners were set up to maintain the vehicle in a fixed attitude with respect to local vertical. Also, an onboard timer initiated the retrofiring sequence after commanding

the attitude to the optimum one for the maneuver. This timer, which was started at lift-off, could be corrected during the flight by command signals from the ground. The astronaut could also control attitude with a hand controller either by using an attitude indicator on the instrument panel or by looking through the window. He could also override the automatic initiation of the retrorocket sequence. The astronaut could crudely determine his position in orbit by comparing his view of the Earth with a clock-driven replica of the Earth's globe.

The control of the mission was performed on the ground. Communications with the spacecraft and tracking data were obtained from a network of stations along the path of the first three orbits. There were 16 different stations so located to allow a maximum of 10 min without communication contact. The data from these stations were sent to the Mercury Control Center at the launch site in Florida. It was on the basis of these processed tracking data that orbital ephemeris was determined. The location of intended splashdown was predetermined to accommodate postretrofiring tracking and thereby to enhance final location for recovery.

During Project Mercury, a total of six manned tests was made. In preparing for these manned tests, a total of 17 unmanned flight tests was made, four of which were made with primates aboard the spacecraft. Only 10 of these 17 tests were successful. In addition to the biomedical studies, these tests provided engineering studies on all the modes of flight such as escape system operation, entry flight, and performance of the various onboard systems. Of the four manned orbital tests, each was of increasing duration starting with the three orbits made by John Glenn and ending with the 22-orbit flight made by Gordon Cooper. A summary of all U.S. manned space flights is listed in Appendix B.

#### Gemini Program

Based on the success of Mercury and undisputed evidence confirmed by the Russian program that man could function comfortably and effectively in the environment of space, it became obvious that the next step would be to develop the role of man in space. Consequently, two new programs were established during 1961 to this end: Gemini, to explore the possibilities and limitations of manned operation in space, and Apollo, a bold program committed to manned exploration of the Moon. In Gemini, the approach was to quickly and, where possible, directly capitalize on what was learned in Project Mercury. However, a vastly improved capability was desired to explore the feasibility of the following operations: maneuvering in orbit, rendezvous and docking with another vehicle, extravehicular activity (EVA) by the astronauts in pressure suits, guided flight during entry to precisely designated target areas, and establishment and refinement of flight operation methods. All these objectives were met with highly positive results that were applied extensively in the Apollo Program and contributed greatly to the success of that program.

The Gemini Program employed the new Titan II launch vehicle. This booster, with roughly twice the capability of the Atlas, made it possible for the Gemini spacecraft to carry a crew of two as well as to incorporate a great number of design features requisite to the program requirements. The Gemini configuration was basically a scaled-up version of the Mercury vehicle. However, the center of mass of the Gemini spacecraft was offset from its centerline. This configuration caused the spacecraft to trim at a slight angle of attack, which produced a small aerodynamic force (lift) normal to the drag force. Although the lift-to-drag ratio ( $L/D$ ) was only slightly greater than 0.1, it was more than sufficient to steer the entry path to a desired touchdown point. Steering was done by controlling the roll attitude of the spacecraft and thereby the lift vector to deflect the reentry path (up, down, right, or left) in the desired manner.

Gemini carried its adapter into orbit. This was the conical structure that connected the Gemini spacecraft to the booster. In this adapter section, which was jettisoned before entry, a great number of systems that were used during orbit were carried. For instance, 16 reaction control thrusters were installed in the adapter. These bipropellant rocket motors provided for both translational and rotational maneuvers. Fuel cells and their hydrogen and oxygen supply tanks were installed in the adapter as were radiators that provided for heat dissipation. The fuel cells and radiators made mission durations as long as 2 weeks possible. The Gemini spacecraft was equipped with an inertial platform and a digital computer to aid in orbital navigation and to provide for reentry navigation and guidance. In the nose of the Gemini capsule, rendezvous radar and docking equipment were installed. Since the hypergolic propellants used in the Titan II launch vehicle were determined to only produce a low-yield explosion in the event of a launch vehicle breakup, ejection seats could be used instead of an escape rocket. Finally, to accommodate extravehicular activity, a hatch which was operable while in orbit was provided over each astronaut's position.

In contrast to the great number of unmanned flights that preceded manned operation in Project Mercury, there were only two unmanned flights in the Gemini Program. The purpose of these flights was primarily to check out system operations and the compatibility between spacecraft and launch vehicle.

The Gemini Program also included a target vehicle to accommodate rendezvous operations. This target vehicle was a converted Agena propulsion stage that was launched into a rendezvous-compatible orbit by an Atlas launch vehicle. The procedure was to launch the target vehicle shortly before the Gemini was launched and then to maneuver the Gemini through a series of rendezvous maneuvers until the Gemini was brought within a few feet of the target vehicle. After inspection, the Gemini was maneuvered into a docking position and the docking maneuver was made. After completion of the docking, the Agena propulsion system could then be controlled by the crew onboard the Gemini. The Agena propulsion stage could be used to make several maneuvers while the Gemini was attached; for instance, raising the orbit's apogee to much higher altitudes than could otherwise be achieved. There were 10 manned flights of Gemini during a 20-month period. All these flights were successfully completed, and a great deal was learned during the Gemini Program that was later to be used in the Apollo Program, particularly for extravehicular activity, rendezvous techniques, mission control procedures, reentry control, and postlanding recovery operations.

#### Apollo Program

Although Project Mercury was still in the early development phase, NASA started considering more advanced manned missions during the spring and summer of 1959. The most obvious and appealing prospect was for some type of lunar mission. Three types of manned missions were considered. They were, in order of increasing difficulty, circumlunar flight, lunar orbit, and lunar landing. Regardless of the ultimate goal, it was generally felt that the first flight would be circumlunar with the spacecraft passing within several hundred kilometers of the lunar far side. After this, orbital flights would be made using the same outbound and homeward navigation techniques proven in the circumlunar missions. Finally, a lunar landing would be made by descending from lunar orbit. By this scheme, each mission would be an extension of the previous one; thus, the overall difficulty of achieving the final goal would be divided into a number of incremental steps, each with a greatly reduced exposure into the unknown. Nevertheless, the first time a plan to make a manned landing by descending from lunar orbit was outlined to NASA management, several in the audience severely questioned the wisdom of not taking advantage of the experience that would be obtained from Surveyor, which was designed to go directly from the Earth straight down to the lunar surface. Clearly, they had not con-

sidered the excitement of the crew during a landing that started at hyperbolic velocity in a near-vertical approach and that would be fully committed before they knew whether the landing rocket would fire. This incident is mentioned to illustrate that at the same time that a manned lunar landing was seriously being debated, the basic understanding of the venture was quite primitive.

Since it required the least amount of propulsion and the least sophistication in navigation and guidance equipment onboard the spacecraft, and since it clearly seemed the least hazardous, the first mission seriously considered was simple circumlunar navigation and return. This seemed to be modest extension of orbital flight; in fact, circumlunar flight is achievable by a highly eccentric Earth orbit of the proper parameters. However, the gravity field of the Moon creates a major influence on such orbits. Consequently, even the smallest error in state vector at the time of translunar injection could not go uncorrected if a safe entry into the Earth's atmosphere was to be made at the end of the mission. The question was how to determine the error and how accurately the corrections could be made. There was real concern that a safe entry into the Earth's atmosphere at lunar-return velocity might be beyond the guidance and navigation "state-of-the-art" technology. At these velocities, during the initial stages of entry, the spacecraft must pull negative lift to skim along a very narrow corridor of the upper layer of the Earth's atmosphere. If the upper boundary of this corridor were exceeded, the spacecraft would skip out of the atmosphere back into a highly eccentric orbit and perhaps expend its supplies before entering the atmosphere a second time. On the other hand, if the corridor boundary were missed on the lower side, the spacecraft would exceed the heating or load limitations of its structure. An associated concern was entry and landing point location. Since the mission was not very well understood, it was conjectured that the time of return might vary greatly from the planned time and, because of the Earth's rotation, the geographical position of the entry might have a large dispersion. For these reasons, configurations with a fairly high lift-to-drag ratio appeared desirable. In summary, the thinking in 1959 was that from a flight control standpoint, the circumlunar mission would be flown using ground-based navigation obtained from tracking data. Guidance instructions would be transmitted to the crew for the necessary midcourse corrections. A budget for velocity changes of 152 m/s (500 ft/s) was initially estimated for this purpose. The aerodynamic configuration of the entry vehicle was estimated to need a lift-to-drag ratio of better than 1 to provide a large maneuvering footprint while safely staying within the entry corridors.

During 1960, enough studies had been performed by NASA and industry to achieve a fairly good understanding of the implications of various manned lunar missions. Lunar-orbit missions with a total flight duration of 2 weeks received a great deal of interest. Such a mission did not appear to be a great deal more difficult than circumlunar flight but would provide much more scientific data. Furthermore, it would provide the means for gaining significant flight operational experience and reconnaissance data that would support a future lunar landing. The biggest hindrance to enthusiastic support for lunar landing was the enormous size launch vehicle that would be required. Another consideration was that features of the lunar surface were known to a resolution of no better than 1 km. Thus, the roughness and soil properties of the surface on which a landing would have to be made were woefully unpredictable.

The Apollo Program got its official blessing and start when, on May 25, 1961, President Kennedy said, "...I believe that this Nation should commit itself to achieving the goal, before this decade is out, of landing a man on the Moon and returning him safely to the Earth." This commitment precipitated a great number of decisions. The guidance and control precision for atmospheric entry at lunar-return velocity was sufficiently well established to commit to an entry configuration that would have an L/D of 0.5 instead of a value of 1 previously mentioned. The selected 0.5 L/D value was compatible with the use of a semiballistic entry configuration de-

sign. Such configurations could achieve the relatively low entry heating of high-drag ballistic designs with a modest amount of lift. Furthermore, these features could be embodied in an axisymmetrical shape, which simplified a number of design, manufacturing, and test considerations. The design chosen for Apollo was a derivative of the Mercury shape. By offsetting the center of mass a distance of 19 cm (7.5 in) from the centerline, Apollo would trim at about 33° angle of attack, which was sufficient to produce the desired L/D of 0.5. With this much lift, Apollo could be confidently guided to land 9260 km (5000 nautical miles) downrange of the entry interface. On the other hand, the landing point could be limited to only 1482 km (800 nautical miles) downrange without exceeding 4g deceleration. However, an interesting thing happened between preliminary design and final assembly. As various equipment was stuffed in the entry capsule, the center of mass inexorably moved toward what one of our engineers called the "idiot point." That is the center of volume. Consequently, the center of mass ended up displaced only a little more than 12.7 cm (5 in) from the centerline and the resulting L/D was 0.35. However, this was more than sufficient. Planned splashdown for all missions actually flown was set for 2593 km (1400 nautical miles) downrange with the never used capability to either decrease it to 1482 km (800 nautical miles) or increase it to 4074 km (2200 nautical miles). It should be mentioned that only once in all the returns from the Moon was it deemed desirable to move the preplanned landing point. It was moved 926 km (500 nautical miles) further downrange to avoid the possibility of predicted bad weather at the previously chosen landing point. However, the decision to relocate was made early and the change was accomplished by a propulsive maneuver during transearth coast a day prior to entry. Thus, the actual entry was a standard one with the nominal downrange distance.

The decision by President Kennedy to make a lunar landing the principal space effort of the decade precipitated the famous debate on the mission scheme to be employed. The NASA, industry, and the U.S. Air Force at that time were studying a great variety of launch vehicles. The most mature of these studies was devoted to the Saturn series of launch vehicles. The NASA was also studying a larger launch vehicle, called the Nova, but this was planned to come after the Saturn.

The first lunar landing missions studied required a launch vehicle more powerful than the largest Saturn under consideration. At that time, it was envisioned that the entire spacecraft would land on the Moon. Since there was insufficient confidence that a Nova-class launch vehicle could be built, attention was turned to employment of rendezvous in lunar orbit or Earth orbit as alternatives that would fit the mission within the capability of the largest feasible Saturn-class launchers. Without getting into the multitude of considerations that finally settled this sticky situation, lunar-orbit rendezvous was chosen. From a mission planning and guidance and navigation standpoint, lunar-orbit rendezvous was completely compatible with all the work that had taken place to the time of that decision. The requirement of rendezvous in lunar orbit during the mission had a major impact on onboard equipment and operational techniques associated with rendezvous navigation. Gemini was extremely valuable as a tool for developing flight control techniques and procedures for orbital rendezvous. Furthermore, the general philosophy of the interplay between the Mission Control Center in Houston and the astronauts in the spacecraft was developed during the Gemini Program.

The decision to employ lunar rendezvous meant that Apollo would require two spacecraft: the command and service module for flight to lunar orbit and back, and the lunar module for descent to the lunar surface and return to lunar orbit. In recognition that the Apollo mission would necessarily be one of extended duration, a crew size of three was chosen, partly to accommodate the 4-hr-on, 8-hr-off duty cycle wherein one crewmember would be on duty at all times. However, based on Gemini experience, the crew stated a decided preference to sleep, eat, and be on duty at the same time. Accommodating this preference to the mission was not considered unsafe. Ground controllers were easily able to monitor the condition of

the spacecraft systems during crew off-duty periods. Nevertheless, the choice of a three-man crew was well suited to the purposes of the program, and, as a matter of fact, it would have been extremely difficult to have performed the necessary tasks with less than three. With this number, one crewman was left in orbit aboard the CSM while the other two members descended to the lunar surface in the LM. It was deemed highly desirable to have two crewmen on the lunar surface particularly during the extravehicular activity since this accommodated the buddy system, which materially added to the safety of that operation.

The arrangement of the CSM was quite similar to that of the Gemini spacecraft. The command module (CM) housed the crew and was the only part designed to reenter the atmosphere. Like Mercury, it was equipped with a launch escape rocket. Although the configuration was similar to Mercury, the conical afterbody was blunter to minimize the heating on this portion at the design angle of attack of 33°. A docking mechanism and tunnel were located at the apex of the cone to accommodate docking with the LM and crew transfer. The command module was much roomier and considerably more comfortable than the Gemini vehicle. The crew of three sat side by side in acceleration couches during launch and entry. During other periods of flight, the center couch was removed to increase room for crew activity. There was sufficient room under the remaining couches for two crewmembers to sleep.

The service module, like the Gemini adapter, provided for the propulsion and electric power during the mission. It was equipped with 16 reaction control thrusters arranged in four identical modules. These provided the thrust for rotational and minor translational maneuvers. The service propulsion system was equivalent to a propulsion stage. It was used during the mission for several minor maneuvers and two major ones: insertion into lunar orbit and departure from lunar orbit into transearth flight. The service module also contained a thermal radiator and a high-gain S-band dish antenna.

The LM was really a two-stage vehicle. The descent stage was equipped with a throttleable descent engine, a landing radar, and a four-legged landing gear. The legs had an extra large spread to accommodate the low impact stability associated with the very low gravity of the Moon. Each leg was tipped with a large dish-shaped pad, which was designed to accommodate the unknown bearing pressure of the lunar soil. The pad was socket-mounted to allow skidding at touchdown in any direction with little chance of tripping. The descent stage was designed to act as the launch platform for the ascent stage at the time of lunar surface departure.

The ascent stage housed the ascent propulsion system and the LM cabin. There were accommodations for two crewmen, who were positioned at their flight stations in the standing position to maximize downward visibility during the landing maneuver. The crewmen slept in hammocks while on the lunar surface. Like the Mercury and Gemini spacecraft, the CSM and the LM had a pure-oxygen atmosphere at a pressure of one-third sea-level pressure. The ascent stage was equipped with a rendezvous radar, 16 reaction control thrusters, and a high-gain S-band dish antenna. The LM was battery-powered and water-cooled. The ascent propulsion was used to launch the ascent stage from the lunar surface into a rendezvous orbit. It was also available to abort the landing maneuver into a rendezvous orbit in the event the landing could not have been made.

After the CSM had established the desired orbit about the Moon, the LM pilot and the commander transferred to the LM and powered it up. The LM was then separated, and the CM pilot was left alone in the CSM until the other crewmen returned from the lunar surface. The LM then made a phasing maneuver, by which its orbit relative to the CSM was lowered. This maneuver placed the LM sufficiently ahead of the CSM so that an aborted powered descent would terminate into a rendezvous-compatible orbit. At the proper time, powered descent was initiated with the descent propulsion system at near full throttle. Slight throttle adjustments were made during

descent to keep the descent trajectory on target. As the targeted landing area was approached, the LM was pitched over while, at the same time, the engine was throttled back. This allowed the crew to achieve visual contact with the landing area several minutes before touchdown. At this time, the commander was able to determine the touchdown point that was being targeted by the onboard guidance system. He did this by looking through a scale painted on his window called the landing point designator. He could then override the guidance system until the landing point designator indicated a desirable landing location. During this period, altitude and descent rate were sensed by the landing radar and control by the guidance system was continued. Hovering flight was achieved several hundred feet above the lunar surface. The final descent and touchdown were then made under control of the commander. As the LM approached the surface, a considerable amount of dust was blown up; but some visibility remained, and good landings were made in all cases. It should be mentioned that several turbofan-powered hovering vehicles designed to mimic the landing characteristics of the LM were built. These lunar landing training vehicles were used extensively by the crew to sharpen their proficiency before each lunar landing mission.

While on the surface of the Moon, the crew used extravehicular mobility units (EMU's) to move about. The EMU's each consisted of a space suit pressurized to 25 kPa (1/4 atm) and a battery-powered backpack. In many ways, the EMU's functioned like miniature manned spacecraft. They provided a cooled, revitalized atmosphere for the astronaut to breathe. The astronaut was kept at a comfortable level by an actively cooled garment through which water was circulated at a regulated temperature. The EMU was equipped with redundant two-way communication links that provided voice communication between the two astronauts and, using the LM communications equipment as a relay station, communication with the Mission Control Center on Earth. Ground controllers were also able to monitor the physical condition of the astronauts and the performance of the equipment by telemetry from the EMU.

The space suits provided sufficient mobility to accommodate a range of activity on the lunar surface such as setting up science stations, taking photographs, and performing short geological traverses for sample selection. On later missions, lunar surface exploration was greatly enhanced by the Rover, an electric-powered four-wheel vehicle that could carry the astronauts and a considerable amount of equipment. The Rover carried two-way communications equipment and a high-gain antenna which provided sufficient bandwidth for a direct link to Earth carrying color or television transmissions as well as relaying the communications from the astronauts to Earth.

It was recognized that both the LM and the CSM must be made considerably more reliable than previous spacecraft for several reasons. Compared to an orbital flight, a journey to the Moon implied a much greater commitment to the hazards of space. As a consequence of the complexity and the time required to abort a lunar mission, both the LM and the CSM had backup navigation and guidance systems, adequate reserve propellant, and levels of redundancy that were carefully determined to be sufficient. All possible failure modes were analyzed so that single-point failures were identified and eliminated or safely accommodated. Furthermore, operational methods were worked out to circumvent failures or to ameliorate the conditions that might be brought about by failures.

Although tracking from the ground was chosen as a primary method of navigation during the mission, it was decided that the spacecraft should be capable of autonomous navigation and guidance to safely return in the event all communication links were lost. The onboard navigation and guidance system was also used in the verification of the accuracy of various maneuvers. Navigation in space, as on the ocean, requires precise position determination. To this purpose, the Apollo command module was equipped with a sextant to measure the angle between a number of preselected landmarks on both the Earth and the Moon and certain catalogued celestial bodies.

Sightings and the time of sightings were fed into the onboard computer, which was preprogramed to solve the navigation problem. The spacecraft were also equipped with inertial measurement units, which, together with the computer, performed the automatic guidance function.

In addition to the launch vehicle and the LM and CSM, there were two other vital elements to successful execution of the lunar missions. These were the Manned Space Flight Network (MSFN) and the Mission Control Center including all the mission procedures and time lines that were worked out by the Mission Control Center personnel. The MSFN was modeled after the Deep Space Information Facility (DSIF) developed by the Jet Propulsion Laboratory. The MSFN consisted of three 25.9-m (85 ft) diameter antennas located at approximately  $120^{\circ}$  intervals of longitude around the Earth to provide coverage of the mission.

Both the MSFN and the hardware onboard Apollo were capable of producing highly accurate navigation data. Data from both sources were compared before making any velocity change maneuver. Also, immediately after the maneuver was made, data were again cross-checked. Navigation done on the ground had the benefit of a large complex of powerful computers. Furthermore, at least two S-band trackers were always available as data sources. The data from the S-band tracker were extremely accurate. In addition to providing a Doppler count for velocity, the carrier signal was also phase-modulated with a pseudorandom noise (PRN) code for range measurement. This digital signal, which was nonrepetitive for 5.5 s, was turned around and retransmitted on another carrier by a transponder on the spacecraft. Distance measurements with an accuracy of about 10 m could be obtained. Velocity measurements were much more useful. High-powered data processing techniques could produce an accuracy better than a millimeter per second by smoothing Doppler data over a period of 1 min. With such data, extremely accurate state vectors could be obtained not only on translunar and transearth flight but also while Apollo was in lunar orbit. This capability was important since lunar gravity anomalies and venting from the spacecraft continually perturbed the orbit. Computational techniques were developed to the point at which tracking data obtained from the lunar module during its landing descent burn could be processed to serve as a sufficiently accurate "tie-breaker" in the event that onboard primary and backup computations produced unexplainable differences.

The general approach to mission planning was to break the mission down into a number of discrete events and periods. Each of these was analyzed in great detail, and a complete model of the mission to great precision was constructed before flight. When flown, the missions would usually duplicate the plan in exact detail. A feature of the planning was the inclusion of time allowances for unexpected events so that the preplanned schedule could be maintained. The advantage was that almost every event or phase of the basic mission was extremely well understood and exercised. In addition to the basic mission plan, a great number of contingency plans were available to cover every rational problem.

All missions were planned to accommodate midcourse corrections both outbound and on return. Specific times were set aside for these maneuvers: four translunar midcourse correction events were allowed and three transearth. However, if the error to be corrected was sufficiently small, the maneuver would not be made, and, as a matter of fact, many missions needed only one corrective maneuver each way. When we first started considering translunar flight in 1959, we budgeted 152.4 m/s (500 ft/s) for midcourse corrections. The estimate was down to 91.4 m/s (300 ft/s) when we actually put the program into gear several years later. When we actually started flying, the  $3\sigma$  estimate was 23.8 m/s (78 ft/s). Actually, most flights required less than 6.1 m/s (20 ft/s). For example, on the last flight, Apollo 17 executed only one correction maneuver each way: translunar, it was 3.2 m/s (10.6 ft/s); for transearth, only 0.6 m/s (2.1 ft/s) was needed.

Mention must also be made of the valuable support obtained from the unmanned lunar exploration programs. The Ranger spacecraft was a probe that transmitted a few closeup images of the lunar surface just before colliding with the Moon at high velocity. The Surveyor was a softlander that made five successful landings on the lunar surface. Subsequent to landing, the Surveyor transmitted pictures of the surrounding features of the moonscape that provided extremely useful information on surface roughness as well as the quantity and size of rocks. Just as important, engineering data obtained from the Surveyor landings was extremely valuable in verifying the firmness of the lunar surface for the landing of the lunar module. The unmanned Lunar Orbiter flights, however, were every bit as valuable to the Apollo Program. They provided high-resolution photographs of the lunar surface that were extremely useful in selection of landing sites. The cartographic quality of the photographs was more than sufficient to make accurate maps of the lunar surface that could be used for orbital navigation and for visual recognition by the astronauts in the terminal phase of their descent. Just as important, analysis of orbital tracking data greatly improved the accuracy of the lunar gravitational constant and provided valuable data on lunar gravitational anomalies, all of which facilitated translunar and lunar-orbit navigation on the Apollo missions. In fact, the accuracy of navigation techniques that were ultimately developed made it possible for Apollo 12, the second landing mission, to come to rest within a short walking distance of Surveyor III, which had landed on the Moon 2-1/2 years previously. The Surveyor's location had been identified on a Lunar Orbiter photograph by patient and meticulous study.

A total of 12 unmanned test flights was made in the Apollo development program. All were successful. Six of these were devoted to qualifying the launch escape and parachute deployment system for the command module. The others were concerned with systems tests of the spacecraft and with compatibility of the launch vehicle and the launch environment. In one of these tests, the service propulsion system was used to drive the command module back into the atmosphere with the velocity and flightpath expected during lunar return.

A total of 11 manned flights was made in the Apollo lunar program. Nine of these were journeys to the Moon, and six were lunar landing missions. These flights are summarized in Appendix B.

On January 27, 1967, the program underwent a critical setback during a countdown rehearsal for what was planned to be the first manned Apollo flight. The interior of the command module suddenly burst into flames, which trapped the crew. Grissom, White, and Chaffee were killed, and the spacecraft was destroyed. Although the immediate cause of this accident was never determined, the general cause was associated with the use of many materials in the cabin interior that were highly flammable in the pure-oxygen cabin atmosphere, which, during prelaunch conditions, was at sea-level pressure. It took the program well over a year to recover from the fire.

Extensive changes in material application were made, and in cases for which suitable replacement materials could not be found, fireproof coatings and coverings were used. Special test programs were made to certify the fireproofing program. It was found that, although good fire protection was achieved for an oxygen atmosphere at the reduced pressure of orbital flight, there were no practical solutions for the prelaunch conditions when the cabin would be at sea-level pressure. Therefore, the oxygen in the cabin was diluted with 40 percent nitrogen during prelaunch activities. After the cabin pressure dropped to one-third of sea-level pressure during ascent, there was still sufficient oxygen for breathing. Later, when the astronauts were ready to further depressurize to space-suit pressure, the nitrogen content of the atmosphere had been sufficiently purged to preclude a bends problem.

With Lovell, Swigert, and Haise on Apollo 13, which was to have been the third lunar landing mission, another accident occurred. During translunar flight as the Moon was being approached, one of the two bottles used to store cryogenic oxygen in the service module exploded. The explosion was sufficiently violent to cause the oxygen in the other bottle to begin leaking, most probably through an external line. With the exception of gaseous oxygen stored in the command module for life support during the short period of reentry flight, there was no other oxygen supply for the CSM. The lost oxygen was to be used both for life support and to power the fuel cells.

This desperate situation was countered by moving the crew to the LM cabin and powering down the CSM. Since the LM was battery-powered for only a planned 3-day period, it was configured for minimum power consumption. After getting everything in order, the crew made a velocity change maneuver to correct the flightpath to a "free-return" trajectory by which the spacecraft would swing around the Moon and head back to Earth on a path consistent with reentry targeting. Another major propulsive maneuver was made after the spacecraft was in transearth flight. That maneuver changed the trajectory to shorten the return time. Although there was sufficient oxygen for the return flight, an improvised method for adapting the command module lithium hydroxide canisters for use in the LM environmental control system had to be devised. This was done by experimenting on the ground and then communicating instructions to the crew. Fortunately, the crippled spacecraft returned safely with an unharmed but disappointed crew.

#### Skylab Program

After the completion of Apollo 17, the Apollo CSM and the Saturn launch vehicles were modified for the Skylab Program. The third stage of the Saturn V was converted to a space laboratory. The hydrogen tank was divided into laboratory, sleeping, eating, recreation, and hygiene compartments. Storage lockers, an environmental control system, and other equipment were added for the comfort and physical well-being of the crew. The oxygen tank was converted to an oversize trash container. A deployable micrometeoroid shield and passive thermal control coatings were added to the skin of the tank. A battery of telescopes designed to study the Sun were mounted on a high-precision pointing platform. There were also an airlock to accommodate EVA and a docking module to which the CSM could dock. In addition to the solar telescopes, a great variety of experimental equipment was carried including a number of Earth-pointed remote-sensing instruments and cameras.

The Skylab orbital workshop was stabilized with three control moment gyros. These could be desaturated from time to time by reaction control jets using gaseous nitrogen as a propellant. Electric power for the Skylab was provided by two photovoltaic solar power systems. One system consisting of four folding arrays was attached to the solar telescope mount. The other system consisted of two wing arrays extending from the converted fuel tank.

The Skylab orbital workshop was launched May 14, 1973, into a 50° inclination orbit. The nearly circular orbit ranged in altitude from 496 to 498 km (268 to 269 nautical miles). During launch, one of the wing solar cell arrays and the micrometeoroid shield were carried away. The other wing was jammed so that it could be only partly deployed. Since the micrometeoroid shield also was important in the passive thermal control system, its absence left the tank skin exposed. This was coated with a highly reflective coating which happened to be biased hot. The result was that the laboratory quickly became overheated reaching internal temperatures greater than 50° K.

The launch of the first crew was delayed while special repair equipment was designed, constructed, and tested on the ground. Eleven days following the launch of

the Skylab orbital workshop, Conrad, Kerwin, and Weitz were launched on Skylab 2. They were able to deploy a parasol-like sunshield over the exposed skin by which the interior temperature was reduced to 24° K. They also freed the jammed wing of the solar cell array. Although the missing wing reduced the power available from that expected, there was sufficient power to perform the planned activity.

The Skylab orbital workshop was occupied by three different crews. Each successive crew stayed a longer period of time. A great number of experiments and science investigations were performed before the Skylab orbital workshop was abandoned on February 8, 1974. It was left in a condition that would allow partial reactivation in a possible visit from the Space Shuttle. Unfortunately, its orbit decayed before that was possible.

#### Apollo-Soyuz Test Project

After a series of meetings in 1971 and 1972, the United States and the U.S.S.R. agreed to a joint manned space mission. This mission became known as the Apollo-Soyuz Test Project (ASTP). Its basic purpose was to produce hardware and operation procedures that would provide the means for one country to work with the other in future manned space missions. Particular emphasis was placed on assistance and rescue missions. Rather than merely trade design and mission control information, it was agreed that the only way to be sure assistance and rescue missions would really be feasible would be to actually make a joint mission.

It was agreed that the American Apollo and the Russian Soyuz would rendezvous and dock during a cooperative mission. To accomplish this objective, a number of technical problems had to be addressed. The primary problems were compatible communications both in space and on the ground, compatible mission control procedures, compatible docking hardware, and accommodations for the different atmosphere constituency in the two spacecraft. The Soyuz cabin was maintained at sea-level pressure with the same 80 percent/20 percent mix of nitrogen and oxygen found in air. The Apollo cabin was at one-third sea-level pressure with an atmosphere of 100 percent oxygen. This incompatibility was solved by use of a docking module that would be carried with the CSM during launch. It was carried in the adapter in a manner similar to that used to carry the LM during the lunar missions. On one end of the docking module was the international docking system, which had been agreed on with the U.S.S.R. The other end, like the LM, had the female portion of the Apollo probe and drogue docking system. The docking module had an internal volume sufficient for as many as three people at the same time. It was also equipped with bottled supplies of both nitrogen and oxygen. Consequently, it could serve as an airlock between the two spacecraft.

The ASTP mission was successfully completed in July with the launch of Leonov and Kubasov in Soyuz followed by Stafford, Brand, and Slayton in Apollo. After rendezvous, both the American and the Russian mechanisms performed successful dockings. After the final undocking, both crews spent additional time in space performing individual experiments.

#### Space Shuttle Program

The Space Shuttle system has been developed to greatly improve accessibility to space and thereby to facilitate many opportunities for space applications that would not be possible otherwise. Through lower cost and improved operational capability and flexibility, the Shuttle will bring about a wide variety of missions and activity in space that previously have not been considered worth the cost and effort. The Shuttle missions primarily fall into three general categories.

1. Missions lasting from one to several weeks and employing one or more of the Spacelab modules being developed by the European Space Agency - These missions will consist of a variety of experiments and observations requiring extensive participation by the crew.
2. Missions employing one or more additional propulsion stages to carry spacecraft to orbits beyond the performance capabilities of the Shuttle alone - Most of these missions will be aimed at geosynchronous orbit, but some planetary exploration missions will also be launched by the Shuttle.
3. Missions in which the Shuttle will deploy satellites directly into orbit - In subsequent flights, the Shuttle may retrieve or service these satellites, thereby adding a new dimension of utility. In some cases, the satellites will be equipped with modest propulsion capability allowing them to move back and forth to orbits beyond the operating altitude of the Shuttle.

Most Shuttle missions probably will be a combination of these three general types to enable more comprehensive use of the cargo load of each flight.

When operational, the Shuttle will be launched from both the east and west coasts of the United States. When launched from Cape Kennedy, orbits with inclinations from 28.5° (the latitude of the launch site) to 56° will be achievable; when launched from Vandenberg Air Force Base (VAFB), launches at inclinations higher than 56° will be achievable. These would include both solar and Sun-synchronous orbits. The Shuttle will be capable of carrying as much as 29,000 kg (32 tons) of payload into low-inclination, low-altitude orbits and about one-half that amount into Sun-synchronous orbits. It will be capable of returning as much as 14,500 kg (16 tons) from orbit.

The Space Shuttle in the launch configuration consists of four major elements: an Orbiter, an external tank, and two solid-rocket boosters. The boosters provide the majority of the thrust for the first 2 min of flight. Each booster weighs about 567,000 kg (1,250,000 lbm) and produces a peak thrust in excess of 11.1 MN (2,500,000 lbf). The boosters are jettisoned after burnout and are recovered for reuse after descending by parachute into the ocean near the launch site.

The external tank contains 530 m<sup>3</sup> (140,000 gal) of liquid oxygen and 1438 m<sup>3</sup> (380,000 gal) of liquid hydrogen. These propellants, with a combined weight in excess of 680,400 kg (1,500,000 lbm), feed the three main engines, which are carried in the Orbiter. Eight minutes after lift-off, when main propulsion burnout occurs, the Shuttle is just short of attaining orbital velocity. The tank is then jettisoned and follows a long, shallow trajectory ending in a remote portion of the Indian Ocean. The tank is almost entirely destroyed during entry, although a few dense components probably survive. After the tanks are jettisoned, the Orbiter employs its orbital maneuvering engines to propel itself into orbit and for other maneuvers in space during the remainder of the mission.

The Orbiter is clearly the most complex part of the Shuttle and provides control of the boosters and the external tank during launch. The Orbiter is designed to be completely reusable with only a minimum amount of refurbishment and maintenance required between flights. Almost all its systems should be capable of functioning reliably after a hundred or more flights. Insofar as practical, development of new systems or technology was avoided in the design of the Orbiter. The Orbiter's electric power is produced from hydrogen and oxygen in fuel cells that are an improved version of those used during the Apollo Program. The basic structure is primarily aluminum with skins, stringers, frames, and ribs in an arrangement typical of any large airplane. The landing gear and the cockpit arrangement are also quite similar to those found on large aircraft. Hydraulic power is used to move flight control surfaces, to lower the gear, and for ground steering and braking. However,

the Orbiter is not by any means just a super-high-speed airplane. It is also the key element of a highly complex launch vehicle and a unique and highly versatile manned spaceship. It is the combination of these functions - launch vehicle, spacecraft, and airplane - that possibly makes the Orbiter the most complex machine ever created by man. Consequently, major advancements in the state of the art were necessary in several areas. These advances were most pronounced in flight control, thermal protection, and liquid-rocket propulsion technology.

The external tank is 45.7 m (150 ft) long and 8.5 m (28 ft) in diameter. The structure of this huge tank weighs only slightly more than 31,750 kg (70,000 lbm) and the tank is consequently quite flexible. Nevertheless, the tank is the central element of the launch configuration, which has a combined weight of nearly 2,000,000 kg (4,500,000 lbm). The consequence is the creation of a number of easily excited low-frequency structural modes. Therefore, the flight control system employs sophisticated mode-suppression techniques in gimbaling the rocket engines as the vehicle is steered during launch.

Flight control during entry is even more complex. Unlike previous manned spacecraft such as Apollo, the Orbiter actually is flying during atmospheric entry and depends on active aerodynamic control surfaces to maintain a stable attitude and control of its flightpath. When aerodynamic flight begins, it is flying four times faster than the X-15, the previous speed record holder.

The Orbiter's configuration features a double delta wing, a high vertical fin, a huge cargo compartment, and a large rear end containing an assembly of rocket engines. This configuration is based on a compromise satisfying a large number of requirements. The cargo bay size was derived from a survey of potential traffic which clearly indicated that space cargo would in general consist of low-density objects. Large main propulsion engines were desirable to minimize the size of the boosters, which represent a major part of the cost per flight. The result was a center of mass that is unusually far aft. Depending on mass and location of cargo, the Orbiter occasionally will encounter neutral to slightly negative static stability about the pitch axis at low speeds. The combination of the large, deep fuselage and the far aft center of mass resulted in the need for a large vertical fin. For control during aerodynamic flight, the Orbiter is equipped with very large elevons on the wing trailing edge, a large flap extending from the rear of the fuselage, and a combination rudder and speedbrake on the trailing edge of the vertical fin.

Return from orbit is initiated by making the typical retrograde propulsive maneuver. The Orbiter then loses altitude and gradually enters the atmosphere. As aerodynamic forces gradually increase, the control of the vehicle becomes manageable using aerodynamic control surfaces, and, with the exception of those controlling yaw, the reaction control jets are deactivated. The rudder, which might be expected to control yaw, is ineffective as a control surface on the Orbiter until it has decelerated to less than one-tenth of orbital velocity. The lack of rudder effectiveness is associated with the Mach number and the angle of attack as well as the peculiarities of the Orbiter's configuration. During the maiden flight of Columbia with Young and Crippen in the cockpit, both yaw jets and the rudder were active for yaw control from Mach 3.5 to Mach 1.0. Below that speed regime, the yaw jets were deactivated and only aerodynamic control surfaces were employed in the flight control system.

Reentry at the end of the flight was made with a preselected angle-of-attack ( $\alpha$ ) profile. An angle of attack of  $40^\circ$  was maintained until Columbia had decelerated to  $M = 11$ . From that point on,  $\alpha$  was decreased continually passing through  $M = 1$  at  $\alpha = 8^\circ$ . The approach and landing maneuvers, in the subsonic flight regime, ranged in  $\alpha$  from  $4^\circ$  to  $10^\circ$ . The acceptable reentry altitude corridor was approximately plus or minus 1500 m (5000 ft) of the nominal flightpath. The Orbiter was

guided to stay within this corridor while flying on the preselected  $\alpha$  profile by varying the bank angle. The nominal bank angle at the start of reentry was  $80^\circ$ . This value gradually diminished to about  $45^\circ$ . During the first flight, the landing site was essentially straight downrange. Consequently, some portions of reentry were flown banked left and others banked right. In this case, bank-reversal maneuvers were made near Mach numbers of 18, 9, 5, and 2.5. In future flights, longer portions of reentry will be spent banked in one direction or the other to accommodate landing at sites which may be displaced as much as 2037 km (1100 nautical miles) on either side of the straight downrange track. Downrange navigation is achieved by flying nearer to the lower or upper bound of the corridor. This can be accomplished by using slightly greater or less than nominal bank angle as the case may require. Except for short periods of control stick steering by the commander, the entire descent was flown by autopilot. John Young went to control stick steering during the last two bank-reversal maneuvers. This allowed him to enter the bank-reversal maneuver more gradually than possible with the maneuver programmed in the autopilot. Recent simulations had shown that the stability margins could be increased by decreasing the roll-rate acceleration during this flight regime. The decision to have the crew make this maneuver avoided modifying mature software. During subsonic flight, Young went once more to control stick steering and made an excellent approach and landing.

The Orbiter configuration has been subjected to more hours of wind tunnel testing than any other flying machine. This was necessary since the Orbiter flies over a greater range of Mach numbers, Reynolds numbers, dynamic pressures, and angles of attack than previous machines. More importantly, during its first flight, it was totally committed to successfully negotiating this wide and diverse range of flight conditions before it could land. This was a unique situation not previously encountered in modern aviation flight testing. High-performance airplanes and in fact virtually all newly designed airplanes are extensively flight-tested before maximum performance is approached. Each flight test is a cautious extension of previously encountered conditions and is followed by a thorough analysis of test results from which the next safe increment in the flight-test program can be defined.

As a substitute for testing in actual flight, the Orbiter's flight control system had been extensively tested in computer-driven simulation facilities. Mathematical models resident in the facility software mimic the atmospheric conditions of flight including gusts and crosswinds. Other models represent the wind-tunnel-derived aerodynamic responses to the simulated flight environment. However, wind tunnel data are of limited precision. Results obtained in different facilities sometimes were significantly different. Furthermore, experience has shown that data obtained in flight often exceed the bounds of data scatter of wind tunnel data. Extensive statistical analyses of these effects on the stability and the control surface effectiveness were made to determine possible worst case aerodynamic qualities. The Orbiter's flight control system successfully "flew" simulated entries in which such worst case conditions were modeled.

The heart of the Orbiter's flight control system is a set of five identical general-purpose computers. Each computer in the set has the capacity to control the entire flight including the guidance and navigation functions without assistance from the others. Under control of the central computers are a great number of subordinate data processors with specialized functions. The central computers formulate steering commands for the gimbal actuators during launch, the reaction control jets during space flight, and the yaw jets and aerodynamic surfaces during aerodynamic flight. These steering commands can originate in response to stick inputs from the crew or be generated within the central computers when in the fully automated flight mode. Redundant sets of rate gyros, accelerometers, inertial measurement units, star trackers, radio navigation aids, radar altimeters, and dynamic and static pressure sensors all feed data to the central computers. The more critical sensors are quadruply redundant allowing at least two and possi-

bly three sensors in a particular set to fail without degrading the performance of the flight control system.

The five centralized computers are divided into a redundant set of four with the fifth computer held in reserve as an independent backup. Normally, the redundant set will control the flight. Each of the four computers in this set is loaded with identical software, and all work in parallel with each other processing data from the sensors and transmitting commands to the controls. If one of the sensors should fail, each computer will detect it and the sensor will be deactivated. Upon completion of every computation cycle, the redundant computers make a simple arithmetic comparison of their output. Any computer failing to agree with its colleagues for two successive cycles is considered to have faulted and is deactivated. The computers also have built-in fault-detection features which should cause them to self-deactivate in most failure cases. Simple logic would indicate that this system with its depth of redundancy and its fault-detection and deactivation features should prove extremely reliable. Nevertheless, there is concern that despite all precautions, an inherent weakness could exist in the software that would ripple through the computers and deactivate them one after another. Should such an event occur, the crew would switch control to the backup computer, which is loaded with software that, although coded differently, is capable of replacing all the functions of the redundant set.

The use of digital computers in the Orbiter for flight control provides a major improvement in versatility and precision over systems used in current aircraft. The digital system facilitates redundancy management, fault detection, and fault isolation and when proven should provide a major improvement in flight safety. Although the inherent complexity of the Shuttle flight control task forced the development of an unusually elaborate flight control system, this pioneering effort may quite likely prove to be a major step toward the production of more economical and safer aircraft.

The propellant combination of liquid hydrogen and oxygen is the most energetic considered practical for use. The Shuttle, like the Saturn launch vehicle upper stages, employs these propellants for its main propulsion. However, a new high-performance engine was developed for the Shuttle (Fig. 2). It features a combustion chamber pressure at full power level of 22,409 kPa (3250 psi), or about four times higher than the Saturn engines. It also incorporates a new cycle that circumvents the wasteful use of some of the propellant just to power the turbine-driven propellant pumps. It might be mentioned that, at full power level, the total horsepower required to pump propellant on the Orbiter is greater than the total propulsion power of a Forrestal-class carrier. The Shuttle rocket engines produce 7 percent more propulsive energy per pound of propellant than the hydrogen-fueled rocket engines used on Saturn. They also have a higher thrust-to-weight ratio. A consequence of these factors is a net gain of 18,144 kg (40,000 lbm) of payload compared to what could have been obtained using Saturn-type engines. The penalty for this performance is a considerably more complex engine that has been more difficult to develop and test.

The Orbiter is easily the largest manmade object to be recovered from space. Seventeen times heavier and with 50 times the surface area, it greatly exceeds the size of the Apollo command module, which previously held this distinction. Not surprisingly, the creation of a thermal protection system suitable for reuse after multiple reentries required a new material technology. Fundamentally, the choice lay between insulating the inner structure from a hot external skin or using external insulation that would maintain the metallic skin cool enough to be part of the primary structure as in conventional aircraft. The second approach was chosen as being both lighter and less costly to produce. The cost savings were primarily associated with the straightforward use of conventional structural materials and manufacturing techniques commonly used in modern large airplanes. An equally im-

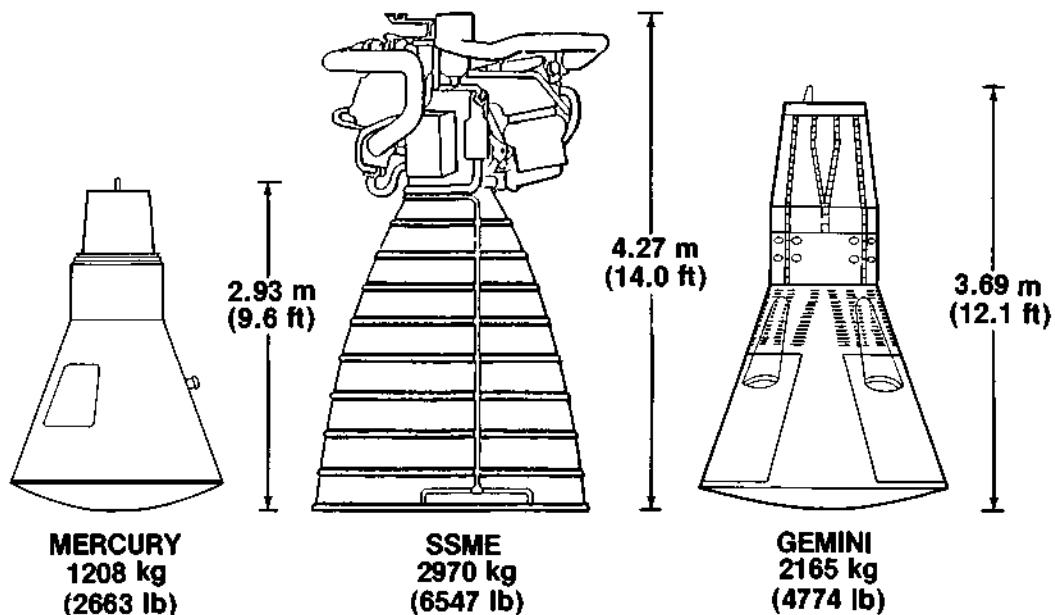


Fig. 2. Comparison of Mercury and Gemini spacecraft with Space Shuttle main engine.

portant benefit was the large data base from which the cost and weight of the major structural assemblies could be estimated.

With the exception of a few regions in which the surface temperature will not exceed 644° K (700° F), the majority of the Orbiter's external surface is covered with lightweight ceramic tiles. There are more than 32,000 tiles, most of which are 15.2- or 20.3-cm (6 or 8 in) squares. The basic material of the tiles consists of a random matrix of fine quartz fibers joined to each other at points of contact. The compaction of the fibers and therefore the density is closely controlled during manufacture. Each tile is coated with a pigmented glaze that provides waterproofing, abrasion resistance, and high heat radiation. Most of the tiles have a density comparable to balsa wood, although some tiles of a higher density are located in regions where expected loads or abuse may be greater. Since the tiles are much less capable of flexing than the aluminum skin to which they are bonded, a felt pad is inserted between the tile and the skin to accommodate relative motion between skin and tile without transmitting unacceptable loads into the tile. The tiles should be suitable for 100-mission reuse if they never exceed a surface temperature of 1533° K (2300° F). Where heating rates are higher, fewer reuses will be possible. One of the very good properties of the tile material is its capability to withstand heating rates almost double its design value and still survive one flight. This was particularly reassuring during the first flight, for which heating rates could only be predicted based on a combination of theoretical analysis and wind tunnel data. The worst case heating predictions fell well below the one-time-use capability of the material.

The Space Shuttle represents a bold and major development in the technology of flight. It will make possible achievements and activities in space that otherwise would not be possible. It could very well be the major steppingstone to a new era of space utilization.

## APPENDIX A - PERTINENT SIZE AND PERFORMANCE COMPARISONS OF U.S. SPACECRAFT

SPACECRAFT ITEM	MERCURY	GEMINI	APOLLO		POST APOLLO		SPACE SHUTTLE
			COMMAND MODULE (S)	LUNAR MODULE	SKYLAB	APOLLO SOYUZ	
<u>WEIGHT (ENTRY)</u> kg (lbm)	1208 (2663)	2165 (4774)	5663 (12,497)	NA	6088 (13,425)	5843 (12,884)	90,249 (199,000)
<u>VOLUME (HABITABLE)</u> m <sup>3</sup> (ft <sup>3</sup> )	1.02 (36)	1.56 (55)	5.94 (210)	4.53 (160)	COMMAND MODULE	COMMAND MODULE	71.46 (2525)
					SKYLAB	DOCKING MODULE	
					344.98 (12,190)	19.81 (700)	
<u>DURATION (MAXIMUM)</u> DAYS	1-1/2	13-3/4	12-1/2	3	84	9	7 TO 30
<u>CREW SIZE</u>	1	2	3	2	3.5	3	2 TO 7
<u>CABIN ATMOSPHERE</u> mm Hg (psi)	100% O <sub>2</sub> AT 250 (5)	100% O <sub>2</sub> AT 250 (5)	100% O <sub>2</sub> AT 250 (5)	100% O <sub>2</sub> AT 258 (5)	COMMAND MODULE 100% O <sub>2</sub> AT 258 (5)	COMMAND MODULE 100% O <sub>2</sub> AT 222 (4.3) TO 296 (5.6)	21% O <sub>2</sub> /79% N <sub>2</sub> AT 760 (14.7)
					ORBITAL WORKSHOP	DOCKING MODULE(1)	
					74% O <sub>2</sub> / 26% N <sub>2</sub> AT 258 (5)	60% O <sub>2</sub> / 40% N <sub>2</sub> 222 (4.3) TO 520 (10)	
<u>SUIT USAGE</u>	CABIN BACKUP	CABIN BACKUP EJECTION EXTRA VEHICULAR ACTIVITY	CABIN BACKUP EXTRA VEHICULAR ACTIVITY CREW TRANSFER	CABIN BACKUP CREW TRANSFER LUNAR SURFACE EXCURSION	CABIN BACKUP CREW TRANSFER EXTRA VEHICULAR ACTIVITY RESCUE	CABIN BACKUP CREW TRANSFER RESCUE	CABIN BACKUP CREW TRANSFER EXTRA VEHICULAR ACTIVITY RESCUE EJECTION (SOFT)
<u>PROPELLION, MAIN MANEUVERING AND RETRO</u> (AV) m/sec (ft/sec)	SOLID RETRO 98.8 (324)	SOLID RETRO 99.1 (325)	SERVICE PROPULSION SYSTEM 1951 (6401)	DESCENT PROPULSION SYSTEM 2.135 (7006)	SERVICE PROPULSION SYSTEM 533.1 (1749)	SERVICE PROPULSION SYSTEM 271.0 (889)	ORBITAL MANEUVERING SYSTEM 304.8 (1000)
				ASCENT PROPULSION SYSTEM 1850 (6070)			
<u>PROPELLION, REACTION CONTROL SYSTEM</u> FOR AUXILIARY MANEUVERS AND ATTITUDE CONTROL (TOTAL IMPULSE) newton - sec (lb - sec)	30,967 (6,900)	ENTRY VEHICLE 90,478 (20,160)	COMMAND MODULE 256,714 (57,200)	782,483 (174,350)	COMMAND MODULE 329,531 (73,425)	SERVICE MODULE 3,517,470 (783,750)	9,236,304 (2,058,000)
<u>LIFT/Drag (ENTRY)</u>	BALLISTIC	0.17 TO 0.09 (mach 24 TO 6)	0.28 TO 0.38 (mach 36 TO 6)	NA	(SAME AS APOLLO)	1.90	
<u>WETTED AREA</u> m <sup>2</sup> (ft <sup>2</sup> )	12.44 (133.9)	19.61 (211.1)	35.97 (387.2)	NA	(SAME AS APOLLO)	1034.5 (11,136)	
<u>W/S</u> kg/m <sup>2</sup> (lbm/ft <sup>2</sup> )	454.1 (93.0)	527.8 (108.1)	471.7 (96.6)	NA	(SAME AS APOLLO)	361.3 (74.0)	

NOTE:

(1) O<sub>2</sub> MINIMUM PARTIAL PRESSURE OF 165 mm Hg (3.2psi), N<sub>2</sub> VARIABLE DURING CREW TRANSFER.

## APPENDIX B - SUMMARY OF AMERICAN MANNED SPACE FLIGHTS

Spacecraft	Launch date	Crew	Flight time, hr:min:sec	Comments
Mercury				
MR-3	May 5, 1961	Shepard	00:15:22	First suborbital flight
MR-4	July 21, 1961	Grisson	00:15:37	Second suborbital flight
MA-6	Feb. 20, 1962	Glenn	04:55:23	First Mercury orbital flight
MA-7	May 24, 1962	Carpenter	04:56:05	Second Mercury orbital flight
MA-8	Oct. 3, 1962	Schirra	09:13:11	Extended duration
MA-9	May 15, 1963	Cooper	34:19:46	Extended duration
Gemini				
GT-3	Mar. 23, 1965	Grisson/Young	04:52:31	First Gemini flight
GT-4	June 3, 1965	McDivitt/White	97:40:01	20-min EVA; handheld maneuvering unit
GT-5	Aug. 21, 1965	Cooper/Conrad	190:55:14	Extended duration
GT-7	Dec. 4, 1965	Borman/Lovell	330:35:01	14-day mission
GT-6A	Dec. 15, 1965	Schirra/Stafford	25:51:21	Rendezvous with GT-7
GT-8	Mar. 16, 1966	Armstrong/Scott	10:41:26	Docked with Agena; Gemini reaction control system (RCS) malfunction; mission terminated
GT-9A	June 3, 1966	Stafford/Cernan	72:20:50	Rendezvous with damaged Agena; 2-hr 7-min EVA
GT-10	July 18, 1966	Young/Collins	70:46:39	Docked with Agena; 3 EVA periods
GT-11	Sept. 12, 1966	Conrad/Gordon	71:17:08	Reached record altitude of 1373 km (741.5 nautical miles) using Agena propulsion
GT-12	Nov. 11, 1966	Lovell/Aldrin	96:34:31	Docked with unusable Agena
Apollo				
Apollo 7	Oct. 11, 1968	Schirra/Eisele/Cunningham	260:09:08	First Apollo flight - Earth orbit
Apollo 8	Dec. 21, 1968	Borman/Lovell/Anders	147:00:42	First manned lunar orbit, no LM
Apollo 9	Mar. 3, 1969	McDivitt/Scott/Schweickart	241:00:53	First manned LM flight - Earth orbit
Apollo 10	May 18, 1969	Stafford/Young/Cernan	192:03:23	LM and CSM rendezvous and docking activity in lunar orbit
Apollo 11	July 16, 1969	Armstrong/Collins/Aldrin	195:18:35	First manned lunar landing, Sea of Tranquillity, time on Moon (T.O.M.) 21:36:21
Apollo 12	Nov. 14, 1969	Conrad/Gordon/Bean	244:36:25	Precise lunar landing near Surveyor, Ocean of Storms, T.O.M. 31:31:12
Apollo 13	Apr. 11, 1970	Lovell/Swigert/Haise	142:54:41	Lunar flyby during emergency return after service module damaged
Apollo 14	Jan. 31, 1971	Shepard/Rosas/Mitchell	216:01:58	Third landing - Fra Mauro, "golf cart" first wheels on Moon; T.O.M. 33:30:31
Apollo 15	July 26, 1971	Scott/Worden/Irwin	295:11:53	Hadley-Apennine region, used Rover - an electric-powered car; T.O.M. 66:54:53
Apollo 16	Apr. 16, 1972	Young/Mattingly/Duke	265:51:05	Descartes region, traveled 27 km in Rover, T.O.M. 71:02:13
Apollo 17	Dec. 6, 1972	Cernan/Evans/Schmitt	301:51:59	Taurus-Littrow region, traveled 35 km in Rover, T.O.M. 74:59:38
Skylab				
Skylab 2	May 25, 1973	Conrad/Kerwin/Weitz	672:49:49	Repaired damage to Skylab; initial space experiment activity
Skylab 3	July 28, 1973	Bean/Garriott/Lousma	1427:09:04	Increased duration; continued experiment activity
Skylab 4	Nov. 16, 1973	Carri/Gibson/Pogue	2017:15:31	Increased duration; completed experiment activity
Apollo-Soyuz Test Project				
ASTP-Apollo	July 15, 1975	Stafford/Brand/Slayton	217:28:24	Rendezvous and docking with Soyuz
Space Shuttle				
STS-1	Apr. 12, 1981	Young/Crippen	54:20:53	Successful first flight of spaceship <u>Columbia</u>

## SYSTEMS ENGINEERING CHALLENGES OF THE SPACE SHUTTLE

L. E. Day

*STS Systems Engineering and Integration, Office of  
Space Transportation Systems, NASA Headquarters, Washington,  
D.C. 20546, USA*

### ABSTRACT

Because of its unique configuration, compared to former and current rocket space vehicles, the Space Shuttle has presented systems engineering challenges. In the past, the analysis and integration of the several elements of rocket space vehicles was greatly aided by the symmetry of the vehicle and the simplicity of the interfaces. The configuration of the Space Shuttle with its winged Orbiter, the large propellant tank, and the two solid rocket motors strapped to the tank presented a complex vehicle for analysis and test. Accordingly, many problems involving propulsion systems, ascent flight control, aerothermodynamics, and structural dynamics have been complicated. Selected technical challenges and their solutions are reviewed.

### KEYWORDS

Space Shuttle, systems engineering of space vehicles, manned space flight, reusable space vehicle

### INTRODUCTION

Two decades of space vehicle development have built on the technology and experience of rocket vehicles originally designed to serve as weapons. Earlier manned space vehicles were adapted from these rockets by increasing their reliability and adding certain safety features such as escape systems for the crew.

The Space Shuttle configuration is a sharp departure from these previous vehicles and has presented unique systems engineering challenges. In the past, the analyses, test, and integration of the several elements of the rocket space vehicles were greatly simplified by the symmetry of the vehicle and the simplicity of the interfaces. After several years of study of reusable space vehicles, the United States began in 1972 the development of the Space Shuttle as we know it today. The configuration of the Space Shuttle with its winged Orbiter, the large propellant tank, and the two Solid Rocket Boosters (SRBs) presented a complex vehicle for design and development. Many problems involving propulsion systems, flight control, aerodynamics, and structural dynamics have been complicated. This paper describes several of these systems engineering challenges which were solved for

the Space Shuttle.

## SYSTEM DESCRIPTION AND OPERATION

The striking difference in configuration between the Space Shuttle and a more conventional multistage rocket vehicle is illustrated in Fig. 1. The Saturn V vehicle used to carry men to and from the moon is compared to the Space Shuttle.



**Fig. 1.** Unusual configuration of the Space Shuttle compared to a more conventional rocket vehicle

The present Space Shuttle configuration evolved from extensive studies trying to meet the requirement of reusability, large payload capacity to and from orbit, and a design flexible enough to perform a variety of missions. The winged Orbiter had to satisfy the conflicting demands of acceptable loads during ascent as well as suitable flying qualities during descent and landing. The SRBs provide the majority of the thrust during the early part of the flight, the liquid hydrogen/liquid oxygen main engines providing the remainder. Following staging of the boosters at about Mach 4.5, the vehicle continues on to orbit with the main engines and the propellant tank. As shown in Fig. 2, the SRBs are parachuted into the ocean and recovered for reuse. The propellant tank, called the External Tank (ET), is separated just short of orbital velocity and reenters the atmosphere and breaks up over a remote ocean area. Following its stay on orbit, the Orbiter reenters and performs a long hypersonic glide to its landing site for reflight on the next mission. (Additional details of the design, operation and first flight are contained in the paper entitled "The Space Shuttle: Orbital Flight Phase Begins," also presented at this Congress.)

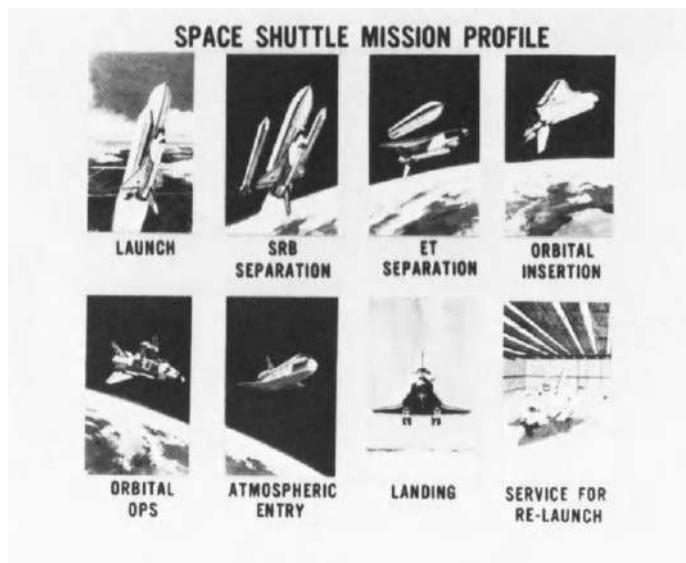


Fig. 2. Mission profile of the Space Shuttle

### ASCENT FLIGHT CONTROL

#### Design Requirements

Ascent flight control for the Space Shuttle is similar in many basic respects to ascent flight control for Mercury, Gemini and Apollo. There are, however, some very unique differences. These differences derive primarily from: (1) the very asymmetric configuration of the Space Shuttle; and (2) the Space Shuttle requirement that the capability for intact abort (runway landing) must exist at all times from liftoff through orbital insertion. These design drivers strongly influence all phases of ascent flight control from liftoff through orbital insertion.

#### Nominal Ascent

The ascent phase for a nominal Space Shuttle flight is performed in an automatic mode. However, if it is determined that the automatic system is not guiding the vehicle on the proper trajectory, it is possible for the crew to use manual guidance and throttling. The manual mode can also be used in abort cases.

The same flight control system is used for both automatic and manual modes. Fig. 3 is a simplified schematic of the ascent guidance, navigation and control system. All elements of the system are two-failure tolerant.

There are four redundant flight computers in the primary system. A fifth computer, with software independently coded, provides a backup flight control system in the unlikely event of a generic software problem in the four primary computers.

**ASCENT GUIDANCE, NAVIGATION,  
AND CONTROL SUBSYSTEM  
FUNCTIONAL BLOCK DIAGRAM**

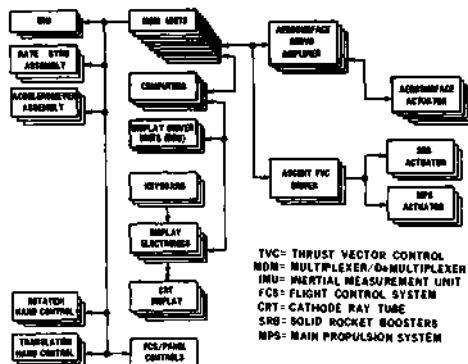


Fig. 3. Simplified block diagram of the ascent guidance, navigation and control subsystem

The ascent flight control sensors include triply-redundant Inertial Measurement Units (IMUs); quadruply-redundant accelerometer assemblies; and quadruply-redundant rate gyros.

Attitude control is maintained by thrust vector control (TVC) of the SRBs and the main engines during first stage, and of the main engines during second stage. The only aerosurface control during ascent is elevon control for load relief during first stage.

**First Stage Flight.** The launch sequence begins with the start of the Space Shuttle Main Engines (SSMEs). When all three engines have achieved a 90 percent of Rated Power Level (RPL) plus 2.75 seconds, the SRBs are ignited. Liftoff occurs when the thrust to weight ratio reaches 1.0 (0.28 seconds after SRB ignition). The vehicle rises vertically until it clears the tower at about eight seconds after liftoff. Vehicle attitude is held constant during this time. Hydraulic actuators on both the main engines and the solid rocket engines drive the nozzles in response to guidance commands steering the vehicle along its flight path. At T+8 seconds, when relative velocity reaches about 107 feet (32.6 m) per second, first stage steering begins with a pre-determined roll, pitch and yaw maneuver which puts the crew heads down and aligns the vehicle with the launch azimuth. As dynamic pressure rises, the loads on the vehicle rise. To provide vehicle load relief, the gains in the flight control system are adjusted to emphasize load relief, which of necessity de-emphasizes the vehicle's ability to follow the reference trajectory. These gain changes are based on sensed normal accelerations and result in the vehicle turning into the wind. Load relief starts at about T+25 seconds at a relative velocity of about 550 feet (167.6 m) per second, and ends at about T+91 seconds at a relative velocity of about 2,600 feet

(792.5 m) per second. In addition to vehicle load relief, it is also necessary to provide load relief for the Orbiter's four elevons (see Fig. 4). To accomplish this, the elevons follow a pre-determined schedule of position based on relative velocity. The elevon position schedule starts at liftoff and ends at about T+91 seconds at a relative velocity of about 2,600 feet (792.5 m) per second. The elevon position schedule is biased by the flight software if the pressure differential on any of the elevon primary actuators exceeds a pre-determined value. Once the excess pressure differential is reduced, the bias is removed. There is no way for the crew to monitor or control either vehicle load relief or elevon load relief.

At about T+48 seconds, the SSMEs are throttled down to 65 percent of RPL to limit maximum dynamic pressure. The SSMEs are throttled back up to 100 percent RPL at about T+64 seconds. Maximum dynamic pressure occurs at about T+70 seconds.

#### WING/ELEVON LOAD LIMITING

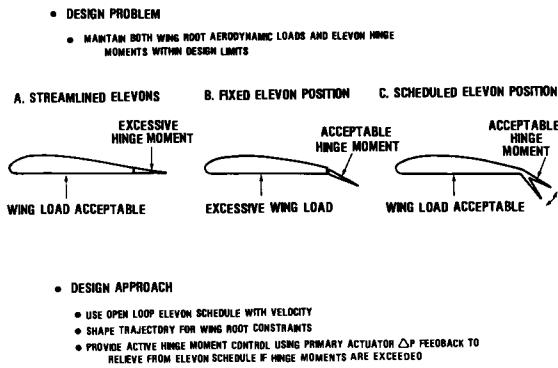


Fig. 4. Load limiting scheme for the wing and elevons

Second Stage Flight. SRB staging occurs at about T+133 seconds. Second stage guidance is engaged four seconds after SRB separation, and continues until the main engine cut-off (MECO) command is sent. Unlike first stage guidance, where a pre-determined attitude profile is followed, second stage guidance is closed loop. The difference between the targeted state vector (MECO target) and the current state vector (computed by navigation software) is cyclically computed, and the optimum vehicle attitude to reach the MECO target is calculated and maintained.

Second stage guidance computations are made assuming three hypothetical thrust phases occur during second stage (see Fig. 5). In the first phase (1), guidance computations are made assuming that an engine will fail at time = T FAIL. T FAIL is defined as the earliest time at which an abort-once-around (AOA) is possible (about four minutes after liftoff). Therefore, during the time prior to T FAIL,

the guidance calculations are made to optimize the second stage trajectory, assuming an engine failure at T FAIL, with two engines remaining for the rest of second stage. The second phase (2) is defined from T FAIL until three "g" acceleration is reached. (A design requirement limits the maximum acceleration during powered flight to 3 g's.) The guidance computations assume constant thrust throughout this phase. Prior to T FAIL, guidance assumes two engines at constant thrust; after T FAIL, and an engine failure has not occurred, guidance assumes three engines at constant thrust.

### HYPOTHETICAL SECOND STAGE THRUST PHASES

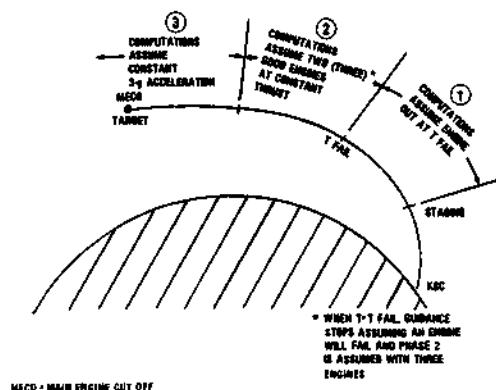


Fig. 5. Second stage guidance phases

Since guidance assumes that an engine failure will occur at T FAIL, a lofted trajectory is commanded prior to T FAIL in order to optimize performance for the reduced thrust expected at T FAIL. When mission elapsed time exceeds T FAIL, with no engine failure having occurred and guidance begins assuming three good SSMEs, the lofted trajectory is no longer optimum. The predicted trajectory with three SSMEs will overshoot the desired MECO target (see Fig. 6). A correction for this is made by commanding a large pitch attitude change ( $\sim 25$  degrees, nose toward earth).

The third phase (3) is defined from the beginning of three "g" acceleration (about 460 seconds after liftoff) to MECO. Guidance assumes three "g" constant acceleration through this phase. The SSMEs are commanded to begin throttling down to 65 percent of RPL at about ten seconds prior to MECO, and this level of thrust is maintained until the targeted velocity is achieved at about 520 seconds after liftoff.

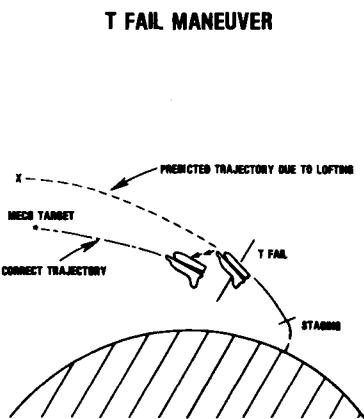


Fig. 6. Corrections to the second stage trajectory to attain the Main Engine Cut-off (MECO) target

#### Abort Modes

Return to Launch Site (RTLS). The three abort modes available during ascent are illustrated in Fig. 7. RTLS abort is the intact abort mode that provides a sub-orbital return capability of the Orbiter, crew and payloads directly to the Shuttle landing field adjacent to the launch pad. This abort mode returns the Orbiter within approximately 22 minutes after liftoff for aborts initiated during the first 260 seconds of ascent. RTLS was primarily designed to protect against failure of an SSME, but its inherent capabilities also provide a three-SSME return capability for time-critical failures requiring an immediate return early in ascent.

Abort-Once-Around (AOA). AOA abort is the intact abort mode used if an SSME failure occurs immediately after RTLS capability is lost at ~260 seconds after lift-off. The capability for three-SSME AOA also exists, and could be used for time-critical failures requiring early mission termination. For the four Orbital Flight Test (OFT) missions, AOA landings would be made at Northrup Strip on the White Sands Missile Range in New Mexico, or at Edwards Air Force Base, California. For operational flights, AOA landings will be made at the landing field adjacent to the launch site, e.g., at the Kennedy Space Center for launches from there.

Abort-To-Orbit (ATO). ATO abort provides the capability to continue to orbit if an SSME failure occurs, there are no time-critical reasons to terminate the mission early, and the capability exists to reach a safe orbit. ATO is the preferred abort mode. The Mission Control Center (MCC) is prime for identifying the abort mode to be selected. The abort mode identification is based on vehicle performance

assessments made in the MCC computers during ascent. Using actual vehicle performance as determined by ground tracking data and vehicle telemetry, the flight controllers make the abort mode call to the flight crew. The flight crew then selects the abort mode to be flown by positioning a cockpit switch.

### ASCENT FLIGHT PHASES

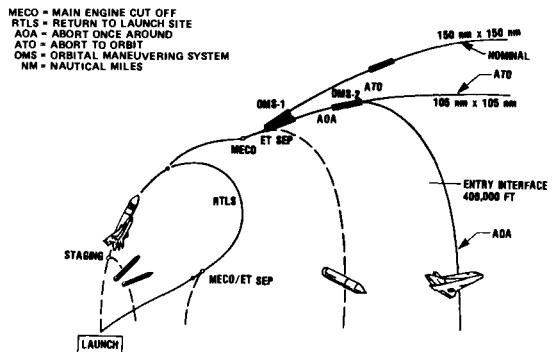


Fig. 7. Intact abort is achievable throughout the ascent flight

#### Test Program

Structural Dynamics. Several major tests were required in the development of Space Shuttle ascent flight control. To verify the math models of the structural dynamic characteristics of the vehicle, a major ground test program was conducted. This program used two test articles, shown in Fig. 8: (1) a high fidelity 1/4-scale model of the total vehicle; and (2) a full-scale test article. The need for the scale model was recognized in the beginning of the Space Shuttle Program since the complexity of the analytical techniques for the multi-body Shuttle configuration would require considerable development and verification before testing of the full-scale article could take place.

The 1/4-scale model was designed to structurally replicate the flight hardware and have a similar dynamic response. The model consisted of an Orbiter, an External Tank (ET), and Solid Rocket Boosters (SRBs). Three sets of scale model SRB cases were provided to represent different propellant load conditions in the flight regime (liftoff, maximum dynamic pressure, and burnout). More test conditions were included in the 1/4-scale program than could be permitted by cost and schedule constraints in full-scale testing. In addition, the data from the scale model were obtained more than a year in advance. This allowed for many refinements in the math models and led to more orderly testing and analysis in the full-scale program.

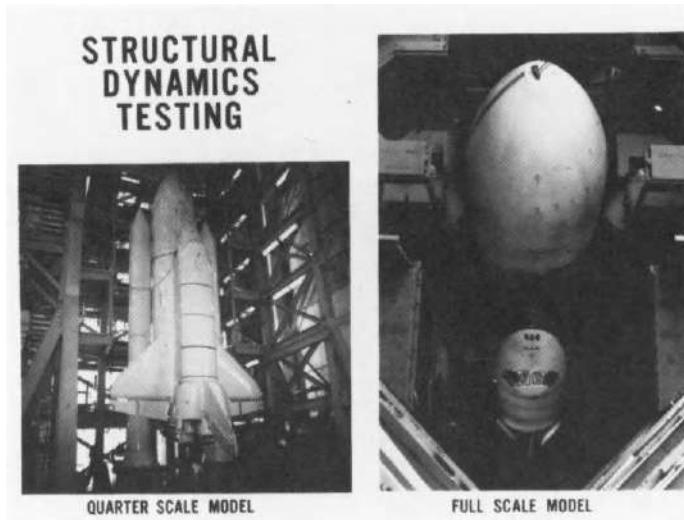


Fig. 8. Ground test articles were used to obtain structural dynamics data

The full-scale test article consisted of Orbiter 101 (previously used for manned approach and landing tests), a flight-weight ET, and flight-weight SRBs (one set of fired cases and one set of inert-loaded cases). The test configuration included Orbiter/ET/SRB in a "liftoff" and SRB "burnout" condition, and the Orbiter/ET configuration under three different propellant load conditions simulating different times in second stage flight. These tests provided structural modal and frequency response data for correlation with scale model testing, and verified the math models used in several system engineering areas: ascent flight control; POGO (interaction of propellants, main engines and structure); and structural loads.

**Avionics.** Ascent flight control avionics and software were tested in the Shuttle Avionics Integration Laboratory (SAIL) at the Johnson Space Center (JSC) in Houston, Texas, shown in Fig. 9.

The SAIL provides full fidelity operation of the flight software and the Orbiter's crew stations and the forward and aft avionics bays. Flight qualifiable avionics boxes are used throughout the facility. Non-avionic subsystems (e.g., propulsion elements) required to interface with flight software, caution and warning, or displays and controls are simulated. The avionics hardware is mounted in the SAIL using flight-type equipment bays and shelves. Flight wire harnesses are used.

#### Ascent Flight Control Summary

Although the Space Shuttle's unique design requirements (e.g., vehicle asymmetry, continuous intact abort capability, two-failure tolerance) posed a number of systems engineering challenges, the flawless performance of the system during STS-1 has demonstrated that a practical, efficient, and flexible ascent flight control

design has been developed.



Fig. 9. Ascent flight control avionics and software were tested in the Shuttle Avionics Integration Laboratory

The Space Shuttle's all-digital design provides the capability for easily "fine-tuning" the design as flight results are evaluated. Any necessary changes in the ascent flight control system parameters (e.g., first stage attitude profile, control system gains) can be readily incorporated in the flight software and quickly verified in the SAIL.

The importance of the test programs which provided input data to the design of the Space Shuttle's ascent flight control system cannot be over-stated. The numerous wind tunnel tests, the 1/4-scale and full-scale structural dynamics tests, and the SAIL avionics and software tests made it possible to fly an extraordinarily complex space vehicle, manned on the first flight, with extremely high confidence that the ascent flight control system would work...which it did, superbly.

#### DEVELOPMENT OF THE MAIN PROPULSION SYSTEM

The Main Propulsion System (MPS) consists of three Space Shuttle Main Engines (SSME), the Propellant Management System (PMS), and the External Tank (ET).

##### Components of the Propulsion System

The SSMEs are mounted in a triangular pattern in the Orbiter aft fuselage. They are reusable, high performance, liquid propellant (oxygen/hydrogen) rocket engines with variable thrust. The SSMEs are ignited on the ground, operate in parallel with the SRBs during the initial two minute ascent phase, and continue to operate

independently for a total of approximately eight minutes (operating time varies with throttle setting). The characteristics and major components of the engines are shown in Fig. 10. The engines are throttleable over a thrust range of 65 to 109 percent of the Rated Power Level (RPL). The throttling capability allows the thrust to be tailored to meet mission needs, to limit maximum dynamic pressure, and to limit the Orbiter's acceleration to 3 g's. The 3-g limit provides an environment which permits participation of personnel not capable of meeting the rigorous standards of previous astronauts and also minimizes structural loads. The engines are gimbaled to provide pitch, yaw, and roll control during ascent.

### SPACE SHUTTLE MAIN ENGINE (SSME)



• DIMENSIONS		
• LENGTH	167 IN.	(424cm)
• NOZZLE EXIT DIA.	94 IN.	(239cm)
• WEIGHT (DRY)	6886 LB	(3126kg)
• THRUST		
• VACUUM	470,000 LB	(2,090,560N)
• CHAMBER PRESSURE	2,970 psia	(2,048N/cm <sup>2</sup> )
• SPECIFIC IMPULSE		
• VACUUM	455.2 Sec	(4,464NSec/kg)
• LIFE	7.5 Hours 55 Starts	7.5 Hours 55 Starts

Fig. 10. Characteristics of the Space Shuttle Main Engine (SSME)

The design useful life for the SSMEs is seven and one-half hours and 55 starts. The ability to be overhauled and recycled is also a requirement.

These main engine requirements are met through the use of a staged combustion, closed power cycle coupled with high combustion pressures (3,000 psi). This results in an overall system efficiency substantially higher than conventional engine systems and an extremely compact engine envelope which minimizes the aerodynamic impact on Orbiter boattail design.

The operation of the SSME is managed by dual redundant digital computers. This electronics package is called a controller and one is mounted on each engine. It automatically performs checkout, start, inflight operations, and engine shutdown. The controller monitors engine operation fifty times each second, adjusts propellant flows and mixture ratios and, in the case of failure, automatically corrects the problem or shuts down the engine safely.

The External Tank (ET) supplies the Orbiter MPS with liquid hydrogen ( $LH_2$ ) and liquid oxygen ( $LO_2$ ) at prescribed pressures, temperatures, and flow rates. Because

the ET is an expendable element, it contains a minimum number of active components and the subsystems are designed for single usage to minimize costs. At main engine cutoff (MECO), the ET is separated from the Orbiter before orbital velocity is obtained. The ET proceeds on a ballistic reentry path and breaks up prior to impact in a remote ocean area.

During SSME burn, propellants under tank pressure flow from the ET to the Orbiter through two umbilicals: one for LH<sub>2</sub> and one for LO<sub>2</sub>. Within the Orbiter, the propellants pass through a system of manifolds, distribution lines and valves to the main engine. This system is the Propellant Management System (PMS). Figure 11 shows the elements of the MPS.

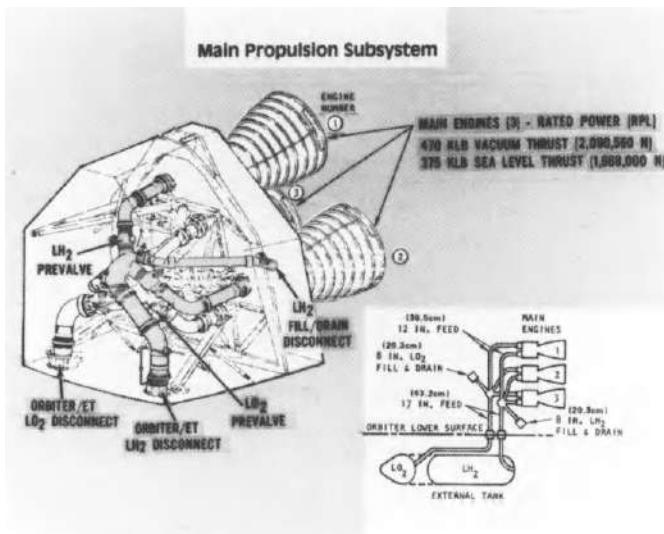


Fig. 11. The Main Propulsion System with its three rocket engines in the aft end of the Orbiter

#### Main Propulsion Test (MPT) Program

In order to qualify the MPS for flight, system tests combining all of the elements previously described were required. These were in addition to the normal qualification tests of the individual elements. The Main Propulsion Test (MPT) Program was a series of cryogenic tankings and 3-engine static firings designed to integrate and evaluate performance of the MPS system, and demonstrate the compatibility of all interfacing elements and subsystems. The program included tests investigating off-nominal conditions and verification of system design changes.

The Main Propulsion Test Article (MPTA) used for MPT consisted of the ET, the Orbiter aft fuselage containing the three SSMEs, and a structural truss arrangement simulating the Orbiter mid-fuselage. This test article is shown in Fig. 12 along with a picture of one of the static firings.

## MAIN PROPULSION TEST (MPT) PROGRAM

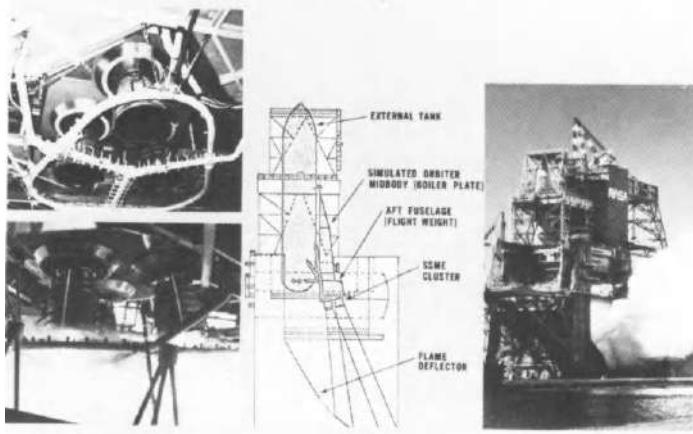


Fig. 12. Three-engine static firings were accomplished on a special test facility to certify the Main Propulsion System for flight

The original MPT test objectives were defined and scheduled for completion in a series of 12 test firings beginning with very short duration (2.35 seconds) firings and low throttle settings (70 percent of RPL), and ending with six full mission duration firings at 100 percent and 102 percent of RPL throttle settings. Completion of all test objectives ultimately required 18 test firings. Seven firings were terminated early due to engine problems (six tests) or facility problems (one test) and only partially met planned objectives. Eleven firings achieved full, planned test duration and met all major objectives. The first test was a 1.0-second firing (terminated early due to a facility wiring problem) in April 1978. The final test, in January 1981, was a 628-second simulated Abort-Once-Around (AOA) mission with one engine out after 239 seconds.

In view of the fact that the MPT Program was conducted in parallel with single-engine development testing, it is not surprising that a number of engine problems occurred during MPT testing. It is somewhat surprising, however, that not one significant systems problem occurred during the entire MPT Program. The MPT Program provided conclusive proof that the MPS design is sound.

The final verification of the flight readiness of the MPS and the related KSC launch facilities was conducted during a 20-second firing of the SSMEs on the launch pad. The vehicle was restrained to the launch deck and SRB ignition was inhibited to prevent liftoff.

### Systems Solutions to Element Problems

While major systems problems were not evidenced during the MPT development program, system solutions were often utilized in the solving of element problems. An

example of this was the solution to an SSME high pressure oxidizer turbopump (HPOTP) overspeed problem.

In mid-1980, an analytical model of the main engine predicted catastrophic overspeed of the HPOP during a main engine oxidizer depletion shutdown. The overspeed results from the net positive suction pressure (NPSP) of the oxidizer dropping to zero as the acceleration head drops to zero. In the case of an oxidizer depletion shutdown, the NPSP decline starts from as low as 6 psi. During the normal shutdown sequence, the oxidizer flow to the oxidizer preburner (OPB) is shut off but the hydrogen flow continues for several seconds. The turbine continues to be driven by the warm hydrogen; and due to the lack of resistance provided by no oxidizer head, the pump overspeeds. The overspeed condition had been recognized previously but prior predictions were within the tolerance range. The condition had not been experienced during ground tests due to the gravitational head.

Several solutions were proposed and investigated. An increase to the ET LO<sub>2</sub> ullage pressure to raise the shutdown NPSP could not be accomplished within the ET structural limitations. A brake on the HPOP, due to design complexity and space limitations, was considered impractical. An OPB fuel valve was considered a viable solution but the development time would not support the early flight schedules.

The solution selected was a combination of changes to the shutdown sequence: (1) the oxidizer pre-valve (OPV), located in the Orbiter upstream of the engine inlet, was commanded closed 1.24 seconds after the engine shutdown command. The shutdown command closes the Main Oxidizer Valve (MOV); (2) helium pressure is applied through the POGO accumulator, pressurizing the LO<sub>2</sub> trapped between the OPV and the MOV; and (3) the OPB helium purge is adjusted to minimize its effect on turbine rotation. The intent was to have the pressurized column of oxidizer provide the resistance needed to prevent the HPOP overspeed.

#### MAIN ENGINE SHUTDOWN FLOW LOOPS

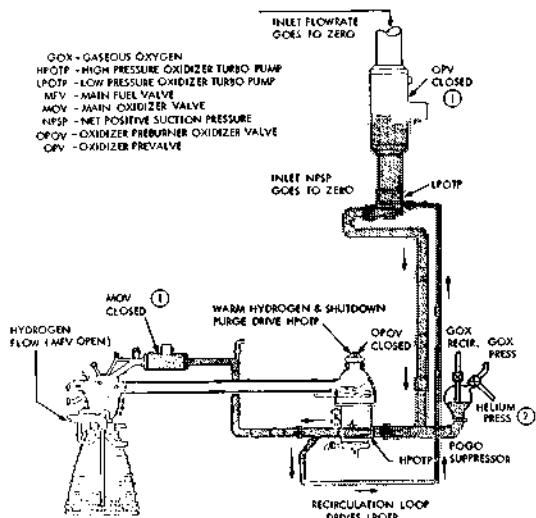


Fig. 13. Pressurization of the oxidizer line at main engine shutdown prevents overspeed of the high pressure oxygen turbopump

Figure 13 is a simplified diagram of those parts of the Main Propulsion System involved in accomplishing this shutdown sequence.

To implement the sequence changes, special testing was necessary. Components were tested to verify tolerances. The selected sequence was verified using the engine analytical model. A special setup was installed on a single-engine test stand to permit engine shutdowns at very low NPSP. The selected procedure was then verified using the MPTA. Finally, the sequence was installed in the Orbiter software and flown successfully on STS-1.

#### SELECTED AERODYNAMIC AND AEROTHERMODYNAMIC CONSIDERATIONS

##### Mated Configuration - Aerodynamic Design and Testing

The Shuttle configuration, as previously noted, is composed of the Orbiter, External Tank (ET), two Solid Rocket Boosters (SRBs), Orbiter/ET Attach Structure, ET/SRB Attach Structure, and miscellaneous protuberances. In order to design a vehicle of this complexity, aerodynamic forces and subsequent moments must be obtained for each item, both as an integrated configuration and for each element. Consequently, an extensive wind tunnel program was required.

By first orbital flight in April 1981, approximately 40,000 total wind tunnel test hours had been conducted for aerodynamics, heat transfer, and structural dynamics. A total of 94 models have been built. Figure 14 shows one such model installed in a wind tunnel located at the Arnold Engineering Development Center.

#### SPACE SHUTTLE WIND TUNNEL MODEL

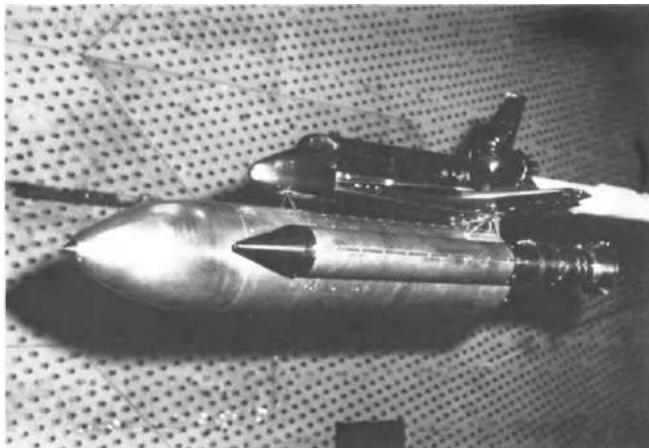


Fig. 14. Extensive wind tunnel testing of the Shuttle configuration was required

Several types of tests were accomplished during the program covering such areas as ground wind loads, first and second stage static aerodynamics without power-on effects, and power-on (simulating SSME and SRB burns) tests for first and second stage. As the test program progressed, the models, data requirements, and complexity increased immensely. Initially, most of the data for the total vehicle was based on very small models (0.4 percent) and typically no data were available for the elements and components. As the program proceeded, pressure models of larger scales (3.0 percent) provided element/component loads data with 1.0 percent scale models used for integrated vehicle dynamics.

As the design of the vehicle evolved to the current configuration, the requirements increased for appropriate model fidelity, more instrumentation, and a higher degree of accuracy. Data from these wind tunnel tests were important inputs to ascent trajectory analyses, ascent loads, subsystem design and analysis, and flight control studies.

#### Thermal Insulation System for External Tank (ET)

The thermal insulation for the ET had to satisfy requirements derived from pre-launch considerations as well as ascent flight. The tank's outside skin is insulated with spray-on polyurethane foam that reduces heat transfer into the tanks that could cause excessive boiling of the cryogenic propellants as well as a reduction in the material strength of the tank primary structure during ascent. It also helps prevent ice buildup during launch preparations that could shake loose inflight and damage the Orbiter. For high heat areas, an ablating material is used under the spray-on foam to protect the tank during ascent through the atmosphere.



Fig. 15. Oil flow technique is used on wind tunnel model to determine areas of high heating

Localized high heating areas encountered during ascent defined the location and thickness of the ablator. During the early program definition phase, the ET design used only the foam insulation. As the configuration evolved and as the results of additional wind tunnel tests became available, excessive interference heating on local areas dictated the utilization of the composite ablator material. Figure 15 shows one of the models used for aerothermodynamics wind tunnel tests. Using the oil flow technique, areas of shock impingement and high interference heating can be determined. Maintaining the primary structure and subsystem components within the design temperature limits and minimizing unusable propellant, as the result of thermal input, were the fundamental considerations. Heat input is obtained from aerodynamic convective flow, the engine plumes, the SRB separation motors and tank pressurization gas. Combining these requirements and the results of wind tunnel tests and analyses resulted in the pattern of insulation on the ET, shown in Fig. 16. Motion pictures were taken of the ET from the Orbiter on the first flight at tank separation. They showed the insulation to be in good condition. The pattern of blackened areas corresponded to the predicted high heat areas where ablation was expected to take place.

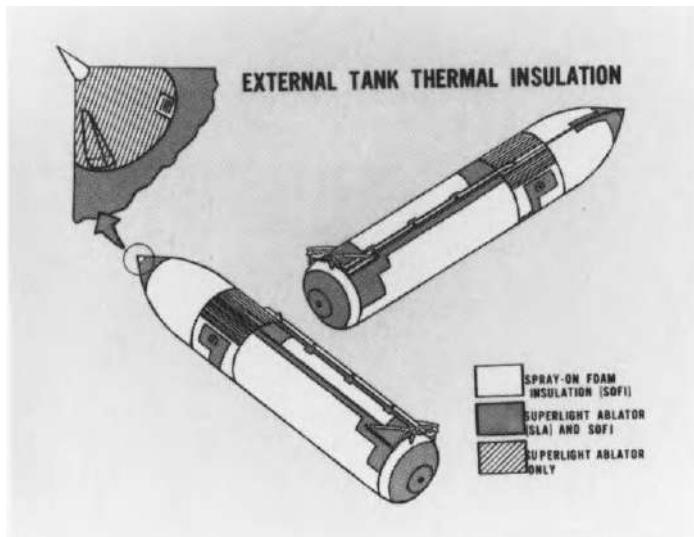


Fig. 16. Combinations of foam insulation and ablator were required to solve the insulation problems of the External Tank

#### Aerodynamic Design of the Orbiter

The Orbiter concept very early in the program was a delta-wing vehicle. Later refinements in the aerodynamic configuration led to a double-delta planform incorporating a more efficient lifting surface (Fig. 17).

Stability, control and performance requirements for aerodynamic configuration design of the Orbiter vehicle are, for the most part, established by the entry and landing phases of flight. Design issues key to achieving the proper aerodynamic balance across the entry/landing flight regime are wing design, wing-body

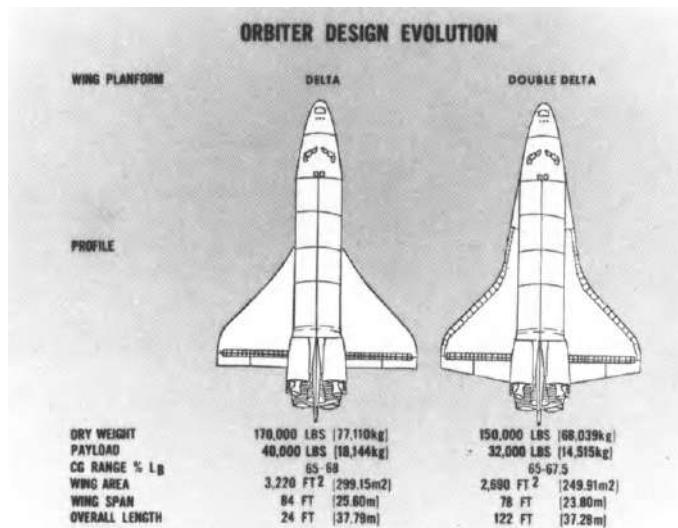


Fig. 17. Refinements in aerodynamic design led to the more efficient double-delta wing form for the Orbiter

integration, and integration of aerodynamic and flight control requirements. Wing design was key because of its influence on vehicle weight, thermal environment, aerodynamic stability, and gliding and landing capability. Wing-body integration was important in obtaining a balanced aerodynamic configuration capable of trim and control over the entire speed range, and in minimizing thermal environment due to interference flow effects. Fuselage dimensions were largely fixed by payload bay size and packaging efficiency while aerodynamic control and aerothermodynamic considerations established forebody shape and local contours. Integration of aerodynamic control requirements was of major importance in meeting flying quality goals and minimizing vehicle weight as affected by control surface arrangement, size, and actuator requirements.

#### Entry Flight Profile

Analyses of the entry phase resulted in a flight trajectory which is accomplished by flying an angle of attack/velocity profile preselected to meet thermal design criteria, and using roll commands for cross-range control. Flight control is achieved in two modes (i.e., spacecraft, aircraft). The spacecraft mode applies from initial through mid-entry phases where the Orbiter is at high angle of attack making the vertical fin and rudder ineffective. The aircraft mode includes mid-entry through approach and landing. Transition between the two modes begins at approximately Mach 5 and is completed at about Mach 1.5. In the spacecraft mode, control in all three axes is initially provided by the aft Reaction Control System jets mounted at the base of the Orbiter on either side of the vertical tail. As control authority of the aerodynamic surfaces become sufficient, the jets are deactivated. The remainder of the flight down to landing is flown in the airplane mode.

### Orbiter Thermal Protection System (TPS)

The Orbiter TPS is designed to be reusable after withstanding the high temperatures experienced during ascent flight and entry from orbit. The airframe is covered with insulation tough enough to last through 100 missions before replacement, and effective enough to protect the aluminum substructure for 500 missions.

The peak heating rates and the longest exposure to these rates occur during entry when surface temperatures may range from 3,000°F (1,649°C) at stagnation points on the nose down to about 600°F (316°C) on leeward surfaces. The TPS is composed of two types of Reusable Surface Insulation (RSI) tiles, a high temperature carbon structure coupled with internal insulation, thermal window panes, and a type of coated felt. The various types of thermal protection used on the Orbiter are shown in Fig. 18.

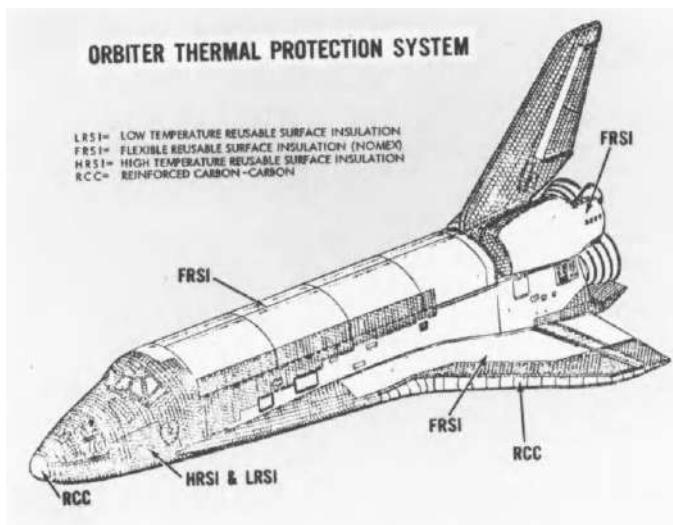


Fig. 18. Several types of insulation are used on the Orbiter to provide an effective thermal protection system

The RSI tiles covering the Orbiter are made of coated silica fiber. The two types of RSI tiles differ only in surface coating to provide protection for different temperature regimes. The Low-temperature Reusable Surface Insulation (LRSI) consists of 8-inch (20.3 cm) square silica tiles, and covers the top of the vehicle where temperatures are less than 1,200°F (649°C). The High-temperature Reusable Surface Insulation (HRSI) is typically 6-inch (15.2 cm) square silica tiles and covers the bottom and some leading edges of the Orbiter where temperatures are below 2,300°F (1,260°C). A high-temperature structure of Reinforced Carbon-Carbon (RCC) is used with internal insulation for the nose cap and wing leading edges where temperatures are greater than 2,300°F (1,260°C). Flexible Reusable Surface Insulation (FRSI), consisting of coated nomex felt, is used on the upper cargo bay door, lower aft fuselage sides, and upper aft wing where temperatures are less than 700°F (371°C).

Performance of the Orbiter TPS was excellent on the first flight requiring a minimum of refurbishment for the next flight. In general, the predictions of the aerodynamic and thermal characteristics of the Space Shuttle were close to the measured flight results, considering the unusual configuration and the speed range from Mach 25 to subsonic flight.

#### SUMMARY

Despite the unorthodox configuration of the Space Shuttle, its first flight results closely matched the predicted characteristics. The critical areas of flight control, aerodynamics, thermal protection and propulsion, as well as other areas, performed extremely well on the first flight of the Space Shuttle. Results of this first flight increase confidence that the Space Shuttle can be brought to an operational status on schedule.

#### ACKNOWLEDGEMENTS

In the preparation of this paper, the author received assistance from William Hamby, Thomas Stuart, William Goldsby, Howard Roseman and Kristie Weingarten, all of NASA.

## THE INERTIAL UPPER STAGE—A SPACE TRANSPORTATION SYSTEM ELEMENT NEARING FIRST FLIGHT

D. J. Rohrbaugh\*, F. J. Redd\*\* and F. van Rensselaer\*\*\*

\*Boeing Aerospace Company, USA

\*\*U.S. Air Force, Space Division, USA

\*\*\*National Aeronautics & Space Administration Headquarters, USA

### ABSTRACT

The Inertial Upper Stage is an upper stage being developed by the United States Air Force with Boeing Aerospace Company as prime contractor. The IUS offers the spacecraft community a highly reliable, cost-effective solid propellant upper stage, with inherent flexibility and adaptability, for integration with the space shuttle. The IUS can deliver spacecraft accurately into a wide range of earth orbits, including geosynchronous.

At the 1979 IAF Congress a review of the IUS program was presented. This paper is intended to update the IAF membership on developments since then as the IUS approaches first launch.

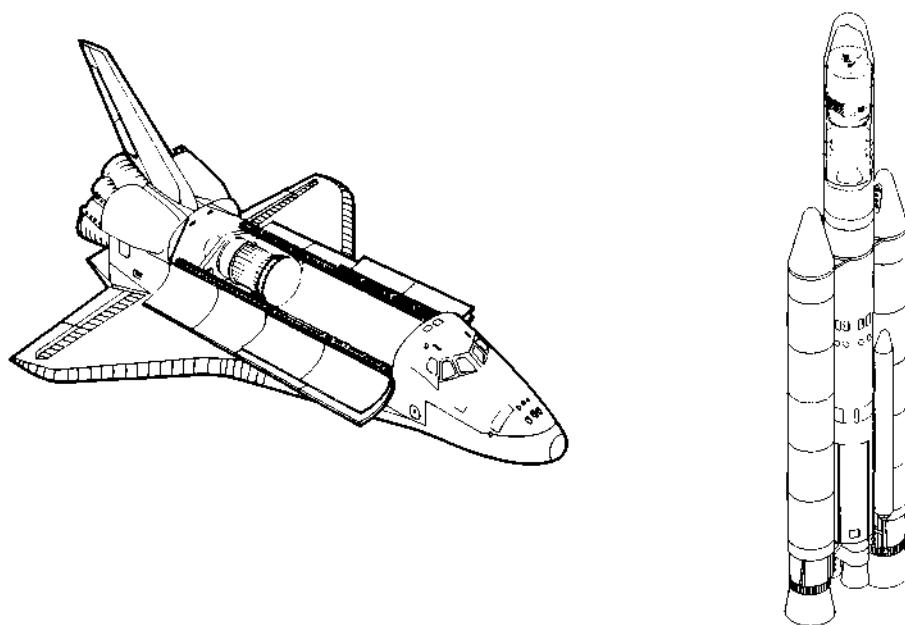
### KEYWORDS

Inertial upper stage, solid propellant rocket motors, avionics, reliable, high performance, software, gamma guidance, development status.

### INTRODUCTION

The IUS is for joint use by the United States Air Force and NASA as a key element of the United State's space transportation system. The IUS is designed to provide an upper stage less likely to fail, than current upper stages for expendable launch vehicles. Solid propellant was chosen to simplify interfaces and integration with the Shuttle Orbiter. The IUS interfaces with spacecraft are intended to ease the transition from current launch vehicles to the Shuttle system.

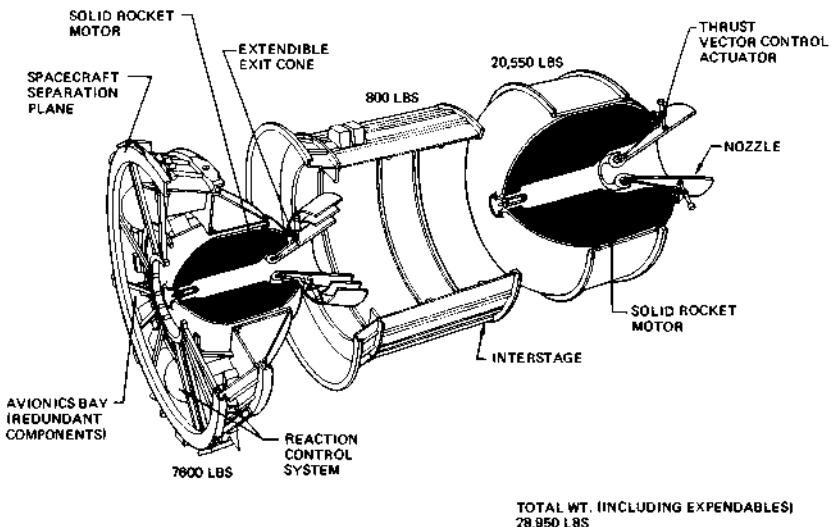
The Boeing IUS will deliver spacecraft of the U.S. Air Force and NASA from the Space Shuttle into a wide range of higher earth orbits. It also will serve as an upper stage for the Air Force's Titan 34D launch vehicle. The IUS was designed, and is being built, under contract to the Air Force Space Division. Boeing is working now toward a Titan/IUS launch date of mid-1982 and an early 1983 date for launch aboard the Shuttle. Fig. 1 shows the IUS configuration with both the Shuttle Orbiter and Titan 34D.



*Figure 1. Shuttle and Titan/IUS Vehicles*

#### IUS DESCRIPTION

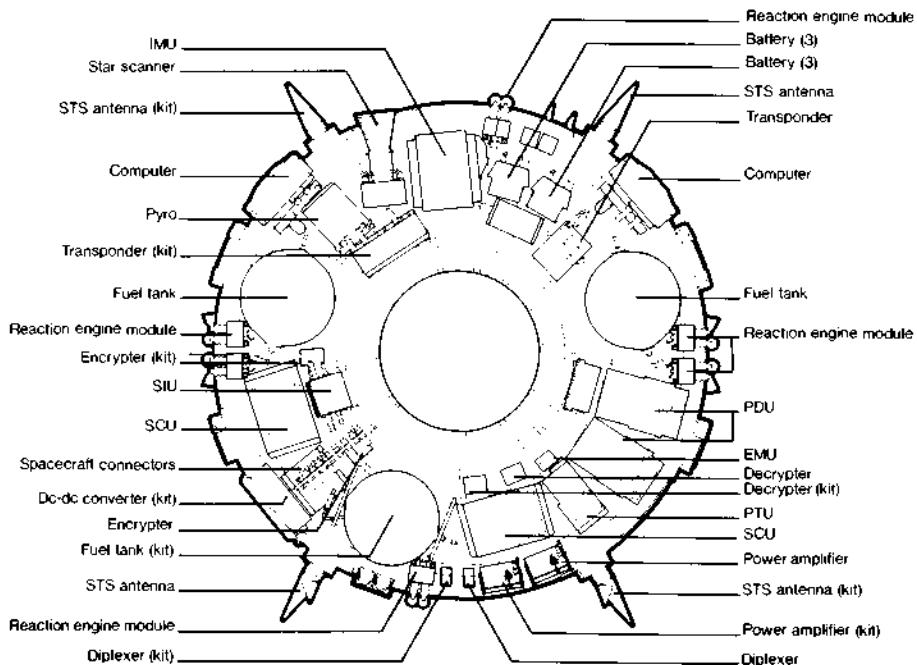
The IUS, which can deliver 2268 Kg (5000 lbs) from the Shuttle to geosynchronous altitude, is a two stage solid propellant vehicle featuring redundant systems that produce the most reliable unmanned launch system ever developed. The IUS, shown in Fig. 2, consists of a 9700 Kg (21,400 lb) propellant weight first stage, an interstage structure, a 2720 Kg (6000 lb) propellant weight second stage and an Equipment Support Section.



*Figure 2. Inertial Upper Stage*

### Equipment Support Section

The Equipment Support Section (ESS) houses most of the IUS avionics and control subsystems. The top of the ESS contains an interface mounting ring and electrical interface connector segment for mating and integration of the spacecraft. Thermal isolation is provided by a multilayer insulation blanket across the interface between the IUS and the spacecraft. All Line Replaceable Units (LRU's) mounted in the ESS can be removed and replaced by means of access doors even with a mated spacecraft. Arrangement and identification of the components are shown in Fig. 3.



*Figure 3. IUS Equipment Support Section*

### Avionics

The avionics system includes the electronic and electrical hardware used to perform all signal conditioning, data processing, and software formatting associated with navigation, guidance, control, data management and redundancy management. The avionics system provides the communication between the orbiter and ground stations. It also provides for electrical power distribution.

The IUS has the most reliable avionics system ever developed for unmanned space applications. The unparalleled reliability is obtained by both stringent parts requirements and complete redundancy of all mission critical command and control equipment as illustrated in Fig. 4.

Avionics consists of these subsystems: Telemetry, Tracking and Command; Guidance and Navigation; Data Management; Thrust Vector Control; and Electrical Power.

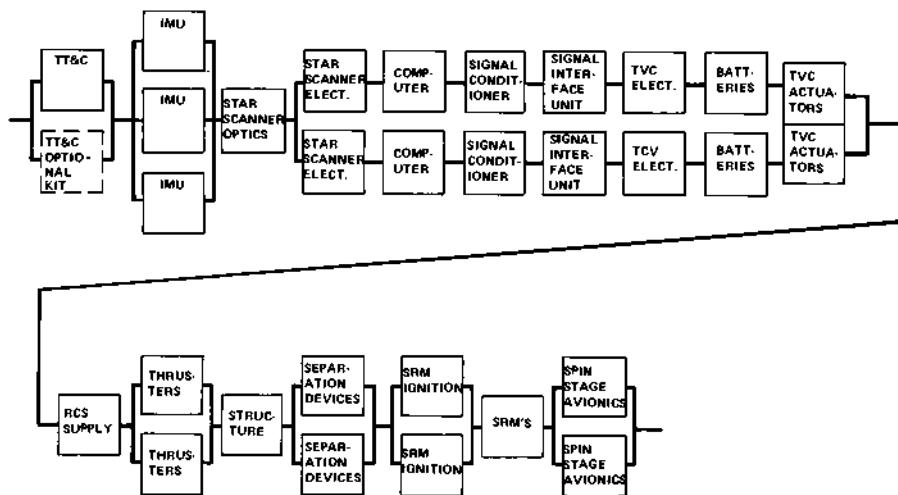


Figure 4. IUS Redundancy Schematic

**Telemetry, Tracking and Command (TT&C).** The Telemetry, Tracking and Command (TT&C) subsystem is designed to be totally redundant. The baseline configuration includes only one set of TT&C components. The redundant set is optional. For telemetry transmission, the IUS has two bit rates as shown in Table 1.

Table 1. IUS Telemetry Capability

	16 KBPS	64 KBPS
Word Length	8 bits	8 bits
Main Frame Length	80 words	160 words
Main Frame Rate	25/sec	50/sec
Sub Frame Length	25 main frames	50 main frames
Sub-Sub Frame Length	(8 sub-frames) (200 main frames)	(4 sub-frames) (200 main frames)
Master Frame Rate (Equiv. to 1 Complete Sub-Sub Frame)	1/8 (sec's)	1/4 (sec's)

**Guidance and Navigation.** The IUS uses a stellar inertial Guidance and Navigation Subsystem (GNS) to provide measurements at angular rates, linear accelerations, and other sensor data to the management subsystem for appropriate processing by software resident in the computers. The GNS consists of a redundant strapped-down Inertial Measurement Unit (IMU) and a star scanner.

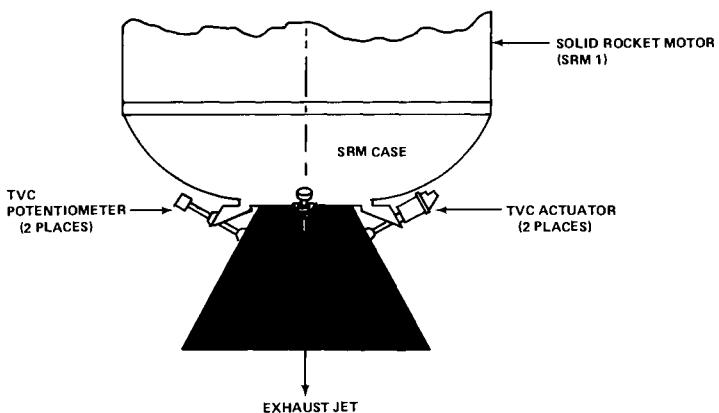
The GNS is calibrated and aligned on the launch pad. The navigation function is initialized at lift off and data from the IMU is integrated in the navigation software to determine the current state vector. Before vehicle deployment, an attitude update maneuver is performed by the orbiter, with the IUS star scanner scanning two stars separated by 60 to 120 degrees, to compensate for accumulated drift errors.

The IUS vehicle uses an explicit guidance algorithm (gamma guidance), to generate thrust steering commands, Solid Rocket Motor (SRM) ignition time and Reaction Control System (RCS) vernier thrust cutoff time. Prior to each SRM ignition and each RCS vernier, the vehicle is oriented to a thrust attitude based on nominal performance of the remaining propulsion stages. During SRM burn, the current state vector determined from the navigation function is compared to the desired state vector, and commanded attitude is adjusted to compensate for the build up of position and velocity errors due to off-nominal SRM performance (thrust, Isp). The primary purpose of the vernier thrust is to compensate for velocity errors resulting from SRM impulse and cutoff time dispersions. However, residual position errors remaining from the SRM burn and position errors introduced by impulse and cutoff time dispersions are also removed by the RCS.

**Data Management.** The Data Management Subsystem (DMS) performs the computation, data processing, and signal conditioning associated with guidance, navigation, and control; safe-arm and firing of the solid rocket motor and electro-explosive devices; command decoding and telemetry formatting; and redundancy management. The DMS also issues spacecraft discretes.

**Thrust Vector Control.** The Thrust Vector Control (TVC) subsystem provides the interface between the IUS guidance, navigation and control system, and the Solid Rocket Motor (SRM) gimbaled nozzle to accomplish powered flight attitude control. Two complete electrically redundant channels are provided to minimize single-point failures.

Power is supplied through the Signal Conditioner Unit (SCU) to the TVC controller which controls the actuators. The controller receives analog pitch and yaw commands, proportional to desired nozzle angles, and converts them to pulse-width, modulated voltages to power the actuator motors. The motor drives a ball screw which extends or retracts the actuator to position the SRM nozzle. Potentiometers provide for servo loop closure and for position instrumentation. A staging command allows switching of the controller outputs from actuators on one stage to actuators on the second stage. The TVC subsystem is illustrated in Fig. 5.



*Figure 5. Thrust Vector Control System*

**Electrical Power.** The power subsystem provides power for the IUS vehicle equipment and for the spacecraft. It can also utilize electrical power supplied via the Airborne Support Equipment, from Ground Support Equipment, or from the orbiter. Provisions are made for the installation of a DC to DC converter production option in the IUS vehicle, if a spacecraft requires regulated power.

This power subsystem provides for switching and distribution of electrical power to the IUS vehicle and spacecraft. Dual fuses ensure that no single-power system failure can disable both A and B avionics

channels. Four batteries are carried in the first stage; five batteries are provided to supply power for the second stage after staging.

Freedom from conducted interference is enhanced by supplying power from a separate utility battery source to Thrust Vector Control motors, Reaction Control System valves, motor driven power transfer switches, and all ordnance devices.

## Software

The on-board STS software is comprised of the Operational Flight Software and the Mission Data Load software.

**Operational Flight Software.** The Operational Flight Software (OFS) has prelaunch/predeployment functions, and controls the IUS in achieving the placement of an attached payload into a desired orbit following deployment from the orbiter. Software functions provide calculation and control capability for these mission operations functional areas: executive, mission sequencing, guidance, attitude control, communications, redundancy management, checkout, and navigation.

One of the unique features of IUS software is the control of IUS redundancy. This redundancy management functional area is responsible for determining the operational status of a number of IUS subsystems and for commanding reconfiguration within milliseconds when failures are detected.

**On-Board Digital Data Load.** The mission On-Board Digital Data Load (ODDL) are the data which are entered into the IUS flight computer memory. It consists of the operational flight program plus mission and vehicle-unique data (Mission Data Load). These data include redundant inertial measurement unit calibration values, guidance parameters, flight control systems gains, thermal maneuver references, mission sequencing event tables, etc. The Mission Data Load is a binary load module form, with all memory address locations defined and ready for extraction. The ODDL is transferred to the flight computer memory in one of two states. The initial load contains all mission and vehicle data plus the flight program. An update ODDL enters changes to the values of certain data parameters by defining data entries to specific memory location.

## Propulsion

The IUS incorporates two propulsion systems: The Solid Rocket Motors for main propulsion, and a redundant Reaction Control System for attitude control during the coast phase of the mission.

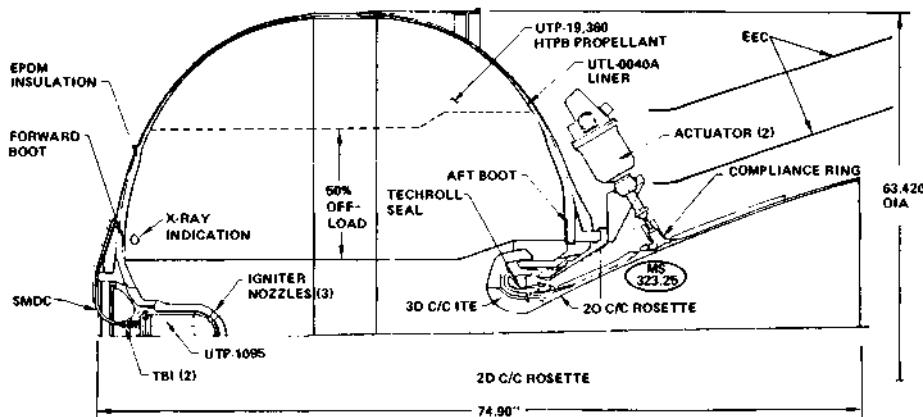
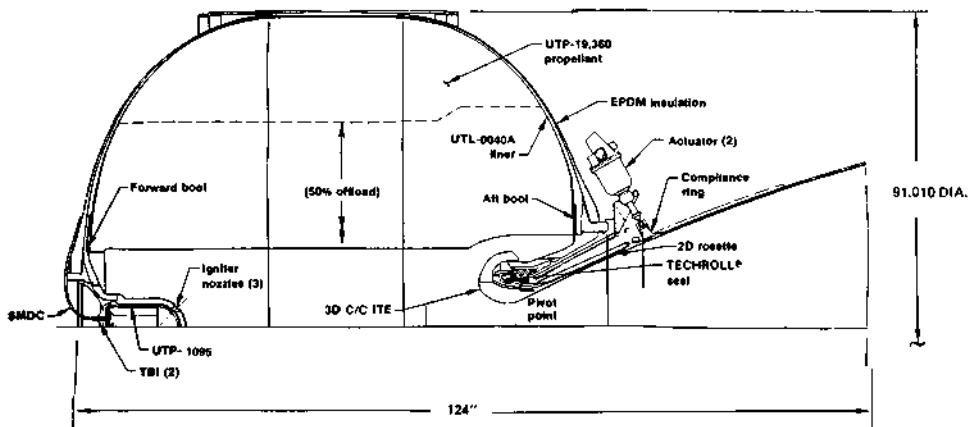


Figure 7. IUS Small Solid Rocket Motor



*Figure 6. IUS Large Solid Rocket Motor*

**Solid Rocket Motors.** The Solid Rocket Motors (SRM's) have been designed to use common approaches between the large and small motors, thereby improving reliability and reducing cost. The IUS two-stage vehicle uses a large solid motor, illustrated in Fig. 6, with 9,700 Kg (21,400 lb) of propellant and a small motor, shown in Fig. 7, with 2,720 Kg (6000 lb) of propellant.

These motors employ movable nozzles for Thrust Vector Control. The nozzles are driven by redundant electro mechanical actuators permitting up to  $4^{\circ}$  of steering on the large motor and  $7^{\circ}$  on the small motor. Kevlar filament wound cases, with integrally wound-in aluminum end rings, provide high strength at minimum weight. The large motor is the longest burning solid propellant space motor ever developed, having a burn time of 145 seconds.

**Reaction Control System.** The Reaction Control System (RCS) is a hydrazine monopropellant propulsion system that controls the IUS during its coast period, the roll during SRM burns, and the delta velocity impulses for accurate orbit injection. Valves and thrusters are redundant, permitting continual operation with a minimum of one failure.

To avoid spacecraft contamination, the IUS has no forward facing thrusters. IUS users with spin stabilized spacecraft can employ the RCS to provide sinup prior to spacecraft separation. The system is also used to provide spacing velocities between multiple spacecraft deployments and for a collision avoidance maneuver after last spacecraft separation.

### Structures and Mechanisms

The IUS structure is capable of transmitting all the loads generated internally and by the cantilevered spacecraft during orbiter operations and IUS free flight. In addition, the structure supports all the equipment and the SRM's within the IUS, and provides the mechanisms for stage separation. The major structural assemblies of the two-stage IUS are the equipment support section, interstage, and aft skirt as shown in Fig. 2 (see page 2). The basic structure is made from aluminum skin-stringer construction with eight longerons and ring frames. The material usage on the IUS is noted in Fig. 8.

### Airborne Support Equipment

The IUS Airborne Support Equipment (ASE) is the unique mechanical, avionics, and structural equipment located in the orbiter. The function of the ASE is to support and provide services to the IUS and its spacecraft while in the orbiter bay, and to provide for positioning of the IUS and spacecraft, in an elevated position, for final checkout prior to deployment from the orbiter. The ASE, shown in Fig. 9,

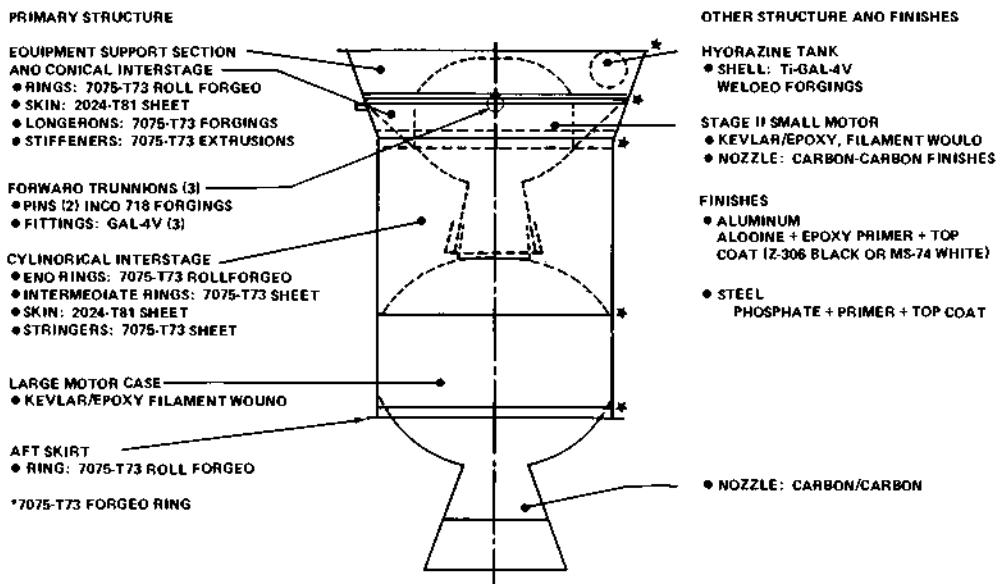


Figure 8. IUS Material Usage

consists of the structure, batteries, electronics, and cabling to support the IUS and spacecraft combination.

The ASE incorporates a low-response spreader beam and torsion bar mechanism that reduces spacecraft dynamic loads to about one-third of what would be experienced without this system. In addition, the forward ASE frame includes a hydraulic load leveler to provide a balanced loading at the forward trunnion fittings.

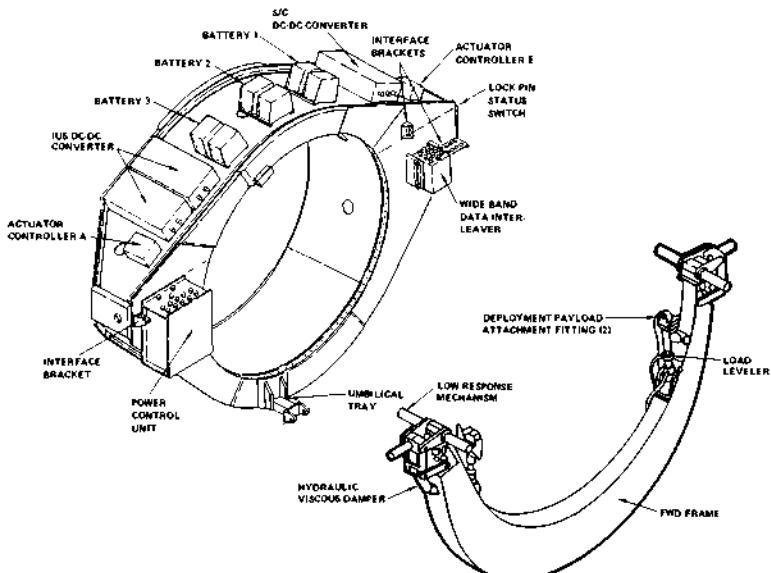


Figure 9. IUS Airborne Support Equipment

## IUS TWO STAGE DEVELOPMENT STATUS

The IUS two stage vehicle is now in an advanced stage of development as it is being readied for its initial capability launch date in mid-1982. Most elements of flight hardware have been fabricated, tested, and made ready for vehicle assembly.

## Testing Requirements

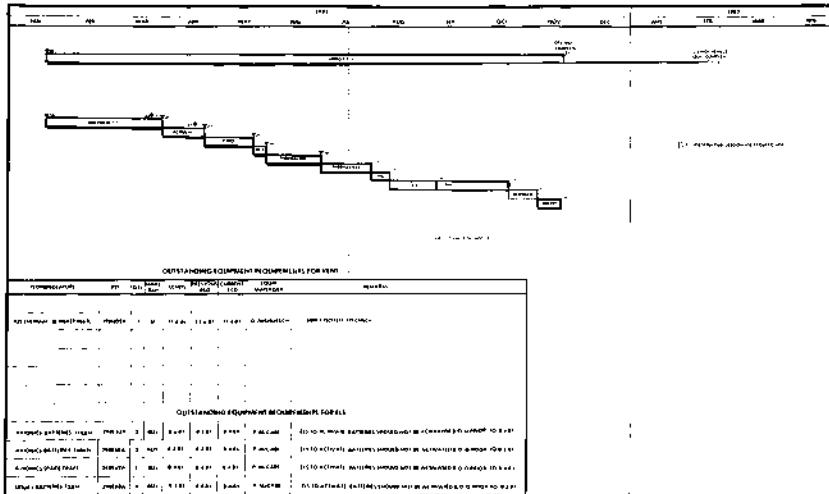
One of the most significant new specification requirements to be placed on the IUS contract was MIL-STD-1540A, which is concerned principally with testing. Although it had been anticipated that this requirement would cause a major impact to the program schedule, the impact greatly exceeded expectations. IUS was the first program to attempt full compliance.

Although MIL-STD-1540A contains numerous testing requirements, perhaps the two most significant to the IUS have been the qualification testing margins for environmental testing and the need to perform repetitive functional checks before, during, and following environmental exposures. The environmental portion of acceptance testing is at flight levels. Qualification and shock vibration testing requirements, for example, are at levels representing twice the energy that the component will experience in flight. Virtually every component tested has failed this test and has required redesign. In many cases, second and third failures have occurred.

The requirement to conduct functional tests before, during, and after an environmental experience resulted in the need to design and develop automated special test equipment. Manual testing would not have been possible. For acceptance testing, the functional test is repeated 55 times for a typical Line Replaceable Unit (LRU); for qualification testing, the functional test is repeated 44 times. If no failures are experienced, a typical LRU is functionally tested 99 times by the time it completes qualification testing. Although stringent and costly this testing significantly increases confidence in the IUS systems.

## **Qualification Test Vehicle**

The IUS vehicle began formal qualification testing last January and all tests are to be completed by November 20. The test schedule for this vehicle is illustrated in Fig. 10.



*Figure 10. Qualification Test Vehicle Schedule*

The avionics components used in the IUS vehicle, for the most part, have been exposed to these environments already. For example 33 of the 40 Line Replaceable Units have completed component qualification testing as shown in Fig. 11.

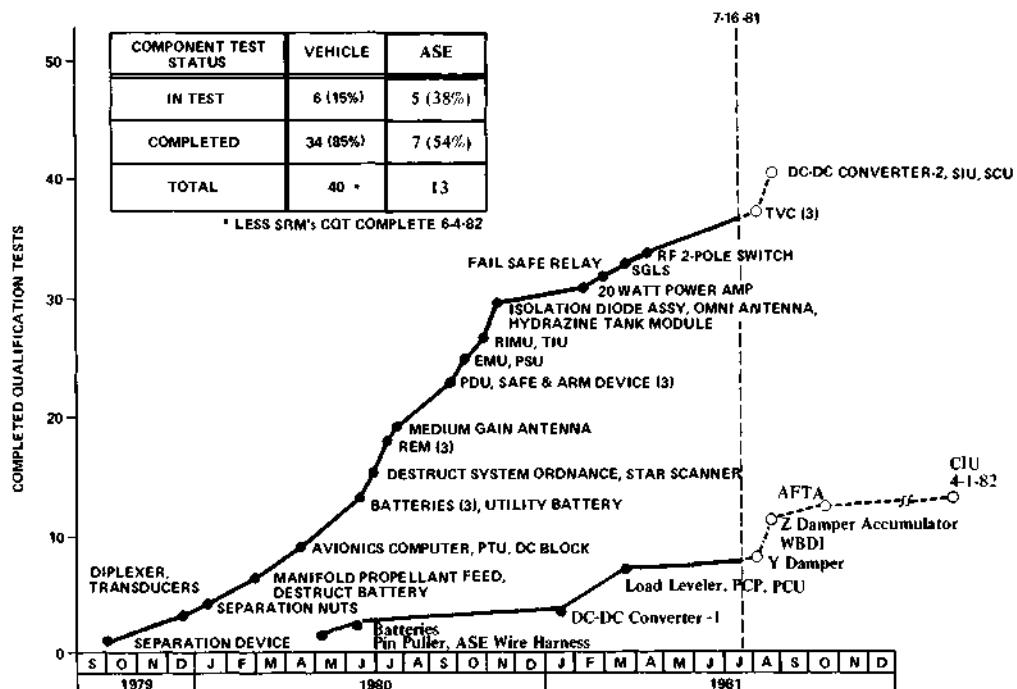


Figure 11. Component Qualification Test Schedule

### Solid Rocket Motors

Solid Rocket Motors of both the large and small size have been tested in simulated altitude firings at the Arnold Engineering and Development Center (AEDC) in Tennessee. The 12 Solid Rocket Motor developmental firings have been completed. Test conditions for these firings have included both high and low temperatures, vectored nozzles, use of Extendible Exit Cone on the small motor, and up to 50% offload on the motors. Qualification test motors will begin test firings in July at AEDC.

### Structures

The IUS vehicle structure for the Titan launch application has been proven by the successful completion of modal survey testing (17 modes mapped) and static loads testing (no yielding or deformation occurred). This structure is essentially the same as that to be used for Shuttle Two-Stage flights. The only difference is that the skin gages for the Titan interstage are lighter to reduce weight for the lower loads experienced in Titan flights. The Equipment Support Section structure is identical for both usages.

### Software

The Operational Flight Software (OFS) for Titan launches has completed coding and is being verified by our subcontractor, TRW, Boeing validation of this software is underway. The Mission Data Load for

the first Titan/IUS mission has been designed and is being merged with the OFS. Flight software and data loads for later missions including the first Shuttle mission are in the design cycle.

### Airborne Support Equipment

The Airborne Support Equipment (ASE) is not required for the first Titan flight and, therefore, lags the vehicle in the developmental/qualification cycle. At the present time, four of 13 components have completed qualification testing and seven more components are in test. (See Fig. 11) ASE interfaces with the orbiter have been worked out over the past four years with JSC and Rockwell and represent minimal orbiter impact. The structural ASE (cradle) has been assembled for the structural test vehicle program.

### GROUND AND LAUNCH OPERATIONS

Typical operations and flow concepts for ground and launch operations at Kennedy Space Center are shown in Fig. 12 and are briefly described herein.

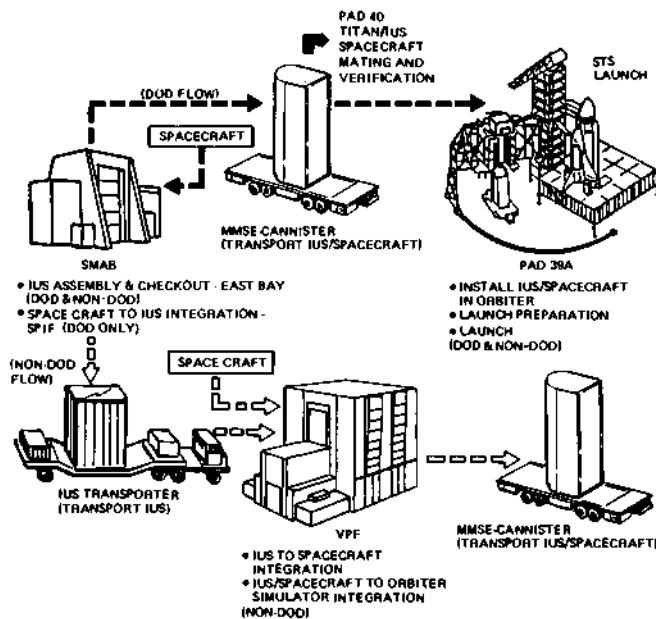


Figure 12. Ground Operations

In the DOD Shuttle flow, the IUS will be assembled and checked all in the east bay of the Solid Motor Assembly Building (SMAB). The IUS will be transported from the SMAB east bay to the Shuttle Payload Integration Facility (SPIF) in the west low bay of the SMAB where it is mated and integrated with the spacecraft. The IUS/spacecraft cargo element is tested as a unit, and cargo element to Orbiter interface checks are made using Orbiter simulation. The IUS/spacecraft cargo element is transported from the SPIF to launch pad 39 as a unit, using the environmentally controlled Multi-Mission Support Equipment (MMSE) transporter/canister. The IUS/spacecraft cargo element is installed in the Rotating Service Structure (RSS) at pad 39 for additional checks and servicing before being installed in the Shuttle Orbiter. Final checkout and servicing of the IUS, spacecraft and Orbiter are accomplished just before launch.

In the NASA Shuttle flow, the IUS will be transported from the SMAB east bay to the Vertical Processing Facility (VPF) where the spacecraft will be integrated with the IUS. The IUS/spaceship to Orbiter interface is checked by means of the orbiter simulator before removal from the VPF. The IUS/spaceship are taken from the VPF to launch pad 39 on the MMSE transporter/canister. The IUS and spacecraft are installed in the RSS for additional checks before being placed in the Orbiter.

In the Titan 34D flow, all integration and servicing activities of the IUS, spacecraft, and Titan are accomplished at launch pad 40. A Mobile Service Tower (MST), which contains an environmentally controlled section that enclosed the IUS/spaceship, is used for prelaunch integration and servicing. The MST is mounted on rails and is rolled back out of the way just before launch.

## CONCLUSION

The Inertial Upper Stage has been a major undertaking by Boeing Aerospace Company since 1973.

The foundation of this program was linked to the premise that the IUS was to be a gap-filler, an interim vehicle that would bridge the gap between the then-existing expendable class of upper stages and the advanced-technology space tug. When it became apparent that the space tug would not be developed in the foreseeable future, the "interim" upper stage began a gradual evolution to the "Inertial Upper Stage". This evolution brought with it a growing sophistication in requirements and design to give the vehicle a viability over a long period of time and to serve a growing list of potential users. The vehicle design transitioned from an elementary growth version of the Burner II upper stage to an all new upper stage. This configuration was dependent upon numerous technology "firsts" and was to be developed in parallel with two boost vehicles, the T34D expendable launch vehicle and the reusable Space Shuttle.

Transition from the initial interim upper stage to the eventual Inertial Upper Stage covered a period of several years. The Burner II, with single-string avionics having a simple guidance timer sequencer, eventually became the multipurpose IUS with redundant avionics and a digital software system. The advanced system is capable of retargeting from low Earth orbit, energy management for the fixed-impulse Solid Rocket Motors, self-test, and redundancy management logic to enable in-flight switchover from a failed system to a redundant operable system. Ultra-high reliability was probably the most significant feature implemented.

The propulsion system is simple, utilizing inherently safe, Solid Rocket Motors with extremely light weight nonmetallic cases and nozzles. It is capable of flying a wide range of missions requiring significantly different propulsive energy management schemes.

The generic thermal design of the IUS envelops thermal environments from the worst-case mission in the mission model. The software design provides for selectable thermal maneuvers (rotisserie, reciprocating, toasting, space facing, or sun facing) to satisfy different payload thermal requirements.

Airborne Support Equipment features a simple interface with the Shuttle astronauts to enable erection, checkout, and deployment of the IUS from the Shuttle. The equipment is structurally designed to accommodate a wide range of spacecraft configurations, including spacecraft designed for expendable launch vehicles and transitioning to the Space Shuttle.

These features ultimately characterized the shift from an interim vehicle to a long-term IUS usable well into the 1990's and beyond.

## DELTA'S ROLE IN REACHING THE FOURTH ENVIRONMENT

D. W. Grimes\* and J. K. Ganoung\*\*

\**Delta Project, NASA/Goddard Space Flight Center, USA*

\*\**Spacecraft Integration, McDonnell Douglas Astronautics Corp., USA*

### ABSTRACT

The Delta launch vehicle has had a significant role in providing access to mankind's fourth environment of space. Since the first Delta launch in 1960, the variety of spacecraft placed into orbit has supported all facets of the peaceful utilization of space including science, meteorology, telecommunications, navigation and earth resources. A high volume period of service through the mid-1980's is currently planned for Delta-class payload launches for both firm missions and for Shuttle backup missions. To enhance the utility of the Delta for these new missions and as a true Shuttle backup operational launch vehicle, various modifications are currently being accomplished.

Included in the paper is a description of (1) the historical role which Delta has played in launching satellites from 1960 to the present, (2) the currently forecasted Delta launch schedule, (3) the vehicle modifications now underway to provide spacecraft interchangeability with the Shuttle and (4) various potential growth versions of the Delta currently being investigated to satisfy the ever-expanding requirements for the further exploration of the fourth environment.

### KEYWORDS

Launch Vehicle, Delta, Shuttle, Space Transportation System, PAM-D, Delta 3920

### INTRODUCTION

Technological developments have historically been at the forefront of acquisition of knowledge and the resultant application of mankind's several environments; the geosphere, the hydrosphere, the atmosphere, and most recently, the fourth environment - SPACE, the theme of the 32nd International Astronautical Congress. This paper addresses one of these developments which continues to provide access for the exploration and utilization of this new environment: a launch vehicle transportation system. Since its first launch in 1960, the Delta launch vehicle has played a significant role as a major space transportation system and has earned the title of "NASA Space Transportation Workhorse." The wide variety of spacecraft placed into orbit has supported all facets of the peaceful utilization of space including science, meteorology, telecommunications, navigation, and earth resources.

Included in this paper will be a brief summary of Delta's past achievements, a description of current vehicle configurations and capabilities, an outline of the program role within the launch vehicle community, and a projection of Delta's utilization in the further development of the space environment.

### PAST UTILIZATION OF DELTA

Of 154 launches since 1960, 143 Deltas have been successful for an overall reliability of 93 percent. Substantial reliability improvements, implemented following the Delta 100 failure in 1974, have subsequently improved the success rate. For the last 52 launches spanning approximately 7 years, 50 Deltas have been successful for a reliability increase from 91 percent for the early period to 96 percent for the current period.

For reference purposes, a brief summary of the growth of the Delta vehicle and its launch record is presented. Figures 1 and 2 summarize the vehicle modifications incorporated between 1960 and 1980 and Fig. 3 illustrates the performance growth history. More detailed information may be found in the list of references.

DELTA MODEL DESIGNATION AND MODIFICATION YEAR													
Delta	A	B	C	D	E	J	M	M6	904	1914	2914	3914	3910 PAM
1960	1962	1963	1964	1965	1968	1969	1970	1971	1972	1973	1975	1980	
① Modified Thor									① Increased Length of Tanks, Isogrid Structure				
② Vanguard Propulsion System									② Added Interstage and Suspended Stage Within				
③ Vanguard X-248									③ Increased Motor Length TE-364-4				
① Updated MB-3 Engine									④ Increased Diameter and Length				
② Lengthened Tankage, Higher Energy Oxidizer, Transistorized Guidance Electronics									① Replaced MB-3 Engine With More Powerful RS-27 Engine				
③ Scout Developed X-258 Replaced X-248									① Increased Size of Solid Rockets C-IV Replaced C-II				
④ Bulbous Fairing									③ Payload Assist Module (PAM)				
① Augmented Thrust Using Thor Developed Strap-on Solid Rocket Motors													
② Larger Diameter Propellant Tanks From Able-Star Stage, Restart Capability													
③ USAF Developed FW-4 Replaced X-258													
④ Larger Fairing From USAF Agena Stage													
① Surveyor Retromotor Modified TE-364-3 Replaced FW-4													
① Lengthened LOX and RP-1 Tanks Increased Diameter of RP-1 Tank													
① Added 3 Solid Rockets to Total 6													
① Introduced UBT to Accommodate 9 Solid Rockets													
② Titan Transtage Engine Strapdown Inertial Guidance													

#### LEGEND

- ① First Stage Modification
- ② Second Stage Modification
- ③ Third Stage Modification
- ④ Payload Fairing Modification

Fig. 1. Delta growth history.

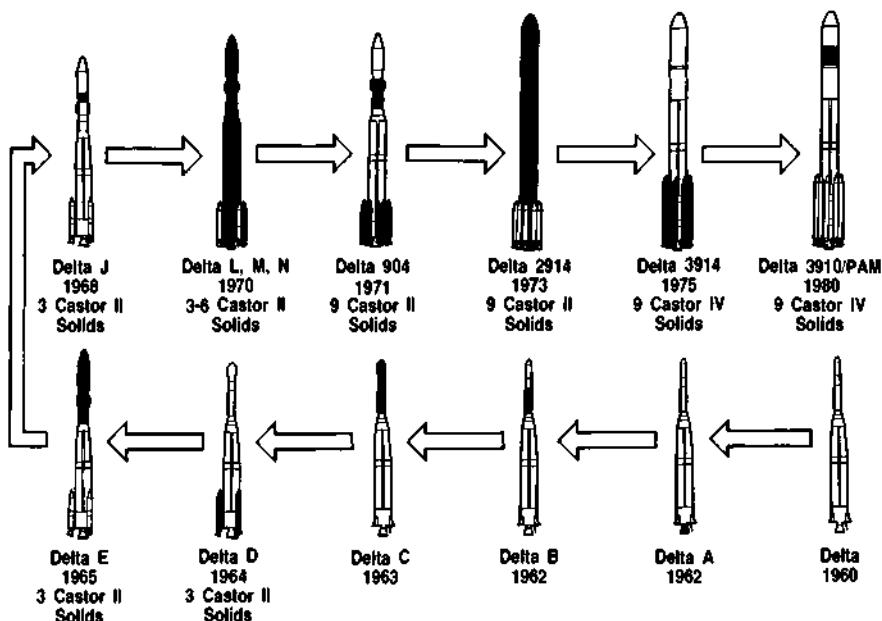


Fig. 2. Delta vehicle history.

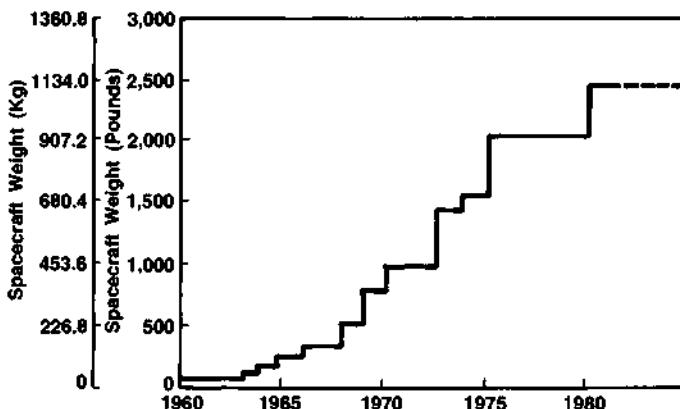


Fig. 3. Delta performance history (geosynchronous transfer missions).

### CURRENT DELTA ROLE

The Delta Program is currently facing the largest number of booked launches ever experienced in the 21-year history. A combination of factors has led to this situation: (1) the planned transition from expendable launch vehicle systems to the reusable Shuttle launch system and the apparent user reluctance to plan to launch during that transition created a discontinuity in service growth patterns, and (2) the rate of demand for the worldwide use of geosynchronous communication and meteorological satellites was unanticipated. These factors have created a demand situation which almost exceeds the capability of both the Shuttle launch system and a fully operational Delta launch system.

The current Delta launch schedule at the time of this paper's publication is shown in Fig. 4. This schedule will be iterated as the Shuttle becomes operational and provides some opportunities for Delta-class payloads. However, the list of customers is still increasing, and the tentative bookings now approach Delta 210.

Delta No.	1981												1982												1983																			
	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A												
154							▲	5-14-81	GOES-E																																			
155													▲	7-31-81	DE																													
156													▲	8-20-81	SBS-B																													
157													▲	9-15-81	SME																													
158													▲	10-15-81	RCA-D																													
159													▲	12-3-81	RCA-C1																													
160															▲	1-7-82	WESTAR-IV																											
161															▲	2-18-82	INSAT-1A																											
162																▲	4-1-82	IRAS																										
163																	▲	5-13-82	TELESAT-E																									
164																		▲	7-1-82	LANDSAT-D																								
165																			▲	8-5-82	TELESAT-F																							
166																				▲	9-16-82	WESTAR-V																						
167																					▲	10-28-82	RCA-E																					
168																						▲	12-2-82	GOES-F																				
169																							▲	1-83	PALAPA-B1	▲																		
170																								▲	2-83	RCA/ASTRO	▲																	
171																									3/83	RCA-F	▲																	
172																										4/83	TELESAT-G	▲																
173																											5/83	HC-A	▲															
174																											6/83	TELSTAR-3A	▲															
175																												TBD	LANDSAT-D	▲														
176																														8/83	INSAT-1B	▲												

Delta No.	1983												1984												1985																	
	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D												
177							▲	9-83	HC-B																																	
178							▲	10-83	NATO-111D																																	
179										▲	11-83	RCA-G																														
180										▲	1-84	PALAPA-B2																														
181										▲	2-84	SPC-A																														
182										▲	3-84	GSTAR-A																														
183										▲	4-84	TELESAT-I																														
184											▲	5-84	TELESTAR-3B																													
185											▲	6-84	HC-C																													
186												▲	8-84	AMPTE																												
187												▲	9-84	GSTAR-B																												
188													▲	10-84	SPC-B																											
189														▲	11-84	NATO-111E																										
190															▲	1-85	RCA-H																									
191																▲	2-85	COMGEN-A																								
192																	▲	3-85	COMGEN-B																							
193																		▲	5-85	SPC-C																						
194																			▲	6-85	WESTAR-VI																					
195																				7-85	WESTAR-VII	▲																				
196																					8-85	ANSCS-A	▲																			
197																						9-85	GOES-G	▲																		
198																							10-85	ANSCS-B	▲																	
199																								11-85	SPC-D	▲																
200																												12-85	GOES-H	▲												

Fig. 4. Delta project master schedule (April 1981).

## DELTA UP RATINGS

As Delta enters into the Space Shuttle era as a companion launch vehicle of the NASA Space Transportation System, additional performance requirements of potential payloads necessitate a review of possible improvements to the launch vehicle. Throughout its history, Delta has been continually called upon for more performance, larger space accommodations, and greater accuracy. Today's requirements are no different. The primary demands are for additional weight and increased volume to provide for the larger diameter spacecraft of the commercial communication satellite customers which comprise the bulk of future Delta users.

Delta is providing the increased weight capability by incorporating several modifications which are discussed below. In order to accommodate the future Delta growth demands, other potential upgrading concepts are being considered and are also discussed.

### Current Modifications

Payload Assist Module. The advent of Shuttle promises an economical capability to carry large and heavy payloads into low-earth orbit. However, only a relatively small number of payloads make use of the low-earth orbit; most spacecraft, especially the communications and meteorological satellites, instead require thrust augmentation to achieve their operational near-earth orbits or their higher geosynchronous operational orbits. This need for an upper stage thrust augmentation system prompted a wide ranging series of studies during the Shuttle conceptual phase covering the adaptation of existing stages and the development of new stages and systems. Current stages now under development are the Inertial Upper Stage (IUS), the Spinning Solid Upper Stage (SSUS)-A and -D stages, and the Wide Body Centaur Stage. For the Delta class payload, a Payload Assist Module (PAM), also called SSUS-D, was proposed by McDonnell Douglas Astronautics Company (MDAC), the Delta prime contractor, to provide an orderly transition from the Delta expendable vehicle to the Shuttle reusable vehicle. An agreement was subsequently reached between the National Aeronautics and Space Administration (NASA) and MDAC which permitted the commercial development of the PAM-D stage for direct procurement by the user community. This system, being adaptable to either an STS or a Delta launch, provides the users with a timely backup capability to offset possible delay in the Shuttle operational readiness date.

A larger system designated PAM-A or SSUS-A is also currently under development by MDAC to support larger (Atlas/Centaur) class payloads being launched by Shuttle.

The PAM-D system elements are shown in Fig. 5, along with their assembled configuration for the Delta and Shuttle vehicles. For Shuttle use, the system has an expendable stage consisting of a spin-stabilized solid-fueled motor, a payload attach fitting (PAF) to mate with the spacecraft, and the necessary timing, sequencing, power, and control assemblies mounted on the PAF. The reusable Airborne Support Equipment consists of the cradle structure for mounting the deployable system in the Shuttle, a spin system to provide the stabilizing rotation, a separation system to release and deploy the stage and spacecraft, and the necessary avionics to control, monitor, and power the system. The PAM-D will also provide a sun-shield for thermal protection while the Orbiter bay doors are open.

The PAM-D stage is supported through the spin table at the base of the motor and through forward restraints at the PAF which are retracted before deployment.

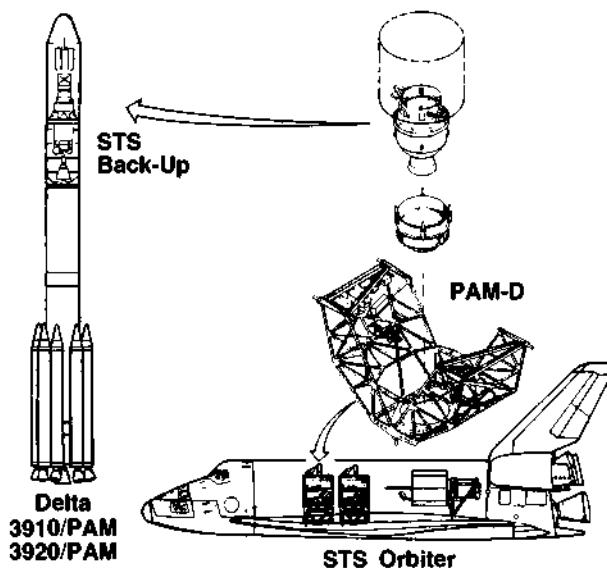


Fig. 5. MDAC Payload Assist Module program (PAM)

When the PAM-D is used with the Delta, the spacecraft interface remains the same but certain changes are made to the stage. The forward restraint fittings are removed, the standard Delta confined detonating fuse (CDF) is employed for sequencing, and the spin table is configured to mount to the Delta second stage and accommodate spin rockets for rotation. The common hardware is shown in Fig. 6.

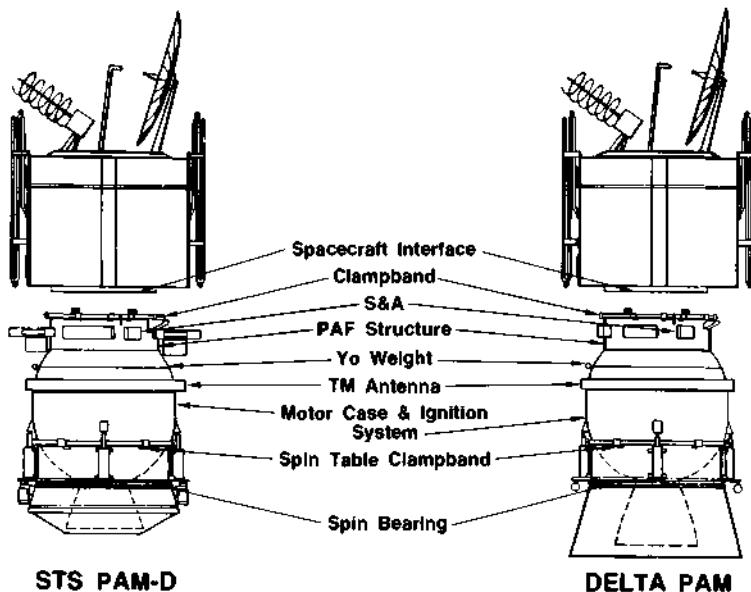


Fig. 6. Common hardware STS PAM-D and Delta PAM.

The PAM-D Thiokol STAR-48 solid rocket motor features a titanium case, 89 percent solid HTPB propellant, a toroidal pyrogen ignition system, a 3-D carbon-carbon throat insert, and a 2-D carbon-carbon exit cone. Fully loaded it carries 4,400 pounds of propellant and offloaded, 3,833 pounds of propellant. The fully loaded impulse of the motor is 1,268,500 lb-sec with an action time of 85.3 seconds.

The first PAM-D was flown in November 1980 when Delta 153 successfully launched the SBS-A spacecraft.

3920 Improved Second Stage. In mid-1979 it became apparent that the expendable to Shuttle transition time would increase over that previously planned. This required not only more Delta launches, but improved performance for the communications and scientific satellites then being designed to take advantage of the greater Shuttle payload capability. Additionally, the Landsat-D Project was having increasing difficulties keeping within the 3910 payload capability, thereby creating a need for additional Delta performance. Landsat-D, with a scheduled launch date of mid-1982, served as the impetus to increase the Delta performance again in a timely and economical manner.

Aerojet Services Company has designed and developed a new second stage for the Japanese N-II launch vehicle, utilizing a Titan transtage engine and increased tankage for delivering more optimum performance. Additionally, the Air Force had funded a development program, called the Improved Transtage Injector Program (ITIP) to increase the performance of the FJ injector. It was decided to utilize the improved ITIP engine and adapt it to the existing N-II stage tanks along with usage of Delta 3910 second stage common components where possible. This plan formed the basis for the new Delta 3920 Improved Second Stage (ISS).

The 3920 stage has the same overall length as the 3910 of 19.6 feet, and is structurally mounted within the existing interstage. The tank diameter has been increased from 54.5 inches to 68.8 inches, and the guidance section and miniskirt assemblies have been strengthened to accept the greater loads. The equipment panel has been redesigned to an isogrid configuration attached to the aft tank section. As can be seen in Fig. 7, the propellant capacity was increased from 10,072 lbs. to 13,291 lbs. and the total impulse was increased from  $3.04 \times 10^6$  lb-sec to  $4.7 \times 10^6$  lb-sec. The ITIP engine is pressure fed similar to the 3910 TR-201 engine.

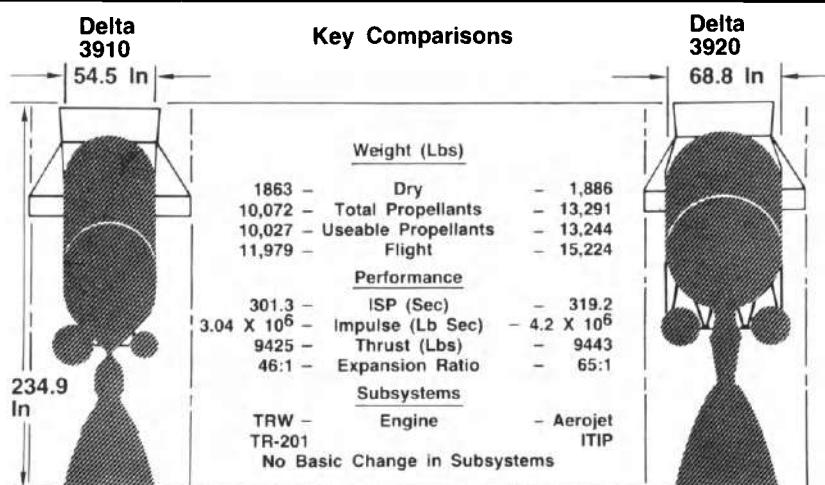


Fig. 7. Delta 3920 configuration Improved Second Stage.

An isometric cutaway view of the 3920 ISS is shown in Fig. 8. As in all past Delta upgrades previously qualified subsystems have been used, where possible, in the interest of economy and reliability. The RACS, hydraulic system, and pneumatic system are as used on the 3910 second stage. Additionally, redundant features previously developed were employed, e.g., dual coil TSPV and actuator pots. The Aerojet-supplied major hardware components are the tank, ITIP engine, and the nitrogen and helium pressure vessels.

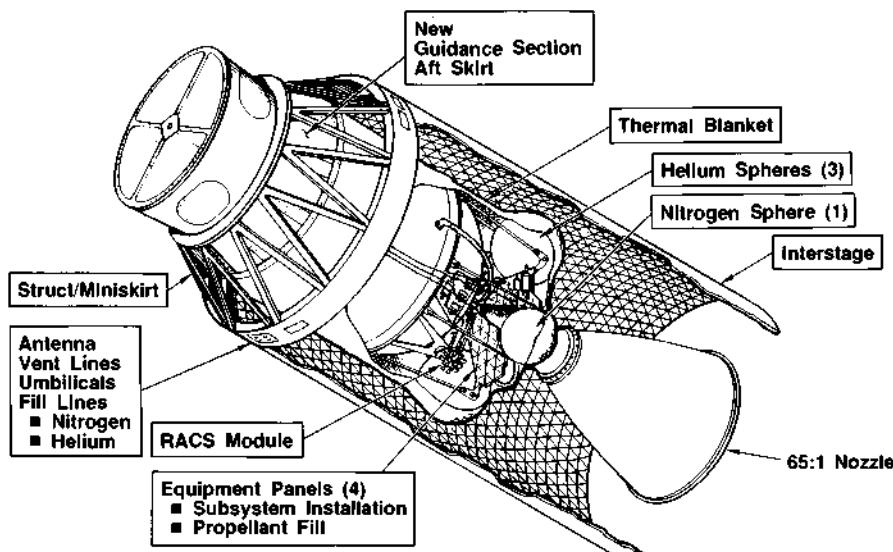


Fig. 8. Delta 3920 Improved Second Stage.

The first 3920 Delta configuration is currently planned for the Landsat-D mission scheduled for launch in July 1982.

Firing Sequence of Solid Rocket Motors. Another modification currently being implemented for performance improvement is the firing resequencing of the Castor-IV solid rocket boost assist motors. Rather than igniting only 5 of the 9 solids at liftoff and then 4 during the ascent trajectory, the sequence change ignites 6 solids on the pad and 3 at altitude. The increase in spacecraft weight in geo-synchronous transfer orbit amounts to approximately 45 pounds for the 3914 configuration and 60 pounds for the 3910/PAM configuration. All 3920/PAM vehicles will use the 6/3 sequence mode. The first operational flight for this new Castor IV sequencing is scheduled for late July, 1981, with fleet change scheduled for early 1982.

Launch Operations. In addition to the launch vehicle developments underway, modifications to the launch facility at the Eastern Test Range have been initiated to accommodate a rate of 10 launches per year. This increased launch rate capability is being effected by reactivation of the second Delta launch pad (Pad 17B). The modifications required are primarily directed toward the accommodation of the Castor-IV solid motor strap-ons, strengthening decks, enlarging cutouts and increasing lifting hoist capacity (Fig. 9). These changes, currently underway, are scheduled for completion by January 1982.

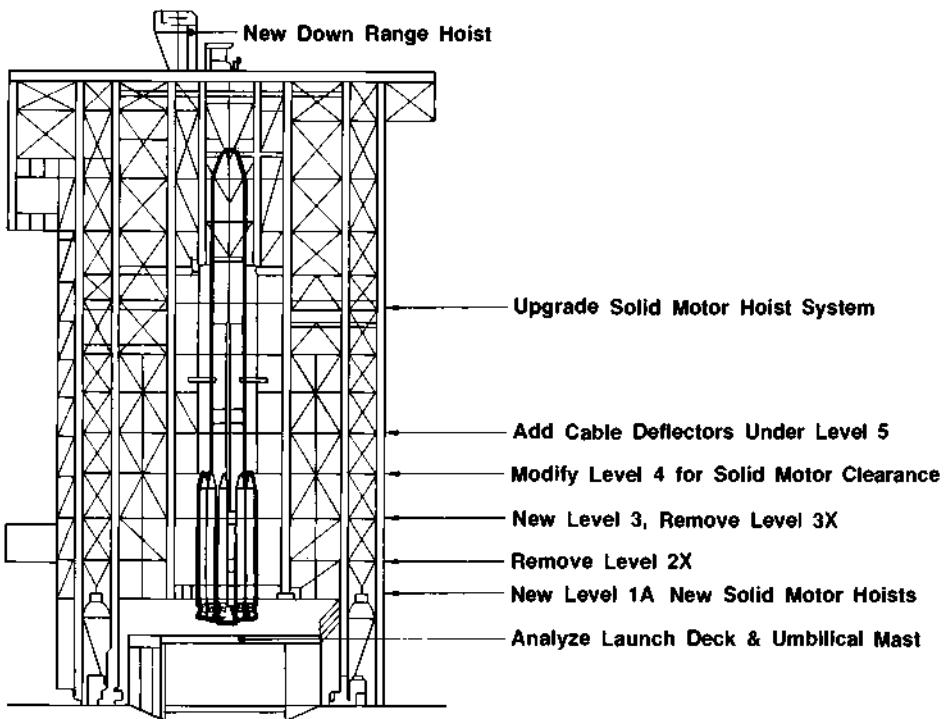


Fig. 9. Launch complex 17B.

#### Potential Future Improvements

**Delta 4920.** A vehicle growth concept recently evaluated by MDAC incorporates a larger payload volume for the spacecraft and provides additional transfer orbit throw weight capability up to 3,000 pounds. This concept, designated the Delta 4920 configuration (Fig. 10) utilizes the MDAC produced Titan IIIC 10-foot diameter fairing design for additional payload volume and a 12-foot lengthening of the booster to provide additional propellant loading. All other elements of the vehicle are essentially the same as the current 3920/PAM configuration with some minor structural modifications to accept the higher loads. The 10-foot diameter fairing (Fig. 11) uses a Titan IIIC nose module, 2 Titan IIIC intermediate modules, and a new base module which adapts to the 8-foot diameter interface. The internal dynamic envelope available to the user would be 108 inches in diameter and 156 inches in height. Missions which are designed for the STS using a PAM-D are required to maintain a maximum vertical height of 101 inches. As seen in the figure, this height is easily accommodated in the fairing design for STS missions requiring a Delta backup. Spacecraft flying a dedicated Delta mission can use the additional vertical height for larger antenna configurations.

**Nine-foot Diameter Fairing.** Another study was recently completed for the purpose of providing additional spacecraft volume by increasing the diameter of the upper portion only of the Delta fairing. This "hammer head" configuration, shown in Fig. 12, results in a usable payload envelope of approximately 98 inches diameter, an increase of about 10 inches over the existing Delta fairing. The payload capability of this vehicle is reduced a modest amount from the standard Delta 3920.

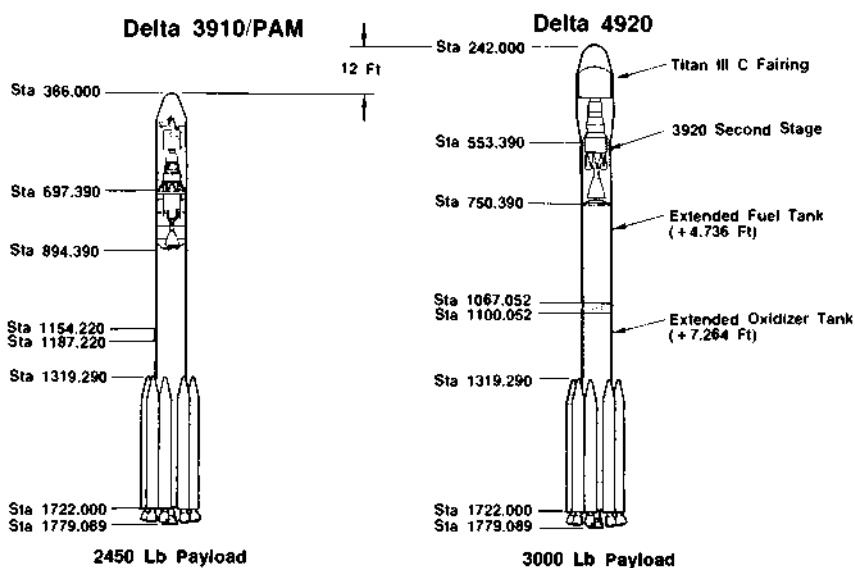


Fig. 10. Delta 4920 configuration (conceptual).

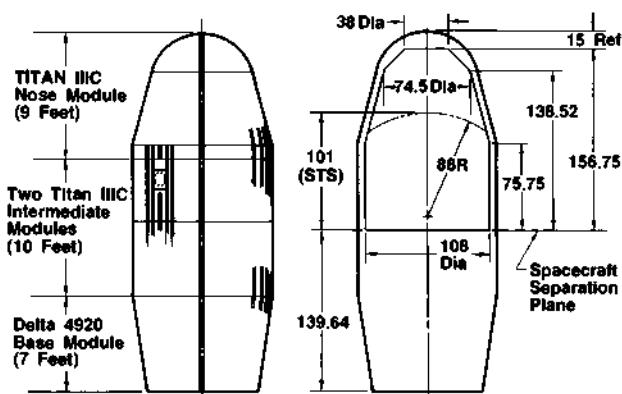


Fig. 11. Delta 4920 fairing (10 ft diameter).

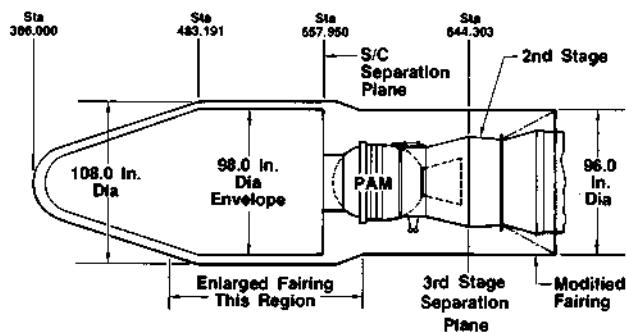


Fig. 12. Nine foot payload fairing concept Delta 3920/PAM.

Booster Engine Performance. In accordance with the Delta Program philosophy of utilizing flight proven, qualified hardware the current first stage engines (Rocketdyne RS-27) are surplus H-1 engines produced for the NASA Saturn Apollo program. Since the supply of these engines will be exhausted in the 1984 time-frame, a new production order will be required for both the engine chambers and new tooling. This new order provides the opportunity to better balance the first stage propulsion systems.

A study has been initiated to determine the impact of implementing a more optimum (higher expansion ratio) chamber on the RS-27 engine. Chambers with expansion ratios of 12, 15, and 21:1 were compared for performance as well as vehicle and facility modifications with the existing 8:1 chamber employed on the current engine.

Performance increases to geosynchronous transfer orbit of greater than 100 and 300 pounds of spacecraft weight can be achieved for the 12:1 and 21:1 chambers, respectively. However, a chamber as large as 21:1 has a significant impact on vehicle and facilities. Therefore, consideration has been centered on chambers in the range of 12:1 to 15:1. Further study is defining the magnitude of vehicle and facility changes required and will result in the selection of a RS-27 chamber expansion ratio for missions subsequent to Delta 185 which optimizes performance and minimizes vehicle impact.

PAM Solid Motor Performance. Another example of growth potential exists within the PAM stage itself. By increasing the length of the motor case, an increase in propellant capacity of approximately 15 percent is possible. With the existing PAM nozzle, the spacecraft separation plane would be relocated approximately 5.4 inches forward. To maintain the existing separation plane the nozzle could be foreshortened similar to the STS PAM application. An increase in booster performance would be required in conjunction with this change.

Universal Second Stage. Preliminary studies have shown that the current TRW second stage engine (TR201) used in conjunction with the 3920 tank could yield a performance increase of approximately 100 pounds over and above the existing 3910 vehicle for a typical geosynchronous mission. The benefit to this modification would be the use of a single propellant tank for all second stage configurations, thus allowing the production line for the existing 3910 propellant tanks to be phased out. The stage would become universal in that propellant tanks, pneumatic system and attitude control systems would be common and only the engine, propellant feed-lines and interfacing hardware would be different.

This modification does require that the TR201 engine operate at a mixture ratio of 1.9 and at lower interface pressures as opposed to a mixture ratio of 1.6 on the current 3910 tankage. TRW is currently under contract to demonstrate the feasibility of this new operating point and should complete testing in the fourth quarter of 1981. It appears, from current planning, that a cost benefit would accrue to the user who could accept this mid-range performance vehicle configuration.

Hydrogen-Oxygen Second Stage (HOSS). Throughout its history Delta has emphasized the use of flight-proven hardware and current technology when upgrading vehicle performance. Numerous studies of a cryogenic propulsion system for the second stage based on the Saturn experience of the 1960's, have been performed. In 1970, at the 19th Congress, the original studies were first reported. Liquid hydrogen and liquid oxygen are used as the propellants with the most recent studies showing an increase in the geosynchronous transfer capability to approximately 3,700 pounds.

Large Strap-On Solids. Current studies show that the use of a lesser number of larger strap-on solids is an alternative approach for achieving increased capability to the 3,700 pound range for future Delta geosynchronous transfer missions. Further engineering and programmatic studies of this approach will determine its benefits as compared to the spectrum of other options.

The continuous series of vehicle upratings, some of which have been discussed, are reflected in Delta's past, current and possible future performance capabilities as presented in Fig. 13.

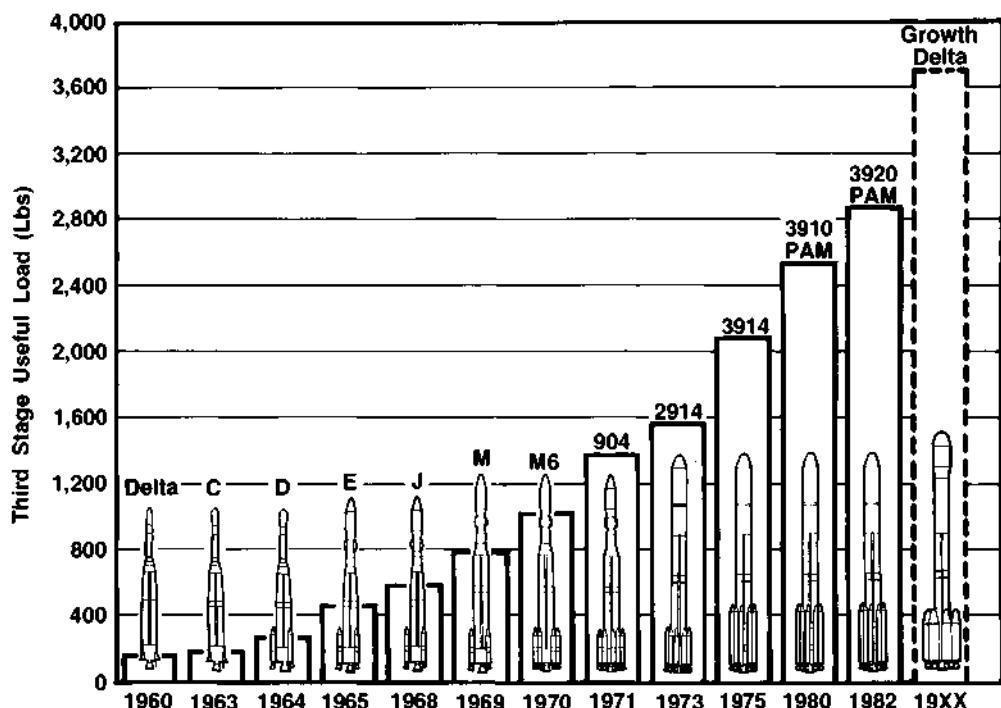


Fig. 13. Delta performance growth (geosynchronous transfer).

## SUMMARY

It has been shown that the ongoing technological development of the Delta launch vehicle to match performance capability to payload requirements may continue well into the transitional period of Shuttle operations. In mankind's quest to explore the fourth environment, SPACE, the payload requirements steadily grow as additional knowledge is gained. To meet these requirements, the Program will constantly be examining new concepts and approaches. This theme of continuity with change has been the hallmark of the Delta Launch Vehicle Program. The third decade of Delta will see a preservation of this theme as the vehicle meets the needs of the user community in exploring and developing the fourth environment.

## REFERENCES

- Bonnett, E. W. (1974). A Cost History of the Thor-Delta Launch Vehicle Family. 25th International Astronautical Congress, Amsterdam, Holland.
- Ganoung, J. K. (1980). Low Cost Transportation of Communication Satellites to Orbit--The Role of Delta. 31st International Astronautical Congress, Tokyo, Japan.
- Ganoung, J. K. and H. Eaton, Jr. (1981). The Delta Launch Vehicle--Past, Present, and Future. Eighteenth Space Congress, Canaveral Council of Technical Societies, Cocoa Beach, Florida.
- Grimes, D. W. and C. J. Arcilesi (1980). An Example of the Communication Network for a Major Launch Vehicle Program.
- Gunn, C. R. (1973). The Delta Launch Vehicle Model 2914 Series, X-470-73-262, Goddard Space Flight Center, Greenbelt, Maryland.
- Kork, J. and W. R. Schindler (1970). The Thor-Delta Launch Vehicle: Past and Future. 19th International Astronautical Congress, Warsaw, Poland.
- Ordahl, C. A. (1980). Shuttle/Spacelab--The New Transportation System and its Utilization. DGTR/AAS Symposium, Hanover, Germany.
- Delta Spacecraft Design Restraints Manual (1980). MDAC Report No. DAG61687D.
- NASA/GSFC Document (19/9). Delta 1960-1979: 150 Missions Exploring Space.

## AN ECONOMICAL AND FLEXIBLE ALTERNATIVE ORBITAL TRANSFER VEHICLE

R. F. Brodsky

*TRW Space Systems Division, Redondo Beach, California, USA*

### ABSTRACT

This paper describes a study to determine the characteristics and costs of an all-liquid low-thrust propulsion module to replace the Inertial Upper Stage (IUS) as a means for obtaining geosynchronous orbit (GEO). The propulsion module proposed by TRW is propelled by low-thrust gimballed engines, using storable bipropellants from a multitank pressure-fed system. It mounts in a cradle in the Space Transportation System (STS) Orbiter cargo bay, and provides a platform for integration of spacecraft subsystems and payload mounting. Performance flexibility is obtained as a result of varying propellant capacity by adjusting the length of the cylindrical portion of the propellant tankage. Although the specific impulse of this system is comparable to that obtainable from the IUS solid motors, important gains in performance and cost come from weight savings and equipment reduction. The latter is achieved by avoiding duplication of many functions readily available from the spacecraft that the propulsion module is transporting.

The TRW propulsion module concept was applied to transfer a growth version of the Tracking and Data Relay Satellite (TDRS) to GEO. The ground rules for the study permitted no changes to TDRS hardware and maximum utilization of its on-orbit subsystems to assist the propulsion module. A novel guidance system which requires little ground assistance was adopted. The study achieved its aim of proving design feasibility and economy.

### KEY WORDS

Propulsion module, orbital transfer vehicles, TDRS, IUS, multiple perigee burns, and low-thrust propulsion.

### INTRODUCTION

The idea of using an integrated propulsion subsystem to assist spacecraft in achieving higher energy orbits above the STS parking orbit has been proposed in various military programs, and by NASA Goddard Space Flight Center (GSFC) in the Multimission Modular Spacecraft (MMS).

TRW had been studying designs for the application of an integral propulsion concept for delivery of heavy spacecraft in the 2200- to 4400-kg class to GEO and similar orbits of high  $\Delta V$  requirements. Recently, a planned growth version of the TDRS

presented an opportunity to make a specific design study. Since the projected weight increase exceeded the capabilities of the IUS which was slated to transfer the TDRS from the STS orbit to GEO, and since early capability also was desired, the proposed propulsion module design draws heavily on the technical background developed in previous company-sponsored study programs. The purpose of this study, then, was to determine the characteristics of a propulsion module to place a growth version of the TDRS into orbit, and to estimate the cost and schedule of such a development. The following ground rules were adopted for the study:

- No change to TDRS hardware
- Payload of 5200 pounds or greater to GEO
- Maximum utilization of available TDRS subsystems
- Mechanical adapter similar to IUS adapter
- Safe disposal of spent propulsion module
- Storable propellants
- Available propulsion systems
- Timely development.

#### THE TDRS STUDY

##### Description of Propulsion Module

The propulsion module selected (Fig. 1) is propelled by two 300-pound-thrust gimballed Viking engines, using storable bipropellants from an unstaged, multi-tanked, pressure-fed system. The design evolved from an integral spacecraft staged propulsion subsystem (Fig. 2) which TRW has been studying for a number of years in an Independent Research and Development (IRAD) program. These IRAD studies

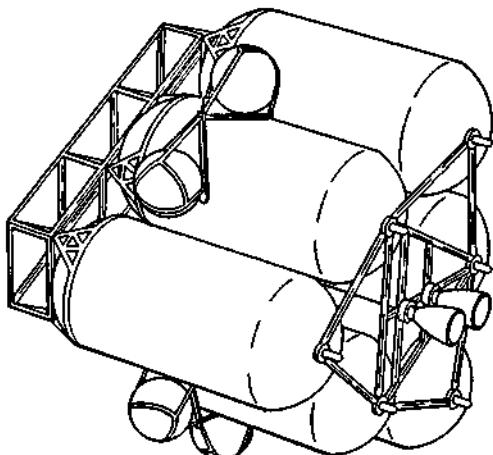


Fig. 1. TDRS propulsion module (unstaged).

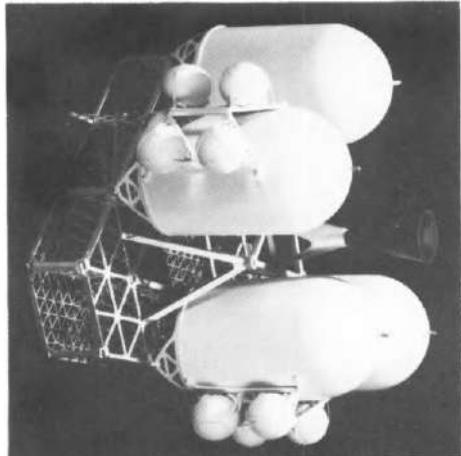


Fig. 2. Early version of propulsion module.

concentrated on low-thrust systems which derive much of their subsystem support from the host spacecraft.<sup>1</sup> Application of these ideas and detailed analyses to the case of the TDRS spacecraft was therefore logical, and influenced many of the salient design features of the Fig. 1 configuration. In particular:

- The six-tank, triangular isogrid configuration was selected because of the considerable analysis done on staged vehicles (dropping one tank pair at a time), with tankage arranged in a triangular fashion to maintain the thrust line cg at the centroid of the triangle (i.e., a heavy inboard NTO tank balanced by its lighter outboard MMH tank pair). This six-tank configuration also makes efficient use of the available frontal area. However, for vehicles where staging will never be considered, other configuration options can be considered.
- The isogrid equipment platform, which provides a mounting structure for both the TDRS-IUS adapter and the propulsion module avionics, was selected because it makes a stiff, strong, lightweight (aluminum) structure amenable to fabrication by computer control. It is easily adaptable to various applications, since individual members can be strengthened to handle local load concentrations.
- Since the designer can make use of the 15-foot cargo bay diameter, a short propulsion module can evolve even for very heavy payloads. For example, propulsion modules on the order of 10 feet in length are practical with storable propellants to boost 10,000-pound-class payloads to GEO. Performance flexibility is obtained as a result of varying propellant capacity by adjusting the length of the cylindrical portion of the propellant tankage.

The TDRS propulsion module replaces the IUS as the means to integrate the flight segment to the STS via a system of Airborne Support Equipment (ASE). The new ASE is designed to use the identical sill and keel fittings used by the IUS. Table 1 provides a weight breakdown of the TDRS propulsion module which meets the twin constraints of transferring a 5200-pound payload to GEO while limiting the STS cargo bay launch payload weight to its early capability of around 41,000 pounds at lift-off.

Past TRW work with large-sized STS-compatible propulsion modules assumed that the module and spacecraft would be designed as an entity, and that consequently, hardmounting of the module to the STS still could be accommodated. However, in the case of cantilever mounting an already-designed TDRS spacecraft to a new structure, load amplification may not permit this simple approach. From structural modeling, it was apparent that two of the IUS' prime payloads (the TDRS and a military payload) could not withstand loads introduced through the STS sill and keel fittings. In the case of the TDRS, the hard-mounting method stressed the farthest extremities of the folded large (16-foot) single-access antennas beyond their design yields. A similar problem occurred with the other payload. Thus, the IUS was forced to adopt a low-frequency ( $\approx 3$  Hz) load alleviation system which added considerable cost,

<sup>1</sup>H. Macklis, R. L. Sackheim, and B. Vogt, "Selection of an Optimized Integrated Propulsion System," AIAA-80-1211, presented at AIAA/SAE/ASME 16th Joint Propulsion Conference, July 1980.

TABLE 1 Weight Breakdown (Pounds)

Payload to GEO	5200
Added monopropellant	40
	<hr/>
Total payload	5240
Dry propulsion module	2941
Propellants	29203
Propellant biases + helium	495
	<hr/>
Total wet module	32639
Flight vehicle	37879
Aerospace Support Equipment (cradle)	1675
Margin*	1446
	<hr/>
Cargo bay at launch	41000

\*Or could add .331 pounds to payload

complexity, and design effort. As indicated above, the TDRS propulsion module adapts to the STS at the same sill and keel stations as the IUS. On the other hand, its structural parameters are quite different; indeed, it appears to be less stiff than the IUS. Consequently, it is not presently clear whether load alleviation would be required for the TDRS propulsion module.

In the present early stages of STS utilization, the analysis of the loads induced throughout launch and abort landing (which in the TDRS case is considerably more benign) must be done by painstaking and costly methods to arrive at answers which are still tentative. Tapes are available which were obtained through a combination of analysis, reduced-scale, and full-scale testing (the latter performed in the vertical dynamic test rig at Huntsville); they simulate sill and keel load input spectra at various discrete sill and keel stations. These forcing functions are translated into compatible machine language and caused to act upon a complete (probably finite element) structural simulation of the spacecraft and transfer vehicle.

It was just such analyses that originally revealed the need for load alleviation by the IUS. Although a proper structural simulation of the TDRS was available for the study, the structural simulation of the propulsion module used an earlier design (i.e., that of Fig. 2), which is dissimilar in structure and weight to the current TDRS module design. Thus, while the simulation did indicate the same high-loading problem on the single-access antennas, only a detailed analysis with a substantially improved propulsion module model could positively identify the problem. The need for load alleviation cannot be taken lightly. If load alleviation is found necessary, then additional design and test programs would add about 3 months to a normal 34 month-to-launch program, while engineering, manufacturing, and testing would add substantial nonrecurring costs, plus modest production unit cost increases.

Fig. 3 shows the weight change to the ASE cradle as a result of adding load alleviation equipment. The addition of this equipment reduces the weight margin of Table 1 by over 500 pounds. Fig. 3 also reveals that it is possible to reduce the weight of the isogrid forward equipment platform by transferring fore and aft loads to the aft sill fitting. This change lowers the weight of the flight segment ejected from the STS, at the expense of increasing the ASE weight. Normally this is a favorable tradeoff, but in view of the considerable detailed effort invested in past analyses and structural models, the potential gain was not taken advantage of in the study. The majority of the ASE weight increase required by the addition of load alleviation equipment is invested in leaf springs, torsion bars, dampers, and load levelers similar to those used by the IUS.<sup>2</sup>

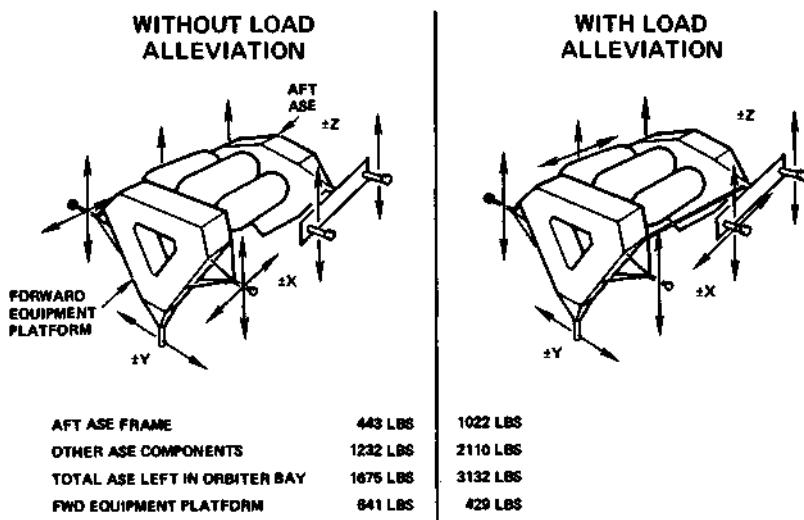


Fig. 3. Retention system changes with and without load alleviation.

The performance gain afforded by the highly conservative TDRS propulsion module design does not come from the 2 to 5 seconds of specific impulse advantage possible from liquid engines over the present IUS motors, but rather from weight savings attributable to nonduplication of many functions that are readily available from the spacecraft payload. The module, using hardline connections required by the IUS stage, derives the following support from the TDRS spacecraft subsystems:

- Power, via battery or solar array.
- Control during coast via the monopropellant RCS system. The normal 1300-pound NH<sub>3</sub> load is supplemented by ~40 pounds for transfer orbit and module disposal utilization. Total TDRS tankage capacity is 1500 pounds.
- Guidance and control information via rate integrating gyro assembly (three-axis orientation and rate signals).

<sup>2</sup>"Inertial Upper Stage User's Guide," Boeing Aerospace Company, D290-10652-1, 1980.

- Fine and coarse solar aspect bearings in north-south plane.
- Complete telemetry, tracking, and command communication links with the NASA STDN stations and STS.

This free support is in contrast to the support afforded by the flying portion of the IUS which supplies the following services to the TDRS during transfer:

- Power, via a battery pack
- Control, via a separate RCS system
- Guidance, via an inertial guidance unit
- Telemetry, tracking, and command services.

The TDRS system being built by TRW for the Space Communications Company (SPACECOM) will be the largest spacecraft in GEO. It spans 75 feet from tip to tip of its solar panel array. Its present weight of around 5000 pounds requires the present-day versions of both the STS and the IUS to operate at top efficiency to place it into orbit with sufficient hydrazine in its monopropellant reaction control system (RCS) tanks to ensure full-term design life in orbit. The TDRS is shown mounted on the propulsion module in Fig. 4, in the view which shows the two 16-foot steerable single-access antennas. There are five other antenna systems, and the rotatable solar panels are shown extending out of the page parallel to the STS' wings.

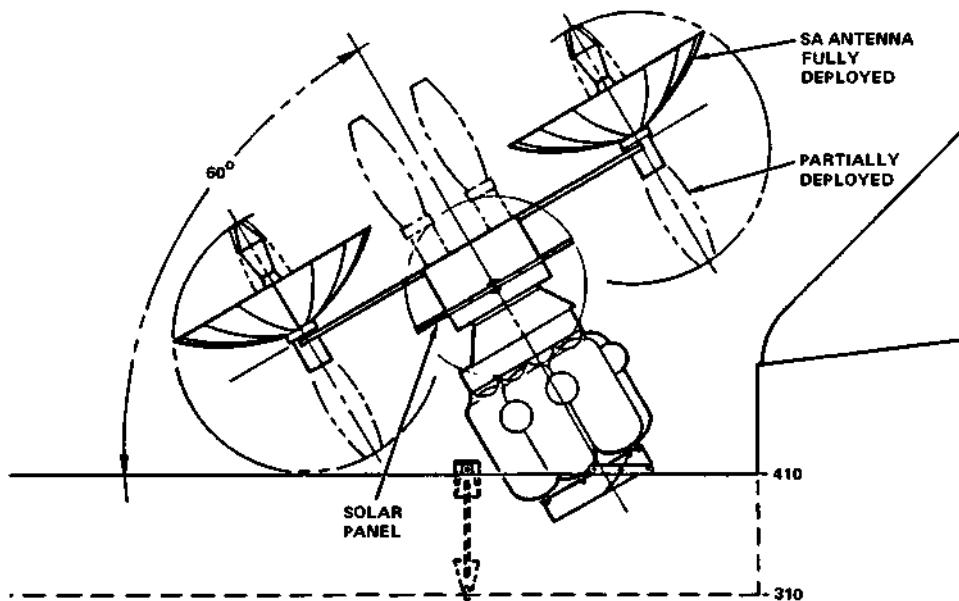


Fig. 4. TDRS deployment from Orbiter bay.

Release of the spacecraft from the STS may be accomplished in several ways. The least controversial and most unimaginative possibilities are to push it away from the STS either sideways from a horizontal position in the cargo bay, or (if it is first erected) along the thrust line. The latter method permits some spacecraft checkout prior to release because the TDRS omni-antenna then can view the earth. On the other hand, even for a spacecraft as large as the TDRS, it is possible to completely deploy its appendages while it is erected (Fig. 4), or on the RMS arm. Problems might arise, however, from contamination of the spacecraft by the STS' RCS system exhaust. Table 2 presents some of the other considerations in this trade-off. It is the author's belief that in order to take full advantage of the versatility of the STS, the deploy-while-attached method ultimately will be accepted in conjunction with low-thrust transfer systems.

TABLE 2 Attached Deployment Considerations

<u>Pro</u>
● Opportunity to fix by command or EVA
● Opportunity to check out key subsystems
● Possibility of restowing and returning
<u>Con</u>
● Safety questions (restowing)
● STS RCS contamination questions
● Ability to refold appendages

#### Selection of Propulsion Subsystem

Three alternative types of propulsion subsystems were considered:

- An all-solid motor system using available IUS or Minuteman-derivative and kick motors
- A hybrid (solid-liquid) propulsion subsystem, as in the Hughes LEASAT
- An all-liquid MMH/NTO pressure-fed bipropellant system using available engines.

A low-thrust all-liquid system was chosen for the following reasons:

- Qualified low-thrust engines are available in short lead times at relatively low cost. These include the Rocketdyne Viking 300 pound-thrust engine; the Marquardt STS R-40 engine, whose nominal 830-pound-thrust range can be varied from approximately 700 to 1200 pounds thrust by varying inlet pressure; and several qualified approximately 100-pound-thrust engines which are available, for example, from Marquardt, MBB, and TRW. The highly efficient Aerojet 6000-pound-thrust STS OMS engine was not considered for several reasons which will become apparent. One important reason is that it costs significantly more than the smaller engines.

- With high-thrust engines/motors in high total-impulse single-burn usage, the requirements on the guidance and control system accuracy are taxing, particularly if initial burn must take place without ground or STS involvement. This has led the IUS to use a self-contained inertial guidance system with three-axis control, and the much less complex LEASAT to use spin stabilization to reduce first stage dispersion. The TDRS was not designed to withstand high roll rates, thus ruling out this option immediately. Although the TDRS does contain a three-axis integrating rate gyro assembly to assist the attitude control system on orbit and while station-keeping, the output of this system, via the telemetry hardline into the IUS, is at a maximum sample rate of 4 Hz. This is not believed fast enough to provide necessary guidance and stabilization data to a propulsion module guidance and control computer during a short, high-thrust, impulsive burn. An equally important reason for not considering solids for this application was the desire, for economy and simplicity, to use the TDRS RCS. It would be difficult for the arrays of 1-pound monopropellant hydrazine jets to maintain roll control during burning of a high-thrust two-axis gimballed solid motor.
- It should be noted, however, that within the stated ground rules the decision could not be made on the basis of propulsion system performance alone. Solid motor designers have managed to remain competitive with the 290- to 309-second specific impulses claimed by the low-thrust liquids with various nozzle area ratios, and have maintained a high degree of motor case structural efficiency.
- However, for the TDRS transfer and similar applications, particularly where the spacecraft is three-axis controlled on orbit, a strong case for the low-thrust liquid propulsion can be made:
  - a) The spacecraft's appendages can be deployed either while still attached to the STS or shortly after it is released from the cargo bay. The TDRS requires around 1400 watts to maintain its subsystems in the standby mode. The IUS supplies this power from its batteries during transfer. Low-thrust propulsion permits elimination of a separate battery system on the propulsion module, since it and the TDRS can use the regular on-orbit spacecraft power system. The TDRS solar panels in their folded position are not designed to provide power.
  - b) Early deployment of the appendages also permits the unobstructed use of the spacecraft's on-orbit omni antenna, which can then be used for telemetry, tracking, and command during transfer. Deployment also exposes the spacecraft's RCS jets, thus permitting three-axis control during transfer coast periods.
  - c) Low-thrust transfer implies repeated perigee burns until the desired apogee is attained, followed by apogee burn(s) for plane change and circularization. This is a leisurely process which does not place any greater requirements on the guidance and control system than would be encountered on orbit. Trajectory errors are easily correctable by subsequent burns, without leading to large increases in transfer  $\Delta V$  requirements.
  - d) If deployment takes place while the spacecraft is still attached to the STS, perhaps the greatest potential advantage is the inherent ability to test out many subsystems, and by work-arounds or EVA (if necessary) to effect fixes or repairs while still possible. Full or partial recovery of the spacecraft can likewise be considered.

There are also other factors that must be considered if low-thrust propulsion is selected:

- In the case of the TDRS, the solar cell area devoted to battery charging was designed for on-orbit GEO operation and therefore is limited in charging rate. This leads the transfer trajectory design to place apogee in the general direction of the sun (perigee therefore taking place during eclipse) and to obtain long time exposure to the sun prior to the next eclipse period as quickly as possible. This need is particularly true for TDRS because of its large standby power requirement. The implication on design, therefore, is to raise apogee as quickly as possible. This can be done by increasing either thrust or burn arc per perigee. The obvious choice is to operate at as high a thrust level as the deployed structure was designed to withstand (0.095 g maximum for the TDRS). This minimizes the burn arc, a desirable effect because long burn arcs result in propellant inefficiencies when compared to impulsive perigee burns. However, since initial orbits are nearly circular, long burn arcs here are more tolerable than they would be for later orbits when the perigee velocity gradients along the flight path are steeper, since energy is most efficiently added at the peak velocity point (i.e., true perigee). Thus, a satisfactory compromise between battery depth of discharge, maximum tolerable thrust levels, and total impulse added during early perigee burns must be found.
- Multiperigee burns also imply that the spacecraft's deployed solar panels will spend considerably more time traversing the Van Allen belts than the single passage possible with a two-impulsive-burn transfer. For example, studies have indicated that for an 11-perigee-burn case, long-term degradation of about 2+ percent is suffered for conventionally designed panels.
- As discussed above, finite perigee burns, and to a much lesser extent<sup>3</sup> finite apogee burns, lead to inefficiencies which increase the  $\Delta V$  requirements above the approximately 14,000 ft/sec ideal GEO  $\Delta V$  required for a Kennedy launch. These losses, which will be detailed for the TDRS case later, can be minimized by keeping burn arcs small and thrust lines along or near the imaginary orbit tangent line. Shorter burn arcs, with low thrusting engines, imply a greater number of perigee burns, with the aforementioned depth of discharge and Van Allen belt traverse degradation tradeoffs open for compromise.

For the case of the TDRS study, the eventual propulsion module dry weight design, combined with the 0.095 g structural design requirement (caused at IUS-TDRS separation by spring-applied force) lead to a maximum permissible thrust level of about 780 pounds. Feasible choices for engines, then, are:

- One Marquardt R-40
- One or two Rocketdyne Viking(s)
- Several 100-pound-thrust engines.

---

<sup>3</sup>Since orbital velocities around GEO apogee are slow, with little velocity gradient along the flight path.

Factors in ultimate engine selection, assuming comparable reliability for all choices, include: thrust/burn time considerations, system reliability/complexity, status of qualification of gimbals and actuators, system acquisition costs, propellant tank pressure (i.e., engine inlet pressure), and engine-out operation. Mission analysis indicated that a 300-pound thrust level with burn times yielding acceptable  $\Delta V$  losses resulted in larger than desired battery depths of discharge and Van Allen belt exposures (e.g., there were around 30 perigee burns). The alternative of using several 100-pound engines was discarded for economic and complexity factors.

The Viking engine and its gimbal/actuator system are presently qualified, while only the Marquardt engine per se is qualified. When both systems are brought to comparable qualification status, there is no overpowering economic advantage between selecting two Viking engines and one Marquardt engine. The former system was tentatively selected because of its qualification status, the thought that the possibility of successful one-engine-out operation results in a system with a higher overall success probability, and the ability of two gimballed engines to supply three-axis control during burning. (It should be noted that in a one engine-out situation, the TDRS RCS roll control jets can adequately handle roll torques caused by small thrust vectors from a 300-pound-thrust engine; however, they would have more difficulty handling roll control with a misaligned 700-pound-thrust engine.) From a performance standpoint, at equal nozzle expansion ratios, both engines have comparable specific impulses. The Marquardt engine has an inherently lower thrust chamber inlet pressure (by approximately 40 psi), which can translate into lower propellant tankage weights.

Two other propulsion subsystem tradeoffs are of interest:

- Earlier studies<sup>4</sup> indicated that significant performance gains (400 to 500 pounds of additional payload to GEO) might be accrued by staging tankage one or two times during transfer to GEO. This possibility strongly influenced the design of the TDRS propulsion module, even though it was decided to forego the possible performance gains for the sake of simplicity and to eliminate jettisoned tankage debris problems. However, if the ultimate in performance is required, then staging of all-liquid systems must be considered, as it is for the jettisoning of empty solid motor cases.
- The choice of tankage materials is clearly an important one, involving tradeoffs in weight, cost, and delivery schedule. It soon became apparent that in the present economic environment titanium tankage would be very expensive, with very long lead times required, and that 6061 aluminum tanks, while requiring relatively short lead times (12 to 14 months), would be very heavy. The compromise was to select 2021 aluminum for the propellant tanks. This selection created the critical path in the development schedule, due to the 6-month delay in delivery of raw stock from England. Pressurant bottle tankage presented less of a scheduling problem. Kevlar-wrapped Inconel bottles were selected, as they are light and more readily available than titanium.

#### Mission Analysis

For multiperigee burn ascents to GEO altitude, the main performance consideration is to avoid excessive  $\Delta V$  losses by compromises between burn arc and thrust level, and between number of burns and time to achieve apogee. These compromises are

---

<sup>4</sup>Macklis, Sackheim, and Vogt.

generally constrained both by availability of ground stations, if commands and telemetry are necessary to the transfer process, and by the above-mentioned specific TDRS problems of Van Allen degradation and spacecraft battery depth of discharge. The compromise considerations are well covered by other authors,<sup>5,6</sup> whose papers provide the following generalizations (Fig. 5 and Fig. 6):

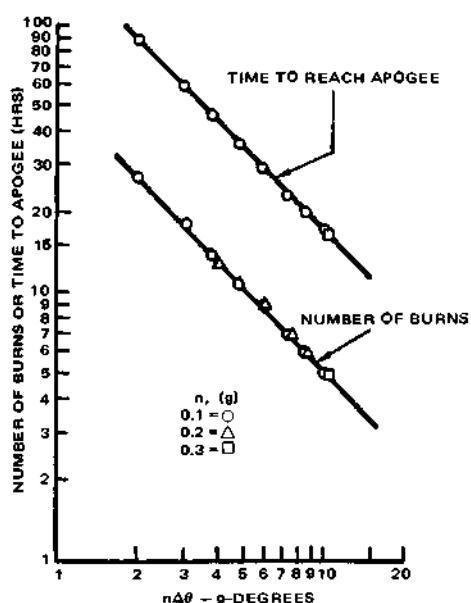


Fig. 5. Time and perigee burns required to attain GEO apogee ( $n$  = maximum acceleration).

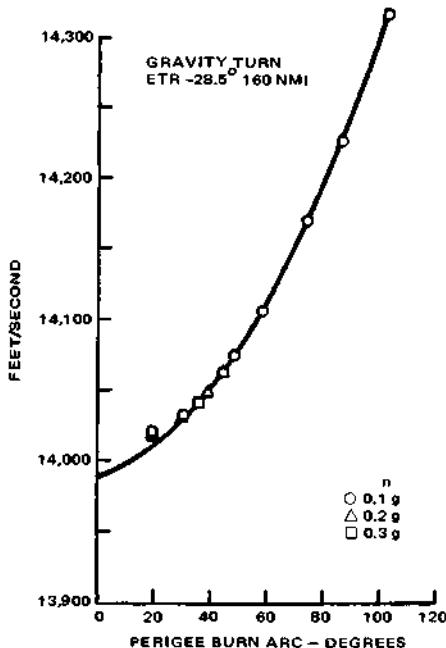


Fig. 6. Velocity loss,  $\Delta V$ , in attaining GEO apogee (no apogee loss included).

- For burn arcs up to at least 120 degrees, absolute  $\Delta V$  losses in attaining GEO are less than  $\approx 300$  ft/sec out of a total  $\Delta V \approx 14,000$  ft/sec.
- Apogee burn  $\Delta V$  losses are trivial — on the order of 10 to 20 ft/sec.
- The losses are relatively insensitive to thrust levels, at least in the range of 0.1 to 0.3 g maximum acceleration.
- Transit time to apogee is roughly a linear function of thrust level at similar  $I_{sp}$ 's for the same burn arcs.

The TDRS study required two special  $\Delta V$  loss investigations. The first of these was needed as an aid to developing the proposed guidance and control system. The information sought was whether large losses would be suffered by perigee burning

<sup>5</sup>Macklis, Sackheim, and Vogt.

<sup>6</sup>M. H. Kaplan and W. Yang, "Finite Burn Effects in Ascent Stage Performance," AAS 81-153, AAS/AIIS Astrodynamics Conference, August 1981.

with the thrust line maintained normal to the nadir as compared to burning tangent to the trajectory. The second study was instigated by the desire to raise apogee as quickly as possible. This could be done by continuous burning normal to the nadir line (i.e., spiralling out), but at what penalty?

Table 3 indicates that there is little penalty for burning normal to the nadir line, and other studies for finite perigee burns corroborated this. The absolute penalty to reach high apogees quickly by continuous burning was severe, and dictated going instead to higher thrust for the TDRS application.

TABLE 3 Finite Burn Penalty (ft/sec) to Achieve Apogee Altitude (Thrust = 300 lbs)

Number of Burns	$\Delta V_{Req}$	Finite Burn Penalty
Impulsive	$\approx 8000$	0
Continuous-spiral out to apogee with thrust line tangent to velocity vector		3557
Continuous-spiral out with thrust line $\perp$ to nadir		3783
Spiral out to 1100 km, then finite perigee burns		665
Spiral out to 2400 km, then finite perigee burns		1393

Table 4 shows the results of these higher-thrust studies. Case 2 was selected as the baseline compromise. With this selection, maximum battery depth of discharge is less than 75 percent, and begins to recover starting at emergence into sunlight following the 7th perigee burn. It should be noted that the velocity losses of Table 4 are somewhat higher than comparable ones derived from Fig. 5 and Fig. 6. This is because the Table 4 trajectories achieved maximum accelerations of only  $\approx 0.066$  (compared to the minimum of 0.1 g of Fig. 5 and Fig. 6). This required uniformly longer burn arcs with consequent increased losses. Moreover, the need to raise apogee quickly in the beginning of ascent dictated burn arcs much larger than optimal.

The propulsion module is disposed of by leaving it in a retrograde equatorial circular orbit 200 km above GEO, consistent with recent recommendations.<sup>7</sup> Similar to the IUS terminal maneuver, the propulsion module places the TDRS spacecraft in its on-orbit attitude, whereupon it separates. Using its normal stationkeeping ability, the TDRS flies back to its ultimate GEO station. This latter maneuver requires about 20 pounds of monopropellant, which is charged against propulsion module weight (Table 1).

<sup>7</sup>M. Hechler and J. C. Van Der Ha, "The Probability of Collisions on the Geostationary Ring," ESA Journal, 1980, Volume 4, p. 277.

TABLE 4 Perigee Burn Studies

Based on 600-pound thrust,  $I_{sp}$  of 296, initial altitude of 298 km, and two timed burns followed by a sequence of variable arc burns.

	<u>Case 1</u>	<u>Case 2</u>	<u>Case 3</u>
<b>Timed burns</b>			
Duration	20 min	30 min	45 min
Initial arc length	80 deg	119 deg	180 deg
<b>Altitude burns</b>			
Number	17	9	5
Final arc length	27 deg	48 deg	74 deg
Final perigee	430 km	667 km	1260 km
<b>Perigee burn totals</b>			
Number	19	11	7
Transfer time to apogee*	68 hours	40 hours	26 hours
Total $\Delta V$ loss (perigee + apogee burns)**	105 ft/sec	225 ft/sec	595 ft/sec

\*Total engine burn time ~4.5 hours

\*\*Includes 15 ft/sec  $\Delta V$  loss for apogee burns; circularization and plane change required an additional 26 to 35 hours, corresponding to 2 to 3 burns

#### Guidance and Control

With three-axis gyro position and rate data available from the TDRS, two schemes for guidance signals in pitch (i.e., motions in the orbital plane) evolved. In both cases, yaw (i.e., maintenance of inclination on the orbit plane) is controlled by error signals from the TDRS gyro assembly, whose drifts are capable of correction using TDRS solar aspect data. The two schemes are:

- Once ground contact can be assured at least once per orbit, a commanded value of pitch rate could be compared with the gyro value to produce an error signal. The pitch rate commanded during coast would be such as to assure a nearly 300-degree turn between burnout and start of next burn. A different pitch rate would be commanded during burning to maintain the thrust line tangent to the trajectory. Such a scheme requires no new instrumentation, but does use much ground station activity. (Early orbits would be preprogrammed until an apogee altitude which assured a ground station contact is attained.) Roll, like yaw, is maintained by gyro assembly error signals.
- Alternately, a simple guidance system can be designed which maintains transfer vehicle attitude in pitch and roll normal to the nadir line during both coast and burning. This attitude has the additional attribute of always maintaining acceptable gain between the TDRS omni antenna and earth stations. Since the TDRS horizon scanners operate only at or near GEO altitude, it was necessary to add a set of conical horizon scanners, with capabilities from LEO to GEO, to the propulsion module. This assembly, with its associated avionics, was the only nonpropulsion support instrumentation necessary to add for the module to complete its mission.

Nadir position from the scanners is available at a 4-Hz sampling rate, as is data from the TDRS gyro assembly on the TDRS telemetry stream. This data is decoded by the propulsion module avionics. These two data sources supply:

- a) Rate data for damping
- b) Position data to generate commands (hardlined back into the TDRS by the discrete channel used by the IUS) to drive the TDRS solar arrays
- c) Yaw position data to maintain the orbital plane inclination
- d) Orientation data for capture following separation or loss of nadir lock.

Engine startup and shutdown (for either scheme) can be commanded, and preprogrammed for several consecutive burns without essential loss of accuracy. Another (proprietary) engine control method was devised which requires no additional instrumentation, and no ground involvement until apogee altitude is attained. The attitude control system block diagram for the relatively unattended system is shown in Fig. 7 and the horizon scanner's view is in Fig. 8. Scanners 1 and 2 are prime, while Scan 3 is the redundant back-up scanner.

The newly developed avionics system for the propulsion module must accomplish the following functions, in addition to normal mode guidance and control:<sup>8</sup>

- Cause uncaging of gyro reference prior to separation via ground or STS command.
- Maintain ground or STS commanded or internally programmed inertial altitude.
- On ground, STS, or internal command, conduct a preprogrammed search mode to establish nadir and yaw angle. Internal program to be initiated when earth lock is lost.
- By internal program, be capable of switching from inertial mode to horizon scanner guidance mode upon capture of earth by horizon scanners.
- Detect and reconfigure for one-engine-out operation.
- Disable RCS in favor of gimbal control during burn. Baseline version has three-axis gimbal control. If one-engine or one-engine-out (baseline version), the system reconfigures to permit RCS control in roll.
- By program or ground command, be capable of gain change for coast phase operations both at perigee and apogee.
- Turn engine(s) on and off (internal and ground command).
- Drive solar panels (internal and by ground and STS command).

<sup>8</sup>It is estimated that these functions can be accomplished with no more than 20 pounds of monopropellant hydrazine added to the TDRS tanks.

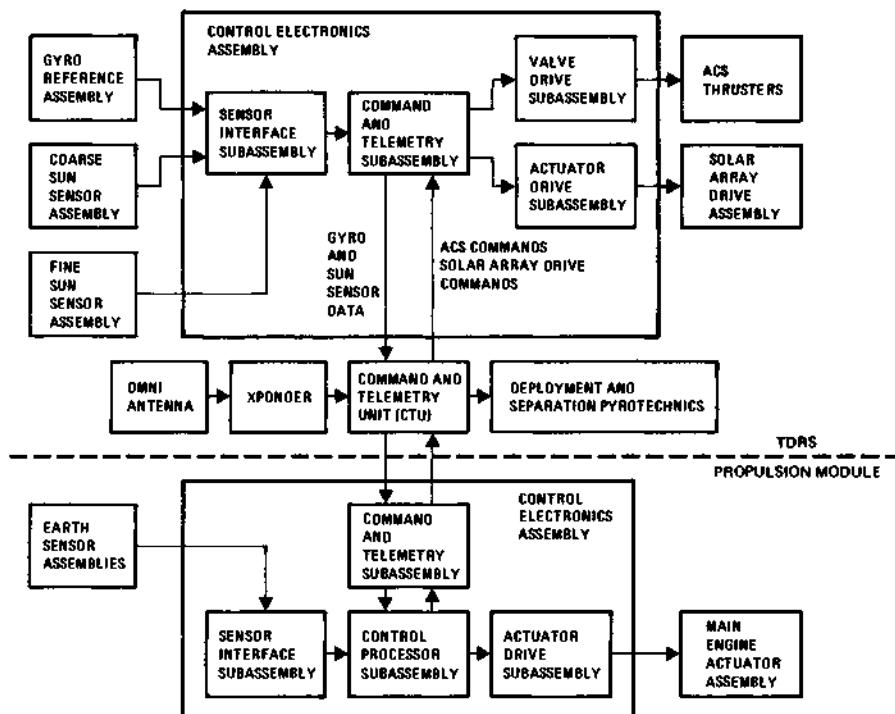


Fig. 7. Control system block diagram.

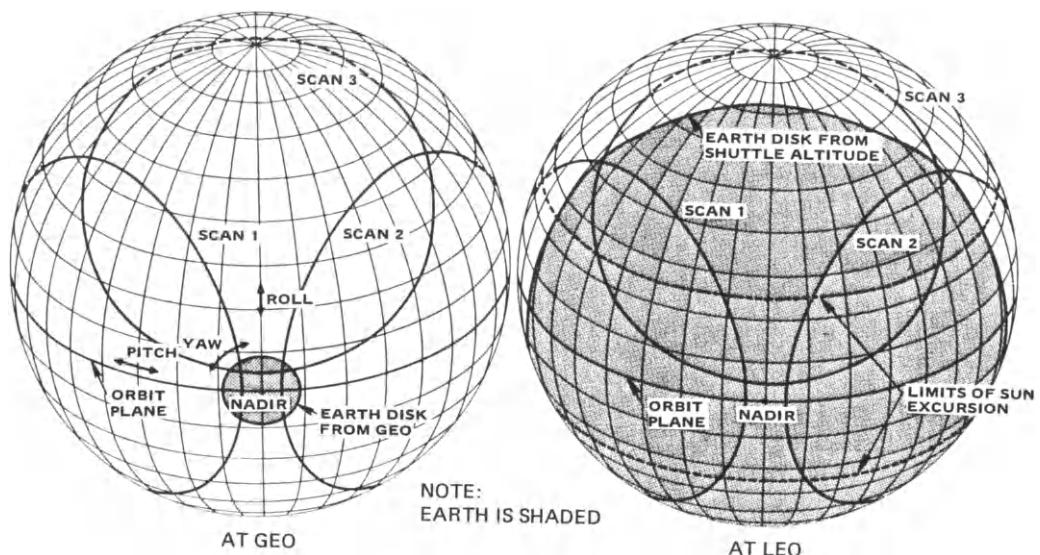


Fig. 8. Horizon scanner view from low Earth orbit (LEO) and GEO.

- Be capable of logic to permit back-up powered flight scheme to be used upon ground command. In particular, accept instructions to maintain commanded pitch rates during coast and/or burn.
- By ground command (and/or with internal logic) correct gyro drift using information from the TDRS solar sensors.

### Telemetry, Tracking, and Command

The key to adapting a new propulsion module without TDRS modification was the suitable use of the many duplicative services that the IUS provides to the TDRS. For example:

- The IUS provides emergency power, via its batteries, to the TDRS in the STS cargo bay as well as standby power during transit. These hardlines can be used to have the TDRS provide power to the propulsion module.
- The IUS provides a command uplink and the ability to send discrete commands to the TDRS. These lines can be used by the propulsion module to command TDRS function. A spare command channel can be used.
- The IUS provides TDRS telemetry transmission service. Thus, a decoder on the propulsion module can extract needed gyro and solar aspect data.
- Commands to the TDRS are echoed on the telemetry stream and can be decoded by the propulsion module. Fig. 9 provides the command block diagram.

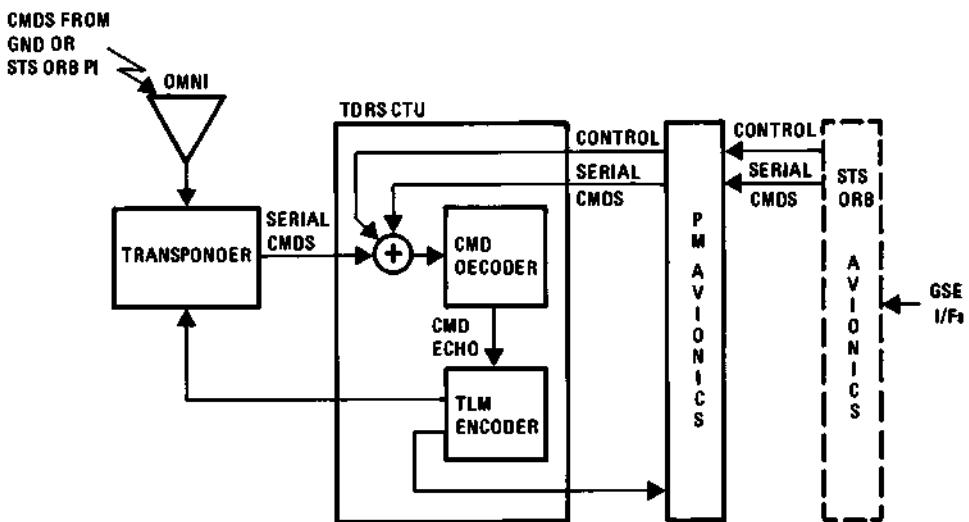


Fig. 9. Command system capability.

Once the TDRS is separated from the STS and deployed, its omni antenna can communicate with both the nearby STS and the ground. Once the horizon scan guidance mode takes over, the omni pattern always has acceptable gain toward nadir. Thus, requirements on the propulsion module avionics subsystem are:

#### Commanding

- Provide propulsion module commands to TDRS
- Accept ground commands via TDRS
- Provide commands for deployment sequence (STS and/or ground initiated)
- Provide TDRS solar array drive commands based on TDRS gyro data

#### Telemetry

- Provide telemetry data to TDRS for transmission to ground
- Provide propulsion module/TDRS telemetry data to STS

#### Power

- Accept STS power in stowed mode
- Accept TDRS power in transfer orbit

#### COST/SCHEDULE

Considerable effort was expended in the study to determine development and production costs of the TDRS propulsion module. Because of its basic simplicity and reliance on qualified components and/or tried and true methodology, the unit costs of the TRW propulsion module not surprisingly turned out to be less than half of the present IUS cost, and nonrecurring RDT&E costs are modest. On an accelerated basis, with tankage as the critical path item, a program of 34 months to first launch appears well within capability. At greater expense and risk a shorter program can be accomplished.

#### CONCLUSIONS

As a result of the study, the following general conclusions can be reported:

- An all-liquid propulsion module could deliver a 5200-pound TDRS spacecraft to GEO at the present-day STS ground launch cargo weight limit of 41,000 pounds with some margin.
- TRW IRAD studies have established basic concept feasibility and resolved all first-order configuration trades leading to a successful all-bipropellant propulsion module.
- No new component technology is required. New tanks, structure, and avionics units are well within the current state of the art.
- Use of an integrated propulsion module can be accomplished without a single identified hardware change to the TDRS space segment.

- The major unresolved risk in development of a new propulsion module lies in the outcome of detailed design studies on the need for load alleviation. However, the additional costs, performance loss, and schedule stretch-out are understood and bounded.

#### ACKNOWLEDGEMENTS

The author wishes to acknowledge the work of the major study participants: R. A. Kaminskas for design, J. R. Wertz for his guidance and control system inventions, C. R. Roberts for detective work in finding dual usage for key circuitry, and D. E. Fritz for assistance in propulsion. Reviewing the manuscript, J. A. Love, A. G. Wells, and J. T. Johnson provided many helpful suggestions.

## GEOSYNCHRONOUS SHUTTLE FOR DIRECT ASCENT FROM EARTH AND DIRECT RETURN TO EARTH

R. Salkeld\* and R. S. Skulsky\*\*

\*System Development Corporation, Santa Monica, California, USA

\*\*Northrop Aircraft Corporation, Hawthorne, California, USA

### Abstract

A concept for a geosynchronous manned or unmanned shuttle for direct return to horizontal Earth landing, is summarized. Previous work is then extended to consider two additional launch modes: air-launch using thrust augmented aircraft pull-up, and ground launch using main propulsion elements derived from the NASA first-generation shuttle program.

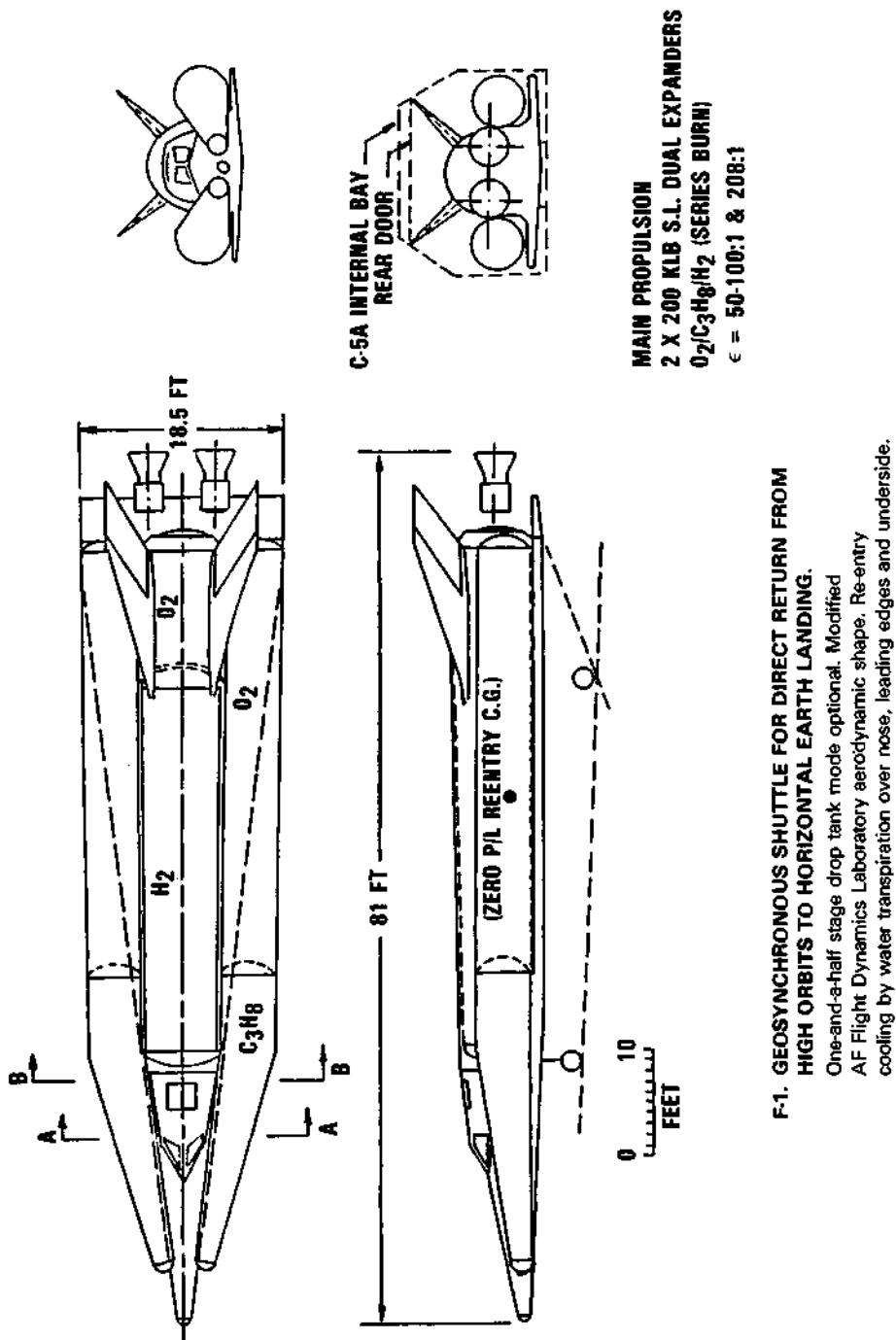
The air-launch pull-up maneuver is found to offer modest orbital performance benefits, and may or may not be justifiable on a performance basis.

The Shuttle-derived ground launch provides capability for direct ascent from Earth to geosynchronous orbit, independent of any operations in low-Earth orbit. Such capability could have important implications and is recommended for further study.

### Background

Technical feasibility of a direct-return geosynchronous shuttle has been explored with favorable results.<sup>1,2</sup> One concept is shown in F-1. As described previously, the vehicle has two O<sub>2</sub>/propane/H<sub>2</sub> dual-expander rocket engines of 200,000 lb. thrust each, a two-man crew station, a 5-by-5-by-10 ft. utility bay, and an all metal transpiration-cooled thermal shield using water as the coolant. It may also operate as an unmanned vehicle. Its two aluminum O<sub>2</sub>/propane tanks can be jettisoned upon exhaustion of the propane or retained during the hydrogen burn into orbit. The vehicle weighs about 20,000 lb. dry and 225,000 lb. fully loaded, including external tanks.

On return the spacecraft pulls out for constant-altitude flight at 170,000 ft. for about 20 min. until it has decelerated to the low-Earth-orbit (LEO) velocity of 26,000 ft./sec., then executes a lifting maneuvering reentry from that



condition, as indicated in F-2. Well-known classical analysis gives maximum temperatures for this reentry as shown in F-3. These are considered conservative and could be reduced by refining the reentry flight strategy. The thermal shield comprises an arrangement of diffusion-bonded individual aluminum platelets with chemically etched internal passages for water distribution, as depicted in F-4 and F-5. Estimated weights for this system are 860 lb. for hardware and 4,000 lb. for water (3,000 lb. for LEO reentry).

Previous investigations of launch modes for this vehicle concept were limited to independent ground-launch and horizontal air-launch from a C-5 class aircraft. These investigations are here extended to include thrust-augmented air-launch employing aircraft pull-up, and ground-launch using propulsion hardware developed in connection with the NASA first-generation shuttle program.

#### Thrust-Augmented Aircraft Pull-up

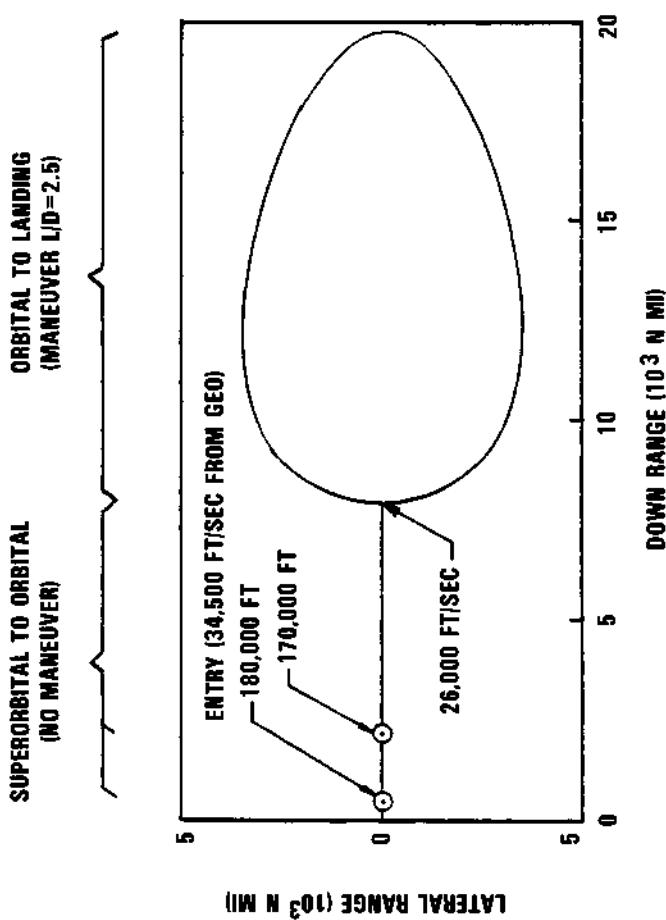
It is postulated that a C-5 class aircraft could be modified to accommodate a liquid rocket propulsion system for thrust augmentation, mounted within the fuselage. If such a system, say using O<sub>2</sub>/kerosene propellants, delivers 1 million lb. of thrust, then the aircraft could be placed in a flight path ascending at an inclination of 50 deg., and would be aerodynamically controllable to altitudes approaching 70,000 ft. A pull-up maneuver of this type is depicted in F-6, wherein it is postulated that the spacecraft is gravity-launched through the open rear cargo door of the aircraft.

Performance calculations show that such a maneuver can increase the spacecraft orbital payload capability to LEO by about 2,000 lb., or 33 percent for a polar launch. This is a smaller increase than we had anticipated, and raises the question to what degree the complication of pull-up modification for such a large aircraft can be justified on a performance basis. A more modest capability for pull-up at lower altitudes might be justified as facilitation of separation of the two vehicles.

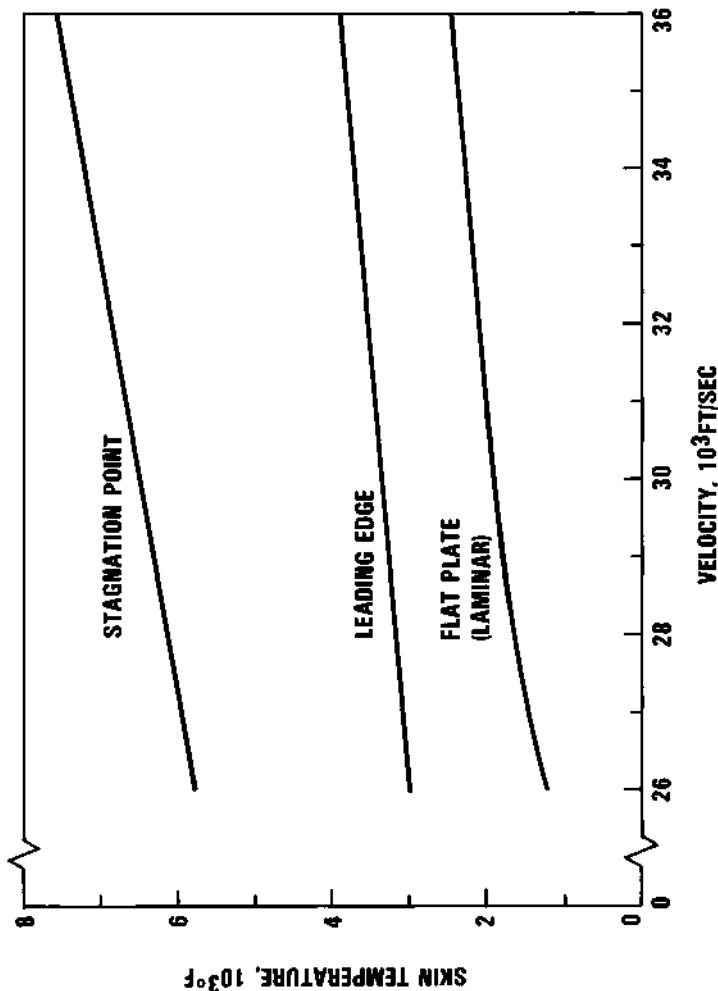
Another use of aircraft pull-up for spacecraft launch involves mounting the spacecraft atop the airplane and flying the aircraft through a convex-upward parabolic trajectory to separate the spacecraft by centrifugal force. Spacecraft performance gains using this approach remain to be assessed, but it appears questionable whether the above pull-up results could be exceeded or even matched in this manner.

#### Direct Ascent to Geosynchronous Orbit

Approximately 95 percent of the gross liftoff weight of the NASA first-generation shuttle comprises propulsion hardware and propellants which can be modularized to form derivative launch vehicles. An early suggestion was to replace the orbiter airframe (except for the three main engines and a recovery pod for them) with a propulsive upper stage to provide a heavy lift capability.<sup>3</sup> The fact that the orbiter weight replaced is very close to the gross weight of the vehicle shown in F-1, in turn suggests that the geosynchronous shuttle shown here could be launched by the same NASA shuttle derivative. The geosynchronous shuttle performance of such an arrangement is attractive, as shown in F-7. Performance is sufficient to permit the number of solid rocket booster segments to be reduced from four to one per booster, with attendant reduction in gross weight and probably in cost per flight.



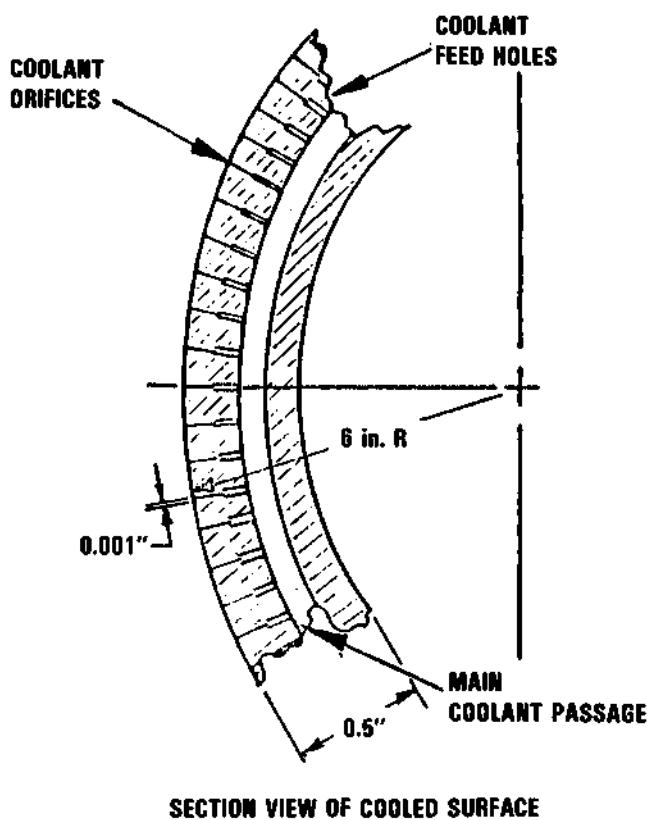
F-2. DECELERATION DISTANCE AND GLIDE FOOTPRINT.



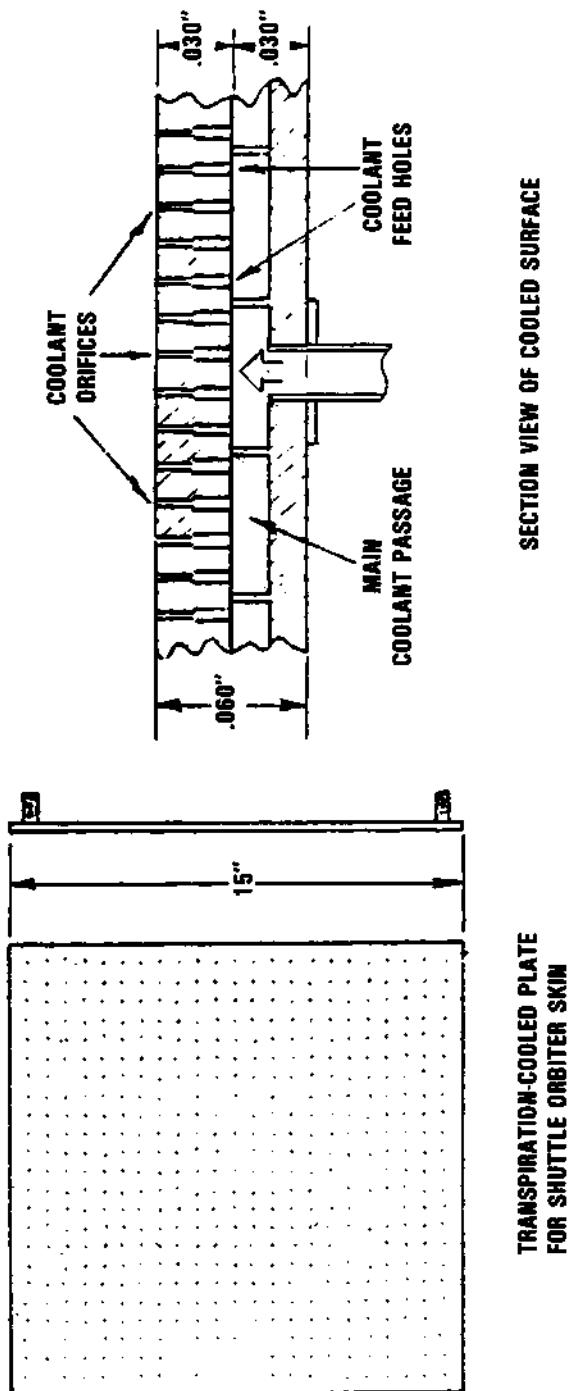
F-3. MAXIMUM SKIN TEMPERATURES FOR  
DECELERATION TO 26,000 FT/SEC

Pull-out altitude: 170,000 FT. Nosecap radius: 0.5 FT.

Leading edge radius: 0.5 FT. Flat plate: 30 FT. AFT of nosecap.

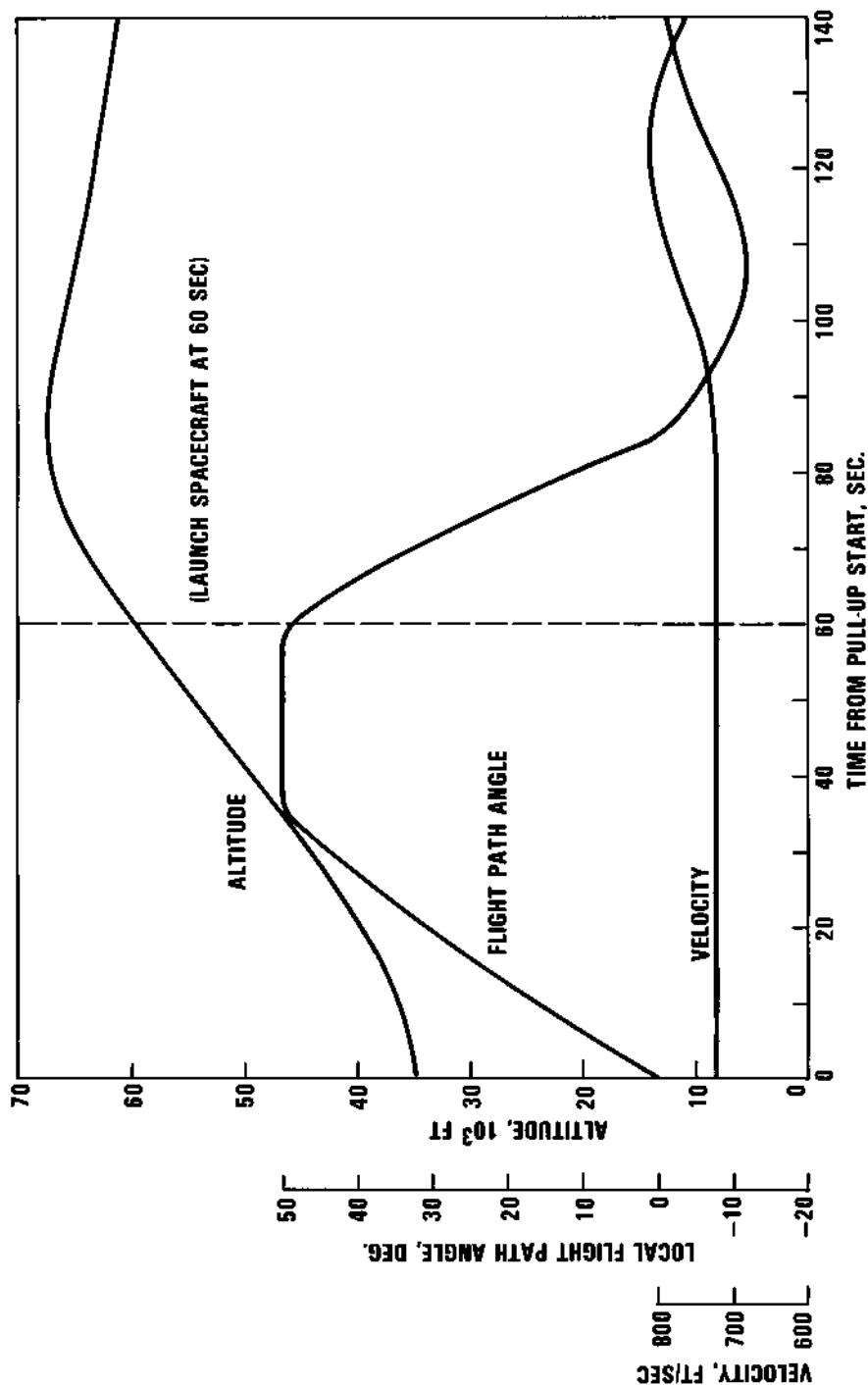


F-4. NOSE CAP CONFIGURATION.



SECTION VIEW OF COOLED SURFACE  
TRANSPIRATION-COOLED PLATE  
FOR SHUTTLE ORBITER SKIN

F5. ACTIVELY COOLED PLATES.



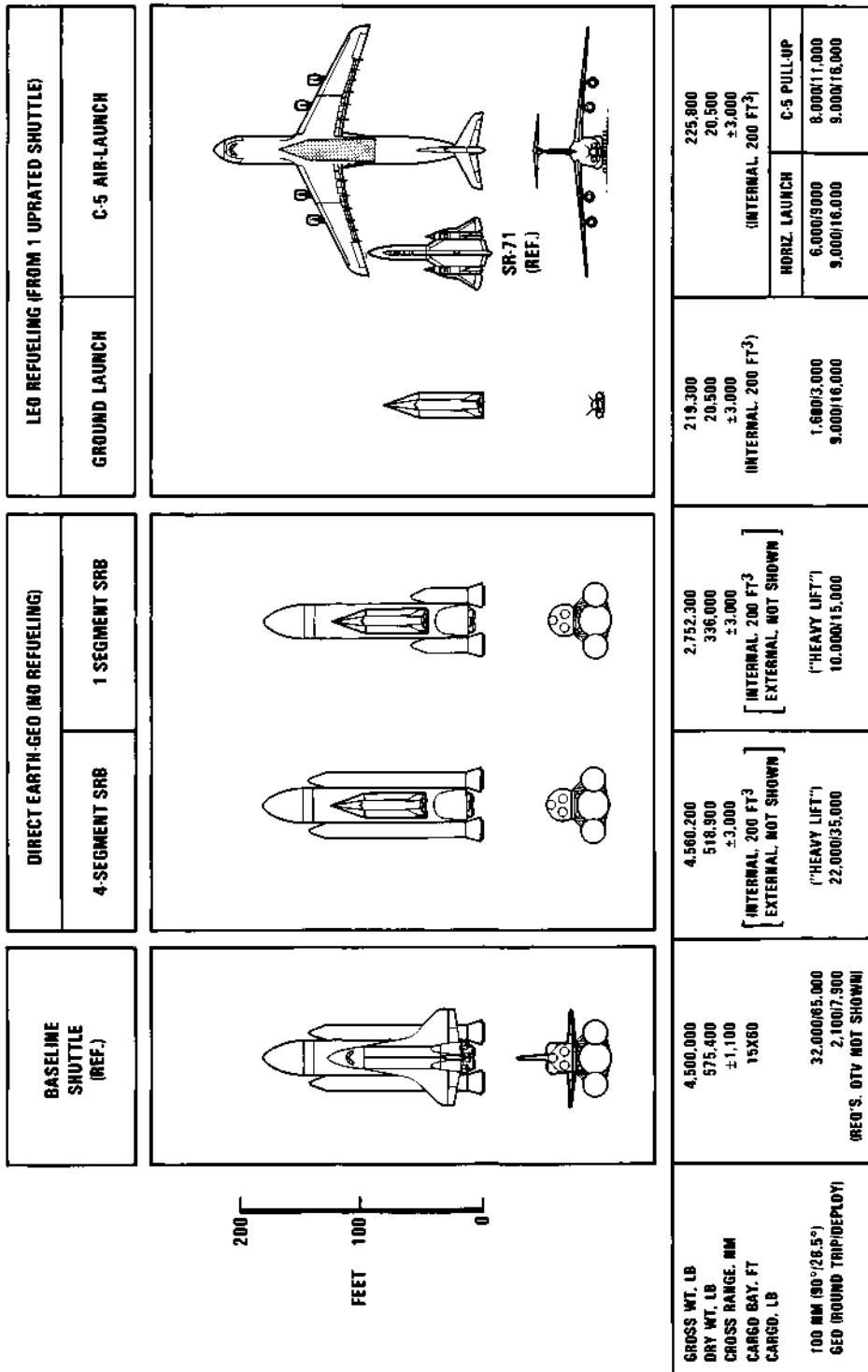
F-6. THRUST-AUGMENTED PULL-UP FOR C-5 CLASS AIRCRAFT

Pocket augmented thrust: 1 million lb. Rocket propellant required: 175,000 LB O<sub>2</sub>/RP-1.

## F-7. POSSIBLE LAUNCH MODES FOR GEOSYNCHRONOUS SHUTTLE

External tanks retained to completion of main propulsion.  
Drop-tank option would increase payloads.

### Geosynchronous Shuttle



This approach thus provides a capability for direct, round-trip, shuttle flights between Earth and geosynchronous orbit independent of any refueling, checkout or other operations in LEO.

Such capability would have fundamental implications for a diversity of future space activities and is commended to all concerned for careful further investigation.

#### Acknowledgement

The authors thank Dr. Malcolm G. Wolfe of The Aerospace Corporation for informal discussions which led to the concept of direct ascent for a geosynchronous shuttle.

#### References

<sup>1</sup>Beichel, R., Salkeld, R., and Skulsky, R., "New Concept for Far-Orbit Transportation," IAF Paper IAF-70-F-243, presented at XXXI IAF Congress, Tokyo, Japan, Sept. 1980.

<sup>2</sup>Salkeld, R., Beichel, R., and Skulsky, R., "A Reusable Space Vehicle for Direct Descent from High Orbits," Astronautics and Aeronautics, April 1981, p. 46.  
(Note: In this reference, T-1 and F-1 erroneously depicted the wrong vehicle. The correct vehicle is depicted in F-1 of the present paper.)

<sup>3</sup>"Use of Orbiter Aft Fuselage as a Heavy Lift Launch Vehicle Propulsion Module," Rockwell International, SSV77-8, Feb. 1977.

## THE NASA TRACKING AND DATA RELAY SATELLITE SYSTEM AND ITS IMPACT ON SPACECRAFT SUPPORT IN THE SPACE TRANSPORTATION SYSTEM ERA

R. E. Smylie and R. O. Aller

*National Aeronautics and Space Administration, Washington, D.C., USA*

### ABSTRACT

This paper discusses the space tracking, data acquisition, and communication network systems and capabilities available to NASA. Emphasis is on the Tracking and Data Relay Satellite System (TDRSS) which is currently under development and scheduled to come into operation in 1983. TDRSS will be the prime support system for communication with the Space Shuttle, Spacelab, and the automated spacecraft to be launched into earth orbit by the Space Transportation System. The NASA Spaceflight Tracking and Data Network (STDN) currently used for support of earth orbital spacecraft is described, and the plans for changes in this network and consolidation of its 26 meter antennas into an expanded Deep Space Network, after TDRSS is operational, are discussed.

The NASA tracking and data systems facilities are frequently employed to support international cooperative space missions, and are also available on a cost-reimbursable basis for support of the space missions of other countries. These practices will continue with TDRSS after it is operational.

### KEY WORDS

Space communication; tracking networks; telemetry acquisition; spacecraft control; satellite relay.

### INTRODUCTION

The National Aeronautics and Space Administration (NASA) operates two worldwide tracking networks which provide tracking, communications, data acquisition, command, and control services for its many science and applications automated spacecraft missions and for the new manned Space Shuttle and other elements of the Space Transportation System (STS).

One of these networks, the Spaceflight Tracking and Data Network (STDN), provides these services for spacecraft in earth orbit, including orbits out to the vicinity of the moon and beyond, such as the on-going International Sun-Earth Explorer (ISEE). The other network, the Deep Space Network (DSN), handles the planetary

exploration spacecraft such as Voyager, requiring a tracking and communication capability to the outer planets - Jupiter, Saturn, and beyond.

Currently under development and scheduled to begin service for NASA in late 1983 is the Tracking and Data Relay Satellite System (TDRSS). This system will consist of two or more specialized communications relay satellites in geosynchronous orbit which will relay data, commands, and voice communication to and from mission spacecraft in low earth orbit and the control centers. TDRSS services will be provided to NASA under a 10-year contract with a communications firm, Space Communications Company (SPACECOM).

TDRSS will take over most of the workload of the STDN in supporting earth orbital spacecraft. NASA is therefore planning to close most of the present STDN ground stations and to consolidate three remaining stations into a single network with the DSN, after TDRSS becomes operational. This consolidated network will then have the facilities and the capabilities to provide tracking, data acquisition, and communications support to all NASA spacecraft which cannot be served by TDRSS; these would be spacecraft in synchronous orbit and high-altitude elliptical orbits as well as deep space and planetary exploration spacecraft.

A world-wide network of terrestrial circuits and communication satellite circuits, leased by NASA from both national and international communications carriers, interconnects all of the tracking stations and the mission control centers for NASA spacecraft. Mission control facilities are located at the Goddard Space Flight Center in Greenbelt, Maryland; the Johnson Space Center in Houston, Texas; and Jet Propulsion Laboratory in Pasadena, Cal.

In this paper, I will expand on these functions and activities and on the principal thrusts of their development for the future. I will provide some background discussion on the types of NASA spaceflight mission requirements which have become the principal drivers for tracking, data acquisition, and space communications capabilities, both in our current network systems and in the upcoming TDRSS. I will indicate the trends we see for these requirements in the future and will discuss the TDRSS system and capabilities which will serve to meet the needs of the Space Transportation System and the spacecraft it carries into orbit, through the remainder of this decade.

#### BACKGROUND

The capabilities and characteristics of the NASA tracking, data acquisition, and communications networks have evolved over many years to meet composite requirements of the science, applications, and manned spacecraft missions which they must support. In recent years, the requirements have become increasingly complex and demanding, both for earth-orbital spacecraft and for the planetary exploration missions.

Data rates continue to grow. Landsat spacecraft currently require a communications capability of 15 megabits per second for the multi-spectral scanner images. In the relatively near future, the new Landsat-D spacecraft with a thematic mapper imaging-system will require data rates of 85 megabits per second. The Spacelab flights will yield telemetry data at rates up to 50 megabits per second.

Many automated spacecraft missions have become highly interactive with their control centers and this trend will continue. The Space Telescope planned to be launched in 1985 will be an observatory in orbit, operated and controlled essentially in real time by astronomers and technicians on the ground. Data must flow from the spacecraft through the TDRSS communications system, to be processed and acted upon in the mission control centers for decisions by the ground operators and for the generation and transmission of commands back to the spacecraft, all in near-real-time (Fig. 1).

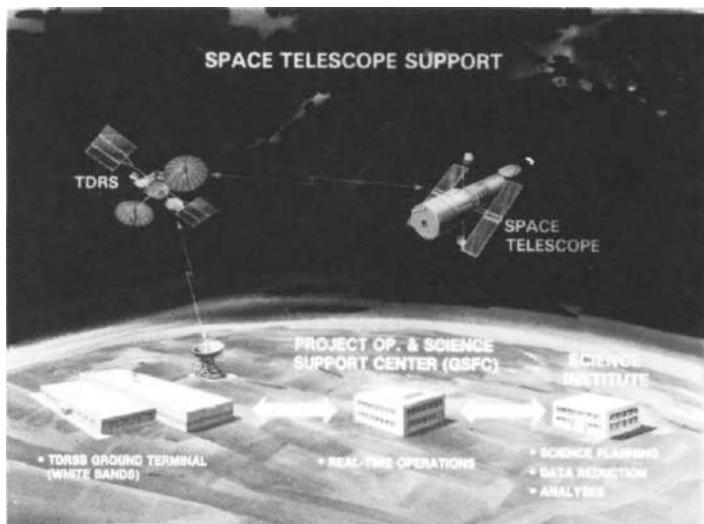


Fig. 1. Space telescope support

It has been the thrust of the TDRSS planning and development activity over the past number of years to assure that it will provide the capabilities to meet these types of emerging requirements. We must also continue to support the on-going missions already in flight, some of which are expected to achieve long lifetimes in orbit. At the same time, we must strive for improved efficiency and effectiveness of operations, so as to minimize the cost of support to all flight missions.

TDRSS, the ground network consolidation, and the new capabilities to be provided in the Deep Space Network along with this consolidation, will comprise our principal efforts in this direction. I will address each of these in the following sections of this paper.

#### SPACEFLIGHT TRACKING AND DATA NETWORK (STDN)

All NASA earth orbital spacecraft are currently supported by the Spaceflight Tracking and Data Network (STDN). This includes spacecraft in geosynchronous orbit, and those in highly elliptical or eccentric orbits extending out past the moon. The recently completed first flight of the U.S. Space Shuttle was supported by this network. I will give a brief description of the capabilities and characteristics of the STDN since they constitute the baseline from which the new capabilities to be provided by TDRSS were measured.

STDN now consists of twelve ground stations, four of which are in the continental United States and the remainder at overseas locations around the world (Fig. 2). The network is managed and operated by the Goddard Space Flight Center at Greenbelt, Maryland, near Washington, D.C. The present network has evolved from earlier network configurations which had more than twenty ground stations and instrumented tracking ships at the time of the early Apollo lunar landing missions. In reducing the network to its present number of stations, which was done in order to reduce costs of network operations, the principal criteria was to maintain a geographic configuration which would provide at least one communication contact per orbit with each satellite. The station at Fairbanks, Alaska, for example, provides a one contact per orbit coverage for several orbits per day of satellites in polar orbit.

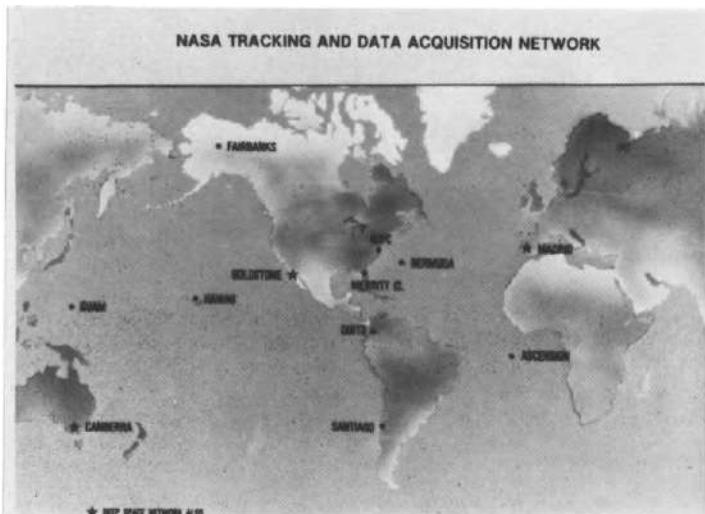


Fig. 2. NASA tracking and data acquisition network

In its present configuration, the network operates almost entirely at S-band, employing 9 meter antennas at most stations and 26 meter antennas at Goldstone, Madrid, and Canberra. A few VHF (136 mhz) facilities have been retained to provide limited support to a few of the older spacecraft, but these are being closed out as requirements diminish. Each station normally functions in a "store and forward" mode; incoming telemetry data from the spacecraft is formatted and buffered, i.e., stored temporarily, before transmission through the communication circuits back to GSFC and the mission operation control centers.

The network has been moving toward a "through-put" or "bent-pipe" mode of data flow, and this capability is available when a real-time operation is required. The Space Shuttle operational communication link on the April flight was handled in this real-time through-put mode (Fig. 3). The next few Shuttle flights will be supported in this same manner by STDN, until TDRSS is operational.

STDN currently supports an average of 30 to 40 spacecraft per day. The oldest of these, Application Technology Satellite (ATS-1) has been in orbit since 1966. After TDRSS becomes operational in 1983, the network stations required only for support of low earth orbital spacecraft will be closed out. Three stations at Goldstone, Madrid, and Canberra required to support high altitude elliptical orbit satellites will be continued and, in the TDRSS era, consolidated with the Deep Space Network under management of the Jet Propulsion Laboratory (JPL) in Pasadena, California. The station at the Kennedy Space Center, Florida, will be retained for launch phase and pre-launch support. Spacecraft in synchronous orbit will be supported by dedicated antenna facilities, usually at the Goddard Space Flight Center.

#### TRACKING AND DATA RELAY SATELLITE SYSTEM (TDRSS)

The Tracking and Data Relay Satellite System (TDRSS) will usher in a new era in NASA space tracking and data systems. For the first time, it will be possible for low earth orbital spacecraft, including the Space Shuttle and Spacelab, to have communi-

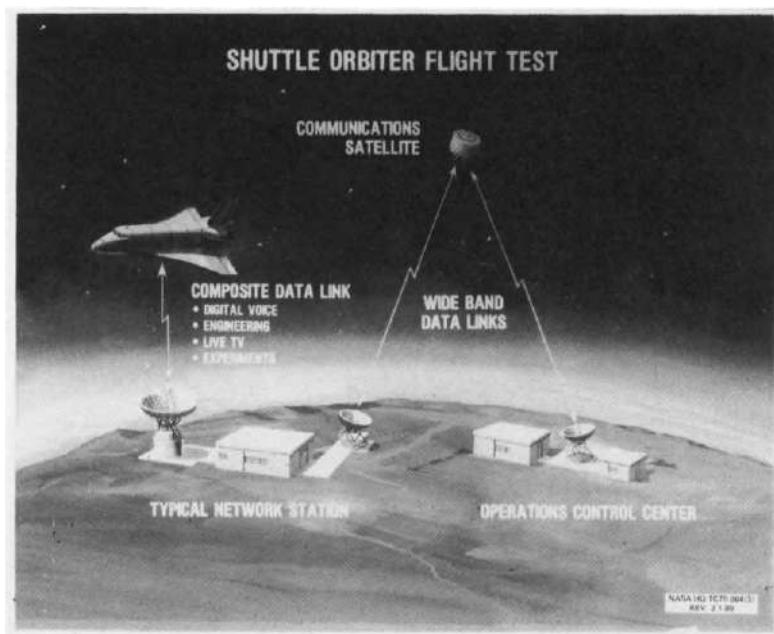


Fig. 3. Shuttle orbiter flight test

cations with the mission control center almost continuously, or at any time, as required. By contrast, the present ground network STDN provides the opportunity for communications with the spacecraft for only about 15% of the time for spacecraft in 200 kilometer orbits.

TDRSS will consist of two specialized data relay satellites in geosynchronous orbit, one stationed over the mid-Atlantic, and the other over the mid-Pacific, about  $130^{\circ}$  apart, with a ground terminal at White Sands, New Mexico. There will also be a spare satellite in orbit approximately midway between the two operational units. The TDRSS spacecraft will relay data between user spacecraft and the White Sands ground terminal and thence to and from the mission control center.

The system is being developed to NASA service requirements by an industry team headed by Space Communications Company (SPACECOM), of Gaithersburg, Maryland. SPACECOM is jointly owned by Western Union Corporation, Fairchild Industries, and Continental Telephone Company. The SPACECOM team includes TRW, Inc. and Harris Corporation. Under contract with NASA, SPACECOM will own and operate the system to provide TDRSS service to NASA for a period of ten years.

The NASA contract with SPACECOM for TDRSS service is based on the concept of a "shared system", in which the space segment will have the capability to provide commercial communications satellite services as well as TDRSS service to NASA. A fourth spacecraft (Advanced Westar) in geosynchronous orbit will be used by SPACECOM for these communication satellite services. The total space segment constellation will thus consist of four satellites: two dedicated to NASA TDRSS service, one to Advanced Westar service, and one spare that can substitute for any one of the three operational satellites (Fig. 4).

The TDRSS spacecraft is a 3-axis stabilized configuration, with sun-oriented solar panels. Each spacecraft will have two high gain 4.9 meter diameter deployable

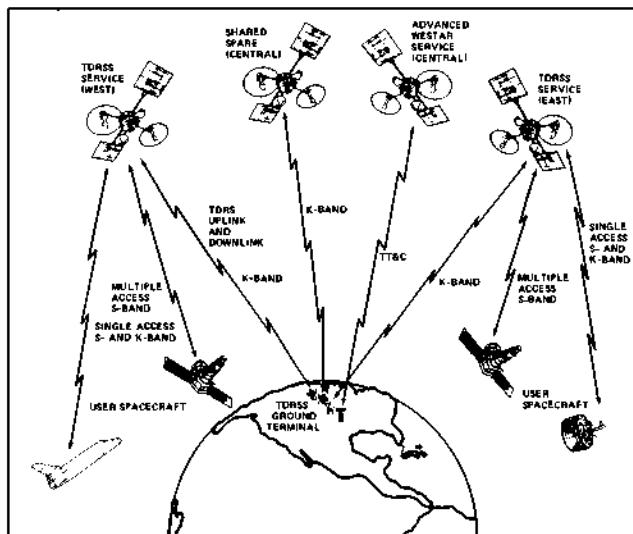


Fig. 4. Tracking data and relay satellite system

parabolic antennas, each of which can provide single access S-band or Ku-band service to a user spacecraft. There is also an electronically steerable phased array antenna, known as the multiple access system, which can communicate with as many as 20 low data rate users concurrently. The multiple access system consists of 30 helical elements in an array, and operates at S-band. Each user is limited to a data rate of 50 kbs. The single access antenna can provide communication at rates as high as 300 megabits per second at Ku-band. A gimballed 2 meter boom-mounted antenna provides Ku-band communication to and from the White Sands ground terminal.

Telemetry, tracking, and command (TT&C) functions are provided by an S-band omnidirectional antenna during the period from launch to synchronous orbit. During the remainder of the TDRSS mission, TT&C functions are provided via the spacecraft-to-ground terminal Ku-band link. Advanced Westar communications use a K-band antenna rigidly mounted on the multiple access antenna ground plane and a boom-mounted C-band shaped beam antenna (Fig. 5).

The spacecraft will be launched on the Space Shuttle and will be transported to synchronous orbit by Boeing's Inertial Upper Stage (IUS). At launch the spacecraft will weigh about 2200 kilograms, including fuel and a spacecraft-to-IUS adapter. Enough station keeping and attitude control fuel will be carried to maintain 10 years of operating life at 0° inclination. When deployed on-orbit, the spacecraft will measure 13 x 17.4 meters and will carry a small solar sail to balance solar pressure disturbance torques (Fig. 6).

The four satellite constellation will be controlled by a ground terminal located at NASA's White Sands Test Facility near Las Cruces, New Mexico. This ground terminal will be equipped for all radio frequency and data processing and handling operations required for tracking and for relaying digital communications to and from NASA's earth orbiting satellites. It will also provide simulation and verification capabilities and will function as the operations control center for all four of the orbiting TDRSS spacecraft. The ground terminal will be primarily for NASA services; the commercial Advanced Westar services will originate and terminate in SPACECOM'S facilities in other locations and will not be controlled from this ground terminal (Fig. 7).

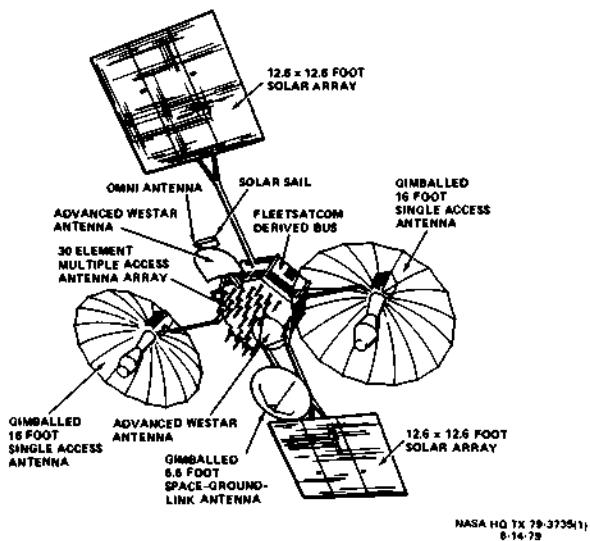
**TDRSS SATELLITE**

Fig. 5. TDRSS satellite

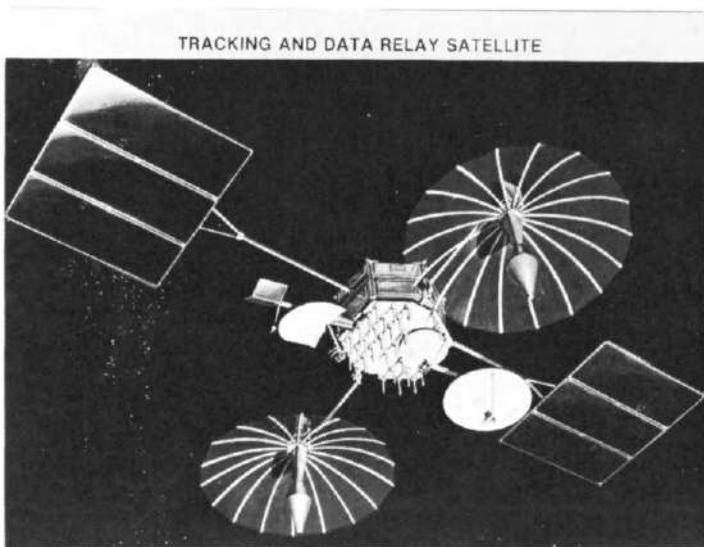


Fig. 6. Tracking and data relay satellite



Fig. 7. TDRSS ground terminal White Sands, New Mexico

NASA's contract with SPACECOM provides for ten years of service. To assure continuity of both NASA and Advanced Westar services, a total of six spacecraft are being built, - four to establish the initial four satellite constellation and two as ground spares which will be launched as necessary to compensate for failures in the orbiting satellites. Although there is high confidence that six spacecraft are all that will be needed for the 10 years of service, parts will be ordered for a seventh spacecraft so that it can be added if required.

#### TDRSS OPERATIONS

Conceptually, the TDRSS satellites function as "bent-pipe" repeaters relaying Forward Link commands and voice from the ground terminal to user spacecraft, and relaying Return Link telemetry, voice and scientific data from the user spacecraft to the ground terminal. The forward link data destined for user spacecraft are frequency-division multiplexed by the ground terminal and transmitted as a composite uplink to the TDRSS satellite in the 14.6 to 15.225 GHz frequency band. The TDRSS demultiplexes the composite uplink by translating each subchannel to a frequency which is compatible with the user spacecraft receiver and then transmits it to the user via a high gain single access antenna or via the multiple access antenna. The return links transmitted by user spacecraft to TDRSS for relay to the ground terminal are received by the TDRSS antennas and translated to an appropriate frequency in the 13.4-14.05 GHz band for frequency-division multiplexed transmission to the ground terminal via the two meter space-ground-link antenna.

The ground terminal at White Sands provides Ku-band communications to and from the East, West and Spare TDRSS satellites through three 60-foot antennas. Each of these large antennas interface with radio frequency equipment, computers, displays, and control equipment within the ground terminal to sustain the ground-space link operations for the operational TDRSS satellites. The ground terminal also generates and transmits commands and analyzes telemetry as necessary to control the TDRSS space-craft.

Systems operations supporting user spacecraft will be scheduled by NASA from a new Network Control Center at the Goddard Space Flight Center. This control center will be the focal point for operational control of TDRSS user spacecraft support services and the ground communication system that interfaces user spacecraft control and data processing facilities with the TDRSS ground terminal. User spacecraft support requirements will be supplied by project operations control centers and correlated with ephemerides to generate conflict-free schedules for the TDRSS relay of user spacecraft command, telemetry, tracking, and simulation data. These schedules will be sent from the Network Control Center as digital messages to the ground terminal where specific TDRSS operations will be controlled by a TDRSS control center. User spacecraft ephemerides, state vectors, and other necessary information will also be provided in this message. From beginning to end, the system will be highly automated and will be able to respond quickly to change.

#### TDRSS SERVICE

After it becomes operational in 1983, the TDRSS will be a major national resource for command, control, and data operations with user spacecraft. The TDRSS will significantly improve NASA's capabilities to access spacecraft during critical mission periods. For example, the Space Shuttle will be in view of TDRSS nearly full time during launch and landing operations and during on-orbit operations. Key events in the mission can be conducted with full participation of ground based flight controllers. This capability is also available for Spacelab missions where considerable real time ground-based experiment control and monitoring may be necessary to complete the multitude of experiments planned for each mission.

NASA's space data return capabilities will be greatly increased not only as a result of long contact periods, but also because the data rate capabilities will be greatly improved over current systems. In addition to as many as twenty multiple access system users, each with data rates up to 50 KBPS, each TDRSS satellite will also be capable of relaying from two single access user spacecraft at both S-band and Ku-band frequencies with composite data rates up to 300 MBPS. Two of the early users of the TDRSS single access service will put these high data rate capabilities to good use. Spacelab will transmit scientific data rates up to 50 MBPS and Landsat-D will transmit earth-imaging data at rates up to 100 MBPS or more.

Perhaps our principal challenge over the next decade will be to exploit these new capabilities brought into the space tracking and data system arena by TDRSS. The capacity for very high rate data flow on a near-continuous or on-demand basis, and the ability to have command access to the spacecraft at any time in orbit, will introduce a new dimension into the conduct of space missions in low earth orbit. The capability will exist for any earth orbital spacecraft to function in a highly interactive mode between the spacecraft/experiment operation and the ground-based experimenters and controllers, a mode of operation such as is now available, for example, to the International Ultraviolet Explorer (IUE) in its synchronous orbit location. In Spacelab missions, it will allow the mission and payload specialists on board to maintain continuing communication with ground-based scientists and experimenters.

R. E. Smylie and R. O. Aller  
TRACKING, DATA ACQUISITION, AND COMMUNICATIONS NETWORKS  
IN THE TDRSS ERA

We plan to begin TDRSS support of NASA spacecraft, particularly the space shuttle, as soon as the first TDRSS is fully checked out in orbit, perhaps sixty to ninety days after launch. Spacecraft currently in low earth orbit but not compatible with TDRSS will continue to be supported by the ground network STDN for a transition period, until the second TDRSS spacecraft is operational in orbit and the spare has been launched. We then will close down the STDN stations required solely to support low-earth orbital spacecraft.

#### Launch Phase Support

The STDN station at the Kennedy Space Center, Florida, will be continued in operation as a launch phase support station. Since this station will function as an alternate to TDRSS for support of the space shuttle during the pre-launch and the immediate post-launch period, it must be configured to "look like" TDRSS to the shuttle communication system. A similar capability is available to interface in a TDRSS-compatible mode with automated spacecraft for on-the-pad pre-launch check-outs.

#### Network Consolidation and the Deep Space Network

Three STDN 26 meter antennas at Goldstone, Madrid, and Canberra will be retained to support high altitude elliptical orbit spacecraft beyond the range of TDRSS - for example, the on-going International Sun-Earth Explorer (ISEE) mission. These three antennas will be merged into the Deep Space Network (DSN), under management of the Jet Propulsion Laboratory (JPL). This consolidation will provide an opportunity to enhance the DSN capabilities for greater data return from planetary missions. We are planning to convert two 26 meter S-band antennas at each station to S/X band systems. We will then have a 64 meter and three 34 meter antennas at each station, which can be arrayed as necessary for critical phases of a planetary mission to provide an improvement of nearly a factor of two in data return from the spacecraft.

The use of antenna arraying to enhance signal reception from distant spacecraft has been evolving in the DSN during recent years. An array of the 64 meter and 34 meter antennas was tested and demonstrated on the Voyager fly-by of Jupiter and was employed operationally on the recent Voyager fly-by of Saturn, to improve the image data from those two encounters. Development of this technique will continue, and it is expected that by the time of the Voyager encounter of Uranus in 1986, additional antennas will have been brought into the array, to provide a data reception capability of about 30,000 bits per second from the spacecraft at Uranus distance - some 3 billion kilometers from Earth.

At the same time, it is planned to introduce some reconfiguration of the DSN station systems and facilities and to automate these systems so as to improve the efficiency of operation and reduce operation and maintenance manpower. We expect in this way to achieve a reduction of better than 25% in operation and maintenance costs for the new single network from the present levels.

#### NASA Communications Network (NASCOM)

Communications and data flow between the network ground stations, the TDRSS Ground Terminal, mission control centers, and the data processing centers are handled by a network of terrestrial and satellite communication circuits leased from the commercial common carriers, both national and international (Fig. 8). The end terminals of these circuits at the stations or other NASA facilities are comprised of NASA-owned computer-controlled switching centers, data modems, and other terminal equipment. The overall system is known as NASCOM (NASA Communications) and is managed and operated by the Goddard Space Flight Center.

### NASA TRACKING, DATA ACQUISITION, AND COMMUNICATIONS NETWORK

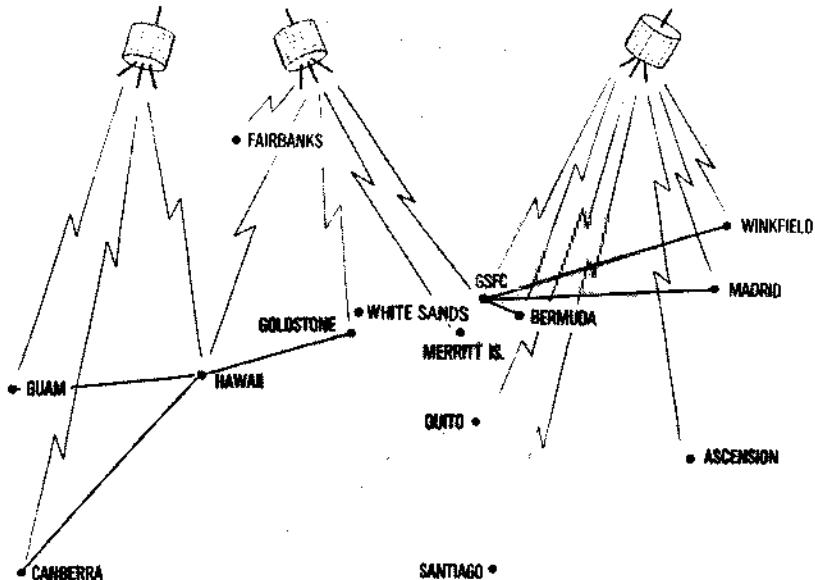


Fig. 8. NASA tracking, data acquisition, and communications network

Most of the communication between the network stations and the control and data processing centers now flow through communication satellites, both domestic satellites and Intelsats. The domestic carriers providing these services have their own terminals at the various NASA centers or facilities, as required.

#### INTERNATIONAL COOPERATIVE MISSION SUPPORT

The NASA network facilities, including the communications and mission control system, provide support for a number of international cooperative space missions. This practice is expected to continue with the Tracking and Data Relay Satellite System. The International Sun Earth Explorer mission (ISEE) and the International Ultraviolet Explorer (IUE) are examples of such missions. ISEE is a joint program between NASA and the European Space Agency (ESA). The mission involves three spacecraft, two of them in highly eccentric earth orbits and one in a halo orbit about a libration point between the Earth and the Sun, some 800,000 kilometers from Earth. NASA conducts the tracking, telemetry acquisition, and control of the spacecraft, and performs preliminary processing of the data which is then transmitted to science investigators at the Paris Observatory, the Max Plank Institute, the University of Iowa, and Stanford University.

The International Ultraviolet Explorer, a joint NASA/ESA/UK mission, is a spacecraft in near-synchronous orbit. This spacecraft is supported jointly by a NASA ground station at the Goddard Space Flight Center and an ESA ground station in Spain.

The NASA network facilities are also available on a cost-reimbursable basis for support of spacecraft of other countries. This practice also will continue with TDRSS. A number of such arrangements have been negotiated and established over the past several years between NASA and the space agencies of other countries, and more are expected in the future. Reimbursement policy and rate structures for the use of TDRSS are currently being developed.

#### SUMMARY

In summary, the Tracking and Data Relay Satellite System (TDRSS) is scheduled to come into service in 1983 and will begin support of the Space Shuttle, Spacelab, and the automated spacecraft launched into low earth orbit by the Space Transportation System. Present ground network stations supporting only low earth orbital missions will be closed out starting in 1984. Three remaining 26 meter antenna stations will be merged into the Deep Space Network, and this consolidated network will support high-altitude elliptical orbit spacecraft as well as the planetary exploration missions. The network stations, the TDRSS ground terminal, and the control centers will be interconnected by leased communication satellite circuits. This overall tracking, data acquisition, and communications configuration is depicted in Fig. 9.

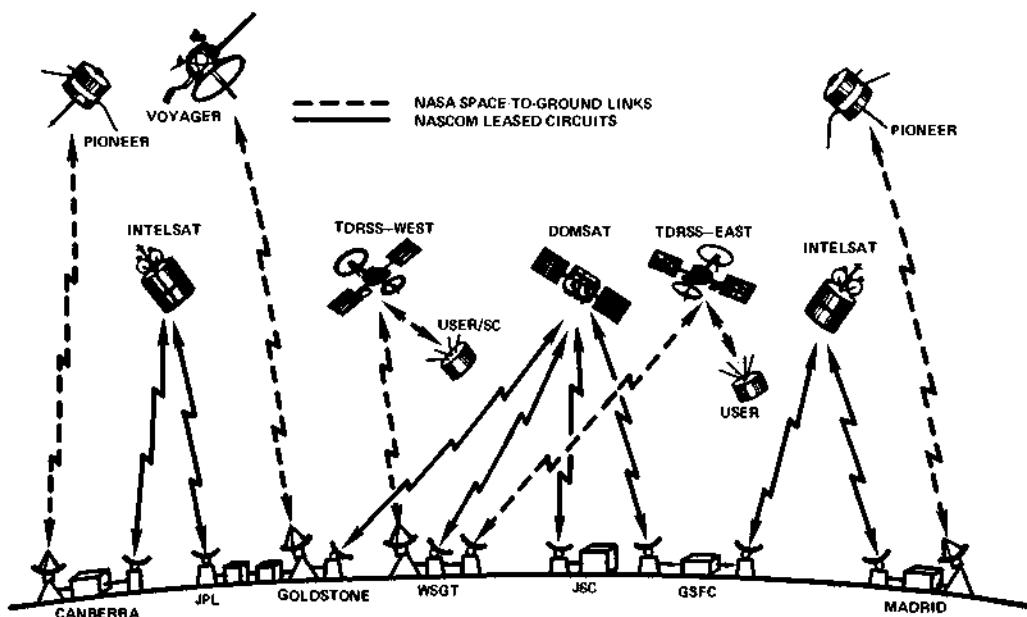


Fig. 9. The NASA space tracking and data systems network and TDRSS configuration for the 1980's

## INTELSAT VI SS-TDMA

G. Forcina, B. Pontano and R. Colby

*International Telecommunications Satellite Organization,  
Washington, D.C., USA*

### ABSTRACT

The basic principles of the beamswitched SS-TDMA approach adopted in the INTELSAT VI satellite are illustrated. The concepts of Switch Frame, Frame Unit and Switch State are discussed. The paper then outlines the general SS-TDMA system configuration obtained from previous system definition studies. A block diagram of the SS-TDMA sub-system meeting these general requirements is described. The paper then proceeds to examine specific requirements and constraints which have been taken into account in developing the performance requirements for the SS-TDMA sub-system.

### KEYWORDS

**INTELSAT VI; SS-TDMA; Microwave Switch Matrix; Distribution and Control Unit; Timing Source; Switch Frame; Switch Master Frame.**

### INTRODUCTION

The new satellites of the INTELSAT VI series will have the maximum capacity of 37,000 voice circuits, which is three times the capacity of the satellites of the previous generation. This increase in capacity is achieved by providing a multiple beam coverage which permits to obtain a six-time frequency reuse at 6/4 GHz and two time frequency reuse at 14/11 GHz (Perillan, 1981). A typical INTELSAT VI beam coverage pattern is shown in Fig. 1. In addition, a better efficiency of bandwidth utilization is obtained by employing a Time Division Multiple Access (TDMA) scheme in several of the satellite transponders. While a multiple beam coverage permits to increase the satellite capacity through frequency reuse, it also restricts the traffic connectivity unless some means to dynamically alter the beam interconnection is provided. In the INTELSAT VI satellite, this capability is provided by a switching device which routes the TDMA traffic to the appropriate coverage areas. This arrangement is referred to as Satellite Switched TDMA (SS-TDMA).

Two types of SS-TDMA exist, i.e., beam-switched SS-TDMA and baseband-switched SS-TDMA.

In the first type, an RF switch matrix is used to connect N up-beams to N down-beams. The switch matrix consists of an array of fast acting switching elements which can establish any set of connections between the input and the output beams on a pre-programmed cyclic basis.

In the baseband-switched SS-TDMA, baseband data are extracted from the received TDMA bursts by on-board demodulators, sorted out according to destination and routed digitally to the appropriate down-beams. Fig. 2 illustrates the two methods.

A beam-switched SS-TDMA approach was selected for the INTELSAT VI since this technology was considered to be more mature than the on-board processing technology required by the baseband-switched SS-TDMA method.

#### SS-TDMA SUB-SYSTEM CONFIGURATION

The selected configuration for the INTELSAT VI SS-TDMA sub-system is shown in Fig. 3. This figure shows two Microwave Switch Matrices (MSM), both providing a dynamic interconnection between the six 6 GHz up-beams and the six 4 GHz down-beams of the INTELSAT VI Satellite. Each of the two matrices routes up-link RF signals to a separate bank of six transponders connected to the 4 GHz down-beams. The transponders of both banks have the same nominal bandwidth, 80 MHz, but the transponder center frequency is different for the two banks. The two frequency bands allocated to the two banks of transponders are referred to as channel 1-2 and channel 3-4, respectively. Fig. 4 shows the frequency allocation and the coverages associated with the two transponder banks.

Each MSM is driven by a dedicated Distribution and Control Unit (DCU). The DCU contains in its memory a pre-programmed set of beam-interconnections (also referred to as Switch States) and provides a drive signal to the MSM for the execution of these interconnections. The execution of the Switch States is periodically repeated, the duration of the repetition cycle being the Switch Frame, defined later. The time of execution of each Switch State, expressed in Frame Units, is also stored in the DCU memory. The Frame Unit is a sub-multiple of the Switch Frame, i.e.,

$$1 \text{ Frame Unit} = \frac{1 \text{ Switch Frame}}{N}$$

where: N is an integer number (N = 1,888 in the INTELSAT VI specification).

The Timing Source provides timing signals to both DCU's. Since both DCU's operate on a common timing, execution of the Switch States in the two MSM's will be performed in a synchronous fashion, even though the Switch States of the two MSM's will in general be different. This permits a single TDMA terminal to hop between channel 1-2 and channel 3-4 (frequency hopping).

#### TIMING REQUIREMENTS FOR THE SS-TDMA SUB-SYSTEM

This section defines the basic timing parameters of the SS-TDMA subsystem and discusses their relationship with the timing parameters of the terrestrial TDMA network.

### Switch Frame

A basic requirement for the SS-TDMA sub-system is that it must be compatible with the TDMA terrestrial network developed for use with the fixed beam INTELSAT V satellite. The TDMA network operates with a frame rate of 2 ms. The TDMA frame consists of traffic bursts and of two reference bursts, as shown in Fig. 5. The first reference burst (RB1) designates the beginning of the frame, while RB2 is used to derive the start-of-frame when RB1 fails (Forcina, 1981). Fig. 6 shows the reference stations 1 and 2 generating reference bursts RB1 and RB2, respectively. In the picture, it is assumed that both reference stations are contained in the NW Zone beam. The reference bursts transmitted into the up-link NW Zone beam must be distributed to all down-beams to provide the start-of-frame to the traffic terminals of the network. The MSM must change beam connections during the TDMA frame in such a way that the TDMA bursts are routed to the appropriate down-beams as shown in Fig. 7.

The sequence of Switch States implemented by the MSM is periodically repeated in a fixed duration cycle which constitutes the Switch Frame. The Switch Frame is independently generated by the on-board Timing Source, however, synchronization with the TDMA frame is required for proper SS-TDMA operation. It is the task of the terrestrial TDMA network to adjust its frame timing in order to synchronize the TDMA frame with the Switch Frame. This function is performed by the Acquisition and Synchronization Unit (ASU), which is located in the reference stations (Campanella, 1981). The ASU adjusts the transmission timing of a special burst (metering burst) to maintain the center of this burst lined up with the trailing edge of a narrow loop-back window generated by the MSM especially for this purpose. Fig. 8 illustrates this concept.

### Switch Master Frame

The structure of the TDMA frame (burst time plan) is changed periodically to accommodate variations of traffic, such as new bursts added, existing bursts expanded or contracted or moved in the frame. In several cases, it may be necessary to change the Switch States to accommodate the new burst time plan. It is a system requirement that changes of the Switch States sequence be synchronized with changes of the burst time plan so that no traffic interruption would result. This requires that the first TDMA frame organized in accordance with the new time plan arrives at the satellite exactly at the time of implementation of the new set of Switch States in the MSM.

In order to achieve this purpose, the Switch Master Frame is introduced. The Switch Master Frame is a fixed repetitive cycle whose period contains an integer number of Switch Frames. A new set of Switch States can only be implemented at the MSM at the beginning of a Switch Master Frame. In order to implement a synchronized traffic change, is necessary to have knowledge, at the reference station, of the time of occurrence of the start of the Switch Master Frame. This is obtained by doubling the size of the loop-back window (used by the ASU for frame synchronization) at the beginning of the Switch Master Frame. When this widened loopback window occurs at the satellite, a portion of the metering burst, which is normally cut-off by the satellite switch, is received by the ASU at the reference station (See Fig. 8). This signals the occurrence of the start of Switch Master Frame at the satellite. The reference station makes use of this information to synchronize the change of burst time plan with the change of Switch State sequence in the MSM.

### Timing Accuracy

It has been shown that the TDMA frame of the terrestrial network is synchronized to the satellite Switch Frame. Therefore, the frame rate accuracy of the TDMA network is the same as the accuracy of the Switch Frame rate. The interfacing requirement of the TDMA network with the terrestrial digital networks are such that a TDMA frame rate with an accuracy of one part in  $10^{11}$  is required. For the reason discussed above, the same accuracy would be required for the Switch Frame generated by the satellite Timing Source. This accuracy could only be obtained if an on-board atomic clock (e.g., a cesium beam frequency standard) was used. The reliability of this type of clocks was not considered compatible with the overall spacecraft reliability requirements so that a different approach was chosen. In this approach, the frequency of the on-board Timing Source is controlled from the ground and maintained to its preassigned value within the limits of the specified accuracy. A possible way of controlling the frequency of the Timing Source is to send a special command from the ground for the frequency adjustment of an on-board variable frequency oscillator.

### SS-TDMA SUB-SYSTEM SPECIFICATIONS

In the previous sections, the basic concepts of the SS-TDMA operation have been outlined and some considerations on the SS-TDMA interface with the terrestrial TDMA network have been presented. In this section, the major specification requirements for the SS-TDMA sub-system components are summarized.

#### Timing Source

The Timing Source is required to generate three timing signals which are provided to the DCU.

- a) Switch Frame Signal. Its period is 2 ms (same as the TDMA frame).
- b) Switch Frame Unit Signal. A Frame Unit is one of the 1,888 equal intervals into which the Switch Frame can be divided. Beam connections can be established or deleted only at the Start of a Frame Unit. Therefore, the minimum duration of any Switch State is one Frame Unit.
- c) Switch Master Frame Signal. Its period consists of 8,192 Switch Frames.

The frequency of the Timing Source is adjustable from the ground. Triple redundancy is required.

#### Distribution and Control Unit

The set of required Switch States is programmed into the DCU from the ground via a command link. Forty-eight Switch States can be stored in any of the DCU memories. The DCU has three memories:

- a) On-line memory. It controls the MSM, and contains the Switch States currently being implemented.
- b) Off-line memory. It is used to store a new set of Switch States. At a preassigned time (start of a designated Switch Master Frame), the off-line and on-line memories switch roles.

- c) Stand-by memory. It can replace any of the other two memories in case of failure.

Full redundancy is required for the DCU.

#### Microwave Switch Matrix

The MSM provides dynamic interconnectivity among six up and six down beams. For reliability reasons, the MSM is equipped with extra rows and extra columns.

One up-beam can be connected with one or more (up to six) down-beams. Two different up-beams cannot be connected to the same down-beam.

When the DCU is programmed with only one Switch State, the MSM acts as a static switch.

The rise and fall time of the MSM switching elements is 100 ns (this requirement insures proper operation of the ASU).

When a switching element is in off-condition, it provides an attenuation of 50 dB.

#### CONCLUSIONS

The SS-TDMA sub-system performance requirements have been developed for inclusion in the overall INTELSAT VI satellite specification.

The Request for Proposal for the procurement of this satellite has been released, proposals have been received and, at a time of writing, the proposal evaluation work is underway.

#### ACKNOWLEDGEMENT

The INTELSAT VI SS-TDMA system architecture, from which the specification for the SS-TDMA sub-system was derived, was developed with the cooperation of many contributors to whom we want to express our grateful acknowledgement. In particular, we wish to thank Dr. S. J. Campanella, Executive Director, Communications Technology Division, COMSAT Laboratories for his key contribution to this effort.

#### REFERENCES

Perillan, L. and T. R. Rowbotham (1981). Proc. 5th International Conference on Digital Satellite Communications, Genoa, Italy, 411-418.

Forcina, G., B. Pontano, and R. J. Colby (1981). Proc. 5th International Conference on Digital Satellite Communications, Genoa, Italy, 377-386.

Campanella, S. Joseph and Roger J. Colby (1981). Proc. 5th International Conference on Digital Satellite Communications, Genoa, Italy, 335-343.

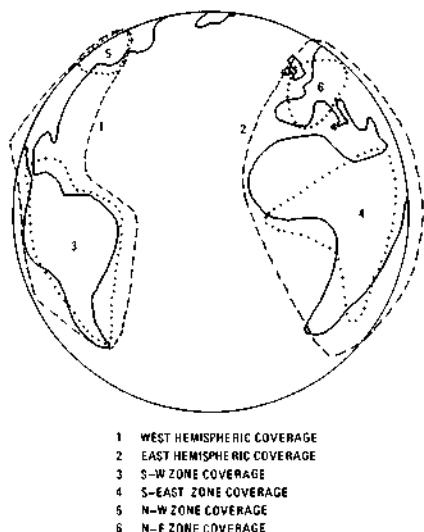


Fig. 1 INTELSAT VI Coverage Areas (Atlantic Ocean Region, C-Band)

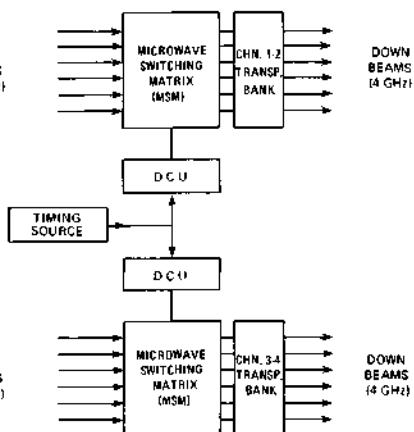


Fig. 3 INTELSAT VI SS-TDMA Sub-System Configuration

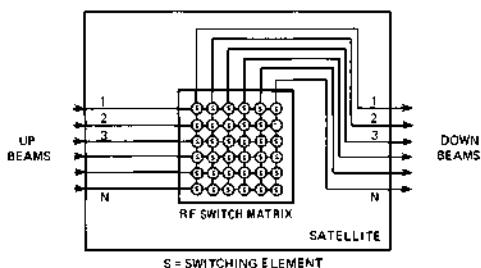


Fig. 2a Beam-Switched SS-TDMA Concept

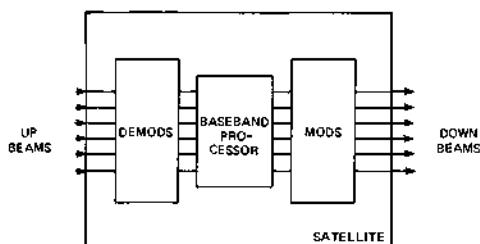


Fig. 2b Baseband-Switched SS-TDMA Concept

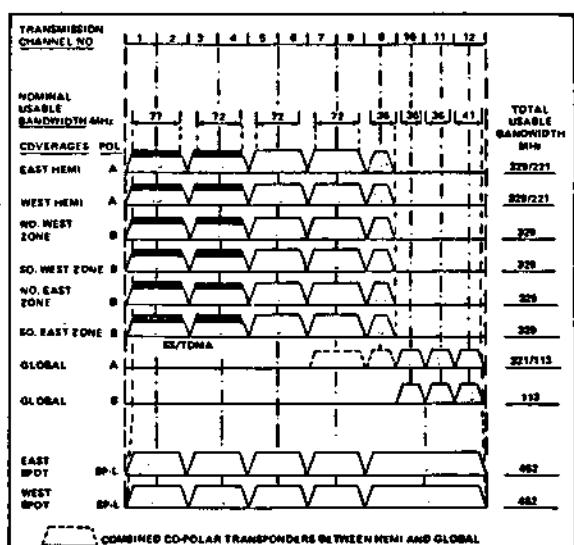


Fig. 4 Transponder Frequency Allocation and Coverages

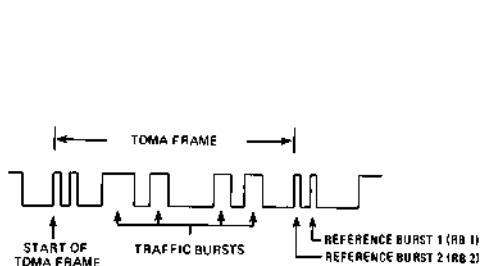


Fig. 5 Example of TDMA Frame

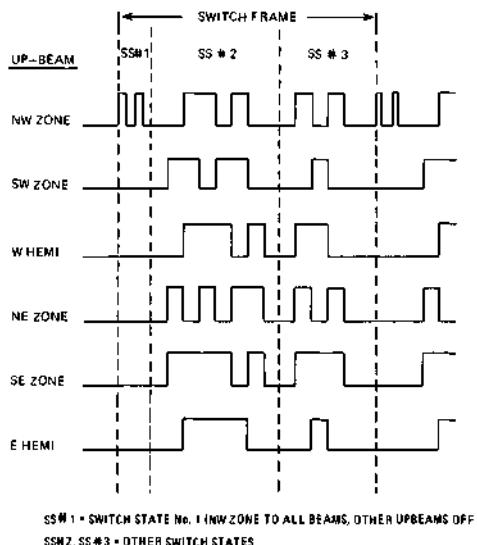


Fig. 7 Example of Switch State Sequence

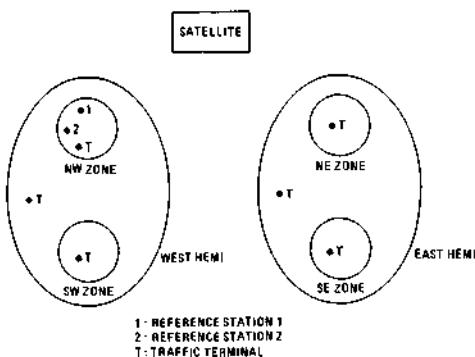


Fig. 6 Possible Reference Station and Traffic Terminal Distribution

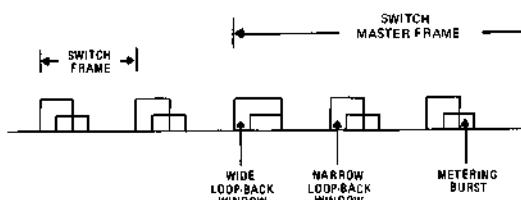


Fig. 8 ASU Frame Synchronization Concept

## ITALY IN MILLIMETER WAVES ACTIVITIES FOR SPACE COMMUNICATION

G. Manoni and S. Ferri

*Selenia S.p.A., Rome, Italy*

### ABSTRACT

The growth of Communications via Satellites requires , mandatoryly, the utilization of high frequencies, above 10 GHz, to avoid bandwidth saturation and interference problems. This paper aims to summarize some key italian activities, and achievements, in this field. The first started in the 1968 with the successful experimental satellite SIRIO 1 working at 12-18 GHz.

The ESA's H-SAT payloads definition took place in 1977-78.

A domestic satellite working at 20/30 GHz is presently being considered in Italy, for serving the Intercity communication needs in the decade 1990-2000. To achieve system concept validation through in orbit tests, a preoperational satellite is foreseen for 1986.

Part of these studies are conducted under ESA coordination in the frame of ASTP program. Italy is contributing to two payloads (Television Broadcasting and 20/30 GHz multimission communication experiment) on-board ESA's L-SAT: this will enable to further exploit the potential benefits of these bands and to develop critical hardware also vital for other missions.

### KEYWORDS

Millimeter Waves, Space Communication, Broadcasting via Satellite, Point-to-Point Communication, Multiple Beam Antenna, Regenerative Repeater.

### INTRODUCTION

The constant growth of the traffic to be routed via communication satellites, mandatory requires to use frequencies above 10 GHz in order to avoid bandwidth saturation and interference problems.

Italy has pioneered the use of high frequencies for communication satellites: the SIRIO project aiming to experiment with 12 and 18 GHz bands was infact established in 1968, developed through 1976, with a satellite launch in 1977, and is operative since then. Incidentally the frequencies chosen are those now used for down and uplinks of the preoperational TV Broadcasting Satellites: the propagation data gathered at 18 GHz can be considered very use-

ful for uplink dimensioning.

The success of SIRIO has encouraged the Italian authorities to promote a multi year National Space Programme, which has a high technology content specially for what concerns the development of satellite communication systems operating in the 20/30 GHz bands, and possibility to explore the 40-50 GHz bands. The communication oriented part of the N.S.P. is centered about the development of a preoperational satellite (ITALSAT), aimed for a launch in 1986.

The National Space Program has obviously a complementary nature to the international ones, specially vis-a-vis of the ESA programs, to which Italy contributes.

At present advanced activities in Space Communication with ESA, mainly concern :

- L-SAT experimental payload at 20/30 GHz;
- ASTP: developments of critical hardware for a multispot beam antenna and regenerative repeater for point to point communications at 20/30 GHz.

Earth Resources Observation from satellites is another exciting field where much remain to be done on sensors, specially using Microwaves.

Passive Radiometry at S.H.F., in particular at millimeter wavelengths (above 60 GHz and up to 180 GHz), should offer a potential for certain accurate and sensitive earth observation. Some initial, and limited, work on the possible use of these frequencies has been also conducted.

In the following the historical background and an outline of ongoing activities in this field are presented.

#### THE "PAST"

"Past" activities on 20/30 GHz date back to 1968 and end in 1980. They concern mainly:

- The SIRIO program
- ESA's H-SAT
- ASTP system studies oriented to the identification of mission needing the 20/30 GHz bands.

#### The Sirio Program

The Sirio program was originated in 1969; its main purpose was to develop an Italian capability in Space Systems. A high technology content - for that time - was given to the program: to experiment with 12 and 18 GHz frequencies.

Selenia was the payload contractor responsible for the development of a SHF repeater and a mechanically despun antenna, to which many Italian industries and some U.S.A. suppliers did contribute.

The satellite was launched from Cape Canaveral on August 25th, 1977. Its operation, more than 4 years after the launch, is still nominal and it should be mentioned that the contractual life was 2 years.

The payload was designed for two types of experiments: (Fig. 1)

- propagation and
- communication (wide band and narrow band)

The main propagation measurements are absolute and relative attenuation, phase distortion

and depolarization characteristics of the transmission medium.

The propagation data so far obtained are the basis for the design of present domestic satellite system and have demonstrated the applicability of high frequencies for satellite communications in Europe.

In the communication mode various type of experiments have been performed among which:

- wide band FM Television transmission: to determine the best modulation parameters, for standard color television signals, to achieve the standard quality with the minimum power transmission.
- Digital TV and data transmissions at 2 Mbit/sec were experimented.

A large number of international institutions are carrying out experiments with SIRIO since the very beginning of the program. The studies and the experiments made have encouraged the high frequency utilization in space communication.

#### The H-SAT Program

The European Space Agency launched in 1977, the H-SAT (Heavy Satellite) program with the following payloads elements:

- a television broadcasting system operating at 11-14 GHz
- propagation experiment at 20/30 GHz over a European coverage
- communication experiment at 20/30 GHz

The two 20/30 GHz experiments (Fig. 2) were proposed to verify the feasibility of an operational Regional communication satellite system for the European traffic requirements of this decade.

Selenia was again responsible for both the 20/30 GHz communication and propagation payloads (Fig. 3). The purpose of the first was to verify the effectiveness of space diversity at 20/30 GHz, to experiment and develop SS-TDMA techniques, and develop millimeter waves technologies in Europe.

The purpose of the second was to make available, to all European Countries, stable satellite beacons at 20/30 GHz to collect long term statistics of the propagation characteristics at the likely sites of ground terminals for future communication satellite systems.

Complete system definition was achieved in 1978: however the H-SAT program was terminated before commencing the real hardware phase for political reasons.

It has to be noted that a payload very much similar to that envisaged for H-SAT 20/30 GHz communication experiment, is now being considered in Germany for experiments at these bands in the frame of their National Space Program.

#### ASTP Activities

In the frame of ESA's ASTP, the study of applications of the 20/30 GHz bands to an overall System Architecture, and the design and development of advanced 20/30 GHz repeaters and multiple spot beam antenna have been awarded to Italy.

A domestic satellite working at 20/30 GHz for serving the Italian intercity telephony communication needs in the decade 1990-2000 was the selected Reference Application.

The work was divided into two main phases: the first aimed to arrive at a system concept identification and preliminary critical hardware breadboard realization; the second was de-

voted to hardware implementation.

During phase 1, which was concluded in 1980, three parallel activities have been performed:

- Overall system study - by Telespazio
- Repeater definition - by Selenia
- Antenna definition - by Selenia

A mission outline, which did serve as a basis for the satellite payload definition, is summarized in table 1. Only the key features established during the mission study for an Italian intercity trunk telephony communication satellite system are reported.

The definition of the 20/30 GHz payload (repeater and antenna configuration) performed by Selenia translated the mission requirements into a set of payload requirements achieving an optimum configuration with respect to performance, reliability, hardware implementation, mass and DC power consumption.

The main outcome of this work was to allow the identification of critical hardware and technological requirements, to be implemented during the succeeding phases.

The payload key characteristics are summarized in table 2.

During phase 1 breadboard activities concerned: the front end section, with a L.N.A. and 30/12.5 GHz down-converter by G.T.E.; for the local oscillator, study and testing of a FET oscillator at about 9 GHz by FIAR; development of a 40PSK modulator at 20 GHz by Selenia (Fig. 5).

TABLE 1 Key Feature of 20/30 GHz ASTP Mission

<u>Frequency band</u>	Up-link      27.5 ÷ 30 GHz Down-link    17.7 ÷ 20.2 GHz
<u>Coverage</u> (Fig. 4)	Spot coverage with one station per transponder Two multiple spot beam antennas are required
<u>Access method</u>	SS-TDMD (Satellite Switched Time Division Multiple Destination). SS is performed by means of a base band switching matrix
<u>Bit rate and modulation</u>	360 Mb/sec, modulation 40PSK. A total of 17 carriers are required to meet the telephony traffic projections for the 1990-2000
<u>Required bandwidth</u>	4.2 GHz with DS1. The required frequency reuse will be done by spatial discrimination
<u>Station/Spot</u>	Any spot (total 13) corresponds to one station only

TABLE 2 Key Features of 20/30 GHz Payload defined under ASTP

<u>ANTENNA</u>	
Diameter:	4 mt (2 antennas)
Beamwidth:	0.24° at 20 GHz; 0.17° at 30 GHz
Peak gain:	55 dB at 20 GHz; 59 dB at 30 GHz
Beam pointing accuracy:	0.03° achieved with an APM (Antenna Pointing Mechanism) slaved to a RF sensor system
Beam to beam spacing:	1.3 beamwidth minimum
Number of spot beams:	total of 13 (6 for one antenna; 7 for the other)
Polarization:	linear vertical both for up and down link
<u>REPEATER</u>	
Type:	Regenerative repeater with demodulation at an IF of 12.5 GHz. Direct modulation at 20 GHz. The satellite switching function is performed at base band
Channel bandwidth:	250 MHz
Number of channels:	17 for a total of 4.2 GHz occupied bandwidth
Frequency reuse:	by spatial discrimination
RX noise figure:	5 dB (paramp at 30 GHz is required)
TX power:	up to 20 W

### THE PRESENT

Present activities, from second half of 1980 onwards, deal mainly with:

- 1) ESA's L-SAT Payloads
- 2) Continuation of the ESA's ASTP advanced hardware development
- 3) Start of the Italsat predefinition phase

### The L-SAT Program

A new program was defined by ESA in 1980 to develop a large multipurpose platform, L-SAT, to be launched in 1984, and to experiment new services. Technological viability and users interest is also to be ascertained.

Four different missions have been chosen, for the first flight, namely:

- 20/30 GHz communication experiments
- Direct Television Broadcasting at 12/18 GHz
- Specialized Services at 12/14 GHz
- 20/30 GHz propagation experiment

Italy is naturally interested in the 20/30 GHz communication systems; furthermore the development of a TVB system is also in line with the long term goals of the National Space Program, which accounts for the important financing, by Italy, of these two payloads. Selenia has been selected as responsible for both italian payloads, while the user require-

ments are coordinated by Telespazio and RAI respectively for 20/30 GHz communication and TV broadcasting missions.

The 20/30 GHz communication payload will enable to perform different experiments and services demonstration:

- Point-to-point video teleconference using digital transmission at 8.448 mb/s. The mission requires two spot coverage areas (Fig. 6) and two repeater chains each one connecting one spot to the other.
- Multipoint video teleconference, again using digital transmission at the same bit rate. Three repeater chains are envisaged and the required coverage area is shown in Fig. 6.
- Tele-education.
- Communication experiments with very wide bandwidth signals (up to about 700 MHz).
- High speed data transmission up to 72 Mbit/sec for VLBI (Very Large Base Interferometry) experiments.

The provision of the above services to small terminals on the premises of private organizations is likely to lead to important reduction of business travels and economies in time and energy.

The TV broadcasting mission requirements call for providing two high power 12 GHz channels:

- a pre-operational channel primarily for the Italian territory;
- an experimental channel available for the European Broadcasting Union (E.B.U.) country members

The areas to be served are:

- a) Italy with the Italian channel (see Fig. 7) and,
- b) on a time sharing basis by repointing the transmit antenna on area going from Iceland down to the Canary Islands and across to Finland and Lebanon with the experimental European channel.

The access (up-link) to the repeaters is performed through a common wide-beam receive antenna for the EBU channel and through the same antenna (transmitting and receiving) for the Italian channel.

### The ASTP Phase 2

Phase 2 of ASTP, devoted to critical hardware implementation, is just started. The objectives of this phase are mainly concentrated on:

- a) Repeater
- b) Antenna

#### a) Repeater

Activities aim:

- 1) to prove the feasibility of a 20/30 GHz regenerative repeater operating at high bit rate
- 2) to demonstrate acceptable system performances and,
- 3) to provide confidence for subsequent steps in the hardware development.

The final outcome, planned for the end of 1982, will be two integrated and tested repeater chains, fully characterized with RF and digital signals.

The characterizing repeater equipments to be designed and developed during this phase are:

- An improved version of 40PSK direct modulator, by Selenia
- Two demodulators of different types will be manufactured: a differentially coherent by Selenia and a coherent by GTE, to experimentally verify the performance on both types.
- A digital test set will be developed by GTE.
- A parallel ASTP activity is devoted to the feasibility study and subsequent implementation of the base band switching matrix and its control. CSELT (Centro Studi e Laboratorio Telecomunicazioni) is in charge of the work under Selenia supervision.

b) Antenna

Present activities follow two paths:

- definition and development of a high gain multibeam antenna
- definition and development of critical hardware for the R.F. sensor

Both activities are performed by Selenia.

The multiple spot beam antenna contract aims to design and breadboard the key elements of an antenna of 4 mt. diameter capable to meet the requirements of a 20/30 GHz European trunk/video communication missions. The antenna will have the capability to provide a high number of contiguous beams with a great number of effective frequency reuses within the coverage area and small angular separation between beams, high isolation and high gain within coverage.

A parallel activity is foreseen for an R.F. sensor system capable to guarantee the required antenna pointing accuracy of 0.03°.

The RF sensor feasibility study will allow to identify the most suitable system with respect to pointing accuracy, implementation problems, reliability and flexibility for use with different missions.

Some peculiar aspects of a RF sensor system, applied to a multiple beam antenna, are to be evaluated such as: residual pointing errors due to satellite yaw axis perturbation, reflector thermal distortion and defocussing compensation.

A critical hardware breadboard activity is also foreseen; the error detector unit is to be developed within 1981.

The succeeding phase activities will allow the completion of critical item construction and then a validation of RF sensor system closed loop operation via both simulation and actual tests.

The Italsat Program

In the frame of the National Space Program a point-to-point telephony trunk satellite at 20/30 GHz is being studied for domestic needs.

The operational mission configuration of ITALSAT, to be launched at the end of this decade, is under definition as it is considered preparatory to the definition of required technologies to both the satellite and the earth segment.

A preoperational mission, aimed for a launch in the 1986, has been already preliminarily defined taking into account the representativity of an operational mission concept even if

with reduced capacity.

The preoperational satellite foresees three payload elements:

- 20/30 GHz communication repeater for trunking telephony and data
- 20/30 GHz repeater for specialized services
- 40/50 GHz propagation experiment

The 20/30 GHz communication mission is outlined in table 3. Regenerative repeaters, with the switching function performed at baseband, will be used. The 2 mt. antennas provide a coverage of territory of Italy as shown in fig. 8.

The required antenna pointing accuracy of 0.03° is met by the use of an Antenna Pointing Mechanism slaved to a R.F. sensor system.

The 20/30 GHz specialized service mission will see transparent double conversion repeaters: two chains each one capable to handle digital signals at about 25 Mb/sec are presently envisaged.

The Special Services coverage, shown in Fig. 9, requires a dedicated antenna.

The propagation experiment foresees the transmission of a 40 GHz phase modulated carrier and a 50 GHz unmodulated carrier over a European coverage reported in Fig. 10.

TABLE 3 Preoperational Italsat Key Characteristics

Number of antenna:	2 of 2 mt diameter
Number of spot beams:	9
Number of station/spot:	more than 1 station/spot
Access, modulation:	SS-TDMA; 40PSK
Bit rate:	120 mb/sec and 360 mb/sec; bit rate conversion on board is foreseen
Channel bandwidth:	83 MHz and 250 MHz
Channel number:	9, 1 channel per spot
Total occupied bandwidth:	1000 MHz, no need of frequency reuse

#### Activity on Perspective Application of S.H.F. Spaceborne Radiometers to Earth Observation

Feasibility study of passive radiometer at millimeter wavelength from 90 GHz to 180 GHz was financed by CNR (Consiglio Nazionale delle Ricerche) and performed by Selenia. The study, completed at the end of 1980, has been done starting from the identification of operational requirements and satellite constraints.

A possible mission outline and key characteristics of SHF passive radiometer for oceanography application was the outcome of the study.

#### CONCLUSIONS

Italian activities for space application of frequencies above 10 GHz have been summarized and some key achievements described.

These activities are performed in the frame of both the National Space Program and European projects.

It results that the Italian Industry is making up the capability to be in a competitive international position in this advanced field.

#### ACKNOWLEDGEMENT

The authors are grateful to Mr. G. Perrotta, L. Di Fiore, A. Florio, for their useful suggestions.

#### REFERENCES

- Alta Frequenza, Vol. XLVII n. 4 (1978). Special Issue on the SIRIO Program.
- Lopriore, M. and Perrotta, G.(1977). The Millimeter wave communication transponder for the H-SAT experiment. Symposium on Advanced Satellite Communication Systems.
- Manoni, G. and Perrotta, G. System trade off on 20/30 GHz multispot beam antenna requirements for a domestic communication satellite.
- Lopriore, M. and Manoni, G. (1981). The design of a 30/20 GHz regenerative payload for satellite application. Fifth International Conference on Digital Satellite Communication.
- Tirrò, S. (1981). Possibilities offered by digital satellite communications for national communications in a european country. Fifth International Conference on Digital Satellite Communication.

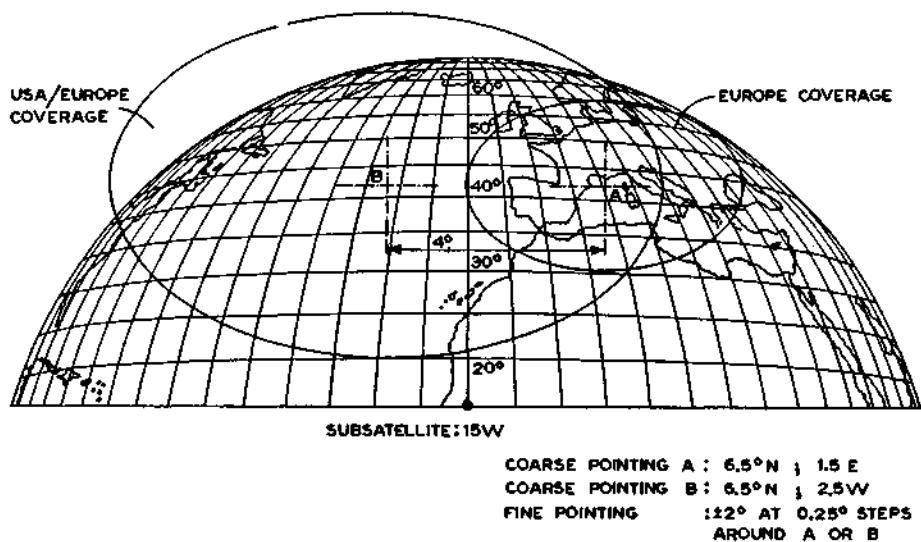


Fig. 1. SIRIO SHF Experiments coverage.

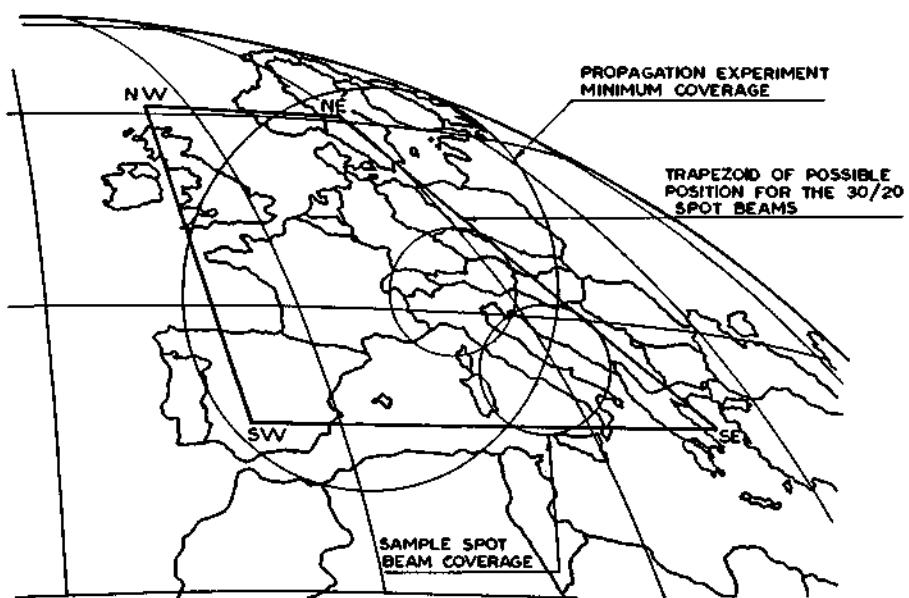


Fig. 2. H-SAT 20/30 GHz experimental payload coverage.

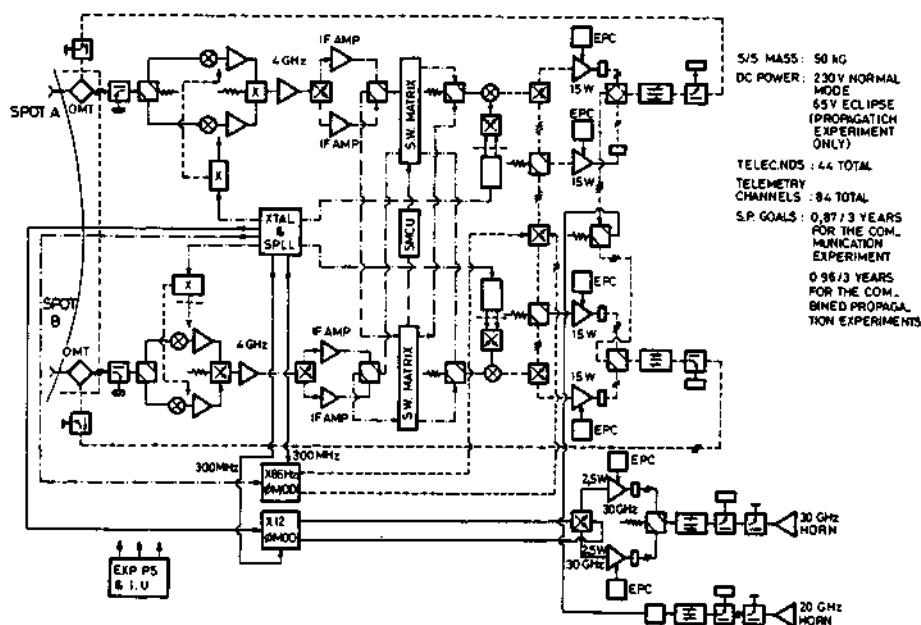


Fig. 3. H-SAT 20/30 GHz package block diagram.

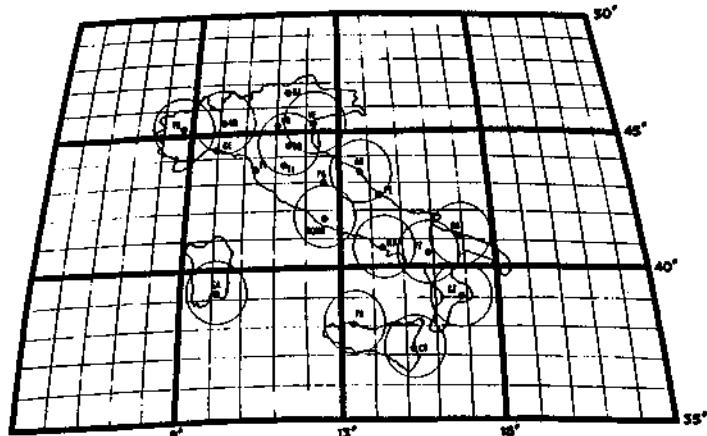


Fig. 4. ASTP 20/30 GHz Italian coverage

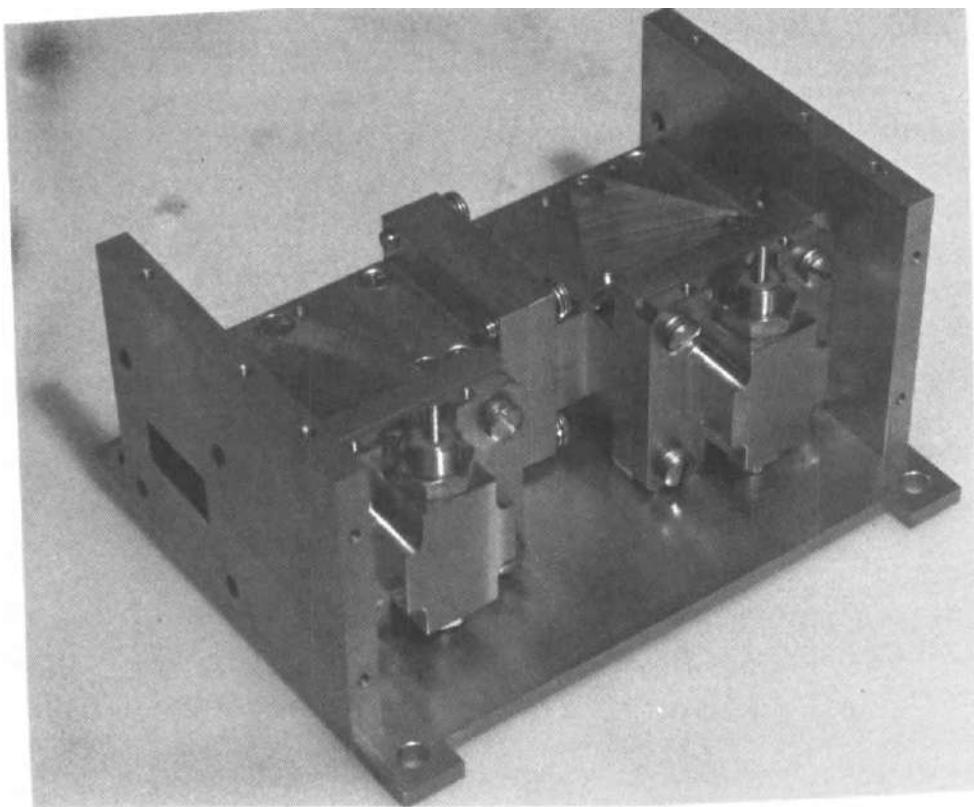


Fig. 5. ASTP - 40PSK modulator at 20 GHz.

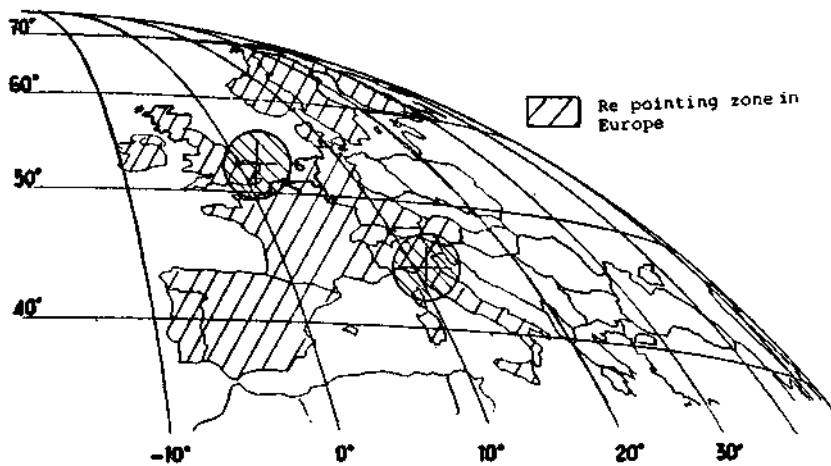


Fig. 6. L-SAT ~ 20/30 GHz communication experiment coverage.

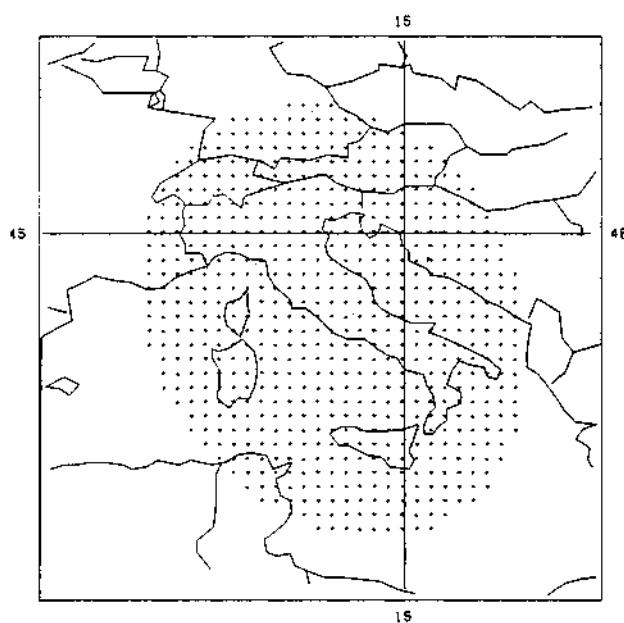


Fig. 7. L-SAT ~ TVBS coverage area of italian channel.

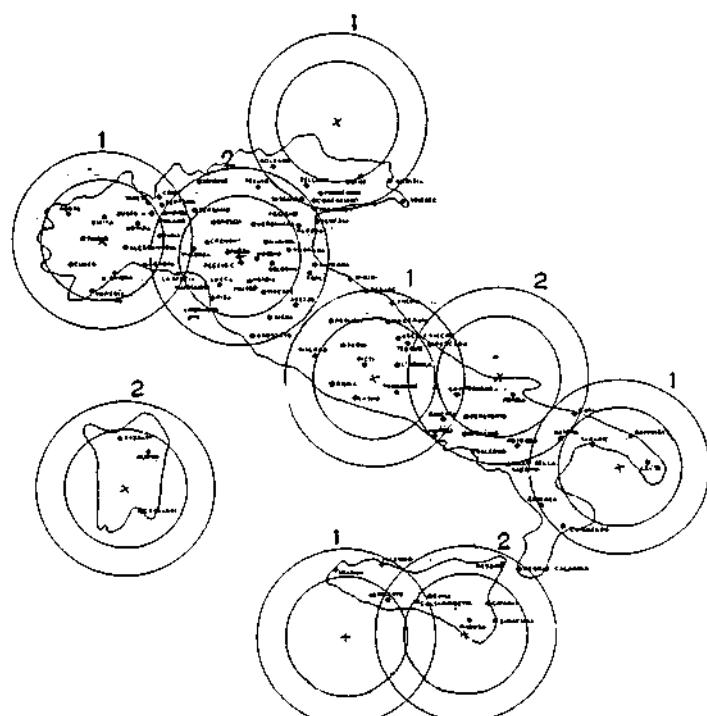


Fig. 8. ITALSAT - 20/30 GHz communication coverage.

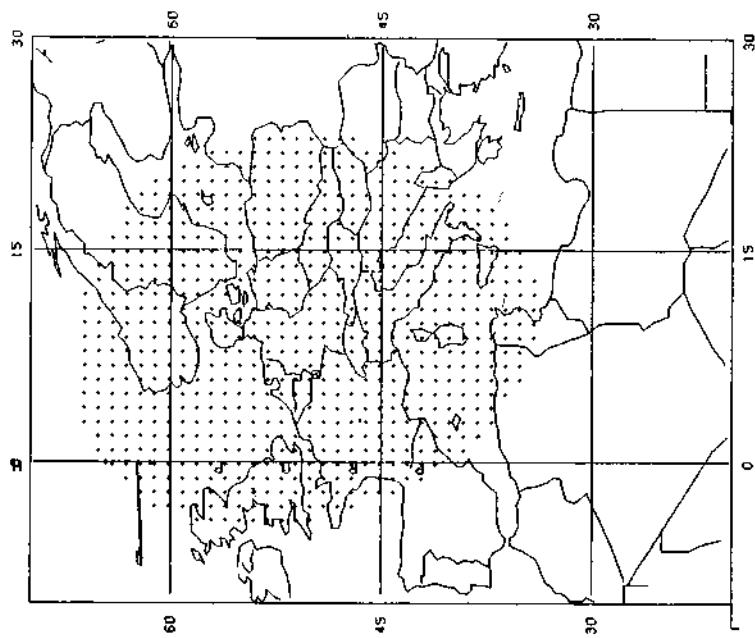


Fig. 10. ITALSAT- 20/50 GHz propagation coverage

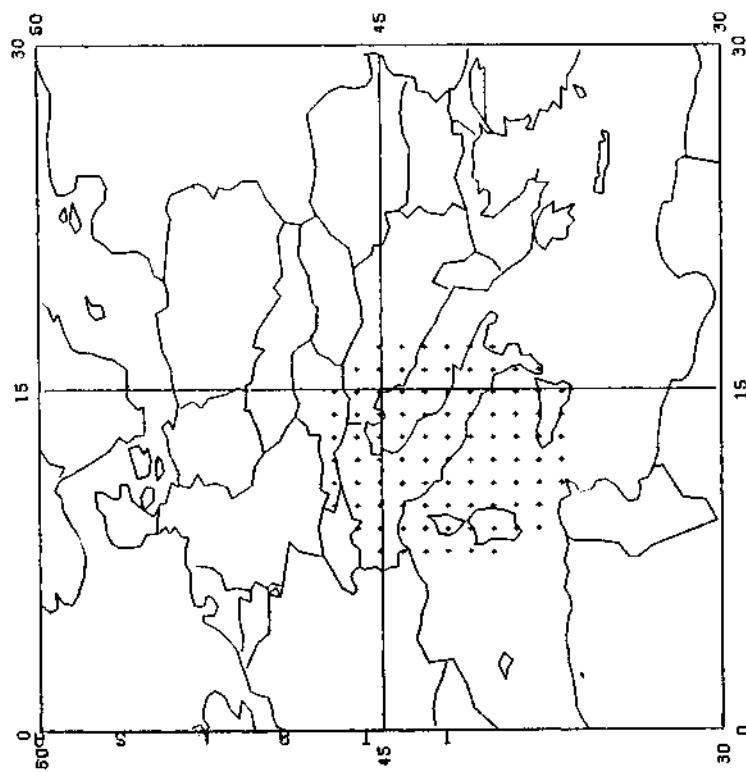


Fig. 9. ITALSAT- 20/30 GHz specialized service coverage

## INTELSAT V SYSTEM SUMMARY AND INITIAL LAUNCH OPERATIONS

J. T. Neer and C. F. Hoeber

*Ford Aerospace & Communications Corporation,  
Palo Alto, California, USA*

### ABSTRACT

This paper describes the technical and operational aspects of the first 2, in a series of 12, INTELSAT V satellites built by Ford Aerospace & Communications Corporation (FACC) for the International Telecommunications Satellite Organization (INTELSAT). The successful launch of the FM-2 and FM-1 satellites took place on 6 December 1980 and 23 May 1981, respectively. Following nominal transfer orbit operations, the satellites were placed into geosynchronous orbit by successfully firing solid apogee kick motors. After one drift orbit, the satellites were maneuvered and reconfigured for synchronous orbit operation. Transition to the three-axis-stabilized mode was successfully executed as planned. Following solar array and antenna deployments, a series of detailed functional tests were performed to validate the state of health and performance characteristics of the satellites. All redundant spacecraft hardware, mechanisms, and critical operational modes were verified to be working properly and all test results agreed quite well with prelaunch factory tests. These tests confirmed there had been no infant mortality during the launch and orbit injection process of either satellite. Communications tests validated that the critical payload elements (antennas, feeds, and transponders) had experienced no adverse effects by launch.

This paper will review the INTELSAT V satellite design, launch preparations, subsequent flight operations, and results of on-orbit testing. The paper will review the operational plans to transition from the INTELSAT IV-A to the INTELSAT V series and the trend in international communications satellites.

### KEYWORDS

INTELSAT V; communications satellites system; satellite operations; on-orbit testing.

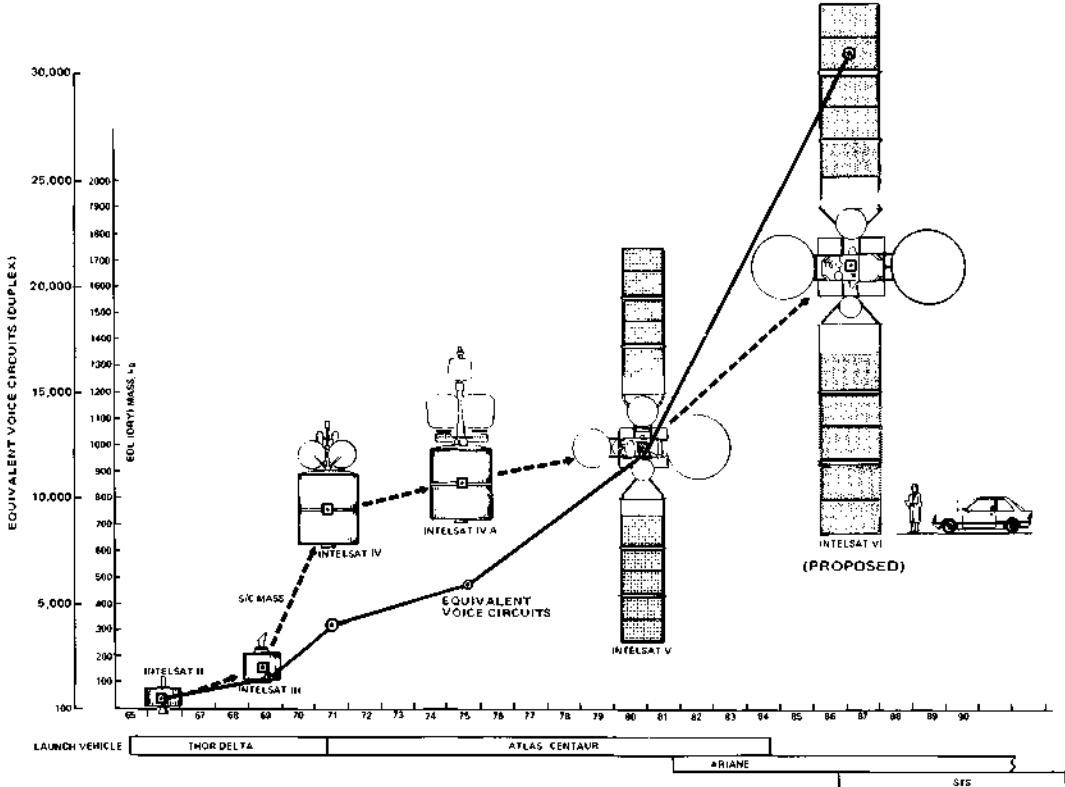
### INTRODUCTION

INTELSAT V is the latest in a series of commercial communications satellites designed to relay telephony and television signals among the member INTELSAT nations. The INTELSAT global network employs satellites in equatorial orbit above the Atlantic, Pacific, and Indian Oceans. Satellites in these three locations are capable of linking virtually all the inhabited areas of the world.

"This paper is based upon work performed under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT). Any views expressed are not necessarily those of INTELSAT".

The INTELSAT system has grown with the advances in communications satellites since the launch of INTELSAT I in 1965<sup>1</sup>. Ford Aerospace & Communications Corporation was selected as prime contractor to produce initially seven but subsequently twelve flight spacecraft for the INTELSAT V mission. The first spacecraft was delivered in October 1980, with launch on an Atlas-Centaur vehicle in early December.

INTELSAT spacecraft growth is illustrated by Fig. 1. The large electrical power requirements and the payload complexity of the INTELSAT V mission led to selection of the body-stabilized configuration shown. INTELSAT V will be the largest commercial communication satellite ever built. The solar arrays alone extend 15.7 m from tip to tip (the height of a five-story building), and the antenna tower looms 6.5 m above the spacecraft base.



*Fig. 1. INTELSAT evolution past, present, and future.*

Modular construction is employed to simplify assembly. The antenna module consists of seven communications antennas and six telemetry, command, and ranging (TC&R) antennas, mounted to a truss. All microwave communications electronics, filters, and switches are attached to the communications module, which constitutes half of the spacecraft body. Housekeeping functions are installed in the matching half of the spacecraft body, designated the support subsystems module.

New technologies have been balanced with flight-proven hardware to meet the major design constraints of reduction in mass and program risk. The communications receivers and traveling wave tubes are based on existing designs and technologies that have been used by FACC since 1970. The body-stabilized attitude control system draws heavily on the control electronics and momentum wheels flown on the Symphonie satellite. The solar arrays are derived directly from the Orbital Test

Satellite (OTS). To reduce mass, graphite fiber reinforced plastic (GFRP) materials are used extensively in fabricating the antenna reflectors, antenna feed array, feed support structure truss, microwave filters, waveguide, and solar array. A new contiguous-band output multiplexer reduces the number of 4 GHz transmit antennas from four to two. Electrothermal hydrazine propulsion, with a specific impulse greater than 290 seconds, is employed to provide north-south stationkeeping for extended mission lifetime.

### INTELSAT V Foreign Contributions

INTELSAT V was designed and manufactured by a strong association of international firms with enviable reputations for space hardware. Eleven companies from six foreign countries are supplying major spacecraft components. Particularly noteworthy is the association of six of these contractors that are formally teamed with FACC. INTELSAT V is truly an international venture; each team member shares in the orbital incentive payment plan. The role of each team member has been selected to emphasize the technical capabilities and strengths of that company. Each is responsible for design, manufacture, and testing of hardware. The team members and their contributions are summarized in the following paragraphs.

*Aerospatiale (France)*. Aerospatiale initiated the structural concept that forms the basis of the spacecraft modular design. It is supplying the main body structure and thermal control.

*Messerschmitt-Bolkow-Blohm (Germany)*. MBB designed and produced the attitude control subsystem and the solar arrays.

*Selenia (Italy)*. Selenia designed and built the six telemetry, command, and ranging antennas, two 11 GHz beacon antennas, and two 14/11 GHz spot beam antennas. It also built the command receiver and the telemetry transmitter, which combine to form a ranging transponder for determination of spacecraft position in transfer orbit.

*Mitsubishi Electric Corporation (Japan)*. Mitsubishi is responsible for both the 6 GHz and the 4 GHz earth coverage antennas, and also manufactures the power control electronics and the telemetry and command digital units from a FACC design.

*Thomson-CSF (France)*. Thomson built the 10 W, 11 GHz TWTs (10 per spacecraft).

*GEC-Marconi (England)*. Marconi produces the 11 GHz beacon transmitter used for earth station antenna tracking.

## SPACECRAFT DESIGN SUMMARY

The INTELSAT V spacecraft is a high-capacity, commercial communications satellite. Each satellite will be a radio-frequency relay, the space links in the vast INTELSAT communications network. As many as 6 INTELSAT V satellites will be operated simultaneously to interconnect more than 300 INTELSAT earth terminals. Depending on the operational configuration employed by INTELSAT, each satellite will carry up to 12,000 two-way telephone circuits and two color television transmissions.

The powerful communications transmitters, sensitive communications receivers, and r.f. upconverters require nearly 800 W of electrical power. Consequently, a large solar array area of nearly 20 m<sup>2</sup> is required to provide electrical power for the communications and supporting subsystems. The solar array area necessitates a body-stabilized spacecraft configuration with deployable, sun-oriented solar panels. The beginning of life (BOL) available power is approximately 1800 W with a conservative 7-year degradation estimate to about 1300 W at end of life (EOL). The total of the spacecraft (S/C) loads including bus housekeeping functions is about 1000 W.

The spacecraft three-axis-stabilized design is composed of a box-shaped main body 1.65 x 2.01 x 1.77 m, containing the electronics and propulsion subsystems, and a truss-type tower holding the

antennas. The spacecraft (Fig. 2) is oriented in space with the 2.01 x 1.77 m sides facing north and south. The solar arrays extend from this surface approximately 7.8 m each side of the spacecraft. The antennas are oriented with the large 4 and 6 GHz reflectors on the east and west sides.

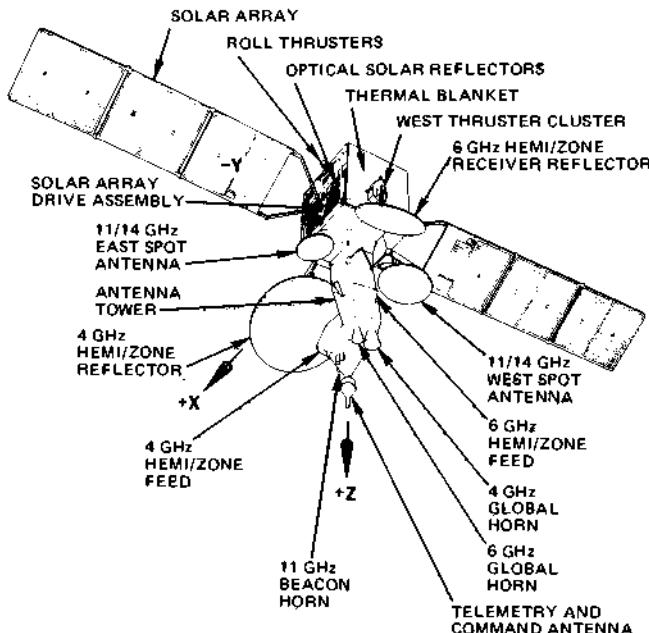


Fig. 2. Spacecraft configuration.

During synchronous orbit operation antenna pointing control is maintained to an accuracy of better than  $0.2^\circ$  in pitch and roll and  $0.5^\circ$  in yaw. Performance measured on orbit with FM-2 and FM-1 is noticeably better than this worst case prediction.

The S/C dry mass at EOL is about 825 kg for the nonmaritime payload version, i.e., FM-1 and FM-2. For these S/C, a hydrazine fuel load of 227 kg is carried to ensure operating lifetime in excess of 7 years. With a fully loaded apogee motor, the S/C weighs 1870 kg at launch.

A more detailed discussion of the spacecraft design including a description of all subsystems is found in references 1 and 2. Reference 3 shows in detail the communication subsystem characteristics and is an excellent presentation of how the requirements and technologies evolved and were ultimately realized in the highly successful FM-1 and FM-2 launches.

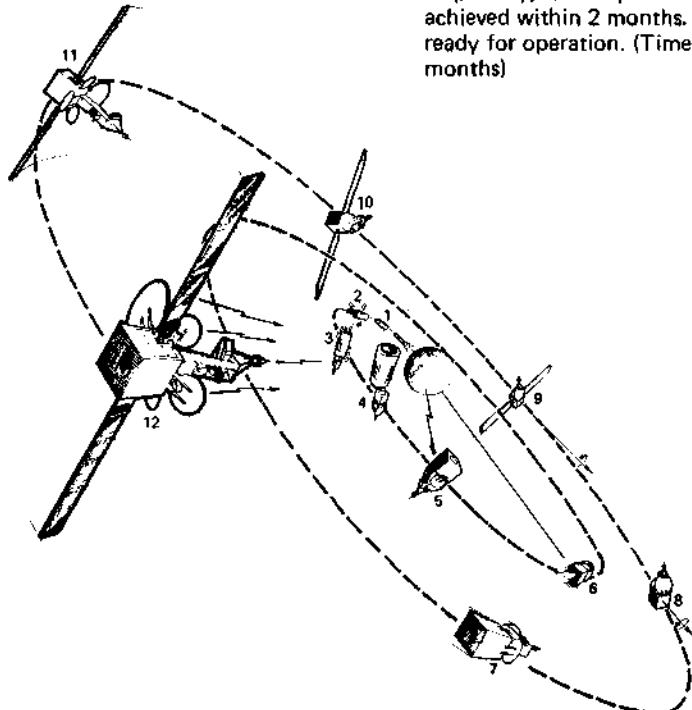
### SPACECRAFT OPERATIONS

Preparations for the launch operation began early in the program with FACC engineers working closely with INTELSAT and COMSAT operational personnel to develop the sequence of events (SOE) for the mission. Figure 3 shows the major events of the Atlas-Centaur launch. There are four distinct orbit phases:

- a. Launch and ascent into a low parking orbit
- b. Transfer orbit characterized by an 11-hour orbit period with a near synchronous altitude apogee
- c. A drift orbit phase to allow for quick station acquisition
- d. Final equatorial synchronous orbit at the desired test longitude

**MISSION SEQUENCE DESCRIPTION**

1. **Launch:** An Atlas Centaur launch vehicle will boost the satellite into a circular parking orbit. Launch times are restricted to approximately 2 hours around midnight GMT. (Time = 0)
2. **Jettison fairing:** During the first Centaur burn, the fairing is jettisoned exposing the spacecraft to a space environment. (Time = 3.5 min)
3. **Transfer orbit injection:** The second Centaur burn injects the spacecraft into an elliptical transfer orbit. (Time = 27 min)
4. **Separation:** Following attitude orientation, the spacecraft will be separated from the launch vehicle and spin stabilized. (Time = 29 min)
5. **Establish contact with Carnarvon, Australia:** Command and telemetry link is established. (Time = 43 min)
6. **Apogee motor ignition:** Apogee motor firing will input the spacecraft into a nearly synchronous drift orbit, 22,243 miles above the earth. (Time = 37 h)
7. **Orbit correction:** Velocity augmentation maneuver to circularize the orbit. (Time = 57 h)
8. **Sun acquisition:** Following despin, the spacecraft is placed in a sun-oriented stable attitude. (Time = 84 h)
9. **Deploy solar arrays:** Solar arrays are deployed and full array power is available. (Time = 85 h)
10. **Earth acquisition:** Spacecraft achieves complete 3-axis stabilized attitude. (Time = 87 h)
11. **Deploy antennas:** Communication antennas deployed. (Time = 93 h)
12. **Station acquisition:** With a drift rate of 1° per day, station position can be achieved within 2 months. Satellite ready for operation. (Time = 1 to 2 months)



*Fig. 3. Mission sequence description.*

Preparatory to the launch, FACC, working under contract for INTELSAT, developed detailed operational plans and procedures to support the launch of INTELSAT V. Since this was the first three-axis-stabilized satellite launched and operated by INTELSAT, a new chapter in satellite communication history had to be written. Just as in the early days of INTELSAT, a significant new stride in technology was being taken by the launch of INTELSAT V. There was thus a need for a well disciplined approach to the launch preparation, planning, and execution.

FACC delivered to INTELSAT documentation critical to the launch operation. These documents included an Orbital Operations Handbook (OOH) and a detailed set of contingency plans. The OOH described operational procedures and constraints necessary to properly operate the satellite. The contingency plans addressed possible "what if" conditions and developed the logic flow through the problem resolution process as well as specifying the corrective action.

The contingency plans were extensively reviewed along with the SOE in prelaunch rehearsals. The rehearsals were planned and executed to simulate the critical launch timeline and ensure adequate allowance for spacecraft diagnostics and problem resolutions. The rehearsals were held at the COMSAT Launch Control Center (LCC) leased by INTELSAT for the INTELSAT V launches. These rehearsals proved invaluable in familiarizing key spacecraft specialists with the telemetry processing and data display formats that would be available during launch.

The INTELSAT V satellite is injected into a transfer orbit similar to all previous INTELSAT satellites. The FM-1 and FM-2 satellites were successfully launched by Atlas-Centaur launch vehicles. The spacecraft is first acquired by a telemetry, tracking, and command (TT&C) station at Zamengoe in West Africa at or near spacecraft separation. Data from Zamengoe were used to confirm separation and initiation of spinup. For both launches, Zamengoe performed well and provided critical data for specialists at the LCC to be confident that the S/C was operating nominally on its own following separation.

Approximately 20 minutes after Zamengoe acquisition, the Carnarvon Australia TT&C station acquires and tracks the S/C for the next several hours. At Carnarvon acquisition, the S/C state of health (SOH) is confirmed and in the case of both FM-1 and FM-2, no anomalous S/C condition or behavior was detected.

Subsequent to Carnarvon acquisition and some preliminary S/C reconfiguration, an attitude change is performed to reorient the spacecraft to the apogee motor firing attitude. Prior to the reorientation maneuver, the S/C is in an attitude normal to the transfer orbit plane to collect attitude sensor data. All S/C maneuvering prior to the apogee motor fire (AMF) was performed nominally on both S/C and the subsequent transfer orbits were used to refine both the S/C attitude and orbit determination prior to AMF. A perigee raising maneuver was executed to minimize the attitude disturbance during perigee passage. Dynamic stability of the S/C was checked and determined to agree extremely well with prediction and test. The spacecraft nutation divergence time constant was measured to be  $25 \pm 2$  minutes, well within the measurement accuracy of the flight and ground test.

For both FM-1 and FM-2, AMF occurred on fourth apogee since these S/C were programmed for the Atlantic Ocean region (AOR). The AMF on both missions was quite nominal, exceeding the required delivered velocity increment. TABLE 1 summarizes the critical performance of the FM-2 S/C through AMF. FM-1 flight data compared quite well with the FM-2 data. Both S/C were within the  $1\sigma$  performance estimates.

Due to nominal performance and conservative assumptions in both design and analysis, the fuel used to place the S/C into operational orbit was less than predicted. The projected lifetime of both FM-2 and FM-1 exceed the predicted 7-year requirement. With fully operational electrothermal thrusters the lifetime should exceed 10 and 8 years respectively.

Following the AMF, the S/C is reoriented back to orbit normal attitude to provide continuous TT&C coverage throughout the drift orbit. During the first drift orbit, all of the acquisition electronics, sensors, and thrusters are checked out and verified to be working normally. These tests allow time to modify the SOE during the critical three-axis acquisition maneuver if any unit failed during launch. Since no hardware failures occurred on either FM-1 or FM-2, the nominal SOE was followed. Following a maneuver to reduce the drift rate and circularize the drift orbit, preparations

TABLE 1 Systems Analysis

Transfer Orbit Trajectory	Predict	Actual	Comment
Apogee radius (km)	$42166 \pm 296$	$42105 (-61)$	
Perigee radius (km)	$6545 \pm 5.6$	$6547 (+2)$	
Inclination ( $^{\circ}$ )	23.75	23.93	
Separation/spinup			
Attitude error ( $^{\circ}$ )	6.4 (3 $\sigma$ )	2.2	
Spin speed (r/min)	$47.5 \pm 5$	46.5	
Final nutation ( $^{\circ}$ )	0.5 (3 $\sigma$ )	0.18	
Spin stability			
Diverging time constant, minimum	25-30	$25 \pm 1$	
ANCE firings per hour	25 max	2-3	
Spin wobble ( $^{\circ}$ )	0.35	0.25	
From attitude data			
AMF performance			
Fire time, s	42	42	
$\Delta V$ , ft/s	$6000 \pm 60$	6020	
Attitude error ( $^{\circ}$ )	2.1 (3 $\sigma$ )	0.3 $^{\circ}$	

Launch: 2330 Z 6 Dec 81, FM-2

5 minutes into first launch window

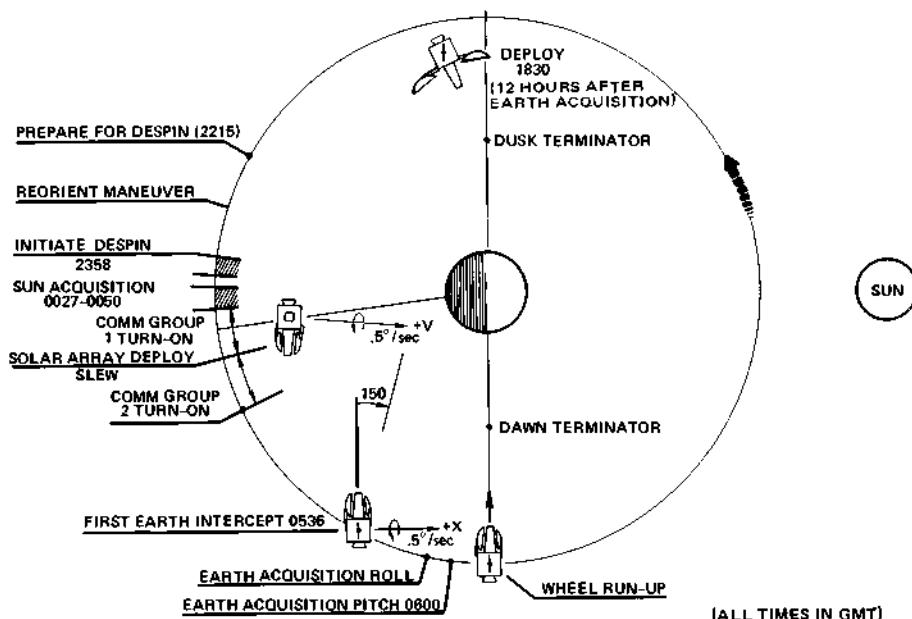


Fig. 4. Dawn acquisition scenario.

to transition to three-axis stabilizations were made. Figure 4 shows the acquisition scenario used on both FM-1 and FM-2. The acquisition begins around satellite midnight with the ground command to despin the satellite to about 20 r/min. Initially the satellite spins about 45 r/min for gyroscopic stability during AMF. When 20 r/min is reached, rate damping mode is commanded and the attitude determination and control system (ADCS) enters a closed loop mode of control using rate sensors (angular rate assemblies, ARAs) to achieve a fixed body rate of 0.5°/s about the pitch and roll axes. During this despin maneuver on both FM-1 and FM-2, the dynamic stability was so well behaved that continuous S/C telemetry was received throughout the event.

At the end of rate damping, the sun acquisition mode is ground commanded. The S/C then performs a three-phase sun acquisition maneuver to point and maintain the S/C +X axis toward the sun in preparation for the dawn acquisition.

Following sun acquisition the communications equipment is powered up to achieve the desired thermal balance for three-axis synchronous orbit operation. Immediately following turnon of half of the communications equipment, the solar arrays are deployed and slewed to the sun line. At this point power is removed from the batteries and placed on the solar arrays. Following solar array deployment the remaining communications equipment is turned on to ensure proper thermal equilibrium of the spacecraft.

The spacecraft power, thermal, and control functions are now in an extremely stable mode. Several hours later the infrared earth sensors begin to intercept the earth as the S/C approaches the dawn terminator. After properly configuring the S/C for earth acquisition, the roll acquisition is initiated close to the dawn terminator. The S/C removes the 0.5°/s roll rate and locks on to the earth. Subsequently the earth is acquired in pitch. At this point the momentum wheel is spun up and provides the desired gyroscopic stability. After wheel runup to nominally 3500 r/min, the S/C is configured for normal synchronous orbit operation. The communications antenna's are deployed after a period of thermal conditioning. At the end of antenna deployment the S/C is ready for functional tests. At this point in the mission, the satellite is ready for communications operation.

Both FM-1 and FM-2 reached this critical initial orbit configuration (IOC) within 4 days of liftoff. The smoothness and success of the missions are directly associated with a conservatively designed S/C and a well planned and thoroughly rehearsed flight plan.

#### Bus Functional Tests

Immediately following IOC the S/C is subjected to a series of tests. All redundant electronic, sensors, actuators, and critical telemetry and command functions are verified. These tests on FM1 and FM-2 disclosed no hardware failures or mission limiting anomalies.

The tests included controls gain, mode logic, and redundant unit tests. Antenna positioners were moved and redundant solar array drive electronics (SADE) were actuated. All spacecraft power and thermal management functions were checked out and found to be performing nominally. Preparations were made for the detailed communications tests and all bus tests were completed.

### IN-ORBIT COMMUNICATION TEST

#### On-Orbit Payload Performance

Prior to entering service, the INTELSAT V FM-2 and FM-1 communications payloads were the subject of a series of performance measurements from a specially built ground station in Fucino, Italy. The results of the performance measurements are summarized in TABLE 2; a more detailed description of the ground and in-orbit test program follows:

**TABLE 2 Communications Performance Summary**

Primary Parameters	Comment
Gain G/T EIRP	Generally within 0.5 dB of prediction Exceptions are probably due to test conditions at time of test (all within 1dB)
<b>Passband Characteristics</b>	
Amplitude Group delay	Nominal to better than ground measurements
<b>Secondary Functions</b>	
All switches Receiver attenuators Multiplexer attenuators L.O. frequencies Beacons	Nominal performance
<b>Antenna Parameters</b>	
Main beam Sidelobe C/I Crosspol C/I	Main beam within measurement accuracy of predictions. Sidelobe and crosspol nominal.

**Payload Description.** The INTELSAT V payload, illustrated in simplified form in Fig. 5, consists of 7 antennas that provide separate receive and transmit beams plus an 11 GHz beam for ground station tracking and the electronic hardware for 27 simultaneously operating transponders. Although redundancy is not shown in Fig. 5, there are a total of 43 TWTAs, of which 27 are normally operating; 15 receivers, of which 7 are normally operating; and 10 upconverters, of which 6 are normally operating. The switch matrices consist of 53 transfer and SPDT coaxial switches and 6 hybrid power splitters that allow interconnection of all beams on a channel by channel basis. The switch configuration allows any interconnection to be made even in the event of a single switch failure per channel.

The transponders, which operate in both the 6/4 GHz and 14/11 GHz bands, vary in bandwidth from 34 to 241 MHz and provide a total usable bandwidth of 2137 MHz through four-fold frequency reuse at 6/4 GHz and two-fold frequency reuse at 14/11 GHz. The multiple frequency reuse requires state-of-the-art control of antenna sidelobe and cross-polarization performance as well as spacecraft attitude determination and control. Tests verifying this performance will be discussed in detail.

**Objective of Tests.** A great deal of importance was attached to the performance measurements of the first INTELSAT V flight spacecraft. There are two fundamental objectives of the test:

- To verify that payload hardware is unchanged due to the launch and space environment.
- To validate the prelaunch data base for on-orbit traffic modeling and traffic assignment.

**Test Summary.** As indicated in TABLE 2, the correlation to predictions based on ground test results is superb. Tests designed to isolate the performance of every active component, all microwave filters, and all switching functions indicate that the satellite performs exactly as it did in prelaunch tests. Antenna performance and alignment is nominal and pointing stability measurements to date indicate an accuracy of a factor of two better than budgeted. The prelaunch data base has been validated and INTELSAT expects to achieve an in-service capacity equivalent to performance 2 dB better than specification level in all channels.

## INPUT MULTIPLEXERS      TWTAs      OUTPUT MULTIPLEXERS

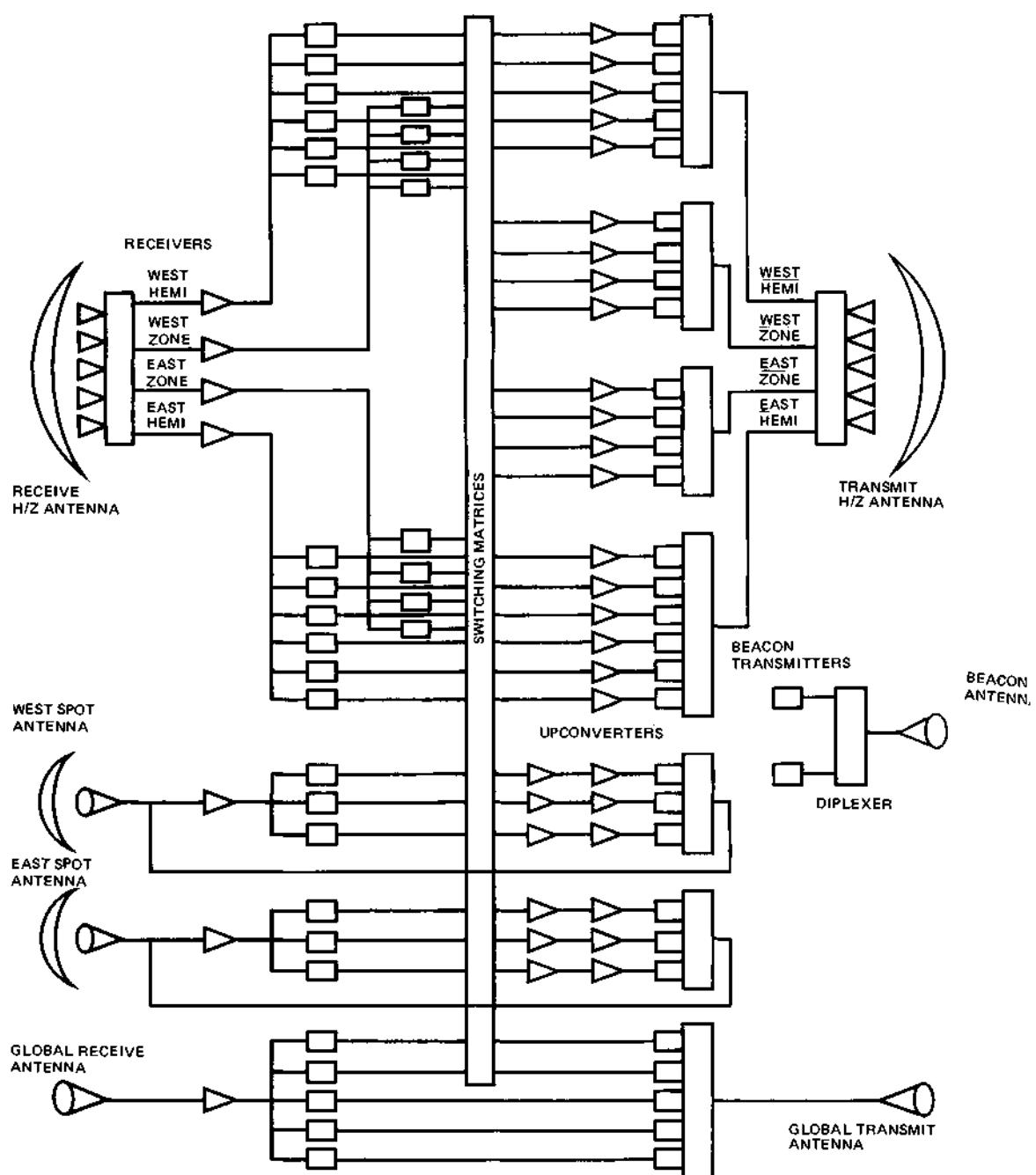


Fig. 5. Simplified INTELSAT V communications payload block diagram.

One minor anomaly is the spontaneous tripoff of three 11 GHz TWTAs. This was not unexpected due to a ground history of such events. One known cause of such events is a microdischarge within the tube, and several design improvements have been made to minimize the occurrence of such events effective on the third INTELSAT V to be launched. Investigation is also under way (utilizing the prototype spacecraft) of the possibility of turnoffs induced by bus voltage variations. The first two on-orbit tripoffs occurred approximately 2 weeks after launch, the third approximately 6 months later. All three TWTAs were subsequently checked and found to be operating normally. It is worth noting that the FM-2 spacecraft has passed through a period of very intense solar activity with no evidence of any static discharge events and no perturbations of digital circuitry which have occurred on previous INTELSAT spacecraft.

Prelaunch Tests. A major reason for the on-orbit success is FACC's thorough ground test program. Both the antenna and transponder subsystems go through extensive thermal tests prior to their final acceptance tests. Not only is the axial ratio of every feed element measured over temperature, but the entire 90 element hemi/zone feed arrays are measured over 10 thermal cycles. The minor calculated variation in antenna gain at all INTELSAT ground stations is computed.

After antenna module integration, the performance of all antennas is measured at FACC's Calaveras test site. The critical hemi/zone antennas have their co-polarized and cross-polarized radiation contours mapped in half degree increments over the earth's field of view at 40 discrete frequencies (1,613,760 data points). In addition, the swept frequency performance is measured at critical sidelobe and cross-polarized locations.

Transponder testing is similarly thorough. A week long thermal test thoroughly exercises all components and the hundreds of coaxial interconnections. There are 588 discrete transponder paths, and test matrices are established such that every possible path is tested in a logical fashion so that any anomalous performance can be immediately traced to a specific component.

Once the spacecraft is integrated, the key performance tests are the reference and final slant range tests performed in a 250 ft antenna range adjacent to FACC's assembly building, and the extensive thermal vacuum performance measurements. The key to the smooth operation of the spacecraft level tests is that all measurements are subsets of antenna or transponder acceptance tests, using modifications of the subsystem software, and with identical measurement techniques and display formats. For example, at the slant range the hemi/zone contours are measured at the center frequency of each channel; a subset of the original 40 frequencies measured at Calaveras. By utilizing common software, the subsystem data base is easily loaded into the system test computers for instant data comparison and trend analysis.

On-Orbit Communications Measurement. The on-orbit measurements can be split into those that are primarily transponder related, and those that are strictly antenna related. In the first group are the primary parameters that determine spacecraft capacity; namely, gain, G/T, and EIRP. TABLE 3 presents the EIRP measurements at 4 GHz. In the interest of brevity, only the results for the primary TWTAs are shown. For the zone beams, both the Atlantic/Pacific and Indian Ocean configurations were tested; the results shown are for the Atlantic/Pacific configuration.

As can be seen, the correlation is excellent, well within the measurement accuracy for both the ground and flight data of  $\pm 0.5$  dB. Note that the global beam correlation is somewhat better than the hemi/zone data. This is because the beam is broader and the measurements are more tolerant of pointing uncertainties during the offset pointing required to make measurements from Fucino. All measurements are performed near the beam peak in the plane containing Fucino ( $6.5^\circ\text{N}$ ). No attempt is made to compare the results directly to the edge of coverage specification since this is done prior to launch and the intent during the on-orbit tests is primarily to look for changes.

TABLE 3. Primary Parameter Measurometer Summary

Beam	Channel	EIRP, 4 GHz			Flux Density to Saturate 6 GHz (1)		G/T, 6 GHz		
		Measured EIRP dBW	Predicted EIRP dBW	Spec EIRP dBW	Measured Gain dBW/m <sup>2</sup>	Predicted Gain dBW/m <sup>2</sup>	Measured G/T dB/K	Predicted G/T dB/K	Spec G/T dB/K
West Hemi	1-2	33.4	33.13	29	-74.8	-75.1	-5.3	-5.8	-11.6
	3-4	32.8	33.1	29	-74.8	-74.9	-5.7	-5.9	-11.6
	5-6	33.2	33.05	29	-74.3	-74.9	-5.7	-5.9	-11.6
	7	32.5	32.71	29	-74.6	-74.1	-5.8	-6.5	-11.6
	8	32.2	32.76	29	-74.4	-74.8	-5.7	-6.1	-11.6
	9	29.4	30.04	26	-77.5	-77.4	-6.9	-6.4	-11.6
East Hemi	1-2	33.2	33.13	29	-76.1	-76.5	-5.5	-5.1	-11.6
	3-4	33.1	33.27	29	-75.4	-75.7	-5.8	-5.6	-11.6
	5-6	33.1	33.37	29	-76.1	-76.1	-5.6	-5.3	-11.6
	7	32.4	32.77	29	-75.9	-75.6	-5.4	-5.1	-11.6
	8	32.3	32.41	29	-75.7	-75.2	-5.6	-5.4	-11.6
	9	28.9	28.99	26	-78.9	-78.4	-6.0	-5.5	-11.6
West Zone	1-2	34.5	33.87	29	-74.0	-74.3	-5.1	-4.65	-8.6
	3-4	34.6	34.12	29	-74.2	-75.0	-4.8	-4.28	-8.6
	5-6	34.9	34.66	29	-74.7	-75.2	-4.6	-4.27	-8.6
	7-8	34.9	34.51	29	-74.4	-74.7	-5.0	-4.95	-8.6
East Zone	1-2	33.1	32.39	29	-74.2	-74.8	-3.5	-3.4	-8.6
	3-4	32.9	32.03	29	-74.6	-75.0	-3.4	-3.2	-8.6
	5-6	32.6	31.99	29	-74.4	-74.5	-3.0	-3.0	-8.6
	7-8	32.5	32.21	29	-73.6	-74.3	-3.2	-3.1	-8.6
Global	7-8	29.0	28.9	26.5	-74.9	-75.1	-12.0	-13.04	-18.6
	9	26.1	26.3	23.5	-78.5	-78.5	-12.2	-13.0	-18.6
	10	26.9	26.8	23.5	-79.0	-78.2	-11.9	-12.61	-18.6
	11	27.1	26.8	23.5	-79.3	-78.5	-12.1	-12.76	-18.6
	12	26.8	26.8	23.5	-79.1	-78.5	-12.1	-12.76	-18.6

NOTE 1: Spec value is a nominal at edge of coverage,  
not a minimum. It is not shown since it  
is not applicable at beam peak.

Some difficulties were encountered in making 14/11 GHz spot beam measurements. The IOT antenna used for these tests experienced calibration difficulties and the measurements were performed during periods of snow, which causes a significant loss at these frequencies. Therefore, no spot beam data are presented here. It should be noted, however, that the spot beam channels have been successfully carrying traffic for some time and there is no indication that the spacecraft is other than nominal.

TABLE 3 also presents the measured gain and G/T at 6 GHz. Again only the primary receiver and TWTA are shown. In addition to the data presented herein, in-band and out-of-band amplitude response and group delay measurements were performed in every channel. Sample plots are shown in Fig. 6. Typical in-band response was better than prelaunch, probably due to nearfield effects on the slant range. Typical group delay measurements, which included doppler compensation, were direct overlays of prelaunch data. Typical out-of-band responses were also direct overlays of the prelaunch data, within the 50 dB dynamic range of the IOT equipment.

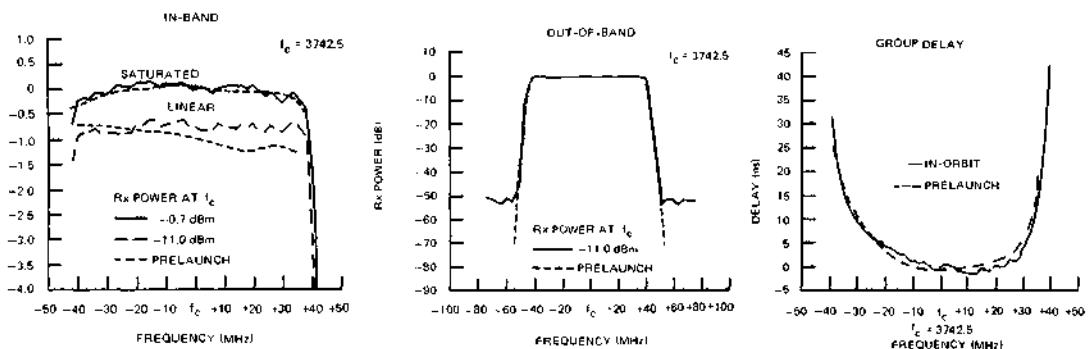


Fig. 6. Swept frequency measurements.

**Antenna Performance.** The antenna performance must be considered a primary performance parameter determining the satellite capacity, along with gain, G/T, and EIRP. In fact, as available frequency band reuse increases, modern telecommunication satellite capacity is limited more by antenna carrier-to-interference (C/I) performance than carrier-to-noise (C/N) ratios. Over half of the on-orbit payload test time was devoted to verifying the performance of the hemi/zone antennas.

The hemi/zone shaped beams, illustrated in Fig. 7, are formed by 90 element feed arrays fed by an airline feed network with arbitrary amplitude and phase at each feed. To measure the antenna patterns over the entire field of view, a series of roll maneuvers, illustrated by Fig. 8, was performed. Each day, a roll maneuver was performed at time  $t_0$  with the spacecraft yaw and roll control loops disabled. Over a 24-hour day, an undisturbed spacecraft would exhibit a sinusoidally varying roll error. Solar pressure torques, however, modify this behavior, and a detailed computer program is required to predict the spacecraft motion. As shown in Fig. 8, there is a period 12 hours after the initial maneuver where the roll angle is nearly constant and yaw is near zero. During this period, the spacecraft is stepped in half degree pitch increments to obtain antenna patterns versus pitch angle, for a given roll offset. For Flight 2, the patterns at five roll angles were obtained. Main beam co-polarized and crosspolarized patterns as well as sidelobe patterns were obtained for each beam. The satellite control system is designed to allow maintenance of spacecraft attitude with large roll biases; however, the above measurement sequence was employed to eliminate any extra fuel expenditure.

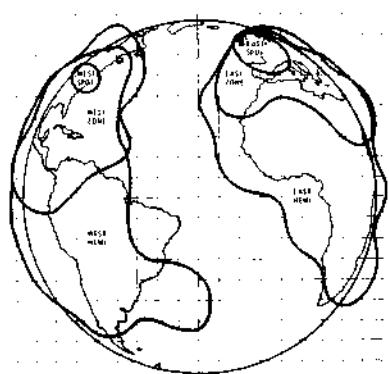


Fig. 7. Antenna pattern results.

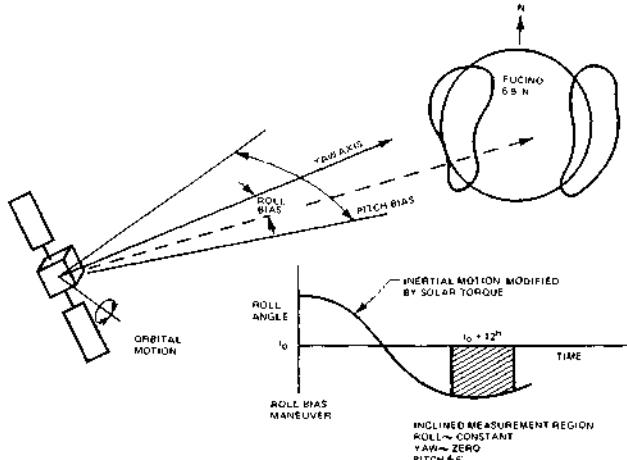


Fig. 8. Communication antenna mapping technique.

Before commenting on the test accuracies, it is important to note the test limitations. During the course of a pattern measurement, roll and yaw could vary by nearly  $0.5^\circ$ . Accounting for the angular variation at every data point is a monumental data reduction task and has not been performed although the IOT software has been modified to perform this task for FM-1. Within this limitation, all patterns agree with prelaunch predictions within  $\pm 0.25^\circ$ . The variations from pattern to pattern appear to be random, and there are absolutely no indications of any alignment biases. A typical co-polarized main beam plot is shown in Fig. 9.

Typical measured cross-polarized and sidelobe performance are also shown in Fig. 9. Minor out of specification values were recorded, but it should be noted that data taking was concentrated in the areas of anticipated specification level performance. In evaluating C/I data, numerous measurement error sources must be considered. For example, at the specification level, the ground station axial ratio of 0.15 dB can contribute over  $\pm 3$  dB of error to cross-polarized C/I measurements. The discrepancies of Fig. 9 are within the bounds of this error source alone.

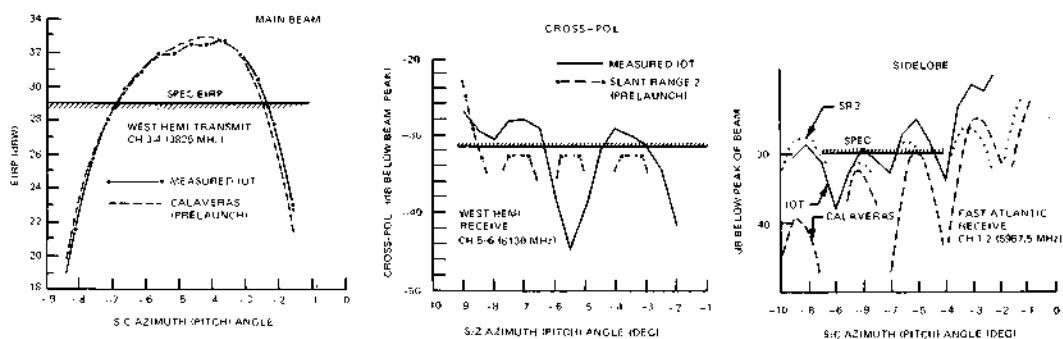


Fig. 9. Antenna pattern results

Figure 9 compares sidelobe performance to both the final spacecraft slant range tests and the antenna acceptance test data at Calaveras. All measurements were evaluated with respect to both prelaunch data sources. The slant range data account for certain minor contributors to antenna C/I, such as intratransponder leakage, although they suffer some degradation due to nearfield effects. Exhaustive data analysis leads to the conclusion that at the specification level, either data base (supplied on computer tape to INTELSAT) can be used with an accuracy of  $\pm 3$  dB at specification level.

The foregoing discussion exemplifies that state-of-the-art frequency reuse satellite antennas are limited as much by ability to perform measurements as by theoretical antenna analysis techniques. FACC has identified the error sources in its antenna ranges and is systematically eliminating them for the design of the next generation of frequency reuse antennas.

**Platform Stability Tests.** Prior to launch, considerable concern over the effects of spacecraft motion was expressed due to the large gain slopes at the antenna edge of coverages. To verify stability, a test was performed in which two r.f. signals were passed through the satellite. The first signal was in the east hemi beam, which has a nearly east-west gradient as seen from Fucino. Any pitch motion would therefore be detected as an r.f. signal variation. Similarly, the second signal was in an east spot beam channel and the spot beam antenna was steered so that its gradient was north-south as viewed from Fucino. This signal served as a roll motion detector. To eliminate any atmospheric effects or effects of test equipment drift, these signals were displayed on a strip chart alongside the 4 and 11 GHz beacon signals. The beacon antenna beams are very broad and thus insensitive to small spacecraft motions. This test was performed during stationkeeping maneuvers and eclipse.

Measured stationkeeping errors agreed with ADCS telemetry and were well within budget. Absolutely no signs of short or long term thermal effects or alignment biases could be detected during the test. Communication performance in and out of eclipse in the Spring of 1981 showed no evidence of thermally induced pointing transients.

### INTELSAT V OPERATIONAL DEPLOYMENT

With the success of both FM-2 and FM-1, INTELSAT has targeted for an operational transition to the INTELSAT V service in August. Figure 10 shows the deployment schedule for both satellites. INTELSAT V FM-1 and FM-2 were spatially collocated in the orbital arc with INTELSAT V-A FM-1 and FM-2 respectively, at the end of the FM-1 IOT program in June 1981. The collocation permits minimum disruption to the customer as transponder reassessments are made. By positioning both IV-A and V satellites to within  $0.06^\circ$  of each other, the spacecraft will be within the ground antenna beamwidth at the same time. The collocation maneuver was practiced on the Indian Ocean INTELSAT IVs and demonstrated to be practical.

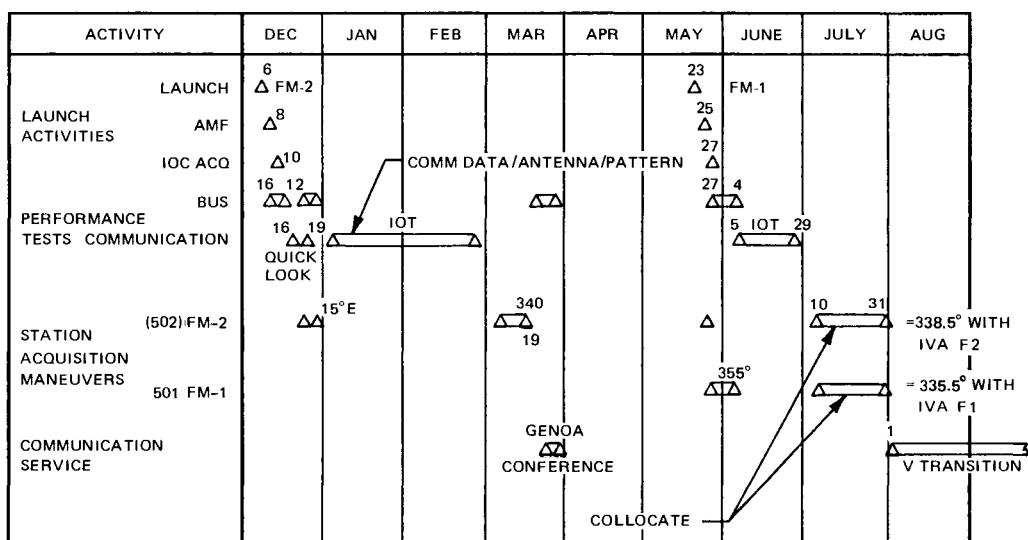
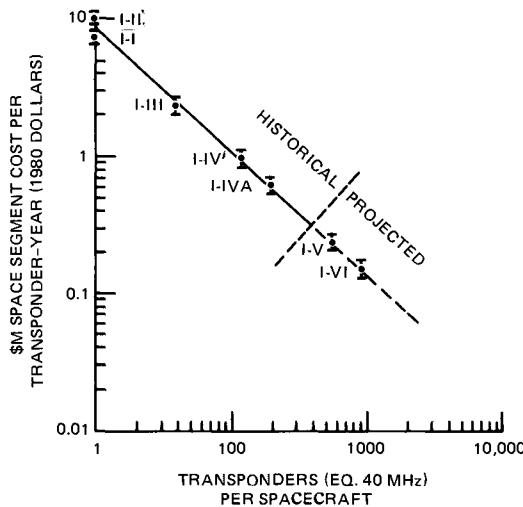


Fig. 10. INTELSAT V initial orbit configuration (IOC) schedule.

### INTELSAT EVOLUTION AND FUTURE

The present INTELSAT V satellite series will clearly establish a new standard for satellite communications. Perhaps Fig. 1 best portrays what significant strides have been made in a relatively short time span. In less than one human generation, i.e., 25 years, the INTELSAT series will have evolved through six generations. Presently, INTELSAT is in evaluation of the INTELSAT VI proposals. When the program is awarded, it will pave the way for a single geostationary platform to carry nearly 33,000 equivalent voice circuits. This increase of over two orders of magnitude in capacity from INTELSAT II has been made possible by the advanced booster capabilities and satellite system technologies discussed in the introduction.

Perhaps more dramatic is the consistent rate reduction for the INTELSAT lease channel service. Figure 11 shows the reduction of costs in 1980 dollars. Thus, even in the face of spiraling inflation, satellite communication service has experienced a cost trend unprecedented in today's economic environment.



Note: From General Dynamic Study, NAS 8-33527,  
 "Geostationary Platform System Concept  
 Definition Study," June 18, 1974.

*Fig. 11. INTELSAT series demonstrates economy of scale in communications spacecraft.*

Much has been said about the concept of economy of scale, but in few other industries has this economic principle been more clearly demonstrated. INTELSAT V will soon add on FM-5 the maritime payload for use by the International Maritime Satellite (INMARSAT) Organization. This auxiliary payload will add L-band for ship-to-shore and shore-to-ship communication service. Starting with FM-10, INTELSAT V will expand its capacity once again to 15,000 equivalent voice circuits and thereby bridge the gap until INTELSAT VI is launched in 1986. FM-10 INTELSAT V-A will incorporate dual polarized global communications service and the addition of C-band feeds to illuminate the existing K-band steerable reflectors. This capability, a straightforward low risk modification to the basic INTELSAT V satellite, will enhance the domestic lease capacity of INTELSAT.

In the immediate future INTELSAT VI is estimated to carry nearly 33,000 voice circuits. Through the use of satellite switched time division multiple access (SSTDMA) digital transmission techniques and six-time frequency reuse of the standard 500 MHz C-band slot, the satellite will certainly be one impressive "switchboard in the sky." Weighing twice that of INTELSAT V, launched on either the STS or Ariane 4, the satellite will dwarf its predecessors.

INTELSAT VI is therefore planned to fully exploit its immediate heritage and programmatic base. Cost-effective, low risk, long life, flexible, and adaptable to new and expanded requirements, INTELSAT VI will be an elegant instrument of man to effectively and efficiently utilize the limited and valuable natural resources, frequency spectrum, and geosynchronous orbital arc. Future programs must carefully consider these nonreplaceable resources when conducting advanced system tradeoffs. The era of dedicated, special purpose satellites is behind us. The future satellites will be large high-capacity multimission communication platforms. Until the time of cost-effective and reliable space construction and transfer to geosynchronous orbit, INTELSAT VI and its derivatives will dominate the communication satellite field well into the 1990s.

## CONCLUDING REMARKS

The first two INTELSAT V's have been successfully launched and placed into geosynchronous orbit. The spacecraft were reconfigured for routine communication service following transition to the three-axis-stabilized mode of control. All performance tests, both bus and payload, have been completed, demonstrating close agreement with prelaunch test data. All mission objectives have been met or surpassed. No significant deviations from the flight plan were required by hardware malfunctions. All spacecraft components survived launch. Spacecraft fuel lifetime remaining exceeds the 7-year requirement by 1 to 2 years. The success of FM-1 and FM-2, the first of a major new satellite series, has been noteworthy in the rapidly advancing communication satellite industry.

This paper has summarized many design and operational aspects of the INTELSAT V satellite series. To make INTELSAT V possible, many "firsts" have been utilized to meet INTELSAT's growing communication requirements. The remarkable success of the first INTELSAT V has marked a real advance in the space communication era. FACC and all of its team members are proud of this accomplishment.

## ACKNOWLEDGMENT

This work is based upon work performed under the sponsorship of the International Telecommunications Satellite (INTELSAT) Organizations. The outstanding success of the first two INTELSAT Vs will be long recognized as a major achievement in the maturing aerospace industry. Ford Aerospace & Communications Corporation and its team members are proud and privileged to have played such a major role in the program. The authors would like to acknowledge the "Team" and all who contributed to the back-to-back successes. To our colleagues at INTELSAT and COMSAT, may we express our gratitude for all your assistance in making INTELSAT V a historic event.

## REFERENCES

1. Rusch, R. J., J. T. Johnson, and W. Baer (1978) INTELSAT V Spacecraft Design Summary
2. Hoeber, C. F. (1977) INTELSAT V System Design. *WESCON Conference Records*.
3. Quaglione, G. (1980) Evolution of the INTELSAT System from INTELSAT IV to INTELSAT V. *AIAA 80-4024*.

## THE EUROPEAN LARGE TELECOMMUNICATION SATELLITE (L-SAT) PROGRAMME: DEMONSTRATION MISSION AND FUTURE PERSPECTIVES

B. L. Herdan

*Directorate of Application Satellites Programmes, European Space Agency, ESTEC, Noordwijk, The Netherlands*

### ABSTRACT

The European Space Agency Large Telecommunications Satellite (L-SAT) programme has been conceived against a background of a predicted rapid growth in satellite communications and a trend towards larger and more powerful satellite designs. The programme objectives are the development of a multipurpose platform matched to a range of future mission needs, the development of a multi-element advanced communications payload, and the launch of a single flight model carrying this payload to support a range of communications missions. This paper outlines the main features of the satellite design and then goes in more detail into the payloads and their foreseen utilisation for tests, demonstrations and preoperational systems.

### INTRODUCTION

In 1979, the European Space Agency undertook a study whose objectives were fundamentally to take stock of the dynamic global growth situation in satellite-based telecommunications and assess the implications on the future market for communications satellites of European origin. It was foreseen that conclusions would be employed to derive the correct orientations for any near-term development programmes to be sponsored by the Agency in this domain and it was as a result of this study, matched by the results of parallel investigations performed in industry, that the ESA Member States decided to embark on the definition phase of the L-SAT programme. British Aerospace was selected as Prime Contractor at the outset of the definition phase. The technical orientation of the programme originated from the results of the prior studies, but more work has been done since and details of the programme's technical direction have been carefully adapted in the light of more recent inputs and developments in the user community.

At the time of preparation of this paper, the L-SAT definition phase has practically been completed and the programme is in the course of making the transition into the main development phase. The paper describes the programme objectives and outlines the satellite design. It then concentrates on the four demonstration missions to be supported by the first flight. Payload characteristics are described and plans for utilisation of the satellite are outlined. The paper concludes with a summary of the programme status and of the perspectives for future L-SAT derivatives.

## PROGRAMME BACKGROUND AND OBJECTIVES

The L-SAT programme has been conceived with a view to development of a satellite design matched to the needs of the future market. The market prospects have been evaluated considering in particular the following market sectors :

- European fixed services including the next generation of Eutelsat trunk telecommunications, specialised or business service systems and national trunk services;
- European broadcast services that embrace national and regional direct broadcasting, community broadcasting and programme distribution;
- Global telecommunications trunk services in the context of the next-generation Intelsat systems;
- Global transponder leasing services for national and regional domestic application;
- Mobile service for future generations of the Inmarsat maritime system and for future regional or global aeronautical and land mobile systems;
- Future non-European regional and domestic mixed services, including both fixed and broadcast services with particular reference to developing countries and regions, as well as developed countries not in a position to entirely satisfy their own needs on an economic basis;
- Other services such as data relay and data dissemination.

Predicted requirements in each sector have been analysed and special attention has been paid to those sectors where rapid growth is foreseen. The results of this work have, among other things, led to a series of conclusions concerning two major points, namely :-

- the magnitude of the market and the scope of its penetration by European industry.
- the main characteristics of the satellite types that will be needed in the near future.

As concerns the magnitude of the potential market, suffice to state that the conclusions are encouraging, several different approaches to the analysis having demonstrated quantitatively that a competitive product will have strong sales prospects in most of the above market sectors and that governmental and industrial investment in the development of larger and more powerful satellite design is likely to be readily justified by subsequent successful commercial exploitation.

As for the characteristics and size of satellites that will be required to meet identified market requirements, it has been found that the medium-sized, so-called 'Delta-class', satellite will be insufficient for about 85% of the future applications investigated and that larger satellites will be needed.

Against this background, the L-SAT programme was established with the following basic objectives :-

- development and in-orbit demonstration of a large multipurpose platform able to satisfy the range of foreseen future mission requirements on a cost-competitive basis.

- development of communications-payload hardware related to future mission needs and in-orbit demonstration of new types of communications systems by a multi-element payload conceived to stimulate the introduction of new satellite-based services and techniques.

As concerns the first of these two objectives, a platform design concept has emerged as a result of the satellite definition phase which is briefly described below and which is considered to be of a truly flexible and multipurpose nature, being readily adaptable to support a wide range of foreseen future communications payloads.

To meet the second of the objectives, it was decided to elaborate a multiple payload for the first "demonstration" L-SAT flight opportunity with elements selected on the basis of :-

- relevance to future market opportunities,
- applicability of payload technology developments to the needs of future operational missions,
- potential impact on the user community and consequent benefit to industry,
- potential for provision of capacity to facilitate a "lead-in" to subsequent operational systems,
- mutual compatibility of payloads and compatibility with a single orbit slot,
- compatibility with a reasonably early launch date.

After a period of payload evaluation and selection early in the L-SAT definition phase it was decided in principle that the following four payloads should be carried on the first flight :-

- a payload to support direct broadcast in the 12 GHz band in view of the strong market prospects in this domain and need to acquire and demonstrate certain new technologies to satisfy future requirements;
- a payload for specialised or business services in the European zone, also downlinking in the 12 GHz band, to develop certain new technologies of relevance to the next generation systems (onboard satellite switching and multibeam antennas) and to demonstrate current and new applications;
- a payload consisting of beacon generators at the 20 GHz and 30 GHz frequencies to support the collection of propagation data essential to the design of future systems operating in this band;
- a payload consisting of transparent 20/30 GHz transponder to demonstrate systems expected to operate in this band - including videoconference where a strong market demand is anticipated - and to develop the necessary technological elements which will be needed as "building blocks" in more complex operational systems.

### SATELLITE DESIGN SUMMARY

Before proceeding with the description of the missions and payloads for the first L-SAT flight, it is considered instructive to give an outline of the overall L-SAT design approach and configuration.

As indicated earlier in the context of the overall programme objectives, it was seen to be of fundamental importance that the satellite platform have the flexibility to accommodate a wide range of communications payload types with differing characteristics. This relates on the one hand to the accommodation of the repeater equipment needed for the full range of mission types and a configuration which permits adequate thermal dissipation from repeater equipment whose output stages may use solid state or TWT technology. On the other hand, flexibility to accommodate all types and dimensions of antennas is essential noting that there is a trend towards more and larger antennas with accurate pointing requirements either to permit narrow beam high flux level operation and/or on account of frequency spectrum conservation requirements.

The foreseen payload mass and power needs were derived as a result of the investigations of market prospects described earlier and can be summarised by the following main parameters :

- |                               |                |
|-------------------------------|----------------|
| - Sunlight Power              | - Up to 7 kW   |
| - Eclipse Power               | - Up to 2.5 kW |
| - Communications Payload Mass | - Up to 500 kg |

Launcher characteristics and development plans presented a further input to considerations of platform configuration concepts. To meet the possible desires of future customers it was considered essential for commercial reasons that the platform be compatible with S.T.S. as well as Ariane. In addition, in order to minimise launch costs, it is obviously desirable to use the most efficient version of a launcher for a particular mission. The platform should thus be suitable for launch on Ariane 2, Ariane 3, or Ariane 4 (both single or double launch), without incurring major penalties. Also as concerns STS launch it is important to select a configuration which, matched to a suitable perigee stage to give equivalent capability into orbit to Ariane 2/3/4, will permit a cost effective launch based on the existing and predicted STS charging policy. In summary, the most significant launcher constraints are :-

- Mass capability
- Environment
- Shroud dimensions
- Safety requirements
- S.T.S. charging policy
- S.T.S. perigee stage concepts

The communications payload requirements of the future missions and the need for a cost effective launch have been drivers for the L-SAT configuration and the inbuilt flexibility concepts for the L-SAT platform design.

The geometrical configuration of the L-SAT satellite has been selected to meet the payload and launcher requirements in an optimum manner. Its characteristics are in summary :

- A North and South radiating area of  $9m^2$  adequate to handle the largest foreseen payload power dissipation requirements while maintaining appropriate operating temperature.

- A core body size of minimum dimensions compatible with the Ariane shroud : the main dimensions are 1.75 x 2.1 metres cross section x 3.5 metres high.
- The capability of mounting antennas from the East and West faces of the body (as well as on the earth pointing face) in order to avoid tall antenna towers adding excess length and hence cost in S.T.S.

In effect, the satellite body acts as an antenna tower for side mounted antennas. Secondly, it is built around the propulsion system which is vertically accommodated within the load-bearing central tube to minimise satellite cross section and structural mass. Adequate volume has been allowed within the service module to accommodate the largest batteries which could be needed for any future mission. A highly modular satellite configuration has been adopted to give maximum flexibility for accommodation of variations in payloads from mission to mission and to minimise assembly and integration durations through parallel working on different modules. The satellite geometry and modular configuration are illustrated in Figures 1 and 2, which also show the accommodation of the L-SAT 1 multi-element payload. The satellite is stabilised in three axes during all mission phases employing a single nominal mode of operation throughout. A basic body earth-pointing accuracy of  $0.1^\circ$  is provided. The platform is designed to meet a 10 year lifetime requirement with high reliability.

The mass of communications payload which can be flown on the L-SAT platform will depend on three main features :-

- the capability of the selected launcher;
- the payload power demand, since the mass of solar array, battery and thermal control depend on this;
- the mission lifetime, since fuel and payload mass are broadly interchangeable;

Figure 3 shows the approximate capability of the L-SAT platform for the intended range of launchers and payload power, for the case of a typical 7 year mission duration. The point representing the L-SAT 1 payload mass/power combination is shown for information.

## THE L-SAT 1 COMMUNICATIONS PAYLOADS

As explained earlier, it was decided at the outset of the programme definition phase that the L-SAT platform should carry a multiple-element payload for the first demonstration flight, and that selection of the payload elements should be based on the strength of market predictions for each mission type, the interest in demonstrating and promoting certain new services, payload technology benefits and availability of financing for space and ground segment. This approach is aimed at leading to a concrete demonstration of the platform multi-mission capabilities as well as promoting the European payload industry. Four missions have been defined in order to arrive at a suitable demonstration payload combination compatible with these broad objectives.

### The Broadcast Mission

L-SAT 1 will carry a two channel payload suitable for direct-to-home broadcasting in the 11.7 - 12.5 GHz band. Circular polarisation will be employed. It is foreseen that one channel will be mainly employed for pre-operational services in Italy while the other channel will be steerable to support experiments and demonstrations over the whole of Europe on a time sharing basis. Facilities for transmission of a single "European" programme over the whole of Europe for reception by community/cable head installations will also be provided. The satellite will receive in the 17 GHz band using a single European coverage receive antenna and will be compatible with uplinks from outside-broadcast stations. The payload will employ a 230 Watt travelling wave tube amplifier in each 27 MHz bandwidth channel to give an output e.i.r.p. of about 61 dBW. Foreseen transmit antenna beamwidths are  $1.6^{\circ} \times 1.6^{\circ}$  and  $1.0^{\circ} \times 2.4^{\circ}$ , and a closed loop RF sensing /antenna pointing system will demonstrate high accuracy pointing for the narrower beam antenna. Figure 4 shows the payload block diagram and Figure 5 shows a typical set of transmit beam coverage contours for a given setting of the two antenna pointing mechanisms. The possible use of this payload for disaster communications is described in another paper at this conference.

### The specialised Services Mission

The second L-SAT mission is concerned with the provision of pilot specialised or business services to the whole European zone, so extending and/or improving on the possibilities provided by the existing generation systems and preparing the ground for the next systems. Operation with small terminals with an antenna diameter of 3 metres or less is foreseen, uplink being in the 14 or 13 GHz bands and downlink in the 12.5 - 12.75 GHz band. A steerable multiple beam antenna is foreseen with  $5, 1.3^{\circ}$  diameter, spotbeams covering the whole of Europe with frequency reuse between beams. The satellite e.i.r.p. will be around 45 dBW over each coverage zone involving the use of 30W TWTA's. The operating system will employ SS TDMA for which purpose the payload incorporates a  $4 \times 4$  switching matrix. Such a concept is intended to prefigure a future operational European system and to prove the key technologies and systems which will be needed.

Further objectives of the payload are the provision of an experimental facility for the demonstration of multipoint video-teleconference by satellite at 12/14 GHz, television programme distribution to cable networks, and provision of some back up for the specialised services package of ECS. Moreover, since the antenna is mounted on a repointing mechanism, it may also be used for demonstrations and experiments of communications at 12/14 GHz to any desired region visible from the spacecraft station, particularly in view of possible promotion of satellite communications systems in Africa and South America.

Figure 6 shows the block diagram for this payload and Figure 7 shows coverage contours for the multibeam antenna.

20/30 GHz Communications Mission

The L-SAT 20/30 GHz Communications Transponder has been conceived as a flexible element providing a number of different experiments and service demonstrations :

- video teleconference (point-to-point) between two simple earth stations located anywhere within Europe using digital transmission of video plus sound at 8.44 Mbit/s through separate repeater chains in each direction.
- video teleconference (multi-point) between three or more earth stations located within Western Europe using digital transmission of two video plus sound signals at 8.44 Mbit/s through separate repeater chains plus auxiliary control and/or sound channels either at channel edges or through a third repeater chain.
- tele-education from one location to a number of others using simple earth stations located within a limited zone with video plus sound transmission using digital or analog (FM) modulation, with (if feasible) audio return links.
- data and video transmission between simple earth stations within Western Europe using digital transmission at, for example, 2.048 Mbit/s.
- communications experiments with wide bandwidth (e.g. pulsed) signals, probably using one earth station in loop-back configuration.

These facilities will be provided by operating in various combinations a payload comprising two independently-steerable transmit/receive antennas, each generating a single 1.2 degree coverage spot beam, two wide-band receive chains, three transmit chains, and a comprehensive arrangement of switches and filters.

Figure 8 shows the block diagram for this payload and Figure 9 is a typical coverage diagram for one setting of the pointing mechanisms.

12/20/30 GHz Propagation Package

Within the general frame of radio propagation research for space telecommunications applications, ESA and national bodies have identified the need for producing reliable models for signal loss, depolarisation, and any other channel impairment introduced by the modes of propagation of radio waves between earth and space. These models are needed in a relatively short time at frequencies of 20 GHz and 30 GHz to support the design of future satellite communications systems. Using the L-SAT 1 propagation package these models can be developed for the European region and will reflect the variability of the climate encountered within Europe. An experimenter group already exists in which more than 40 organisations are represented.

The 20/30 GHz propagation payload will incorporate 2 transmitters (at approx. 20 GHz and 30 GHz). The signals of the propagation package will be of high performance regarding :

- amplitude stability within the specified coverage,
- polarisation purity within the specified coverage,
- the frequency stability and frequency jitter.

The 20 GHz beacon can be set to either one of two orthogonal polarisations or switched at 933 Hz between the two polarisations.

The propagation package will also include a 12 GHz beacon transmitting an unmodulated carrier in order to provide a tracking facility for all L-SAT communications earth stations. This beacon can also be used for propagation measurements.

Figure 10 shows the block diagram for this payload and Figure 11 shows the coverage requirements established in view of identified experimenters participating in the programme.

Overall Payload Accommodation

The ability of the L-SAT multipurpose platform to accommodate the four payload elements, in itself a significant challenge, has been demonstrated during the Definition Phase and the platform/payload combination has been designed in preparation for the Development Phase.

Figures 1 and 2 have shown the L-SAT configuration and how the four payload elements are accommodated on the L-SAT platform. Two of the Broadcast payload antennas will be mounted on the earth pointing face while the European beam transmit antenna is deployed from the West satellite wall. The other equipment is mounted on the South face. Also mounted on this face and on the inside of the earth pointing face is the propagation package equipment.

The Specialised Services payload equipment is mounted with the 20/30 GHz payload equipment on the North face. The multispot beam antenna is deployed from the East wall of the satellite. The 20/30 payload antennas are mounted on the earth pointing face.

The main system parameters of the L-SAT 1 satellite are listed on Table 1.

<u>MASS</u>	(Kg)
Payload (Repeaters and Antennas)	314
Platform	847
Margin	83
Total Dry Mass	1244
Propellant	1178
Mass into Transfer Orbit	2422
<u>POWER</u>	(W)
Payload	2141
Platform	719
Battery Charging	175
Total Power Required	3035
Solar Array installed power	3300
Eclipse Power Capacity	1050
<u>DIMENSIONS</u>	
Body Dimensions	2.1m x 1.75m x 3.5m
Array Span Tip-to-Tip	27m
<u>PAYOUT CONTENT</u>	
Total Number of installed Transponders	9 (plus 3 beacons)
Total Number of Antennas	10 (5 steerable)
<u>LIFETIME</u>	
5 Year Mission duration (Propellant Loading)	
<u>LAUNCH</u>	
Ariane 3 (STS back-up)	
<u>STATION</u>	
Satellite Station : 19°W	Station-keeping ± 0.07°

TABLE 1 L-SAT 1 - System Characteristics

### CURRENT PROGRAMME STATUS AND DEVELOPMENT PLANS

The start-up of the L-SAT definition phase was decided in mid-1979 following a period of preliminary studies in ESA and in industry along the lines indicated earlier in this paper. A total of nine countries are participating in the definition phase programme, namely Austria, Belgium, Canada, Denmark, Italy, The Netherlands, Spain, Switzerland and the United Kingdom. During the second half of 1979, a competition was undertaken for the selection of the Prime Contractor for the satellite, and this was won by British Aerospace. The definition phase work got underway in industry at the beginning of 1980 and was completed this summer, culminating in the submission of a detailed and mainly fixed price proposal for the execution of the main development phase which is currently under negotiation, in part with a view to effecting a degree of cost reduction. A bridging phase has been established to cover the period between completion of the definition phase and start of the main development phase in order to provide for continued forward progress on schedule critical activities while allowing adequate time for negotiation of the development phase contract between ESA and industry and for the nine governments to reach the decision to proceed with the main development phase of the programme. This latter is confidently expected to be achieved by December 1981.

It is probably of interest to note that the L-SAT definition and bridging phases will, in total, have taken almost two years to complete and have involved a total expenditure by the participating governments of some \$40 million. The result of this large amount of time and effort expended to date has been to generate a rather well defined satellite design which ESA and the industrial team believe to be solidly substantiated by the extensive tradeoffs of which some examples have been outlined in this paper. Moreover, almost half of the total sum mentioned above has, or is to be, employed for schedule critical detailed design elaboration, including significant hardware tasks (breadboarding of equipments or subassemblies), and for initiation of long lead procurement activities to lay a sound basis for the start of the main development phase.

The total cost-to-completion of the main development phase, including the satellite development and flight model plus spares procurement by industry, the ground segment development, the launch cost, the cost of the injection and operations for five years, and the cost of ESA programme management, is currently estimated at \$520 million at 1980 price levels.

### FUTURE PERSPECTIVES

The emphasis of this paper has been mainly put on the L-SAT satellite design and the communications missions for the first flight. However, it would not be complete without some commentary concerning future perspectives. As has been stressed earlier the L-SAT programme has its origins in the extremely encouraging results of market survey investigations, and the satellite design has been mainly driven by the requirements predicted for future missions. This is not a mere generality but a summary statement of a situation supported by a wealth of detailed study work and design activities. Some twenty typical and probable medium term future missions have been identified and corresponding L-SAT derivative designs outlined. A sample of six typical derivative types has been elaborated in more detail to verify rather precisely the correct and direct adaptability of the L-SAT design to meet their requirements.

As concerns the nearer term and prospects for more immediate sales, it can be mentioned that, among the countries participating in the L-SAT programme, three of the larger contributors (Canada, Italy and the United Kingdom) have identified probable near-term domestic requirements for satellites of a pre-operational or operational nature derived from L-SAT and each of these is the subject of an ongoing

study programme. There seems little doubt that, within the next year or two, at least one of these will come to fruition and become the "launch customer" for the series of L-SAT derivatives. Further ahead, the industry marketing efforts have identified a series of promising market prospects both within and beyond the participant countries, and this has added further impetus to the programme. As the programme proceeds, work on the identification and design of derivatives will be energetically pursued and it can confidently be predicted that several will be in an advanced state of design or even development by the time that L-SAT 1 is launched early in 1986.

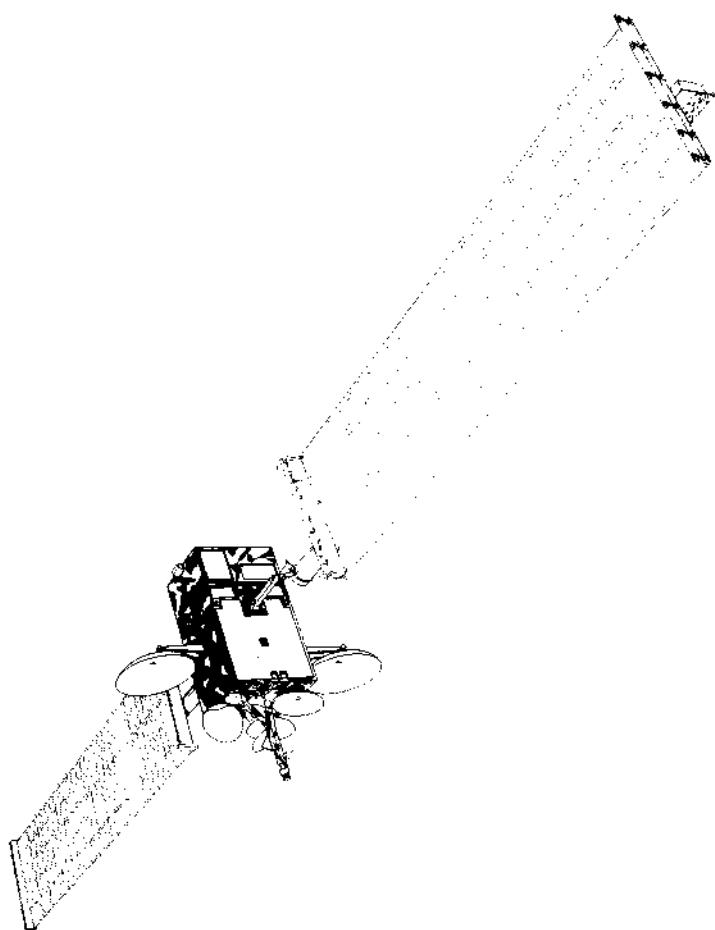


Fig. 1. L-SAT In Orbit Configuration

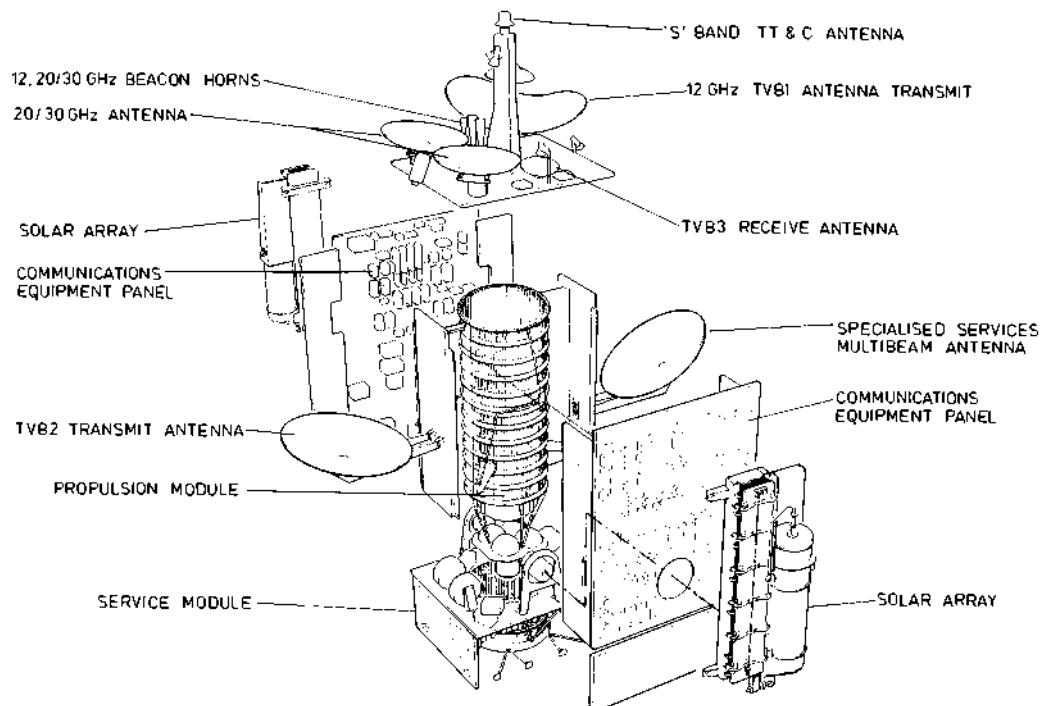
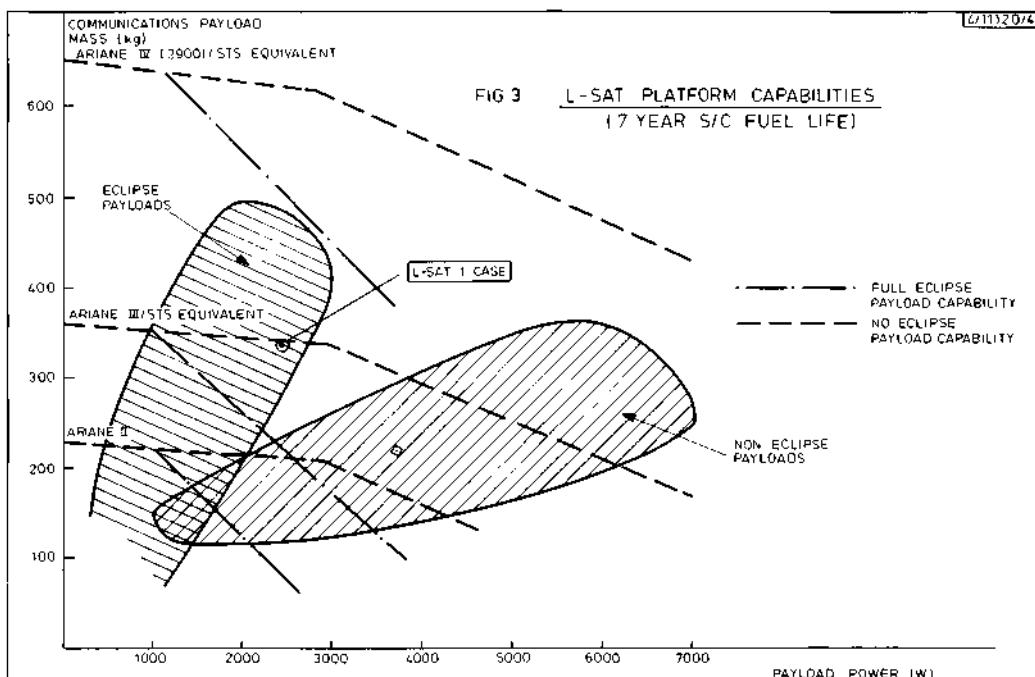


Fig. 2 L-SAT Modular Arrangement (Exploded Diagram)



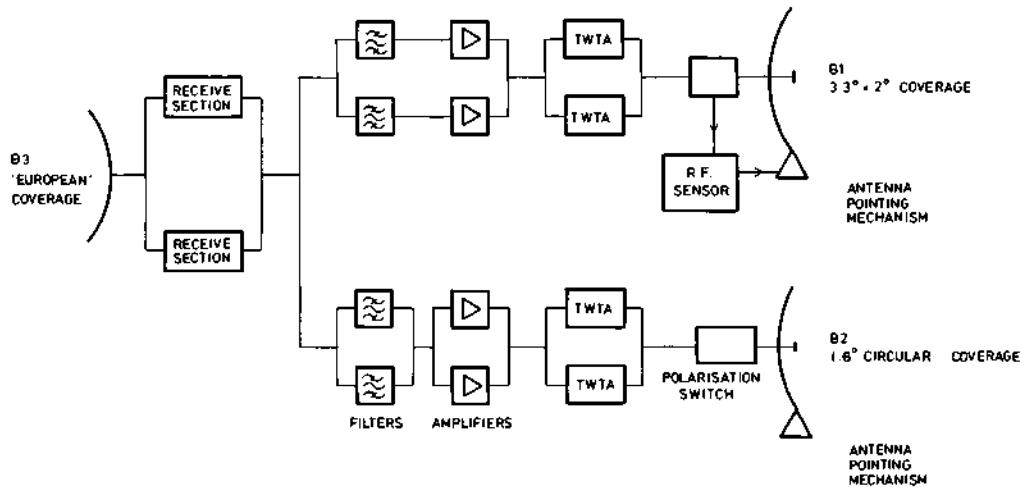
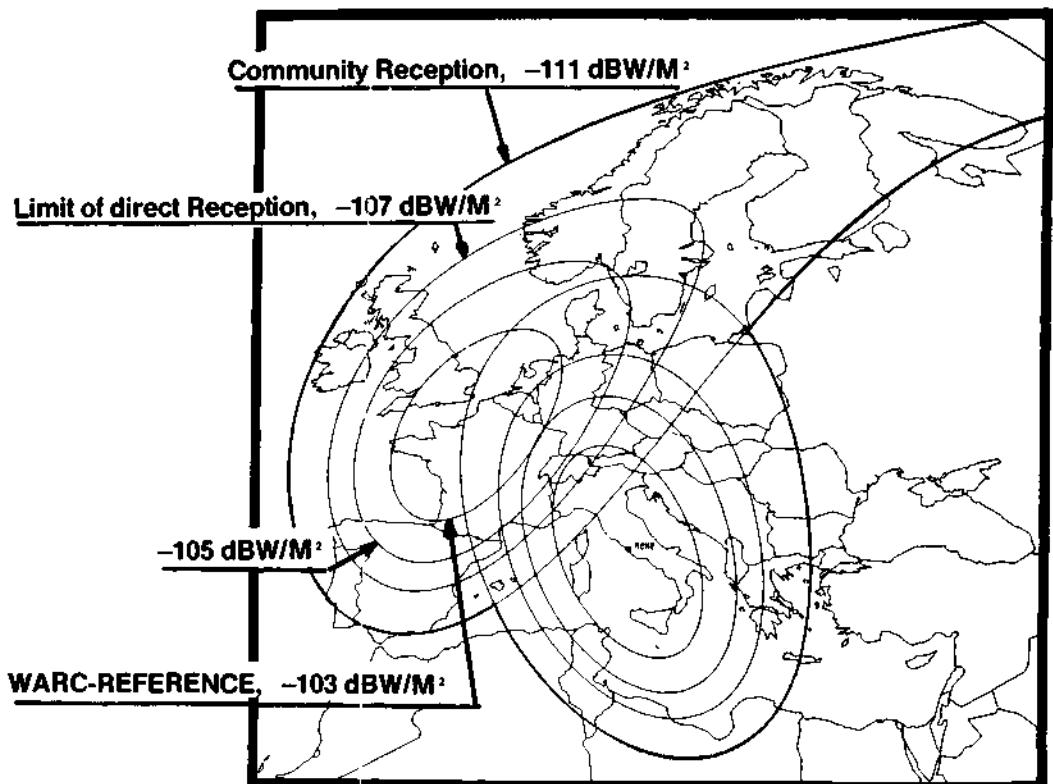


Fig. 4 L-SAT T.V. BROADCAST PAYLOAD

Fig. 5. Broadcast Mission  
Set Of Transmit Beam Coverage Contours

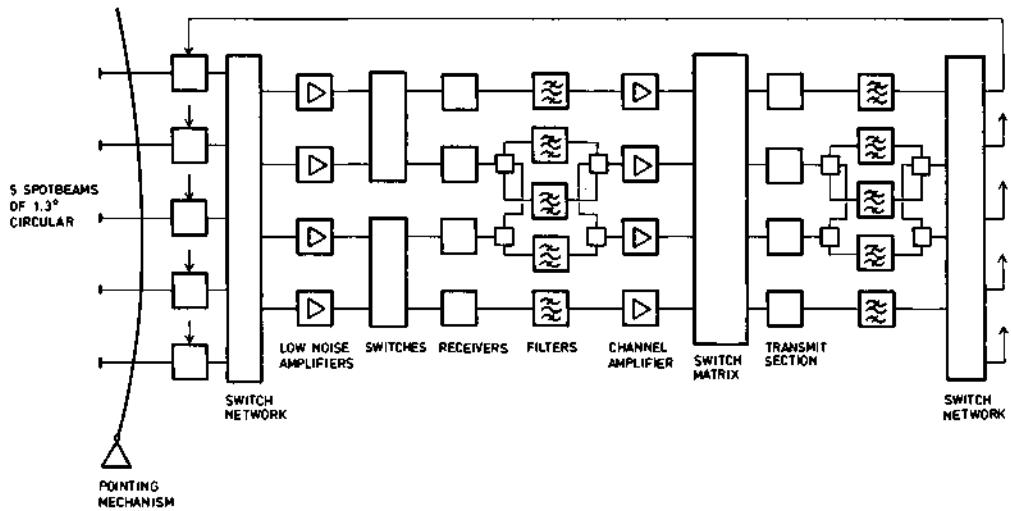
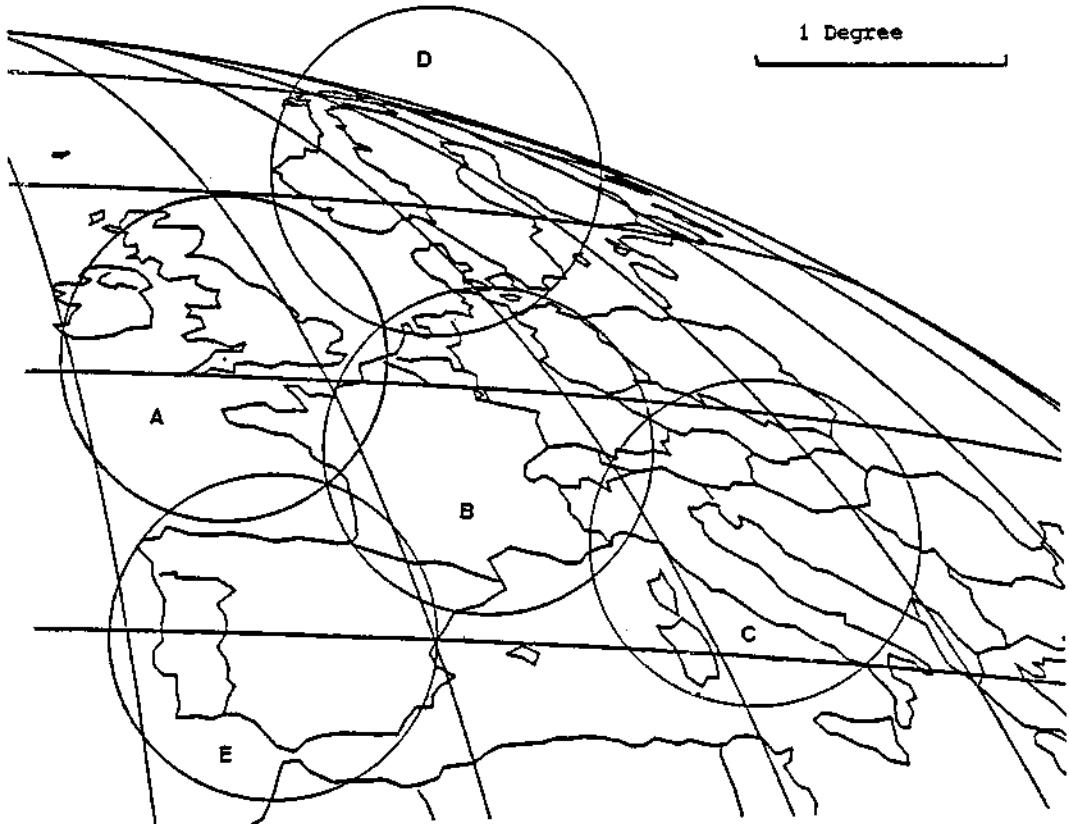


Fig. 6 L-SAT Specialised Services Payload

Fig. 7. Specialised Services Mission  
Typical Coverage Contours For The Multibeam Antenna

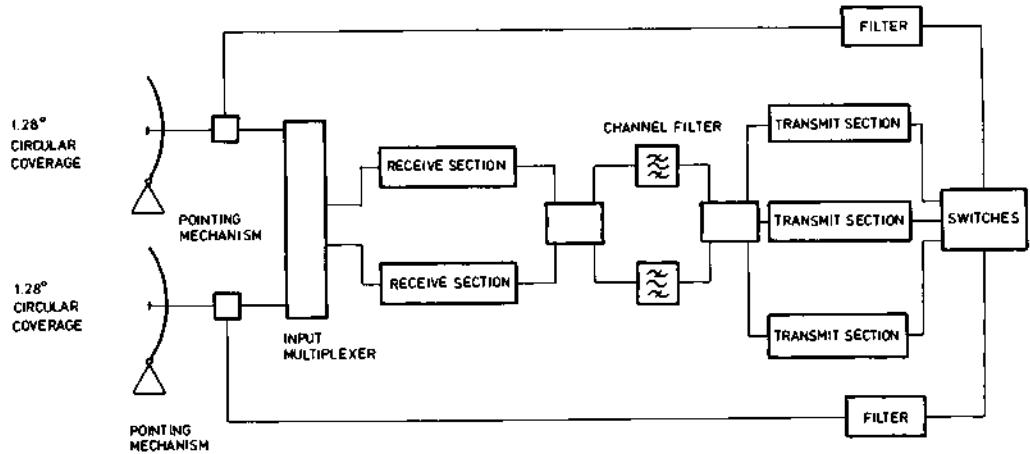


Fig. 8 L-SAT 20/30 GHz Communication Payload



Fig. 9. 20/30 GHz Communications Mission  
Typical Coverage Diagram For One Setting Of The Antenna Pointing Mechanisms

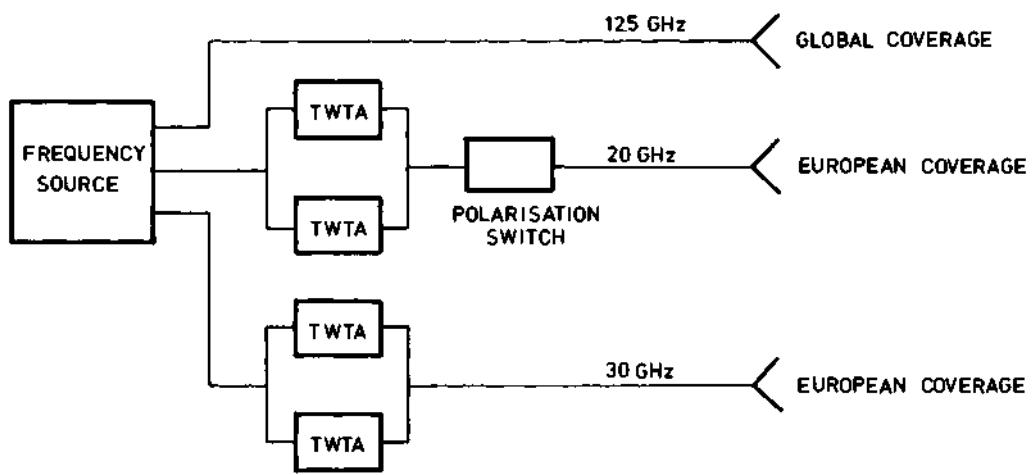


Fig. 10 12/20/30 GHz Propagation Payload

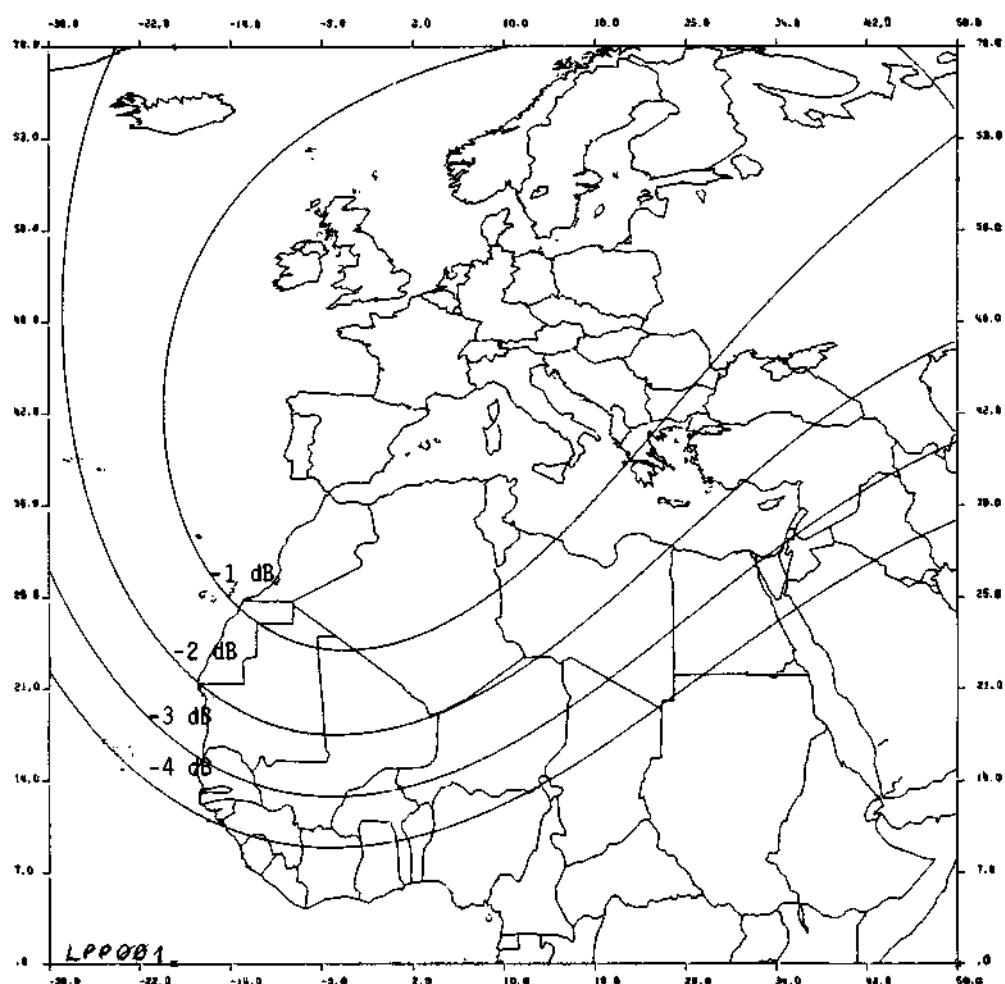


Fig. 11. 12/20/30 GHz Propagation Mission  
Coverage Requirements

## PLANNING FOR THE BROADCASTING-SATELLITE SERVICE IN THE UNITED STATES OF AMERICA

J. F. Clark

RCA Corp., P.O. Box 432, Princeton, N.J. 08540, USA

### ABSTRACT

The 1983 Regional Administrative Radio Conference (RARC-83) for Region 2 is the next in a series of International Telecommunications Union (ITU) meetings which will shape the Broadcasting-Satellite Service (BSS) in the western hemisphere. This paper provides insight into some opportunities and associated problems presented by BSS in the United States, and mechanisms by which the Federal Communications Commission (FCC) is integrating advice from the private sector into its planning for RARC-83. Eight proposals, to provide Direct Broadcast Satellite (DBS) service in the 12.2 - 12.7 GHz band, were accepted for filing by the FCC in October 1981, and provide important inputs to the RARC-83 preparations.

### KEYWORDS

Broadcasting; broadcasting-satellite service; direct broadcast satellite; Federal Communications Commission; International Telecommunications Union.

### BACKGROUND

It was at the 1963 Extraordinary Administrative Radio Conference (EARC) that the BSS was formally recognized as a distinct space service. BSS received frequency allocations in bands near 0.7, 2.6, 12, 23, 42, and 85 GHz at the 1971 World Administrative Radio Conference (WARC) for Space Telecommunications.

At WARC-77, a plan was developed for the BSS in the 12-GHz band. This plan allotted one of the 35 orbital positions to each of the 252 service areas of the 150 Administrations and Territories of ITU Regions 1 and 3 (Europe, Africa, Asia, and Australia)[1]. Each service area was allotted from one to eight frequencies, with an average number of nearly four.

In contrast to the detailed a priori planning for Regions 1 and 3 at WARC-77, the countries of Region 2 (the Americas and Greenland) postponed detailed BSS planning until the RARC which is now scheduled for 1983. Specifically, WARC-77 categorized BSS in Region 2 as experimental until RARC-83 and confined such satellites to two limited orbital arc segments. The remainder of the orbital arc was allocated to the Fixed-Satellite Service (FSS), which shared the Region 2 BSS allocation of 11.7 - 12.2 GHz.

Two years later, it had become apparent that arc segmentation between the BSS and FSS provided sufficient capacity for neither service. Significant provisions of WARC-79 eliminated arc segmentation in Region 2. The band 11.7 - 12.1 GHz was allocated to the FSS as the primary space service, 12.3 - 12.7 GHz to the BSS as the primary space service, and the intervening 200 MHz will be divided between the BSS and FSS at RARC-83.

International preparations for RARC-83 are well underway. The CCIR has established an Interim Working Party, IWP 10-11/2, to provide a technical basis for the RARC in the form of a report to be given to all conference delegates. Membership of this IWP was drawn partly from CCIR Study Groups 10 (Sound Broadcasting) and 11 (Television Broadcasting), as indicated by its designation.

Another important CCIR activity is a Joint Working Group, JWG 10-11S, where the S designates satellite. Its last meeting took place in Geneva in the fall of 1981, in part concurrently with IWP 10-11/2.

#### U.S. PLANNING FOR RARC-83

In the United States the FCC released a Notice of Inquiry (NOI) in July 1980 requesting public comment on the BSS in preparation for RARC-83[2]. Major areas addressed in the NOI are basic service requirements, technical specifications, sharing criteria, and planning principles and procedures. In February 1981 the FCC ordered the establishment of an advisory committee to assist in these areas. Its first meeting was held in May. In June the FCC issued its second NOI, based on 21 responses to its first inquiry.

With regard to basic service requirements, the FCC preliminarily suggested that the U. S. proposal to RARC-83 divide the shared band at 12.2 GHz. The BSS would be based on individual (rather than community) reception. A four-service-area (time-zone) approach would be selected for the 48 contiguous United States (CONUS). A larger number of channels from a more limited number of orbital locations would be expected to provide a better operational environment. Finally, eclipse protection would be required in first generation BSS satellites.

Concerning technical parameters, the FCC favored circular polarization over linear for BSS transmissions. A minimum single-entry carrier-to-interference ratio (C/I) between 29 and 31 dB and a minimum total-entry C/I between 24 and 26 dB were suggested. Possible need for a channel bandwidth greater than the 18 MHz specified by the 1977 WARC-BS was cited, particularly for ancillary services such as dual language, stereo, and teletext, and for high definition television. The FCC viewed the receiver gain-over-temperature ratio (G/T) of 6 dB/K, specified by the WARC-BS, as too conservative, and requested comments on the relative costs of earth terminals having antenna diameters between 0.6 and 1.0 m, and with G/T values between 6 and 12 dB/K. Power flux density (PFD) was viewed as a system-dependent parameter to be considered together with earth terminal G/T and necessary signal bandwidth.

The FCC cited the issue of detailed a priori planning versus more limited (flexible) planning as a very critical one. Its preliminary view was that RARC-83 should pursue only a limited planning approach, with reserve orbit and spectrum capacity to be accessed later, as needed, through an established set of procedures.

#### ADVANCING TECHNOLOGY AND PLANNING

Hindsight verifies the wisdom of more flexible planning, reserving detailed planning until the orbit or spectrum capacity must actually be used. One example

of a foreclosed technological advance is the WARC-77 specified value of earth terminal G/T of 6 dB/K in Regions 1 and 3. A present-day low-noise gallium-arsenide field-effect transistor (GaAs FET) input stage, following a 0.6 to 1.0 m antenna, provides a very cost-effective G/T 3 to 6 dB better than the WARC-77 value. This G/T margin might have been used to reduce the BSS transmitter power by 50 to 75 percent. Or, in a more speculative vein, the advantages of an elliptical (rather than circular) ground-terminal parabolic antenna reflector seem not to have been considered. By orienting its major axis of at least 1.0°m east-west (relative to the BSS satellite longitude), the resultant fan beam 1.8° or less in width would decrease the required spacing between adjacent BSS satellites. To decrease the cost of the ground-terminal antenna, its minor (north-south) reflector axis could be reduced in length so long as the resultant G/T did not fall below its required value. This trade-off could open the still more speculative possibility that the satellite might be launched into a geosynchronous orbit which is inclined to the equator. The satellite's narrow "figure-eight" track would remain within the fan beam of the ground-terminal antenna. Such an orbit would permit a heavier payload to be launched by a given launch vehicle from a non-equatorial site, and could also reduce the required amount of north-south satellite station-keeping propellant required for a specified satellite lifetime. This last advantage can also be achieved on more conventional geostationary satellites by employing station-keeping propellants of high specific impulse. Electrically-heated monopropellant and ion thrusters are two examples of propellant systems having successively higher specific impulse characteristics. Such techniques represent a very small portion of the technological advances which will provide ever more efficient utilization of BSS spectrum and orbit, improved BSS service, and reduced BSS system cost.

#### INTERIM DBS PROPOSALS

An FCC June 1, 1981 Notice of Proposed Policy Statement and Rulemaking set an important deadline for potential applicants who wished to provide interim DBS service in the United States[3]. All such applicants who desired to have their filings considered with the same priority afforded the December 1980 application of the Satellite Television Corporation were required to submit their interim DBS applications no later than July 16, 1981. In response, 11 more applications were received before the deadline, and two additional applications were submitted within the following 24 hours. They ranged from one page to many chapters in length.

Eight of these proposals, to provide service in the 12.2 - 12.7 GHz band, were found acceptable for filing by the FCC on October 22, 1981. These are, alphabetically: Columbia Broadcasting System, Inc. (CBS)[4]; Direct Broadcast Satellite Corporation (DBSC)[5]; Graphic Scanning Corporation (GSC)[6]; RCA American Communications, Inc. (RCA)[7]; Satellite Television Corporation (STC)[8]; United States Satellite Broadcasting Company, Inc. (USSB)[9]; Video Satellite Systems, Inc. (VSS)[10]; and Western Union Telegraph Company (WUTC)[11].

Table 1 provides a summary of some key parameters of these eight proposed systems in their fully-deployed versions. Despite many common elements, such as individual reception of the DBS transmissions, the dominant impression projected by this flood of DBS proposals is their innovative diversity of purpose and, to a lesser extent, of implementation.

Only one of the new applicants, GSC, followed the lead of STC in providing an exclusive multi-channel, scrambled, subscription-supported DBS service. DBSC was alone in requesting common carrier status for its ambitious system, which is also the only one of these proposed systems to use spot beams in addition to CONUS coverage. RCA also proposed to lease its channels, but reserved the right to

TABLE 1 Key DBS Parameters

Applicant	Noteworthy Characteristics	CONUS Zones	No. Channels	Orbital Locations	TWTA Watts	EIRP dBW	Bandwidth MHz	Polarization	Launch Vehicle Class
CBS	Sponsored; Rebroadcast	4	3	"RARC-83"	400	60.4	27	Circular	Shuttle/SSUS-A
DBSC	Common Carrier	3	[+8]1	6 143° 143°	200 [20]1	56	22.5	Circular	Shuttle/ Centauro
GSC	Subscription	2	2	115°, 143°	300	53.7	18	Circular	Shuttle/ SSUS-D
RCA	Leased Channels	4	6	110°, 125° 140°, 155°	230	58	25 [72]3	Linear	Shuttle/ SSUS-A
STC	Subscription	4	3	115°, 135° 155°, 175°	185	57	16 [28,100]3	Circular	Shuttle/ SSUS-D
USSB	Sponsored; Rebroadcast	4	3-6	115°, 135° [+2]2	230	57	16	Circular	Shuttle/ SSUS-A
VSS	Sponsored; Rebroadcast	4	2	115°, 135° 155°, 175°	150	56	18	Circular	Shuttle/ SSUS-D
WUTC	Leased Channels	4	4	80°, 100° 120°, 140°	100	55.5	16	Linear	Shuttle/ SSUS-D

1 - Spot beams

2 - Undesignated

3 - HDTV

retain some capacity for its own use. WUTC proposed complete control of the selection of program suppliers for its system, and described a complex automated ground-terminal switching scheme to insert local programming received from existing VHF/UHF terrestrial stations.

USSB would bridge DBS and conventional broadcasting by retransmitting the unscrambled commercially sponsored DBS programming from one terrestrial member station in each geographic area. This would permit the member stations to share in the provision of the DBS programming via local feeder links and to share in the commercial sponsorship revenue. The public could receive the same programming using either a conventional VHF/UHF receiver or a 12-GHz DBS terminal. A second USSB channel would provide a 24-hour news and information service. VSS also proposed to distribute free commercially sponsored programming, but its scrambled signals would be available to the public only through designated terrestrial affiliated stations within their service areas. Elsewhere free decoders would enable direct DBS reception by individual viewers and by CATV systems not carrying any VSS affiliate. CBS proposed a wide-screen high-definition television (HDTV) satellite system to serve viewers through their local affiliated CBS stations. Like VSS, CBS would encourage individual reception of their network HDTV programming directly from their DBS only outside affiliated-station service areas. CBS also proposed two additional scrambled HDTV subscription channels for institutional, business, and residential users. Five other applicants proposed more limited HDTV uses: DBSC, RCA, STC, USSB, and VSS.

The major common element of purpose in these eight proposals is DBS service to viewers in all parts of CONUS (and to at least some parts of Alaska, Hawaii, Puerto Rico, and the Virgin Islands) by a system which is economically viable, supported by some combination of subscriber and/or advertiser revenue. No government subsidy is contemplated.

#### PROPOSED DBS TECHNICAL CHARACTERISTICS

Six of the eight proposed systems followed the STC lead in dividing CONUS into four service areas which are approximately equal to the four CONUS time zones. Their areas are adjusted to provide stronger signals in regions of climatologically heavier precipitation, at the expense of drier areas where lower signal margins are sufficient. The exceptions are DBSC, which proposed three service areas; and GSC, which proposed two. All of the fully-deployed system proposals feature one operating satellite per service area. CBS, GSC, RCA, and STC proposed to deploy their systems one satellite (service area) at a time. DBSC proposed to cover all three CONUS service areas from its first operating satellite. USSB, VSS, and WUTC each proposed to cover all four CONUS zones with their first two satellites (two zones per satellite). In their fully-deployed systems, GSC and VSS would provide two channels everywhere in CONUS; CBS and STC, three channels; WUTC, four; USSB, three to six; and DBSC and RCA, six. DBSC would also provide four channels in each of two spot beams in each of their three service areas.

WUTC proposed the only system which would provide full service during DBS satellite eclipse periods. The proposed orbital locations between 80° and 140° west longitude are almost due south of their service areas, thereby insuring high elevation angles (33° - 55°) of the ground terminal antennas. DBSC proposed to place their three satellites between 103° and 143°, but with operation of only about one-third of their capacity during eclipse periods. The six remaining applicants proposed to place their satellites between 110° and 175°. By providing little or no capability for operation during the delayed early morning eclipse periods, significant satellite-battery weight savings would be achieved.

All proposed DBS transponders would use travelling-wave tube amplifiers (TWTAs), ranging from 100 to 400 W output power. After normalizing for the area of coverage, most of the proposed values of effective isotropically radiated power (EIRP) at the edge of coverage fall between 56 and 58 dBW. One exception is that of GSC, which would provide less than 54 dBW because of the unusually large half-CONUS coverage per beam. The other exception is CBS, which would provide more than 60 dBW EIRP for a more adequate carrier-to-noise ratio (C/N) for HDTV transmission.

Suggested channel bandwidths varied from 16 MHz (STC, USSB, WUTC) to 24 MHz (RCA-NTSC) and 27 MHz (CBS-HDTV). STC proposed HDTV experiments using 28 and 100 MHz, and RCA proposed 72 MHz (three 24-MHz channels) for similar purposes. Circular downlink polarization was suggested by six of the applicants, but RCA and WUTC defended vigorously their choice of orthogonal linear horizontal and vertical polarization for maximum cross-polarization discrimination (XPD) in the presence of rain.

Four proposed spacecraft (GSC, STC, VSS, WUTC) are in the Shuttle/spin-stabilized upper stage - Delta (SSUS-D) weight class. The total power output of their operating TWTAs would range between 300 and 600 W. Three spacecraft (CBS, RCA, USSB) would be in the SSUS-A category, producing between 1200 and 1400 W of output power. The DBSC spacecraft would require a launch vehicle of even greater weight capability (e.g., a modified Centaur liquid-propelled upper stage) because of the 14 TWTAs operating at a total output power of nearly 1400 W. Many of the proposers (CBS, USSB, VSS, WUTC) regard an appropriate model of Ariane as a possible alternative launcher.

In general, most of the NTSC system proposers advocated ground terminal antennas from 0.6 to 1.0 m in diameter, a signal-to-noise ratio (S/N) of at least 42 dB, and a C/N at least 4 dB above threshold in fair weather. Such a signal would be rated excellent by at least half of the viewers, and would be available for at least 99% of the worst month of the year throughout CONUS.

#### THE FCC ADVISORY COMMITTEE ON RARC-83

Table 2 presents the organization of the Advisory Committee. Receipt and publication of the DBS proposals in July, 1981, by the FCC provided key inputs to all of the Working Groups. Inclusion in one or more of those proposals is neither a necessary nor a sufficient condition for qualification as a BSS service requirement. However, it is obvious that those proposals provide a very useful source of diverse system concepts and requirements. Another principal source of information for many of the Working Groups is the documentation produced by several CCIR activities, particularly JWG 10-11/S and IWP 10-11/2. Substantial common membership between these CCIR activities and the FCC Working Groups facilitates rapid application of the CCIR documentation to FCC tasks.

Ideally, the Advisory Committee would first search out and document all BSS service requirements through Subcommittee 1. Then Subcommittee 2 could translate these requirements into recommended technical parameters of the DBS systems. Only then could Subcommittee 3 commence its work on inter-service sharing. In practice, we could never "get there from here" in time to meet the deadline of April, 1982, for the final report to the FCC, in the sequential manner just described. Instead, a "parallel processing" scheme is being used, with feedback loops from Subcommittee 3 to 2, and 2 to 1.

TABLE 2 Organization of FCC Advisory Committee on Preparations  
for the ITU RARC-83 BSS Planning Conference

Committee Chairman - J. F. Clark

Subcommittee 1 on Service Requirements - S. E. Doyle

Working Group 1A: Conventional Television Service Requirements  
Working Group 1B: High Definition Television Service Requirements  
Working Group 1C: International (non-U.S.) Service Requirements  
Working Group 1D: Other Related Service Requirements  
Working Group 1E: Public Service Requirements

Subcommittee 2 on Technical Parameters - E. E. Reinhart

Working Group 2A: Planning Parameters  
Working Group 2B: Planning Approaches and Modification Procedures

Subcommittee 3 on Inter-Service Sharing - J. J. Kelleher

Working Group 3A: Sharing in the 12 GHz Band  
Working Group 3B: Sharing in the 17 GHz Feeder Links  
Working Group 3C: Spurious Emissions  
Working Group 3D: Interface between ITU Regions

#### THE FCC ADVISORY COMMITTEE REPORT

There are two basic types of advisory committee reports. The first type is motivated by an overwhelming desire to achieve consensus, regardless of the cost. Unfortunately, more important and controversial subjects generate more disagreement, and consensus may be gained only by emasculating the product report. The second type of report is motivated by an underwhelming desire to achieve consensus. Controversy is endemic, and the recommendations may become so diffuse and contradictory that they are also virtually worthless. Obviously, one should seek the happy medium. Consensus without emasculation is the target. But a solid, defensible minority view is sometimes as valuable as a meaningful consensus. Both will be welcomed as we formulate our conclusions and recommendations.

#### REFERENCES

- [1] Reinhart, E.E., and others. The Impact of WARC-79 on the Broadcasting-Satellite Service. IEEE Trans. Commun., COM-29, pp. 1193-1209.
- [2] FCC 80-417, General Docket No. 80-398, Notice of Inquiry, adopted July 17, 1980, released July 25, 1980.
- [3] FCC 81-181, General Docket No. 80-603, Notice of Proposed Policy Statement and Rulemaking, adopted April 21, 1981, released June 1, 1981.
- [4] CBS application to FCC for authority to construct an experimental and developmental high-definition television satellite system in the 12 GHz band, July 16, 1981.

- [5] DBSC application to FCC for authority to construct and operate a common carrier transmission service by direct-satellite-to-home broadcasting facilities, July 16, 1981.
- [6] GSC application to FCC for interim authority to construct and operate a direct broadcast satellite, July 15, 1981.
- [7] RCA application to FCC for authorization of a satellite communications system under interim direct broadcast satellite rules, July 16, 1981.
- [8] STC application to FCC for authority to construct an experimental DBS system for use in providing subscription television service, December 17, 1980.
- [9] USSB application to FCC for authorization of an interim DBS system, April 30, 1981, as amended July 15, 1981.
- [10] VSS application to FCC for authority to implement phase I of an experimental and developmental system to operate a DBS service, July 14, 1981.
- [11] WUTC application to FCC for authority to construct a new DBS system and for modification of authority to lease capacity in Advanced Westar for use in providing an interim experimental DBS system, July 16, 1981.

## THE FRENCH BROADCASTING SATELLITE PROGRAM

A. Pouzet

*Centre National d'Etudes Spatiales, France*

### ABSTRACT

France decided, together with Germany, to start the development of a preoperational broadcasting satellite system. A bilateral program was undertaken for the development fabrication and launch of two satellites. This paper describes the main objectives, the system requirements and the basic technical characteristics of the French TDF 1 Broadcasting Satellite.

### KEYWORDS

Broadcasting Satellite; mission requirements; satellite modules; modular configuration; growth potential; mass budget; power budget.

### INTRODUCTION

Work in the field of TV broadcasting satellites has been conducted in France since the early 1970's.

First studies were concentrated especially in technology research and development, in particular in the sector of high power amplifiers in the 12 GHz band and of high power and high efficiency solar generators which are the two main characteristics of a broadcasting satellite compared to a classical telecommunications satellite.

Separate national feasibility studies in France and Germany lead to the conclusions that the technological progress and experience gained by both parties in previous telecommunication satellite projects as SYMPHONIE, give sufficient trust to start the development of preoperational systems and to go by steps to future operational systems.

Considering the similar objectives of their national broadcasting satellite programs and the advantages obtained by national cooperation between both countries, France and Germany decided in 1978 to investigate the possibilities of a bilateral program in broadcasting satellite.

France and Germany have decided, during the top level meeting of October 1979 and February 1980 between President Giscard d'ESTAING and Chancellor SCHMIDT, to under-

take a cooperative program of broadcasting satellites.

A memorandum of understanding was signed by the relevant ministers on April 29, 1980, giving the thrust to a program which has as main objectives to start French and German preoperational national broadcasting satellite systems and to form a joint French-German industrial team to succeed in fulfilling the foreseen exportation market in the field of broadcasting satellites.

This paper describes the objectives the system requirements and the basic technical characteristic of the French program.

#### THE OBJECTIVES OF THE BILATERAL PROGRAM

The program will encompass the development, fabrication and launch of two preoperational satellites (besides a spare satellite stored on ground for back-up). This satellite will call upon common technical solutions and will be identical, except for the equipments which have to be specific due to the differences in the mission requirements. These requirements were defined by the national administrations which will use and operate the satellites.

One of the satellites shall be used by the German (TV-SAT), the other by the french (TDF 1), both in conformity with the WARC (World Administrative Radiocommunication Conference of Geneva 1977 and 1979) allocation for national utilization.

Each satellite shall be able to transmit three broadcasting channels. Both shall be launched by the ARIANE launcher in 1984.

The main objective of the program is :

- to acquire the experience for the realization and the utilization of this new type of satellite
- to create the technical and industrial conditions for the future fabrication and sale of this new product on a market which will obviously soon develop in the world.

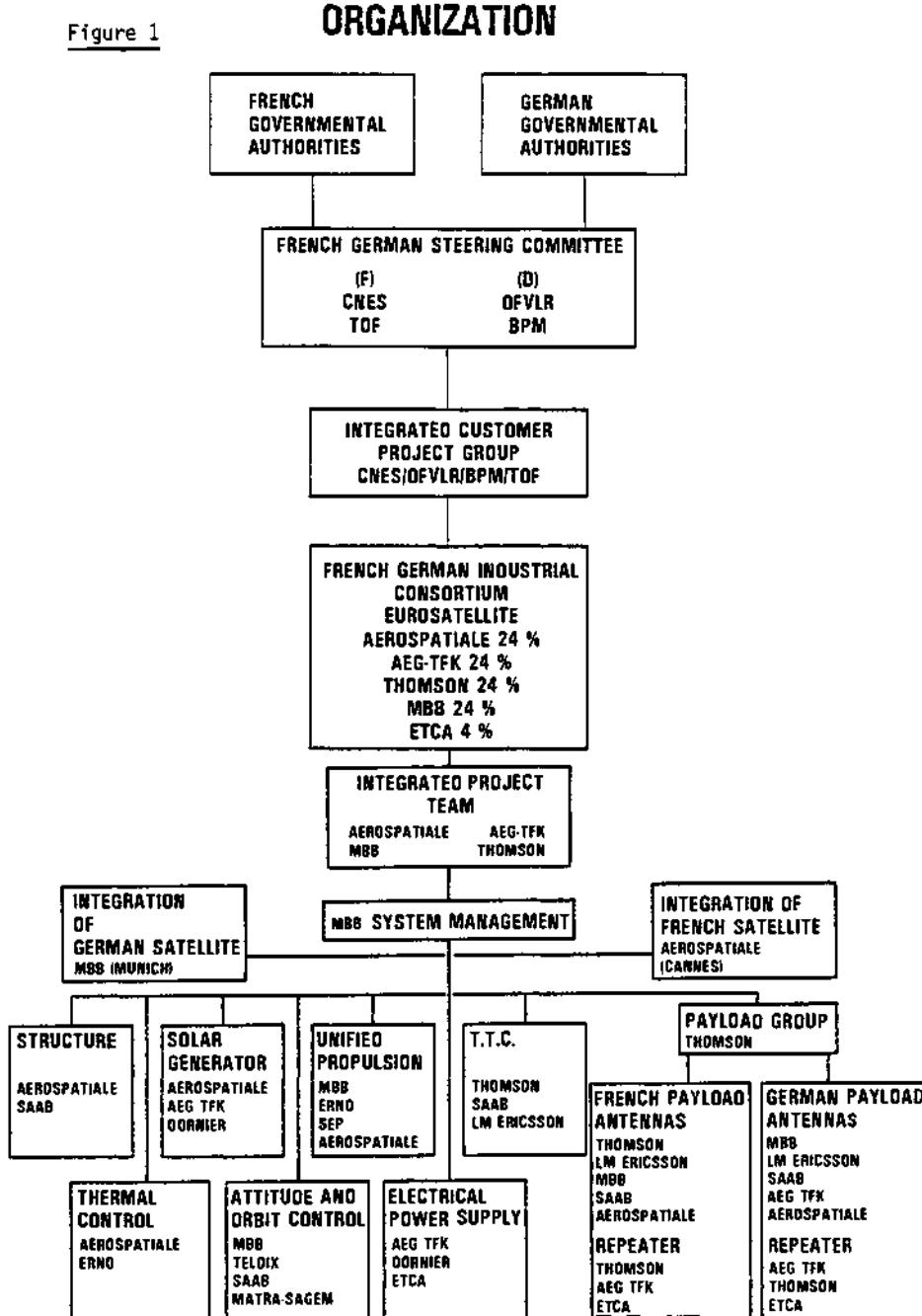
On the industrial level, main French and German firms involved in the development and fabrication of satellites, form a joint team on the basis of equally shared duties and rights between members for the industrial phase. For the development phase of the preoperational system, the overall system and the communication payload responsibilities are attached to a German respectively French company. Major industrial firms have decided to join their effort ; a consortium EUROSATELLITE has been built with MBB and AEG/TELEFUNKEN for Germany, with AEROSPATIALE and THOMSON for France, and with ETCA for Belgium.

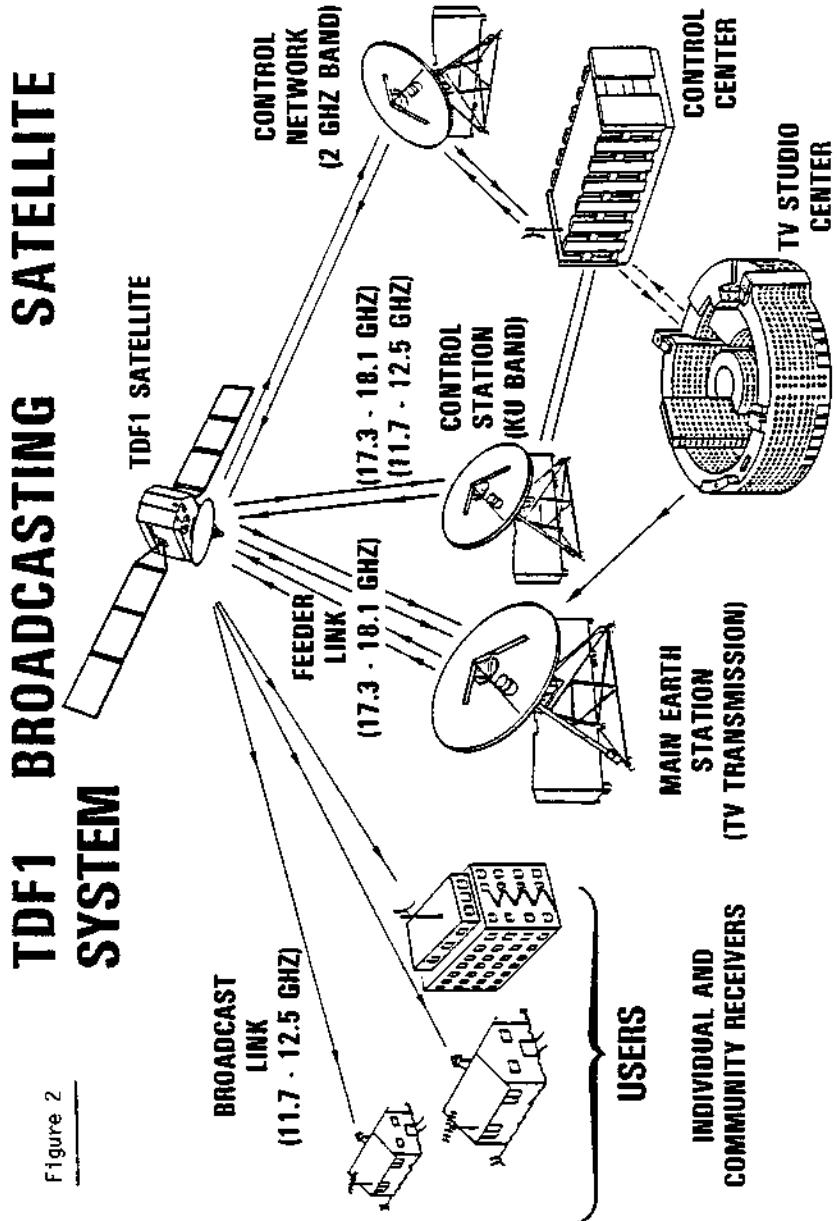
The funding of the preoperational spacecraft will be shared in equal parts between France and Germany. The industrial organization, contractor for the satellite shall be controlled by official Franco-German authorities.

Due to the higher participation of France in the Ariane program, the geographical distribution of work for the satellite program will be on the basis of an industrial return of 54 % for Germany, and 46 % for France.

On the administration level, the overall responsibility of the program is devoted to a French-German authority composed of French and German representatives : the Steering Committee has four members from TELEDIFFUSION de FRANCE and CENTRE NATIONAL D'ETUDES SPATIALES for the French side, and from DEUTSCHE FORSCHUNGS UND VERSUCHSSTALT FUR LUFT UND RAUMFAHRT and BUNDES MINISTERIUM FUR POST for the German side.

Figure 1





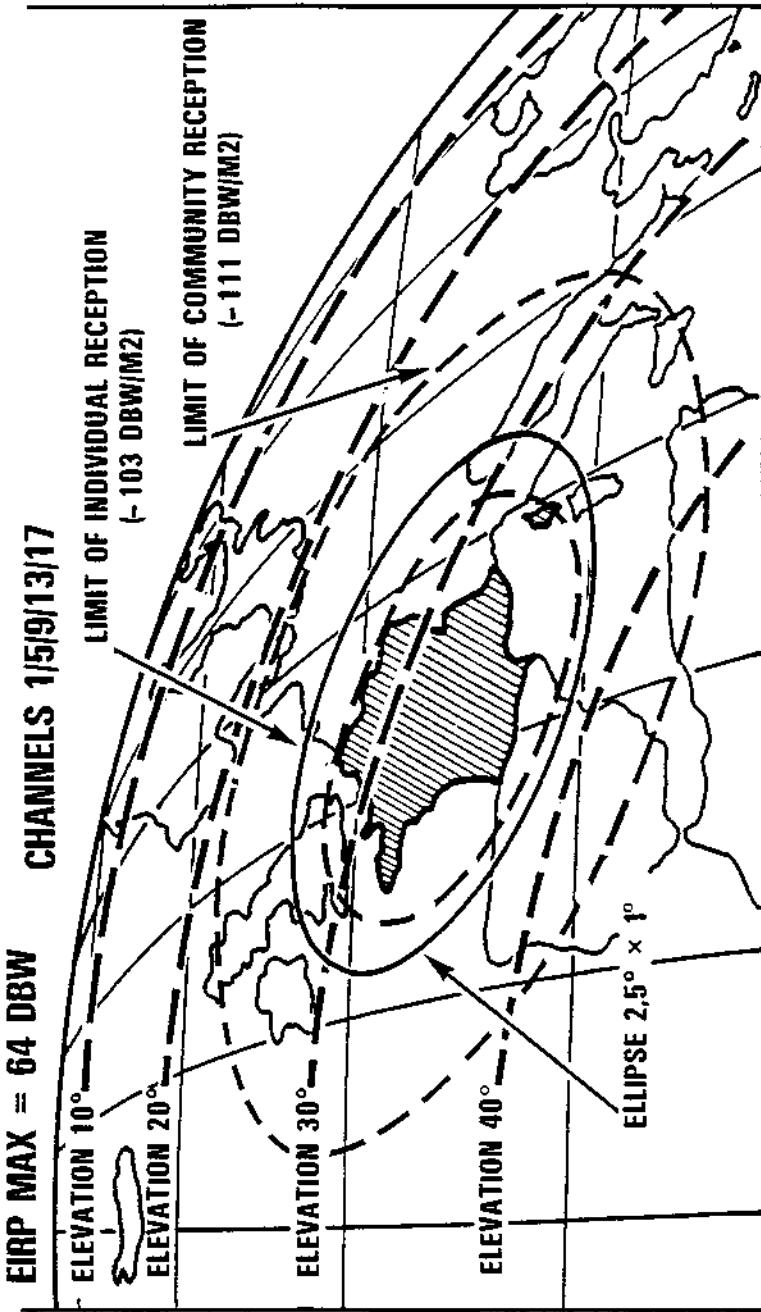
## TDF : SATELLITE REQUIREMENTS

Satellite requirements are derived for the mission from, on one side the WARC regulations, and on the other side the specific user's needs. The Geneva Plan for broadcasting satellite has defined the specifications concerning national allocations for French coverage: ( Figure 3 )

BROADCASTING CHANNELS IN THE 11.7 - 12.5 GHZ BAND	1    5    9    13    17
E.I.R.P. IN DBW	63.8    63.8    63.9    64    64
POLARIZATION	RIGHT HAND CIRCULAR
NOMINAL ORBITAL POSITION	19° WEST
STATION KEEPING	± 0.1° E-W ± 0.1 N-S
LIFETIME	7 YEARS
Tx COVERAGE AREA HALF POWER BEAM WIDTH BORESIGHT COORDINATES ORIENTATION OF THE ELLIPSE POINTING ACCURACY	ELLIPTICAL 2.5° × 0.98° 45.9° NORTH 2.6° EAST 160° < 0.1°
Rx POWER FLUX DENSITY FREQUENCY BAND HALF POWER BEAMWIDTH	FROM -94 TO -76 DBW/M <sup>2</sup> 17.3 / 18.1 GHZ 0.7° x 0.7°
FORESEEN RELIABILITY (AFTER 7 YEARS) SINGLE FAILURE POINT PROB	0.8 < 0.01
TTC FREQUENCIES - TRANSFER PHASE - ON STATION PHASE	2.0 / 2.3 GHZ 17.3 - 18.1 / 11.7 - 12.5 GHZ
LAUNCHER	ARIANE 2

# FRENCH COVERAGE OF BROADCASTING SATELLITE AT 12 GHz

Figure 3



#### DESCRIPTION OF THE SATELLITE

A modular approach has been selected leading to the five following modules (Figure 4) :

- antenna module (AM)
- communications module (CM)
- service module (SM)
- solar array module (SAM)
- propulsion module (PM)

Also after integration of the modules, equipment inside the satellite remains accessible by dismounting the covers.

Figure 5 shows the integrated satellite in launch configuration and transfer configuration, Figure 6 with deployed solar arrays on station.

##### Antenna Module

The satellite TDF 1 uses separate receive and transmit antenna systems. The antenna dishes are mounted on an antenna platform with an antenna tower supporting the RX and TX feed systems. Transmit antenna uses an elliptical reflector of 2,4 x 0,9 m and receive antenna uses a circular reflector of 2,1 m. Transmit antenna is equipped with a multifeed array and a RF sensor according to the monopulse concept. To provide WARC requirements concerning pointing accuracy, Rx and Tx antennas are mounted on APM's with the Rx antenna slaved to the Tx antenna system which lead to a pointing accuracy of 0.05° for Tx beam.

##### Communication Module

The satellite is equipped with a redundant receiver, five non redundant channel amplifiers at all WARC allocated channel frequencies, four non redundant and one redundant transmitter stages according to Figure 7. Design of the power system is such that only three of the five channels can be operated simultaneously. In case that one of the channels fails, another one has to be switched on. Since this will lead to a change of the transmitted channel redundancy concepts but enables experiments to be performed at all frequencies allocated by WARC. The satellite will be equipped equally with TWT's manufactured by AEG/TELEFUNKEN and THOMSON-CSF.

##### Service Module

Attitude and Orbit Control. The Attitude and Orbit Control Subsystem (AOCS) proposed for TDF 1 spacecraft uses the momentum bias principle for 3-axis stabilization. The design utilizes a fixed momentum wheel, sun and infrared earth sensors, gyros, a bi-propellant thruster system (UPS) and an attitude control electronic (ACE), for data handling, command distribution, control law implementation, mode control and surveillance.

Power conditionning. Due to the power range of the payload a sequential digital shunt concept was chosen giving very good efficiency, low dissipation and a high current handling capability.

## TDF1 MODULAR CONFIGURATION

Figure 4

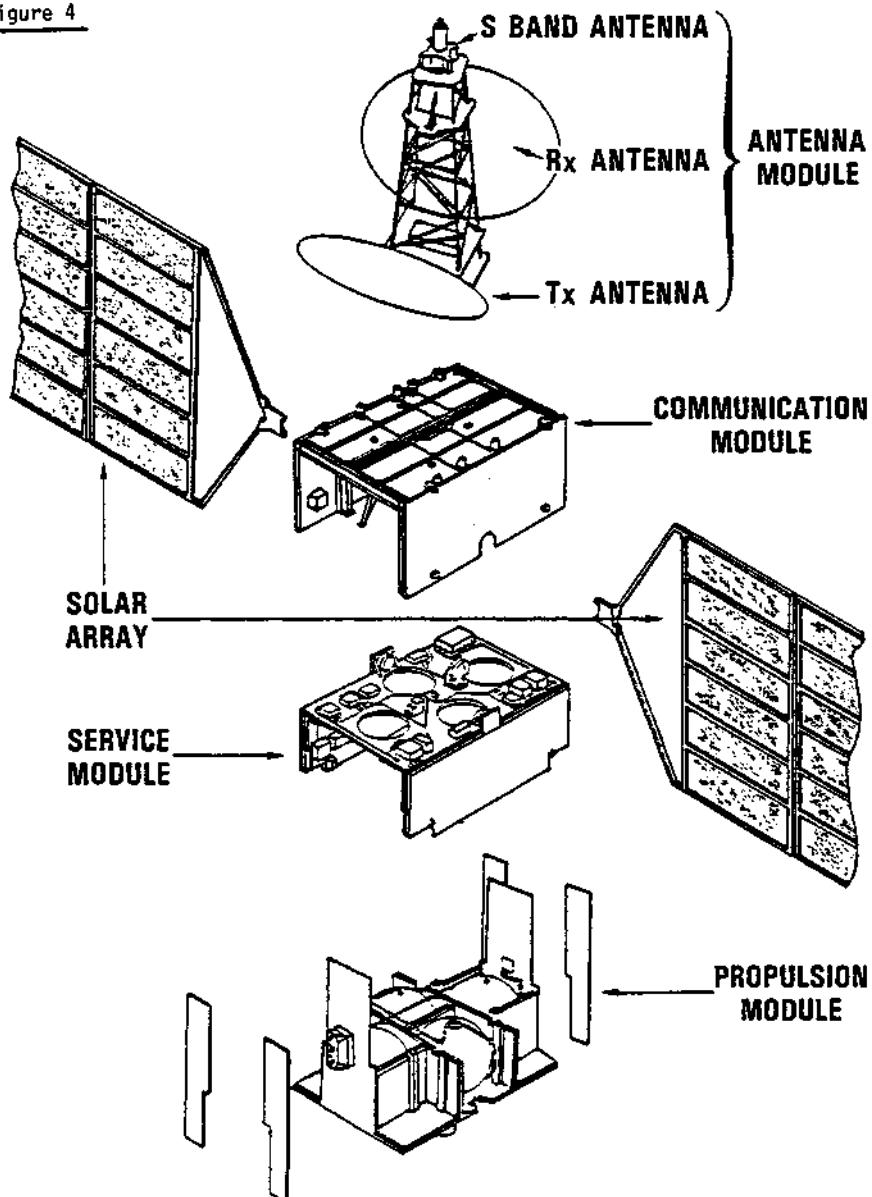


Figure 5

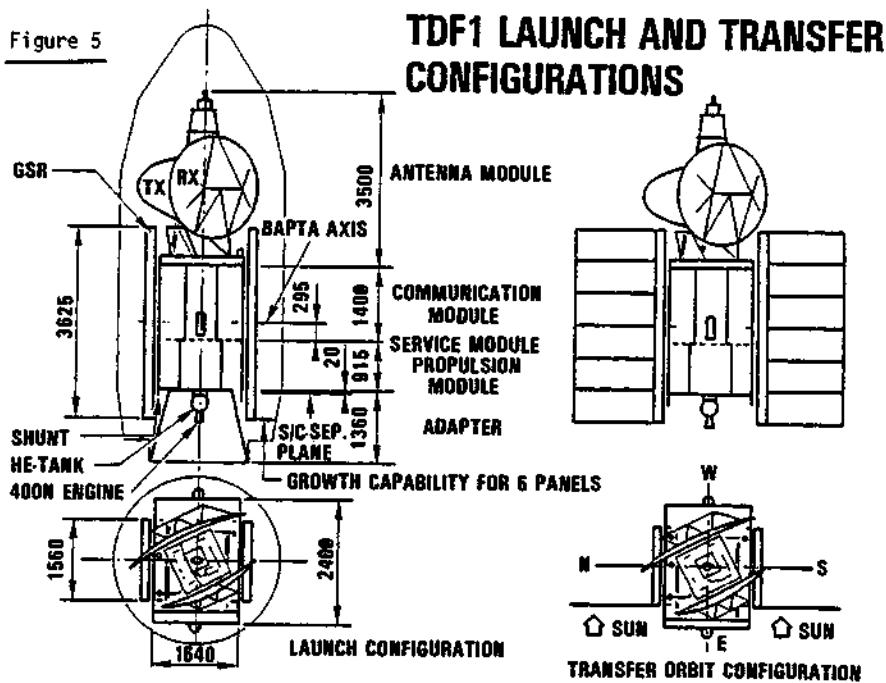


Figure 6

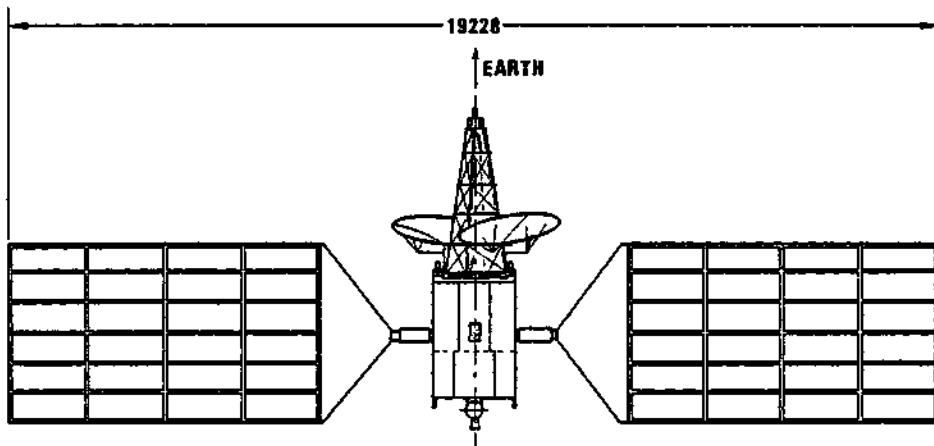
**TDF1 ORBIT CONFIGURATION**

Figure 7

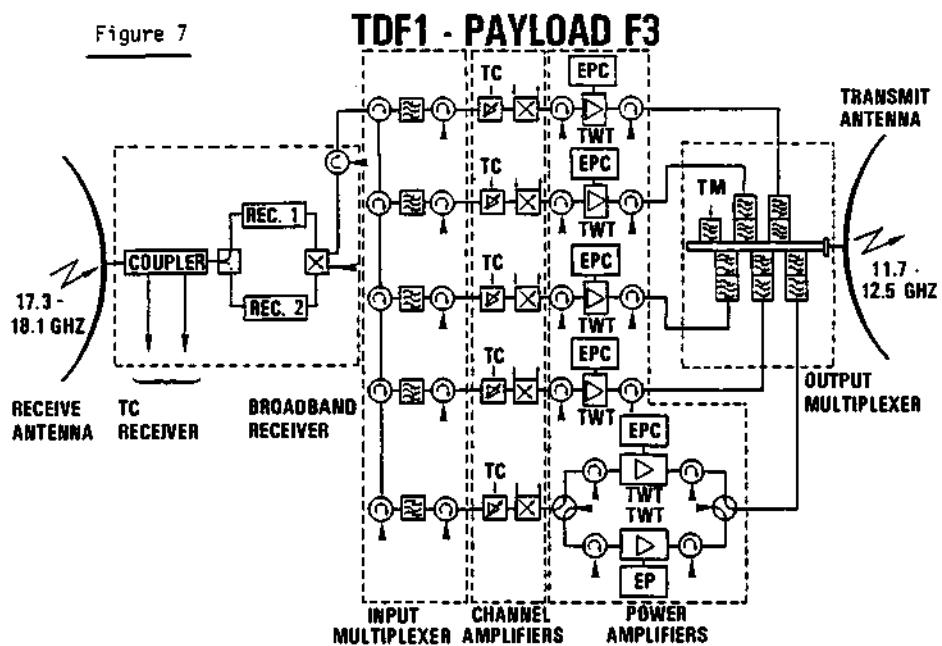
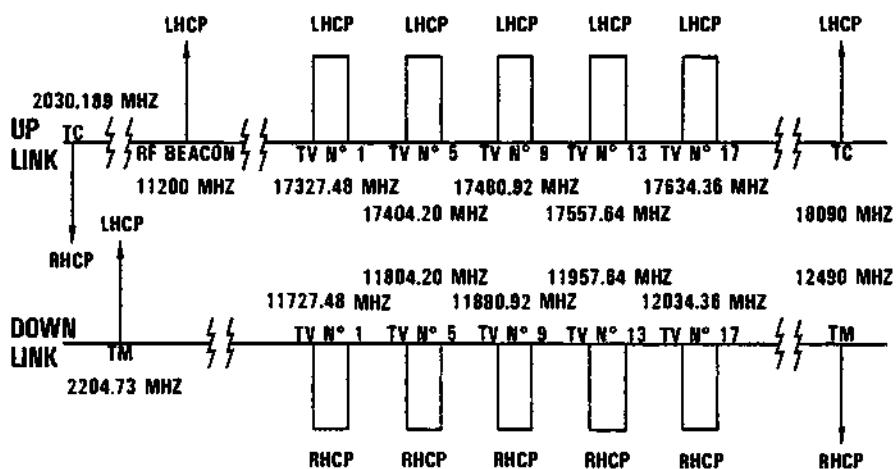


Figure 8

**TDF1 FREQUENCY PLAN**

RHCP : RIGHT HAND CIRCULAR POLARIZATION

LHCP : LEFT HAND CIRCULAR POLARIZATION

According to the mission requirements two separate busses are provided : the Main Bus to the payload (high power during sun phases) and the Permanent Bus to the platform during all phases, including eclipses.

A battery supplies the platform subsystems during launch, transfer and eclipse phases.

TT and C . The telemetry, tracking and command operations of the satellite are available on two links :

- Ku-band through the Payload Antenna and cold redundant Ku-Rx and Ku-Tx during the nominal mission
- S-band through the quasi-omnidirectional antenna (Rx/Tx) at the top of the antenna tower during transfer phase and as a back-up. The command reception is kept completely hot-redundant.

#### Solar Array Module

The technology of the TDF 1 solar arrays is based on the lightweight TELECOM 1 design.

The dimensions of the solar array panels are optimized in terms of maximum active surface versus available volume under the ARIANE shroud, which leads to a 1.6 x 3.6 m size.

The power capability at end of life (7 years) for the basic set of 4 panels per wing is more than 3 060 W.

The growth capability of the solar array extends to 2 x 9 panels. A trade-off study for the transfer orbit configuration has indicated that a partial deployment of 2 x 1 panels is adequate.

#### Propulsion Module, Unified Propulsion System

All necessary propulsion maneuvers (apogee maneuver, station-acquisition and keeping, attitude control torques) will be performed by one propulsion system, the Unified Propulsion Subsystem (UPS).

The UPS is based on SYMPHONIE propulsion technology for the 400 Newton thruster for the AEM and for the 10 N thrusters which are used for all other propulsion needs in a redundant configuration. A four tank (surface tension tank) arrangement is installed in the Propulsion Module to store the necessary propellants (MMH and N<sub>2</sub>O<sub>4</sub>). Two high pressure tanks contain helium for pressurisation. The tanks can be easily adapted to ARIANE 2 or ARIANE 3 capability by introducing a cylindrical segment in the propellant tanks and varying wall thickness of the high pressure tanks.

#### Thermal Control

Due to the high dissipation in the Communication Module equipment and the power subsystem, only the extensive use of heat pipes allows the height of the satellite to be kept to a minimum.

The Communication Module is thermally decoupled from the Service and Propulsion Modules to minimize reciprocal action between the modules and achieve a modular thermal control.

The high variations of dissipations in the Communication Module also leads to the use of actively simulated loads when the travelling wave tube amplifiers are swit-

ched off. In addition some specific equipments need an automatic thermal control.

#### Structure

The main feature of the structure is its modular concept. Three modules compose the satellite body structure (excluding solar array and antennas).

- Communication Module
- Service Module
- Propulsion Module.

The structure is designed to accomodate a Communication Module with height and mass capabilities adapted to more powerful missions.

#### Growth potential

The design of the French-German platform is fitted to the three operating channels satellites TDF 1/F3 and TVSAT/D3 but is conceived to allow with only minor changes more important missions. In particular, five operating channels satellites for French coverage (F5) as well as for German coverage (D5) shall be accommodated by the common platform.

### MAIN BUDGETS

#### Mass Budgets (in kg)

	F3	F5
Repeater	111.9	178
Antenna	92	92
Power	83.2	98.2
Solar Array	149.3	215.7
TTC	33.5	33.5
AOCS	48.3	48.3
UPS	106.8	113.8
Structure	171.6	171.6
Thermal	75	83
Elec. Ins. Syst.	65.9	74.7
Balance	5	10
Total dry	942.17	1 118.8
Propellants	1 003.8	1 207.1
Adaptor	47.5	47.5
Launch	2 065	2 425
Margin	71 : 7.5 %	51.6 : 4.6 %

## Power Budgets (in W)

	SUMMER SOLSTICE EOL		EQUINOX EOL	
	F3 3 channels	F5 5 channels	F3 3 channels	F5 5 channels
BUS	590	680	590	680
Thermal	140	80	280	230
Battery Charging	20	20	95	95
Payload	2 125	3 575	2 125	3 575
TOTAL	2 875	4 355	3 090	4 580
NB panels	2 x 4	2 x 6	2 x 4	2 x 6
SA Power	3 060	5 490	3 312	4 968
Margin	6.5 %	5.4 %	7.2 %	8.5 %

TDF<sub>1</sub> - TV SAT FORESEEN DEVELOPMENT SCHEDULE

	1981	1982	1983	1984
MILESTONES	PDR CDR MDMR /OTDR FAR LAUNCH			
LLI.				
G.S.E.				
STM (STM-STM-S-STNL)	EQTS REALIZATION AIT			
EMT	EQTS REALIZATION AIT			
PFM 1	EQTS REALIZATION MODULE AIT SYSTEM AIT SHIPMENT - PREPARATION & LAUNCH			
PFM 2	EQTS REALIZATION MODULE AIT SYSTEM AIT SHIPMENT - PREPARATION & LAUNCH			

## OPERATIONAL BROADCASTING SATELLITE PROGRAM IN JAPAN

Y. Ueda\*, K. Ishibashi\*\* and Y. Ichikawa\*\*\*

\**Telecommunications Satellite Corporation of Japan, 2-12-18 Shiba-Daimon,  
Minato-Ku, Tokyo 105, Japan*  
\*\**Japan Broadcasting Corporation (NHK), 2-2-1 Jinnan, Shibuya-Ku,  
Tokyo 150, Japan*  
\*\*\**National Space Development Agency of Japan, 2-4-1 Hamamatsu-cho,  
Minato-Ku, Tokyo 105, Japan*

### ABSTRACT

The Broadcasting Satellite-2 (BS-2) Program is intended to establish the first Japanese operational domestic satellite broadcasting system using K-band frequencies allocated to Japanese domestic satellite broadcasting service by 1977-WARC-BS. This paper presents the outline of the BS-2 program, including mission objectives, spacecraft, ground system and home receiver.

### KEYWORDS

Broadcasting Satellite-2 (BS-2); Medium-scale Broadcasting Satellite for Experimental Purpose (BSE); DBS; Home Receiver;

### INTRODUCTION

The BS-2 program, in request of Japan Broadcasting Corporation - Nippon Hoso Kyokai (NHK), has been planned by the Ministry of Posts and Telecommunications (MOTP) and authorized by the Space Activities Commission (SAC) as a successor to the Japanese Medium-scale Broadcasting Satellite for Experimental Purpose (BSE).

The space segment of the BS-2 system consists of BS-2a and BS-2b, each positioned at 110°E longitude. The BS-2a and BS-2b are now under development by National Space Development Agency of Japan (NASDA) and will be launched in early 1984 and mid 1985 respectively at Tanegashima Space Center of NASDA by a Japanese N-II Launch Vehicle.

NASDA is responsible for launching, positioning on the geostationary orbit and initial checking up of the BS-2. The operation and control of the BS-2 are then handed over to Telecommunications Satellite Corporation of Japan (Telesat Japan).

NHK plans to establish operational direct broadcasting satellite services to solve difficulties of terrestrial television signal receptions and also to aim at opening new area for future applications of DBS.

### MISSION OBJECTIVES

The mission objectives of the BS-2 program are;

- a. to perform the operational broadcasting satellite service mainly as to solve TV reception difficulties of about 400 thousands households in mountainous areas and remote islands of Japan,
- b. to develop some advanced service with the satellite,
- c. to acquire the technology of the broadcasting satellite.

As shown in Fig. 1, the satellite broadcasting system of the BS-2 is composed of NASDA's tracking & control station, spacecraft of BS-2a and BS-2b, NHK's earth stations and related facilities, Telesat Japan's operation and control station and numerous home receivers of the general public.

The spacecraft of BS-2a and BS-2b now under development by NASDA are scheduled to be launched by Japanese N-II launch vehicle from NASDA's Tanegashima Space Center in early 1984 and mid 1985 respectively. BS-2b, having the same capabilities as BS-2a, is to back up the former spacecraft BS-2a in orbit.

Spacecraft's operation and control succeeding the launch is to be carried out by NASDA up to completion of the initial performance check up of the spacecraft in orbit, and then the operation and control be handed over to Telesat Japan which undertakes the whole responsibilities of the operation and control of the spacecraft throughout the life.

The whole broadcasting transponders onboard are to be utilized by NHK, using its feeder link earth station with telemetry and command function for the transponder operation. The general public could enjoy NHK's two channels of color TV broadcasting satellite program with their simple home receivers.

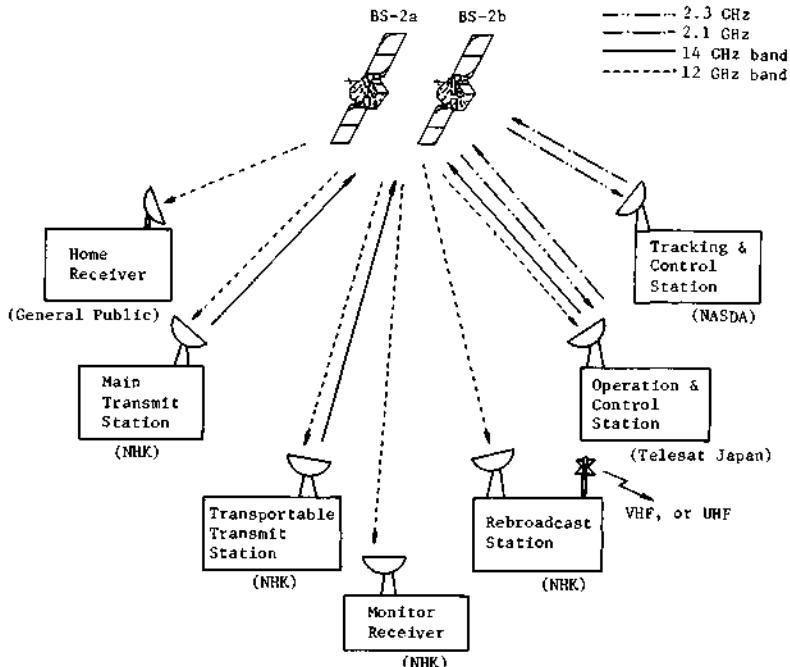


Fig. 1 The satellite broadcasting system of the BS-2

### SATELLITE SYSTEM

The spacecraft orbital configuration is shown in Fig. 2. The spacecraft is three axis stabilized and provides a fixed antenna platform, North/South equipment panels and a sun oriented solar array for power generation.

This design is almost same as former launched BSE except for the following items.

- a. WARC requirements concerning frequency assignment, polarization, RF beam pointing accuracy and sidelobe characteristics of transmitting antenna.
- b. Launch using the Japanese N-II Rocket instead of Delta 2914 launch vehicle.
- c. Five years design life compared with three years of BSE.

The rectangular center equipment bay has a dimension of 120 cm and 132 cm, and total length from the top of antenna support to the nozzle of the apogee kick motor is 283 cm. The total length of the solar array, which is folded at launch and deployed in orbit, is 890 cm. The weight of the BS-2 is 670 kg at launch and 350 kg after firing the apogee kick motor propellant.

The spacecraft is composed of the following subsystems; Communication subsystem, Tracking Telemetry and Command subsystem, Electrical Power subsystem, Attitude Control subsystem, Apogee Kick Motor, Secondary Propulsion subsystem, Structure and Thermal Control subsystem.

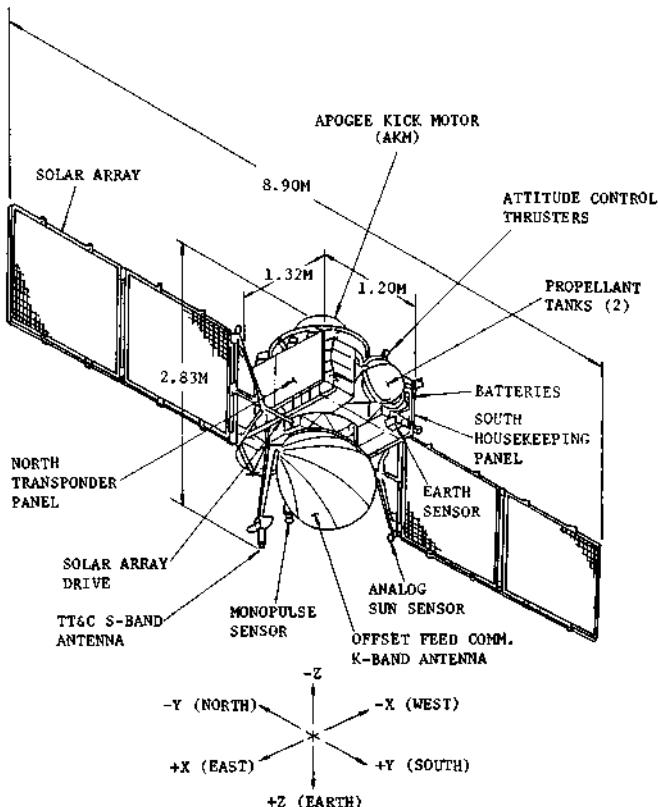


Fig. 2 Spacecraft orbital configuration

Communication equipments are mounted on the north panel and bus equipments such as telemetry and command, attitude control and power control are mounted on the south panel of the spacecraft. The functional block diagram of the BS-2 is shown in Fig. 3 and the summary description of the BS-2 is shown in Table 1.

Communication subsystem of the BS-2 is composed of a communication transponder and a high gain antenna. Channel 11 and 15 of the 12 GHz band, assigned to Japan in WARC BS, are used for down link frequency, and 14 GHz band which has frequency difference of about 2.3 GHz is used for up link. The planned K-band frequency assignments are summarized in Fig. 4 including TT&C frequencies.

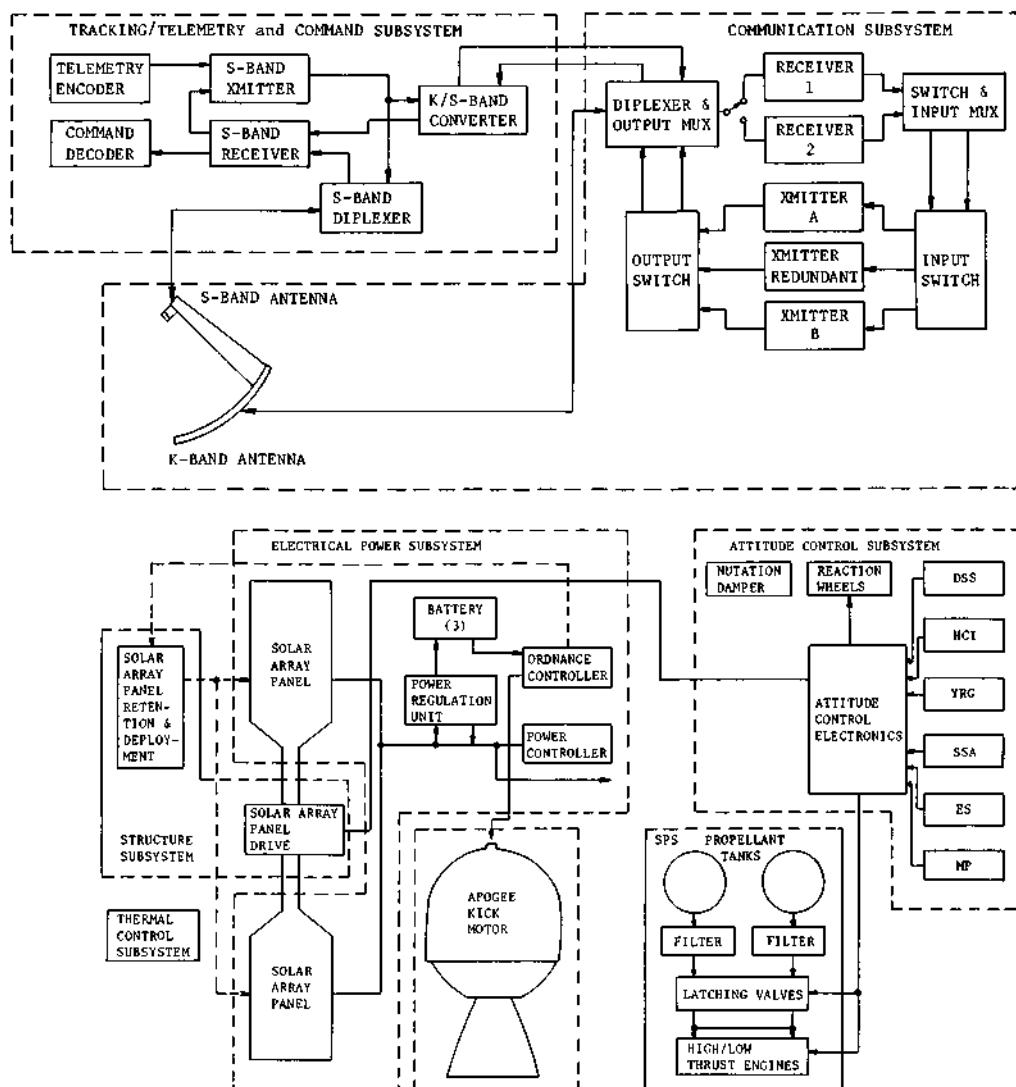


Fig. 3 Functional block diagram of the BS-2

Table 1 BS-2 Summary

Physical Configuration and Dimensions	Fixed antenna, deployed solar arrays, modular assembly spacecraft body : 120 cm x 132 cm x 283 cm deployed solar array : 890 cm
Weight	670 kg (at launch) 350 kg (on station)
Stabilization	3-axis stabilization (zero momentum)
Communications	
Antenna	Offset feed multi-horn shaped beam antenna
Transponders	14/12 GHz (K-band) 100 W, 2 channels
TT & C	
Antennas	S-band Omni-directional K-band Communication antenna
TT & C Equipments	NASA STDN compatible S-band transponders K-band/S-band converter, telemetry encoder and command decoder
Electrical Power	
Solar Array	BOL (winter/summer solstice) 900 W EOL (winter/summer solstice) 825 W
Batteries	4AH x 3
Communication Antenna	
Pointing Accuracy	Less than $\pm 0.1$ deg ( $3\sigma$ )
Beam Rotation	Less than $\pm 0.6$ deg ( $3\sigma$ )
Thermal Control	Passive system, augmented by heaters
Reaction Control	Hydrazine monopropellant system
Apogee Kick Motor	Thiokol STAR-27
Orbit Position	
BS-2a and BS-2b	110 deg, East Longitude
Station Keeping	Latitude less than $\pm 0.1$ deg Longitude less than $\pm 0.1$ deg
Life	Five years (goal)
Launch Vehicle	Japanese N-II Rocket

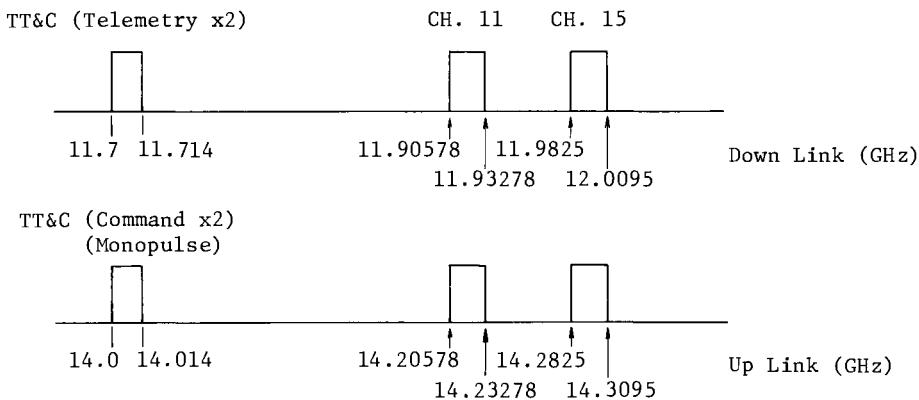


Fig. 4 Planned frequency assignment (K-band)

The transmitting subassembly of the communication transponder has 2 prime and 1 redundant channels of 100 watt helix TWT amplifier. Receiving subassembly has fully redundant GaAs FET amplifiers. The block diagram of communication subsystem is shown in Fig. 5.

The communication antenna is an offset feed circular polarized type and power splitting 3 horn feed provides the shaped beam gain pattern of 38 dB towards the Japanese mainland and 27 dB towards the Japanese territory including remote islands. The antenna coverage is shown in Fig. 6.

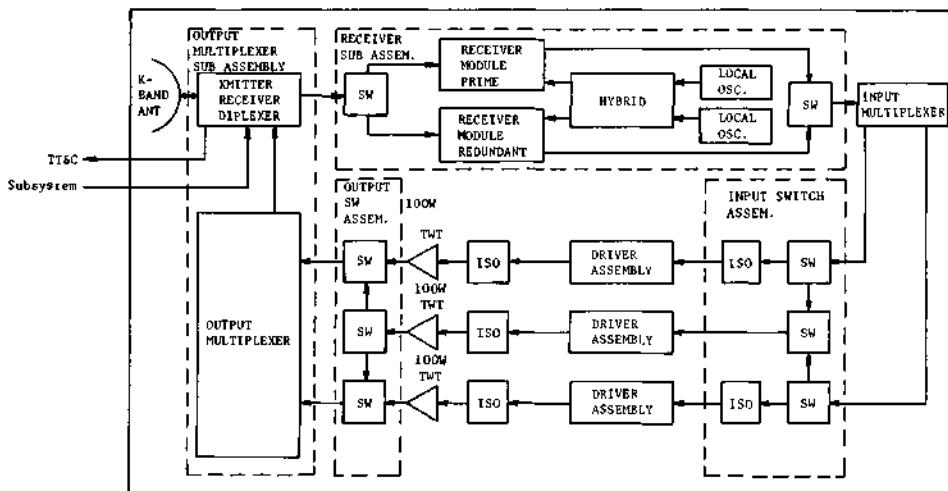


Fig. 5 Block diagram of communication transponder

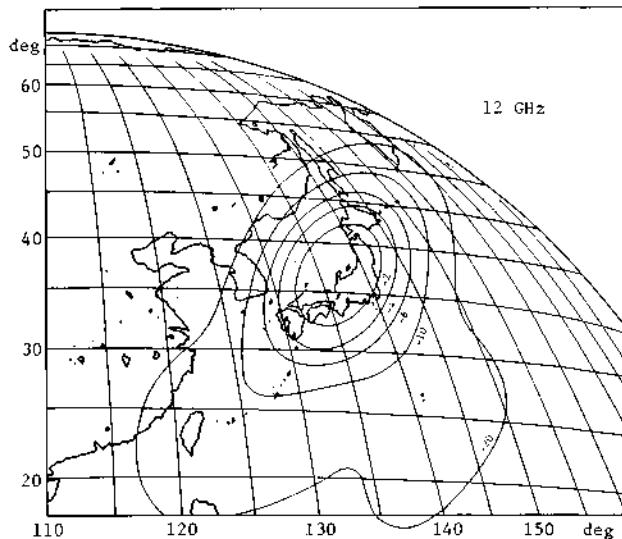


Fig. 6 BS-2 antenna coverage

#### NHK'S EARTH STATIONS AND FACILITIES

NHK plans to construct its earth stations and related facilities so as to achieve the following;

- a. to perform the operational broadcasting satellite service to solve TV reception difficulties,
- b. to secure the nationwide terrestrial TV broadcasting network even when some terrestrial microwave links be snapped by a natural disaster,
- c. to gather and relay directly programs with a transportable earth station from any spot including mountainous area and far remote islands where program relay might be impossible with the conventional terrestrial means of communication.

NHK also conceives to have some additional apparatus so as to develop some advanced service with satellite in such new broadcast systems as the still picture broadcasting, the high definition TV and the high fidelity sound broadcasting.

The construction plan includes such facilities as a program switching and transmission facilities in the broadcasting center in Tokyo, a main transmit station with a feeder link of two TV programs also in the broadcasting center, a few transportable transmit stations with a feeder link of a single TV program, about 60 monitor receivers scattered throughout Japan and some translator stations in far remote islands.

The main transmit station might have an additional site diversity antenna in the later phase after detailed investigation into the rainfall attenuation and receiver dissemination.

Prototype transportable transmit stations of van type (type B) and easy to set up type (type A), which have proved to be substantially operational through BSE experiments started in 1973, are to receive some modifications in frequencies and others for operational use. Table 2 gives an outline of the earth stations and the related facilities.

#### SATELLITE HOME RECEIVER

##### Present status of home receiver development

A satellite home receiver, catching DBS signal with a small antenna, requires low noise in low cost. Since 1971, NHK Technical Research Laboratories had succeeded in developing a low noise satellite receiver with the state of the art technology of the planar circuit which is quite suitable for mass production. In 1976, CTS experiment had proved high rank of it's capability. Furthermore, two years of the BSE experiment since 1978, together with tests by domestic manufacturers, have proved the receiver to be suitable for operational use in such as low noise, immunity against various types of interference and long term stability of the characteristics.

##### Outline of the receiver

A typical home receiver is composed of a parabolic antenna of about 1 meter in diameter, an outdoor unit of SHF/UHF down converter directly attached at the back of the antenna, an indoor unit of FM/AM converter and a conventional home TV receiver. A typical characteristics of the home receiver having the planar circuit is shown in Table 3.

Table 2 NHK's Earth Stations and Related Facilities

Station & Facility	Capability	Location
Main Transmit Station	<ul style="list-style-type: none"> <li>- To transmit two channels of color TV programs with the feeder links in 14 GHz band,</li> <li>- To command the onboard broadcast transponders on/off and to collect telemetry data,</li> <li>- To have an order wire channel,</li> <li>- Transmit Antenna —— about 6 m diameter Transmit Power —— about 2 kW Receiver G/T —— min. 26 dB/K</li> </ul>	NHK Broadcasting Center (Tokyo)
Program Switch and Transmission Facility	<ul style="list-style-type: none"> <li>- To handle two channels of color TV programs with multiple access,</li> <li>- On Air Programs Switching Matrix, VTR, Flying Spot Scanner, Announce Machine, etc.</li> </ul>	NHK Broadcasting Center (Tokyo)
Transportable Transmit Station	<ul style="list-style-type: none"> <li>- To transmit single color TV program with the feeder link in 14 GHz band,</li> <li>- To have multiple access function to the spacecraft,</li> <li>- To have an order wire channel,</li> <li>- Transmit Antenna —— 2 ~ 3 m diameter Transmit Power —— about 2 kW Receiver G/T —— min. 18 dB/K</li> </ul>	To park usually at NHK B.C., to move out nationwide
Monitor Receiver	<ul style="list-style-type: none"> <li>- To monitor signal strength variation of DBS signal in 12 GHz band,</li> <li>- To monitor video/audio signal quality,</li> <li>- To transfer received broadcast signal level, rain fall data, to the Main Transmit Station</li> <li>- Receive Antenna —— 1 ~ 3 m diameter Receiver G/T —— 12 ~ 21 dB/K</li> </ul>	Regional stations of NHK (about 60)
Rebroadcast Station	<ul style="list-style-type: none"> <li>- To receive two channels of broadcast signals in 12 GHz band and to rebroadcast it in VHF or UHF band,</li> <li>- Receive Antenna —— about 4.5 m diameter Receiver G/T —— min. 24 dB/K Rebroadcast Power —— 100 W in VHF 1 W in UHF</li> </ul>	Minami-daito Isl. (about 1,300 km SW of Tokyo)  Bonin Isl. (2) (about 1,000 km S of Tokyo)

Table 3 A Typical Characteristics of a Home Receiver

Noise Figure (Bandwidth)	----- 3.4 dB (300 MHz) ~ 4 dB (800 MHz)
Image Rejection	----- more than 40 dB

## OPERATION AND CONTROL SYSTEM

NASDA

The operation and control system of NASDA is composed of Tsukuba Space Center and three tracking and control stations located at Katsura, Masuda and Okinawa in Japan. Tracking and control during launch and initial station acquisition phase will be supported by several NASA STDN stations.

After the transfer orbit injection of the spacecraft, the maneuver of the spacecraft attitude and apogee motor firing will be conducted and injected into the drift orbit. In the drift orbit, 3 axis attitude acquisition and geostationary orbit acquisition will be conducted. During the first three months after launch, the operation and control of the satellite will be carried out by NASDA and continued up to completion of the initial performance check up of the spacecraft.

Telesat Japan

The Telesat Japan is now constructing the operation and control earth station at Kimitsu city in Chiba prefecture, about 130 km south-east of Tokyo.

The operation and control system for BS-2 program has the following function;

- a. to maintain the satellites (BS-2a, BS-2b) in the same orbital position of 110°E within a specified limit,
- b. to control the satellite subsystem equipments, including attitude of satellite, and
- c. to check up the characteristics of the satellite subsystems periodically.

Table 4 shows an outline of the operation and control facilities planned at Telesat Japan for BS-2 program.

Two K-band antennas with a diameter of 10 m and an S-band antenna with a diameter of 18 m will be used for the TT&C operation. Each of these 10 m antennas will be pointed to BS-2a and BS-2b respectively.

A K-band antenna with a diameter of 5 m will also be added for transmitting monopulse signal to both BS-2a and BS-2b.

Another K-band antenna with a diameter of 10 m will also be used for precise angle measurement in order to improve the accuracy of orbit determination.

## ACKNOWLEDGEMENT

The authors would like to express their gratitude to the staff of MOPT and other agencies participated in the initiation of the BS-2 program.

Table 4 Outline of the Operation and Control Facilities planned at Telesat Japan

Facilities	Number of antenna	Mount system	Function	Characteristics	
				e.i.r.p. (dBm)	G/T (dB/K)
18 mΦ antenna*	1	AZ-EL	S-band TT&C/Ranging  ( used in common for both trans- mitting and receiving )	~ 106.5	~ 27.9
10 mΦ antenna*	1	AZ-EL	K-band Angle measure- ment  ( used only for ) ( receiving )	—	~ 30.5 (at K-band)
10 mΦ antenna	2 each for BS-2a and BS-2b re- spectively	AZ-EL	K-band TT&C/Ranging  ( used in common for both trans- mitting and receiving )	~ 108.1	~ 30.5
5 mΦ antenna	1	Semi-fixed	K-band Monopulse  ( used only for ) ( transmitting )	~ 106.9	—

\* : used in common with CS-2 (Communications Satellite-2)

## PASSIVE MICROWAVE SYSTEMS

E. Schanda

*Institute of Applied Physics, University of Bern, Switzerland*

### ABSTRACT

Passive microwave sensors are presently applied on spaceborne platforms for investigations of large scale phenomena in geoscience and hydrology as well for the regular observation of meteorological features. Results of ground-based and airborne experiments over well-controlled test areas together with the present experience in synoptic earth observations from space can now be used for defining sensor systems for operational application in various fields of geoscience. Two examples are discussed: First the elements of a microwave payload for the global observation of the extent, the development and the hydrological state of the snow cover. Second a payload for the measurement of the three-dimensional distribution of several minor constituents and of the temperature of the middle atmosphere by technique of limbsounding from Spacelab.

### KEYWORDS

Remote sensing, microwave, radiometry, snow, atmosphere

### ORBITING PASSIVE MICROWAVE SENSORS

The obvious advantages of microwave radiation data for various earth-sciences and the relative ease of constructing passive microwave sensors are the motivations since more than a decade to put passive microwave sensors into orbit. These sensors have delivered large amounts of observational data for many successful applications in earth-research. After the first successful Soviet satellites with passive microwave sensors on board, Cosmos 243 and Cosmos 384 (Basharinov et al. 1971 and 1972), the series of Nimbus - Satellites (5, 6 and 7 with imaging radiometers ESMR, SMMR, and frequency selective spectroradiometers e.g. NEMS) by NASA became the bestknown and widely-used sources of microwave radiometric observations of the earth and its atmosphere (e.g. Staelin et al. 1973, as one of the first publications on this series). Presently orbiting since 1978 is Nimbus 7 with the Scanning Multichannel Microwave Radiometer (SMMR). Figure 1 presents the instrument configuration of the SMMR (Nelsen, 1978).

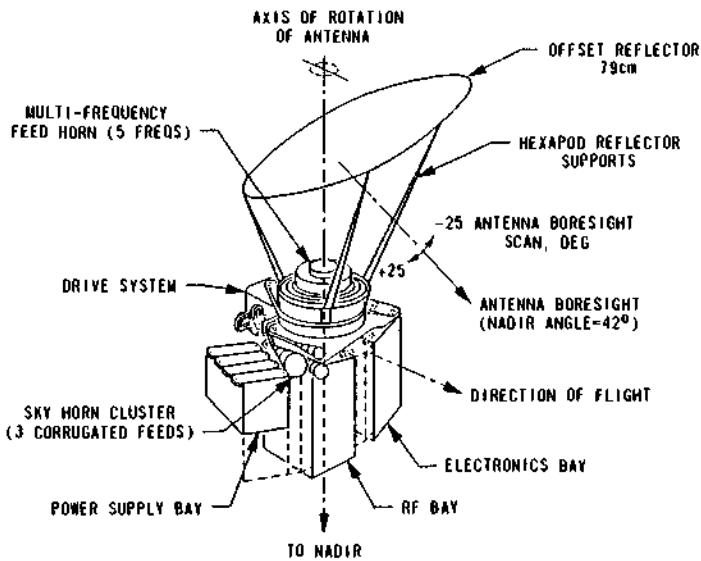


Fig. 1 SMMR Instrument Configuration

SMMR is a five-channel imaging radiometer designed for the determination of environmental parameters as ocean surface wind stress, rainfall rate, atmospheric water vapor and liquid water content, sea ice coverage and dynamics, sea surface temperature and snow cover characteristics over land.

A conical scan antenna with forward viewing incidence angle of 50 degrees and swath of about 800 km on earth symmetrically around the sub-orbital track, yields orthogonal polarisation measurements on all channels. Table 1 gives the most important performance characteristics. Use of one common reflector antenna causes different foot print sizes for the different frequency channels. The absolute accuracy of the brightness temperature measurements is claimed to be 2 degrees.

TABLE I SMMR Performance Characteristics (Nelsen, 1978)

Frequency (GHz)	6.6	10.69	18.00	21.00	37.00
Integration Time (ms) (appr.)	126	62	62	62	30
Dynamic Range (K)	10-330	10-330	10-330	10-330	10-330
Temperature Resolution, $\Delta T_{rms}$ ( $^{\circ}$ K) (per IFOV*)	0.9	0.9	1.2	1.5	1.5
Footprint, km (major/minor)*151/97	91/59	55/41	46/30	27/18	

\*IFOV's are remapped to form equal sized cells (150, 90, 50km) across the swath prior to retrieval of geophysical parameters; the  $\Delta T$  rms's are correspondingly lower.

Among the various investigations utilizing the SMMR microwave brightness temperature data there is one on global mapping of the snow coverage. It is well known from previous ground based and satellite experiments that the snow cover has a pronounced effect on the surface emissivities at microwave frequencies. Therefore SMMR-data can be used to map the snow cover over land and to retrieve snow parameters such as water equivalent and onset of melting. First results of this investigation are now in the process of publication (Künzi, 1981, Künzi et al. 1981, Rott et al. 1981). Three quantities of the snow cover: extent, water equivalent and melting have been retrieved for two time periods in October/ November 1978 and February/March 1979.

The SMMR brightness temperatures measured at 5 frequencies and 2 polarisations were subject to correlation and discriminant analyses with respect to the above snow cover parameters taken from ground-based observations. The statistical analyses resulted in the selection of a simple algorithm based on the spectral gradient of the 18 and 37 GHz brightness temperatures. These are used for discriminating the snow parameters. The global snow cover extent as derived from SMMR data is in good agreement with the snow cover maps produced by the U.S. National Environmental Satellite Service (NESS) based on visible and infrared satellite data. Early in the winter season large snow layers which lasted only for a few days are mapped by the SMMR but are not shown on the NESS-maps. However, the SMMR will neither detect very thin snow layers (< 5 cm) nor - with the present algorithm - wet snow, therefore the total extent of the snow covered region is typically somewhat smaller than given by the NESS-maps. Figure 2 shows the SMMR-derived global snow cover for two periods (three-days-averages each) in spring 1979.

Over land the white area is snow covered. In figure 2A the effect of snow melt can be seen. In southern Scandinavia and in Russia ( $30^{\circ}$ E) no snow is shown; from meteorological data it is known that the snow surface was melting in these regions. In figure 2B the same area appears snow covered; from meteorological data low temperatures are reported. For snow layer thickness of up to  $\sim 80$  cm (corresponding to a water equivalent of  $\sim 20\text{g/cm}^2$ ) the water equivalent can be determined. The accuracy of this parameter is low because of the averaging over the whole pixel size ( $\sim 50$  km diameter) and the fact that this measurement has to be compared with usually one or two point measurements on the ground within this pixel. For selected test sites in uniform planes with a dense network of reporting ground stations a correlation coefficient between ground based and SMMR water equivalent of up to more than 80% can be achieved.

The onset of melting can be obtained from a sudden change of the microwave spectral behavior. There is a difference of brightness temperature gradients  $[T_B(37\text{GHz}) - T_B(18\text{GHz})]$  between day and night of more than 6 K for melting areas, regionally the difference can be more than 20 K.

Concerning passive microwave sensors in orbit, a remark on an operational system: The meteorological satellites NOAA 6 and 7 have four-channal microwave radiometers on board for measurements in the 5 to 6 millimeter wavelength oxygen absorbing region (Schwalb, 1978). Two antennas scan  $\pm 47^{\circ}$  either side of nadir covering a swath width of more than 2000km on the ground, the suborbital ground resolution is 109km. The radiometric sensitivity is 0.3 K within a dynamic range of 0 K to 350 K.

Finally the Indian Space Research Organisation has put into a 500km height orbit its first Satellite for Earth Observation (SEO) with radiometers at 19.3 GHz and 22.2 GHz on board. This system is success-

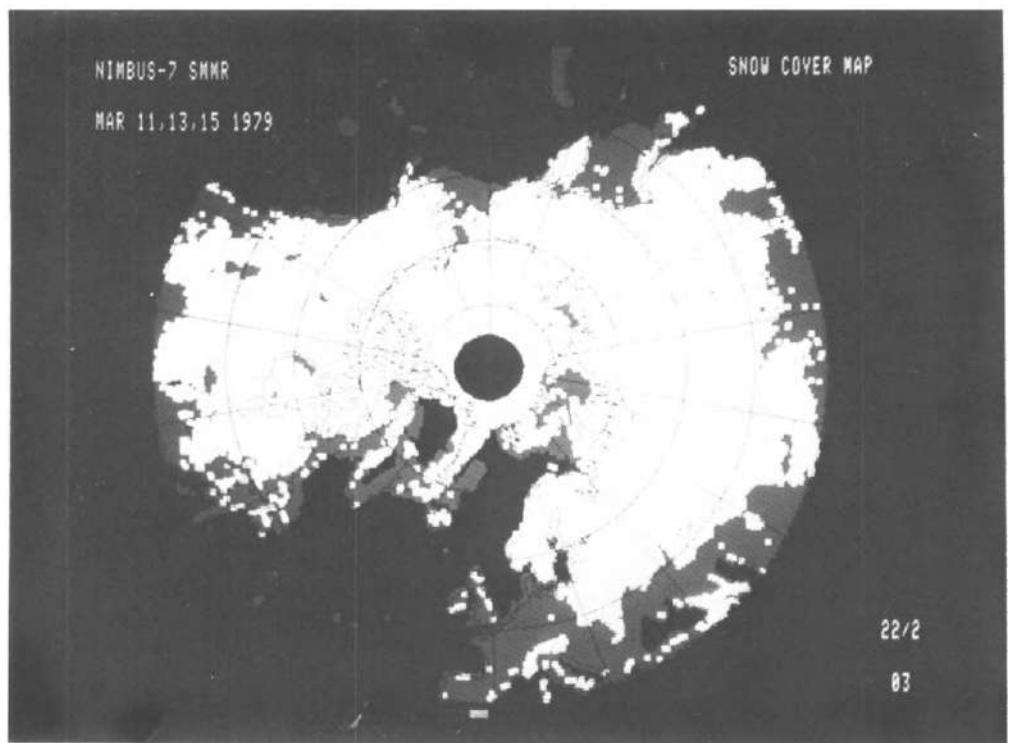
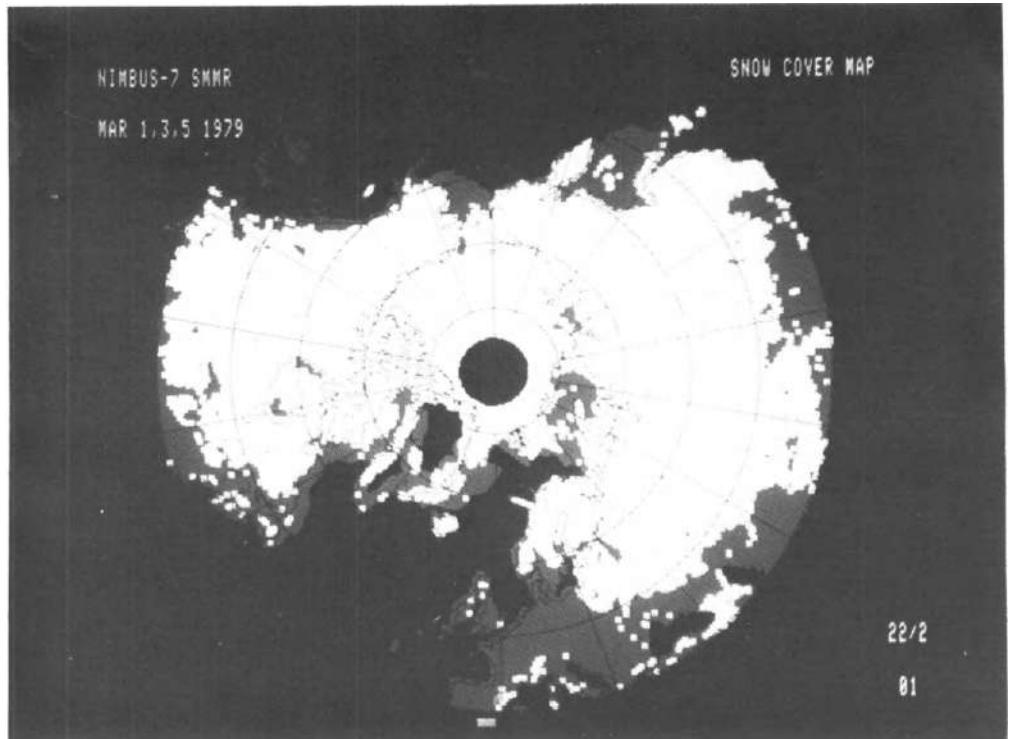


Fig. 2 : Global snow cover derived from NIMBUS-7 SMMR-data

fully monitoring sea state, sea surface temperature, brightness temperature over land and water vapor content of the air, all data are intended to be used for meteorological purposes (Incospas 1979).

#### ON THE DEFINITION OF AN OPTIMUM SNOW PAYLOAD BY GROUND BASED MICROWAVE MEASUREMENTS

Besides the utilization of the Nimbus 7 passive microwave data for snow mapping, an extensive ground-based microwave measurement program is going on for the determination of the emissivity and backscatter properties of snow during all seasons with widely varying snow conditions. A multi-frequency radiometer within the frequency limits 1.8 GHz and 94 GHz and a scatterometer at X-band are yielding data to determine the relationship between brightness temperature and the snow parameters as moisture content, water equivalent, onset of melting, development of the snow metamorphosis.

In the following paragraph a short review of several of the most pronounced results is presented, based on earlier publications (Schanda and Hofer, 1977; Hofer and Mätzler, 1980; Mätzler, Schanda, Hofer and Good, 1980; Schanda and Mätzler, 1981) in particular by Mätzler, Schanda, Good (1982). Figure 3 shows an average spectrum of brightness temperatures for horizontal and vertical polarisation of typical winter conditions i.e. without any melting metamorphism of the snow cover. The averaging is obtained by interpolating the measured values of three successive winters to a common value of the water equivalent, namely 48 cm, corresponding to a snow depth of approximately 160 cm. The brightness temperature of the horizontal polarisation is typically 20 K to 30 K lower than that of vertical polarisation and both are decreasing with increasing frequency from about 20 GHz onward.

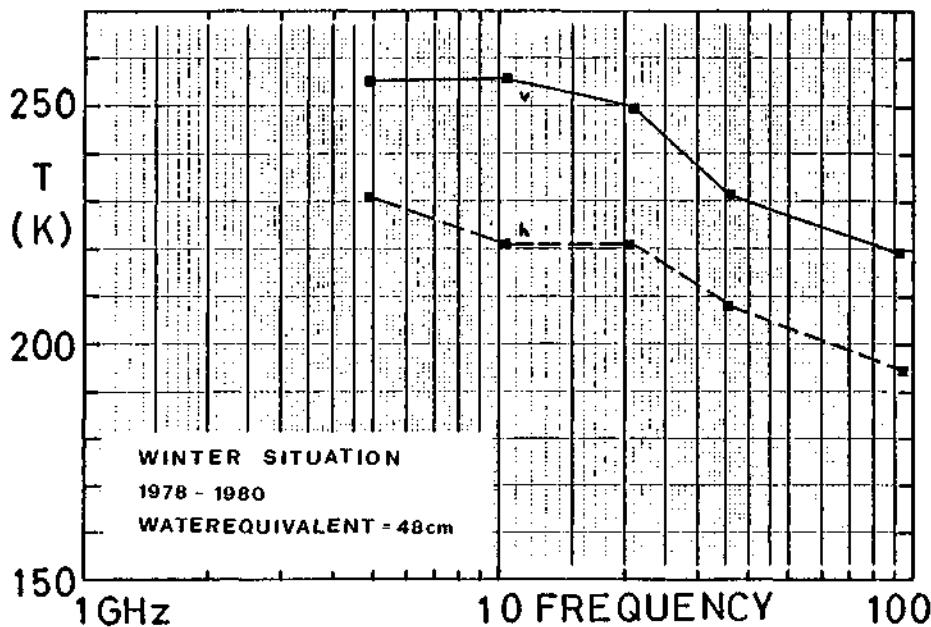


Fig. 3. Average spectra of brightness temperatures at 50° nadir angle for horizontal and vertical polarization (full line), during winter conditions at a water equivalent of 48 cm.

Figure 4 shows the brightness temperature at 36 GHz at both polarisations and the difference  $T_B(21 \text{ GHz}) - T_B(36 \text{ GHz})$  for vertical polarisation, all at  $50^\circ$  nadir angle as a function of the water equivalent. Below 20 cm the value of water equivalent has a pronounced effect on the microwave emissivity. Above this value a partial reversal of this effect causes an ambiguity in the interpretation of the emission properties in terms of water equivalent. The observed data points are from 4 different years, this demonstrates that the effect of water equivalent as given in Figure 4 is a characteristic property of the microwave emissivity.

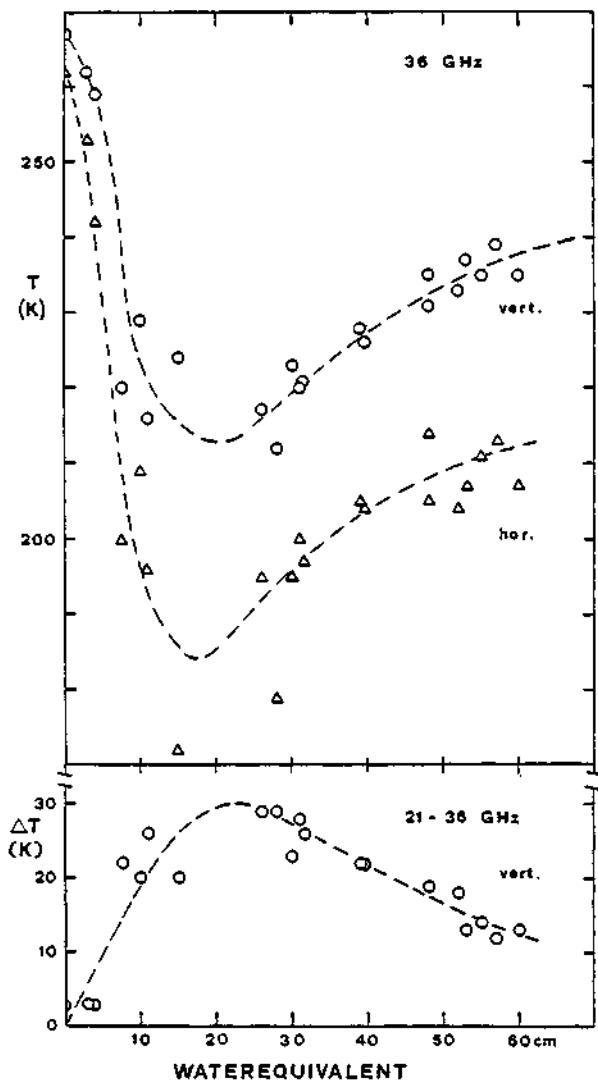


Fig. 4 The effect of water equivalents on brightness temperature  $T_B$  at 36 GHz and on the difference  $T_B(21 \text{ GHz}) - T_B(36 \text{ GHz})$ , all at  $50^\circ$  nadir angle in winter conditions. (Mätzler et al., 1982)

In Figure 5 the spectral behavior is shown of a thick firn layer of wet quasispherical ice crystals (1 to 3 mm diameter) which are formed

at temperatures above freezing point during daytime in spring, in contrast to the refrozen crust developing during clear cold nights. These spectra are characteristic for 'springtime snow' and are very different from the winter spectra.

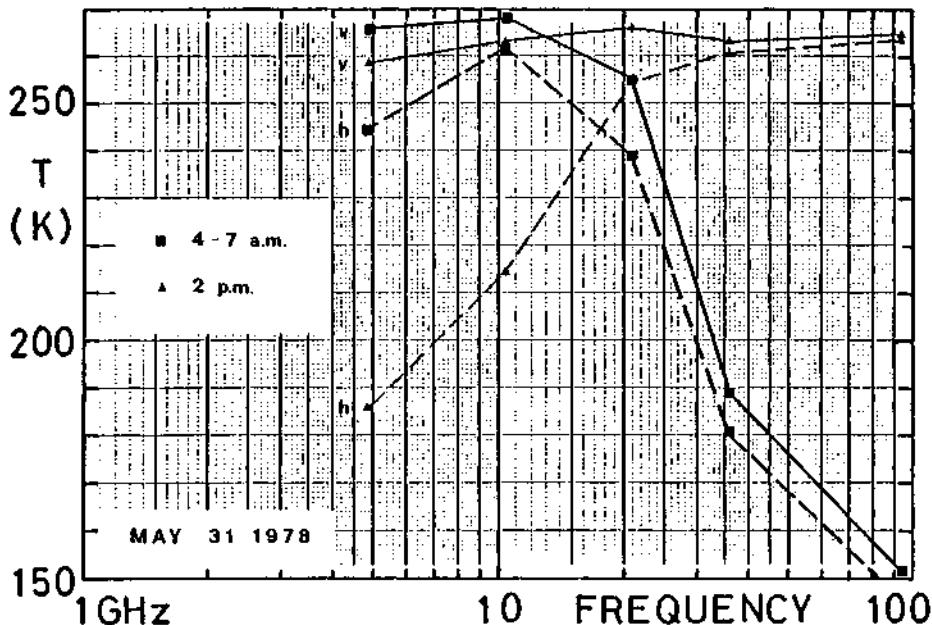


Fig. 5 Brightness temperatures at  $50^{\circ}$  nadir angle of a thick crust of refrozen firn (large symbols) and of wet firn (small symbols). (Mätzler et al., 1982)

The large contrast of mm-wave brightness temperatures between snow and snow-free regions disappears when the snow becomes wet, the uncertainty in the discrimination of wet snow from soil increases even more due to the relatively wide range of brightness temperatures caused by soil moisture and soil roughness. This discrimination problem may be solved by the measurement of the backscatter coefficient at a single frequency. Investigations at X-band have shown (Mätzler et al. 1982) that for wet snow a strong decrease of the backscattered signal at nadir angles larger than  $30^{\circ}$  occurs. In a long term recording at  $50^{\circ}$  nadir angle it has been demonstrated that wet snow exhibits a backscattered intensity of only 0.1 or even 0.02 times that of dry snow or bare ground.

The complex radiometric and ground truth data are not easily interpretable at first sight. In order to gain a systematic understanding of the experimental data the calculus of multivariate data analysis was applied. Figure 6 displays the 128 data points of complete sets of observations on more than 60 different days in the projection on the first and second factor axis after a factorial analysis.

The different symbols refer to different clusters found on the basis of an unsupervised cluster analysis. The largest cluster A together with a single element of cluster B represent winter snow. The lowest brightness temperatures found in cluster E are samples of refrozen

snow in April and in early winter. Dry snow is also found in the re-frozen firn in May with a strong negative gradient  $\Delta T_B / \Delta v$  (cluster G).

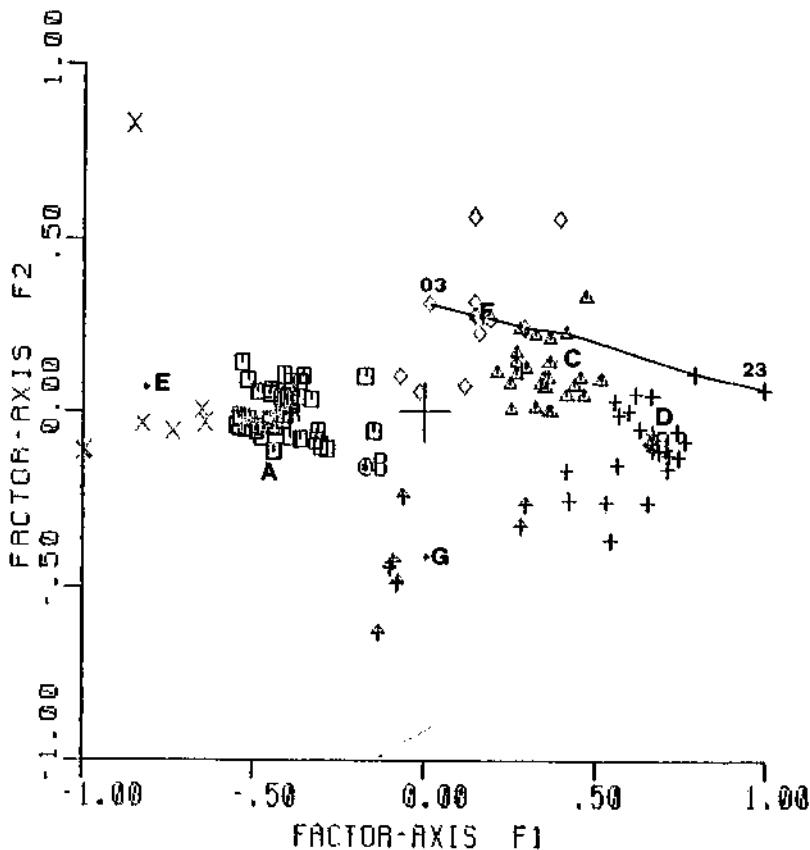


Fig. 6 Plot of the 128 radiometric measurements projected on the first and second factor axes. The points connected by a solid line in clusters C, D, F show the measurements of the snow free test field at temperatures between 3°C and 23°C.

The clusters D, F, C represent data of snow containing wetness or snow free situations, where D contains humid snow, C wet snow and F results from a mixture of very inhomogeneous situations and snow free data points.

As a conclusion of this outcome the following recommendations for microwave sensors for remote sensing of snow can be given (Mätzler et al. 1982): The sensor has to be adapted to the user requirements, and therefore different examples with increasing sophistication will be presented:

- Mapping of dry snow should be done with a radiometer of the highest possible frequency with acceptable atmospheric transparency. Depending on the atmospheric requirements the frequency lies between 25 and 45 GHz or in the 90 GHz window. Look angle and polarization are of secondary importance. An improved performance could be obtained by using a dual polarized radiometer and a second radiometer at a different fre-

quency to extract information about the type of snow.

b) Mapping of wet snow must be done with a scatterometer or with an imaging radar looking off-nadir with more than 30° incidence angle. Optimum frequencies are expected to be higher than 10 GHz. The polarization is not critical.

c) The combination of a) and b) is already a powerful sensor of the snowcover especially when applied to the melt and refreeze cycle with sufficient time resolution. The required "a-priori information" for the retrieval of the water equivalent can very likely be found with this sensor. Further information such as the energy exchange between the snow cover and the atmosphere can easily be extracted by studying the time variation of the data.

At this point it should be emphasized that the special combination of a passive microwave radiometer with a Synthetic Aperture Radar (SAR) is a very powerful instrument, because the high spatial resolution of the SAR can be combined with the high temperature resolution of the radiometer in a mixed signature algorithm.

d) Mapping of snow wetness for the observation of the melting process or of rain requires a radiometer at about 5 GHz. The highest sensitivity is obtained at large incidence angles.

With the present knowledge the optimum microwave remote sensor of snow is defined by combining c) and d) in a specific example. It is required that all instruments are looking in the same direction, approximately 45° off nadir:

- Imaging radiometers at 4.9 GHz horizontal polarization, 31.4 GHz horizontal and vertical, 45 GHz vertical polarization, and a
- Imaging SAR at 11 GHz with any like, linear polarization.

This instrument would be a valuable contribution to geophysical investigations with its capability to observe during day and night and largely unaffected by cloud cover.

#### THE PROJECT OF A PASSIVE MICROWAVE ATMOSPHERIC SOUNDER (MAS)

Passive microwave sensors can be designed not only for the detection of spectral continuum radiation as received from the surface of the earth and the oceans but by applying special frequency filtering it becomes feasible to measure the intensity and spectral width of the line radiation caused by molecular resonances of atmospheric constituents. Since the days of the earlier mentioned Nimbus 5 satellite the remote probing of atmospheric parameters as temperature and relative density of minor constituents has been performed very successfully (Waters et al., 1975; Staelin et al., 1976). The microwave atmospheric sensors to date in use on satellites are essentially vertically viewing instruments. This mode of operation suffers from a reduced sensitivity to detect minor constituents of the strato- and mesosphere due to the strong contributions by the troposphere and the height resolution is limited to at least 7 km due to the effect of transfer of radiation. A considerably higher sensitivity to detect weak line radiation from low-abundance trace gases and an improved height resolution can be achieved by the limb sounding principle. This technique has been re-

viewed and advantages and limitations for investigation of the middle atmosphere (the region comprised of stratosphere and mesosphere) have been discussed earlier (Schanda et al. 1976, Künzi, 1980). The present understanding of this part of the atmosphere is very incomplete and large consideration has been given recently to the investigation of this important portion of the atmosphere e.g. in the Middle Atmospheric Program (MAP), a major international cooperative effort which is planned for the 1982 - 85 time frame.

In this situation a definition study "phase B" has been initiated on a Microwave Atmospheric Sounder (MAS) to be realized as 4th mode of the Microwave Remote Sensing Experiment (MRSE) for its second flight on Spacelab. The study is conducted under contract of the Ministry for Research and Technology of the Federal Republic of Germany (B M.F.T.) by Dornier Systems together with experimenter teams of Max Planck Institute for Aeronomy, FRG, and University of Bern, Switzerland.

The main objective of the MAS experiment is to study the composition and dynamic structure of the stratosphere, mesosphere and lower thermosphere (height range of 20 to 100 km). Many atmospheric constituents show molecular resonances in the frequency range below 300 GHz. Examples for electric dipole moment transitions are H<sub>2</sub>O, N<sub>2</sub>O, CO, O<sub>3</sub>, ClO, the last two are of special interest in the ozone photochemistry and for the detection of possible anthropogenic effects on the ozone layer. The intensity and line shape of the resonance line of a particular molecule allows to determine the height profile of the mixing ratio of this particular constituent. The oxygen molecule (O<sub>2</sub>), in contrast to the molecules mentioned earlier, has no electric but a magnetic dipole moment. All O<sub>2</sub>-resonances below 300 GHz are fine structure transitions between different orientations of the electron spin and the molecular rotation. O<sub>2</sub> has a constant mixing ratio in the atmosphere up to ~90 km, therefore the measured thermal radiation can be used to determine the atmospheric temperature profile. A major advantage for temperature measurements at microwave frequencies is the fact that the condition of local-thermodynamic-equilibrium (LTE) is valid up to <100 km, this condition has to be fulfilled for a correct retrieving of the physical temperature from the measured radiation temperature. Furthermore some of the higher rotational transitions are insensitive to temperature variations and because the linewidth is dominated by pressure broadening up to 80 km height, radiation from these lines can be used to derive atmospheric pressure profiles.

Another objective is the determination of sources and sinks of water vapor in the middle atmosphere. The effect of chlorine on the ozone chemistry is also an urgent question. Man-made CFCl<sub>3</sub> and CF<sub>2</sub>Cl<sub>2</sub> used in refrigerators and spray-cans are most likely main sources of stratospheric Cl. Global data on ozone and ClO, the most important product in the catalytic destruction of ozone by Cl, will help to answer the question of how seriously the ozone layer is affected by chlorine.

From these scientific objectives the following systems design of the MAS has been derived (see also Künzi, 1980):

The geometry of the earth limb observation from the envisaged Spacelab orbit is shown in Figure 7. The process of radiative transfer causes the lower limit of the height resolution to be ~3km for an ideal pencil beam antenna. A 1 meter antenna on a 250 km orbit, as available with MRSE/MAS, yields a height resolution of 10 km at 60 GHz (channel 1) at the lower (20 km) tangent point and better than 4 km at frequencies above 180 GHz (Channel 2).

The parameters to be measured by MAS, the relevant height ranges and

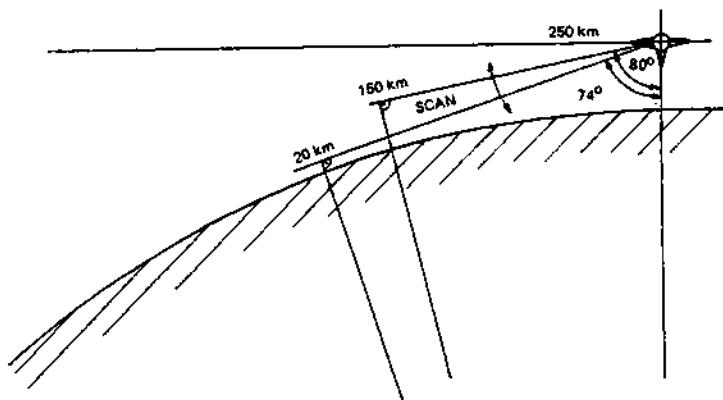


Fig. 7 Geometry of Earth-Limb Sounding from Spacelab.

expected accuracies are given in Table 2, the main characteristics of the MAS are summarized in Table 3.

TABLE 2 Main scientific objectives of the MAS experiment

Frequency (GHz)	Constituent or Parameter	Height Range (km)	Accuracy	Integration Time (s)
Channel 1:	(oxygen resonance)			
61.151	Kinetic temperature	20-100	2 K	2-10
62.998 63.568	Pressure	35- 70	1 %	2
Channel 2:	Water Vapor H <sub>2</sub> O	20- 90	2·10 <sup>-7</sup> VMR*)	2-100
183.310	Ozone O <sub>3</sub>	20- 90	2·10 <sup>-7</sup> VMR	2
184.370 204.352	Chlorine Monoxide ClO	30- 45	2·10 <sup>-10</sup> VMR	100

\*) VMR = Volume Mixing Ratio

TABLE 3 Main characteristics of the MAS

Measurement Frequencies :	62, 184, 204 GHz
Spectral Resolution :	138 Channels, 10 bits each
Data Rate:	< 45 kbit/s
Angular Resolution :	0.03°
Scan Velocity :	1,25 °/s
Angular Range :	10° to 16° (Elevation).

The MAS will operate in two submodes namely the pointing mode (PM) and the Elevation Scan Mode (ESM). During the PM the attitude is fixed with respect to the pallet, the attitude angles can be selected by telecommand in the range  $10^\circ$  to  $15^\circ$  elevation ( $80^\circ$  to  $75^\circ$  nadir angle respectively). This mode is specifically designed for long radiometer integration times (up to 100 seconds) in order to achieve sufficient sensitivity for detecting the weak radiation by  $\text{H}_2\text{O}$  and  $\text{ClO}$  with sufficient reliability. In the ESM the antenna scans with constant velocity of  $1.25^\circ/\text{s}$  through the elevation range from  $10^\circ$  to  $16^\circ$ , the azimuth angle is fixed at  $0^\circ$ . The time diagrams of both modes allow for calibration of the radiometers by a cold space view and a hot load resp. between each scan or pointing position. The microwave receivers at 62,184 and 204 GHz are total power radiometers. They downconvert the high frequency spectrum to an intermediate frequency (IF) of ~10 GHz. The IF-spectrum is analyzed in a real time spectrum analyzer based on chirp compressive receivers with 138 channels each with 10 bit resolution.

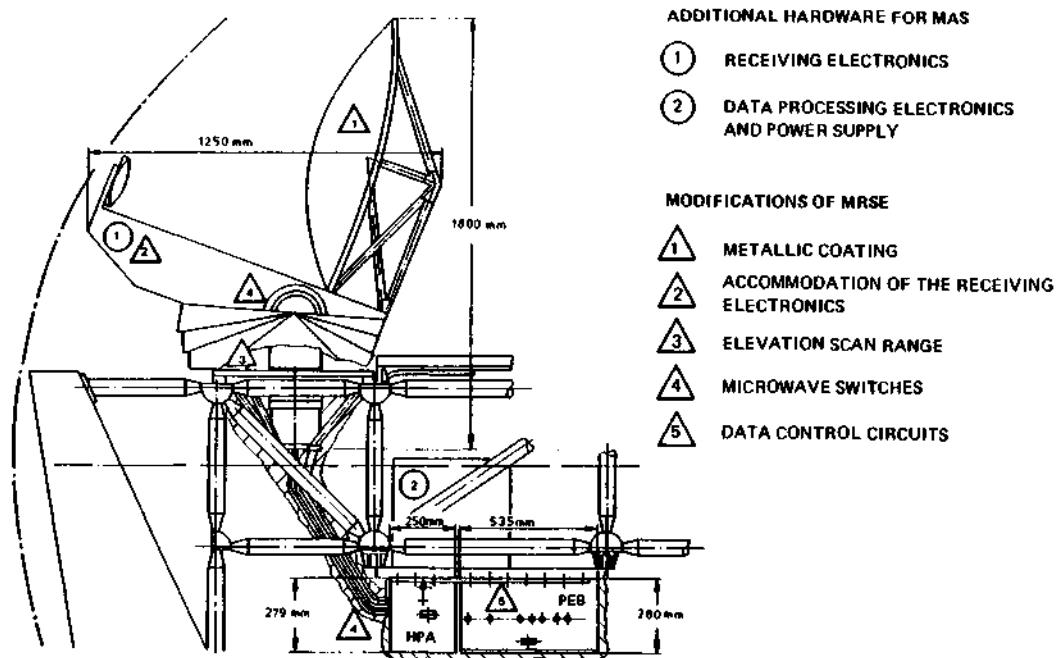


Fig. 8 Additional Hardware and Modification of MRSE for MAS

The configuration of a part of MRSE mounted on the European Bridge is shown in Figure 8. Additional hardware for MAS and modification of the MRSE for accommodation of MAS are marked.

The phase B study is presently in its final stage. After the mission for MAS is known the accommodation of the instrument MAS/MRSE will be defined in a phase B'. Thereafter the phase C/D for MAS will start with the procurement of MAS components and development of receiving and data processing electronics. The modification of MRSE and implementation of MAS as 4th mode of MRSE is planned after mid 1983. The flight instrument will be delivered in late 1985 to be ready for launch at the second flight of MRSE on Spacelab.

## REFERENCES

- Basharinov A.E.; A.S. Gurvitch, S.T. Igorov (1971). Proc. 7th Symp. Rem. Sens. Env. (Ann Arbor), 119-131
- Basharinov A.E.; A.S. Gurvitch, A.K. Gorodezky, S.T. Egorov, B.G. Kutuza, A.A. Kurskaya, D.T. Matreev, A.P. Orlow, A.M. Shutko. (1972) Proc. 8th Symp. Rem. Sens. Env. (Ann Arbor) 291-269.
- Hofer, R., Ch. Mätzler (1980), J. Geoph. Research, 85 C1, 453-460.
- Incospa (Indian National Committee for Space Research), Report to COSPAR (1979)
- Künzi, K.F., (1980), In Space Dev. for the Future of Mankind, ed. L.G. Napolitano, Pergamon Press , 127-132.
- Künzi, K.F., (1981) 5th Gen. Ass. EARSEL, (Voss)
- Künzi, K.F., S. Patil, H. Rott (1981) Internat. Geoscience and Remote Sensing Symp., (Washington)
- Mätzler, Ch., E. Schanda, R. Hofer, W. Good, (1980), NASA Conf. Publ. 2153, Microwave Remote Sensing of Snowpack Properties, Ft. Collins, 203-223.
- Mätzler, Ch., E. Schanda, W. Good, (1982) Trans. IEEE, Geoscience and Remote Sensing, 20 in press
- MAS Project and Program Description Microwave Atmospheric Sounder, (1981) ed. Dornier Syst. Friederichshafen.
- Nelsen, E.D., (1978), Data Plan, NASA Goddard Space Fl. Ct. Greenbelt
- Patil, S., K.F. Künzi, H.Rott, (1981), 11th Europ. Microwave Conf., (Amsterdam)
- Rott, H., K.F. Künzi, S. Patil, (1981) URSI Symp. on Remote Sensing, (Washington)
- Schanda, E., J. Fulde, K.F. Künzi, (1976) In: Atm. Phys. from Spacelab, ed. J.J. Burger, D. Reidel Publ., 135-146.
- Schanda, E., R. Hofer, (1977) Proc. 11th Internat. Symp. Remote Sensing of Environment, (Ann Arbor), 601-607.
- Schanda, E., Ch. Mätzler, (1981), Adv. Space Res. 1, 151-162.
- Schwalb, A., (1978), NOAA Techn. Memorandum NESS 95,
- Staelin, D.H., A.H. Barrett, J.W. Waters, F.T. Barath, F.J. Johnston, P.W. Rosenkranz, N.E. Gaut, W.B. Lenoir, (1973) Science 182, 1339-1341.
- Staelin, D.H., K.F. Künzi, R.L. Pettyjohn, R.K.L. Poon, J.W. Waters, R.W. Wilcox, (1976), J. Appl. Meteorology 15, 1204-1214.
- Waters, J.W., K.F. Künzi, R.L. Pettyjohn, R.K.L. Poon, D.H. Staelin, (1975), J. Atmosph. Sc. 32, 1953-1969.

## ACTIVE MICROWAVE SYSTEMS

W. Keydel

*Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt E.V.,  
Institut für Hochfrequenztechnik, 8031 Oberpfaffenhofen, Federal  
Republic of Germany*

### ABSTRACT

Most important radar systems for earth- and ocean observation are altimeters, scatterometers and imaging radars with real and synthetic apertures as well. These radar types will be considered principally with respect to physical and technical basics and the respective applications. The most important systems, realized or planned will be shown with respect to the respective data and user aims and an outlook will be given on the work necessary and possible for the future.

### KEYWORDS

Radar, Altimeter, Scatterometer, Synthetic Aperture Radar (SAR), Radar Cross Section, Reflectivity  $\sigma_0$ .

### 1. INTRODUCTION

Active microwave remote sensing systems and techniques are synonymous with Radar (Radio detection and ranging). Radar uses basically the propagation and the reflection of electromagnetic waves for target detection and target analysis, it measures principally the dielectric behaviour of targets in a small frequency band. The aim of this very presentation is to make a survey on the present state of the art of active microwave remote sensing for earth and ocean observation with respect to specifications and applications of existing systems and systems under development, to give insight into respective problems and to show some prospects for the future.

The advantages of radar over passive microwave systems and techniques are mostly in the possibility of distance measurements ("ranging") and therefore, of covering 1, 2 and 3 dimensions, in the better geometrical resolution achievable with special techniques and in the fact that the required signal to noise ratio and herewith the geometrical and radiometric resolution and the measurement accuracy of radar is widely dependent on the transmitter power, wave form design and other radar design parameters so that in principle no theoretical limit exists for the optimisation of radar systems. However, the limits are practical ones such as power limitations, the inability to meet tolerances or to handle large structures etc. The disadvantages of radar over passive MW-systems are in possible environmental disturbances, in the principal necessity of power and in the limited frequency range available due to political and frequency economical reasons.

The main advantages of microwave systems over optical systems are in the all-weather capability of radar and the possible penetration depth below the surface. However, the microwave systems have the disadvantages against optical systems of having a bad geometrical resolution, which is dependent on the ratio of wavelength to objective and antenna diameter respectively and of having a small bandwidth. Therefore, at clear weather an optical system has better capabilities than a microwave system. This holds principally for active and passive systems as well.

All these considerations should be seen under the aspect that earth and ocean observation with radar is still in a development stage. Radar is a tool for remote sensing which is in wide parts not yet developed completely. Especially the use of the SAR (Synthetic Aperture Radar) and the interpretation of the respective measurements and pictures is widely uncertain and it is to be assumed, that these techniques open a wide field of new applications and advantages of active microwave systems for land and sea applications.

Center of consideration are here remote sensing possibilities by satellites. Therefore, primarily spaceborne radar systems are in the focal point of consideration and the most important systems are here radar altimeter, radar scatterometer and imaging radar with real and synthetic aperture as well. These radar types will be mainly considered in this paper.

## 2. PHYSICAL FUNDAMENTALS AND TECHNICAL BASICS

### 2.1 RADAR PRINCIPLES AND BASIC EQUATIONS

Radar uses information on electromagnetic waves characterized with amplitude  $E$ , phase  $\phi$ , frequency  $f$ , the polarization which characterizes the vector character and the signal delay time  $\tau$ ; it works basically on the echo principle: an electromagnetic wave is transmitted by a focussing antenna onto targets within the antenna beam and partly reflected back by these targets to a receiving antenna which is usually but not necessarily colocated with the transmitter antenna. The received signal is of the kind of

$$\vec{E}_r = \sum_{i=1}^n \vec{E}_i e^{j(\omega_i t_i + \phi_i)} . \quad (1)$$

Here is  $n$  the total number of scatterers contributing to the signal,  $\vec{E}_i$  the vector-field scattered by the  $i$ th scatterer,  $\vec{E}_r$  the totally received field,  $\omega_i = 2\pi f$ , the circular frequency and  $t_i = t + \tau_i$ , the observation time with the respective delay time  $\tau_i$ . The radar receiver in connection with the data processing part has the task to transform this complex signal to an observable signal proportional to  $|\vec{E}_r|^2$  and to extract the information on distance, velocity, behaviour, shape, quality and other respective characteristics of the observed targets.

Each radar of the here discussed types can principally be described by the bloc diagramme in fig. 1. However, there are big differences possible with respect to coherency, interconnections between transmitter and receiver, different modulation shemes and other characteristic blocs, units and techniques which are not to be discussed within this very paper.

The measurement of echo power proportional to  $|\vec{E}|^2$  will be used in altimeters, scatterometers and imaging radars with real and synthetic aperture as well for special target detection, surface roughness and structure estimation and measurements of the dielectric properties like permitivity and conductivity; the echopower and its fluctuations estimate the tone and the texture of an radar-image that means the image quality and the brightness. The measurements of power spectra of frequency and its variations enables normally the estimation of target velocities, of

wind and wave directions, seaway spectra and current velocities within ocean observations with scatterometers and SAR and they lead to the necessary doppler coding for SAR images. The latter is identical with phase measurements for SAR.

A very important part plays the term "coherency" which will be used for techniques, instruments and for scattering mechanisms as well, it means the phase stability and the statistical behaviour of phase of electromagnetic waves over the observation time. For complete incoherent scattering, this means if there are many independent scatterers within the beam and for a pure statistical phase distribution between the scattered signals the performance of radar can be described with good validation by the radar equation [1].

For point targets and a pulse radar holds:

$$(S/N)_r = \frac{P_{ave} G^2 \lambda^2 \sigma}{(4\pi)^3 R^4 kT_o B F_n \cdot (PRF) \Delta t L} . \quad (2)$$

Herin is:  $(S/N)_r$  = received signal to noise ratio,  $P_{ave}$  = average transmitter power,  $G$  = antenna gain,  $\lambda$  = wavelength,  $\sigma$  = radar cross-section,  $R$  = distance radar-target,  $k$  = Boltzmann Constant,  $T_o$  = noise reference temperature usually 290 K,  $B$  = receiver bandwidth,  $F_n$  = receiver noise figure, PRF = pulse repetition frequency,  $\Delta t$  = pulse length,  $L$  = losses. The factor  $L$  in equation (2) represents all losses in the total system containing the radar, the propagation path, the target and the signal evaluation. With exception of part of these losses equation (2) contains only one factor which is not under control of the radar designer namely the radar cross section  $\sigma$ , it is a decisive key factor within the radar area. For ground clutter considerations, surface observations, and imaging radar as well instead of  $\sigma$  another factor  $\sigma_o$  will be introduced, the radar cross-section per unit area, defined by

$$\sigma_o = \frac{d\sigma}{dF} . \quad (3)$$

Here represents  $F$  the area of the observed surface. This  $\sigma_o$  is a further decisive factor, it is responsible for the quality of imaging radars.

## 2.2 RADAR TECHNIQUES FOR SPACE APPLICATIONS

Most important radar systems for earth- and ocean observation by satellites are radar altimeter scatterometer and imaging radar with real and synthetic aperture as well.

### 2.2.1 THE RADAR ALTIMETER

A radar altimeter is basically a high-resolution pulse radar measuring the distance between the satellite and the ocean surface in nadir direction. A sophisticated echo signal processing allows the estimation of a number of important ocean parameters like average ocean profile, significant wave height, wind speed et al. and rain detection as well. Furthermore current detection and probably an estimation of ocean floor topography is possible. Altimeters are also applicable for geoid determination, for land topography and for the estimation of geologic structures. The radar transmits a pulse and if one knows the satellite height over the ocean surface and the satellite orbit respectively, the time delay between the transmitted pulse and the half power point of the leading edge of the averaged echo pulse allows the determination of the average ocean profile. The shape of an assumed rectangular transmitter pulse will be changed due to the irregularities of the ocean surface and the echo pulse has a finite slope which is related to the significant wave height and the mean power of the received pulse is related to the sea

surface reflectivity and herewith to the wind speed. Fig. 2 shows these radar altimeter principles schematically together with results obtained by the skylab altimeter [17]. For those measurements a very high accuracy is required, state of the art is at present 5 cm up to 10 cm. This entails a very high resolution radar technique, normally the pulse compression or the use of very short pulses depending on the altitude of the respective satellite. A comparison of the different altimeters used in Skylab, Geos-C and Seasat and planned for ERS-1 is given in table 1 [2].

	Skylab	Geos-C	Seasat	ERS-1
Mean altitude/km	485	840	800	650
Antenna beamwidth	1.5°	2.6°	1.6°	1.5°
Frequency/GHz	13.9	13.9	13.5	13.5
Peak power/kW	2	2	2	0.5
Average power/W	0.05	0.24	6.5	1.5
Uncompressed pulselength	100 ns	1 μs	3.2 μs	3 μs
Compressed pulselength	-	12.5 ns	3.1 ns	3 ns
PRF/Hz	250	100	1020	1000
Footprint/km	8	3.6	1.7	1.5
Altitude precision (rms)	< 1 m	< 0.5 m	< 0.1 m	< 0.1 m (0.05 m)
Status	Mission complete	Mission	Mission finished	planned for 1986

Table 1 Comparision of different satellite-altimeters

### 2.2.2 RADAR SCATTEROMETER

A scatterometer is a special radar for measurement of backscattered echo power dependent on the incidence angle, the received echo variations allow estimations of decisive surface parameters like surface roughness, texture, structur with respect to shape, directions and scale, material, velocities and movements. Scatterometer use primarily pulse modulation and frequency modulation respectively for measurements on stationary targets and dopplershift measurements for moving targets. Scatterometric sensors can principally operate with more than one frequency and with different polarizations at the same time and there are coherent and incoherent systems. Usually scatterometers are downward looking systems measuring surface characteristics. Due to the small resolution capabilities these sensors will preferably be used for observations and measurements on homogeneous areas like seasurfaces. Therefore, are scatterometers an important tool for oceanography. Scatterometer measurements enable the estimation of important oceanographic parameters like waterwave-spectra, water current velocity, wind direction, et al. For oceanographic applications wave- and wind-measurements are decisive. The basis for scatterometric wave measurements is the bragg-effect describing the interaction between waterwavelengths and their periodically structures with radar wavelengths corresponding to frequency or difference frequencies in multi-frequency radars respectively. Those interactions are in principle interference effects. Fig. 3 to fig. 5 show results of measurements conducted from a platform in the northern sea with a two frequency L-band-scatterometer [3]. Fig. 3 shows the measurement geometry, fig. 4 gives an example for wavespectra measured with a scatterometer and compared with ground trouth where the peak position corresponds to the measured wavelength and fig. 5 shows results of water current velocity measurements representing the variations and the respective period. The agreement between radar remote sensing results and direct measurements with current meters in different depths is excellent.

Table 2 gives a list of several satellite scatterometers with its respective stage.

	Frequency	Resolution	Swath	Antenna	Pol	Height	Stage
Skylab	13.9 GHz	8 km	8-10 km	1.2 m	VV,HH	435 km	c
Seasat	14.6 GHz	Res. $\leq$ 50 km acc. Wind $\pm 2 \text{ ms}^{-1}$	2x750 km 1x140 km sub track	4 x 0.06mx3.0m	VV,HH	800 km	f
MRSE	9.65 GHz	7.5 km	21 km	1 m x 2 m	VV,HH	250 km	p 83
ERS-1 Wind	5.36 GHz	accuracy $2 \text{ m s}^{-1}$ (10 %)	400 km	2 x 0.3mx3.6m	VV, HH	650 km	p 86
ERS-1 Wave	5.3 GHz	100 m	400 km	1 m x 10 m	HH	650 km	p 86

Table 2 Data of Satellite Scatterometer; c = complete, f = finished, p = planned

### 2.2.3 IMAGING RADAR

The described altimeter and scatterometer techniques are principally one dimensional measurements using sophisticated data evaluation like Fourier spectral analysis et al. for surface properties estimation. For image-construction the second dimension has to be added and this will be done principally by scanning the antenna beam in earthbound fixed systems and by moving the radar platform of side looking systems in aircrafts or satellites respectively (fig. 6). The latter will be discussed in the following. The range resolution of such systems is usually dependent on the pulselength and the azimuth resolution is dependent on the antenna aperture which estimates the half power beamwidth of the radar beam. Here is a principal difference between radars with real and with synthetic apertures. Real aperture radars like ordinary side looking radars have a range dependent azimuth resolution which is normally not sufficient for satellite applications (see table 5). However, sidelooking systems in aircrafts, especially in low flying aircraft are useful instruments for all kind of remote sensing on land and sea surfaces and signatures. Fig. 7 and fig. 8 show images recorded with a very simple X-band sidelooking aperture radar which are representative for the possibilities of real aperture side looking radars [4].

A synthetic aperture radar (SAR) is basically a coherent scatterometer or real aperture radar with very high sophisticated data evaluation and image processing. One point of gravity is here the coherence of the system. However, data storage, evaluation and processing are the key points for SAR. The use of a synthetic aperture in connection with extremely high range resolution methods like pulskompression requires a high degree of coherency and frequency stability. This holds also if pulse compression techniques for scatterometer will be used. However, the combination of synthetic apertures with pulse compression entail requirements for frequency adjustable oscillators for example and other high sophisticated components and this makes evident: SAR requires extreme effort not only with respect to software but also to hardware.

Basic of SAR is the construction of a very long antenna along the flight path (here assumed as strongly linear) by means of data processing. Along the flight path are the measuring points for amplitude, phase and frequency of the backscattered signal (fig. 6). This is principally a normal array. However, in a conventional array all signals arrive at the same time and will be added at the receiver input. The synthetic aperture may exist of one single element receiving the signals one after the other at the respective positions and store them correctly with respect to amplitude, phase and position. In such a way the SAR antenna becomes the single element of a large (synthetic) array antenna. The stored echos can be added by a complicated

data processing procedure and in the image processor the SAR image can be produced.

Basically there are two SAR-techniques called the focussed and the unfocussed techniques. The unfocussed SAR integrates the different incoming coherent signals neglecting the inherent phase differences occurring along the linear flightpath with respect to a fixed point target. This phase mismatch limits the possible synthetic aperture length and leads to a maximum synthetic aperture length  $L_{\max} = \sqrt{\lambda R}$  and a respective azimuth resolution of

$$\delta_{az\ max} = \frac{1}{2} \sqrt{\lambda R} . \quad (4)$$

R is here the distance measured from radar [5]. Remarkable is that the resolution is independent of the real antenna aperture. In the case of a focussed synthetic aperture radar the recorded phases will be shifted before signal summation, the phases will be matched, this leads to a maximum synthetic aperture length

$$L_{\max} = \frac{\lambda}{D} \cdot R . \quad (5)$$

One obtains an azimuth resolution of

$$\delta_{az} = \frac{D}{2} . \quad (6)$$

The resolution of a focussed SAR is independent of wavelength and distance R and furthermore a better resolution can be reached with smaller real antennas and not with larger antennas, this is opposite to real aperture radars and other optical systems.

SAR-pictures are different from normal photographies in many details. Most remarkable is the lack of perspective and the distance independent resolution. The pictures have land-map-character. They look like a photography taken vertical from a large height although they represent a sidelooking perspective. This side looking perspective leads to shadow-effects and to the well known characteristic plasticity of SAR-pictures, a typical example for plasticity gives fig. 9. The shadows can principally be used for height estimation of trees, rocks and other shadow producing elements in a SAR imagery.

Keypoint of a SAR-system is its phase coherency. High quality images can be produced only if the target phase history is precisely known and this is not always the case. Phase errors play an important part. Main sources for phase errors are radar platform instabilities, target motion and acceleration, atmospheric and ionospheric influences on wave-propagation, orbit eccentricity of satellites, earth rotation effects and instabilities in the SAR electronic, for instance due to oscillator drift or pulse time jitter. Normally holds: phase jitter may not exceed a value of  $\pi/8$  rad of the RF-cycle. Phase errors entail a decrease in image quality, most important effects are defocussing in azimuth and range, a loss in the mainlobe response and an increase in image sidelobe level. Usually a SAR signal-processor locates the position of a stationary target at an image place where its doppler frequency is zero. If the target is moving with a radial velocity component  $v_{tr}$  than results a cross range displacement  $\Delta x$

$$\Delta x = - R_o \cdot \frac{v_{tr}}{v} . \quad (7)$$

Here is v the platform-velocity. That means a radial velocity towards the radar will cause a target location rearward to the flight direction and vice versa. An example for image shift due to the radial component of target motion shows fig. 10. Ships with velocity components radial to the seasat SAR velocity have an ima-

ge displacement across their wake direction. Table 3 gives some data of Seasat SAR and the planned MRSE- and ERS-1-SAR-systems as examples for satellite SAR-systems.

	Frequency	Resolution	Swath	Antenna	Pol	Height
Seasat	1.28 GHz	25 m	70 km	2.2x10.7 m	HH	800 km
MRSE	9.62 GHz	25 m	8.5 km	1m x 2m	HH,VV	250 km
ERS-1	5.3 GHz	30 m (100 m)	75 km	1m x 10m	HH	650 km

Table 3 Data of satellite SAR

## 2.2.4 COMPARISON OF DIFFERENT RADAR TYPES

A decisive factor for all radar considerations is the signal to noise ratio, S/N, introduced in the radar equation (2), which holds for point targets. For the detection of a point target the radar signal of its background, the clutter and the respective contrast between signal and clutter is most important. However, each imaging radar measures and uses the clutter as primary signal. Therefore, the clutter to noise ratio, C/N, is for imaging radars decisive. This leads to the fact that for point target estimation the contrast, signal to clutter, is very important.

The introduction of  $\sigma_0$  in (3) leads to the radar equation for area targets with real and synthetic aperture as well. Table 4 shows the different equations for area and point-targets and for real and synthetic apertures and the respective contrasts. For contrast computation the losses in all equations have been assumed as identical and equal. However, this is normally not the case and it shall be remarked that normally special system losses are dependent on the special applied techniques; but for simplicity it shall be allowed here to take all losses out of consideration in order to reach the simple comparision and contrast equations in table 4.

	Area Target	Point Target	Contrast
Real Aperture	$(C/N)_{real} = \frac{P_{ave} G^2 \lambda^3 \sigma_0 \delta_{rg}}{(4\pi)^3 R^3 kT_o F (PRF) D L}$	$(S/N)_{real} = \frac{P_{ave} G^2 \lambda^2 \sigma}{(4\pi)^3 R^4 kT_o F \cdot B(PR) \Delta t \cdot L}$	$(S/C)_{real} = \frac{\sigma \cdot D}{\sigma_0 \cdot R \cdot \lambda \cdot \delta_{rg} \cdot B \cdot \Delta t}$
Synthetic Aperture	$(C/N)_{SAR} = \frac{P_{ave} G^2 \lambda^3 \sigma_0 \delta_{rg}}{(4\pi)^3 R^3 kT_o F 2V \cdot L}$	$(S/N)_{SAR} = \frac{P_{ave} G^2 \lambda^3 \sigma}{(4\pi)^3 R^3 kT_o F \cdot 2V \cdot \delta_{az} \cdot L}$	$(S/C)_{SAR} = \frac{\sigma}{\sigma_0 \delta_{rg} \cdot \delta_{az}}$
Improvement Factor	$(C/N)_{SAR}/(C/N)_{real} = \frac{(PRF) \cdot D}{2V}$	$(S/N)_{SAR}/(S/N)_{real} = \frac{B \cdot \Delta t \cdot R \cdot (PRF)}{\delta_{az} \cdot 2V}$	$(S/C)_{SAR}/(S/C)_{real} = \frac{B \cdot \Delta t \cdot R \cdot \lambda}{\delta_{az} \cdot D}$

Table 4 Radar- and contrast equations for real aperture radars and SAR

$(C/N)$  = received clutter to noise ratio,  $(S/N)$  = received signal to noise ratio,  $P_{ave}$  = average transmitter power,  $G$  = antenna gain,  $\lambda$  = wavelength,  $R$  = distance,  $\sigma$  = radar cross section,  $\sigma_0$  = radar cross section per unit area,  $h$  = Boltzmann constant,  $T_o = 290$  K,  $B$  = receiver bandwidth,  $F$  = receiver noise figure,  $PRF$  = pulse

repetition frequency,  $\Delta t$  = pulselength,  $L$  = losses,  $V$  = SAR-velocity,  $D$  = azimuth diameter of SAR antenna,  $\delta_{rg}$  = ground range resolution,  $\delta_{az}$  = processed azimuth resolution.

Table 5 compares azimuth resolution and maximum synthetic aperture length for imaging radars with real aperture, focussed and unfocussed synthetic aperture for an azimuthal real antennalength  $D$ .

	Real Aperture	Unfocussed SAR	Focussed SAR
max. synthetic aperture length	$D$	$\sqrt{(\lambda + R)}$	$\lambda + R/D$
azimuth resolution	$\lambda + R/D$	$1/2 \sqrt{(\lambda + R)}$	$D/2$

Table 5

Fig. 5 represents the azimuth resolution of different radar systems versus height for a 10 m antenna typically used for satellite borne SAR as it is planned for ERS-1.

### 2.3 TARGET CHARACTERISTICS

Radar targets can be specified and characterized through the radar-cross-section (RCS)  $\sigma$ . The scalar RCS is defined as quotient of the power scattered per unit solid angle into receiver direction over the power per unit area (power density) incident at the target. A mathematical formulation of this definition is:

$$\sigma = 4\pi r^2 \frac{\text{Re}(|\vec{E}_2 \times \vec{H}_2|)}{\text{Re}(|\vec{E}_1 \times \vec{H}_1|)} . \quad (8)$$

Herin is  $\vec{E}_2$  and  $\vec{H}_2$  the scattered and  $\vec{E}_1$  and  $\vec{H}_1$  the incident electric and magnetic field respectively. Under farfield conditions (for large distances) holds:

$$\sigma = \lim_{r \rightarrow \infty} 4\pi r^2 \frac{|\vec{E}_2|^2}{|\vec{E}_1|^2} = \lim_{r \rightarrow \infty} 4\pi r^2 \frac{|\vec{H}_2|^2}{|\vec{H}_1|^2} . \quad (9)$$

However,  $\sigma$  depends on the frequency of the incident radiation, the shape, the material and the orientation of the target relative to the antennas and on the polarization of the transmitting and receiving antenna. Equation (8) and (9) are scalar descriptions of  $\sigma$ . Normally the polarization dependencies can be described by the so called scattering matrix, taking into account the vector character of the fields in (8) and (9) and that each scatterer acts as a polarization transformer.

In principle a complete description of a radar target can only be given if all copolar and crosspolar amplitudes and the respective phases of the radar signal are known. A radar being able to measure all these factors is called a "complete radar", it gives all information on a target possible within the relative small bandwidth of the radar carrier frequency [6].

Equation (1) shows the statistical nature of the radar response signal due to the target and scatter statistic. Normally complex radar targets like cars, trees and areas etc. have not only one but many scattering centers which contribute to the radar response signal; (exceptions are very simple targets only like spheres, corners, cylinders, plates). Fig. 12 shows as an example the distribution of such centers on a car [7]. Fig. 13 shows the radar cross section amplitude distribution on a tree.

Under the assumption of statistical distributed amplitudes  $E_i$  and phases  $\phi_i$ , (1) becomes a statistical phasor the quality of which is dependent on the respective distribution functions. For uniform distributed phases  $\phi_i$  and Rayleigh distributed amplitudes results an exponential distribution for the amplitude square of (1). That means in general if homogeneous areas of the earth are illuminated by a coherent radar the backscattered signal amplitudes of the single observations (pixels) are statistically distributed and this is the reason for the speckle typical for each radar image also for images of very homogeneous areas. There is no possibility to characterize a radar target by measurement of one pixel only. Principally this speckle can be reduced by an increase of the integration time with the use of multi looks for instance, this reduces the standard deviation proportional to the root of the look number and smoothes the speckle. This speckle can be seen in all radar images of this very paper.

The radar cross section per unit area  $\sigma_0$  introduced in (3) is a decisive environmental factor for all earth observing radars. Each radar image is in fact the result of  $\sigma_0$  estimations and their combinations. However,  $\sigma_0$  and  $\sigma$  respectively is the only factor in the radar equations which cannot be influenced by the radar designer and this points out the necessity for deep going research in this area.  $\sigma_0$  is not only dependent on the nature of the surface like behaviour, material and structure but also on observation parameters of the radar, mainly on frequency, polarization and look-angle.

The principal look-angle dependency can be described rather exactly with empirical models, the frequency dependency is principally to be estimated with measurements experimentally for each surface.

Experimental data show that for near nadir incidence  $\sigma_0$  does not depend significantly on polarization but for grazing incidence it depends on them very strongly [8]. However, the polarization behaviour of  $\sigma_0$  is widely still to be cleared, here is a wide field for future studies, especially the position of reflection characteristics on the poincaree-spere seems to be a widely unknown and therefore a very important area for future research.

The material characteristics of the earth surface with respect to radar reflection may be expressed by three constants, the relative permeability, the dielectric constant and the conductivity as it is principally the case with respect to any other medium.

The relative permeability can normally be regarded as unity so there remain the dielectricity constant and the conductivity only responsible for  $\sigma_0$  and its frequency dependency. However, the frequency behaviour of  $\sigma_0$  is also dependent on the incidence angle. Normally the dependency increases with increasing incidence angle. The frequency dependency is very strong near grazing incidence, it depends there mostly on the surface roughness and its geometrical and structural properties [8]. The same holds principally for the penetration depth of electromagnetic waves into the surface. It depends on roughness, coverage and density as well as on the surface dielectricity constant and conductivity, the latter depends on frequency as shown in fig. 14 for different water and soil conditions [9]. Very important is here the soil moisture for ground areas and the salinity of water. The results in fig. 15 [9] show the fact that a reasonable penetration depth into seawater cannot be reached with electromagnetic waves. However, in the frequency range between about 100 MHz and 1000 MHz a reasonable penetration depth into dry and wet land surfaces seems to be possible (between 1 m for wet ground and about 100 m for very dry ground) whilst for fresh water ice penetration depths of 1 m and more depending on ice temperatures can be reached with frequencies up to 20 GHz. This leads to the conclusion that radar methods in the respective frequency ranges are very good tools for ice-observation. Using the advantage of small antenna apertures SAR's in frequency ranges between 10 and several 100 MHz seem to be very useful for observation of the

first meters underneath the earth surface. This holds principally for bare soil and for vegetation covered soil as well. However, vegetation has a lower density than soil and that enables principally a better penetration through vegetation like foliage, bushes and crop from the top to the bottom. Here is a main advantage of m- and cm-waves against optical wavelengths. But here is the penetration depth normally strongly dependent on the incidence angle. A near nadir incidence allows principally a better sight to the bottom of the vegetation than grazing incidence does. In the limit, grazing incidence does not allow a penetration anyway. Fig. 16 shows the comparison of radar measurements at 10 GHz and 17 GHz over corn covered soil and bare soil. This measurement has been conducted with a pulsed radar at 20° incidence of nadir and 1 ns pulselength [10]. Given is here the logarithm of measured backscattered intensity versus distance from radar for both, corn covered soil (continuous line) and bare soil (dotted line). The first signal is from direct spillover between receiver and transmitter antenna the second signal is the radar reflection. The results show that the width of the radar echo corresponds to mean height of the corn and that main reflections result from top of the corn (first peak) from a middle region (second peak) and from the bottom (third peak). However, the shape of this radar echo depends strongly on the incidence angle. Furthermore it can be seen that the penetration into the crop of 10 GHz waves seems to be better than the penetration of 17 GHz waves. Many other results show principally that for crop covered surfaces a frequency > 8 GHz VV polarization and an incidence angle of > 40° is an optimum [11]. For crop identification multiday observations are necessary and because of the fact that vegetated surfaces show significant temporal changes an exact crop classification can be performed only by comparing  $\sigma_0$ -values measured over the same area at different times during the growing seasons. That requires all weather capability and implies the necessity of microwave use. Experimental results show that the accuracy of classification conducted with Landsat can be improved by about 10 percent up to 15 percent by adding radar [12].

A very important task for remote sensing is soil moisture estimation and therefore is the dependency of  $\sigma_0$  on the soil moisture decisive for many observations. Experimental data show that radar backscatter is maximally correlated and sensitive to near surface soil moisture at C-band frequencies with like polarization between 10° and 20° incidence [16].

However, fig. 17 shows on an X-band imagery (9.8 GHz) the same scene before and after rain and on the bottom the histogram (the measured probability density) of  $\sigma_0$ . Significant differences can be seen here. After rain, for wet ground, the distribution of  $\sigma_0$  is much smaller around the mean value than before rain.

Microwave remote sensing methods for oceanography are in a high developed state of the art, due to the fact, that the ocean surface is principally less complicated than land surfaces.

The radar backscatter from ocean surface depends strongly on the surface roughness and its structure due to wind speed and wind direction. Generally, the backscatter increases with windspeed and it has maxima parallel to wind direction and minima in cross wind direction the upwind maximum is principally higher than the downwind maximum. These facts are represented in fig. 18 showing the  $\sigma_0$  dependency on the azimuthal look angle  $\phi$ . This behaviour can be described by an empirical formula [13]:

$$\sigma_0 = a_0 V^{\gamma_0} + a_1 V^{\gamma_1} \cos\phi + a_2 V^{\gamma_2} \cos 2\phi \quad (10)$$

It seems to be possible to combine Moore's Formula (10) with another empiric model for the mean of  $\sigma_0$  [15] describing  $\sigma_0$  dependent on the grazing angle  $\delta$  and neglecting the dependency on  $\phi$ :

$$\bar{\sigma}_o(\delta) = \begin{cases} \gamma_n \frac{\sin^{v+1} \delta}{\cos^\mu \delta} & \text{for } 0 \leq \delta \leq \delta_g \\ \gamma_\mu \frac{\sin^{v+1} \delta}{\cos^\mu \delta} & \text{for } \delta_g \leq \delta \leq 90^\circ \end{cases} . \quad (11)$$

The constants  $\mu$ ,  $v$ ,  $\gamma_n$ ,  $\gamma_\mu$ ,  $\delta_g$  are arbitrary and have to be estimated empirically. This formula is applicable for sea- and landsurfaces as well and its validity is experimentally proved as well as formula (10) [16]. A combination of (10) and (11) leads to an expression for  $\sigma_o$ :

$$\sigma_o = \bar{\sigma}_o [V^{\gamma_0} + b_1 V^{\gamma_1} \cos\varphi + b_2 V^{\gamma_2} \cos 2\varphi] . \quad (12)$$

All coefficients in (12) have to be estimated experimentally.

#### 2.4 FREQUENCY CONSIDERATIONS

The choice of frequency for remote sensing radars depends on many factors. Rain, clouds and other atmospheric influences produce attenuations, false echos and phase errors. These effects increase with increasing frequency and are therefore one reason for upper frequency limits. The reflectivity  $\sigma_o$  is strongly frequency dependent. The frequency estimates the gain and herewith the dimension of the antenna, the available frequency bands are limited due to international contracts and due to the present state of the art of respective technology.

At present technological limitations in the efficiency of radio frequency power generators and in power supply are essential. There is a gap between power and frequency requirements and realization possibilities. Resolution and image quality requirements and requirements of surface structures to be observed in oceanography (small ripplewaves for instance) call for higher frequencies. In principle the frequency range applicable is limited due to atmospheric losses and disturbances and motion compensation problems. From physical aspects the upper possible frequency limit seems to be at the 20/30 GHz region. However, a look on the radar equations in table 4 shows an rapid increase of the average power necessary for higher frequency ranges. For example: An overall systems consideration including and comparing a X- and C-band SAR on a satellite leads to an average power of 87 W for a C-band and 3319 W for X-Band and the latter does not fit with the present state of the art of power supplies with solar array. This is for example one reason for the choice of C-band for the ERS-1.

#### 3. APPLICATIONS OF ACTIVE MICROWAVE REMOTE SENSING

In spite of the relatively long wavelengths, the relatively small bandwidth and the limited channel capability radar has a wide field of applications in the remote sensing area. Points of gravity are landuse and cartographic applications, agriculture and forestry, ice observation, ocean observation, mineral prospection and geologic applications, weather forecast, natural disasters etc.

Ocean observation is necessary not only for scientific but also for economic reasons. The knowledge of seastate and iceberg location is essential for planning offshore activities like search and exploitation of gas and for the optimization of ship routes as well. Especially the latter requires all weather capability of the observation system. The efficiency of fishing fleets can be improved due to observation of fish concentrations and movements and the respective currents and upwellings of the ocean surface which can be registered in principle by altimeters,

scatterometers and imaging radars as well. Costal processes can be observed by SAR and oil and pollution detection is also possible. Fig. 19 shows an example for detection of oil films on the sea surface with Seasat SAR. Ice conditions can be determined with radar, possible pools and channels of open water can be located and the limits of pac ice can be estimated, this is important for navigation and other human operations in ice covered waters.

Radar imagery can be used to prepare planimetric topographic maps of uncharted cloud or jungle covered areas. An example for cartography and its evaluation is given in fig. 7 which is an example for the application of radar in the field of geologic analysis. Shown is a radar map taken with a real aperture imaging radar at X-band, the respective geological evaluation and for control and comparision a normal geological map of the same area, the lakes near DFVLR facility. Complex geologic structures appear sometimes more distinctly at radar than at optical photography because of the illumination angle and the radar produced shadows. The all weather capability of radar enables rapid and efficient cartographic versions. As already mentioned radar can be used for agricultural and forest surveys. The penetration of electromagnetic waves enables principally the estimation of crop and timber-height. Grey tone differences show size, shape and in the future also the use of different fields.

The allweather and day and night capabilities of radar are the basis for proper natural disaster assessments, especially water disaster like flooding and fire disaster change significantly the radar signature of an area. However, this requires a permanent covering and entails the requirement for operational systems.

#### 4. CONCLUSIONS AND OUTLOOK

All facts shown in the previous chapters show: radar is a very important and usefull tool for satellite remote sensing. The main advantages are allweather and penetration capability, the disadvantages are the principal limited resolution and - in some cases - the underdeveloped state of the art with respect to technological, physical and systemtechnical aspects. However, the aim should be to develop operational remote sensing satellites with multifrequency and multipolarization high resolution radars, synthetic aperture radars, with variable incidence for land, ice and sea observations and for atmosphere remote sensing, with target identification and classification capability. In order to reach that goal the following studies on the respective areas are recommended.

Studies under physical aspects.

In center of interest are here theoretical and experimental studies on the reflectivity  $\sigma_0$ . More collection of statistical material on  $\sigma_0$  and comparative measurements on same areas but with different frequencies and polarizations are necessary. The scattering mechanisms on single targets like trees, bushes et al. have to be studied and more exact models for rough surface scattering have to be developed in order to understand the scattering mechanisms.

Especially in the near future C-band (4.5 GHz) measurements are necessary in order to prepare the scientific field for ERS-1.

Theories have to be developed and respective experiments have to be conducted for the development of an optimum polarization concept in radar remote sensing including the complete description of the scattering matrix with respect to amplitudes and phases of the matrix elements.

The penetration depth has to be studied in detail and with respect to radars at low frequency ranges.

**Studies under technical aspects.**

The most interest is here on the fields of resolution and accuracy improvement especially on the area of SAR including the wide field of SAR image evaluation interpretation and object classification. The resolution of different radars should be optimized with respect to different application requirements.

Measuring of scattering coefficients for image and classification purposes requires accurately calibrated radars. Especially the problem of area calibration has to be solved, and the problem of radiometric resolution its definition and application is included here.

Modelling is a wide field applicable for the solution of special SAR problems.

Motion and phase error compensations are a wide field for further studies in order to improve the present state of the art of SAR.

Stereo techniques should be applied and so called stereo SAR should be developed.

The lower frequency limits for acceptable SAR application have to be estimated and radars for foliage, vegetation, soil, ice and fresh water penetration have to be developed.

Multifrequency and multichannel systems should be developed and combinations of active microwave, passive microwave and optical techniques should be considered.

**Studies under technological aspects.**

Higher frequency ranges should be exploited up to 20/30 GHz especially with respect to power generator efficiency. The improvement of the state of the art of power supply in space with solar arrays is here a conditio sine qua non.

The short pulse technology has to be extended for pulses less than second. Special pulse shape and filter techniques have to be developed and adapted on remote sensing requirements.

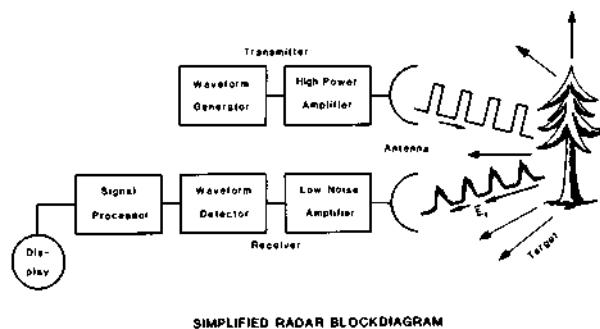
The data processing and data transmission has to be improved especially with respect to capacity and velocity.

Last but not least a very important system aspect should be mentioned. Each technique and technology does not end in itself, it has to have a strong user and application aspect and therefore, the user should have the last decision what is necessary, what systems should be built and - what it may cost. The radar scientists and technicians can evaluate and demonstrate the principal possibilities only.

**REFERENCES**

- [1] Skolnik, I (1970). Radar Handbook. McGraw Hill Company, pp. 1.4 ff.
- [2] McArthur J.L. (1976). Design of the Seasat - A Radar Altimeter. Oceans' '76, pp. 10B1-10B8.
- [3] Alpers, W., Schröter, J., Schlude, F., Müller, H.J. and Koltermann, K.P. (1980). Ocean Surface Current Measurements by an L-Band Two Frequency Microwave Scatterometer. Radio Science Vol. 16, No. 1, pp. 93-100.
- [4] Nithack, J. (1980). Visuelle Auswertung von E-SLAR-Aufnahmen. Interner Bericht DFVLR-Institut für Hochfrequenztechnik, IB 551-80/13.

- [5] Tomiyasu, K. (1978). Tutorial Review of Synthetic Aperture Radar (SAR) with Applications to Imaging of the Ocean Surface. Proc. IEEE Vol. 66, No. 5, pp. 563-583.
- [6] Börner, W. (1981). Optimal Polarization Concept in Radar Scattering from Rough Surfaces and Vegetated Areas. Proc. of ESA-Workshop Alpbach, Austria, ES-SP-AGG, pp. 129-142.
- [7] Graf, G. (1978). High Resolution Imaging of Radar Targets with Microwaves. Conf. Proc. Military Microwaves '78, London/Engl., 25.-27.10.1978. Microwave Exhibitions and Publishers Ltd., S. 295-302.
- [8] Beckmann P. and Spizzichino, A. (1963). The Scattering of Electromagnetic Waves from Rough Surfaces. Pergamon Press, pp. 401.
- [9] CCIR (1980). Conclusions of the Interim Meeting of Study Group 5 (Propagation in Nonionized Media) Geneva 16. June - 4. July 1980, Doc 5/206E, 29.8.1980. Recommendation 527, pp. 38-44.
- [10] Kleintz, M., Röde, B. and Sieber, A. (1981). High Resolution Radar Scatterometer Measurements. Paper held on: URSI Commission F Symposium on Signature Problems in Microwave Remote Sensing of the Surface of the Earth, Jan. 5-8, 1981, The University of Kansas, Lawrence, Kansas, USA.
- [11] Ulaby, F.T., Bush, T.F. and Batlivala, P.P. (1975). Radar Response to Vegetation II: 8-18 GHz-Band. IEEE Transactions on Antennas and Propagation, AP-23, No. 5, pp. 608-618.
- [12] Ulaby, F.T., Li, R.Y., and Shanmugam, K. (1981). Crop Classification Using Airborne Radar and Landsat Data. AgRISTARS Report, NASA, Nr. RSL TR 360-13; SR-KI-04043.
- [13] Moore, R.K. and Fung, A.K. (1979). Radar Determination of Winds at Sea Proceedings of the IEEE, Vol. 67, No. 11, pp. 1505-1521.
- [14] Keydel, W. (1976). An Empirical Model for Average Scattering Cross Section Computations for Land and Sea Surfaces. AGARD Conf. Proc. Nr. 208, pp. 6.1-6.15.
- [15] Keydel, W. (1981). Application and Experimental Verification of an Empirical Bac Scattering Cross Section Model for Earth's Surface. To be published in Special ISSUE of IEEE trans. on Geoscience and Remote Sensing. Paper held on: URSI Commission F. Symposium on Signature Problems in Microwave Remote Sensing of the Surface of the Earth, Jan. 5-8, 1981, The University of Kansas, Lawrence, Kansas, USA.
- [16] Bradly, G.A. and Ulaby, F.T. (1981). Agriculture Soil Moisture Experiment (ASME). AgRISTARS-Report, NASA, No. RSL TR 264-31; SM-KI-04035.
- [17] Mc Googan, J.T., Miller, L.S., Brown, G.S. and Hayne G.S. (1974). The S-193 Radar Altimeter Experiment. Proc. IEEE Vol. 62, No. 6, June 1974.



SIMPLIFIED RADAR BLOCKDIAGRAM

Fig. 1

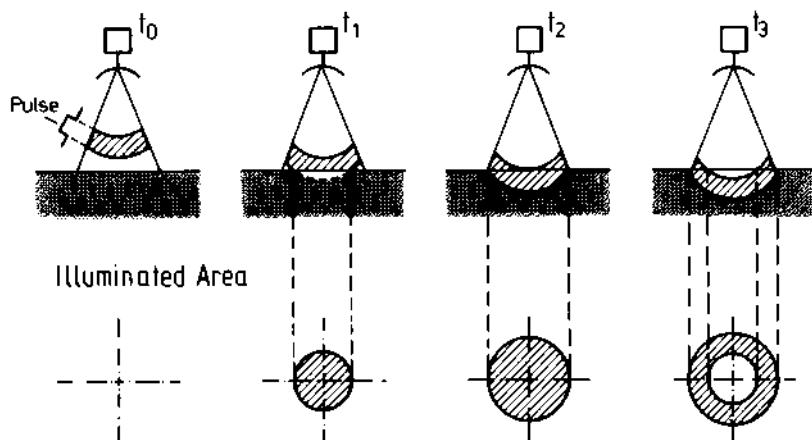
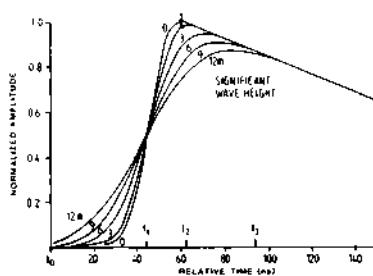
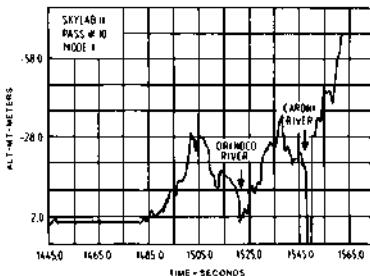
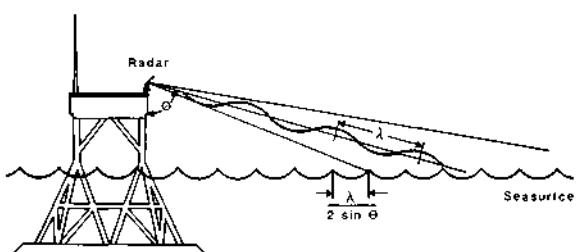
**Pulse Transmission****Typical Return for Different Waveheights****Altimeter Altitude Data over Venezuela**

Fig. 2 Radar Altimeter Principles



2FS - MEASUREMENT GEOMETRY  
(MARSEN EXPERIMENT)

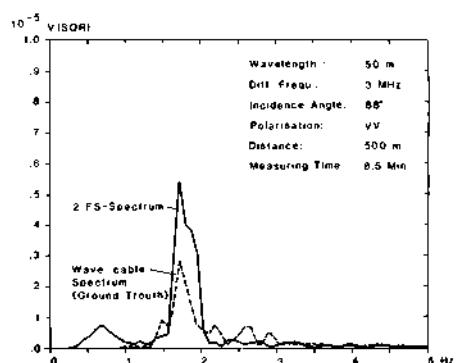
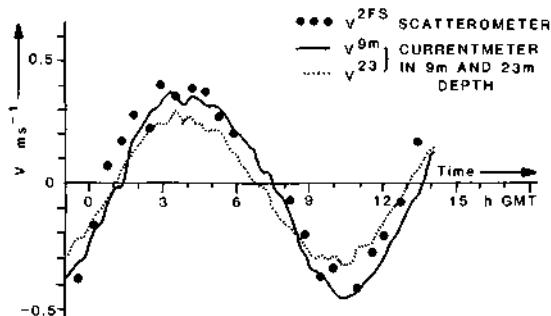


Fig. 3

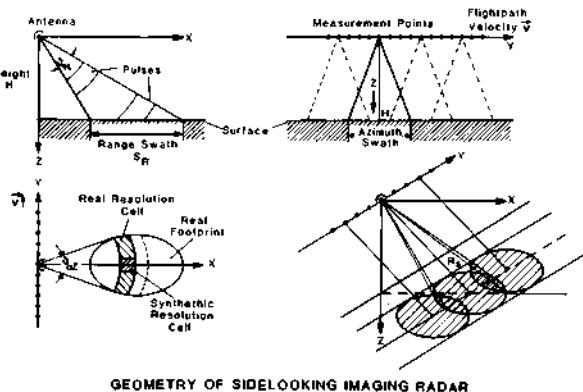
POWERSPECTRUM OF 2 FS - MEASUREMENT

Fig. 4



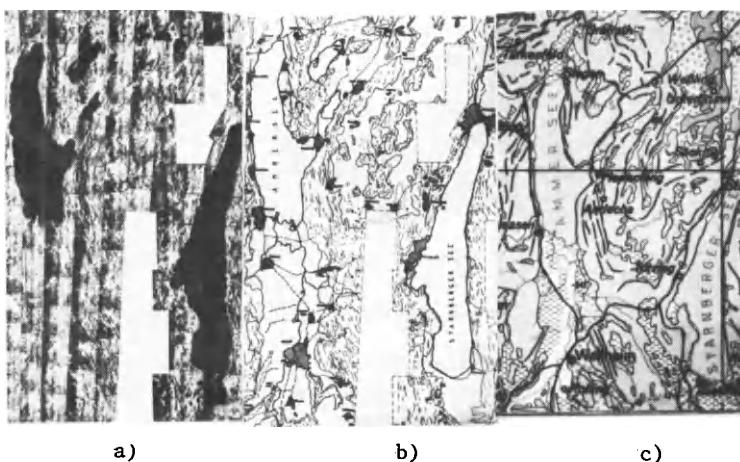
CURRENT VELOCITY VERSUS TIME

Fig. 5



GEOMETRY OF SIDELOOKING IMAGING RADAR

Fig. 6



**Fig. 7**  
**Cartography with radar.**  
**Lake district near DFVLR-**  
**Facility with different**  
**moraines of Würm ice age**  
 a) Radar mosaic  
 b) Geologic morphologic  
     interpretation  
 c) Geologic map  
 Measurement and interpre-  
 tation: Nithack, DFVLR  
 [4]



Fig. 8  
Example for real aperture  
imaging, X-band measure-  
ment E-SLAR, DFVLR



Fig. 9 Seasat SAR example  
for plasticity (Proc. DFVLR)

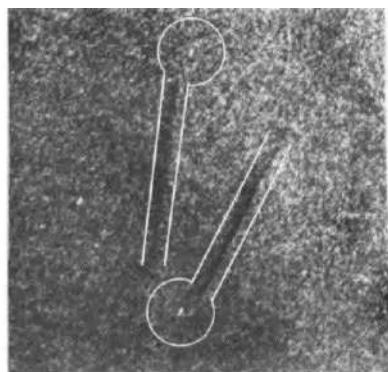


Fig. 10 Seasat SAR example for velocity displacement of ships (Proc. DFVLR)

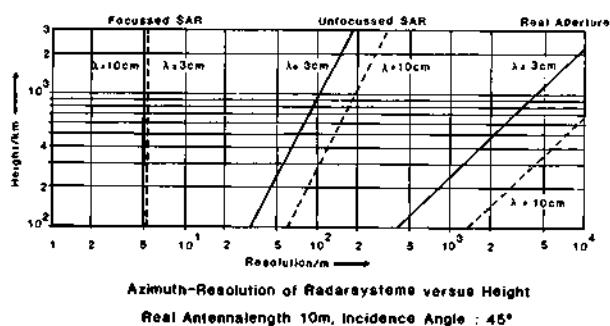


Fig. 11

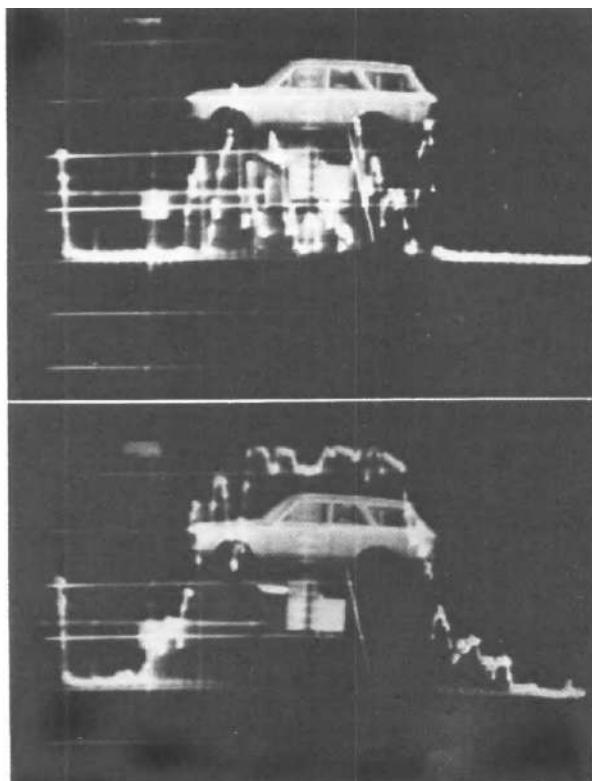


Fig. 12a Radar cross section distribution on a car for 2 different aspect angles (total view, 1 dimensional distribution) [7]  
(measurement DFVLR, Graf)

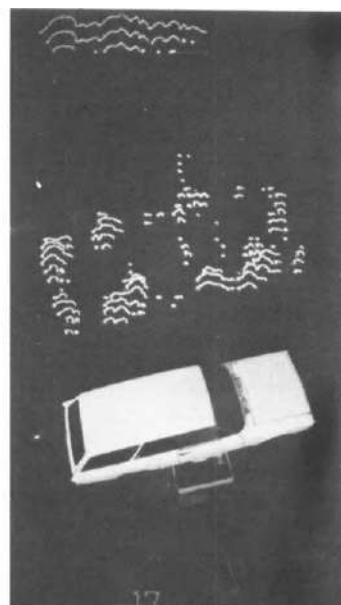


Fig. 12b Radar cross section of a car, 2 dimensional distribution of scattering centers (measurement DFVLR, Graf [7])



Fig. 13 2 dimensional scattering center distribution on a tree  
(measurement DFVLR, Graf, Röde)

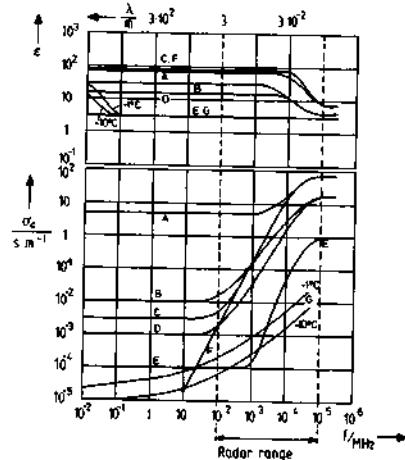


Fig. 14

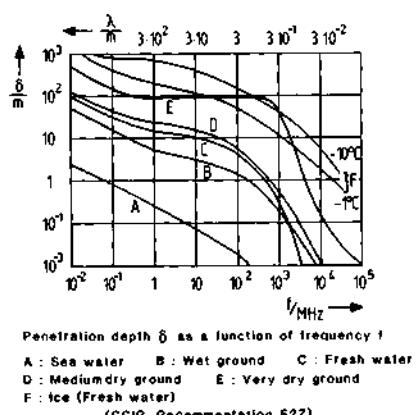


Fig. 15

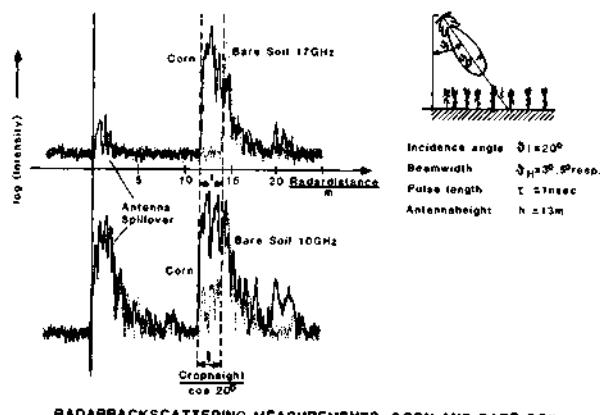


Fig. 16 (Measurement DFVLR, Röde, Sieber [10])

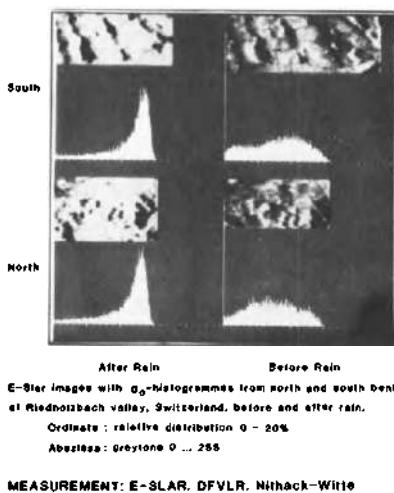


Fig. 17 (Measurements after rain left side)

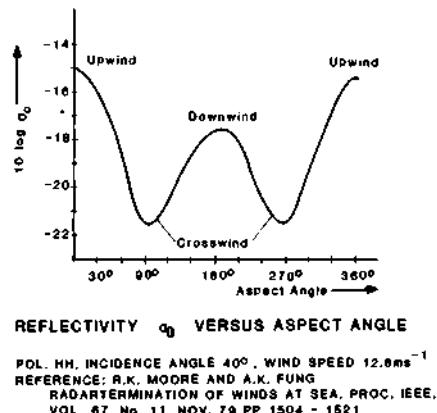


Fig. 18



Fig. 19 Seasat SAR image,  
example for oil pollution  
(Processing: DFVLR)

## ADVANCED SENSORS FOR SPACEBORNE MEASUREMENTS OF THE EARTH'S ATMOSPHERE

E. D. Hinkley

*Jet Propulsion Laboratory, California Institute of Technology,  
Pasadena, California 91109, USA*

### ABSTRACT

The need to measure meteorological and chemical properties of the Earth's atmosphere on a global basis is becoming increasingly important in view of concerns about air quality, depletion of the ozone layer, changes in the Earth's overall temperature, and weather patterns. Advanced techniques are being developed at several research centers which will enable key measurements to be made which are now either impossible to make or in need of improvement. Several laser systems are being developed for spaceborne measurements of winds, pressure, trace species, and aerosol particles. Among the passive instruments, a high-spectral-resolution infrared sensor may eventually provide global measurements of air temperature from sea level through the stratosphere, and an advanced microwave sounding unit is under consideration for the measurement of temperature and moisture profiles as well as precipitation intensity for operational weather forecasting. Optical and microwave instruments are being developed for wind measurements in the stratosphere and mesosphere, respectively; microwave techniques also show promise for measuring meteorological parameters at the air-sea interface. The evolution and current status of such advanced instrumentation for future measurements from space are described in this paper.

### KEYWORDS

Remote sensing; global measurements; atmosphere; meteorology; pollution; infrared, microwave; ultraviolet; lasers.

### INTRODUCTION

Concern about a variety of changes taking place in the Earth's atmosphere has resulted in efforts to develop new techniques and instrumentation to measure its meteorological and chemical properties on a global basis. Major emphasis is being placed on providing data which will improve our understanding of the present status and evolution of chemical species and weather drivers in the troposphere as well as possible changes in the stratospheric ozone concentration and Earth's radiation balance. These properties relate to such phenomena as air pollution, acid rain, skin cancer, species extinction, crop damage, weather, and climate.

Advanced sensors under development at several research centers will enable key measurements to be made which are now either impossible to make or in need of improvement. Specific examples of atmospheric parameters whose measurement on a global basis will be either enabled or improved substantially by new technology are: tropospheric winds, temperature, and pressure; tropospheric O<sub>3</sub>, H<sub>2</sub>O, CO, and NH<sub>3</sub> (variability with time and location); stratospheric ClO, ClONO<sub>2</sub>, OH, and sulfate particulates; and the atmospheric CO<sub>2</sub> concentration (precise measurements of change in the average global value with time). The techniques involve transmission and/or detection of electromagnetic radiation using principles of absorption, emission, fluorescence, and scattering. The wavelength range extends from the microwave to the ultraviolet, employing passive as well as active instruments.

Figure 1 is a schematic representation of the global measurement perspective from a Spacelab/Shuttle-type orbit at 250 km altitude and a free-flyer orbit at 800 km. The locations of the tropopause and stratopause are indicated to scale. The top of the stratosphere (stratopause) at 43 km is barely distinguishable from the inner circular line whose thickness is about the same as the entire troposphere. The straight line a-d passes through the troposphere and indicates the trajectory for limb occultation measurements. These are the only types of orbits considered for the advanced instrumentation described below. An Earth-synchronous orbit would be off the page; although it would permit the use of long integration times to improve the signal-to-noise ratio for a specific measurement, the Earth-synchronous orbit would not provide the global coverage needed for an understanding of the dynamics and transport processes.

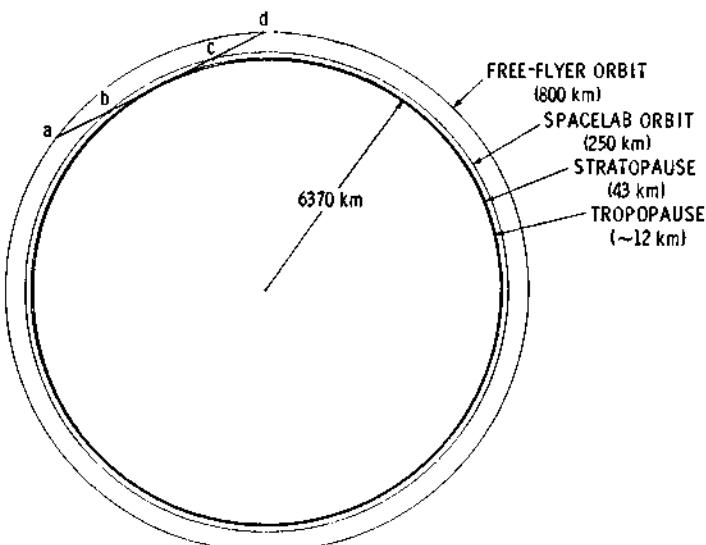


Fig. 1. Global Measurement Perspective.

Recent publications have covered in detail the feasibility of using advanced instruments for global measurements from space. These include the "NASA Shuttle Atmospheric Lidar Research Program Report" (Browell, 1979), the "NASA Upper Atmospheric Research Satellite Program Report" (Huntress, 1978), the report of the NASA Working Group on Tropospheric Pollution Planning" (Seinfeld, 1981), and the "Global Wind Workshop Summary Report" (Hinkley, 1979). From these publications it is clear, for example, that infrared, visible, and ultraviolet laser systems may be useful for measuring tropospheric meteorological parameters as well as trace

species in the troposphere and stratosphere. Laser technology is now ready to be applied to a few global measurements (Hinkley, 1976), but is in need of significant research and development for others such as wind sensing (NASA, 1981). With regard to other types of instruments, a high-spectral-resolution, passive infrared sensor shows promise for accurate measurements of temperature from sea level through the stratosphere, and an advanced microwave sounding unit is under consideration for the measurement of temperature and moisture profiles as well as precipitation intensity for operational weather forecasting. Active microwave techniques are also useful for measuring meteorological parameters at the air-sea interface. The basic technologies for future global measurements of the Earth's atmosphere from spaceborne platforms include differential absorption LIDAR (laser radar), fluorescence LIDAR, heterodyne Doppler LIDAR, long-path laser absorption, microwave scatterometry, and high-resolution microwave and infrared radiometry.

This paper includes a discussion of specific measurements of meteorological variables (winds, temperature, pressure) and chemical species (trace molecules, atoms, and aerosol particles). Where possible, comparisons are made with existing techniques and estimates are given of the expected measurement accuracy, spatial and temporal resolution, and altitude of optimum application. Most of the emphasis is placed on measurements of the troposphere as opposed to the upper atmosphere because spaceborne stratospheric instrumentation is already fairly well along in the development cycle.

#### PRINCIPLES OF REMOTE SENSING

The basic physical principles on which most remote sensing instruments for species and meteorological measurements are based are (1) absorption, (2) emission, (3) fluorescence, and (4) scattering. Several of the techniques to be described involve two of these mechanisms simultaneously. A brief review of the basic principles is given in this section.

##### Absorption

Absorption of electromagnetic radiation by atoms or molecules serves as the basis for remote measurements of several key species. Since nearly all molecules have rich spectra in the infrared, between approximately 2 and 20  $\mu\text{m}$ , and pressure-broadening does not produce as severe a change in spectral signature as it does at longer wavelengths, this region is generally recognized as being the most useful one for species measurements based upon absorption. The Earth's atmosphere itself also absorbs radiation (although the absorption is higher in the ultraviolet than in the infrared), and this, together with some other considerations to be mentioned below, make other wavelengths (e.g., microwave or ultraviolet) more attractive for certain species.

Mathematically, for radiation of wavelength  $\lambda$ , the presence of a molecular species with a density  $N(\text{cm}^{-3})$  and optical cross-section  $\sigma(\text{cm}^2)$  at wavelength  $\lambda$ , the transmittance,  $T$ , of the radiation over a pathlength  $L$  (cm) is given by the Beer-Lambert equation:

$$T = \exp(-N\sigma L). \quad (1)$$

Equation (1) holds for narrow-band instruments, such as most lasers, which have emitting bandwidths much narrower than the width of the spectral line being interrogated. For other instruments, such as some spectrometers, Eq. (1) must be modified, resulting in a smaller effective cross section, with concomitant loss in sensitivity and specificity.

Instruments based upon the absorption of radiation through the atmosphere can be either active (e.g., laser or microwave sources) or passive (e.g., spectrometers, interferometers), and the detection process may involve either direct or heterodyne techniques.

#### Emission

All atoms and molecules emit electromagnetic radiation, and this can be measured and quantified with appropriate instrumentation and circumstances. Thus, the detection of this radiation has been incorporated into several passive spaceborne instruments and is being proposed for several more. Measurements of meteorological parameters as well as species can be made by detecting radiation emitted by the atmosphere.

The submillimeter and microwave regions are especially useful for detecting upper atmospheric species where pressure broadening is low and background transmission high. Detection of infrared emission lines from certain minor species can also permit the measurement of such meteorological parameters as temperature and winds, based in the first case on a comparison of emission from bands in different wavelength regions, and in the second case by subtle shifts in emitted wavelength caused by the Doppler effect.

#### Fluorescence

If radiation of an appropriately short wavelength (high energy) impinges upon certain atoms and molecules, they can be made to fluoresce. Fluorescence measurements are specific in two ways: (1) the wavelength of the incident electromagnetic radiation must coincide with an absorbing wavelength of the species; and (2) the fluorescent emission wavelength is characteristic of the species as well. Laser radiation in the visible and ultraviolet regions of the spectrum have been most useful in enabling the remote sensing of atoms and molecules using fluorescence. Because fluorescence is quenched as background pressure increases, and to avoid absorption by oxygen and other major components, the use of this technique to detect trace atoms and molecules is generally limited to the upper atmosphere.

#### Scattering

Scattering of electromagnetic radiation by atmospheric constituents can provide information about the nature of the scatterers as well as the intervening atmosphere. Multiple-wavelength measurements of backscattering of laser radiation from atmospheric aerosols in the ultraviolet through infrared regions of the spectrum can provide information on the concentration and size distribution of the scattering centers. Research is also proceeding toward the development of techniques to remotely measure the chemical constituency of aerosol particles, and scattering from such particles serves as the basis for Differential Absorption Lidar (DIAL), reflecting the laser radiation back toward the receiver. For DIAL, the measurement technique is based on the principle of absorption, discussed above, with time-gating of the return pulses used to determine distance.

### GLOBAL METEOROLOGICAL MEASUREMENTS

In this section are descriptions of advanced instrumentation for remote measurements of tropospheric winds, temperature, and pressure from an orbiting spacecraft. In the troposphere, which extends from sea level upward to 12-15 km, at-

mospheric pressure diminishes from 1 atm to approximately 0.15 atm at the tropopause. This variation is important for certain of the measurement techniques to be described. In the stratosphere, pressure varies from 0.15 atm to 0.04 atm, yielding near-Doppler limited molecular spectra in the infrared for gaseous species near the stratopause.

### Winds

Global windspeed information is urgently needed in order to understand the transport and dynamics of the atmosphere, which bear strongly on the formation and intensity of pollution episodes as well as on weather and climate.

Infrared Laser Instrumentation to Measure Tropospheric Winds. Although cloud-tracking from orbiting satellites has been used in the past to indicate wind fields, there are two factors which limit the usefulness of cloud-derived windspeed information: (1) the altitudes of the clouds are unknown and do not cover all regions of interest; (2) the relationship between cloud drifting rate and windspeed is not well defined (Hubert and Thomasell, Jr., 1979). Even in the equatorial region, which is the main driver of atmospheric dynamics, surface measurements are not made on a routine basis. There is, therefore, a definite need for a new technique to measure tropospheric winds from an orbiting spacecraft. The required accuracy is 1-2 m/sec, with anywhere from 10 km to 400 km horizontal resolution, depending on whether the application is operational forecasting or wide-scale modeling (see, e.g. Hinkley, 1979).

Huffaker (1978, 1980) has proposed that a CO<sub>2</sub>-based infrared laser heterodyne system could provide the necessary measurements of tropospheric winds from an orbiting spacecraft -- either Shuttle at an altitude of 250-300 km, or an operational satellite at 800 km altitude. A schematic representation of the WINDSAT concept, for which the lidar system operates in a conical scanning mode, is shown in Fig. 2. An airborne CO<sub>2</sub> laser system using the same principle of detecting the Doppler shift of backscattered laser radiation has been operating at the Marshall Space Flight Center for several years (Bilbro, 1979); thus, the technique has been



Fig. 2. Schematic drawing of the NOAA WINDSAT concept (after Huffaker, 1978).

demonstrated on a small scale. For particles with an average velocity component  $v$  along the direction of the laser beam, the Doppler shift of the backscattered radiation (scattered by aerosol particles moving with the wind field) is:

$$v = 2v_o(v/c), \quad (2)$$

where  $v_o$  is the original laser frequency and  $c$  the speed of light. A spaceborne system to measure tropospheric winds would need a 10-joule  $\text{CO}_2$  laser operating at 20 Hz pulse repetition rate, with a frequency stability during each pulse of 50 kHz or better (Huffaker, 1978, 1980; Browell, 1979). The main difficulty is the large power requirement of the laser itself, of several kilowatts. Although this power may be available on a 1-2 week Shuttle flight, for flights of longer duration it will require advances in energy storage or production onboard the spacecraft, as well as improvements in laser system efficiency.

Abreu (1979) has studied an alternative technique to measure winds from a satellite platform using a lidar system operating in and near the visible portion of the spectrum, and incoherent (rather than heterodyne) detection. Because of the enhanced backscatter coefficient at shorter wavelengths, such measurements can extend to higher altitudes, well into the stratosphere, whereas spaceborne lidar systems operating in the 10- $\mu\text{m}$  region will generally be limited to the troposphere.

#### Temperature

During the past decade, numerical weather prediction models have evolved far more rapidly than the capability of satellite-borne temperature sounders to supply appropriate input data. The current generation of passive infrared sounders is capable of measuring tropospheric temperature with a vertical resolution of only 5-6 km, whereas 2 km or better is needed. This limitation in vertical resolution is caused by the broadness of the weighting functions of current instruments. (When the weighting functions are broad, the emitted energy reaching the satellite will have components originating from a wide region of the atmosphere, thereby making reconstruction of fine-scale vertical details practically impossible.) Because of this, and cloud contamination and surface emissivity effects, the rms errors in the retrieved temperature profiles are around 2.5K -- well above what is required by the circulation models. The current generation of microwave sounding units for tropospheric temperature has a vertical resolution of 8 km, which is worse than the infrared. The microwave technique is better, however, in the upper atmosphere. Design studies and numerical simulations have shown that an advanced instrument can be developed which is capable of retrieving clear-column temperature profiles with a vertical resolution of 2 km in the troposphere and an accuracy of 1.5K even in the presence of multiple layers of broken clouds. This new instrument, proposed by Chahine, Kaplan, and Susskind (1979), is called the Advanced Meteorological Temperature Sounder (AMTS).

AMTS is an infrared sounder for which the desired temperature accuracy and vertical resolution are achieved by careful choice of narrow-band channels in the 4.3- $\mu\text{m}$  and 15- $\mu\text{m}$  bands of  $\text{CO}_2$ . For temperatures in the troposphere, this can be met by a spectral resolution of  $2 \text{ cm}^{-1}$  in the high-J lines of the R-branch; and in the upper troposphere and stratosphere, by a complementary set of 15- $\mu\text{m}$  channels with a spectral resolution of  $0.5 \text{ cm}^{-1}$ . Elimination of the effects of clouds is accomplished by taking simultaneous measurements in both bands. A recent study (Chahine, 1980), has shown that AMTS, with twenty-eight appropriately-selected infrared channels, will be able to make the following measurements from a free-flyer spacecraft at an altitude of 800 km:

- Retrieve clear-column temperature profiles in the presence of up to three layers of broken clouds with an average rms error of 1.5K throughout most of the troposphere;
- Simultaneously obtain humidity profiles with an accuracy of 20%;
- Recover day and night surface temperatures of oceans and solid earth with an average absolute accuracy of 1.5K;
- Map the fractional cover and heights of multiple cloud layers globally (as seen from above) with peak-to-peak accuracies of  $\pm 0.05$  and 0.25 km, respectively;
- Determine the location of the tropopause to within 0.5 km.

The Advanced Meteorological Temperature Sounder is being proposed for a NASA free-flyer, Shuttle-launched mission in the mid-1980's. A CO<sub>2</sub> laser technique has also been proposed to measure tropospheric temperature (Murray, Powell, and van der Laan, 1980; Kalshoven and Korb, 1978; Browell, 1979), but it is not as far along as the AMTS in terms of readiness for an operational satellite.

#### Water Vapor

The Advanced Microwave Sounding Unit (AMSU) is a 20-channel microwave radiometer system under consideration by NASA and the National Oceanic and Atmospheric Administration (NOAA) to provide soundings of atmospheric water vapor, temperature profiles, and precipitation distributions for operational weather forecasting. It utilizes rotational water vapor lines at 22.2 and 183.3 GHz for sounding the moisture profiles, and several channels in the 50-60 GHz oxygen band to sound the temperature profiles. The instrument is expected to provide temperature measurements with an average rms accuracy of around 1.5K between ground and 30 km altitude, and slightly less accurate profiles to 50 km altitude. The humidity measurements will have an rms accuracy which is better than or equal to the geometric sum of 0.2 g/cm<sup>2</sup> and 10% of the measured humidity.

The AMSU is the next generation microwave sounder for NOAA operational applications and represents the fourth level of sophistication in NASA systems of this type. Predecessor systems include the Nimbus 5 Microwave Spectrometer (NEMS), Nimbus 6 Scanning Microwave Spectrometer (SCAMS) and the Microwave Sounding Unit (MSU) currently being flown on TIROS-N and the subsequent NOAA series spacecraft. The AMSU will be used aboard the Advanced TIROS-N (ATN) spacecraft beginning with the NOAA-I mission.

#### Pressure

Global measurements of atmospheric pressure are important in synoptic meteorology, numerical weather forecasting, atmospheric dynamics, and climate studies. At the present time, pressure data are gathered principally from land-based weather stations and are supplemented over the oceans by reports from ships and aircraft. The lack of data over large areas of the globe (in particular, over oceans in the southern hemisphere) presents a serious limitation.

The World Meteorological Organization has specified a set of observational requirements for the First Global Experiment of the Global Atmospheric Research Programme (GARP, 1973). Measurements of pressure in data-sparse regions are

required with a horizontal resolution of 500 km and an accuracy of  $\pm 0.3\%$ , equivalent to  $\pm 3$  mb at the surface. A recent survey of user needs (Seasat, 1976) indicates that a slightly higher accuracy of 1-2 mb may be desirable.

**Microwave Pressure Sounder.** The Microwave Pressure Sounder (see Fig. 3) is an active instrument which emits bursts of energy in the 60-GHz frequency region and detects the fraction backscattered from the sea surface. The technique can provide surface pressure measurements over the ocean by measuring beam absorption due to molecular oxygen which is uniformly mixed with other components of the atmosphere, the amount of oxygen being directly proportional to atmospheric surface pressure. Other factors, such as the atmospheric temperature profile, the presence of water vapor and liquid water, and the properties of the ocean surface, also affect the measured absorption; however, the microwave frequencies are selected to be in a region where these other absorptions vary slowly with frequency, and their effects can be removed by making additional measurements outside the oxygen band.

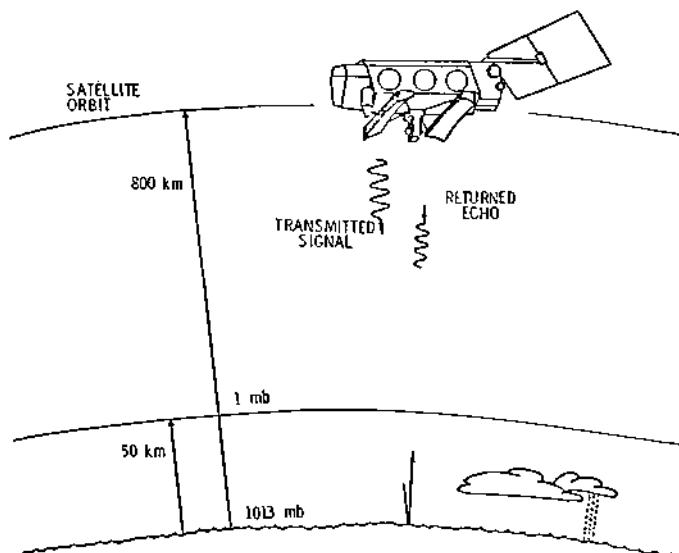


Fig. 3. Schematic drawing of the Microwave Pressure Sounder for sea-surface atmospheric pressure measurements (after Flower and Peckam, 1978).

Spectroscopic calculations (Flower and Peckam, 1978) show that the absorption coefficient for a vertical path through the atmosphere varies as surface pressure raised to a power of between 1.5 and 2 over the frequency range of interest. The nearer the frequency is to the band center, the stronger will be the change in absorption coefficient for a given change in surface pressure. A 6-channel instrument with an emitted power of 2 watts is expected to be able to measure surface pressure with a standard deviation error of between 1 mb (no clouds) and 2 mb. In addition to surface pressure, the instrument will provide estimates of the water vapor and liquid water content of the atmosphere, and of the surface roughness of the sea. The Microwave Pressure Sounder is being proposed as an instrument for a free-flying NASA satellite of the mid-1980's.

Laser Pressure Sounder. Because spectral lines broaden with increasing background pressure, this phenomenon can be used as an indicator of atmospheric pressure. Kalshoven and Korb (1978) have proposed to determine atmospheric pressure on a global basis using a spaceborne laser system in which the transmittance is measured of a laser beam whose wavelength is midway between two absorption lines of molecular oxygen. By using the LIDAR technique whereby only a small region at a known distance from the spacecraft is probed, it should be possible to derive the vertical profile of atmospheric pressure. Laboratory feasibility tests are currently underway to determine whether or not this laser technique can measure pressure with the necessary precision.

#### GLOBAL MEASUREMENTS OF ATMOSPHERIC SPECIES

##### Gaseous Species

Global measurements of several key tropospheric gases are expected to be possible using active laser techniques. Airborne measurements based on the principles of resonance fluorescence (Browell, Carter, and Shipley, 1981) and differential absorption (Menzies and Shumate, 1978; Weiseman and co-workers, 1978) have shown potential for future spaceborne applications.

The simplest approach for this purpose is one based upon differential absorption of laser radiation backscattered from the Earth's surface. Figure 4 illustrates a

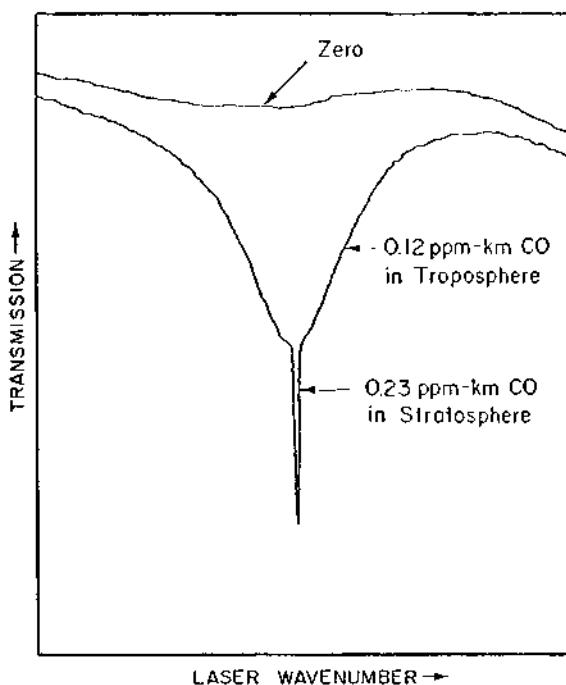


Fig. 4. Tunable diode laser scan of the P(7) line of CO near 4.6- $\mu\text{m}$  wavelength. Using two gas cells, the indicated concentrations of CO in the stratosphere and troposphere (1 atm pressure) were simulated, and the resulting scan performed by passing the laser beam through both cells in series.

tunable laser scan of a spectral line of carbon monoxide, when viewed along the vertical path for a simulated concentration of 0.12 ppm-km in the troposphere and 0.23 ppm-km in the stratosphere. The stratospheric line is narrow (approximately  $0.004 \text{ cm}^{-1}$ ) and strong. By contrast, the tropospheric portion is pressure-broadened to  $0.2 \text{ cm}^{-1}$  and shifted slightly. Although Fig. 4 is a laboratory simulation, it illustrates the type of data possible with long-path measurements of the atmosphere under sufficiently high spectral resolution.

Laser Absorption Spectrometer (LAS). The LAS measures species by differential absorption of laser radiation backscattered from the Earth's surface. Such measurements of tropospheric ozone have been made over the past few years from airplanes. Using two small  $\text{CO}_2$  lasers, the LAS measures the airplane-to-ground absorption for a laser line which is absorbed by ozone, and compares it with absorption of the second laser line which is beyond the ozone-absorbing region. By ratioing the two return signals, effects of turbulence and changing reflectivity of the Earth's surface are largely eliminated.

The present LAS system utilizes two 1-watt, continuous (cw)  $\text{CO}_2$  lasers, and has provided ozone measurements on the west coast (Grant and Shumate, 1980), east coast (Shumate and co-workers, 1981), and midwest (Shumate, 1980). Since range-gating is not possible with a cw system, only the total ozone burden below the airplane is measured. Measurements of species concentration as a function of altitude cannot be made unless the airplane flies at different altitudes. Some altitude information (with a vertical resolution of around 5 km) can be obtained, however, using the pressure dependence of the spectral line shapes (see Fig. 4 for an illustration). Extrapolations to Shuttle indicate that two 10-watt lasers will be able to make species measurements on a global basis. Species which may be measured with this technique include  $\text{O}_3$ ,  $\text{NH}_3$ ,  $\text{C}_2\text{H}_4$ ,  $\text{CCl}_4$ ,  $\text{C}_2\text{H}_3\text{Cl}$ , and  $\text{HNO}_3$ .

DIAL Lidar System. A pulsed, tunable infrared laser can provide range information lacking with a cw system, and has the potential of providing 1 km vertical resolution. For a Spacelab orbit, a laser with the required wavelength coverage, energy (25 J/pulse), pulse repetition frequency (15 Hz), stability, and overall efficiency is not yet available, but development is continuing. Recent airborne measurements of water vapor and ozone with a visible-UV DIAL lidar system have been successful, however, and potential spaceborne systems are further along than in the infrared.

Stratospheric Species Measuring Instruments. Although the present discussion is limited to the lower atmosphere, it may be useful to include a brief listing of instruments being considered for future spaceborne measurements of the stratosphere. For gaseous species, several are in the advanced stage of development -- generally further along than those for tropospheric measurements. The molecular spectra of gases in the stratosphere exhibit narrow infrared absorption lines, very near the Doppler limit of  $0.002\text{--}0.01 \text{ cm}^{-1}$ . Conventional instruments, such as filter channel infrared radiometers, and spectrometers or interferometers, have resolving powers which broaden the spectra, resulting in a potential loss in both specificity and sensitivity. An advanced interferometer called ATMOS is being constructed under a NASA contract to operate on Space Shuttle and provide a rapid scan of the infrared "fingerprint" region, using solar occultation, of stratospheric gases with a resolution of  $0.02 \text{ cm}^{-1}$ . A laser heterodyne spectrometer can operate in a similar solar occultation mode, but with the potential for greater spectral resolution (Menzies and co-workers, 1981; Allario, 1980). At the present time, the limited tuning ranges available from laser local oscillator sources greatly restricts species coverage.

Table 1 is a summary of some of the atmospheric species for which laser remote-sensing instrumentation is being developed. This is only a partial list because

of the wide variety of organizations throughout the world that are developing laser instruments for their own specific applications.

TABLE 1 Atmospheric Species Optimally Measured using Laser Instrumentation

Species	Laser Technique	Status
H <sub>2</sub> O	Absorption; near-IR DIAL	Aircraft flights ongoing
OH	Induced fluorescence	Aircraft flights ongoing
CO	IR absorption	Aircraft flights ongoing
CH <sub>3</sub> Cl	Laser Absorption Spectroscopy	Spectroscopy in progress
O <sub>3</sub>	IR Absorption, DIAL	Aircraft flights ongoing

The Microwave Limb Sounder (MLS) has recently made two successful balloon flights which have provided the first remote, simultaneous measurements of ClO and O<sub>3</sub> (Waters and co-workers, 1981). It is scheduled for incorporation onto the NASA Upper Atmospheric Research Satellite, scheduled for launch in the mid 1980's.

#### Aerosols

Global measurements of atmospheric aerosols can lead to an understanding of the formation and evolution of haze layers in the troposphere as well as enable study of stratospheric heterogeneous reactions relating to the ozone depletion question. In order to delineate more clearly lower atmospheric hazes, a special infrared channel is being installed on Landsat D'. Clusters of sulfuric acid molecules (from sulfur compounds emitted at ground level) are thought to act as condensation nuclei leading to the growth of stratospheric aerosols which are the principal components of the Junge layer. The sink for sulfur occurs when the heavier aerosols settle out of the stratosphere into the troposphere, forming a dilute "acid rain." Consequently, in addition to measuring the concentration and size distributions of aerosols in the stratosphere, it is important to determine their chemical composition. Russell and co-workers (1981) have recently published a comprehensive report on the potential measurements of aerosols from Shuttle.

A High Spectral Resolution Lidar (HSRL) has been developed at the University of Wisconsin by Shipley, Eloranta, and Tracy (1979) to measure the spatial distributions of the atmospheric aerosol optical extinction coefficient on both regional and global scales. It uses a nitrogen UV laser to optically pump a high-spectral-resolution dye laser, the output of which is directed into the atmosphere. The HSRL measures the aerosol optical extinction coefficient by distinguishing light which is backscattered by the aerosol from that backscattered by air molecules. Quantities such as the aerosol optical extinction coefficient, backscattering phase function, aerosol-to-molecular scattering ratio, and visibility can be derived from this information. A 35-cm-diameter receiver telescope collects the light backscattered by aerosol and air molecules, and the return signal is analyzed using a high-spectral-resolution, two-channel Fabry-Perot spectrometer.

Identification of the chemical composition of stratospheric aerosol particles may be possible using a multiwavelength (infrared) backscattering technique called

DISC (Wright, Pollock, and Colburn, 1975). A preliminary study of this approach indicates that sufficient lidar sensitivity can be obtained at 220-km orbital altitude with a 10-J CO<sub>2</sub> laser and a 1-m-diameter collector. The measurement principle is based on the fact that the aerosol particle backscatter coefficient shows a dependence on wavelength that is characteristic of its composition. This could also be a powerful tool for tropospheric aerosol analysis if the backscatter signatures of different tropospheric aerosols are distinctive enough.

#### CONCLUSION

Several key measurements of the Earth's atmosphere which are needed for an understanding of important processes related to weather, climate, and environmental quality cannot be made at the present time due to a lack of suitable instrumentation. As a result of laboratory research into new detection techniques, advanced remote sensors are being developed with the expectation that several of them will eventually be used to measure meteorological variables (winds, temperature, pressure) and species (trace molecules, atoms, and aerosol particles) on a global basis. Several of the techniques described in this paper are listed in Table 2 along with their potential applications. Some of these are expected to be deployed on Shuttle or on free-flying operational spacecraft during the 1980's, and should provide important inputs to our knowledge of the Earth's atmosphere.

Table 2 Summary List of Advanced Sensors and Applications

Application	Technique	Instrument/Sensor
<b>Winds</b>		
- Troposphere	Doppler/IR backscatter	CO <sub>2</sub> Lidar
- Stratosphere	Doppler/VIS backscatter	VIS Lidar
	Correlation spectroscopy	IR Correlator
- Mesosphere	Doppler/emission	Microwave Limb Sounder
<b>Temperature</b>		
	IR emission	Advanced Meteorological Temperature Sounder
	Microwave emission	Advanced Microwave Sounding Unit
<b>Pressure</b>		
- Sea-surface	Microwave absorption	Microwave Pressure Sounder
- Troposphere	IR absorption	Laser Pressure Sounder
<b>Species</b>		
- Troposphere	IR absorption	Laser Absorption Spectrometer; IR DIAL
- Stratosphere	Fluorescence	Laser Fluorosensor
	IR absorption	Laser Heterodyne, ATMOS
- Mesosphere	Microwave emission	Microwave Limb Sounder
<b>Aerosols</b>		
- <0.5 μm dia	UV/VIS scattering	<0.7 μm UV/VIS Lidar
- >0.5 μm dia	IR scattering	>0.7 μm IR Lidar

## ACKNOWLEDGMENT

The author is indebted to many colleagues for providing up-to-date information on specific measurement techniques and applications. The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

## REFERENCES

- Abreu, V. J. (1979). Winds measurements from an orbital platform using a Lidar system with incoherent detection: an analysis. Applied Optics, 18, 2992.
- Allario, F. (1980). NASA Langley Research Center. Private communication.
- Bilbro, J. (1979). NASA Marshall Space Flight Center. Private communication.
- Browell, E. V. (1979). Shuttle Atmospheric Lidar Research Program, NASA Publication SP-433.
- Browell, E. V., A. F. Carter, and S. T. Shipley (1981). Measurements of ozone and aerosol profiles with the NASA Langley airborne DIAL Lidar system, Paper V13, IEEE Conference on Lasers and Electro-optics, Washington, D.C.
- Chahine, M. T., L. D. Kaplan, and J. Susskind (1979). Advanced meteorological temperature sounder. Unpublished.
- Chahine, M. T. (1980). Jet Propulsion Laboratory. Private communication.
- Flower, D. A., and G. E. Peckam (1978). A Microwave Pressure Sounder. JPL Publication 78-68.
- GARP global experiment objective and plans (1973). GARP Publication Series, 11, World meteorological organization, Geneva.
- Grant, W. B., and M. S. Shumate (1980). Remote measurements of ozone burdens in the L. A. basin on September 20 and 24, 1979 using the JPL airborne laser absorption spectrometer. Final report to the NASA Office of Technology Utilization.
- Hinkley, E. D. (1976). Laser Monitoring of the Atmosphere, Springer-Verlag, Berlin.
- Hinkley, E. D. (1979). Global Wind Workshop Summary Report. JPL Publication 715-15.
- Hubert, L., and A. Thomasell, Jr. (1979). Error Characteristics of Satellite Derived Winds. NOAA/NESS Publication.
- Huffaker, R. M. (1978). Feasibility Study of Satellite-Borne Lidar Global Wind-Monitoring System. NOAA TM-ERL WPL-37.
- Huffaker, R. M. (1980). Feasibility of a global wind measuring satellite system (WINDSAT). Paper TuC1 at the topical meeting on coherent laser radar for atmospheric sensing. Aspen, Colorado.
- Huntress, W. T. (1978). Upper Atmosphere Research Satellite Program, JPL Publication, 78-15.
- Kalshoven, J. E., and C. L. Korb (1978). Engineering a Laser Remote Sensor for Atmospheric Pressure and Temperature. NASA TM-795381.
- Menzies, R. T., and M. S. Shumate (1978). Tropospheric ozone distributions measured with an airborne laser absorption spectrometer. J. Geophys. Res., 83, 4039.
- Menzies, R. T., C. W. Rutledge, R. A. Zantesson, and D. L. Spears (1981). Balloon-borne laser heterodyne radiometer for measurement of stratospheric trace species. Appl. Opt., 20, 536.
- Murray, E. R., D. D. Powell, and J. E. van der Laan (1980). Measurement of atmospheric temperature using a CO<sub>2</sub> laser radar. Appl. Opt., 19, 1974.
- NASA advisory committee meeting on the status of CO<sub>2</sub> laser technology for WINDSAT (1981). Washington, D. C.
- Russell, P. B., B. M. Morley, J. M. Livingston, and G. W. Grams (1981). Improved Simulation of Aerosol, Cloud, and Density Measurements by Shuttle Lidar. Final report of SRI International to NASA, Contract NAS1-16052.

- Seasat follow-on user needs report (1976). Jet Propulsion Laboratory Report, 624-1 (internal document).
- Seinfeld, J. H. (1981). NASA Report of the Working Group on Tropospheric Program Planning, NASA Reference Publication 1062.
- Shipley, S. T., E. W. Eloranta, and D. H. Tracy (1979). Measurements of the optical extinction coefficient of atmospheric aerosols by means of a high spectral resolution Lidar. Paper 2-10 presented at the 9th International Laser Radar Conference.
- Shumate, M. S. (1980). Participation of the JPL Laser Absorption Spectrometer in the 1980 PEPE/NEROS Program in Columbus, Ohio. JPL Report, 715-84 (internal document).
- Shumate, M. S., R. T. Menzies, W. B. Grant, and D. S. McDougal (1981). Laser absorption spectrometer: remote measurement of tropospheric ozone. Appl. Opt., 20, 545.
- Waters, J. W., J. C. Hardy, R. F. Jarnot, and H. M. Pickett (1981). Chlorine monoxide radical, ozone, and hydrogen peroxide: stratospheric measurements by microwave limb sounding. Science, 214, 61.
- Weiseman, W., R. Beck, W. Englisch, and K. Gürs (1978). In-flight test of a continuous laser remote sensing system. Appl. Phys., 15, 257.
- Wright, M. L., J. Pollock, and D. S. Colburn (1975). DISC: a technique for remote analysis of aerosols by differential scatter. 7th International Laser Radar Conference. Menlo Park, California.

## THE FIRST ESA REMOTE SENSING SATELLITE (ERS): THE PROGRAMME AND THE SYSTEM

D. Lennertz and C. Honvaut

*European Space Agency (ESA), Earth Observation Programme  
Department, Toulouse, France*

### ABSTRACT

The European Space Agency is preparing the development of the first ESA Remote Sensing Satellite (ERS) System. This first mission is oriented towards ice and ocean monitoring and the main mission objectives are of both scientific and economic nature.

- To increase the scientific understanding of coastal zones and ocean processes;
- to develop and promote economic applications related to a better knowledge of ocean parameters and sea state conditions.

The payload of ERS-1 is composed of

- an Active Microwave Instrumentation (AMI), combining the functions of a Synthetic Aperture Radar (SAR), a Wave Scatterometer and a Wind Scatterometer, primarily for the measurement of wind fields and wave spectra and for all-weather imaging;
- a Radar Altimeter (RA), primarily for the measurement of significant wave heights;
- an Ocean Colour Monitor (OCM), primarily for the measurement of sea surface temperature and ocean colour.

The above payload will be assembled with an existing platform which is being developed in the framework of the French Remote Sensing Programme and will be launched by ARIANE-2 (or 3) in 1987 on a sun-synchronous orbit at 675 km. An overview of the data requirements in terms of quality, coverage, is given as well as the concept of the associated Ground Segment.

### KEYWORDS

Ice and Ocean Monitoring; Scatterometers; Synthetic Aperture Radar; Ocean Colour Monitor; Radar Altimeter.

## 1. INTRODUCTION

Since the early 1970's, a number of remote sensing satellites have been launched, mainly by the USA, and a large number of experiments carried out to assess the value of this new space technology for a wide range of land and ocean applications. Since 1976 the European Space Agency, assisted by various groups of European experts, studied the mission objectives with the view to identifying European needs and eventual contributions to remote sensing satellite programmes.

These activities resulted in a recommendation for the development of two satellite systems: one oriented towards the monitoring of coastal zones and open oceans, and the other oriented towards land applications. A further recommendation was to consider not only the optical instruments for the payloads, but also microwave imaging sensors, e.g. Synthetic Aperture Radar (SAR) providing an all-weather capability. In early 1980, the ESA Member States decided that the first priority should be given to an oceanic mission which would also perform the monitoring of polar regions. The following sections will provide a brief description of the first ESA Remote Sensing satellite programme (ERS-1) in terms of mission objectives, payload configuration, ground segment and development plan.

## 2. THE MISSION OBJECTIVES

The main objectives are of both a scientific and economic nature and aim to:

- increase the scientific understanding of coastal zones and global ocean processes which, together with the monitoring of polar regions, will provide a major contribution to the World Climate Research Programme. ERS data alone or, in conjunction with complementary data from buoys, radio-sondes, research vessels etc, will enable significant advances to be made in physical oceanography, glaciology, climatology and marine biology;
- develop and promote economic/commercial applications related to a better knowledge of ocean parameters and sea-state conditions. This is of importance in view of the increasing development of coastal and offshore activities and the adoption by countries of the 200 nautical mile economic zone. In addition, monitoring of sea-ice and icebergs will be of importance for industrial activities performed at high latitudes.

### 2.1 Scientific Objectives

#### Physical Oceanography and Glaciology

The measurements of wind field, wave spectra and wave height will be used to improve the understanding of wave generation and propagation and to produce more refined models of these dynamic processes. The little-understood mesoscale eddies are detectable by altimetry and by their temperature signatures. Flow visualisation in ocean colour images may also prove to be a useful tool.

A combination of altimetry, infra-red radiometry and microwave imagery can be used for the study of ice sheet profiles and the behaviour of sea-ice in the polar oceans.

#### Climatology

A priority research topic envisaged with the World Climate Research Programme is "the controlling effect of the physics and dynamics of the oceans on the global cycles of heat, water and chemicals (especially carbon) in the climate system". For climate studies, sea-surface temperatures must be measured with a high absolute accuracy (better than 1°C) averaged over periods of a few weeks with a

spatial resolution of the order of 1000 km. Monitoring of the sea ice in the polar oceans is also important because of the strong thermal effects resulting from changing ice cover. Satellite data will be used in conjunction with a wide variety of other data in international experiments that are already being planned, such as CAGE (a study of the heat budget of a well-defined volume of ocean and the atmosphere above it) and the World Ocean Circulation Experiment (WOCE).

#### Marine Biology

Marine biology may benefit from the advanced measurements in physical oceanography discussed above. Of direct biological interest is the distribution of chlorophyll in phyto-plankton near the surface of the ocean. The spatial information from ERS will complement in situ studies of the associated up-welling of nutrients, the time-evolution of the plankton populations, and will contribute to models of primary, secondary and pelagic biological production.

#### 2.2 Applications Objectives

##### Offshore Activities and Ship Routing

Benefits to be expected from a mission such as ERS-1 result essentially from the possibility of generating improved short-term and medium-term forecasts of weather and ocean conditions on a local or global basis for the continuous and reliable monitoring of the ocean surface. The main inputs required are surface wind field measurements, supplemented by measurements of the surface wave energy spectrum and by direct measurements of wave height. All these parameters can be quantitatively measured from space with active microwave instruments.

Industrial activities which will benefit from the improved forecasts include:

(i) offshore activities by providing:

- synoptic data for short-term forecasting for planning and operations during the construction of oil platforms and the exploitation phase of the oil field;
- a continuous monitoring of ocean parameters for the establishment of statistics on wave and wind field for engineering design of oil platforms;

(ii) ship routing by:

- reducing the time on trade routes in optimising ship routes and consequently reducing total fuel consumption;
- improving the safety by reduction in hull damage, cargo damage, marine insurance costs, catastrophic ship losses, etc.

##### Management of Fish Resources

Accurate measurements of sea-surface temperature will allow the determination of temperature fronts. This information, combined with meteorological and oceanic data on ocean currents, will allow the evolution of the thermal fronts to be monitored, which are of prime importance to locate pelagic fish species living in the vicinity of the sea surface, e.g. tuna. This may lead mainly to a more efficient management of the fish resources and to an improvement in the efficiency of fishing fleets.

##### Monitoring of Sea-Ice and Icebergs

Monitoring of sea-ice and icebergs with the high resolution all-weather radar imagery will be of direct interest to the applications mentioned above performed in high latitude areas such as the Arctic region, Greenland waters and Canada.

Pollution Monitoring

Monitoring of oil and chemical pollution is also considered to be a potentially very important application. However, considerable development is still required in the areas of data interpretation and reduction techniques. Moreover, the capability of the ERS instrumentation to measure ocean surface parameters, taken in combination with meteorological information, will provide model inputs for predicting the trajectory (fate) of pollutants and their potential threat to sensitive shore areas.

Other Application Objectives

Imaging microwave and optical instruments have specific value for coastal processes such as bathymetry, surface currents, internal waves and local wind fields, both for applications and for research.

Although the main objective of the mission is to observe the oceans and ice, microwave imaging over land will be performed on an experimental basis in preparation for a later dedicated land observation mission. Some of this data can be used as an all-weather complement to optical data provided by other satellites such as LANDSAT, SPOT.

It should be noted that most of the applications mentioned above are of interest to developing countries, for which the monitoring/control of the marine resources and the development of offshore activities have a high priority.

### 3. PAYLOAD DESCRIPTION

The main consideration in the choice of a payload for ERS-1 has been to achieve as well as possible the mission objectives described in Section 2.

The first priority in the payload has been given to a comprehensive set of active microwave instruments able to observe as completely as possible the surface wind and wave structure over the oceans. The set of instruments consists of the Wind Scatterometer, the SAR employed as a Wave Scatterometer, and the Radar Altimeter. Results from SEASAT have demonstrated the ability of the Wind Scatterometer to provide wind speeds to an accuracy of 10-20% and wind directions to within about 20°, and of the Radar Altimeter to provide significant wave heights to about 10% accuracy. Further, the SAR on SEASAT was successful in imaging waves yielding the directions and wavelengths of the dominant waves and potentially the full two-dimensional wave energy spectrum.

In the selected ERS-1 configuration, SAR small-scene wave images will be interleaved with Scatterometer data of the same microwave wavelength in a global sampling scheme, thereby enabling the development of algorithms including the important coupling between the wind, the short backscattering ripples and the longer modulating waves. The simultaneous global sampling of wind and wave fields will furthermore enable the application of mutually supportive objective analysis schemes in which the wind and wave fields are estimated jointly through the application of a dynamic wave model.

For ERS-1, the Wind Scatterometer, Wave Scatterometer and SAR will be combined as an Active Microwave Instrumentation (AMI). This approach leads to a reduction of the acquired mass volume and cost by sharing common hardware.

The second priority on the payload has been given to the measurement of sea-surface temperature with a number of infrared channels aimed at giving significantly higher accuracy than is currently available, and to the measurement of

ocean colour with the main purpose of locating areas of high chlorophyll content in phyto-plankton near the ocean surface and of visualising the surface flow. The visible and infrared channels are put together in the instrument called the Ocean Colour Monitor (OCM).

To summarise, the ERS-1 payload will consist of the following instruments:(Fig.1)

- A C-band Active Microwave Instrumentation (AMI) combining the functions of a Synthetic Aperture Radar (SAR), a Wave Scatterometer and a Wind Scatterometer, primarily for wind field and wave spectrum measurements and all-weather imaging;
- an Ocean Colour Monitor (OCM), primarily for the measurement of sea-surface temperatures and ocean colours;
- a Radar Altimeter (RA), primarily for the measurement of significant wave heights.

In addition to the instruments specified above, and within the satellite capabilities of available mass (50 kg), volume, field of view clearance, power, data handling (< 100 Kbps), etc., additional instrumentation might be incorporated in the payload. The main purpose of this additional instrumentation is to complement the nominal payload package and the mission by giving access to a new range of environmental parameters or by extending the possibilities of a particular sensor.

Candidate additional instruments are as yet undefined; they are the subject of a "Call for Opportunity" which has been issued by ESA, in the European scientific community.

Examples of possible additional instrumentation are:

- climatological package (scanning radiometer, pyranometer, pyrhéliometer, etc.) for the measurement of the Earth radiation budget,
- High Accuracy Satellite Position measurement package (HASP) to be used in conjunction with the Radar Altimeter in order to improve its ability to monitor the ocean circulation (currents, etc.),
- microwave sounder for the measurement of the water content of the atmosphere, also in support to the Radar Altimeter operations (atmospheric corrections).

The selection (by ESA) of the additional instrumentation is intended to take place before the start of Phase B contract, and the selected additional instruments will be funded and procured by the proposing entities.

The data from these instruments are collected, formatted within the Instrument Data Handling Subsystem before transmission to the ground. Basically, three formats are foreseen:

- Format A: direct read-out of OCM and low data rate instruments  
(data rate < 7 Mbps)
- Format B: raw SAR data imaging mode of AMI  
(data rate about 100 Mbps)
- Format C: playback of recorded data, including OCM and low data rate instruments  
(data rate about 15 Mbps) .

At any time 2 X-band channels are available for transmission of the above formats to the ground

A short description of the main performances and characteristics of the instruments is given hereunder:

Ocean Colour Monitor (OCM)

- . Spectral channels : 800 m in all channels
- . Photometric resolution : Better than  $5 \times 10^{-4}$  ( $\text{NE}\Delta\phi$ ) in channels 1 to 7  
12 bits encoding
- . Radiometric resolution :  $\text{NE}\Delta T$  better than 0.2 K at 290 K for channels 8 to 11
- . Calibration
  - Visible : 1% relative (interchannels)  
2% absolute (target)
  - Infrared : 0.2 K at 290 K (relative)  
0.5 K at 290 K (absolute)
- . Image quality : Channel to channel registration  
VIS/VIS (less than 1/10 pixel)  
VIS/IR (less than 1 pixel)
- . Sunglint avoidance : By orbit phasing
- . Field of view :  $\pm 40^\circ$
- . Aperture : 200 mm
- . Data rate :  $\approx 2.8 \text{ Mb/s}$  average

Synthetic Aperture Radar (SAR)

- . Spatial resolution : 100 m x 100 m or 30 m x 30 m
- . Swath : 75 km
- . Radiometric resolution : 1 dB for 100 x 100 m at  $\sigma_0 = -18 \text{ dB}$   
2.5 dB for 30 x 30 m at  $\sigma_0 = -18 \text{ dB}$
- . Incidence angle (on the ground) :  $\approx 35^\circ$  side-looking
- . Frequency : 5.3 GHz in HH polarisation
- . Mean RF power : < 400 W
- . Data rate :  $\approx 100 \text{ Mb/s}$
- . Antenna size :  $\approx 10 \text{ m} \times 1 \text{ m}$

Wave Scatterometer

- . Spectral samples : From < 100 m to 1000 m in 12 log steps resolution equal to step size
- . Angular samples : Over  $180^\circ$  in azimuth in steps  $< 30^\circ$   
Resolution equal to step size
- . Spatial samples : Each 100 km along track looking at 5 km square
- . Sample accuracy :  $\pm 20^\circ$  spectral energy density of  $\sigma_0$  spectrum (target figures)
- . Incidence angle :  $\approx 35^\circ$  side-looking
- . Frequency : 5.3 GHz in HH polarisation
- . Mean RF power : < 100 W
- . Data rate :  $\approx 300 \text{ Kbps}$

Wind Scatterometer

- . Spatial resolution : 50 km squares
- . Wind speed :  $4 \text{ to } 24 \text{ m s}^{-1}$
- . Accuracy :  $2 \text{ m s}^{-1}$  or 10%
- . Wind direction accuracy :  $< 20^\circ$  ( $1\sigma$ ); acceptable level of ambiguities TBD\*  
(design goal: 0)
- . Swath : 400 km one-sided ( $25^\circ$  to  $55^\circ$  incident)
- . Polarisation : VV & HH, 2 antenna beams
- . Frequency : 5.3 GHz
- . Antenna size : 3.6 m long, 0.3 m high, 2 off
- . Data rate :  $\approx 1 \text{ Kbps}$

Radar Altimeter (RA)

- . Altitude measurement :  $< 10 \text{ cm}$  (goal 5 cm) ( $1\sigma$ , 1 sec)
- . Wave height : 1 ... 20 m ( $\pm 10\%$  or .5 m)
- . Backscatter coefficient measurement :  $\pm 1 \text{ dB}$  ( $1\sigma$ )
- . Frequency : 13.5 GHz
- . Bandwidth : 300 MHz
- . Peak power : 500 W
- . Antenna size : 1 m Ø
- . Data rate : 8.5 Kbps

## 4. LAUNCHER AND SATELLITE PLATFORM

ERS-1 will be launched by ARIANE-2 (or 3) and will use the Multi-Mission Platform (PFM) developed in the framework of the French SPOT programme. The principal characteristics of this Multi-Mission Platform are:

- . Compatibility with sun-synchronous, circular orbits of altitudes between 600 km and 1200 km
- . Compatibility with local time of satellite passes (ascending and descending node: 08.00 to 16.00 hours)
- . Available power: 1.9 KW (beginning of life) provided by a solar generator
- . Power storage capacity: 4 x 23 Ampere hours
- . Attitude control performance : Pointing towards Earth centre with yaw steering capability  
Stability on yaw:  $1.1 \cdot 10^{-3}^\circ/\text{sec}$   
Pitch and roll :  $7 \cdot 10^{-4}^\circ/\text{sec}$   
Accuracy :  $0.15^\circ$  for all axes
- . Attitude measurement accuracy : Better than  $0.15^\circ$   
Use of a Fine Attitude Measurement System (FAMS) with star sensor could improve above performance

---

\*under investigation with simulated data products

- . Orbit control : Parallel and perpendicular to the orbit plane  
Maximum capability 580,000 Nsec. (300 kg of hydrazine)
- . Satellite Management : By an on-board computer with 20 K words (16 bits) memory available for the payload management
- . Communications with ground : Telemetry and telecommand via an S-band transponder compatible with ESA and NASA networks.  
Transmitted power: up to 200 mW  
Data rate : 2 Kbits/sec

### 5. ERS-1 ORBIT

Considering, inter alia, the need to observe:

- high latitude areas,
  - the Earth's surface features by an optical imaging instrument requiring constant illumination conditions,
- a sun-synchronous, quasi-polar orbit has been selected for ERS-1. Its characteristics are:
- 3 day repeat cycle
  - altitude: 675 km
  - ground track spacing at equator: 900 km.

Minor changes of a few km or the orbit altitude will permit major increases in the ground coverage, at the expense of the repetitiveness. The flexibility in the tuning of the orbit provides the possibility, for instance, to achieve a full coverage of Canada and Europe with the all-weather imaging instrument within a few days, or a global coverage of the Atlantic Ocean with the wind and wave instruments and optical instrument.

### 6. DATA REQUIREMENTS

#### Coverage

ERS is intended to provide global worldwide coverage, compatible with the duty cycles of its payload instruments. The Agency will accordingly ensure real time acquisition through its own station network, and effect any necessary coordination with national facilities. In this way, it is expected that the following areas will be covered: the North Sea, the Baltic, the Mediterranean, the waters of Greenland and the Arctic, the East and Central Atlantic, and a part of the Indian Ocean. It is assumed that the coverage will be significantly extended with the utilisation of Canadian stations (Prince Albert and Shoe Cove). Furthermore, the on-board tape recorder(s) will give access to data from any part of the World with the proviso that the global amount of data recorded per recording period will have an upper limit of around 5 Gbits. The playback rate will be about 15 Mbit/sec.

#### Instrument Duty Cycles

The following operational modes and duty cycles have been used for the definition of the ground segment concept:

##### a) SAR, nominal mode (date rate: 100 Mbps)

On-board recording cannot take place in this mode and therefore, data acquisition can only take place from areas in the coverage of available stations. The SAR will operate at a duty cycle of 10%.

b) OCM, nominal mode (data rate: 2.8 Mbps)

Expected duty cycle of the OCM in its nominal (or direct) mode will be 25%. It can be either recorded on-board and played back, or acquired in real time.

c) Low-bit rate mode

During the low bit rate mode, the following sensors operate with a duty cycle of 40%:

- . altimeter (data rate): 8.5 Kbps
- . radar scatterometer (date rate: 1 Kbps)
- . SAR in wave mode (date rate: 600 Kbps)
- . OCM in reduced mode (IR channel) only (data rate: 0.9 Mbps).

N.B. These duty cycles reflect the constraints imposed by the available on-board energy.

The following activities can take place in parallel:

- SAR observation and downlink,
- Real-time acquisition of low-bit rate data and of OCM data (direct and reduced mode) including real-time downlinking of the data),
- Playback of low bit rate data and OCM data recorded earlier.

Acquisition and downlink of data can always be combined with the recording of that data.

Deliverable Products and their Quality

LEVEL A: RAW DATA

Raw data as received from the satellite and recorded on HDDT in satellite format

LEVEL B: QUICK-ANNOTATED RAW DATA/SYSTEM CORRECTED DATA

Raw sensor data, time ordered, time tagged, Earth located. Internal calibration tables are attached but not applied. Geometric auxiliary information will include Earth location data, making use of the latest available orbits and attitude information.

LEVEL C: QUICK-LOOK DATA

RT/TD Data, in digital or photographic form, for payload monitoring, browse facility support and, if requested, fast delivery to users.

LEVEL D: QUICK THEMATIC PRODUCTS

RT/TD Geophysical data, from low bit-rate sensors, after application of instrument transfer functions and removal of environmental effects. They will be extracted mainly from Level B data.

LEVEL E: PREPROCESSED PRODUCTS

DT Preprocessed data, with principal instrumental, radiometric and geometric errors removed, taking advantage when possible of ground truth information, possibly using refined flight dynamics data based on reconstitution.

RT = real-time from 3 to 6 hours

DT = delayed time (from one day up to several weeks)

End to End Data System

The problem of remote sensing data acquisition, processing and dissemination has been identified as a key aspect. In the particular case of oceanic missions for the monitoring of rapidly changing dynamic phenomena, requirements are especially severe and delivery of the information to end users must be performed within a few hours. A figure of 3-6 hours is currently considered adequate.

Therefore, when designing the ERS Ground Segment, it is essential to consider the complete end-to-end data system irrespective of which entities will be in charge of some parts of this system. The end-to-end data system includes all activities (both software and hardware) starting from the acquisition and recording of the raw data at the ground station until delivery of the requested information/data to the end users. This involves a number of activities or functions such as:

- recording of raw data
- data preprocessing and processing (including algorithms development)
- storage/archiving and handling of archives
- generation of products
- data/products transmission, distribution and quality control
- mission control and management.

## 7. THE ERS GROUND SEGMENT

A ground segment fulfilling all operational requirements of ERS-1 will comprise the following elements (see Figure 2).

- a) a Mission Control Centre (MCC) with associated TTC and payload data acquisition facilities. The prime TTC facility should be located at a high northern latitude to cover a high number of passes. The MCC will have responsibility for the platform and payload monitoring, their optimal operations and the implementation of the payload utilisation plans;
- b) a Mission Management Office (MMO) which coordinates the entire ground system operation. It maintains the necessary day-to-day contacts to the users to ensure optimum system usage and also coordinates with other remote sensing system operators as required.
- c) one Playback Data Acquisition Facility (PDAF) which serves as the prime facility for receiving the data stored on the on-board tape recorder. The data will be recorded after reception. There will be no data processing. It can be reasonably expected that in a standard operations scenario all low data rate instruments will principally store their data on the recorder, i.e. the PDAF would be the only station concerned with these types of instruments. The PDAF has to be located at a high northern latitude to ensure frequent contacts with the satellite as the capacity of the on-board recorder is normally limited to one or two orbits depending on the rate of the recording.
- d) one Real-Time Processing Facility (RTPF) which receives the data acquired in the Playback Data Acquisition Facility (PDAF) without delay and generates data products (mainly thematic products) within hours. This facility will provide the products derived from the low data rate instruments for the users requiring services as soon as possible after data acquisition.
- e) several Real-Time Data Acquisition Facilities (PDAFs). They will receive real-time payload data, record them by means of a high-density data tape recorder

and produce quick-look products within hours. These products must be sent directly to the users without delay. The raw data will be transported on tape to the Pre-Processing Facility defined in f) below. The location of the RDAFs is primarily determined by the desired sensor coverage as not much (OCM) or no data at all (SAR full mode) of the high data rate instruments can be recorded on-board.

- f) a Pre-Processing and Archiving Facility (PAF). It receives all raw and some processed data for archiving and further processing. It will perform instrument data processing required for payload operations, generate some special precision products and be in charge of archiving and data delivery to those non real-time users who are not served directly from the data acquisition facilities.

## 8. DEVELOPMENT PROGRAMME

The development schedule for the ERS-1 programme is expected to be as follows:

- i) Approval by ESA Member States of ERS-1 programme (Phase B) before end 1981
- ii) Carrying out of ERS-1 Phase B studies (detailed system and subsystem design) by Industry in 1982;
- iii) Approval and start of ERS-1 Phases C/D in early 1983 (development and manufacture of the spacecraft/instruments, plus the setting up of an adequate ground segment);
- iv) Launch of ERS-1 by ARIANE in early 1987. Lifetime of ERS-1 is expected to be 3 years.

It is worthwhile mentioning that in order to prepare the necessary support technology, the ESA Member States approved in early 1979 the Remote Sensing Preparatory Programme (RSPP) aimed at initiating the development of critical technologies both for optical and microwave payload elements and on-board and ground data management.

## 9. CONCLUSION

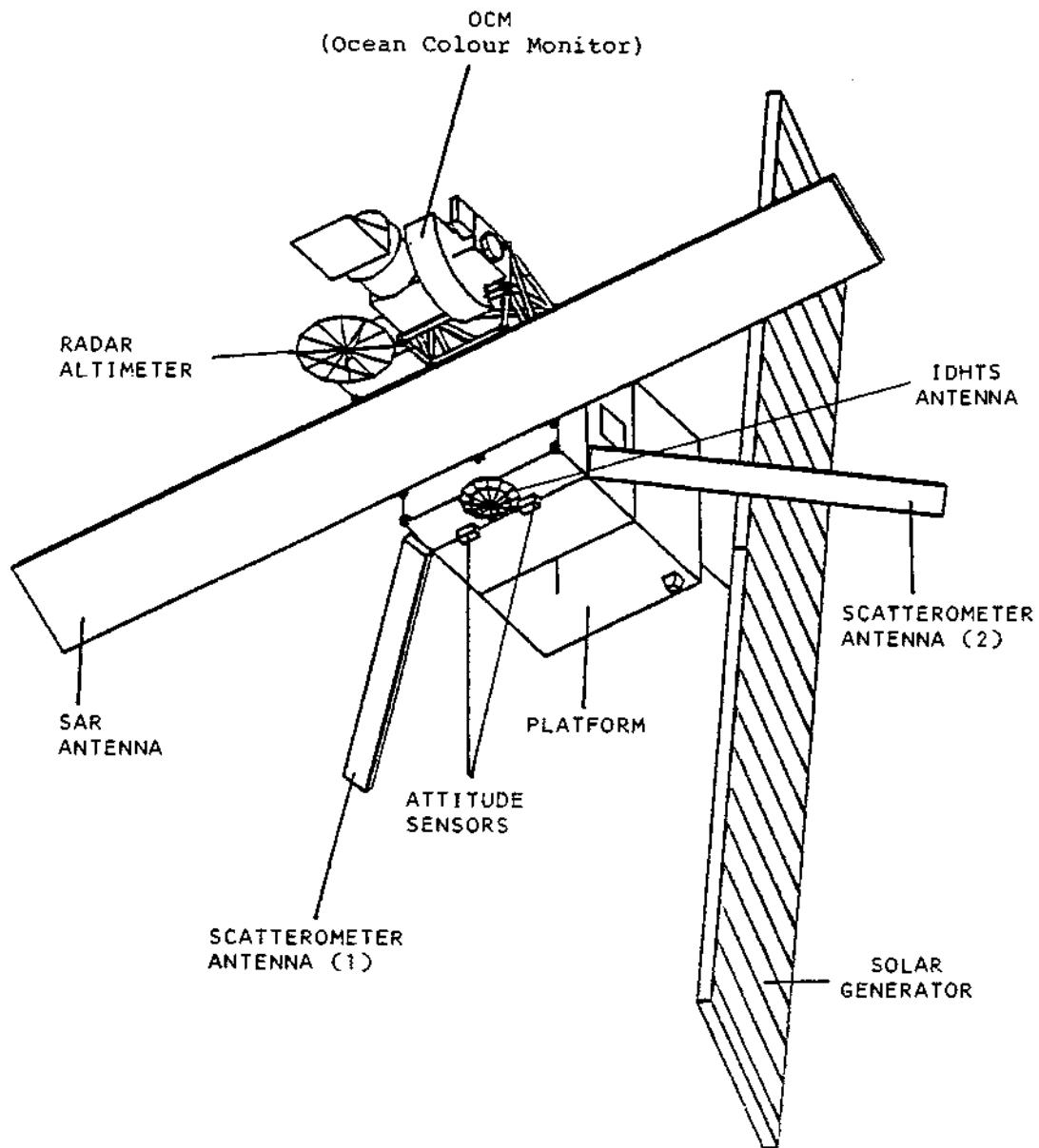
It is expected that ERS-1 will be both an experimental and pre-operational system. It will be experimental since, as the first ESA mission, it will have to demonstrate that the concept and technology for both the space and ground segments are right for the applications envisaged, and that the users are ready and able to use the data generated. This will require a number of activities before and after launch which are very important for the success of the mission, such as:

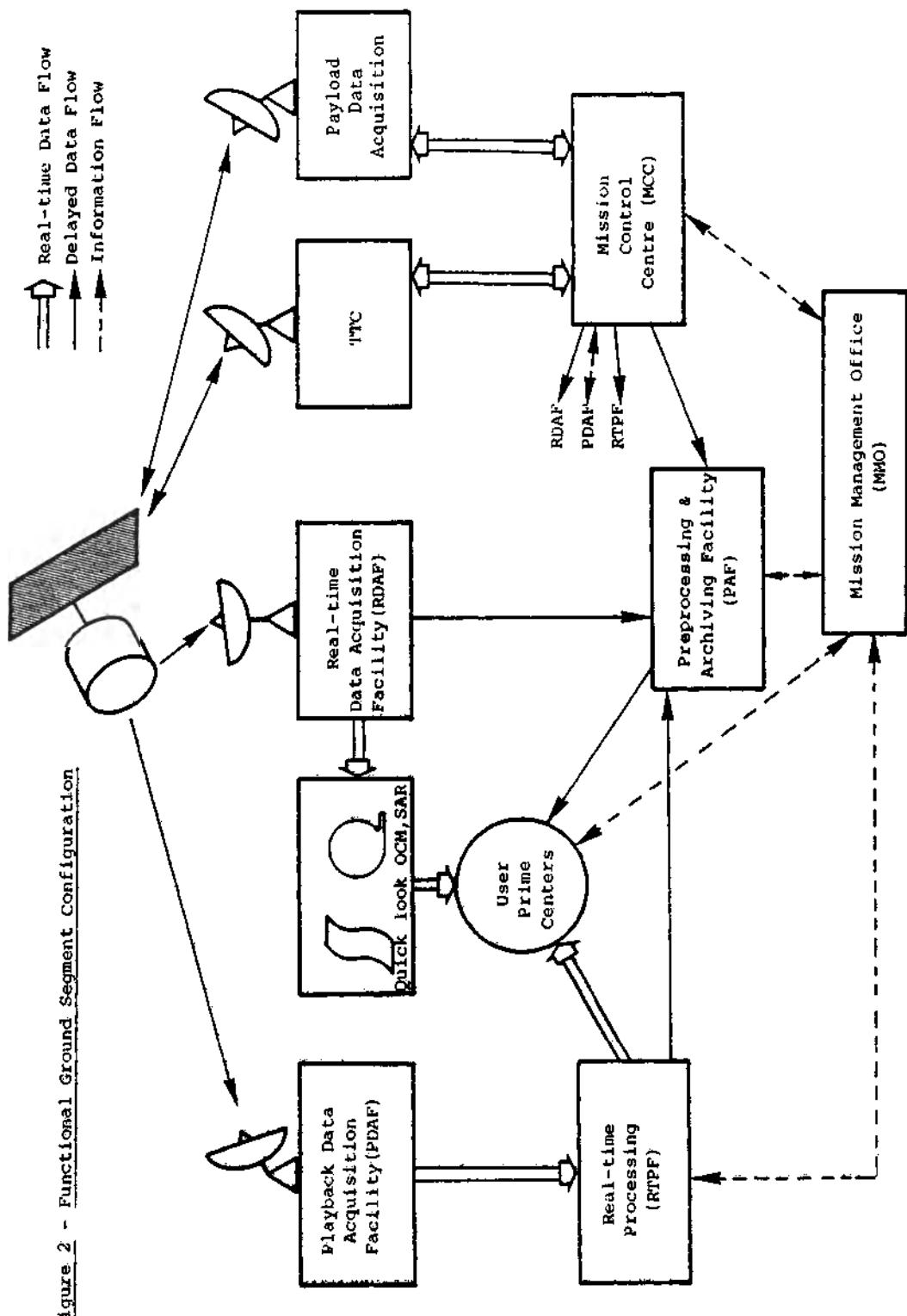
- Simulation and optimisation of sensor performances (airborne testing),
- Development and testing of algorithms and models,
- Setting up and testing of data and products,
- Definition of distribution networks to meet user requirements,
- Development of pilot-projects (small scale) and demonstration missions (large scale),
- Continuation of research and development to optimise the use and value of the data provided by the satellite system.

On the other hand, ERS-1 will have to demonstrate, for some appropriate applications and on a limited scale, an operational capability. This will, of course, require that data or products be delivered in quasi real time to corresponding existing operational services or end users.

In conclusion, the experimental/pre-operational phase now proposed for ERS-1 should be used to evaluate the means of reaching a fully operational system. In this context, Europe will be prepared to contribute to a multi-satellite system for global rather than regional monitoring.

Figure 1 - ERS-1 Baseline Concept





## INVESTIGATION, EVALUATION AND FORECAST OF NEAR-EARTH SPACE RADIATION SITUATION

S. I. Avdjushin, N. K. Pereyaslova and P. M. Svidsky

*Institute of Applied Geophysics, Goskomgidromet, Moscow, USSR*

With the present space technology development and active near-earth space investigation, to provide radiation safety of manned spaceflights has recently become a topical and important task.

Solar-induced corpuscular fluxes and electromagnetic emission represent the main radiation hazard in space. Solar flares generating charged particle fluxes are a particular hazard. The most penetrating and interacting solar radiation component is the solar cosmic ray proton component (Kaffner, 1971; Logachev, Savenko, Sladkova, 1974; Petrov, Kolomensky, Zil, 1979; Silberger, Tsao, 1979).

To develop methods of the evaluation and forecast of space solar cosmic fluxes it is essential to investigate solar cosmic ray characteristics experimentally and to correlate them with a number of heliogeophysical phenomena in space, magnetosphere, atmosphere and at the Earth's surface. Such investigations are very important for solving the problem of establishing cause-conditioned relations between the solar activity and certain meteorological processes on the Earth (Wilcox, 1978; Loginov, Sazonov, 1978; Besprozvannaya, Sazonov, Shchukina, 1972).

Since 1969 Meteor satellite observations of proton fluxes with energies  $E_p > 5, 15, 25, 40, 65$  and 90 MeV have been carried out in high-latitude zones of the Earth's magnetosphere at about 1000 km altitudes. Temporal, spatial and spectral characteristics of solar

and galactic cosmic ray proton fluxes registered at geomagnetic latitudes  $\Phi \gtrsim 67^\circ$  have been studied during the 20th and 21st solar activity cycles.

Experimental data on cosmic ray fluxes obtained from Meteor satellite, as well as from other types of spacecraft like Cosmos, Explorer, Prognoz, Pioneer, Solrad, Theos, etc., are analyzed in correlation with characteristics of heliogeophysical phenomena, such as optical flare, radio and X-ray bursts, polar cap absorption, magnetic field variations, Forbush effect, etc. A number of regularities have been revealed during the investigations. It has been established that there is a dependence of proton event characteristics on heliocoordinates of the flare, which the observed proton fluxes are associated with, as well as on the state of interplanetary medium, in which the particles propagate from the source to the Earth (Belovsky, Ochelkov, 1976; Van Hollebeke, Ma Sung, McDonald, 1975; Pereyaslova, Nazarova, Petrenko, 1980; Akinyan, Fomichev, Chertok, 1977; Vernov and others, 1977). There is a relation of solar cosmic ray characteristics to the solar magnetic field parameters and the sector structure of the interplanetary magnetic field (IMF), in particular, to the origin of the near-earth sector boundary (Pereyaslova and others, 1976; Mikirova, Pereyaslova, 1978; Nesmeyanovich, Nesmeyanovich, 1972; Mikirova, Pereyaslova, 1977). There is a difference between processes of solar cosmic ray propagation from western and eastern flares, and the fact tells on the characteristics of the events considered (Ljubimov, Pereslegina, 1972; Kuzhevsky and others, 1978).

The regularities and relations between solar cosmic ray characteristics and heliogeophysical phenomena characteristics obtained in the course of investigations make it possible to carry out an operative evaluation of radiation situation in near-earth space and along spaceflight routes as well as to develop methods of radiation situation forecast.

To develop the scheme of solar cosmic ray event forecast it is necessary to predict the state of interplanetary medium, as well as flares followed by charged particle release.

The problem of solar flare forecast is being studied in the Soviet Union (the Crimean Astrophysical Observatory), in France (Medon) and in the USA (Boulder).

In the main solar flares occur in regions with strong magnetic fields and high field gradients. The forecast method developed in the Crimean Astrophysical Observatory is based on the study of the character of magnetic field evolution in the active region. A powerful flare can be expected in case of a high longitudinal magnetic field gradient and a complicated crossed magnetic field (Severny, Steshenko, 1970, 1972; Severny, Stepanyan, Steshenko, 1979).

The Medon Observatory predicts active region evolution on the basis of optical observations and magnetic field and centimeter radio burst data: a quiet centre - less than one chromospheric flare a day is expected; an eruptive centre - several chromospheric flares and not more than one radio burst a day; an active centre - at least one geophysical event a day; a proton centre - at least one proton flare (Simon, McIntosh, 1976; Simon, 1979).

In the forecast centre in the USA X-ray emission of flares serves as the main predictor, since there is a reliable correlation between the development of active region groups and X-ray bursts. If the active region is characterized by increased fluctuating X-ray fluxes, solar flare is likely to occur. Proton events with proton fluxes of  $E_p > 10$  MeV per 1 a.e., which are 10 times greater than the background level, are also predicted. 83% of proton event forecast prove to be correct, a number of active region characteristics being taken into account, as well (Simon, McIntosh, 1976; McIntosh, 1970).

To predict radiation situation it is insufficient to predict flare activity. It is essential to predict proton flux release into space, their spectrum, and intensity on the Earth's orbit.

Institute of applied Geophysics is working on the development of radiation situation forecast methods aimed at ensuring radiation safety of cosmic flights. Radiation situation forecast is based on

analyzing the whole complex of interconnected heliogeophysical phenomena. On the basis of aerological stations of the USSR State Committee for Hydrometeorology a network of stratosphere sounding stations was built (Shoina, Cape Schmidt, Cape Cheljuskin, Dalnerechinsk) to control high-energy solar cosmic ray component. To equip these stations a complex radiosound was constructed permitting to obtain meteorological and radiometric data simultaneously.

The investigation of heliogeophysical data received at the Service resulted in developing the technique for forecasting the parameters determining radiation situation. In 1974-1975 the technique for forecasting proton event development in the near-earth space was constructed on the basis of isotropic diffusion model and put into practice in the operative service (Avdjushin and others, 1979; Bezruchenkova, Nazarova, Pereyaslova, 1976). In subsequent years techniques for forecasting proton event parameters determining radiation situation in space were developed using solar activity characteristics, such as electromagnetic radiation and microwave solar bursts (Avdjushin and others, 1979; Bezruchenkova and others, 1977; Bezruchenkova, Pereyaslova, Frolov, 1978; Belovsky and others, 1978). The method of forecasting the time of arrival at the Earth and time profile of proton flux in solar cosmic ray event was also developed on the basis of photosphere solar magnetic field analysis (Avdjushin and others, 1979; Mikirova, 1981).

Modern theoretical models of burst-generated particle propagation were used in forecasting: Lupton-Stone model (Lupton, Stone, 1973), modified Krimigis isotropic diffusion model (Bezruchenkova, Nazarova, Pereyaslova, 1976; Krimigis, 1965), and complex model (Mikirova, 1981). During the analysis of sufficiently large amount of data on solar cosmic rays with proton energy  $E_p > 30$  MeV and small time required to reach the maximum (less than 10-12 hours), it was established that the period (time development of solar cosmic ray events) in which the major part of burst protons ( $\sim 90\%$ ) is registered, is mainly determined by the time required to reach the event maximum. Proceeding from the revealed similarity method a statistical model of forecasting the integral flux in solar cosmic ray event was constructed.

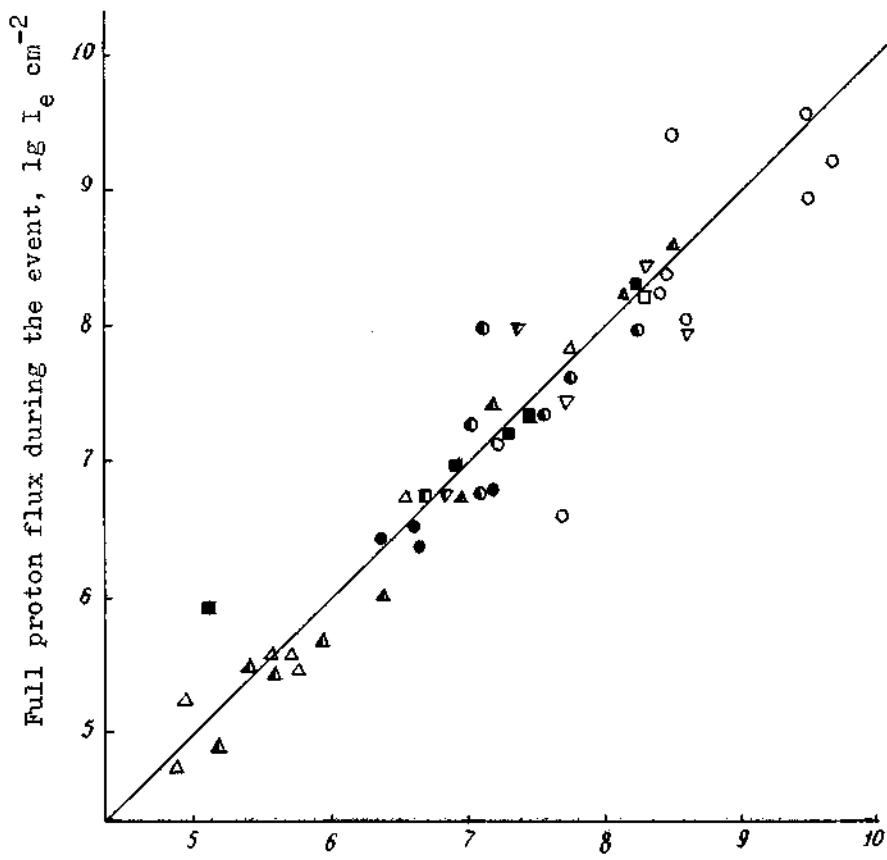
The establishment of relations between solar cosmic ray characteristics and electromagnetic radiation resulted in developing the methods of diagnosis and forecasting the parameters determining radiation situation in the near-earth space and along spacecraft routes: time required to reach the maximum for the proton flux with energies  $E_p > 5$  MeV; the integral proton flux within the whole solar cosmic ray event; the integral dose within the event (Bezruchenkova and others, 1977; Bezruchenkova, Pereyaslova, Frolov, 1978).

During the investigation of burst-generated proton coronal propagation in the energy range  $5 < E_p < 90$  MeV a relationship was established between the observed temporal, spatial and spectral characteristics of solar protons and photosphere magnetic field structure. The obtained regularities were used while developing the method of forecasting the following solar cosmic ray event characteristics: time of arrival of proton flux at the Earth's orbit; solar cosmic ray event duration; decrease constant and time of the maximum arrival; spectral characteristics ( $\gamma_i$ ,  $\gamma_{max}$ ) (Avdjushin and others, 1979; Bezruchenkova, Nazarova, Pereyaslova, 1976).

The technique of proton event express forecast according to burst emission makes it possible to obtain the following: the lower limit of the proton flux with  $E_p > 10$  MeV in the event maximum; the forecast of solar cosmic ray event with  $E_p > 60$  MeV (Avdjushin and others, 1979; Belovsky and others, 1978).

Fig.1 presents the logarithmic scale of forecasted and measured values of the full proton flux for solar cosmic ray events, the date of their registration on the Earth's orbit being given in the Table. The Table also presents the logarithms of registered ( $I_e$ ) and forecasted ( $I_f$ ) integral proton in solar cosmic ray events calculated according to particle transport models, as well as calculated error presented as a ratio of the difference between the experimental and forecasted values of the flux, and the measured value (%):

$$\delta = \frac{I_e - I_f}{I_e} \cdot 100\%$$



Full proton flux during the event,  $\lg I_p \text{ cm}^{-2}$

Fig.1. The correlation of measured full proton flux values with those predicted by models for solar cosmic ray events with different energies.

$E_p$	5 MeV	-	o	Diffusion	$\nabla$	Complex	u	Lapton-30MeVStatic
	25 MeV	-	•	model	v	model	■	Stone
	90 MeV	-	•		▼	model	▲	Static
					■	model	▲	10MeVmodel
					▲		△	

AES	Date of event re- gistra- tion	Ener- gy, MeV	$I_e$ , $\text{cm}^{-2}\text{s}^{-1}\text{sr}^{-1}$	Model Forecast			Statistic $I_f; \delta$
				Isotrop. diffu- sion $I_f; \delta$	Lupton- Stone $I_f; \delta$	Comp- lex $I_f; \delta$	
"Explor- er"	08.02.68	30	5,16				4,90 45
	25.07.68	30	5,59				5,47 24
	04.II.68	30	6,37				6,03 54
		60	5,70				5,50 37
	24.01.71	30	8,53				8,61 -17
		60	7,79				7,84 -II
	I6.05.71	60	4,82				4,75 I5
	01.09.71	30	8,20	8,28	-I8		8,25 -II
	03.I0.71	60	5,59				5,60 -23
"Meteor"	22.08.75	30	5,49				5,50 -23
		60	4,90				5,26 -56
	30.04.76	5	8,28	8,27 I,3	8,26 3,7	8,36-22	
		25	7,02	7,32	-33		
		30	6,89				6,77 24
		40	6,79	7,15	-40	6,72 I3	6,81-6,5
		60	6,59				6,72 -26
		90	6,38	6,41	-3		
	I9.09.77	15	8,02	7,04	95		7,34 78
	I3.02.78	5	9,46	8,50	89		
		25	7,55	7,37	33		
		30	7,15				7,45 -50
	28.04.78	5	9,62	9,54	5,3		
		25	7,74	7,59	29		
		90	6,64	6,28	55		
	07.05.78	5	8,32	8,30	4,3		
		25	7,18	7,16	8,3		
		90	6,59	6,57	4,9		
	22.06.78	15	6,49	6,44	I0		
	23.09.78	5	9,69	9,21	67		
		25	8,25	7,96	49		
		90	7,19	6,80	59		
	I7.02.79	5	7,69	6,60	92		7,49 37
		30	5,90				5,68 40
		60	5,65				5,60 II
	06.06.79	5	8,59	8,03	72		7,96 77

The greatest deviation from the measured value is observed when diffusion and complex models are used, however in case of the proton flux with the threshold energy exceeding 25 MeV, the forecasted value precision increases and is slightly variable for different models.

The obtained result also account for the difference in the propagation laws of protons with the energy  $E_p > 5$  MeV and  $E_p > 25$  MeV. The best precision is obtained when Lupton-Stone model is used, as the model considers the three main processes of solar cosmic ray particle interaction with interplanetary medium: anisotropic diffusion, convective transport and adiabatic cooling. Yet at present it appears impossible to use this model in operative practice. Statistic model is the most rational and acceptable.

The conducted developments resulted in creating the complex of fully automatized methods of forecasting the parameters determining radiation situation.

An important role in forecasting is played by international helio-geophysical information exchange, carried out within the International Ursigram and the World Days Service, as well as through bilateral exchange between prognostic centres. Data on active centres, burst and subsequent event characteristics, as well as current forecasts are being regularly obtained from the Space Environment Services Center, Boulder (Heckman, 1979). These data are used as input information in radiation situation forecasts.

More perfect methods of forecasting radiation situation in the near-earth space and along spacecraft routes will be developed in future on the basis of further investigations of cause-conditioned relations between the near-earth radiation situation and solar activity, interplanetary media and the Earth's magnetosphere. Further experimental and theoretical investigations will help to construct a model of the near-earth radiation situation, which will improve the precision of forecasts, as well as create the premises of solving the problem of solar-terrestrial relationships.

- Akinyan S.T., V.V. Fomichev, I.M. Chertok (1977). Opredelenie parametrov solnechnykh protonov v okrestnosti Zemli po radiovspleskam. "Funktsiya dolgotnogo oslableniya". Geomagnetizm i aeronomiya, 17, 177-183.
- Avdyushin S.I., N.K. Pereyaslova, F.L. Dlikman, Yu.M. Kulagin (1979). Forecastings of solar energetic radiation at the forecast center of the Institute Of Applied Geophysics. "Solar-Terr. Predictions Proc. and Workshop Program."(Boulder), 1, 89-103.
- Belovsky M.N., Yu.P. Ochelkov (1976). Zavisimost nabljudajemoi pikovoi intensivnosti solnechnykh kosmicheskikh luchei ot gelio-graficheskoi dolgoty. Solnechnye dannyje, 12, 73-76.
- Belovsky M.N., Yu.P. Ochelkov, T.S. Podstrigach, M.A. Fasakhova (1978). O prognozirovaniii protonnykh yavleniy po radiovspleskam. In Radioizluchenie Solntsa, 4, LGU, 44-51.
- Besprozvannaya A.S., B.I. Sazonov, T.M. Schuka (1972). Svjaz zemnykh poteplenij v stratosfere s protonnymi vspyshkami. Tr.Arkt. i antarkt. NII, 311, 34-39.
- Bezruchenkova T.M., M.N. Nazarova, N.K. Pereyaslova (1976). Primenenie modeli izoptropnoy diffuzii dlja prognozirovaniya razvitiya potokov solnechnykh kosmicheskikh luchey v vysokoshirotnykh zonakh magnitosfery Zemli. Geomagnetizm i aeronomiya, 16, 592-598.
- Bezruchenkova T.M., N.A. Mikrjukova, N.K. Pereyaslova, S.G. Frolov (1977). Diagnostika protonov solnechnykh vspyshek po soprovozh-dajuschemu electromagnitnomu izlucheniju. Geomagnetizm i aeronomiya, 17, 820-825.
- Bezruchenkova T.M., N.K. Pereyaslova, S.G. Frolov (1978). Diagnosticheskoje prognozirovaniye radiatsionnykh kharakteristik solnechnykh vspyshek po parametram electromagnitnogo izluchenija. Geomagnetizm i aeronomiya, 18, 992-997.
- Cuffner J.(1971). Yadernoje izluchenije i zaschita v kosmose. Atomizdat, Moscow.
- Heckman G.R. (1980). Predictions of the space environment Services Center. Solar-Terr. Predictions Proc. and Workshop Program, Boulder, Colo., 1979, 1. Washington, D.C., 322-349.
- Kostjuchenko I.G., N.B. Malyshev, P.M. Svidsky, S.G. Frolov (1981). O vlijanii vozmuschennosti magnitosfery na shirotnoje raspredele-nije potokov SKL. In Tr.IPG "Voprosy fiziki atmosfery i radia-tsvyonov kosmofiziki", 50, 35-40.

- Krimigis S.M. (1965). Interplanetary diffusion model for the time behavior of intensity in a solar cosmic ray events. J. Geophys. Res., 70, 2913-2960.
- Kuzhevskiy B.M. and co-workers (1978). Issledovaniye rasprostraneniya SKL ot vostochnykh vspyshek May 28, June 15, August 2, 1972 na sputnike "Prognoz". "Model rasprostranenija solnechnykh chasits v mezhplanetnoy srede". Kosmicheskie issledovaniya, 16, 250-256.
- Ljubimov G.P., N.V. Pereslegina, N.N. Kontor (1972). O rasprostranenii solnechnykh kosmicheskikh luchey v mezhplanetnoy srede. Izvestija AN SSSR, 36, 2297-2303.
- Logachev Yu.I., I.A. Savenko, A.I. Sladkova (1974). Dozy radiatsii v mezhplanetnom prostranstve po dannym sputnikov "Prognoz". Kosmicheskie issledovaniya, 12, 789-791.
- Loginov V.F., E.I. Sazopov (1978). Kosmicheskie luchi i troposfernaja tsirkuljatsiya. Phys. Solariterr., 85-92.
- Lupton J.R., E.C. Stone (1973). Solar flare particle propagation of a new analitic solution with spacecraft measurements. J. Geophys. Res., 78, 1007-1018.
- McIntosh P.S. (1970). Techniques of solar flare forecasting. In V. Agg (Ed.), Ionospheric forecasting, AGARD Conf. Proc., No.49, Tech. ed. Rept. Ltd, London, 81.
- Mikirova N.A., N.K. Pereyaslova (1977). Vlijaniye fotosfernykh magnitnykh poley na osobennosti rasprostranenija protonov solnechnykh vspyshek. DAN SSSR, 234, 798-801.
- Mikirova N.A., Pereyaslova N.K. (1978). Osobennosti vremennykh i spektralnykh kharakteristik kosmicheskikh luchey, obuslovленnye vlijaniem sektornoy struktury mezhplanetnogo magnitnogo polja. Geomagnetizm i aeronomiya, 18, 583-589.
- Mikirova N.A. (1981). Fizicheskie osnovy metoda prognozirovaniya vremennykh i spektralnykh kharakteristik SKL na osnove fotosfernykh magnitnykh poley. In Trudy IPG, "Voprosy fiziki atmosfery i radiatsionnoy kosmofiziki", 50, 35-40. Gidrometizdat, Moscow.
- Nesmeyanovich A.T., E.I. Nesmeyanovich (1972). Khromosfernyje vspyshki i udarnyje volny v mezhplanetnom prostranstve. In, Problemy kosmicheskoy fiziki, 2.
- Pereyaslova N.K., M.N. Nazarova, S.M. Mansurov, L.G. Mansurova (1976). Nekotoryje zakonomernosti svjazi solnechnykh kosmiches-

- kikh luchey v poljarnykh oblastjakh s sektornoy strukturoy mezh-planetnogo polya. Geomagnetizm i aeronomiya, 16, 407-412.
- Pereyaslova N.K., M.N. Nazarova, I.E. Petrenko (1980). Proyavlenija solnechnoy aktivnosti na vетvi rosta 21-go tsikla po issledovanijam kosmicheskikh luchey na ISZ "Meteor". Geomagnetizm i aeronomiya, 20, 977-981.
- Petrov V.M., A.V. Kolomensky, M.V. Zil (1979). Radiatsionnaya opasnost solnechnykh vspyshek v okolozemnom kosmicheskom prostranstve. 2. Metodika otseki dozy i riska ejo prevyshenija. Kosmicheskie issledovaniya, 17, 122-126.
- Severny A.B., N.V. Steshenko (1970, 1972). Method for forecasting of solar flares. Solar Terrestrial Physics, part 1, 173.
- Severny A.B., N.N. Stepanyan, N.V. Steshenko (s.a.). Soviet short-term forecasts of active region evaluation and flare activity. Solar-Terr. Predictions Proc. and Workshop Program. Boulder, Colo., 1979, 1, 89-103. Washington, D.C.
- Silberger R., C.H. Tsao (1979). Biological hazards associated with cosmic-ray and solar flare exposures. 16 Int. Cosmic Ray Conf., Kyoto, 1979, 5, SP.Sess, 317-322. S.I.
- Simon P. (s.a.). The forecasting center of Meudon, France. Solar-Terr. Predictions Proc. and Workshop Program. Boulder, Colo., 1979, 1, 1-11. Washington, D.C.
- Simon P., E.C. McIntosh (1976). Obzor sovremennykh solnechnykh progностических центров. In, "Nabлюдения и прогноз солнечной активности", 289-301. Mir, Moscow.
- Smart D.F., M.A. Shea (1980). A computerized "event mode" solar proton forecasting technique. Solar-Terr. Predictions Proc. and Workshop Program. Boulder, Colo., 1979, 1, 406-427. Washington, D.C.
- Van Hollebeke M.A., L.S. Ma Sung, F.B. McDonald (1975). The variation of solar proton energy spectra and size distribution with heliolongitude. Solar Phys., 41, 189-223.
- Vernov S.N. and co-workers (1977). Proton fluxes in interplanetary space relevant to the active processes on the Sun. "15th Int. Cosm. Ray Conf. Plovdiv, 1977", 3, 142-149". Sofia.
- Wilcox J.M. (1979). Solar activity and changes in atmospheric circulation. Solar-Terr. Phys. Proc. 5th Int. Symp., Insbruck, 29 May - 3 June, 1978. J. Atmos. and Terr. Phys., 41, 753-763.

## MAJOR MEDICAL RESULTS OF THE SALYUT-6—SOYUZ 185-DAY SPACE FLIGHT

O. G. Gazenko, A. M. Genin and A. D. Egorov

*Moscow, USSR*

### ABSTRACT

The major task of medical investigations in the 185-day manned space flight was to accumulate information about reactions of the vital systems of the body to a prolonged exposure to weightlessness. Inflight, cardiovascular responses, changes in body mass and leg volume were investigated in detail. Postflight, a wide range of studies, including motor function, cardiovascular activity, hematology, biochemistry, and immunology, was carried out. These investigations once again demonstrated that man can well adapt to a 6-month exposure to orbital flight, maintaining good health condition and high work capacity.

The changes observed were adequate reactions to the environmental effects, they were reversible and returned to normal after a relatively short period of readaptation. Medical control over the use of countermeasures and rational arrangement of the work-rest cycle contributed significantly to the maintenance of good health condition and high work capacity as well as to the facilitation of readaptation after recovery. The medical investigations performed during and after flight showed that the duration of manned space missions can be gradually increased.

### KEYWORDS

Space flight, medical investigations, body mass, cardiovascular system, provocative tests, motor activity, vestibular function, biochemical, immunological and hematological parameters, readaptation.

### INTRODUCTION

In 1980, the fourth expedition performed a 185-day space flight onboard the orbital complex Salyut-6 - Soyuz. The crew consisted of the Commander L.I. Popov (Cr-4) and Flight-Engineer V.V.Ryumin

(FE-4) who participated in the 175-day flight the previous year.

The highlights of the 185-day space flight were four visiting crews (three international) and three unmanned cargo vehicles Progress-9, 10, 11 that visited the station.

In the 185-day space flight medical investigations of human responses to the effects of prolonged weightlessness were continued and enlarged.

#### INFLIGHT OBSERVATIONS AND INVESTIGATIONS

Upon insertion into orbit both cosmonauts showed moderate sensations of blood rush to the head, puffiness of the facial tissues, noticeable nasal voice. Cr reported illusionary displacement of the panel. He also mentioned that his face was hot. These sensations diminished by the end of the first week and disappeared entirely by the 17th flight day. FE emphasized that in this flight his adaptation to zero-g continued for a shorter period of time than in the preceding 175-day flight. Following the adaptation period, the health state of the crewmembers was good throughout the flight. Occasionally, they indicated the feeling of lassitude which was relieved by the night sleep. Their appetite, urination and defecation were normal. They slept for about 7-8 hours. Cr-4 fell asleep easily, sometimes having dreams. After 2 months FE-4 sometimes lay awake for long and in this case he used to take sleeping pills.

Anthropometric studies indicated that body mass of both cosmonauts was higher than preflight during the entire flight (except for days 4 and 7 in FE-4 whose body mass decreased by 0.6 and 0.2 kg, respectively) and tended to grow linearly with flight time (Fig.1). Body mass plateaued around day 140 when it exceeded the preflight level by 1.9-3.2 kg and 4.3-4.7 kg in Cr-4 and FE-4, respectively.

Thus, inflight increase of body mass of both cosmonauts points out that metabolic losses can be made up by proper diets and adequate water consumption.

Leg volume of both cosmonauts decreased significantly during the first 10 days: at that time the decrease was 7.4-9.7% ( $156-219 \text{ cm}^3$ ). Later, the losses progressed (Fig. 2) and by day 70 reached 15-16% ( $312-380 \text{ cm}^3$ ). In the second half of the flight the leg volume tended to increase, especially in FE-4. At the end of the flight the leg volume deficiency in Cr-4 and FE-4 was 5.8% ( $120 \text{ cm}^3$ ) and 13.8% ( $327 \text{ cm}^3$ ), respectively.

Thus, the type of changes in leg volume was similar in both crew-members. The difference in the level of recovery of the parameter in the second half of the flight can be attributed to the different scope and structure of exercises the cosmonauts performed.

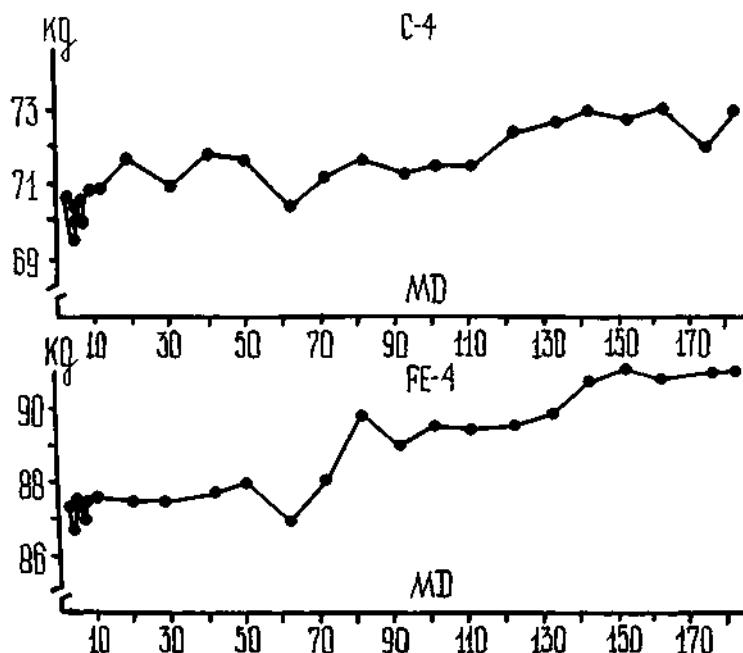


Fig. 1 Body mass of the Salyut-6 - Soyuz 185-day crewmembers. MD=mission day. Dashed line indicate mean value preflight, broken line - mean value inflight. Cr-4=Commander. FE-4=Flight-engineer.

Cardiovascular studies revealed the following changes at rest:

- increase in heart rate of Cr-4 during the entire flight and of FE-4 during the first four months that was higher than the mean preflight values (in FE-4 only during the 2nd-5th months) (Fig. 3);
- statistically significant shortening of isometric contraction time in FE-4 and isometric relaxation time in both cosmonauts (Fig. 4);
- noticeable shortening of the ejection time in Cr-4 and increase of the fast filling time in both cosmonauts (in FE-4 only at the end of the flight);
- increase of cardiac output in both cosmonauts during most measurements;

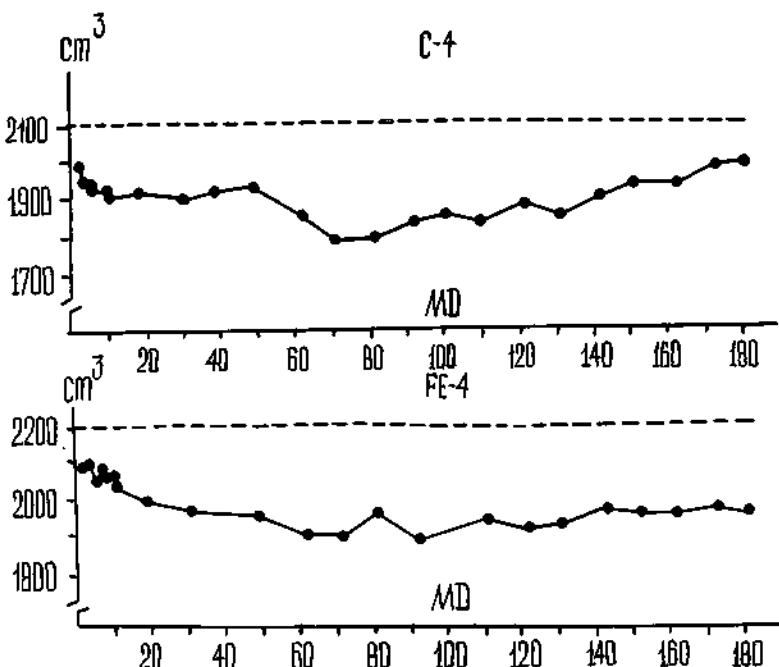


Fig. 2 Leg volume of the Salyut-6 - Soyuz 185-day crewmembers. MD=mission day. Dashed line indicate mean value pre-flight, broken line - mean value inflight. C-4=Commander. FE-4=Flight-engineer.

- decrease of minimal in Cr-4 and side systolic and mean pressure in FE-4;
- decrease of specific peripheral resistance and its ratio to the normal values;
- increase in the velocity of pulse wave propagation along the aorta;
- decrease of pulse blood filling of leg vessels;
- variations of pulse blood filling of cerebral vessels within preflight limits;
- development of asymmetric blood filling of the large hemispheres (in Cr-4 at the end of and in FE-4 during the entire flight) and in some cases of venous waves in cerebral rheograms;

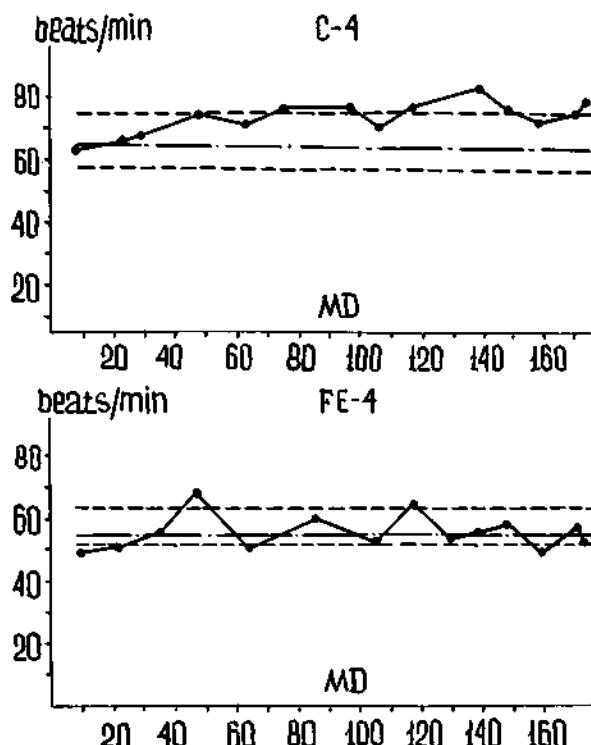


Fig. 3 Heart rate of the Salyut-6 - Soyuz 185-day crewmembers at rest. MD=mission day. Broken line indicate mean value inflight, solid line - mean value preflight, dashed lines - limits of variations preflight. C-4=Commander. FE-4=Flight-engineer.

- decline of the tone of small cerebral vessels (more expressed in FE-4) and development of asymmetry of the parameter;
- decline of venous pressure in lower extremities (from 22-32 mm Hg preflight to 7.5-12.5 mm Hg inflight) approximating that in upper extremities;
- increase of pressure in jugular veins;

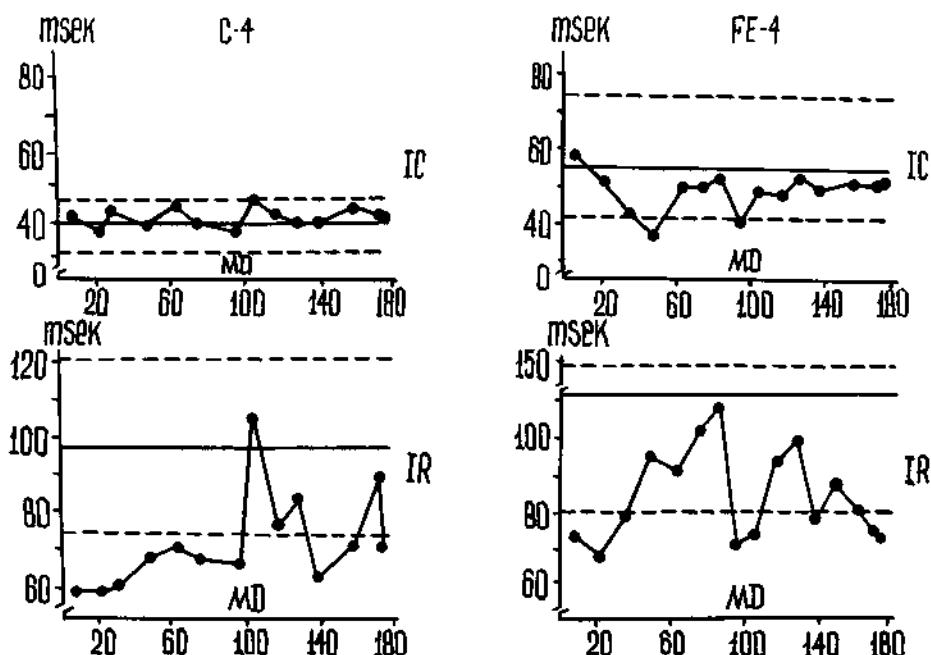


Fig. 4 Time of isometric contraction and isometric relaxation of the left ventricle in the Salyut-6 - Soyuz 185-day crew-members at rest. MD=Mission day. IC=Isometric contraction. IR=Isometric relaxation. Broken line indicate mean value inflight, solid line - mean value pre-flight, dashed lines - limits of variations preflight. C-4=Commander. FE-4=Flight-engineer.

- reduction of the venous pressure gradient in the upper and lower body;
- increase in distensibility and capacity of leg veins (as estimated by the plateauing of the plethysmographic curve) and simultaneous decrease of contractility (decline of venous tone);
- increase in contractility of forearm veins and their essentially unchanged distensibility;
- decrease of arterial inflow to the leg vessels and increase of that to the forearm vessels.

The above hemodynamic changes seem to result from cranial blood shifts in weightlessness that lead to increased venous return and preload of the heart. It is also probable that decrease in the postural-tonic activity and load upon the antigravitational musculature in weightlessness and constant deficiency of muscle work in prolonged space flight reduce the work of the peripheral muscle heart (Arinchin and Nedvetskaya, 1974) involved in blood displacement and increase the role of the systolic work of the heart and its sucking function (active diastole) in hemodynamics. It is likely that all these factors generate a certain load on the heart and increase the strength of cardiac contraction which manifest in rearrangement of the phasic structure of cardiac cycle. Increased blood filling of the cranial compartments seems to be responsible for involvement of reflex mechanisms from receptors of central veins, atria and cardiopulmonary area (Chernigovsky, 1960; Marshall and Shepherd, 1968; Parin, 1946), thus resulting in partial compensation of the changes (decrease of the vascular tone, some parameters of arterial pressure and peripheral resistance, fluid losses, tendency for normalization of the phasic structure of the cardiac cycle).

Changes in bioelectrical activity inflight included an insignificant shortening of the intraventricular and increase of the atrio-ventricular conductance (no more than by 0.02 sec), increase of the R/T ratio and displacement (only in FE-4) of the electrical axis to the left.

Postflight, both crewmembers displayed a reduction of the repolarization vector.

Thus, changes in the myocardial bioelectrical activity were insignificant and seemed to be induced by positional changes, development of autonomic dysbalance and metabolic changes in the myocardium.

The provocative test with lower body negative pressure (at -25 mm Hg for 2 min and at -35 mm Hg for 3 min) inflight brought about the following cardiovascular changes as compared to the preflight exposure:

- greater increase of the absolute value of heart rate (Fig. 5);
- smaller increase of isometric contraction and relaxation time (Fig. 6);
- greater decrease of the absolute value of the ejection time, fast filling time and, as a rule, of the intrasystolic index (Fig. 7);
- greater decrease of stroke volume;
- greater decrease of the minimal arterial pressure;

- similar decline of pulse blood filling of cerebral vessels;
- greater increase of the velocity of pulse wave propagation along the aorta.

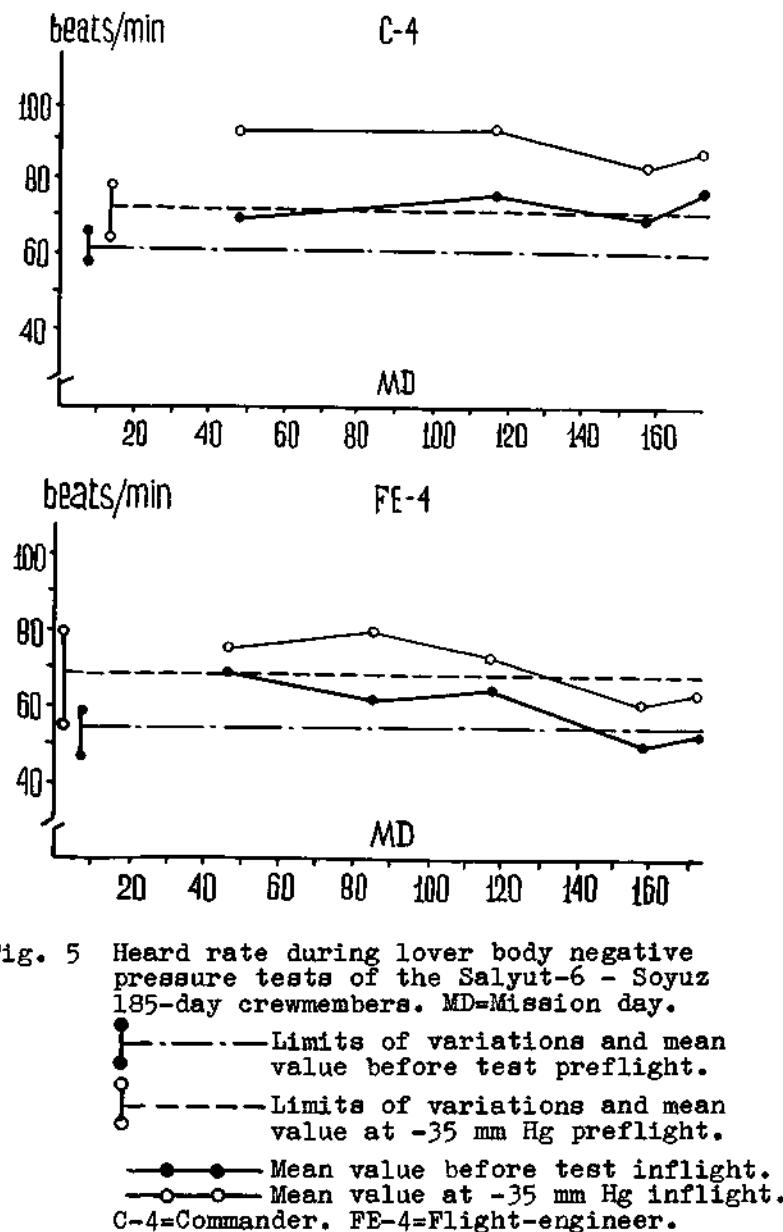


Fig. 5 Heard rate during lower body negative pressure tests of the Salyut-6 - Soyuz 185-day crewmembers. MD=Mission day.

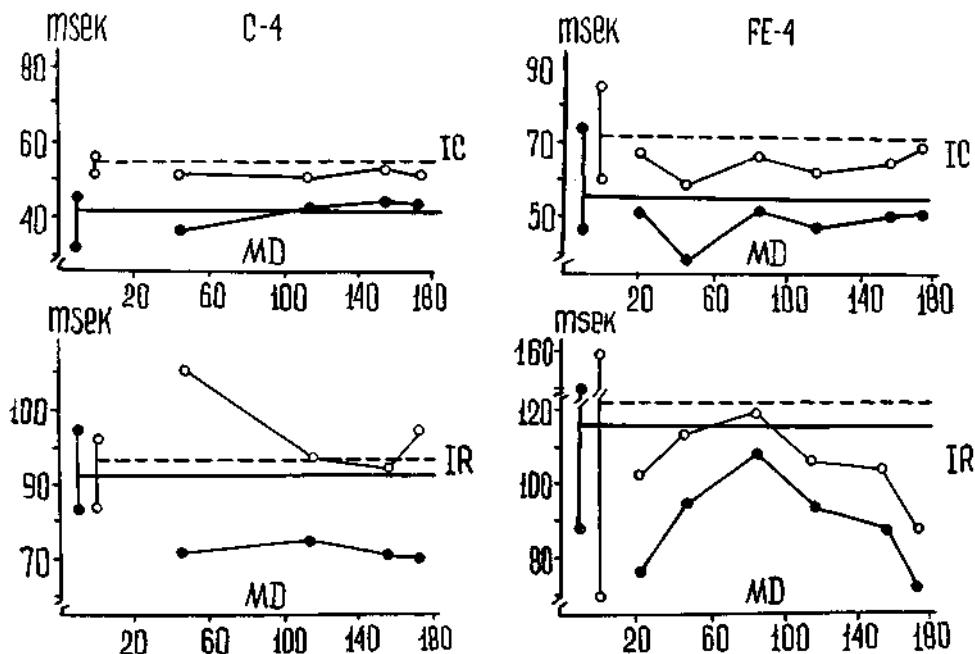


Fig. 6 Time of isometric contraction and isometric relaxation of the left ventricle during lower body negative pressure tests in the Salyut-6 - Soyuz 185-day crewmembers. MD=Mission day. IC=Isometric contraction. IR=Isometric relaxation.

● ————— Limits of variations and mean value before test preflight.  
 ○ ————— Limits of variations and mean value at -35 mm Hg preflight.  
 —●—●— Mean value before test inflight.  
 —○—○— Mean value at -35 mm Hg inflight.  
 C-4=Commander. FE-4=Flight-engineer.

The exercise test (pedaling of the bicycle ergometer at a rate of 60 rpm and with a workload of 750 kgm for 5 min) inflight gives rise to cardiovascular reactions that differ from those prelaunch:

- greater increase of the absolute value of heart rate in Cr-4 and smaller increase of the parameter in FE-4 (Fig. 8);
- greater decrease in most test of the absolute value of isometric contraction and relaxation time;

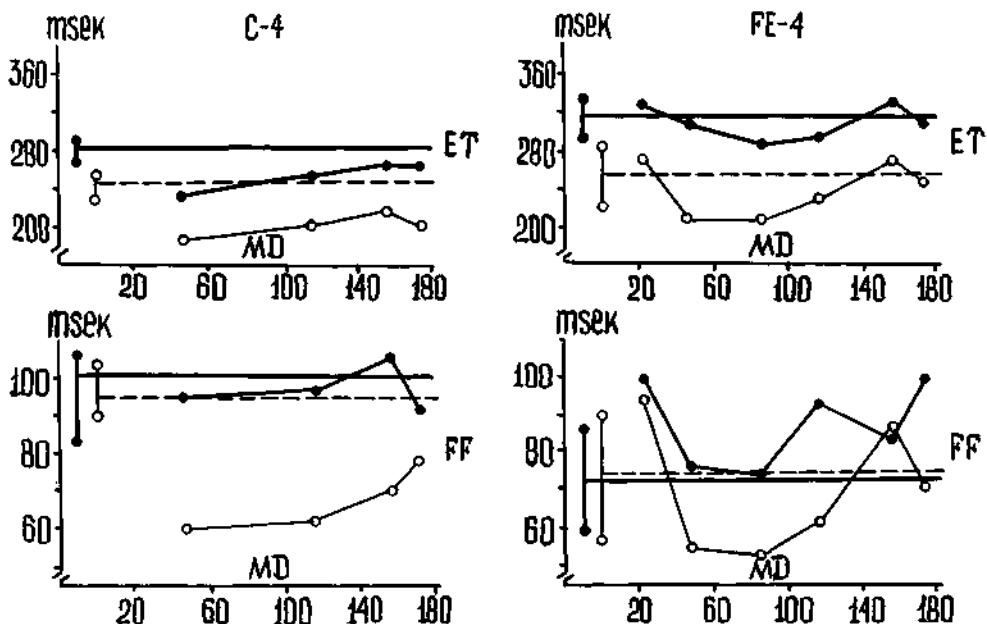


Fig. 7 Left ventricular ejection time and fast filling time during lower body negative pressure tests in the Salyut-6 - Soyuz 185-day crewmembers. MD=Mission day.  
 ET=Ejection time. FF=fast filling of ventricles.

- — Limits of variations and mean value before test preflight.
- — Limits of variations and mean value at -35 mm Hg preflight.
- Mean value before test inflight.
- Mean value at -35 mm Hg inflight.

 C-4=Commander. FE-4=Flight-engineer.

- lack of changes or decrease of stroke volume and increment of cardiac output at the expense of enhanced chronotropic influences;
- increase of the minimal pressure in Cr-4 in some tests;
- more marked increase of the absolute value of the velocity of the pulse wave propagation along the aorta.<sup>1</sup>

<sup>1</sup> All the parameters, except for heart rate, were examined at the so-called stage of early recording, i.e. during the first minute after the test was completed.

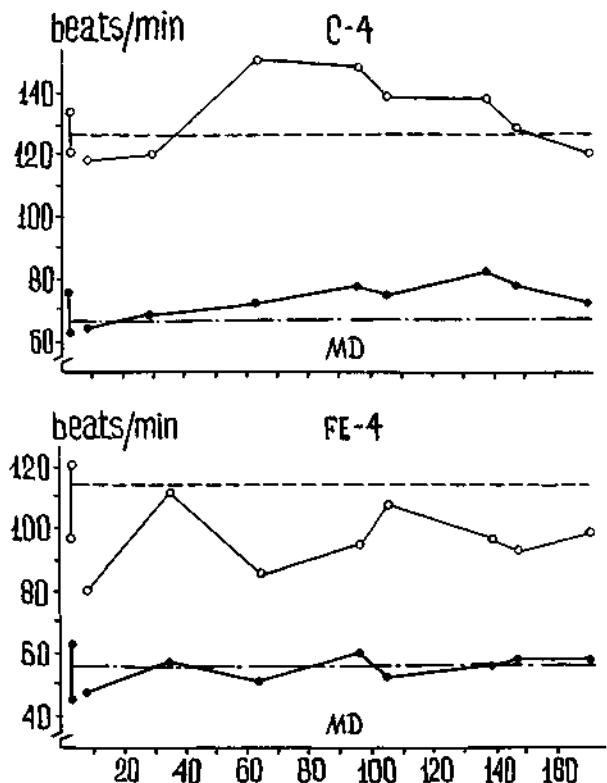


Fig. 8 Heart rate during exercise tests of the Salyut-6 - Soyuz 185-day crewmembers.  
MD=Mission day.

- — Limits of variations and mean value before test preflight.
  - — Limits of variations and mean value during test preflight.
  - Mean value before test inflight.
  - Maximum value during test inflight.
- C-4=Commander. FE-4=Flight-engineer.

Thus, these data indicate that physical fitness declined slightly in Cr-4 and remained at a high level in FE-4 by the end of the flight.

The proposed scheme of mechanisms of cardiovascular changes is as follows.

The decrease of the circulating blood volume in weightlessness and blood pooling in arterioles of the working muscles during exercise as well as simultaneous increase of the tension of vein walls in the working and resting extremities (Bevegard and Shepherd, 1965) seem to be responsible for the lack of an increase in the stroke volume.

Exercises produce blood redistribution in the cardiopulmonary area, resulting in an increase of the lung volume (Harrison, Goldblatt and Braunwald, 1963; Ruosteenoja, Linko and Lind, 1958) in weightlessness; this may cause blood displacement into the vessels of the leg working muscles, including the area of collapsed veins or of free compatibility that forms in weightlessness.

Due to the hypothesized greater than on the Earth decline of the pulmonary pressure, reflex-mediated adrenergic effects are enhanced, thus leading to increased strength of cardiac contractions and more marked shortening of the isometric contraction time.

It can be thus inferred that specific changes of cardiovascular reactions to LBNP tests inflight may depend on the larger decline of stroke volume and enhancement of adrenergic reactions due to the greater than on the Earth blood shifts to the abdominal cavity and lower extremities. Specific changes of cardiovascular reactions to exercise tests inflight are associated with a greater shift of blood to the working muscles and reflex-mediated increase of inotropic influences.

#### POSTFLIGHT OBSERVATIONS AND INVESTIGATIONS

After recovery (at the landing site) the crewmembers reported that they felt tired and that the weight of their bodies and objects with which they manipulated was seemingly increased. This feeling disappeared 2 or 3 days postflight.

The examinations at the recovery site revealed in both cosmonauts lassitude, lability of heart rate and arterial pressure, skin paleness and increased perspiration, diminished locomotor function, and disturbed movement coordination when walking. FE-4 had a pasty and puffy face. Both cosmonauts tolerated well without complaints or difficulties the standing 5 min test.

Study of the motor function demonstrated various changes: decrement of the volume and circumference of lower extremities, slight muscle mass losses which were masked by increased accumulation of the adipose tissue. The body weight of Cr-4 was similar to the prelaunch value and that of FE-4 increased by 3.5 kg. The femur and calf circumference decreased in Cr-4 by 2.3 and 3.3 cm and in FE-4 by 3.8 and 3.4 cm, respectively. The examination on day 4 postflight showed subatrophic changes of the long muscles of the back, chest and abdominal muscles; their size, velocity and strength characteristics as well as their tone declined.

During the first postflight days changes in the mechanisms of co-ordination of the posture, walking and other motor acts were seen which could be associated with the involvement of postural and antigravitational mechanisms on the Earth.

The time of recovery of the equilibrium in response to perturbations applied from the outside increased, thus indicating a mismatch of postural synergisms. On the whole, the level of motor changes was low, being in some respects less marked than after short-term flights.

Cardiovascular studies demonstrated instability of the heart rate and arterial pressure during the first postflight days. Echocardiographic examination performed at R + 0 and later revealed no significant changes in the heart size and function.

Electrocardiographic study showed metabolic changes of the myocardium. The provocative tests (exercise test, tilt test) were well tolerated, suggesting high reserve capabilities of the body post-flight.

Study of hemodynamics and phase structure of the cardiac cycle during exercise tests<sup>2</sup> at R + 7 found no pathologies during and after exercises. However, both cosmonauts displayed decreased circulation regulation. In Cr-4, this included a decrease of the circulation volume and development of the phasic syndrome of myocardial hypodynamics (elongation of isometric contraction time, and shortening of the ejection time), and in FE-4 this included increase in heart rate and signs of the phasic syndrome of myocardial hyperdynamics. The myocardial hypodynamic syndrome in Cr-4 was probably associated with the reduced venous return.

Study of fluid-electrolyte metabolism and renal function indicated that at R + 0 renal excretion of sodium and potassium remained essentially unaltered in both cosmonauts which could have been a result of water-salt supplements taken before reentry.

In the first and second postflight days renal excretion of sodium and potassium in Cr-4 decreased, with potassium excretion predominating over sodium excretion. The decline in sodium excretion was likely to be associated with changes in sodium transport in tubules since its content in the diet was adequate.

From the first postflight days FE-4 showed a high sodium concentration in the urine; together with high fluid excretion this induced increased sodium excretion both in absolute values and as related to its entry into the body. The Na/K ratio decreased in Cr-4 and increased in FE-4. These data gave indirect evidence for an altered activity of mineralocorticoids. At R + 0, Cr-4 displayed a marked hypokalemia and FE-4 a decreased potassium concentration in blood. The serum content of total calcium and its ionized fraction did not grow in the two cosmonauts, in contrast to other

<sup>2</sup> The exercise test was pedaling of the bicycle ergometer at a rate of 60 rpm for three min at each step with a workload of 450, 600, 750, 900, 1050 kgm/min.

crewmembers who participated in previous long-term flights. At R + 0 these parameters were below the preflight level. In both cosmonauts, the blood content of PTH was lower than the mean values, and the content of calcitonin in Cr-4 was equal to the mean level and in FE-4 was below it. In both crewmembers, the renal excretion of calcium was higher than preflight, as was the case in previous flights. Increased renal excretion of calcium, decreased blood content of ionized calcium and essentially unchanged glomerular filtration could be related to the reduced calcium reabsorption in tubules.

The potassium chloride loading test brought about increased potassium excretion. However, these changes were less expressed than after the 175-day flight. During the test, excretion of divalent ions at the stage of maximum diuresis also increased.

Blood examinations showed a decrease in the red blood cell mass (by 1.0-1.2 mln) and hemoglobin concentration in both crewmembers which continued to decline for a certain period of time because the liquid portion of blood recovered faster than formed elements. At the same time the count of reticulocytes, red blood precursors increased. Later, the red blood cell count and hemoglobin concentration returned to normal.

Biochemical investigations postflight were performed to measure activity of blood enzymes, hormones, fat and carbohydrate metabolism as well as to evaluate activity of the sympathoadrenal system.

The study of the Krebs cycle enzymes characterizing energy metabolism showed a decrease in total activity of isocitrate dehydrogenase and (only in Cr-4) malate dehydrogenase with changes in its isoenzymes (a marked decrease of cytoplasmic fractions of MDH<sub>2</sub> and a simultaneous increase of the mitochondrial fraction of MDH<sub>3</sub>), thus suggesting a decline in the rate of oxidative processes. The reduced activity of malate dehydrogenase also occurred after 96-, 140- and 175-day flights.

In both cosmonauts the total activity of lactate dehydrogenase increased, and in FE-4 the activity of cardiac isoenzyme LDH<sub>1</sub> also grew.

The total activity of creatine phosphokinase as well as the activity of skeletal-muscular and cardiac isoforms increased in both crewmembers. The increase in CPK activity at the expense of its skeletal-muscular isoform seems to be related to the changes in the musculo-skeletal system. More marked changes of CPK in FE-4 in this flight as compared to the similar data in his previous 175-day flight are indicative of greater metabolic changes in skeletal myocytes. This can be associated with changes in the permeability of myocyte structures. The increase in total activity of lactate dehydrogenase of both cosmonauts as well as the increment of the total activity of LDH<sub>1</sub> and the cardiac isoforms of CPK in FE-4 are possible reflections of metabolic changes in the myocardium or permeability of its membrane structures. The increase in alanine aminotransferase (mostly in FE-4) and simultaneous decrease of aspartate aminotransferase and hepatic alkaline phosphatase isoenzyme may suggest alterations in liver metabolism (Komarov and others, 1976) due to weightlessness-induced circula-

tion changes in this organ (Yarullin and Vasilyeva, 1979). The increase in the activity of bone alkaline phosphatase isoenzyme may be a result of an enhanced osteoblastic function. Postflight, blood lipid changes were different, remaining essentially within the normal limits. Mention should be made of a significant decline of triglycerides and an increase of nonesterified fatty acids in Cr-4 as a specific feature. Carbohydrate changes included a noticeable increase in the pyruvate concentration, a slight increase in the glucose concentration and a simultaneous decline of the lactate content.

Thus, the 185-day flight brought about certain metabolic changes aimed at maintaining homeostasis in an unusual environment with no pronounced metabolic changes.

Hormonal studies demonstrated different changes in the hormone content in blood. Preflight, renin activity in Cr-4 and aldosterone content in FE-4 increased; the ratio of pressor and depressor prostaglandins changed and renal excretion of total 17-HOCS increased; these features normally accompany an emotional stress (Markov, 1978; Tigranyan and others, 1980). Similar changes were also seen before other space missions (Kalita and others, 1979).

Postflight, changes in the renin-angiotensin-aldosterone system were seen which included increases in the renin activity and aldosterone content in Cr-4 and a decrease in the aldosterone content in FE-4. In general, these changes can be regarded as those meant to maintain fluid-electrolyte balance. At R + 1, there was a significant increase in the blood content of cyclic adenosine monophosphates in both cosmonauts and guanosine monophosphates in Cr-4. The increase in the blood content of cyclic nucleotides occurred due to a release of various hormones into blood. The blood content of pressor prostaglandins decreased and that of depressor prostaglandins increased. The postflight changes in the prostaglandins of pressor and depressor groups are individual and can be considered as compensatory reactions aimed at increasing the vascular tone and preventing orthostatic intolerance.

Investigations of the sympatho-adrenal system showed a postflight increase in catecholamines which may reflect readaptation of the human body to the Earth environment after a prolonged exposure to weightlessness. Similar changes in the sympathoadrenal system were seen after the 140- and 175-day flights. Variations in the biochemical parameters observed after flights of different duration can be associated with the wave-pattern of adaptation to weightlessness.

Studies of amino acid metabolism demonstrated a decline of the amino acid pool (14 amino acids were measured) in blood of both cosmonauts at R + 1. The total amino acid content decreased in Cr-4 from 17.5 mg% preflight to 14.4 mg% postflight, and in FE-4 from 21.4 mg% preflight to 10.3% postflight. On the 8th day post-flight the parameter returned to normal in FE-4 and remained reduced in Cr-4. The decline in the total amino acids seems to be related to a specific rearrangement and development of a new level of amino acid metabolism in space flight.

Studies of the cholinergic system of blood that plays an important part in the entire complex of adaptive reactions revealed no deviations in the content of acetyl choline in whole blood and acetyl cholinesterase activity in red blood cells of both cosmonauts. At R + 1, activity of nonspecific cholinesterase was reduced by 56% in both crewmembers; this may be connected with the liver malfunction.

Immunological and allergological investigations showed a decrease in the content of thymus-dependent lymphocytes in Cr-4 and a drastic decline of the functional activity of these immunocompetent cells in FE-4 at R + 1. On the 9th day postflight these parameters returned to the preflight level. On the 1st and 9th postflight days the helper activity of T-lymphocytes in Cr-4 decreased, as follows from the study of the functional activity of the T-lymphocyte subpopulation. The investigation of lymphocytes - natural killers - pointed at a decline of natural cytotoxicity and its normalization by the 9th postflight day. The content of immunoglobulins A, G, M remained unaltered.

Allergological examinations demonstrated sensitization of Cr-4 to streptococcal and of FE-4 to staphylococcal agents.

As indicated earlier, these data give evidence for changes in immunological reactivity under the influence of flight effects that gradually returned to normal in the postflight period.

Microbiological investigations pointed out that in the course of flight the microbial count of the mucosa of the upper respiratory tract of both cosmonauts did not grow. In some areas the microbial count even decreased. However, during flight (in Cr-4 on MD 169 and in FE-4 on MD 54) and after flight pathogenic staphylococci belonging to the phage type 6147 phage group III and being of high toxigenic activity were found in the nasal cavity. It can be assumed that this colonization developed as a result of the transfer of pathogenic staphylococci from the Commander of the first visiting crew who exhibited the microorganisms in his oral and nasal mucosa pre- and postflight. However, the pathogenic staphylococci acquired through the exchange failed to multiply. Postflight, both crewmembers also displayed enterococci in the oral cavity which normally do not inhabit this area; this may be due to dysbacteriotic changes.

Examinations of the intestinal microflora postflight demonstrated changes in the intestinal microbiocenosis with a decrease of bifidoflora and increase in the number of conditionally pathogenic enterobacteria as well as enlargement of their species composition and appearance of *Pseudomonas* bacteria that are not typical of the intestinal microbiocenosis. On the 9th postflight day these changes weakened. After flight, the resistance of the conditionally pathogenic enterobacteria of the *E. cloacae* species to antibacterial drugs increased.

## CONCLUSION

One of the major results of scientific investigations onboard the orbital complex Salyut-6 - Soyuz is the demonstration of the fact that man can perform motivated and productive actions, including operator's function, in space flights of long duration (up to 6 months). Another important inference is that a long exposure to space flight effects brings about no changes in the health condition of the crewmembers, provided that they employ special countermeasures.

It should be emphasized that FE-4 made his 185-day flight 6 months after his 175-day flight. His health status and work capacity were essentially identical in both missions. This flight also gives evidence for the lack of stable aftereffects of the stress exposures occurring in space flights of long duration. The 6 month interval proved long enough for different functions to recover and another longer-term flight to be made possible.

The changes seen in different systems at rest and during provocative tests during and after flight were adaptive, adequate to the effects applied, and did not influence the work capacity of the crewmembers or their implementation of the flight program.

Medical investigations were concentrated on those systems of the human body which, as expected before flight, could drastically change in the course of a long-term flight.

Cardiovascular studies showed that weightlessness caused blood redistribution in different segments of the body which tended to normalize with flight time. Cardiovascular reactions at rest and during provocative tests inflight were relatively stable. However, postflight orthostatic and exercise deconditioning was observed; these changes were, nevertheless, transient and their level was similar to those that followed short-term flights. One of the most important findings of the 185-day flight was the fact that in the forearm, calf and jugular veins the venous pressure was almost identical to that inflight. These data give support to the hypothesis derived from theoretical speculations that in weightlessness the venous pressure reaches the level of the right arterial or central venous pressure and that the role of a reduced venous pressure gradient in hemodynamics decreases substantially. These results are important to gain a better insight into the hemodynamic changes in weightlessness and to prevent them properly.

The obvious consequences of the effects of weightlessness on the motor system included slight muscle mass losses, visible atrophy of leg muscles, long and wide back muscles, and decrease of the muscle tone.

Bone losses of calcium and other mineral components could become a limitation of the duration of future space missions. It was shown that after 6-month flights mineral losses from the heel bone were 3.2-8.3%, this being significantly lower than after prolonged bed rest; the small loss can be attributed to the application of various countermeasures. The values are close to mineral losses after

3-month flights. Thus, it can be concluded that active exercises may stabilize this important parameter.

Of great interest are blood studies in view of previous data that demonstrated a significant red blood cell mass loss postflight. Since the life time of red blood cells averages 120 days, they would have been completely renewed within 140-180-day flights. If weightlessness could disturb the process of red blood cell formation and development, then severe anemia would have developed.

As follows from our studies, postflight there was a slight decrease in the erythrocyte count which returned to normal within a month and a half postflight. In consistency with modern concepts, these shifts can be induced by the compensatory decrease of the circulating blood volume inflight and a much more rapid recovery of plasma than of the red blood cell count postflight. The changes in the erythrocyte structure were very moderate and did not progress with flight time. These findings allow optimistical evaluation of possibilities of blood adaptation to prolonged space flight and recovery after flight.

Thus, the medical investigations revealed no important changes in the vital systems of the human body that could prevent further increase in flight time.

Maintenance of good health condition and high performance in prolonged space flights as well as fairly rapid and undisturbed development of readaptation after the 185-day flight are a convincing indication of the success of the medical control, application of countermeasures and regular medical examinations inflight. The achievements can also be associated with an adequate work-rest cycle, nutrition, water supply and sleep.

#### ACKNOWLEDGEMENT

The authors express their sincere gratitude to Drs. O.Yu.At'kov, A.V.Beregovkin, V.S.Georgievsky, A.I.Grigoriev, S.N.Zaloguev, V.V.Kalinichenko, O.P.Kozerenko, I.V.Konstantinova, I.B.Kozlovskaya, L.N.Kornilova, V.I.Legen'kov, N.N.Liz'ko, V.I.Stepantsov, V.A. Tishler, R.A.Tigranyan, A.S.Ushakov and their colleagues who have supplied the data to be incorporated into this paper.

#### REFERENCES

- Arinchin, N.I., and G.D. Nedvetskaya (1974). Intramuscular Peripheral Heart. Minsk.
- Bevegard, B.S., and J.T. Shepherd (1965). Effect of local exercise of forearm muscles on forearm capacitance vessels. *J.Appl.Phys.*, 20, 968-974.
- Chernigovsky, V.N. (1960). Interoceptors. Moscow.
- Harrison, D.C., A. Goldblatt, E. Braunwald (1963). Studies on cardiac dimensions in intact, unanesthetized man. I. Description of techniques and their validation. II. Effects of respiration. III. Effects of muscular exercise. *Circulation Research*, 13, 448-467.

- Kalita, N.F., E.A. Pavlova, B.V. Afonin, and V.M. Ivanov (1979). Normal reaction of cosmonauts after 7-day space flights. In O.G. Gazeiko (Ed.) Aviakosm. Med. Moscow-Kaluga, Vol. 2, 193-194.
- Komarov, F.I., B.F. Korovkin, and V.V. Menshikov (1976). Clinical Biochemical Investigations. Leningrad.
- Markov, Kh.M. (1978). Prostaglandins and arterial hypertension. In Kh.M. Markov (Ed.), Prostaglandins in the Experiment and Clinic. Abstr. All-Union First Conf., Moscow, pp.26-29.
- Marshall, R.J., and J.T. Shepherd (1968). Cardiac Function in Health and Disease. W.B. Saunders Company, Philadelphia-London-Toronto.
- Parin, V.V. (1946). The Role of Lung Vessels in Reflex Regulation of Circulation. Moscow.
- Ruosteenoja, R., E. Linko, and J. Lind (1958). Heat volume changes at rest and during exercise. Acta Med.Scand., 162, 263-275.
- Tigranyan, R.A., L.L. Orloff, N.F. Kalita, N.A. Davydova, and E.A. Pavlova (1980). Changes in blood levels of several hormones, catecholamines, prostaglandins, electrolytes and cAMP in man during emotional stress. Endocrin. Exptl., 14, 2, 101-112.
- Yarullin, Kh.Kh., and T.D. Vasiliyeva (1979). Regulation of Regional Hemodynamics in Prolonged Space Flights. In O.G. Gazeiko (Ed.), Aviakosm. Med., Kaluga, pp. 10-11.

ИМЕИ МЗ СССР. Зак. № 2/94.

## U.S. PROGRAM ASSESSING NUCLEAR WASTE DISPOSAL IN SPACE: A 1981 STATUS REPORT

E. E. Rice\*, D. S. Edgecombe\*, P. R. Compton\*\*  
and R. E. Best\*\*\*

\**Space Systems and Applications Section, Battelle's Columbus Laboratories,  
Columbus, Ohio, USA*

\*\**Space Utilization Systems, NASA Headquarters Energy Systems Division,  
Washington, D.C., USA*

\*\*\**Office of Nuclear Waste Terminal Storage Program (NWTS) Integration,  
Battelle Project Management Division, Columbus, Ohio, USA*

### ABSTRACT

The importance of nuclear waste management dictates that studies of reasonable disposal alternatives be carried to the point that determinations can be made regarding their relative merit. Various concepts for the space disposal of nuclear waste have been studied by NASA. Studies sponsored by the U.S. Department of Energy, Office of Nuclear Waste Isolation (ONWI), and in coordination with NASA's Energy Systems Division are continuing to assess the desirability of a space option. The unique feature of a space option is the potential for total and permanent separation of certain wastes from the human environment. This paper provides a status report on the U.S. program and provides important insights into the current and planned investigations to determine the viability of the space disposal of selected nuclear wastes. Studies to date have indicated the economic infeasibility of disposing of spent fuel rods in space; however, space disposal is still under consideration because of its potential for complementing terrestrial disposal methods (namely, mined geologic repositories) by disposing of selected radionuclides that might be more effectively disposed of in space. The basic issue to be resolved by current and planned investigations is whether any health risk benefit can be realized with space disposal complementing terrestrial disposal. If a potential health risk benefit is identified, the next step would be to determine the cost of this benefit.

### KEYWORDS

Nuclear energy; radioactivity; nuclear waste; space disposal; Space Shuttle; international; space benefits; program planning; risk analysis; decision analysis; NASA; DOE.

### INTRODUCTION

The concept of space disposal of nuclear waste has been under investigation in the U.S. since the early 1970's (see reference list). However, because of

demonstrable technical feasibility, risk, economic, societal and other considerations, the United States has selected the mined geologic repository (MGR) as the preferred method for disposing of high-level nuclear wastes (U.S. DOE, 1981). In making this selection, the U.S. Department of Energy gave consideration to a number of alternative technologies, including space disposal. Each of the alternatives was evaluated against waste management objectives. However, studies of space disposal are being continued to assess its potential for complementing the mined geologic repository by disposing of selected radionuclides in space and thereby enhancing the overall performance of the U.S. nuclear waste management system.

The basic issue to be resolved by current U.S. investigations is whether any health risk benefit can be realized with space disposal as a complement to terrestrial disposal. If a potential health risk benefit is identified, the next step would be to determine the cost of this benefit. This paper discusses the current U.S. investigations to further evaluate the viability of disposing of selected nuclear wastes in space.

Most of the recent investigations for space disposal have been sponsored by the U.S. Department of Energy (DOE) through the Office of Nuclear Waste Isolation (ONWI), in cooperation with the National Aeronautics and Space Administration (NASA). Studies of the space systems considered for space disposal have been coordinated by NASA Headquarters' Energy Systems Division.

#### CONCEPT SUMMARY

Space disposal offers the option of permanently removing part of the nuclear wastes from the Earth's environment. In this concept the waste of interest would be packaged in a special flight container for insertion into a solar orbit, where it would remain for at least 1 million years (Friedlander and Davis, 1979).

NASA has studied many space disposal options since the early 1970's (e.g., Burns, 1978). A reference concept using an Upated Space Shuttle utilizing reusable liquid rocket boosters (LRBs) has been defined (Rice, et al, 1980a and 1980b). In this concept, the Upated Space Shuttle would carry the waste package to a low-Earth orbit. An orbit transfer vehicle would then separate from the Shuttle Orbiter to place the waste package and another propulsion stage into an Earth escape trajectory. The transfer vehicle would return to the Shuttle Orbiter, while the remaining propulsive stage would insert the waste into solar orbit. Figure 1 shows (1980 Reference Concept) the orbital operations required to place the waste package in a 0.85 AU heliocentric orbit, the currently preferred space destination. This figure also shows the associated ground operations for disposal of high-level waste.\*

To place the space disposal concept into perspective from a total system viewpoint, Fig. 2 shows the location and process flow details of the space disposal of high-level waste within the total waste management system. Two points are apparent from this figure: (1) chemical processing would definitely be required

---

\*DOE defines high-level waste to include high-level liquid wastes, products from solidification of high-level liquid waste, and irradiated fuel elements, if discarded without reprocessing. A proposed U.S. Nuclear Regulatory Commission (NRC) regulation (10 CFR 60.3) defines high-level waste to include irradiated fuel, high-level liquid waste, and products from its solidification.

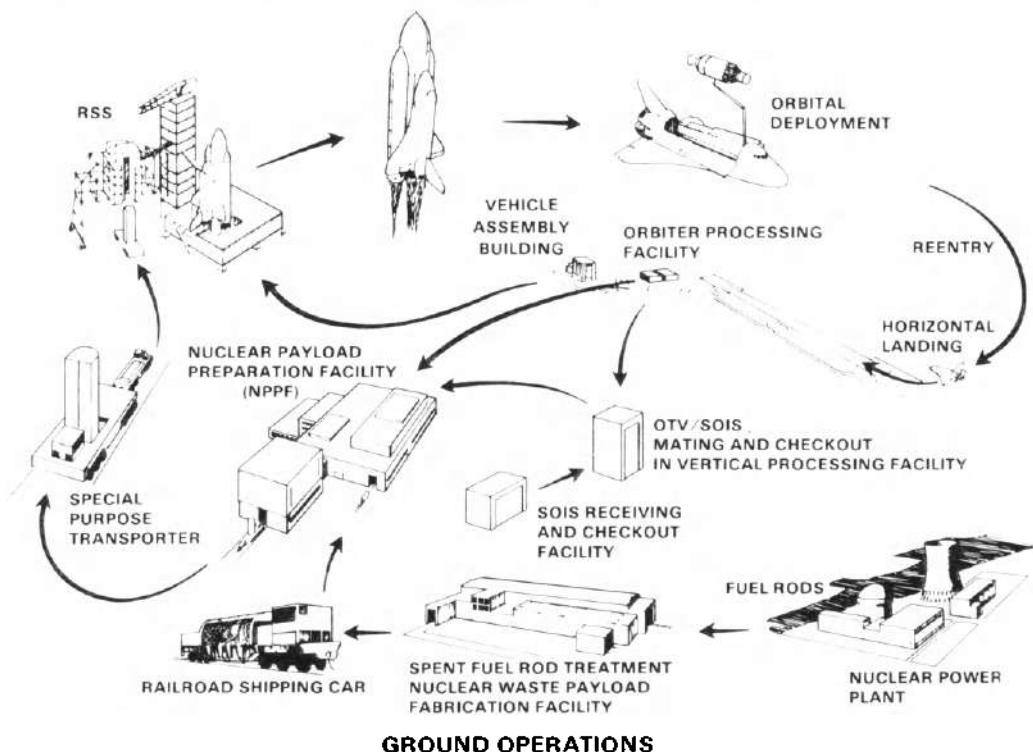
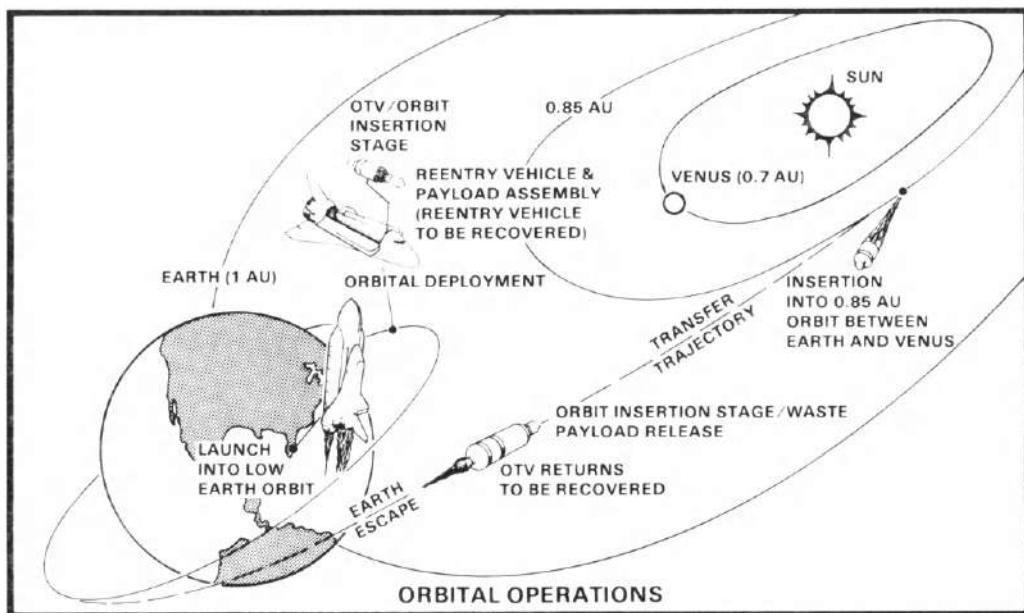


Fig. 1. Ground and space operations for 1980 reference space disposal mission.

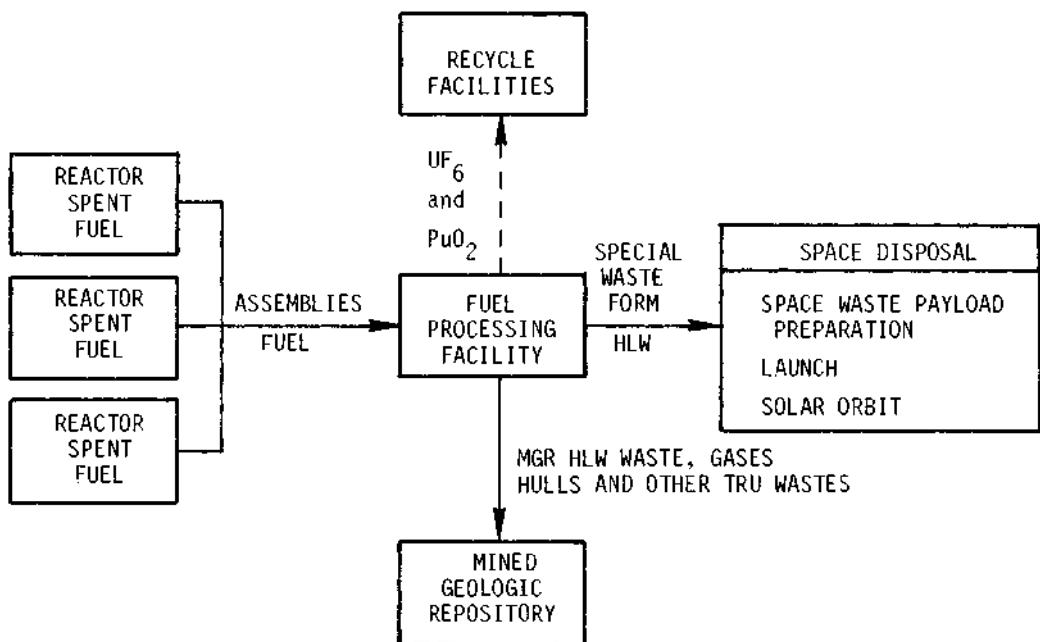


Fig. 2. Fuel cycle diagram with space disposal complement.

for space disposal of waste, and (2) mined geologic repositories would be part of the total system.

Chemical processing is necessary because space disposal of unrepurposed spent fuel rods would result in an excessive number of launches (Pardue, et al, 1977). This would result in high energy requirements (perhaps 25% of the energy originally produced by the nuclear reactors), high costs, and increased environmental effects. However, for high-level waste, a much lower number of Space Shuttle flights would be required. Approximately one launch per week to dispose of all high-level waste from 5000 MT of heavy metal resulting from annual operations of approximately 200 GWe nuclear power capacity is projected (Yates and Park, 1979). The remaining transuranic (TRU) contaminated process wastes and gases would still require terrestrial disposal. The mined geologic repository is the preferred method for disposing of TRU waste.

Of interest at this time, however, is the possibility that the space option could be used as a complement to rid the Earth of selected fractions of radioactive wastes, thereby optimizing benefits to the total nuclear waste management system. These selected fractions might include long-lived isotopes, such as the actinides, iodine and technetium. Space disposal might offer risk benefits for these long-lived wastes because it is extremely remote that physical forces would cause the disposed radioactive waste to migrate to the Earth before the material has decayed. Consequently, for a package properly placed in solar orbit, there would be virtually no long-term risk (Rice, et al, 1980a) or surveillance requirements, as in terrestrial disposal. However, the health risk of launch pad accidents or low-Earth orbit failures must be evaluated and weighed against the long-term risk of release from mined geologic repositories. The studies being carried out in 1981 under the direction of the Office of NWTS Integration and NASA's Marshall Space Flight Center to help resolve these issues of risk benefit are described below.

#### CURRENT STUDIES

Current and near-term planned studies are concerned with the following:

- (1) Identifying candidate waste mixes and forms for space disposal, based on benefits to the terrestrial disposal system risk
- (2) Selection of the space transportation system for each candidate waste mix
- (3) Developing a preliminary space systems risk assessment to aid in assessing the overall risk benefit of space disposal.

The interrelationships of these programs and the activities to which they will supply inputs are shown in Fig. 3. These studies are expected to provide the following:

- Evaluation of the benefits to terrestrial disposal in mined geologic repositories by removing selected waste fractions for disposal or elimination by other means. These studies will consider state-of-the-art or near-term process technology and will address radiation exposure risks and benefits.

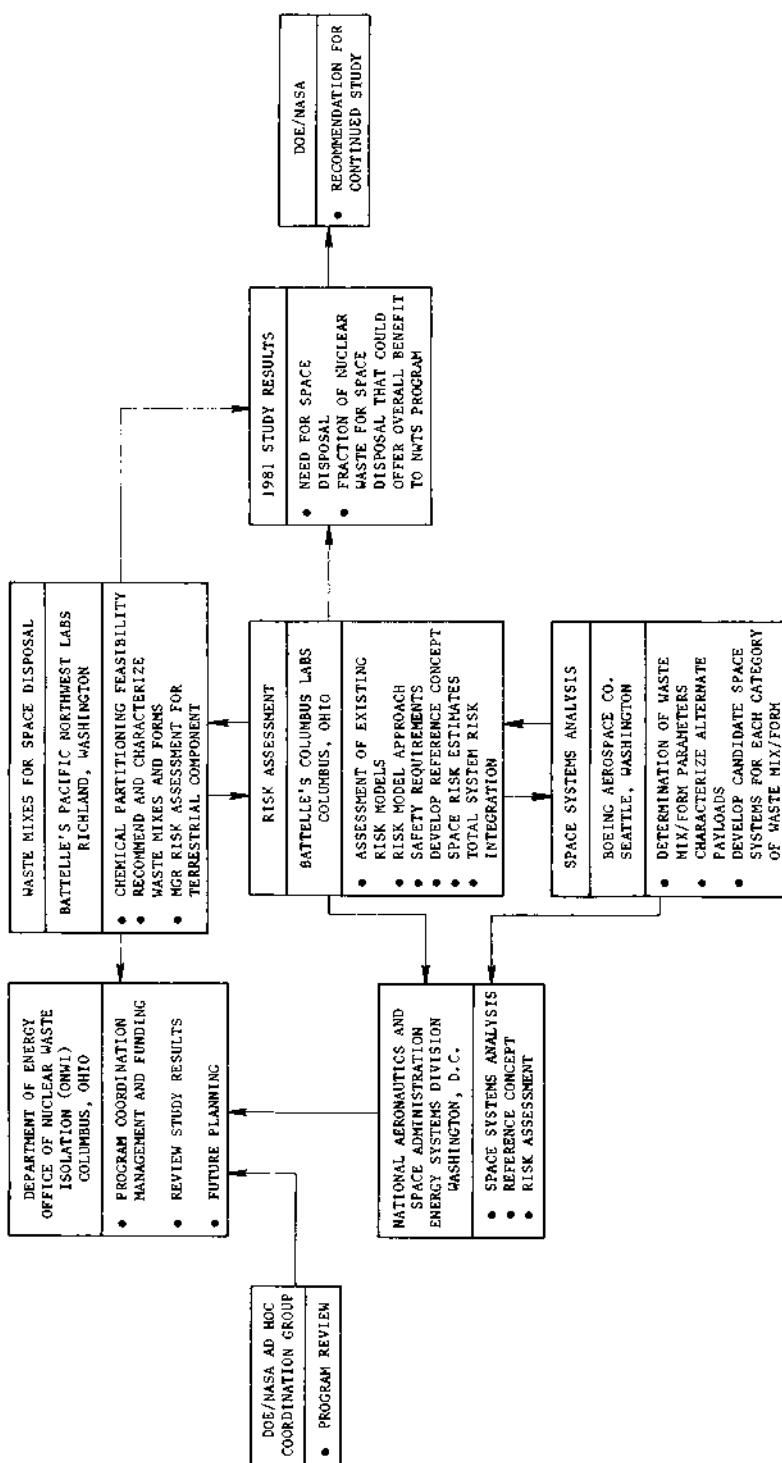


Fig. 3. 1981 study activity for U.S. nuclear waste disposal in space program.

- Identification of the space systems required for use in disposing of possible waste mixes other than the total high-level wastes from the Purex process
- Evaluation of the risks of placement in space of selected waste mixes. The terrestrial and space disposal risks will be integrated to prepare an estimate of total system risk for a selected group of waste mixes (for space disposal). Based on the results, a recommendation for a waste mix to be used as a reference space disposal concept will be made.

The following discussion provides more detail regarding the three 1981 study areas.

#### Waste Mixes

The objective of this study is to identify U.S. commercial nuclear waste components that, when removed from wastes to be disposed of in mined geologic repositories and disposed of by space disposal or other means, would have a potential to result in reduced risk for nuclear waste disposal.

To do this, theoretical long-term impacts/benefits to mined repository performance that will result from removal of candidate isotopes from commercial wastes have been identified. Consideration has been given to toxicity, half-lives, pathways, and repository failure scenarios. Also of interest is the identification of the capability of chemical process technology to remove selected species. The side waste streams and effluents that would result, for different degrees of recovery, are to be estimated, as well additional waste species that would be carried with the primary species of interest. The characteristics of the output streams (waste mixes) for space disposal are to be estimated and related to possible waste forms. Examples of waste mixes and forms being considered are shown in Fig. 4.

After consideration of the above and other significant factors, this study will recommend and characterize waste mixes for commercial waste components, that if disposed of by space disposal, would offer benefits to long-term risk.

This study program is being conducted by the Battelle's Northwest Laboratories (BNWL) of Battelle Memorial Institute, Columbus, Ohio.

#### Space Transportation System Concepts

Space transportation systems concepts are being evaluated by Boeing Aerospace Company, Seattle, Washington, under subcontract to NASA. This activity concerned with selection of the space transportation systems concepts for use in disposal of the waste mixes and forms defined by BNWL. The space transportation system study is providing the following:

- (1) Identification of waste mix/form parameters that are important to the definition of the candidate space systems concepts and the establishment of the range of relevant parameter values that bound the possible waste mixes and forms (payloads) to be used for space disposal. Determination of categories of waste payloads for use in parametric analyses of the space systems.

<u>MGR DISPOSAL WASTE MIX</u>	<u>SPACE DISPOSAL WASTE MIX</u>	<u>LIKELY SPACE DISPOSAL WASTE FORMS</u>
1. ALL TRU, HLW, GASES	NONE	- -
2. ALL TRU, HLW, GASES, LESS I <sup>139</sup>	IODINE - 129	Ba(I <sub>03</sub> ) <sub>2</sub> , PbI <sub>2</sub>
3. ALL TRU, HLW, GASES, LESS Tc <sup>99</sup>	TECHNETIUM - 99	Tc METAL, TcO <sub>2</sub>
4. ALL TRU, GASES, 95 PERCENT OF Cs <sup>137</sup> AND Sr <sup>90</sup>	ALL HLW, PLUS 5 PERCENT Cs <sup>137</sup> AND Sr <sup>90</sup>	CERMET
5. ALL TRU, GASES	ALL HLW (DECAYED)	CERMET

Fig. 4. Waste mix options being considered for space disposal complementing disposal in MGR.

- (2) Parametric systems analyses to determine changes in system requirements, system designs and mission operations that would result from changes in the characteristics of the waste payload. Emphasis will be given to waste payload protection system concepts and space transportation system alternatives for the 0.85 AU disposal destination.
- (3) Recommendation of a candidate space system for each category of waste payload.

#### Space Systems Risk Assessment

This study is being performed by the Columbus Laboratories of Battelle Memorial Institute, Columbus, Ohio, and is concerned with developing preliminary quantitative risk data for disposing of waste in space. This activity follows work previously done by Rice, et al (1980a), Edgecombe, et al (1979), and Pardue, et al (1977). Integration requirements for coupling nuclear waste systems risk assessments (terrestrial disposal) with space system risk assessments are to be defined to allow proper comparison.

During 1981, the specific activities will be to:

- Identify and review existing risk model approaches and data bases that exist in the literature
- Formulate a preliminary space systems risk model approach
- Improve safety requirements and guidelines previously defined
- Develop a new Reference Concept for Disposal
- Prepare estimates of space system long- and short-term risk and risk uncertainty for selected waste payload (mix/form) categories, as developed in Boeing's 1981 space system studies
- Prepare preliminary quantitative estimates and uncertainty of total waste disposal system risk for selected waste mixes/forms, including the mined geologic repository and space system components
- Recommend a waste mix/form for use as a reference in continued studies.

#### POSSIBLE STUDY RESULTS

Possible study results of the 1981 studies are:

- (1) Space disposal is demonstrated to offer an expected net risk advantage when used as a complement to terrestrial disposal for nuclear wastes
- (2) Space disposal is demonstrated to offer no expected net risk advantage or disadvantage
- (3) Space disposal is demonstrated to offer an expected net risk disadvantage

- (4) Results of 1981 studies are uncertain to the extent that no clear advantage, disadvantage, or other expected risk relative to terrestrial disposal can be determined.

Each of the possible study results will have an associated uncertainty. For this reason, the term "expected" is used to describe the relative position of the most likely value of difference in risk between terrestrial disposal in mined geologic repositories and terrestrial disposal in mined geologic repositories complemented by space disposal.

#### DECISION ALTERNATIVES

Given the above possible study results of 1981 studies, the possible decisions regarding the DOE funded space disposal program are:

- (1) If positive risk benefits are projected, research on space disposal will likely be continued
- (2) If no risk benefit is projected, space disposal related technology will likely be tracked by the U.S. DOE
- (3) If negative risk benefits are projected, the U.S. DOE funded space disposal program will be discontinued
- (4) If the results are indeterminant, space disposal related technology will likely be tracked by the U.S. DOE.

#### SUMMARY

In summary, various space disposal concepts have been developed over a number of years by NASA and its contractors. Significant recent efforts by NASA, Battelle's Columbus Laboratories and the Boeing Aerospace Company have provided detailed definition to the space disposal concept. Investigations being carried out in 1981 will provide a basis for estimating the total systems health risk benefit, if space disposal is employed as a complement to mined geologic repositories for nuclear waste disposal. If significant health risk benefits are identified, a basis will have been established for continued investigations of space disposal of certain types of nuclear wastes.

## REFERENCES

- Burns, R. E. (1978). Nuclear waste disposal in space. NASA-TP-1225, NASA/ Marshall Space Flight Center, Huntsville, Alabama. USA.
- Edgecombe, D. S., E. E. Rice, N. E. Miller, K. R. Yates, and R. J. Conlon (1979). Evaluation of the space disposal of defense nuclear waste - phase II. NASA CR-161101, 161102, and 161103, Battelle's Columbus Laboratories, Columbus, Ohio. USA.
- Friedlander, A. L., and D. R. Davis (1979). Long-term risk analysis associated with nuclear waste disposal in space. AAS/AIAA Astrodynamics Specialists Conference, Paper No. 79-175, Provincetown, Massachusetts. USA.
- Pardue, W. M., E. E. Rice, N. E. Miller, and D. K. Davis (1977). Preliminary evaluation of the space disposal of nuclear waste. BCL-8-32391(100), NASA CR-150440, Battelle Memorial Institute, Columbus, Ohio. USA.
- Priest, C. C., R. F. Nixon, and E. E. Rice (1980). Space disposal of nuclear wastes. AIAA-Astronautics and Aeronautics, April 1980, 26-35.
- Rice, E. E. (1978). Energy impact assessment of NASA's past, present, and future space launch vehicles. AIAA-Journal of Energy, Volume 2, No. 3, 182-188.
- Rice, E. E., N. E. Miller, K. R. Yates, W. E. Martin, and A. L. Friedlander (1980a). Analysis of nuclear waste disposal in space - phase III. NASA CR 161418 and 161419, Battelle's Columbus Laboratories, Columbus, Ohio. USA.
- Rice, E. E., C. C. Priest, and A. L. Friedlander (1980b). U.S. program assessing nuclear waste disposal in space: a status report. 80-IAA-50, IAF XXXI Congress, Tokyo, Japan, September 21-28, 1980.
- U.S. DOE (1981). Program of research and development for management of disposal of commercially generated radioactive wastes; record of decision. U.S. Federal Register, Vol. 46, NO. 93, May 14, 1981, 26677.
- Yates, K. R. and U. Y. Park (1979). Projection of commercial nuclear capacity and spent fuel accumulation in the United States. Transactions of American Nuclear Society, June 1979, 350-352.

## CURRENT TOPICS OF SPS REALIZATION FROM A EUROPEAN VIEWPOINT

W. Westphal

*AEG-Telefunken, Wedel, Federal Republic of Germany*

### ABSTRACT

Although having yet a relatively high standard of investigation and software analysis about SPS by several dedicated studies and by the "Concept Development and Evaluation Program" from NASA and the US-DOE, there is no forcing policy at the moment to pursue research and development to potentially meeting the expected demand threshold just behind the turn of the century.

But apart from actual energy policies, especially in Europe there is undoubtedly an urgent need to create back-up options for the coming energy supply transition and to lead them up to a higher level of judgement.

The very limited number of presently identified alternatives should justify an increasing European interest in SPS.

Because Europe's potential of SPS application is critically imposed by the rectenna area requirements of a fully deployed system and by a variety of big volumed advancement needs for reasonable contributions in development and implementation, it can take advantage by own efforts only.

So, to uphold a minimum optional status of SPS for Europe, activities on scientific and basic technological research should be started now. These should currently focus on better solutions for the power transmission from space to evolve basic feasibility conditions for Europe by significantly reduced rectenna sizes, and on strategic solar generator development to meet the high requirements for SPS production participation on a leading position.

### KEYWORDS

Satellite Power System / Space Policy / Energy Policy /  
Photovoltaic Space Generator

## INTRODUCTION

The investigation of satellite-based energy supply systems in the U. S. during the recent years has been focussed in 1980 on the Program Review of the SPS "Concept Development and Evaluation Program". Since then, an adversary discussion has been started and support for further research has been reduced stringently.

European efforts on SPS investigations, which have been on a much lower extent, mainly dealing with evaluations of special domestic conditions, are nearly dropped at this time.

The main reason for this reserve seems to be a certain confusion arising from the so-called SPS-Reference System Concept, which misleadingly has been understood widely as an engineering concept. In particular, its bold dimensions in size and cost, featuring a hypothetical but nevertheless technologically conservative system design, generated an understandable reluctance. The most issues of scepticism have been grown out of the fact that the currently deducible but also partially very conservative assumptions of the reference concept led to highly complex operations and huge configurations of all the main components. That, in turn, created high uncertainty of the cost estimations and, of course, of the performance targets which have to be acquired. Especially the early fixing of the parameters of the microwave power transmission system postulated an economically viable order of magnitude, which is far beyond the envisaged space opportunities deployment for the next two decades.

Nevertheless, this energy supply option should deserve far more interest and support in Europe simply because of the fact that our present general reliance on fossil primary energy is threatened by exhaustion due to still growing consumption rates and a scarce backup of own indigenous resources.

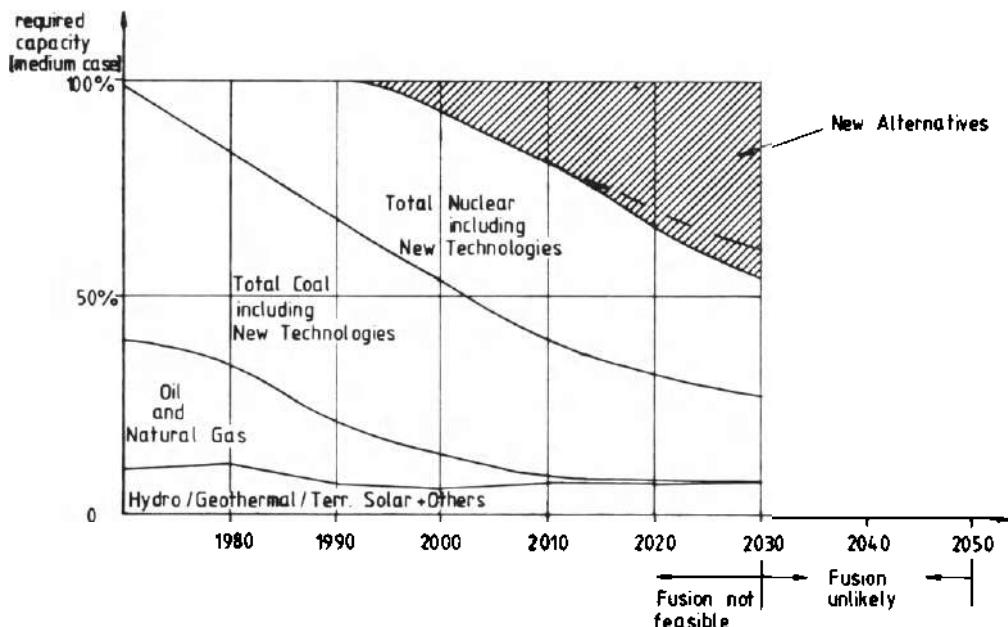
Moreover, there is a lack of soothng number of alternatives for the coming energy transition need which possibly are required early in the next century.

In the following, the possible European energy transition need will be depicted, reasonably urging to uphold the optional status of SPS. Therefore, at least both present topics should be pursued : Scientific research on the special European earth/space interface conditions and at reduced rectenna site areas and commencement on systematical technology research for future space solar generator developments.

### POTENTIAL EUROPEAN INTEREST IN SPS

In several recent studies /1; 2/, the future energy supply needs for Western Europe have been analyzed, applying different technology scenarios of basic developments and sets of political and economical constraints. Generally, there have been considered modifications of current energy options in quality - for example : Coal liquefaction and coal gasification - and extending them by some new options like "nuclear fission fast breeder reactors", and to a very small extent, so-called "renewables", (which are: wind, waves, tidal, geothermal and terrestrial solar options) (see Fig. 1).

But at this time, no real alternatives have been predicted that can reach baseload capability. Especially the European lack of considerable indigenous energy resources will possibly effect a shortfall starting early in the 1990s resulting from the scissors of growing consumption needs and decreasing availability of fossil primary energy.



**Fig.1: Summary of the predicted demand structure for the EC electricity supply**

The following rationale, infencing the European forecast situation, can be drawn from the above quoted studies /1; 2/ :

- the fusion technology will emerge beyond or far beyond 2030 if success can be assured (Fig. 1)
- hydro, ground solar, geothermal and other renewable resources are unlikely to supply more than 5-6% by the year 2000 (Fig. 1)
- between 2000 and 2030, the severe decrease of oil supply cannot be substituted by existing alternatives and their evolving new technologies only, but a transition gap must be filled from about 10% up to 40-45% (Fig. 1), whilst the existing oil and alternative supply will decrease from about 40% to 15% (not shown)
- "A common denominator to nuclear technology, renewable technology and fusion is the production of electricity. This indicates that a flexible response to these problems must of necessity result in a society which is largely based upon a comprehensive electrical network" /1/
- "There is an urgent need to implement new RD & D policies now in order to provide the Community with the political flexibility necessary for the successful transition to the post-oil society of the next century" /1/

In the conclusion of the quoted studies it is demonstrated that even if the most dynamic utilization of EC-indigenous resources takes place, and even if realistic nuclear growth rates of breeder technology are assumed, a reasonable balance can only be met by alternative sources of new technology.

It is a clear outcome of the above that the SPS concept represents a rather suitable option to fill the European "new non-fossil technology" gap for several reasons :

- The state of knowledge about SPS is already on a relatively high degree of maturity among all non-realized energy options of large capacity and future technology.
- This knowledge indicates no unsurmountable technological hurdles so far, although any potentially preventing items have been investigated at high effort compared to alternative energy systems at that early or even higher developed state.

- This knowledge even indicates compatibility in time and in capacity to fill the analyzed gap beyond the turn of the century.
- Fusion as well as fission energy options of any kind are seen to be no real concurrency but a complement in the 2000 - 2050 regime (or vice versa, respectively).
- The required implementation ratio of SPS should be manageable
  - especially if other suitable energy options relax the energy investment constraints.
- Utilizing a "cost-free" primary energy, the SPS potentially can substitute fuels by delivering electricity economically for hydrogen production. This potential typically cannot be met by resource fuel-fed electricity power plants.

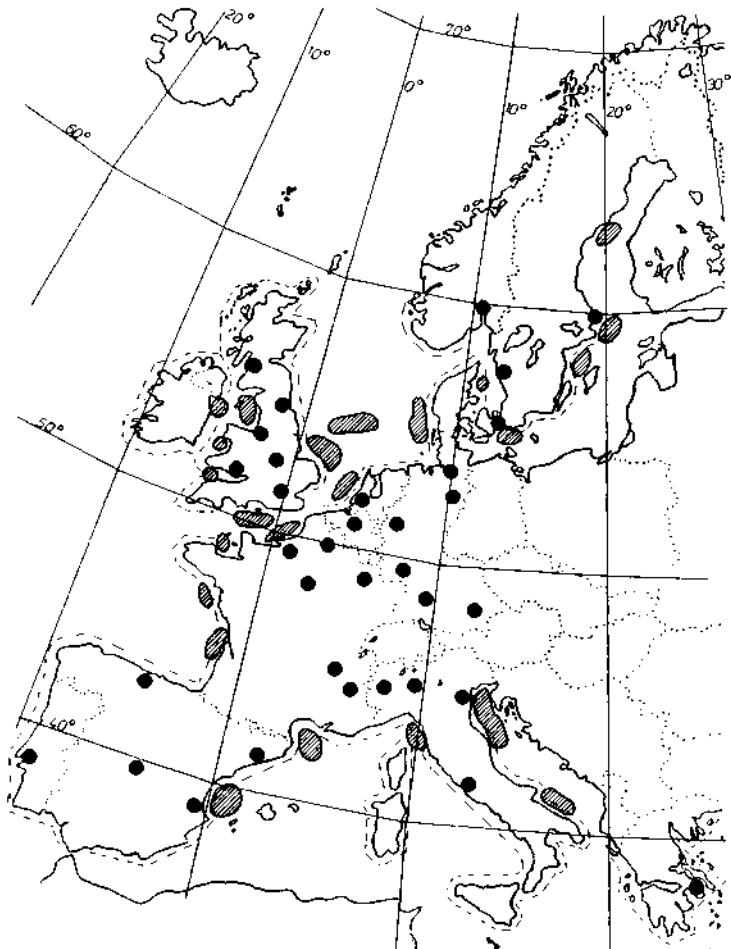
#### CURRENT EUROPEAN PRIORITIES DIRECTED TOWARDS SPS

##### A. Basic Research on Present Major Problems

The most severe difference between European and the U.S. application of SPS results from the bold dimensions of the rectenna, which are required for the microwave power transmission system of the reference concept.

It is clear that Western Europe does not dispose of sufficient land areas for the placement of about 40 5-GW reference rectennas meaning at least 50,000 - 60,000 km<sup>2</sup> in total due to European latitudes (including safety areas /3/).

Thus, a recent study intended by ESA /4/ was pointed towards the European off-shore siting potential (see Fig. 2). A variety of off-shore constructions was examined out of which a polder construction has been analyzed due to cost and requirements on shallow coastal regions. (A polder is a ring dike with dewatered inner area by a permanent drainage system).



● Identified electrical consumption areas at and above 3GW currently

Fig. 2 : Identified Potential for Off-Shore Rectenna Sites

The total construction cost of such one off-shore rectenna has been estimated to accrue to about  $5 \times 10^9$  \$ (1980). So this could mean about 1.4 to 1.8 \$/kWh delivered regarding the conservative reference conditions and a 30-year-life-time of the rectenna.

Although these issues could become competitive under certain conditions, the credibility of SPS for European application is faced by a thorough review of the concept which should reduce economically viable the presently unacceptable ground station dimensions.

Several basic research efforts in this area which, if started now on relatively low budgets, can help to assure the further optional status of SPS for Europe. Engaging scientific institutes, the many basic options for the microwave power transmission should be investigated incorporating higher carrier frequencies. 5.8 GHz could meet reasonable attenuation limits and allow the receiving area to be reduced by a factor of 4 compared to the reference system. Scaling down the power per beam (possibly 2 to below 0.7 GW) can be economically met by solid state microwave amplifiers being on a very preliminary state of investigation. Multiple transmitting antennas per satellite, which avoid overlap of the beams in the ionosphere allow for higher power and/or smaller size per rectenna. Also, the application of laser power transmission, which could reduce the rectenna area dramatically, is on a very early state of research. Finally, the reference system limit of microwaves on  $23 \text{ mW/cm}^2$  for the central beam intensity, respecting ionosphere interactions, can possibly be enhanced due to recent investigations. That may allow for higher power levels per beam or reduced rectenna areas, too.

TABLE 1

Issues for potential reduction of rectenna dimensions

- Application of higher microwave beam carrier frequency (5.8 GHz)
- Scaling down of transmission power (2 to 3 GW)
- Using multiple transmitter beams per satellite
- Application of laser power transmission
- Increasing the  $23 \text{ mW/cm}^2$  -limit for ionosphere

### B. Advancement Strategies on Relevant Technologies

Considering the space segment of a future SPS system, it is clear that development and implementation need new dimensions of economic efforts and international interactions. Therefore, Europe's competence is requested as a full-scale contributor, if SPS realization will be commencing. Reasonably, the most favorable converter system will be the photovoltaic solar array, meaning the biggest space segment subsystem (> 50%).

Since the early steps into space in the 1960ies, a variety of European contributions in space projects have evolved a very high technology standard concerning the space-oriented solar generator. These contributions, not lastly from AEG-TELEFUNKEN, have assured a firm leading position on several state-of-art issues. Even continuous efforts oriented towards low cost photovoltaic devices, based on mass production technologies, which are specifically qualified for terrestrial applications, let European technology be seen on a leading position.

Nevertheless, the challenge to conserve SPS optional status regarding the converter as its main determinant calls for new development strategies, which incorporate several identified basic development directions already in the near-term generator projects (Fig. 3).

Especially the mentioned dual experience of AEG concerning mass production and high efficiency technology of photovoltaic devices offers an advancement potential directed to high qualified low cost generators for future space applications, regarding significantly growing power levels (Fig. 3).

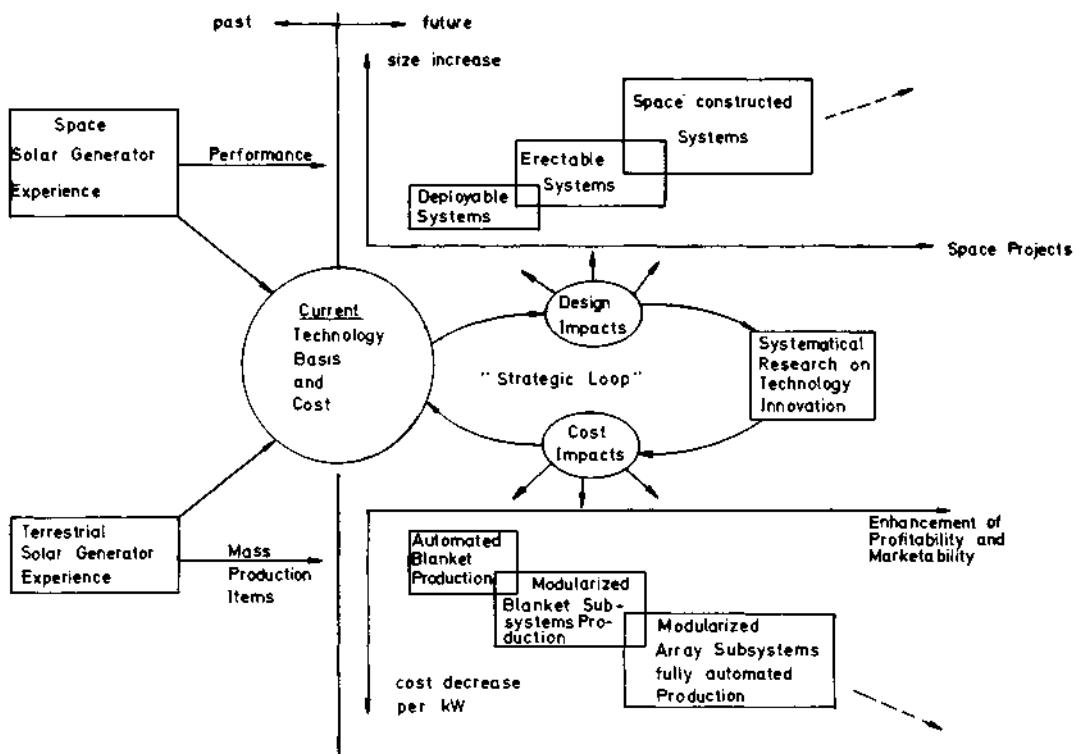


Fig.3: Evolution process on future space solar arrays

In Table 2, several objectives are summarizing currently envisaged directions and tasks of space solar generator development also relevant for a SPS policy in Europe :

TABLE 2

Solar array development directions relevant for optional SPS technology are to:

- Increase modularity of panels/arrays
- Increase standardisation of mechanical interfaces
- Develop high voltage blanket-technology
- Decrease system sensitivity on thermal loading
- Improve cost compromise between solar cell/panel performance and mass fabrication technology
- Provide evolution of array structure for growing system sizes
- Minimize development needs between sequent projects

#### CONCLUSION

The lack of sufficient amounts of indigenous primary energy resources is facing Europe with a potential supply shortfall possibly before the year 2000. Compared to the U.S., this threat is very close. So the SPS application for Europe (and, for example, for Japan) should logically deserve much more interest than for the U.S. A reasonable energy policy should at least provide the optional status of SPS, which could be acquired presently at relatively small budgeting (Fig. 4).

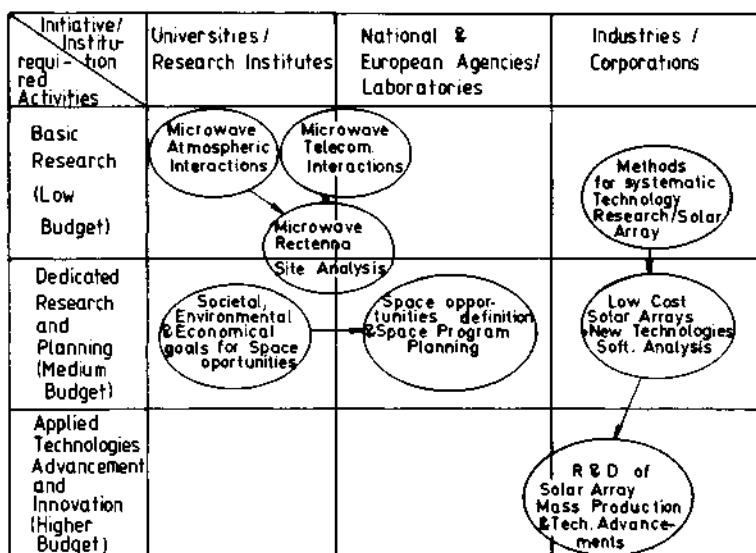


Fig: 4 Near-term efforts required to uphold SPS optional status

For the next 5 to 10 years, this status can be held by the intensive scientific research on European rectenna conditions on one hand and on the other by the continuous technology research affecting the required photovoltaic space and terrestrial devices deployment.

More critical is the necessary decision preparation for the future European space opportunities planning, which mainly deals with cooperation between Europe and the U.S.. There is an additional challenge for Europe to transit from a "junior partner" to a "full-scale partner" of the U.S. as space is turning into a big market.

## REFERENCES

1. Crucial Choices for the Energy Transition,  
Commission of the European Communities  
Report EUR 6610 (1980)
2. The Solar Satellite Power System as a Future  
European Energy Source  
H. Stoewer, D. Kassing, K. K. Reinhartz  
ESTeC
3. Study on European Aspects of Solar Power  
Satellites, Final Report  
J. Ruth, W. Westphal, TU Berlin (D)  
ESA Contract Report 3705/78/F/DK (SC) (1979)
4. Study on Infrastructure Considerations for  
SPS-Offshore Rectennas in W. Europe,  
Final Report  
Hydronamic B. V., Sliedrecht (NL)  
ESTeC Contract No. 4382/80/NL/PP (SC) (1980)

## VOYAGER AT SATURN\*

R. L. Heacock

*Jet Propulsion Laboratory, 4800 Oak Grove Drive,  
Pasadena, California 91109, USA*

### ABSTRACT

Two Voyager spacecraft were launched in August and September of 1977 on scientific exploration missions of the Jupiter and Saturn systems and the intervening interplanetary space. The spacecraft encounters with Jupiter were in March and July of 1979. The spacecraft returned spectacular photographs and other scientific data which revealed several fascinating discoveries. The tremendous volume of data from the Jupiter encounter will require years of study. The Saturn encounters were in November 1980 for Voyager 1 and August 1981 for Voyager 2. The data returned from the Saturn encounters proved to be equally spectacular with unique new data about the Saturn system.

An overview of the Voyager Project is provided with emphasis on the key features of the spacecraft design and the in-flight operations which contributed most to the outstanding success of the mission. The key scientific discoveries will be highlighted with emphasis on the Voyager Saturn results. The new findings from the Voyager 2 August 25, 1981 encounter will be presented.

The Voyager spacecraft missions continue and are described. Voyager 1 is continuing to make fields and particle measurements on its flight toward interstellar space. Barring in-flight failures, Voyager 2's trajectory past Saturn will provide an encounter with Uranus in January 1986 and an encounter with Neptune in August 1989 before it also heads into interstellar space.

### KEYWORDS

Space exploration; outer solar system; Jupiter; Saturn; Voyager; spacecraft; science instruments; mission operation.

\*The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

## INTRODUCTION

When the Voyager 2 spacecraft completes its post-encounter observations of Saturn at the end of September 1981, the Voyager Project will have completed its primary mission of exploring the Jupiter and Saturn systems. The Project has experienced outstanding in-flight success with the Voyager spacecraft and their operation including the required precision navigation and science instrument observations. The photographs of the Jupiter and Saturn systems have been particularly well received because of their high quality, near real time availability and the fascinating character and sheer beauty of the planets, satellites and rings.

The Voyager Project is carried out for the National Aeronautics and Space Administration (NASA) by the California Institute of Technology's Jet Propulsion Laboratory (JPL). Many thousands of people have contributed to the success of Voyager through the design, development, and test of the required systems and through support of the flight operations. The Deep Space Network with receiving stations in Australia, Spain, and California has been essential to the recovery of the high volume of science data which has been made available for analysis by approximately one hundred scientists who support the Project.

## BACKGROUND

Mission studies into the exploration of the outer solar system were initiated in the mid-1960's with the discovery that the outer planets (Jupiter, Saturn, Uranus and Neptune) would be uniquely aligned in the late 1970's and 1980's such that a single spacecraft launched in 1977 or 1978 could explore all four of these large planetary systems. This so-called "Outer Planets Grand Tour" requires a "gravity-assist" trajectory in which the gravity field of one planet is used to couple some of that planet's orbital velocity to the spacecraft and thereby accelerate it on to the next planet in the sequence. This unique alignment of the outer planets occurs approximately every 175 years.

In 1968, NASA initiated at JPL an advanced technology project to investigate the detailed requirements for an outer planets mission. The TOPS (Thermo-electric Outer Planets Spacecraft) Project (1) investigated detailed mission design constraints, science investigation requirements and spacecraft design and implementing technology requirements. In addition in 1968, NASA initiated the Pioneer 10 and 11 spacecraft program (2) for a preliminary exploration mission to Jupiter. The Pioneer 10 and 11 spacecraft were very successful and provided invaluable data to the Voyager Project on the Asteroid Belt hazard and the Jovian radiation environment. Pioneer 11 was able to use a gravity assist flyby of Jupiter to go to Saturn for a very successful encounter which provided excellent science results and invaluable pathfinder data on the ring hazards for Voyager 2's flyby of Saturn on its Uranus trajectory.

In 1970, NASA initiated the Outer Planets Mission (OPM) Project at JPL in order to proceed with the detailed planning and costing of the exploration of the outer solar system. The key people and concepts from the TOPS Project were utilized on the OPM Project. The Project costs were estimated at \$750 million in order to achieve the long mission life (over ten years) and satisfy the support requirements of the sophisticated science investigations. This high cost and the competition for unmanned exploration funds resulted in NASA cancelling the OPM Project in December 1971. Fortunately, the Project team working with NASA and the science community was able to develop a lower cost, limited scope mission called the Mariner Jupiter/Saturn 1977 (MJS 77) Project. The focus on just Jupiter and Saturn shortened the mission lifetime requirement to four years thus permitting the use

of earlier Mariner and Viking Orbiter based technology and hardware designs.

The MJS 77 Project name was changed to Voyager in March 1977 in order to provide the Project its own, separate identity.

#### MISSION DESIGN (3)

The Jupiter and Saturn systems presented significant challenges to the Voyager mission designers and science investigators due to the large number of separate targets and types of observations of interest to the science investigators. Jupiter and Saturn are large planets with immense magnetospheres, large numbers of satellites and in the case of Saturn, a complex ring system. In order to select the best flight trajectories for the mission, it was necessary to "fly" thousands of trajectories in the computers and analyze their specific characteristics in terms of their potential science return. The most favorable trajectories were then interacted with the science investigators in making the final selection of the trajectories shown in Fig. 1.

Voyager 2 was launched first on August 20, 1977 on a lower energy trajectory which would provide the "Outer Planet Grand Tour" trajectory to Jupiter, Saturn, Uranus and Neptune. The trajectory is characterized by a relatively distant flyby of Jupiter (722,000 km) with several good satellite encounters before Jupiter closest approach, a relatively close flyby of Saturn (161,000 km) and several of its satellites, a flyby of Uranus with a close encounter with the satellite Miranda and a close flyby over the north pole of Neptune in order to obtain a close encounter with the satellite Triton after Neptune closest approach.

With the successful launch of Voyager 2, Voyager 1 was launched on a different, higher energy trajectory on September 5, 1977. Voyager 1 passed Voyager 2 in December 1977. Its trajectory is characterized by a closer flyby of Jupiter (374,000 km) in order to achieve a close flyby of Io (21,000 km) including flying through the Io flux tube, several good satellite encounters after Jupiter closest approach, a close flyby of Saturn's satellite Titan (7,000 km), a close flyby of Saturn (184,000 km) and several of its satellites, and radio signal occultations of Titan, Saturn, and Saturn's rings. Diagrams of the Voyager 1 and Voyager 2 Saturn trajectories are shown in Figs. 2 and 3. The Titan III E Centaur D-IT and a small solid rocket kick stage integrated into the spacecraft were necessary to provide the launch energies required by the Voyager trajectories.

#### SPACECRAFT (4)

The Voyager spacecraft made use of earlier Mariner and Viking Orbiter concepts and hardware designs, but several new capabilities were required to satisfy the mission requirements. The exploration of the outer solar system required the use

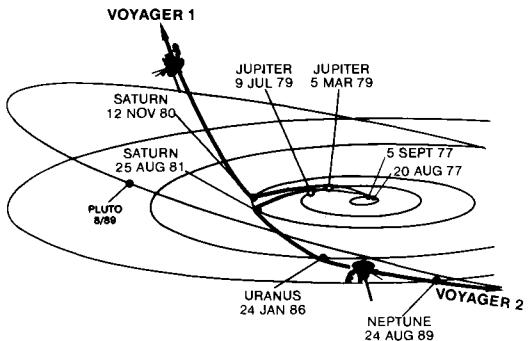
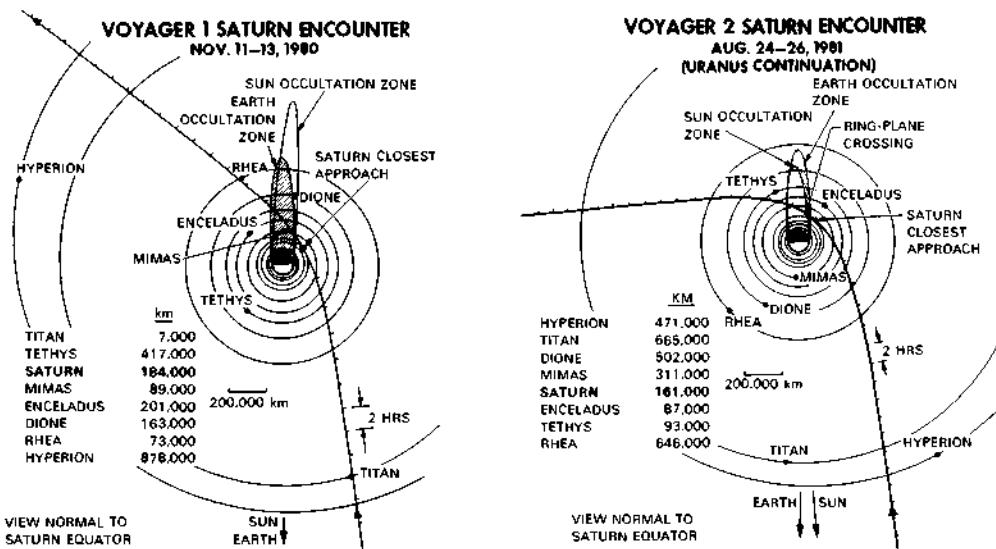


Fig. 1. Voyager 1 and 2 trajectories



Figs. 2 & 3. Voyager 1 & Voyager 2 spacecraft Saturn near encounter trajectories

of solar independent power. The extreme distances required increased long range communications capability. The long roundtrip light times in communicating with the spacecraft (2 hours 50 minutes at Saturn and 5 hours 20 minutes at Uranus) and the requirement for long periods of unattended operation required a high degree of spacecraft autonomy. A configuration drawing of the Voyager spacecraft is shown in Fig. 4. The spacecraft injected mass was 826 kilograms and nominally requires 374 watts of power. Its unusual appearance is driven primarily by the special requirements of solar independent power, long range communications and integration requirements of the science payload. The science payload mass is 117 kilograms and requires 91 watts independent of temperature control power.

Radioisotope thermoelectric generators (RTG's) were used to provide the solar independent power. The three RTG's were mounted in a linear stack and oriented such that self shielding would minimize the neutron and gamma ray radiation effects upon the spacecraft equipment and particularly the radiation sensing science instruments. The three RTG's provided approximately 430 electrical watts at Saturn.

The increased long range communications capability was achieved by increasing the capabilities on both the ground and in the spacecraft. The Deep Space Network's

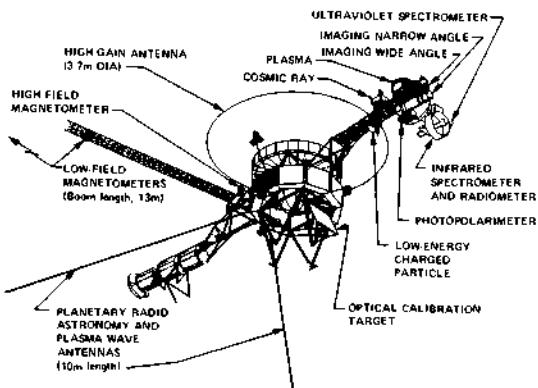


Fig. 4. Voyager spacecraft in-flight configuration

stations' performance was improved, including additional improvements for Saturn involving the use of super cooled masers and the routine arraying of the 64 and 34 meter antennas at each site. In the spacecraft a high gain antenna (3.65 meters was used to transmit on either S-band ( $\sim 2300$  MHz) or X-band ( $\sim 8400$  MHz)). At X-band with a 21.6 watt output traveling wave tube amplifier, 115.2 kilobits per second and 44.8 kilobits per second was received from Jupiter and Saturn respectively. Additional ground station arraying is planned to enhance the performance capabilities further for the Voyager 2 Uranus encounter in January 1986.

Autonomous operation of the Voyager spacecraft (5) is achieved by special purpose on-board computers which execute stored sequences for automatic operation of the spacecraft and special stored algorithms in response to on-board fault indications in order to protect the spacecraft. The on-board fault response actions generally put the spacecraft into a safe configuration pending intervention by the operations personnel via radio commands. However, certain faults can be corrected without ground intervention.

The Voyager Project carries out the eleven science investigations listed in the following table.

#### Voyager Project Science Investigations

<u>Investigation</u>	<u>Principal Investigator/Organization</u>
Imaging Science	B. Smith, University of Arizona (Team Leader)
Infrared Spectroscopy and Radiometry	R. Hanel, Goddard Space Flight Center
Photopolarimetry	A. Lane, Jet Propulsion Laboratory
Ultraviolet Spectroscopy	L. Broadfoot, University of Southern California
Radio Science	G. Tyler, Stanford University (Team Leader)
Magnetic Fields	N. Ness, Goddard Space Flight Center
Plasma	H. Bridge, Massachusetts Institute of Technology
Plasma Wave	F. Scarf, TRW
Planetary Radio Astronomy	J. Warwick, Radiophysics, Inc.
Low-Energy Charged Particles	S. Krimigis, Johns Hopkins University
Cosmic Ray	R. Vogt, California Institute of Technology

The Imaging Science Subsystem, Infrared Interferometer Spectrometer, Ultraviolet Spectrometer and Photopolarimeter are mounted on an articulated scan platform. The platform can point these instruments at selected targets with an accuracy of better than 2.0 milliradians. The required pointing instructions are included in the on-board stored sequences. The Cosmic Ray, Low Energy Charged Particle, Magnetometer, Planetary Radio Astronomy, Plasma, and Plasma Wave instruments were integrated in such a way as to satisfy their sensor systems' requirements and minimize instrument to instrument and spacecraft to instrument interferences. For example, the two low field magnetometer sensor systems were mounted on a 13-meter boom with one at the tip and one at 0.6 of the length in order to minimize and measure interfering spacecraft magnetic fields.

## OPERATIONS

The Voyager operations system consists of a Ground Data System, a Mission Operations System and institutional support by the Mission Control and Computing Center (MCCC) and by Tracking and Data Acquisition (TDA) (Deep Space Network stations in Australia, California, and Spain). The Ground Data System provides the necessary hardware and software to handle the received data in "real time" for operations control and "non-real time" for engineering and science data processing and analysis. Experiment data records are prepared for each investigation and provided to the investigation teams for their detailed analysis and interpretation. The Ground Data System also provides the tools required for the generation and validation of uplink commands including the sequence loads.

The Mission Operations System provides the organization, procedures, and knowledge for the planning and subsequent execution of the desired operation of the spacecraft and its science instruments. In order to carry out such activity, the Mission Operations System brings together in a controlled, working environment, the scientists, science planners, sequence designers, spacecraft experts, navigators, mission controllers and appropriate ground system interface personnel including the MCCC and TDA to plan and execute the necessary operations.

The sequence loads which are developed for automatic operation of the spacecraft consist of approximately 2000 command words per load. During the cruise between planets when the activity level is relatively low, a sequence load is sent to the spacecraft every four weeks. However, during the peak activity period of a planetary encounter, a sequence load may execute in less than twenty-four hours due to the higher activity level. The care put into the design of these sequence loads has been a key contributor to the success of Voyager. Without the attention to detail and the dedicated effort of the total operations team including the supporting scientists, serious compromises in the spectacular results achieved from the Jupiter and Saturn encounters might well have occurred.

## JUPITER (6) (7) (8)

The Voyager spacecraft were highly successful in their exploration of the Jovian system in 1977 and the results have been reported in the literature. New and unique data were provided for the study of the magnetosphere, atmosphere and satellites of Jupiter. A few highlights from the findings are provided.

The magnetosphere of Jupiter is an immense structure which varies in size with variations in the solar wind pressure. The two spacecraft each made multiple crossings of the bowshock and magnetopause before finally entering the magnetosphere on the approach to Jupiter. Within the magnetosphere, the plasma of electrically charged particles is flattened into a disk more than 4.8 million kilometers in diameter. The disk is coupled to Jupiter's magnetic field and rotates around Jupiter. Energetic ions of sulfur and oxygen were detected in the outer magnetosphere. In addition to the Io flux tube in which a current of  $2.5 \times 10^6$  amperes was measured, a plasma torus was discovered at the orbit of Io which contained highly ionized sulfur and oxygen (SIII, SIV and OIII).

The atmosphere of Jupiter (Fig. 5) proved to be far more turbulent than expected. The cloud structures in the northern and southern hemispheres are distinctly different with extensive cyclonic and anticyclonic activity. The Great Red Spot is an anticyclonic feature rotating in an anticyclonic shear zone. The white



Fig. 5, Voyager 1 photo of Jupiter  
Jan. 24, 1979 at 40 million km

ovals and other anticyclonic spots at mid-latitudes resemble the Great Red Spot. Extensive lightning and aurora activity was detected in Jupiter's atmosphere. The Galilean satellites of Jupiter (Fig. 6) proved to be quite varied with each satellite having unique characteristics and features. Callisto presented an ancient, heavily cratered crust with vestigial rings of enormous impact basins since erased by flow of the thick ice crust. Ganymede, however, had both highly cratered terrain and much younger, grooved terrain suggesting that some type of global tectonic processes have occurred. Europa appears to be virtually craterless with numerous intersecting linear features possibly caused by crustal rifting and/or tectonic processes. Driven by tidal heating, eight active volcanoes were discovered on Io. These volcanoes are the source of the sulfur and oxygen ions in the Io torus and outer magnetosphere since sulfur and sulfur dioxide are the principal ingredients of the volcano plumes. Amalthea was also photographed and three new satellites and a ring of very fine material inside Amalthea's orbit were discovered.

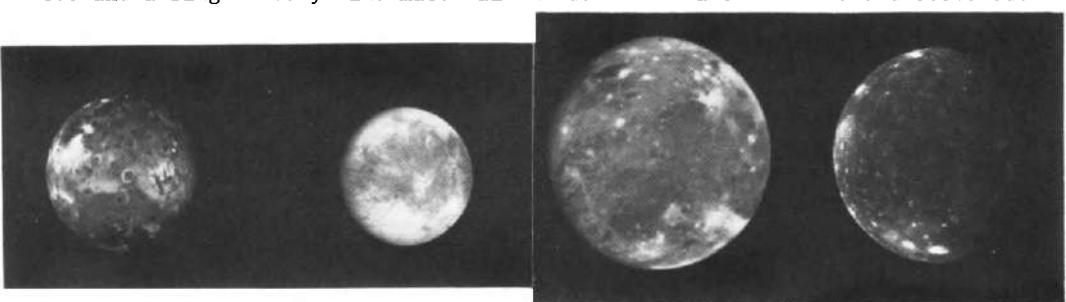


Fig. 6, Voyager 1 photo of Galilean satellites

#### SATURN ENCOUNTER

The Voyager exploration of the Saturn system presented special challenges in designing the encounter sequence loads. The maximum communications rate from Saturn of 44.8 kilobits per second vs. 115.2 kilobits per second from Jupiter resulted in minimum imaging frame times of 144 seconds versus the 48 seconds at Jupiter. The Saturn system is smaller than the Jupiter system and at the encountering velocity only 37 hours were involved in the Voyager 1 near encounter versus 200 hours at Jupiter. In addition, eight versus four satellites were mapped, five versus one additional satellites were imaged and 19 hours versus 0.6 hours were spent on ring studies in the plus or minus 24 hours around closest approach.

The encounter is divided into the observatory, far encounter 1, far encounter 2, near encounter and post encounter phases. For Voyager 1, these phases began as follows: Observatory - August 22, 1980; Far Encounter 1 - October 24, 1980; Far

Encounter 2 - November 2, 1980; Near Encounter - November 11, 1980, and Post Encounter - November 13, 1980 to December 16, 1980. The corresponding Voyager dates are: Observatory - June 5, 1981; Far Encounter 1 - July 31, 1981; Far Encounter 2 - August 11, 1981; Near Encounter - August 25, 1981 and Post Encounter August 27, 1981 to September 28, 1981. The observatory phase runs until 2x2 mosaics are necessary to capture the entire planet and ring system. Far encounter 2 starts when the 2x2 mosaics end and area targeting with occasional larger area mosaics are used. The near encounter phase covers the period of close planet and satellite flybys. Post encounter is a repeat of the observatory phase type of periodic activity as the spacecraft leaves the Saturn system. Spectacular data was provided during the Voyager 1 encounter and through its observatory phase Voyager 2 has been doing equally well. The Voyager 2 cameras are more sensitive than the ones on Voyager 1 and have been providing better images at equivalent ranges.

#### SATURN (9)

Saturn is smaller and colder than Jupiter and its atmospheric features are less visible because of a thick haze layer above the visible cloud tops. Figs. 7 and 8 show views of Saturn by Voyager 1 and 2 from equivalent ranges.

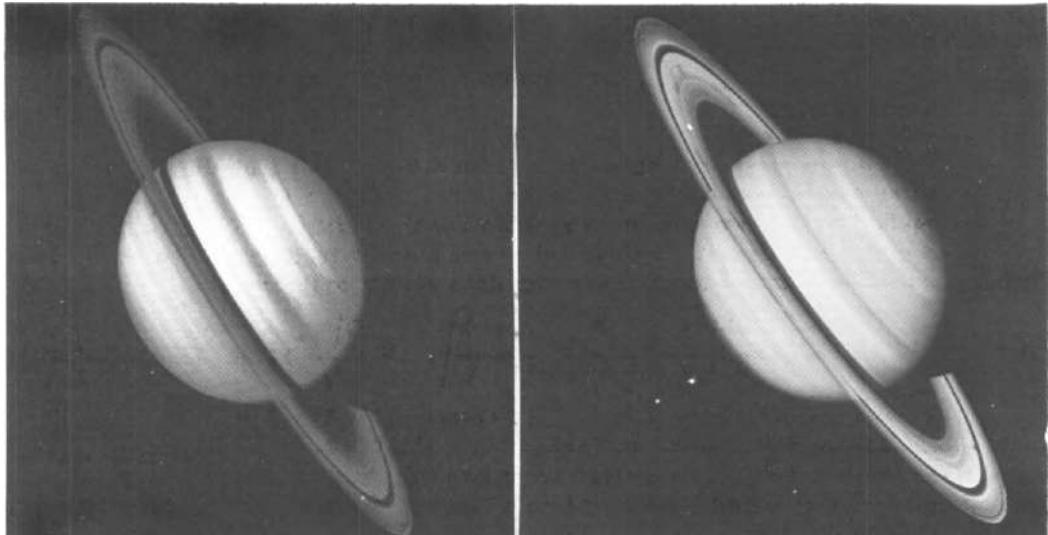


Fig. 7, Voyager 1 photo of Saturn  
Oct. 18, 1980 at 34 million km

Fig. 8, Voyager 2 photo of Saturn  
July 20, 1981 at 33.9 million km

Saturn's rings were edge-on to the sun in March, 1980 and as Saturn moves into the northern hemisphere spring, the rings become more fully illuminated. In the Voyager 2 picture the rings are almost as bright as the planet. Several ring and planet atmospheric features are clearly visible in Fig. 8.

While the atmosphere of Saturn has a similar appearance to that of Jupiter, the wind speeds are substantially higher and do not seem to be closely tied to the belt/zone boundaries. The highest wind speeds (approximately 1770 kilometers per hour) blow to the east at the equator (four times higher than on Jupiter). The

wind speeds appear to fall off smoothly to nearly zero at 40° north and south latitude.

Voyager 1 measured aurora-like emissions near the illuminated limbs and auroras in ultra-violet in a band near 80° south latitude. Efforts to observe lightning on the dark side of Saturn were unsuccessful due to the high levels of light reflected onto the dark side from the illuminated rings. Radio emissions typical of lightning were measured, but the discharges appeared to come from the rings rather than from the planet.

Radio emissions from the north polar region and from near 90 degrees longitude indicated that the body of Saturn and its magnetosphere rotate with a period of 10 hours, 39 minutes and 26 seconds compared to the classical 10 hours, 14 minutes.

As illustrated in Fig. 9, the traditional A-, B- and C- rings actually consist of hundreds of narrower rings. The computer enhancement normalizes the brightness across the rings and brings out the ring structure. Even the classical gaps such as the Cassini Division are seen to contain rings. The radio occultation data indicated effective particle sizes of 2 meters in the C-ring; 8 meters in Cassini Division near A-ring and 10 meters in the A-ring.

The F-ring discovered by Pioneer 11 in 1979 can be seen just outside the edge of the A-ring.

Fig. 10 shows a closeup view of the F-ring indicating that it really consists of several rings which appear to be intertwined. Voyager 2 will make observations of the F-ring designed to provide data on its structure and dynamics. Voyager 1 discovered and photographed three "shepherd" satellites associated with the outside edge of the A-ring (1980S28-30km Diameter), the inside edge of the F-ring (1980S26-200 km Diameter), and outside the edge of the F-ring (1980S27-220 km Diameter). These shepherds interact with the material in the rings in such a way as to keep the material confined to the ring.

Voyager 1 also photographed the D- and E- rings during its passage through Saturn's shadow and confirmed a new ring near 2.8 Saturn radii (from Saturn's center) which had been postulated by Pioneer 11 fields and particles data.

Fig. 11 shows a series of photographs taken at approximately 15-minute intervals which show spoke-like features which rotate with the rings. The spokes appear dark when photographed on approach but bright when photographed leaving Saturn. This suggests that the spokes are extremely small particles with strong forward scattering properties. It is possible that the fine material becomes charged and then levitated above the rings. The lightning-like electrical discharges in the rings may be associated with these spoke-like features.

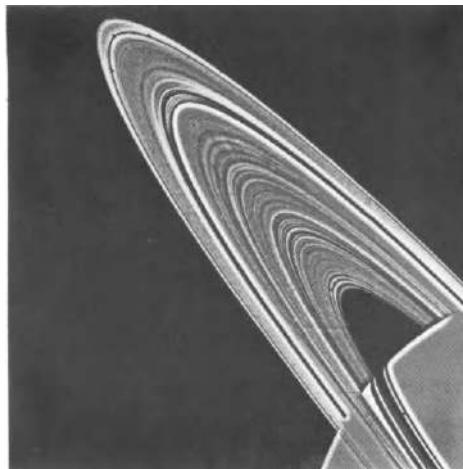


Fig. 9. Voyager 1 mosaic - Nov. 6, 1980  
at 8 million km

Voyager 1 photographed all of Saturn's known satellites except for Phoebe. Saturn's satellites are low density and therefore made up mostly of water ice.

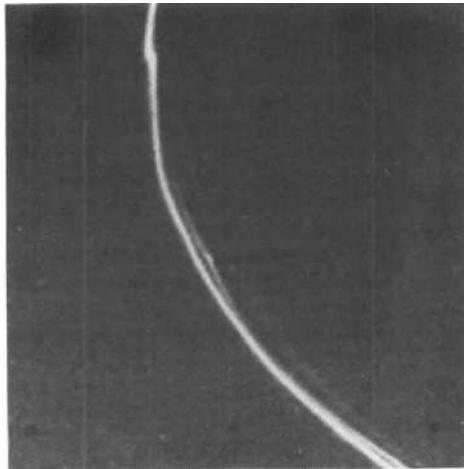


Fig. 10. Voyager 1 photo of F-ring  
Nov. 12, 1980 at 750,000 km.

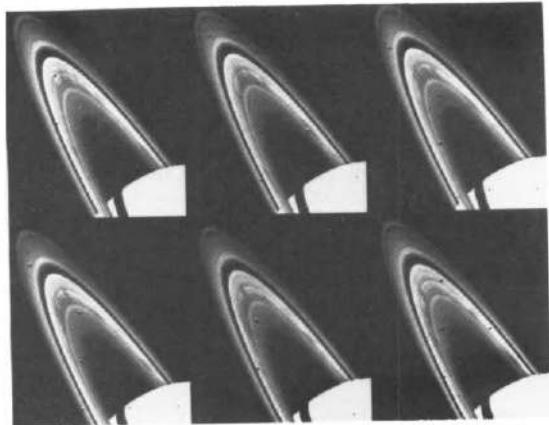


Fig. 11. Voyager 1 photo sequence  
Oct. 25, 1980 at 24 million km  
showing spoke movement

However, the satellites are so cold that the ice exhibits higher strength. As a result, the satellites exhibit significant elevation relief and have features such as rebound peaks in the center of craters.

Figures 12 and 13 picture the satellites Mimas and Dione. Mimas is approximately 390 km in diameter. The large crater is almost 100 km in diameter. Dione is larger at 1120 km in diameter and is highly cratered. The largest crater shown is slightly less than 100 km in diameter. The sinuous valley seen near the terminator may be a great fracture in the crust of the satellite.

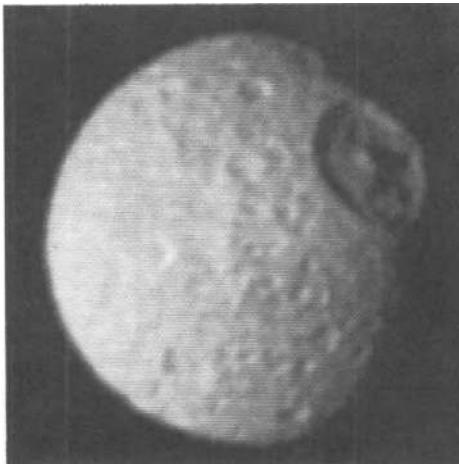


Fig. 12. Voyager 1 photo of Mimas  
Nov. 12, 1980 at 425,000 km

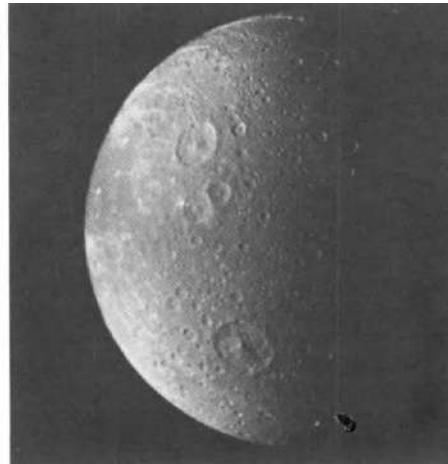


Fig. 13. Voyager 1 photo of Dione  
Nov. 12, 1980 at 162,000 km

Voyager 1 made a close flyby of Rhea and photographed a surface which was saturated with craters. Enceladus, however, showed no evidence of cratering at 12 kilometers resolution. There is speculation that Enceladus may be the source of the E-ring material since its orbit is at the maximum intensity of the E-ring. Enceladus may well experience tidal stressing similar to Io and/or Europa thus having a mechanism to erase cratering. Voyager 2 will provide much higher resolution photographs of Enceladus.

Titan has proved to be the most fascinating satellite in our solar system. Its unique atmosphere shrouded its surface from view by the Voyager cameras. However, other investigations were able to answer many questions about Titan. Titan had long been thought to be the largest satellite in our solar system, but the radio occultation data indicated that its diameter is 5140 km (slightly smaller than Jupiter's Ganymede). Titan's density is therefore about twice that of water requiring that Titan be composed of nearly equal amounts of rock and ice, as is Ganymede.

The atmosphere has a dense, optically thick photochemical haze layer. Several distinct, detached haze layers were photographed which merged into a thick hood at the north pole. The Radio Science occultation data indicated that the atmospheric pressure near the surface was 1.6 bars (60 percent higher than the Earth). Methane had been detected from the Earth, but the Voyager instruments indicated that nitrogen was the principal ingredient with ethane, acetylene, ethylene, and hydrogen cyanide being positively identified in addition to methane.

The surface temperature appears to be about 93 degrees Kelvin which is close to the triple-point for methane. This suggests that methane may play the role in Titan's atmosphere that water plays in the Earth's atmosphere. Fig. 14 makes a comparison between the Earth's and Titan's atmospheres. Methane may exist as liquid on the surface in rivers and lakes or as rain, as solid on the surface as snow or ice and in the "air" as ice crystals in the clouds or snow, or as a gas or vapor in the "air." Titan could have a varied surface with rivers, lakes, glaciers, etc. depending upon the temperature, elevation, latitude and perhaps season. Because of the photochemistry high in the atmosphere, other hydrocarbons may well have been raining down on the surface creating a very thick layer of such material and unimaginable surface conditions. Titan should be given serious consideration in planning future outer solar system missions.

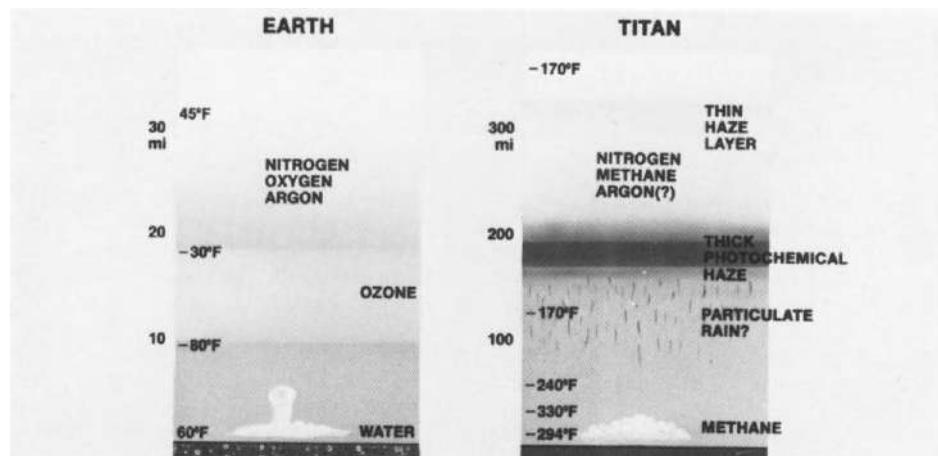


Fig. 14. Comparison of the atmospheres of Titan and the Earth

Saturn's magnetosphere is about one-third the size of Jupiter's extending nearly 1.6 million kilometers. The Voyager instruments detected a cloud of uncharged hydrogen surrounding Titan and its orbit and extending in to Rhea's orbit forming an enormous hydrogen torus. A disk of plasma composed of hydrogen and possibly oxygen ions was also detected which extended from outside Tethys orbit almost to Titan's orbit. The plasma is nearly in full co-rotation with Saturn's magnetic field.

The Voyager 2 spacecraft will build on the findings of Voyager 1 and particularly focus on answering questions raised by the new discoveries such as the ring spokes and the "intertwined" F-rings. Voyager 2 will also provide higher resolution imaging of Enceladus and Tethys. Perhaps the images of Enceladus will give some clue as to the possible interaction with the E-ring.

Voyager 2 can, barring failures, provide equally spectacular data on the distant Uranus and Neptune systems. Ample fuel exists and no life limiting characteristics are known to be a problem. Fig. 15 illustrates the Uranus encounter in January 1986.

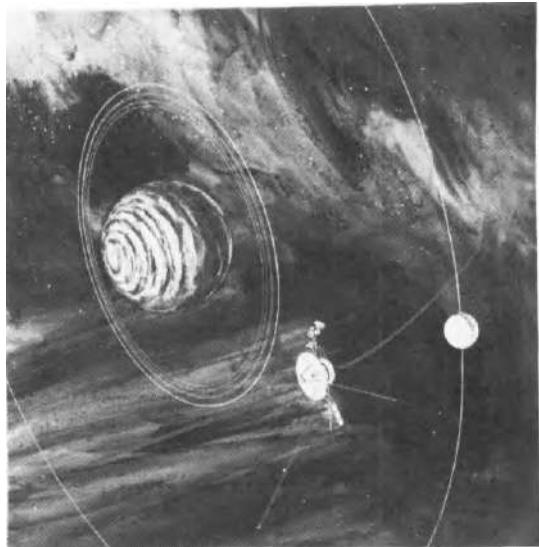


Fig. 15, Uranus encounter in Jan. 1986 with a close flyby of satellite Miranda

#### VOYAGER 2 ADDENDUM:

The Voyager 2 spacecraft was highly successful in its encounter with Saturn. New information was obtained on Saturn's magnetosphere, satellites, rings and atmosphere. The Voyager 2 scan platform did have problems which caused the loss of some remote sensing data but the results still exceeded expectations.

The Voyager 2 photopolarimetry, unlike the Voyager 1 unit, was functional and made a wide variety of observations including a star ring occultation. The Voyager 1 cameras had revealed that hundreds of rings existed. However, the Voyager 2 photopolarimeter star occultation revealed that thousands and perhaps tens of thousands of rings exist.

Voyager 2 provided complementary satellite coverage to that provided by Voyager 1. Excellent photographs of Enceladus, Iapetus and Hyperion were provided. Enceladus was observed to have a complex surface somewhat resembling Ganymede even though it is a tenth the size of Ganymede. Hyperion is very unusual in appearance with a very irregular shape like a thick (210 km) irregular disk (360 km).

The scan platform problem began while the spacecraft was in occultation and involved erratic motion of the azimuth axis causing the loss of many planned observations. Normal scan platform sequencing was terminated for the balance of the encounter. The Voyager operations personnel were able to generate modified sequences in order to continue the observations of Saturn during the post encounter period and photograph Phoebe during the spacecraft's closest approach. Laboratory tests with a spare actuator indicate that the problem is probably due to the loss

and/or degradation of the lubrication associated with one or more gears in the actuator gear train. Laboratory tests are continuing.

Despite the scan platform problem the outstanding results from the Voyager 2 encounter have been quoted as providing 200% of the expected results. The spacecraft is successfully on its planned trajectory toward Uranus with ample fuel reserves. The Voyager operations team will plan and work to optimize the return at Uranus taking into account the scan platform problem and a planned enhancement of the science data return capability. Given no disastrous problems, Voyager 2 can be expected to provide outstanding and undoubtedly unexpected results during its January 1986 Uranus encounter.

#### REFERENCES

- (1) (1970). Tops-Outer Planet Spacecraft. Astronautics and Aeronautics, 8-9.
- (2) (1974). Pioneer 10 Mission, Jupiter Encounter. Journal of Geophysical Review, 79-25.
- (3) Kohlhase, C. E. (1977). Mariner Jupiter/Saturn 1977 Mission Design Tradeoffs. 13th Annual AIAA Meeting, AIAA Paper No. 77-281.
- (4) Heacock, R. L. (1980). The Voyager Spacecraft, Proceedings of The Institution of Mechanical Engineers, 194-28.
- (5) Jones, C. P. (1979). Automatic Fault Protection in the Voyager Spacecraft. AIAA Paper No. 79-1919.
- (6) (1979). Voyager 1 Encounter with the Jovian System. Science, 204-4396.
- (7) (1979). Jupiter and Io. A Special Supplement. Nature, 280-5725.
- (8) (1979). Mission to Jupiter and Its Satellites: Voyager 2. Science, 206-4421.
- (9) (1981). Voyager 1 Mission to Saturn. Science, 212-4491.

## GALILEO: MISSION TO JUPITER

J. R. Casani

*Jet Propulsion Laboratory, California Institute of Technology,  
Pasadena, California 91109, USA*

### ABSTRACT

The Galileo mission to Jupiter in the latter half of this decade is NASA's next step in the exploration of Jupiter. Galileo will explore in depth many unanswered and newly suggested questions which remain after the Voyager mission. The primary science objectives are to study the satellites, the magnetosphere, and Jupiter's atmosphere. In general, there are four characteristics of the Galileo mission which provide the capability to address questions not answered by Voyager: (1) long-term observations--at least eleven orbits of Jupiter in nearly two years--will allow temporal studies of Io vulcanism, for example, as well as of interactions between the satellites and magnetosphere; (2) very close satellite flybys--at a distance less than 1000 kilometers at each Galilean satellite--will allow in-depth studies of these satellites at distances 20 to 100 times closer than Voyager achieved; (3) an atmospheric entry probe will measure composition, structure, temperature, energy balance, cloud layer locations and structure, and particle size distribution in Jupiter's atmosphere; and (4) the Orbiter's advanced instrumentation will allow higher resolution and more detailed studies of the satellites, atmosphere, and magnetosphere. The Galileo Project is managed for NASA by the Jet Propulsion Laboratory, California Institute of Technology.

### KEYWORDS

Project Galileo; Jupiter; atmospheric entry probe; planetary orbiter; solar system exploration.

### INTRODUCTION

Project Galileo is the next step in the exploration of the giant planet Jupiter, its "mini-solar system" of satellites, its tenuous ring of dust, and its extensive magnetic field. Four missions--Pioneer 10 (1973), Pioneer 11 (1974), and two Voyagers (1979)--have made initial reconnaissance passages through the Jupiter system. The wealth of data returned by these flyby missions

emphasizes the need for Project Galileo, which will send a probe into the planet's atmosphere to sample it directly and an orbiter to tour the satellite system for an extended period of nearly two years, sampling the magnetic field and mapping the satellites. The mission is designed to return the maximum amount of information from what is our only mission to Jupiter in the foreseeable future. This can be achieved not by a brief flyby mission, but by a mission which remains in the Jupiter system and actually probes the atmosphere of our solar system's largest planet.

Galileo is designed to fulfill the three major objectives for Jupiter exploration as defined by the National Academy of Sciences: (1) to determine the chemical composition and physical state of Jupiter's atmosphere, (2) to determine the chemical composition and physical states of Jupiter's satellites, and (3) to characterize Jupiter's magnetic field and charged particle environment.

Jupiter has characteristics in common with Earth, the Sun, and astrophysical magnetic objects such as pulsars. Galileo provides an unparalleled opportunity to study the Jupiter system in depth, adding to our basic understanding of the origin and evolution of the solar system and to our understanding of the Earth and its meteorological, geological, and magnetic processes.

Galileo takes the exploration of the Jupiter system beyond the capabilities of previous missions in four ways: (1) extended observations--nearly two years of operations as the Orbiter tours the system, (2) closer flybys--at distances 20 to 100 times closer to the major satellites than achieved by the stunning Voyager mission, (3) direct sampling of Jupiter's atmosphere with an entry probe, and (4) advanced instrumentation.

Galileo is an ambitious undertaking; indeed, all of our planetary missions have been ambitious--and highly successful.

#### MISSION DESCRIPTION

The baseline mission plan is to launch the combined Orbiter and Probe spacecraft aboard the Space Shuttle in April 1985. Figure 1 shows the trajectory to Jupiter. The mated spacecraft will be attached to the top of a modified Centaur upper stage rocket, and all three will nestle inside the Shuttle's payload bay. Once in Earth orbit, the payload bay doors will open and the Galileo spacecraft and Centaur will undergo an extensive engineering checkout. Provided all systems survived the rigors of the launch, the tilt table holding the spacecraft and upper stage will rotate 45 degrees, pyrotechnic devices will fire, and steel springs will gently push the Centaur-Galileo away from the Shuttle at a rate of about 1.6 m/s. The Shuttle will fly away from Galileo and return to Earth. At the end of its fourth Earth orbit, the Centaur main engines will ignite to inject the Galileo spacecraft into its interplanetary trajectory toward Jupiter. The power supply and science booms will deploy prior to Centaur separation, and the umbrella-like high-gain antenna will unfurl after separation (Yeates and Clarke, 1981).

The Earth-to-Jupiter cruise phase will be supported with tracking coverage from the Federal Republic of Germany (FRG) and NASA's Deep Space Network. During cruise, all instruments aboard the Orbiter will be calibrated, and the fields and particles instruments will conduct cruise science, sampling the solar wind periodically.

About 10 months after launch, the spacecraft will perform a heliocentric plane change, or "broken plane" maneuver. This maneuver reduces both the required declination of the Earth departure asymptote and the launch energy. The 1985

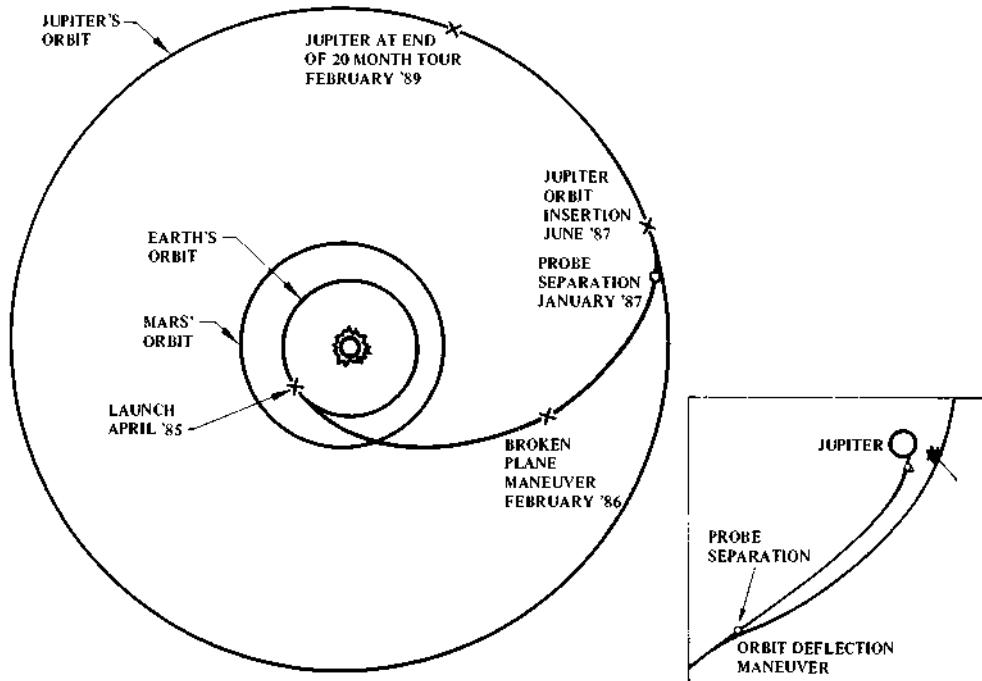


Fig. 1. Galileo to Jupiter.

launch opportunity involves high declination of the Earth departure asymptote, which increases performance demands on the launch vehicle or injection stages. By using the 10-Newton thrusters of the Orbiter's retropropulsion module (supplied by FRG) to perform a broken plane maneuver, the additional demand on the launch system can be somewhat alleviated (Farless, 1981).

The exact flight path and Jupiter arrival date depend on the exact launch date and the launch energy. There is a primary launch period of ten days. With each day into the launch period, the Jupiter arrival date moves back about two weeks. While the Sun-relative orientation angles during approach and in the Orbiter's initial orbit are a strong function of launch and arrival dates, early or later arrivals do not degrade the science value of the mission. Therefore, the launch period is being chosen primarily to reduce the total required launch performance.

Early in 1987, about five months before Jupiter arrival, the Orbiter will release the Probe and then deflect itself away from the Probe trajectory. Both will continue along independent trajectories toward Jupiter (see Fig. 2). The Probe spins at 10 revolutions per minute to maintain a stable attitude. The Probe has no maneuvering capability, so it must be accurately aimed and oriented by the Orbiter prior to release. The angle of the Probe's entry path into the atmosphere is very critical. If the Probe enters the atmosphere too steeply, the aerodynamic forces will destroy it. Conversely, if the entry angle is too shallow, the Probe will ricochet off the top of the atmosphere, skipping back into space.

The Probe is targeted to enter the atmosphere within either of two narrow bands from 1.0 to 5.5 degrees north or south latitude. This is within Jupiter's light-colored equatorial zone where the topmost clouds are believed to consist of ammonia. By entering through the highest cloud region, the Probe will be able to make diagnostic measurements of all the important cloud layers.

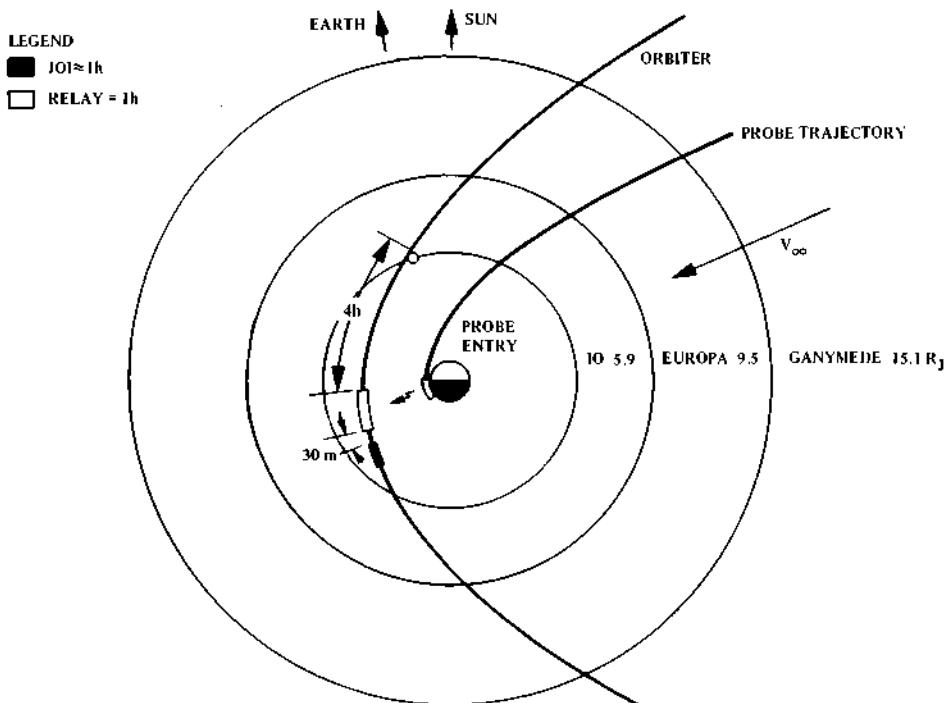


Fig. 2. Typical Jupiter arrival geometry.

Hurtling toward Jupiter, the Probe will enter the atmosphere at a speed of about 48 kilometers per second (over 100,000 miles per hour). This will be a critical point, since no other man-made object has entered an atmosphere at such speeds (neither the Apollo returns from the Moon to Earth nor the Venus probes). In addition, the chemical processes of the gases in the Jovian clouds are not well understood: whether the processes absorb or radiate heat is not known. The Probe will experience a deceleration force of almost 400 G's. Most of its protective heatshield--accounting for almost 50 percent of the Probe's total weight--will glow brightly and burn away, carrying with it the fierce heat caused by the friction of entry. These extreme conditions will pass within a few seconds. Then the aft cover of the Probe will be jettisoned, and a drogue parachute will pop open to slow the Probe even more. Within seconds, the main parachute will pop open and the forward heatshield will be jettisoned. The Probe will now be at the top layers of the clouds, about two minutes after entering the upper reaches of the atmosphere. As the Probe drifts downward, suspended on its parachute, it will transmit scientific measurements to the Orbiter, which is passing overhead. The Probe is expected to send data for about 60 minutes. About 40 minutes after entry, the Probe will have penetrated to a depth of about 10 bars, believed to be below the lowest water cloud layers. The temperature will be about 73°C. At 60 minutes, the Probe will be at the 15 to 20 bar level--100 to 130 kilometers below the cloudtops. Below this level, the Probe is not expected to survive--silenced by a combination of increasing temperature and weakened radio signals (Galileo to Jupiter, 1979). Figure 3 shows the Probe descent.

A few hours before Probe entry and the Orbiter's closest approach to Jupiter, the Orbiter will fly within 1000 kilometers of the volcanic satellite Io--twenty times closer than Voyager 1's closest approach to this dynamic satellite. This will be the Orbiter's only flyby of Io since increased time in the intense

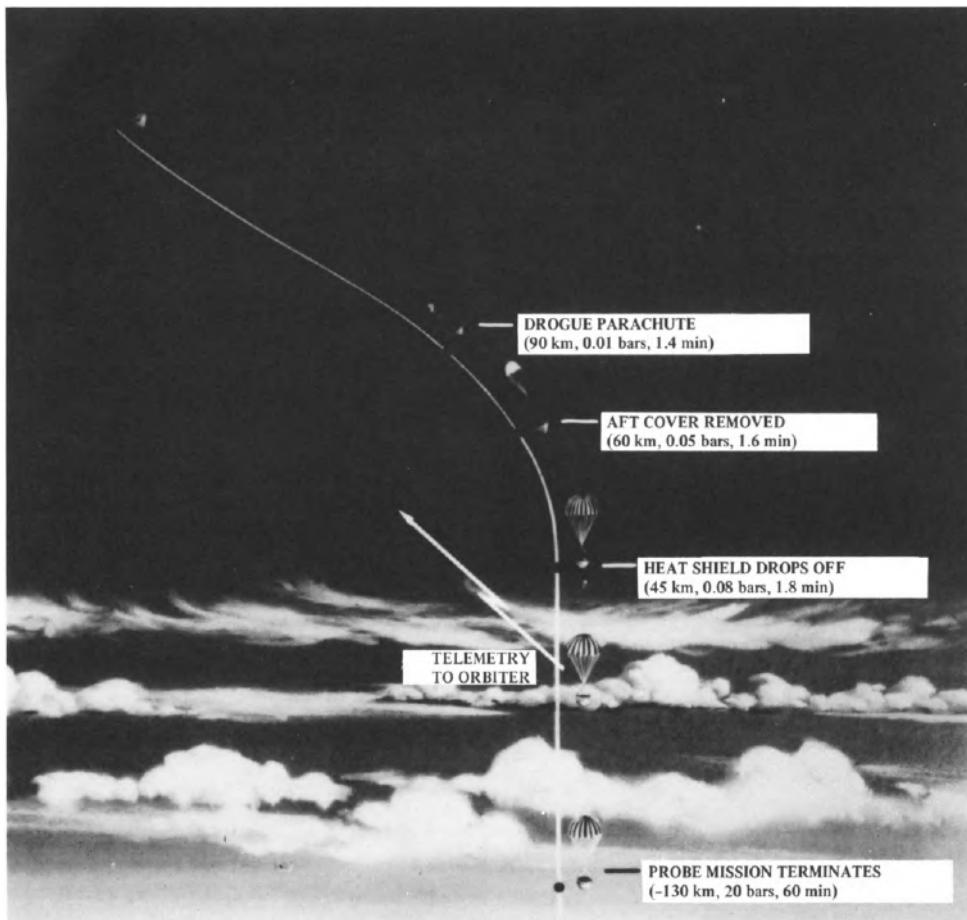


Fig. 3. Probe descent.

radiation environment near Jupiter could severely damage the spacecraft. The gravitational attraction of Io will slow the Orbiter, reducing the amount of propellant required for the Jupiter orbit insertion maneuver after Probe entry. This Io gravity assist will reduce the total velocity change requirement by about 150 m/s.

The Orbiter will be about 200,000 kilometers above Jupiter's clouds as the Probe makes its entry. Data from the Probe will be relayed to the Orbiter and then from the Orbiter to Earth. Shortly after the Probe mission is complete, the Orbiter will fire the 400-Newton engine of its retropropulsion module for about one hour to complete the slowdown necessary to place it in an elliptical orbit about Jupiter. Initially, the orbit will range from 200,000 to 15 million kilometers above Jupiter. While the Orbiter will continue to orbit forever, the Galileo mission baseline design includes eleven orbits over the next 20 months, with a close encounter to Europa, Ganymede, or Callisto on each orbit. An extended mission is possible.

Galileo will be the first mission to use satellite gravity assists to tour a planetary system. The Orbiter will use the gravity of the satellites to bend its own orbit each time it nears a satellite. The use of the gravity assist technique to increase or decrease the energy of a trajectory requires the

interaction of three bodies--the spacecraft, an intermediate body with respect to which there is no energy change, and the central body about which the energy is being controlled (Mitchell, 1980). For additional details on gravity assist trajectory design, Nock and Uphoff (1979) give an excellent treatment of the subject, including theory and applications.

A number of interplanetary spaceflights have successfully used the gravity assist of an intermediate planet to reach subsequent targets (Mitchell, 1980). Galileo is using this concept more times than any of these missions, both to aid in orbit insertion, and to tour the satellites. A propulsive budget of 205 m/s is currently planned for the tour and is to be used for navigation purposes as well as for minor deterministic maneuvers to get from one satellite to another. If the spacecraft were to provide the total velocity change required to complete the tour, the additional capability required would be around 6 km/s (Mitchell, 1980).

Parameters such as solar phase angle, imaging resolution, and the extent to which coverage is complementary to Voyager coverage can be specified in terms of encounter distance, whether the encounter is pre- or post-spacecraft passage of perijove, and the orientation of the approach direction to the satellite with respect to the direction to the Sun (Yeates, private communication).

Figure 4 shows possible orbital tours for the Galileo mission, depending upon early or late arrivals at Jupiter. Tour design studies will continue until shortly before Jupiter arrival at which time the actual tour will be chosen.

The Orbiter's science requirements on the mission are:

- (1) To fly by Io. Although the altitude is still to be determined, the current plan is to fly about 1000 kilometers from Io near Jupiter orbit insertion for performance reasons, and this would meet the science needs. However, if this plan were abandoned for any reason (spacecraft safety, for example), the level to which science needs would be allowed to influence this plan is still an open item.

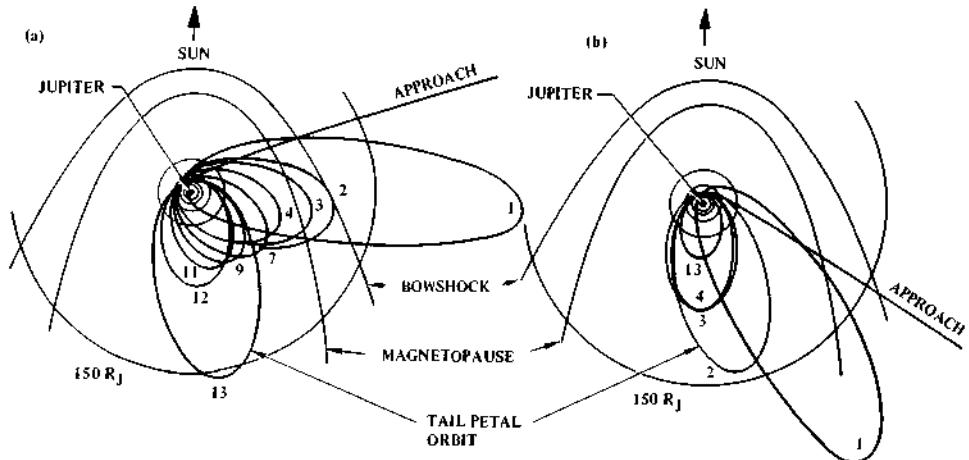


Fig. 4. Satellite tour examples: (a) early arrival, and (b) late arrival.

- (2) To be capable of taking advantage of nontargeted satellite encounters (i.e., flybys of satellites other than the primary satellite encounter on a given orbit), for science purposes, as close as 25,000 kilometers.
- (3) To make the following detailed satellite investigations:
  - (a) To obtain at least 50 percent solar-illuminated coverage of Europa, Ganymede, and Callisto with resolution  $\leq$ 1 kilometer.
  - (b) To map all the Galilean satellites, at high resolution, with moderate to low Sun angles.
  - (c) To obtain global coverage of the Galilean satellites at high and low Sun angles.
  - (d) To make synoptic observations of Io.
  - (e) To achieve at least one close flyby ( $\leq$ 1000 kilometers) of Europa and Ganymede for gravity and magnetic field studies.
  - (f) To pass through the Europa flux tube.
  - (g) To pass through the wakes of Europa, Ganymede, and Callisto at distances of less than 5000 kilometers from the surface.
  - (h) To achieve a radio occultation of each of the Galilean satellites.
  - (i) To achieve at least one occultation of the Io torus.
- (4) To make the following detailed Jupiter investigations:
  - (a) To perform observations of the Jovian ring.
  - (b) To obtain three Jovian occultations at ranges  $< 10^6$  kilometers at varying latitudes.
  - (c) To achieve at least one polar near occultation so that the spacecraft line of sight to Earth passes within  $0.1 R_J$  (Jupiter radii) of Jupiter.
  - (d) To achieve at least one Jovian ring occultation.
  - (e) For magnetospheric investigations, to achieve one orbit in the tour with an apojoove altitude  $> 150 R_J$  and with a Sun-Jupiter-spacecraft angle at apojoove of  $180^\circ \pm 15^\circ$ .

The Probe's science requirements on the mission are:

- (1) To enter at a Jovian latitude within a band between  $1^\circ$  and  $5.5^\circ$ N or between  $1^\circ$  and  $5.5^\circ$ S.
- (2) To allow acquisition of at least 60 minutes of Probe data after Probe entry.

In addition to the 114 international scientists selected for the investigation teams, fifteen interdisciplinary scientists have been selected. They will work with the data from several experiments and provide a broad link among the many disciplines involved in Project Galileo and its studies of the Jupiter system.

## SPACECRAFT DESCRIPTION

Orbiter

The Galileo Orbiter (Fig. 5) is a dual-spin spacecraft. Part of the spacecraft will be three-axis stabilized to provide a steady base for the remote-sensing instruments. These instruments must be precisely pointed. The despun section carries its related electronics. The main portion of the Orbiter will spin at three revolutions per minute to provide stability and to allow its instruments to continuously "sweep" the sky to make their measurements. The spun section contains both the high- and low-gain antennas, the retropropulsion module (RPM) for all propulsive and attitude maneuvers, the nuclear power sources, most of the electronics, most of the command and data equipment, and the fields and particles science instruments.

The RPM is being built by Messerschmitt-Bölkow-Blohm GmbH near Munich, Federal Republic of Germany, and is managed by Deutsche Forschungs-und-Versuchsanstalt für Luft- und Raumfahrt e.V. (DFVLR), a German research agency under the aegis of the German ministry for research and technology. The RPM has one 400-Newton engine and two clusters of 10-Newton thrusters that are used for attitude control and the smaller propulsive maneuvers. Because of minimum burn size and total wetted lifetime constraints, the 400-Newton engine will be used only for the deflection maneuver after Probe release, orbit insertion, and the perijove raise maneuver. Interplanetary maneuvers will be done using the 10-Newton thrusters (The retro-propulsion module, July 1981).

Galileo will use a 4.8-meter diameter (16-foot) furlable antenna to communicate with Earth. This antenna is similar to the one under development for NASA's Tracking and Data Relay Satellites. A small (1-meter) relay antenna will ride on the despun portion of the Orbiter to receive data from the Probe for relay to Earth.

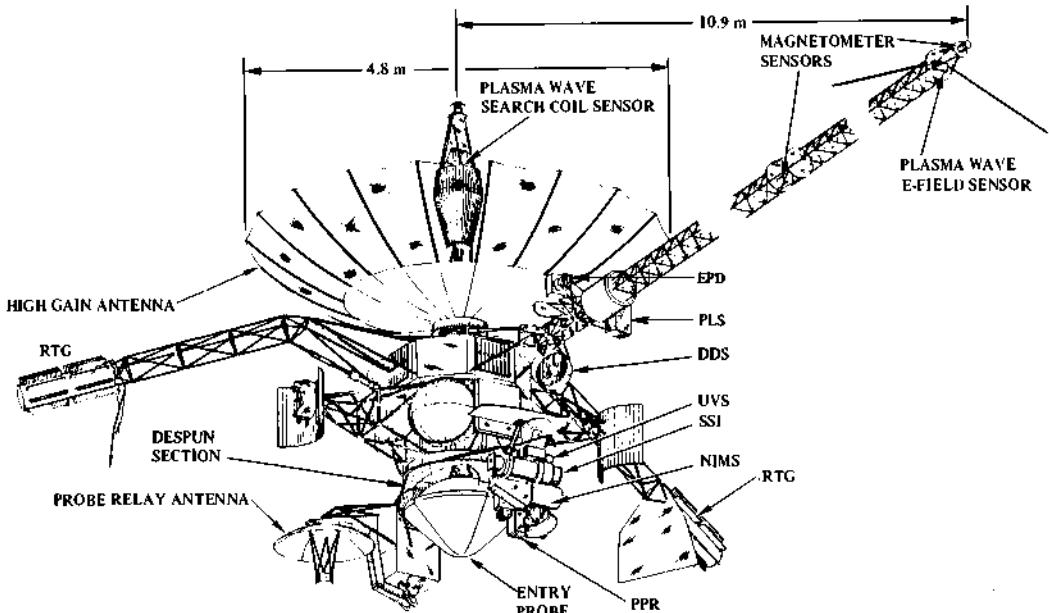


Fig. 5. Galileo spacecraft.

Since Jupiter is too far from the Sun for solar cells to provide enough electrical power, Galileo will use radioisotope thermoelectric generators (RTGs) similar to those flown on the two Voyagers. These nuclear power sources are developed by General Electric Company for the U.S. Department of Energy.

Table 1 lists the Orbiter's science investigations.

Despun Section. The despun section of the Orbiter carries four remote-sensing instruments: a solid-state imaging system, a near-infrared mapping spectrometer, a photopolarimeter radiometer, and an ultraviolet spectrometer.

The solid-state imaging (SSI) camera system uses a solid-state charge coupled device (CCD) instead of the vidicon tubes flown on previous planetary missions. The CCD system can provide broader spectral response and can take pictures every two seconds--a great advance over Voyager's 48-second picture-taking capability. This will allow more pictures to be taken while the Orbiter is at close range, providing more high-resolution coverage. A goal of Galileo is to map 50 percent of the surfaces of Europa, Ganymede, and Callisto at resolutions better than 1 kilometer. Several encounters with these large satellites should yield pictures with a best resolution of about 30 to 50 meters. All of the Galilean satellites will be observed with resolutions 20 to 100 times better than Voyager achieved.

The SSI is based on an 800 x 800-element silicon CCD. The camera head--composed of a radiation-shielded, radiatively cooled CCD, an eight-position filter wheel (ranging from <4450 to >9680 Å plus clear), and supporting electronics--is coupled to an optical telescope with 1.5-meter focal length and an aperture of f/8.5.

During the nominal 20-month mission, a total of 40,000 images of Jupiter and its satellites are planned. This coverage at higher resolution, over wider areas, and for longer periods of time than available in flyby missions will aid in understanding unusual geological formations on the satellite surfaces, Io's active volcanoes, Jupiter's atmospheric structure and dynamics, Jupiter's auroras and lightning storms, and the dynamics of Jupiter's tenuous ring.

The near-infrared mapping spectrometer (NIMS) will determine the chemical composition of the satellite surfaces and the characteristics of Jupiter's cloud layers, the spatial and temporal variations in the atmospheric constituents, and the temperature vs. altitude profiles in the region between 1 and 5 bars. The NIMS is capable of recognizing and classifying silicates, carbonates, nitrates, and other compounds that might account for the appearance of the satellite surfaces. The NIMS data will be compiled into geological maps of the satellite surfaces and will be compared to other moons and even asteroids in an effort to understand evolutionary processes in the solar system. At Jupiter, NIMS should be able to define cloud thicknesses and identify layers as a function of depth in the atmosphere. NIMS will learn about the atmospheric chemistry and dynamics by studying the variations in abundance of free molecular species such as water, ammonia, phosphine, and germane as a function of time of day and Jovian season.

The Galileo NIMS is a new instrument that has never been flown in space. The near-infrared region of the electromagnetic spectrum is powerfully diagnostic for gaseous species and solid-state reflection characteristics. The instrument primarily consists of a telescope and a spectrometer. The spectrometer features an array of 17 detectors, each sampling a limited spectral region.

Galileo's photopolarimeter radiometer (PPR) will measure temperature profiles and the energy balance of Jupiter's atmosphere as well as the composition and characteristics of the planet's clouds. The photometric, polarimetric, and

TABLE 1 Orbiter Science Investigations

Instrument Name	Principal Investigator/ Organization	Objectives
<b>Remote Sensing Instruments</b>		
Solid State Imaging (SSI)	M.J.S. Belton (Team Leader)/Kitt Peak National Observatory	Image Jupiter and its satellites for studies of atmospheric dynamics and physical geology.
Near Infrared Mapping Spectrometer (NIMS)	R.W. Carlson/Jet Propulsion Laboratory	Study mineralogy of satellite surfaces, as well as morphology and structure of Jovian clouds.
Photopolarimeter Radiometer (PPR)	J. Hansen/Goddard Institute for Space Studies	Study photometric and thermal properties of satellite surfaces as well as cloud and haze properties in Jovian atmosphere.
Ultraviolet Spectrometer (UVS)	C.W. Hord/LASP (University of Colorado)	Study composition and structure of high neutral atmospheres of Jupiter and Galilean satellites.
<b>Fields and Particles Instruments</b>		
Magnetometer (MAG)	M.G. Kivelson/ University of California, Los Angeles	Study magnetic field of Jupiter and search for magnetic fields associated with the satellites.
Plasma (PLS)	L.A. Frank/University of Iowa	Study Jovian plasma.
Plasma Wave (PWS)	D.A. Gurnett/University of Iowa	Study time-varying electric and magnetic waves in the Jovian plasma.
Energetic Particles Detector (EPD)	D.J. Williams/NOAA Space Environment Laboratory	Measure detailed energy and angular distribution of protons, electrons, and ions.
Dust Detector (DDS)	E. Grün/Max-Planck-Institut für Kernphysik (FRG)	Study physical and dynamical properties of small dust particles in the Jovian environment.
<b>Radio Science (RSS)</b>		
Celestial Mechanics	J.D. Anderson (Team Leader)/Jet Propulsion Laboratory	Study gravity fields of Jupiter and its satellites, as well as the space environment.
Radio Propagation	H. Howard/(Team Leader)/Stanford University	Study structure of atmospheres of Jupiter and satellites by use of radio signals from Orbiter and Probe.

radiometric measurements will provide spatial resolution, absolute accuracy, phase angle, and zenith angle coverage that is not possible with ground-based observations. The measurements are designed to learn about atmospheric dynamics, cloud heights, particle sizes and shapes, variations in the radiation energy budget (reflected vs. emitted radiation) on a three-dimensional scale, and the surface properties of the satellites.

The PPR consists primarily of a 10-cm diameter telescope mirror, a 32-position filter/retarder wheel for polarimetry, two silicon photodiode detectors for polarimetry, and a lithium tantalate pyroelectric detector for radiometric measurements, as well as a space-view telescope.

The Galileo ultraviolet spectrometer (UVS) will study composition and structure of the upper atmosphere of Jupiter and its satellites. It will also study the huge doughnut-shaped cloud of charged atoms (the torus) which is injected into the magnetosphere by Io's volcanoes.

The UVS will measure the composition and scattering properties of Jupiter's atmosphere at and above the cloudtops as a function of position on the planet. Ammonia and other UV-active molecules will be studied in the lower atmosphere. Atomic and molecular hydrogen will be studied in the upper atmosphere. The UVS will detect and explore any auroral emissions and map the distribution of UV-absorbing aerosols.

At the satellites, the UVS will scan the atmospheres of Europa, Ganymede, and Callisto to determine the existence of neutral and atomic molecular species of hydrogen, oxygen, nitrogen, carbon, sulfur, calcium, lithium, magnesium, molecular nitrogen, nitric oxide, carbon monoxide, sulfur dioxide, hydroxyl and cyano radicals, as well as ions of molecular nitrogen, carbon monoxide, carbon dioxide, and magnesium. The UVS will also look for condensates of ammonia and sulfur dioxide on the satellite surfaces.

The combination of the imaging, infrared, and ultraviolet data will yield valuable information on the surface coverage, the bulk and impurity composition, and the evolutionary history of the ices. The UVS instrument should also learn something about the extent to which the evolution of the satellite surfaces is being modified by proximity to Jupiter and its magnetic field.

The UVS instrument consists of a telescope, a spectrometer, and three detector channels. The spectral range is 1150 to 4300 Å. The telescope focal length is 250 mm.

Spinning Section. Five instruments--a magnetometer and instruments to study plasma, energetic particles, plasma waves, as well as a dust detector--ride the spinning section of the Orbiter. In addition, the spacecraft radio is used to perform celestial mechanics and radio wave propagation experiments.

The magnetometer (MAG) will measure magnetic fields and the ways they change near Jupiter and its satellites, including changes caused by the satellites. It will also search for possible magnetic fields originating in the satellites themselves. The magnetometer will map the structure of the magnetosphere, analyze its dynamics, study the properties of the satellites and their interaction with the magnetosphere, investigate magnetosphere-planetary ionosphere coupling, measure magnetic fluctuations and refine models of the main Jovian field.

The magnetometer consists of two sets of sensors, inboard and outboard. They are mounted on a boom extending from the spinning section of the Orbiter. The

outboard sensor operates in the range of  $\pm 16$  gamma and  $\pm 512$  gamma, while the inboard sensor operates in the range of  $\pm 512$  gamma and  $\pm 16,384$  gamma.

The plasma instrument (PLS) is designed to (1) establish the sources of Jovian plasmas, i.e., ionosphere, solar wind, and/or the satellites; (2) study the plasma interactions with the satellites; (3) investigate the role of plasmas as a source for the energetic charged particles of the radiation zones; (4) determine the nature of the equatorial current sheet; and (5) evaluate the roles of magnetic merging, corotational forces, and field-aligned currents in the dynamics of the Jovian magnetosphere. The instrument measures directional, differential energy spectra of positive ion and electron intensities over the energy range of 1 eV to 50 keV. Miniature mass spectrometers on the two electrostatic analyzers have sufficient resolution to identify ionized hydrogen, helium, sodium, potassium, and sulfur.

The energetic particles detector (EPD) will (1) measure energetic electrons and protons as a function of energy, angle, and particle species; (2) measure composition, distribution, and energy spectra of high-energy particles trapped in Jupiter's magnetic field; (3) determine the structure of the magnetosphere; (4) explore and determine the dominant plasma processes going on in the magnetosphere; (5) sample Io's wake through the magnetosphere during the Jupiter orbit insertion; and (6) study Jupiter's magnetotail. The effects of the Jovian satellites on the magnetospheric particle population (source or loss) will be studied, and the processes of Jupiter's plasma processes will be compared to those at Earth to achieve a more universal understanding of these phenomena under vastly different astrophysical conditions.

The EPD uses two separate solid-state detector telescopes, one that measures low-energy electrons and protons in Jupiter's magnetosphere and one that measures composition of energetic particles in the magnetosphere.

A plasma wave instrument (PWS) will investigate plasma waves generated inside Jupiter's magnetosphere as well as waves radiated from lightning discharges in the planet's atmosphere. It will provide measurements of the spectral characteristics of electric and magnetic fields in the frequency range from 5 Hz to 5.65 MHz. Wave-particle interactions are believed to play an important role in controlling the overall dynamics of the Jovian magnetosphere.

The PWS instrument consists of two electric dipole antennas 1.5 m in length, a search coil assembly, and a main electronics assembly. The PWS shares the high-rate telemetry used by the imaging instrument, similar to the PWS on the Voyager spacecraft.

The dust detection instrument (DDS) will determine size, speed, and charge of small particles--such as micrometeorites--near Jupiter and its satellites. The experiment will seek to identify sources of dust around Jupiter which comes from interplanetary space or the satellites themselves. The DDS will also investigate Jupiter's thin ring discovered by Voyager.

Negatively charged particles entering the dust sensor are detected by the charge which they induce when flying into the sensor. Charged or uncharged dust particles are detected by the plasma produced when they impact a sensor. The ions and electrons of this plasma are separated by an electric field and then accumulated by charge-sensitive amplifiers which deliver two pulses of opposite polarity. The sensor can determine the mass and speed of the impacting particle.

A celestial mechanics experiment will use spacecraft tracking data to measure the gravity fields of Jupiter and its satellites and to search for gravity waves

propagating through interstellar space. This experiment will also improve the ephemerides of Jupiter and its satellites and measure relativistic red shift in Jupiter's gravity field and the general relativistic time delay during solar conjunctions.

The scientific data of the radio science investigations are contained in measurement of the amplitude, frequency, phase, and polarization of the downlink signals at S- and X-band.

A radio propagation experiment will use radio signals from the Orbiter and Probe to study the structure of the atmospheres and ionospheres of Jupiter and its satellites. This experiment will measure the radii of Jupiter and the Galilean satellites at occultation points, map the Jovian magnetic environment, study turbulence and winds in Jupiter's atmosphere and ionosphere, investigate interactions between ionospheres and the magnetosphere, and, if geometric opportunities are available, conduct occultation investigations of Jupiter's ring and the Io torus.

#### Probe

The Galileo Probe (Fig. 6) carries seven science experiments. Table 2 lists the Probe science investigations.

The atmospheric structure instrument (ASI) will determine the pressure, temperature, and density of the Jovian atmosphere as the Probe descends to about the 10 bar level. The data should provide answers to questions about the existence of local turbulent patches in the atmosphere, the magnitudes of vertical winds, and the altitudes at which phase changes (from gas to liquid or vice versa) of the key cloud-forming constituents take place. It may be possible to detect gravity waves if stable regions in the atmosphere can be identified. The experiment uses three sensors to determine pressure, temperature, and acceleration. The exact time and number of observations taken by each sensor are carefully established so the fine structure of the atmosphere can be deduced.

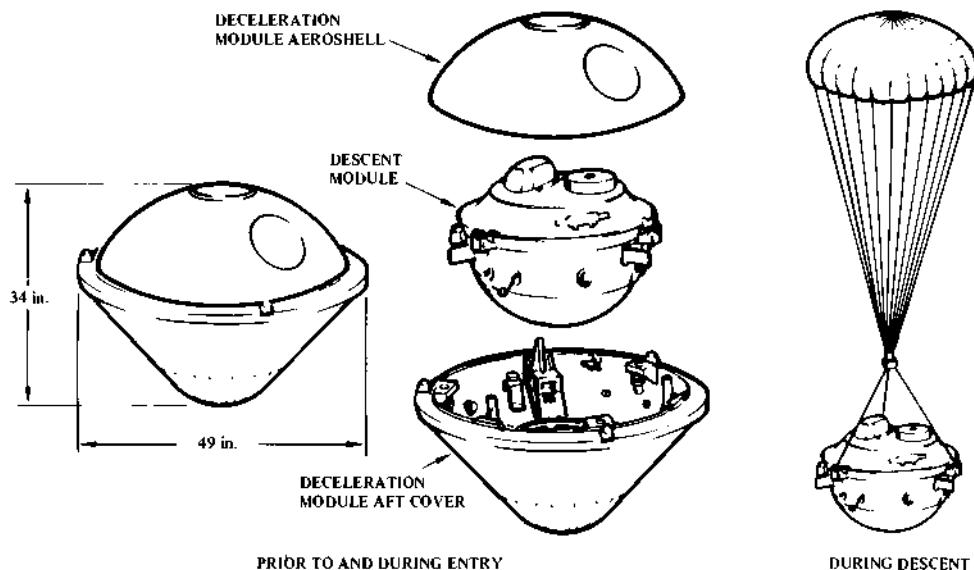


Fig. 6. Galileo Probe.

TABLE 2 Probe Science Investigations

Instrument Name	Principal Investigator/ Organization	Objectives
Atmospheric Structure Instrument (ASI)	A. Seiff/Ames Research Center	Determine state properties (temperature, pressure, density and molecular weight) of Jovian atmosphere.
Neutral Mass Spectrometer (NMS)	H.B. Niemann/Goddard Space Flight Center	Determine chemical and iso- topic composition of Jovian atmosphere.
Helium Abundance Detector (HAD)	U. von Zahn/Bonn University (FRG)	Perform precision determina- tion of helium abundance mea- surement in Jovian atmosphere.
Nephelometer (NEP)	B. Ragent/Ames Research Center	Determine microphysical char- acteristics (particle size distribution, number, density, and physical structure) of Jovian clouds.
Net Flux Radiometer (NFR)	R.W. Boese/Ames Research Center	Measure vertical distribution of net flux of solar energy and planetary emissions.
Lightning and Radio Emission/Energetic Particle Detector (LRD/EPI)	LRD - L.J. Lanzerotti/ Bell Labs/University of Florida and K. Rinnert/ Max Planck Institute (FRG)	Study lightning in the Jovian atmosphere and energetic particles near Jupiter.
	EPI - H. Fischer/ University of Kiel (FRG) and J. Mihalov/ Ames Research Center	
Radio Science	H. Howard/Stanford University	Study composition and struc- ture of Jovian atmosphere.

The nephelometer (NEP) will determine the location of cloud layers in Jupiter's atmosphere at altitude resolutions and intervals of less than 1 kilometer. It will also determine particle sizes and densities at each of these cloud layers, and determine sphericity or nonsphericity of the particles (implying liquid or solid particle states). Using forward and backscattering sampling units, the instrument measures the light-scattering characteristics of aerosols in the Jupiter atmosphere.

The helium abundance detector (HAD) will determine with about 0.1 percent precision the relative abundance of helium in the Jovian troposphere in the 2.5- to 10-bar pressure range. From this, the ratio of hydrogen to helium in Jupiter's atmosphere will be accurately determined. Knowledge of this ratio is important for theories of star and planet formation.

The net flux radiometer (NFR) will determine the energy flux being radiated inward and outward at each level in the atmosphere, detect cloud layer location

and infer composition, determine the relative importance of cloud layers and gaseous constituents for supplying opacity in the thermal infrared, and estimate the amounts of water vapor and ammonia. Knowledge of energy deposits in the atmosphere is important to an understanding of the driving force behind Jupiter's atmospheric dynamics. The NFR views the Jovian atmosphere in the infrared through six different detector filters, to determine solar and thermal flux, excess thermal radiation, water vapor, ammonia, or cloud presence.

The neutral mass spectrometer (NMS) will measure the composition of gases in Jupiter's atmosphere and their variations at different atmospheric levels from about 0.1 to 17 bars. Mixing ratios of helium will be determined to 1 percent accuracy and to 5 percent accuracy for water, methane, and ammonia. The isotopic ratio of Ne20 to Ne22 will be measured to 2 percent accuracy. All species with mass numbers 1-52 AMU will be measured, with periodic scans to 150 AMU, which includes krypton and xenon.

The NMS consists of a quadrupole mass spectrometer equipped with an electron impact ion source with two electron beam guns of variable kinetic energy and a secondary electron multiplier ion detector.

A lightning and radio emissions detector (LRD) will be coupled with an energetic particles instrument (EPI). The LRD will (1) verify the existence of lightning on Jupiter, (2) measure its characteristics, and (3) relate it to physical properties of the Jovian atmosphere. These properties include cloud electrification and charge separation mechanisms in the atmosphere; the scale size of cloud convection or turbulence; the rate and distribution of lightning activity; evidence for sources of heat and acoustic shock waves in the atmosphere which may affect chemical balance of minor constituents, production of cloud color, and production of precipitation; radio frequency noise levels; magnetic noise characteristics; and the inner Jovian magnetic field.

The EPI will measure electrons and protons as the Probe traverses the unexplored inner regions of Jupiter's radiation belts and will determine the spatial distribution of electrons and protons near the edges of the belts. The EPI will also study the ring material and its relation to trapped radiation and will take measurements at selected points in the magnetosphere as the Probe hurtles toward the planet. It will also study the energetic proton and electron flux impacting the upper Jovian atmosphere.

#### MISSION HISTORY

Galileo became an approved mission in 1977. The basic mission plan remains the same, but the launch date has been slipped twice to accommodate the availability of the Space Shuttle and an upper stage for insertion into the Jupiter-bound trajectory. Figure 7 shows the 1982, 1984, and 1985 mission trajectories.

At its inception in 1977, Galileo was scheduled to be launched in early 1982-- just a few months from now. This would have been a single Shuttle launch of the mated Orbiter and Probe, using the three-stage planetary configuration of the Inertial Upper Stage (IUS) to propel the spacecraft out of Earth orbit toward Mars. The 1982 mission plan called for a close (275 kilometer) flyby of Mars to gain heliocentric energy to reach Jupiter. This Mars gravity assist greatly reduced the required injection energy.

Development problems with the Shuttle necessitated the first mission delay to a 1984 launch. However, the poorer Mars alignment in 1984 necessitated a large ( $\sim 1$  km/s) maneuver near Mars. It was therefore necessary to split the spacecraft into two parts for the 1984 launch -- the Orbiter and Probe to be launched aboard

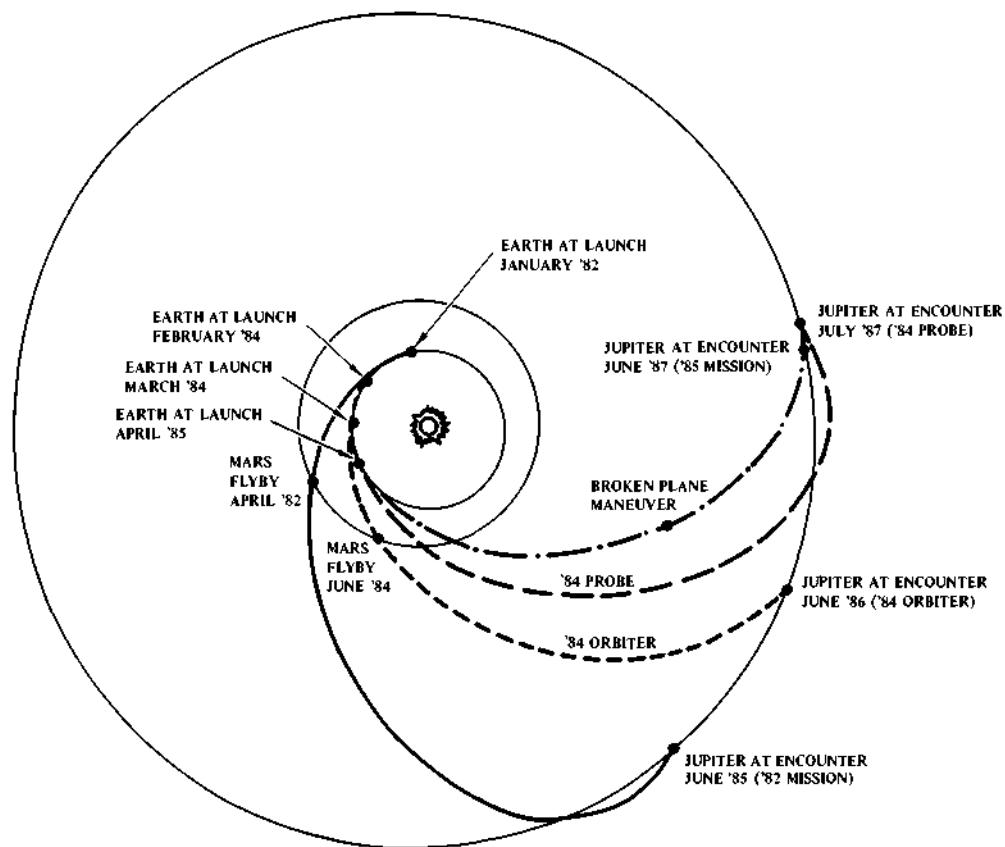


Fig. 7. Galileo interplanetary trajectories.

two separate Shuttles about one month apart. A Mars Flyby Module (MFM), designed to perform approximately the first 700 m/s of the near-Mars maneuver, was introduced in place of the Probe for the Orbiter launch. The remaining 300 m/s or so would be delivered by the Orbiter's retropropulsion module after jettisoning the MFM. The Mars gravity assist would have occurred about three months after launch. A Probe Carrier spacecraft was designed to deliver the Probe on a direct route to Jupiter and to relay the Probe's data to Earth.

With continuing problems in the development of the three-stage IUS, NASA decided in January 1981 to replace the IUS with a modified, wide-bodied Centaur upper stage, and moved the Galileo launch to 1985. The higher performance capability of the Centaur stage allows the Orbiter and Probe to once again be combined in a single launch, as originally planned for the 1982 mission, except that a direct trajectory with no Mars flyby is used. This is the course the Project is currently following. Figure 8 traces the changes in spacecraft configuration for the three launch schedules.

Meanwhile, numerous mission options, in case of further delays or budget cutbacks, have been prepared. Should the Centaur stage be unavailable for the 1985 launch opportunity, the preferred option is to launch with a two-stage IUS and a kick-stage on what is known as a ΔV-EGA trajectory (Fig. 9). This type of trajectory

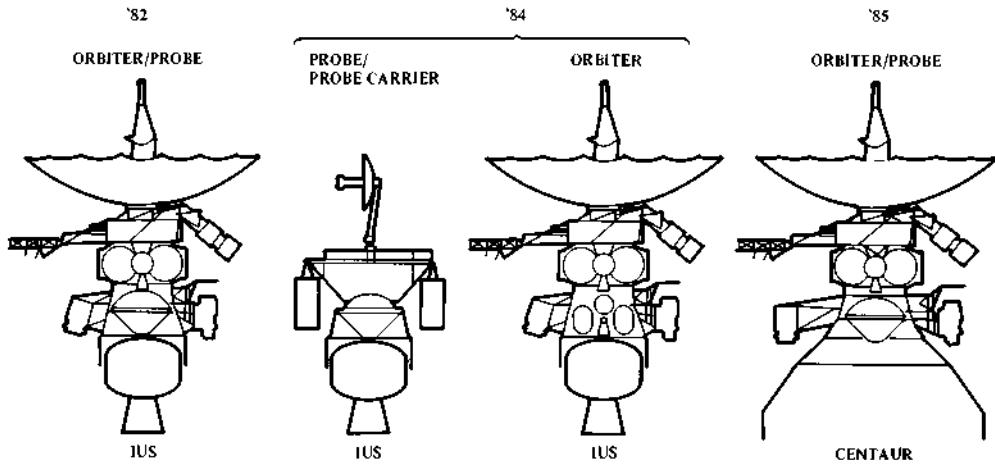


Fig. 8. Configuration contrasts.

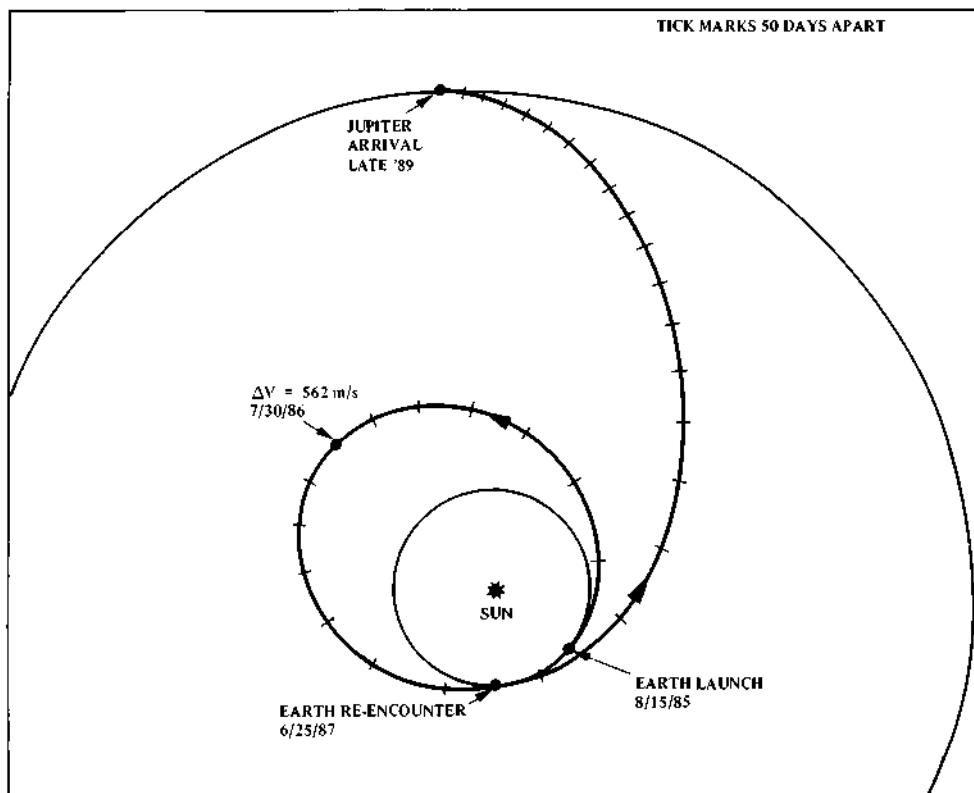


Fig. 9. Galileo 1985 ΔV-EGA.

uses a complete orbit around the Sun, a large  $\Delta V$  maneuver, and an Earth flyby gravity assist in order to establish the Earth to Jupiter trajectory. The Orbiter spacecraft would be launched separately into a two-year heliocentric orbit which would bring it back near Earth about two years later. About one year after launch, near aphelion, a large maneuver of about 0.5 km/s would be performed to set up the Earth re-encounter geometry such that Earth's gravity would accelerate the spacecraft about 4 km/s with respect to the Sun. This would send the spacecraft hurtling toward Jupiter for an arrival about two and a half years later, in late 1989. Meanwhile, the Probe would be launched in 1986, using a carrier spacecraft, into a direct trajectory. The Probe would arrive simultaneously with the Orbiter. This option obviously involves a large increase in cost, a substantial stretchout in the schedule, and a corresponding delay in science return from the mission.

#### CONCLUSION

The Galileo mission is an ambitious undertaking, yet one that is well worth the effort in terms of the return, both in our understanding of physical processes in our solar system and in our perceptions of our place in the universe. The development cost of \$539 million spread over eight years represents an investment of less than 30¢ per American per year.

The next step after initial flybys is an in-depth study. Galileo, with its Orbiter and Probe, is an unparalleled opportunity to study a "mini-solar system" in depth. With extended observations, direct sampling, close flybys, advanced instrumentation, improved navigational techniques, and the knowledge gained from the pathfinder missions, Galileo will provide the most comprehensive exploration of an outer planet ever accomplished. With its complex atmosphere, diverse moons, and far-reaching magnetic field, the Jupiter system is a pristine, untouched laboratory in which we can study processes and dynamics which formed our solar system and are still at work.

#### ACKNOWLEDGMENT

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The Galileo Project is managed for the National Aeronautics and Space Administration by the Jet Propulsion Laboratory, California Institute of Technology. JPL is also designing and building the Orbiter spacecraft. NASA's Ames Research Center, Mountain View, California, has management responsibility for the Probe spacecraft, which is being built by Hughes Aircraft Company, Culver City, California.

#### REFERENCES

- Farless, D. L. (August 1981). Galileo performance assessment report. Jet Propulsion Laboratory Document No. 1625-109 (internal document).
- Mitchell, R. T. (September 1980). Project Galileo mission design. Paper IAF-80-G-291 presented at the XXXI Congress of the International Astronautical Federation, Tokyo, Japan.
- Nock, K. T., and C. W. Uphoff (June 1979). Satellite aided orbit capture. Paper presented at the AAS/AIAA Astrodynamics Specialist Conference, Provincetown, Massachusetts.
- Yeates, C. M. (private communication).
- Yeates, C. M., and T. C. Clarke (in publication, 1981). The Galileo mission to Jupiter. *Astronomy*.

Galileo science requirements document (August 1980). Galileo Project Document No. 625-50, Rev. B, Jet Propulsion Laboratory (internal document).

*Galileo to Jupiter: Probing the Planet and Mapping Its Moons* (July 1979). Jet Propulsion Laboratory Document 400-15 (GPO Order No. 1979-691-547).

The retro-propulsion module (July 1981). *The Galileo Messenger*. Jet Propulsion Laboratory Document No. 1625-101, Issue 2 (internal document).

## NEW COMPLEX FOR IONOSPHERE-MAGNETO- SPHERE STUDY 'INTERCOSMOS— BULGARIA—1300'

K. Serafimov\*, I. Kutiev\*, S. Chapkunov\*,  
M. Gogoshev\*, V. Adasko\*\*, V. Balebanov\*\*, A. Bochev\*,  
Tz. Dachev\*, I. Ivanov\*, A. Josifjan\*\*, I. Podgorny\*\*  
and N. Sheremetjevsky\*\*

\*Central Laboratory for Space Research, Sofia, Bulgaria

\*\*Institute of Space Research, Moscow, USSR

### ABSTRACT

The paper gives an overall review on the original or modified instrumentation designed for the Intercosmos-Bulgaria-1300 experiment. The main scientific fields of research are also discussed. Fundamental space physics problems related with insufficient knowledge or instrumental reliability are also described.

### KEYWORDS

Intercosmos-Bulgaria 1300 experiment, ionospheric-magnetospheric interactions, spaceborne instrumentation.

### INTRODUCTION

The recent decade advance in space research largely contribute to the understanding that without consideration of the earth magnetospheric effects not only the ionospheric balance studies but also the interpretation of the basic structural ionospheric parameters is possible. For example, it was revealed that such processes as the maintenance of the nighttime ionization is due to plasmasphere fluxes. The knowledge on the polar region was enriched. The magnetosphere-ionosphere coupling there is not only of direct manifestation but is also related to large amount of energy investment, measurable with the energy for instance produced by the UV solar radiation in the upper atmosphere during strong magnetic storms. It was discovered that this energy is not absorbed in the polar region only since part of it is transported by various wave-type processes and horizontal wind into the atmosphere at mid and low latitudes. It was found that the large scale magnetospheric electric fields could penetrate not only the polar ionosphere but also the lower latitude regions inducing strong plasma drift and affecting largely the ionospheric dynamics.

The equatorial ionosphere though remoted from the poles expresses strong coupling with the magnetosphere. It is evident that the ring

current in the geomagnetic equator plane at a distance of 4-6 earth radii largely affects the ionosphere. Multiple satellite and rocket borne measurements reveal the fact that the main ring current carriers are the protons, due to charge-exchange processes they may leave the ring current and penetrate the equatorial ionosphere as high energy neutral particles.

The magnetospheric-ionospheric interactions manifest the reverse dependance also: magnetospheric filling and feeding from the ionosphere. Particularly well expressed is this process in the plasma-sphere which is filled by upward hydrogen and helium ion fluxes in the daytime ionosphere. Permanent hydrogen ion diffusion flux from the polar ionosphere upward along the open field lines, i.e. the polar wind also fills the plasma populations in the magnetospheric tail.

The Bulgarian scientists in their capacity of coauthors in the project here discussed, have already established routine patterns and achieved success in the studies of ionosphere and magnetosphere-ionosphere interactions. The participation of the Bulgarian scientists in four satellite and four high altitude rocket experiments within the framework of the Intercosmos program has resulted in significant experience on direct measurement techniques and modes in particular plasma probes and airglow photometers. Extensive investigations in the global scale ion distribution, longitudinal variations of the equatorial region - F<sub>1</sub>, anomaly and neutral wind-plasma interactions have been carried out. The results obtained together with the scientific interests aimed at the detailed interpretation of the processes and phenomena already observed, provided the back-grounds for the Bulgaria-1300 satellite project. The hardware designed for this program is much more sofisticated and comprises new techniques basec on new principal approach.

#### SCIENTIFIC OBJECTIVES

The main objective of Bulgaria-1300 program is to study the ionospheric-magnetospheric interactions. The coupling between the two media of different physical properties is of complex nature and its study is both of fundamental scientific and application importance. We have on one hand the magnetospheric plasma that practically is collisionless and on the other - the ionospheric low temperature plasma which is affected by the neutral atmospheric gas. The coupling between these two media is established mainly through the earth magnetic field, along the highly conductive field lines.

The electric nature of these interactions is the major feature in the magnetospheric-ionospheric relationship. The magnetospheric electric field is projected along the highly conductive field lines in the polar ionosphere. Due to the electric field, the magnetospheric plasma together with the 'infrozen' magnetic field lines drifts from the tail to the sun. Reaching the frontal magnetospheric part the plasma flows back along its outer sides thus producing circular motions. This drift is related to the ionospheric generation of two vortex systems also. When due to certain factors, this field line conductivity decreases, current fluxes to compensate the potential differences appear. It is known that often current fluxes to the magnetosphere are observed also. The current circuitry close through the ionosphere

where the transverse electric conductivity is rather large. When the currents attain some critical value, plasma instabilities are generated and the ionosphere-magnetosphere link is broken down. Locally accelerated charged particles (often electrons) immerge in the potential drop thus generated along the field lines. They produce heating and oscillation in dense atmospheric layers. Certain aurora forms are related to this local accelerations.

Energy and mass transport are both to be observed between magnetosphere and ionosphere. The energy transport is effected from magnetosphere to ionosphere by energized charged particles and electromagnetic waves. This energy dissipates in the ionosphere and activates multiple large scale processes. On the other hand, the ionosphere is the main source of lowenergy ions and electrons to the magnetosphere. Affected by various processes in the magnetosphere, these particles are energized and envolved in the largescale magnetospheric processes.

The problems of magnetospheric-ionospheric coupling still to be revealed are multiple. We can mention some, for which the data expected in this project could be useful:-control of interplanetary magnetic field (IMP) on polar F-layer convection; There is evidence to consider that under a given value of the northern IMF component, the convection may change direction or may significantly be transformed. This problem relates to the fundamental problem of the solar-wind - magnetospheric relationship.

-Regularities in the appearance of localized regions of electric fields parallel to the field lines. There is evidence that these fields generate in the regions of intensive current fluxes and are due to the plasma instability excitation or to the double electric layers.

-Excitation mechanisms on stable auroral red arcs. How the energy of the ring current dissipates in the ionosphere. What is the relationship between the discrete forms appearance of polar aurora and the generation of parallel electric fields. Is the airglow excited by the local acceleration of the electric field charged particles. The global scale neutral wind system in the F-region. What is the type of relationship with the ion species.

## MEASUREMENTS

### Measurement Parameters

The most important feature of Bulgaria-1300 project is the complex nature of the measurements. The scientific task realized with the project is not aimed at the morphological study of phenomena and processes. The large scope of simultaneous in development phenomena and their interactions will be investigated. Therefore, equipment for measurement of a large number of physical parameters will be installed onboard the satellite. These parameters could be divided into four main groups: plasma, highenergy, electric and magnetical, optical. The measurement of these parameters will provide for evidence on relatively complete description of the abovementioned ionospheric-magnetospheric interactions.

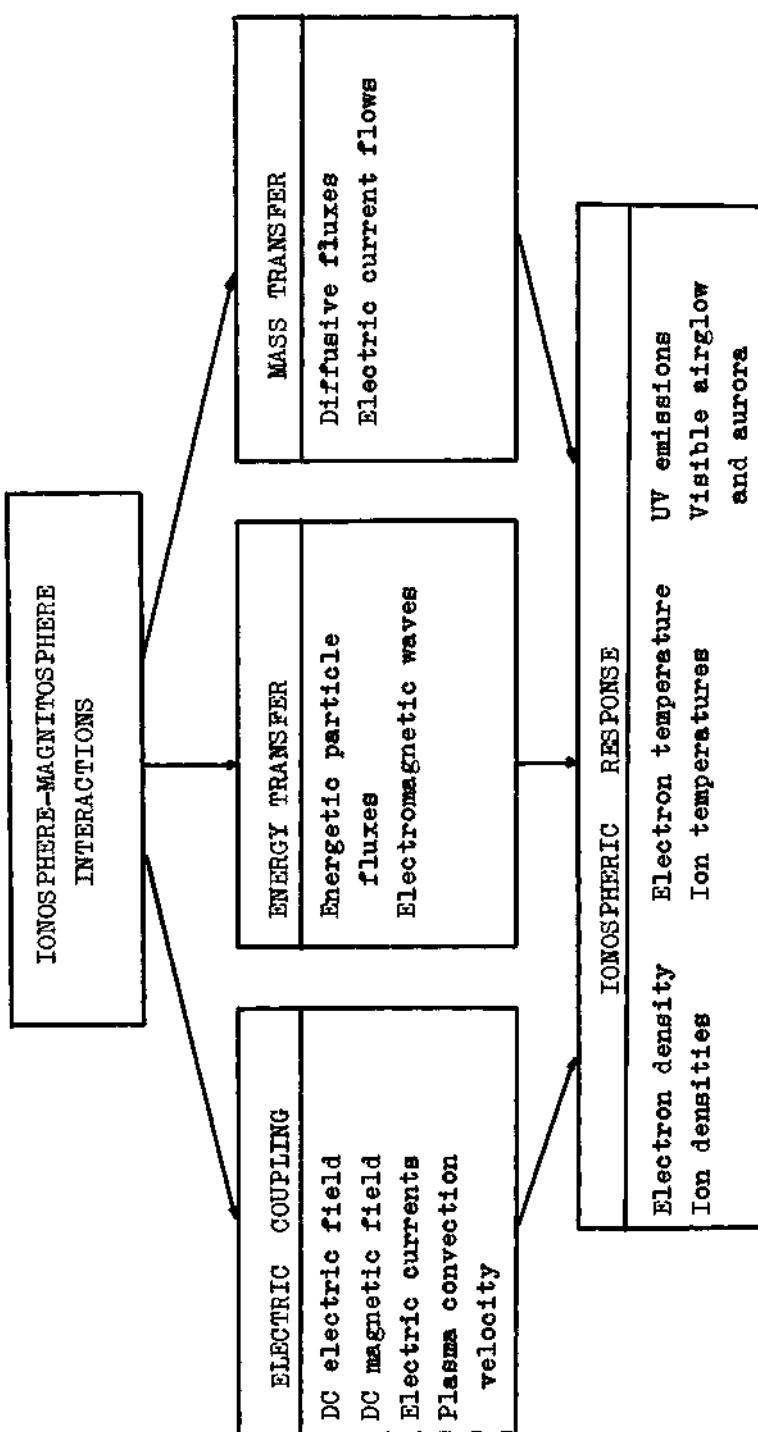


FIG 1

The first group includes: cold plasma parameters: electron and ion density and temperature, energy distribution and bulk ion velocity. The second group incorporates: measurements on highenergy electron and ion distribution, their mass composition and the pitch-angular distribution. The third group provides direct information on the electric and electromagnetic coupling of the ionosphere and the magnetosphere. The three components of the natural electric fileds, the three components of the earth magnetic field and the spectrum of natural electromagnetic emissions within the range of 0.2 Hz - 16 kHz are incorporated also.

The high measurement accuracy is a specific feature of these experiments: for the electric field the accuracy is to be 0.1 % while for the magnetic field it is about 0.002 %. The measurement of the optic parameters provide data on the fundamental aeronomic processes in the F-region, on effects related with proton and electron discipation. The UV emission is to be observed also within the range of 115 - 260 nm and the visible spectrum in 6 bands. The spatial picture of the emitting regions below the satellite will be obtained for the red atomic oxygen line 6300 Å.

The measured parameters and the relationship types to be studied based on the measurements are given in Figure 1. The instruments, measurement parameters and investigators are enlisted in Table 1.

TABLE 1

Instrument and Parameters	Investigators
Analyzer for lowenergy protons and electrons Array of electrostatic analyzers Energy range: 0.2 to 15 keV/unit charge, up to 16 programmable steps; Energy resolution: $E/E = 10\% ; 20\% ; 30\%$ . Flux range: $10^4$ to $10^9$ ( $\text{cm}^2 \cdot \text{sr} \cdot \text{keV} \cdot \text{s}$ ) $^{-1}$ Angular resolution: $7^\circ \times 24^\circ$ along three orthogonal axes.	Tz. Dachev Yu. Matviychuk
Magnetometer Three axes fluxgate Range: $\pm 64000$ nT resolution: $\pm 2$ nT.	A. Bochev
Ion Energy Mass Analyzer Two electrostatic analyzers with magnetic ion mass spectrometer Energy range: 1 to 30 eV/unit charge; 0.25 to 15 keV/unit charge; Mass range: 1 to 64 amu, 1 to 54 amu; Energy resolution: 5.5%; 7%; Angular resolution: $10^\circ \times 60^\circ ; 40^\circ \times 5^\circ$ ; Flux range: $10^5$ to $10^{10}$ ( $\text{cm}^2 \cdot \text{sr} \cdot \text{eV} \cdot \text{s}$ ) $^{-1}$ ; $5 \times 10^5$ to $10^9$ ( $\text{cm}^2 \cdot \text{sr} \cdot \text{keV} \cdot \text{s}$ ) $^{-1}$ .	P. Nenovski R. Koleva
Proton - 1 UV differential and integral photometer Intensity range: 80 R to 80 kR Wavelengths: 1100 to 2600 Å	M. Gogoshev

Angular resolution: cone angle  $\pm 15^\circ$ ;  
 Longitude dependent orientation to nadir

Electron temperature Instrument	V-Markov
Double probe of spherical sensors Temperature range: 500 to 6000°K; Spacecraft potential: -30 to +50 V .	D.Teodosiev
Ion Langmuir Probe	T.Ivanova
Spherical electrostatic probe Ion temperature: 1000 to 5000°K; Ion density: $10^3$ to $10^6$ cm $^{-3}$ Density irregularities	T.Samardzhiev
Electron Langmuir probe	T.Ivanova
Cylindrical electrostatic probe electron temperature: 1000 to 10000°K; electron density: $5 \times 10^2$ to $3 \times 10^5$ cm $^{-3}$ .	K.Georgieva
Visible photometer Wavelengths: 6300 Å , 5577 Å , 4278 Å , 6563 Å ; Sensitivity: 10 R to 100 kR.	Tz.Gogosheva
High Energy Proton Instrument Energy Analyzer on Solid State Detectors Energy range: 50 keV, to 1 MeV, 4 steps; Flux range:	K.Kazakov
Vector Electric Field Instrument	G.Stanev
Three-axial Probe of Spherical Sensors dc electric field:range: +1 V/m resolution: 2.5 mV/m variational electric field: 0.2 Hz to 16 kHz; 0.5 mV/m Hz; range: 80 dB electromagnetic waves: 0.2 to 10 Hz; spacecraft potential: -10 to +2 V.	D.Teodosiev
Ion Drift Meter	L.Bankov
Retarding Potential Analyzer of Segment Collectors B.Kiröv ion velocity: 0.1 to 1.5 km/s; ion density: $10^2$ to $10^6$ cm $^{-3}$ ; ion temperature: 300 to 10000° K; plasma irregularities: 0.1 to 100% ; photoelectron energy: 0.1 to 30 eV; mass range: 1 to 32 amu.	

---

The parameters providing basic information of the ionospheric situation are rather important in this study. Considering the fact that for multiple interactions we judge upon their effects on the ionosphere the importance of its nondisturbed parameters and their variations are of particular importance. There follow the parameters providing

information on the electric nature of the relationship, such as convection characteristics, current flows and natural electric fields. The energy transfer will be studied by the energetic particle fluxes, mainly from magnetosphere and ionosphere and by the dissipation of various electromagnetic waves. The mass transport will be directly extracted in measuring the bulk plasma velocity and indirectly - by the height distribution variations from the hydrostatic equilibrium, generating the diffusion fluxes.

### Selection of Orbit

When the whole complex of scientific tasks and measurements parameters is defined; it is of particular importance to select the most effective type for the satellite and its orbit.

The three-axial satellite stabilization as one of the axes is permanently fixed to the earth center is the most reasonable solution of this requirement. Thus most precise orientation of the plasma measuring instruments will be attained along the ion flow, together with fixed position of the optical instrumentation at the nadir. The measurements enlisted, require satellite stabilization within  $1 - 2^\circ$  along each axis. Additional requirement of precise information on the orientation, imposes the ion bulk velocity measurements. The orientation fixed precision is  $0.5^\circ$  providing possible measurements on the ion drift of 100 m/s, and higher. Many booms have to be extended from the satellite to measure the electric and magnetic fields. The high precision reliability of the plasma experiments require copotentiality between the satellite and the other surfaces. All these requirements were met with the Soviet satellite Meteor type by many modifications and approbations.

The study of the magnetospheric-ionospheric interactions within the project scope will be performed on two main lines: direct measurements of the characteristics of the interactions and measurements on their effects in the ionosphere. Thus except for measurements on the parameters of the direct coupling between ionosphere and magnetosphere, it is necessary to observe also the ionospheric behaviour and hence to extract the ionospheric-magnetospheric interaction effects. The first type of parameters include: electric and magnetic fields, energetic fluxes, electromagnetic emissions. While the ionospheric effects are observed through the optical emissions, the electrodynamic plasma drift, its density and temperature. The satellite orbit must be high in the ionosphere to avoid the influence of the neutral atmosphere and the specific ionospheric photo-chemical processes. In other words the orbit must be at the highest orbit where it is still possible to observe the ionospheric behaviour. With reference to the ion composition, the main ionospheric ion - the atomic oxygen, must be reliably recorded along the orbit. In addition the orbit has to be selected above the main emitting regions of the ionosphere such as the aurora, SAR arcs, tropical arcs, etc. All these requirements are met with a circular orbit of 900 - 1000 km of height. The circular orbit eliminates the altitudinal variations of the ionospheric parameters, which are difficult to be extracted from the spatial-temporal variations due to the outer processes in discussion.

### Scientific Payload

The instrumentation of Intercosmos-Bulgaria 13000 satellite is designed to implement the scientific tasks of the project. Table 1 as given before shows the instrumentation, the measurement parameters and ranges and the individual experiments.

Based on the measurement principle, the instruments may be divided into four equipment complexes: I) plasma, II) high-energy, III) electromagnetic, IV) optical.

The plasma complex incorporates 4 probe instruments and 1 mass-spectrometer. The probes measure the basic ion parameters (P6 and ID) and the electron parameters (P7 and DIET) in the plasma components. The methodology implemented in each individual instrument provides higher accuracy of obtaining the given parameters and possibility to select the other parameters. For instance P6 measures with higher precision the Maxwell ion distribution and hence the temperature of the predominant ion in the medium. ID measures with a good accuracy the ion composition and density. The DIET instrument is designed to measure the electron temperature with high precision and the P7 device provides accurate information on the electron density. The combination of the data obtained from these instruments will permit the precise parametric determination of the environmental thermal plasma.

The AME I instrument is a combination of electrostatic analyzer and a permanent magnet ensuring simultaneous energy and mass analysis of the thermal plasma. The ID equipment is designed to measure the parallel and the perpendicular ion flux velocity components, with reference to the earth corotation at velocities exceeding 100 m/s. This is the principle instrument of F-region convection measurements.

The highenergy complex incorporates the ANEPF and Proton instruments. The first measures electric and proton fluxes of energies 0.2 - 15 keV in three inter-perpendicular directions. At fixed energy of 1 keV, the spatial-temporal flux radiations are measured with high resolution.

The Proton instrument performs measurements on proton fluxes from 50 keV to 1 MeV. Certain modes of operation on AME I and ID equipment also should be mentioned in the highenergy complex. AME I measures the energy spectrum and the mass ion composition within the range of 0.25 - 1.5 keV and the ID performs measurements on the photoelectrons with energies of 1 - 30 eV.

The electromagnetic complex incorporates the IESP and the IMAP instruments. The first one measures the dc electric fields components on double probe principle with area of observation around 10 m. The ac electric fields are measured in the range of 0.2 Hz - 16 kHz.

The IMAP instrument is a three-component fluxgate magnetometer of magnetic field compensation of  $\pm 100$  gamma and resolution of 2 gamma.

The optical complex includes: scan photometer EMO-5, vacuum UV spectrometer Foton-1. The first measures determined spectral strips of emissions related to the aeronomy processes in the ionosphere. EMO-5 incorporates scanning unit which enables the observation of the spatial picture of intensity distribution in the red oxygen line. The Foton-1

instrument scans the UV spectrum between 110 - 260 nm providing information for the aeronomical studies and identification of high energy impact ionization processes.

#### ACKNOWLEDGEMENT

The Authors would like to express their gratitude to the multiple space physics and space engineering Bulgarian and Soviet teams, participating actively in both planned and active phases of the Bulgaria-1300 Project.

## TEN YEARS OF THE INTERNATIONAL REVIEW MEETINGS ON COMMUNICATION WITH EXTRATERRESTRIAL INTELLIGENCE (CETI)

R. Pesek\* and J. Billingham\*\*

\*Czechoslovak Academy of Sciences, Prague, Czechoslovakia

\*\*NASA Ames Research Center, Moffett Field, California, USA

### ABSTRACT

The only regularly scheduled international meetings on CETI and SETI are those held annually by the International Academy of Astronautics. After ten years it is interesting to recall the development of CETI ideas; i.e., to discuss different methods of the detection of planets and of interstellar communication, and to observe the growing number of seminars, conferences and books devoted to CETI.

### KEYWORDS

Communication with extraterrestrial intelligence (CETI); search for extraterrestrial intelligence (SETI); radio frequency interference; life in the universe; observational strategies; planetary detection; multichannel spectrum analyzer; signal processor.

### INTRODUCTION

The topic of communication with extraterrestrial intelligence (CETI) was introduced into the IAA by one of us (Pesek) as long ago as 1965. As a result, a CETI Study Group was formed and met together at the Madrid meeting of the IAA in 1966. After some years of planning, the first CETI Review Session of the IAA, with five papers presented, was held during the Vienna Congress of the IAF in 1972. At the fifth CETI Review Session in 1976, one of us (Pesek) gave a review of all CETI activities in the IAA from 1965 to 1976 (see Billingham and Pesek, 1979a). Now that the tenth CETI Review Sessions are being held at the 1981 IAF Congress in Rome, it seems appropriate to review the events of the last five years, 1977-1981, with regard to CETI activities in the IAA and elsewhere, and to discuss the outlook for CETI in the future.

### CETI IN THE IAA

The IAA CETI Review Sessions have continued to be a regular part of the technical proceedings of the annual Congresses of the IAF. In 1977 the number of CETI papers increased to the point where it was necessary to have two half-day CETI

Review Sessions instead of the one session of previous years. There have been two CETI Review Sessions each year since.

In 1978, we agreed, in concert with the IAA and the Editor-in-Chief of *Acta Astronautica*, to prepare a Special Issue of the Journal on CETI. The Special Issue was duly published (Billingham and Pesek, 1979b). It contained also the first section of a U.S. report entitled, "The Search for Extraterrestrial Intelligence - SETI" (Morrison, Billingham, and Wolfe, 1977), and a Soviet paper on a design concept for a large space radiotelescope (Buyukas, 1979) which had been read at another session of the 1977 IAF meeting. The full table of contents of the Special Issue of *Acta Astronautica* is given in an appendix to this paper (Appendix I).

The present membership of the CETI Committee of the IAA is:

R. Pesek	Czechoslovakia	Chairman
J. Billingham	U.S.A.	Co-Chairman
C. Clemesdon	Sweden	
R. Edelson	U.S.A.	
V. Gogosov	U.S.S.R.	
N. Kardashev	U.S.S.R.	
A. Lawton	U.K.	
G. Marx	Hungary	
C. Ponnampерuma	U.S.A.	
M. Subotowicz	Poland	

In 1976, the CETI Committee had become very concerned about the rapidly increasing use of microwave frequency space by a wide variety of transmissions from many nations. Such transmissions can seriously interfere with SETI searches. The CETI Committee adopted a resolution on "Radio Frequency Protection for a Search for Extraterrestrial Intelligence", and forwarded it to the President of the IAA, asking that it be transmitted by the IAF to the International Telecommunications Union for consideration at the World Administrative Radio Conference in Geneva in 1979. In 1977, the IAA and the IAF approved the resolution and it was forwarded to the ITU. The resolution is included here (Appendix II) because its wording remains valid today. The ITU was also asked by twenty nations to consider allocation of frequency space for SETI listening programs and the subject was considered by the CCIR and taken up at the WARC meeting in Geneva. Unfortunately no space in the spectrum was approved for SETI at the WARC meeting. Although the current allocations do now include a "Footnote" to the effect that SETI activities will be occurring in the water hole and other selected regions of the microwave window. So SETI continues to face an extremely severe problem. The degree of interference will only really become apparent after large scale SETI multichannel spectrum analyzers come into use over the next decade. It is clear that all investigators should continue to argue as strongly as possible for the establishment of some protected regions for SETI. It is not too early to begin preparations for again seeking protection for SETI at the 1999 WARC meeting. The only alternative is to take SETI antennas out into space where they may be shielded from terrestrial radio frequency interference.

In 1982 the United Nations will hold a meeting on the Peaceful Uses of Outer Space, with the title of UNISPACE 82. Various international societies have been asked to contribute non-technical papers on space activitiers as background material for the delegates to the Conference. COSPAR is dealing primarily with the space sciences, and the IAF with space technology. COSPAR was asked to provide the background paper for SETI. They in turn asked IAF to prepare the SETI paper, and IAF asked the IAA CETI Committee chairman to undertake the task.

This we did, with the considerable help of associate contributors from many nations. The co-authors are:

R. Pesek	Czechoslovakia
J. Billingham	U.S.A.
H. Djojodihardjo	Indonesia
F. Drake	U.S.A.
R. Edelson	U.S.A.
A. Lawton	U.K.
G. Marx	Hungary
M. Morimoto	Japan
Y. Pal	India
C. Ponnampерuma	U.S.A.
M. Rees	U.K.
J. Ribes	France
M. Subotowicz	Poland

The report is in the hands of COSPAR. It is too long to include as an appendix to this paper, but it may be obtained from either one of us on request.

The IAF was also asked to prepare a brief background paper on the technology of SETI for the direct IAF input to the United Nations. This was prepared by us and is included here as Appendix III. It is included in the "Review and Projection of Space Technology" - A Background Paper Prepared for the United Nations, Outer Space Affairs Division, by the IAF, January 31, 1981.

#### CETI OUTSIDE THE IAA

In 1976 the Viking spacecraft landed on Mars and began returning voluminous streams of data about the composition of the Martian surface. At first it looked as though the three biology experiments, designed to detect microbial metabolism, were yielding positive results. After much analysis, it is now believed that the results can largely be explained by some unusual chemical characteristics of the surface soil. It is not possible to say of course, that life is not present on Mars, but strong proponents for the presence of life are hard to find. For an excellent summary of the Viking biology experiments see Klein (1979).

In thinking about extraterrestrial intelligent life, our attention in the post-Viking era is inevitably focussing more and more on life outside the solar system. More books and papers are beginning to appear on the topic (Morrison, Billingham and Wolfe, 1977; Papagiannis, 1980; Goldsmith and Owen, 1980). In the summer of 1979 a meeting on Life in the Universe was held at the Ames Research Center in California. It encompassed all aspects of cosmic evolution in a sequential order, culminating with a paper on SETI by Wolfe and others. The proceedings will be published late in 1981 by the MIT Press.

Another possible method of detecting extraterrestrial life is receiving considerable attention in the U.S.A. This is the development of techniques for the detection of extrasolar planetary systems. One approach is reported in Project Orion (Black, 1980), a NASA Ames Research Center Summer Study to design a ground based optical interferometer for discovering extrasolar planetary systems by astrometric techniques. It is conceivable that direct measurements of the atmospheres of extrasolar planets in the infrared may in the next several decades reveal the presence of oxygen or other gases associated with life. The studies are under the jurisdiction of David Black at the Ames Research Center.

In the SETI field, searches continue at a steady pace. A dozen or so have taken place over the last five years, mostly by U.S. and U.S.S.R. observers. A complete list of SETI searches since 1960 has been compiled by Tarter (1981). In the summer of 1977 the SETI Report, prepared for the NASA Ames Research Center by a Science Workshop under the chairmanship of Philip Morrison, appeared as NASA Special Publication No. 419. The Conclusions of the report were:

1. It is both timely and feasible to begin a serious search for extraterrestrial intelligence.
2. A significant SETI program with substantial potential secondary benefits can be undertaken with only modest resources.
3. Large systems of great capability can be built if needed.
4. SETI is intrinsically an international endeavor in which the United States can take a lead.

The report recommended the initiation of SETI now.

In the Soviet Union, SETI activities continue. Troitskii and others (1979) and Gindilis and others (1979) each reported the results of SETI searches in the Special Issue of Acta Astronautica on CETI. Buyakas and many other distinguished authors (1979) described a concept for an "Infinitely expandable space radio telescope" in the same journal. An experimental 10 meter telescope was carried up to the Soviet Salyut vehicle, erected outside the vehicle, and, according to Pravda, used for radioastronomy, SETI, and other observational studies.

There have been two meetings dealing mainly or exclusively with SETI. One at the Jet Propulsion Laboratory in the U.S.A. in 1977, and one consisting of joint sessions of Commissions 16, 40 and 44 of the International Astronomical Union at the IAU General Assembly in 1979 (Papagiannis, 1980). And a symposium entitled "SETI-81" is to be held in Tallin in December of 1981, sponsored by the Scientific Council on Radio Astronomy of the USSR Academy of Sciences and the Astronomical Society of Estonian SSR.

In the USA a NASA SETI Program Team has been established under the direction of one of us, with members from Ames Research Center, the Jet Propulsion Laboratory, and Stanford University. The team is working closely on the research and development of sophisticated multichannel spectrum analyzers and signal processors, and will evaluate these devices on telescopes of the Deep Space Network at Goldstone and Arecibo. Dr. Michael Klein is Chief of the SETI Office at Jet Propulsion Laboratory.

#### THE OUTLOOK FOR CETI

It seems likely that CETI will continue to develop more sophisticated strategies, techniques and instrumentation as the years go by, so that it will become somewhat easier to tackle the extremely difficult task of exploring the multi-dimensional search space. The U.S.A. and U.S.S.R. remain the principal actors on the stage, but there is some interest in carrying out SETI observations in France, Holland and the Federal Republic of Germany. There are as yet no formal arrangements for collaboration between nations, but the level of interest continues to rise. At some appropriate time formal collaborations should be explored, since they could have significant benefits.

In the meantime the scientific, technical and popular interest in SETI continues

to increase. In the U.S.A. there is now a journal devoted entirely to SETI. Titled, "Cosmic Search", it was started in 1978 by Kraus and Dixon of Ohio State University and is written so as to be understood by the intelligent layman. Courses and lectures on SETI are constantly on the increase, and there now exists in the U.S. a CETI Foundation.

The most disturbing problem is radio frequency interference. It may be that SETI is destined to go into space to escape. If that should be the case, in future years it is most proper that the CETI Review Sessions and Committee activities remain firmly entrenched in the IAA, since astronautics will then actually be involved.

In the meantime these sessions continue to provide the only regular international forum for the exchange of scientific and technical data on all aspects of SETI.

#### REFERENCES

The references given below include all those mentioned in this article. Following that is "An Introductory SETI Bibliography", which is a listing of a variety of SETI readings of a general and specific nature, and includes some of the references cited in this article. The Introduction has been prepared, with annotations, by Dr. Charles Seeger of San Francisco State University.

Billingham, J., and R. Pesek (1979a). Communication with Extraterrestrial Intelligence. Acta Astron., 6,1-2, 3-9. Pergamon Press.

Billingham, J., and R. Pesek (1979b). Communication with Extraterrestrial Intelligence. Acta Astron., 6,1-2. Pergamon Press.

Black, D. C. (ed.) (1980). Project Orion: A Design Study of a System for Detecting Extrasolar Planets. Special Pub. SP-436, NASA Ames Research Center, Moffett Field, Calif.

Buyakas, V. I., Yu. I. Danilov, G. A. Dolgopolov, K. P. Feoktistov, L. A. Gorshkov, A. S. Gvamichava, N. S. Kardashev, V. V. Klimashin, V. I. Komarov, N. P. Melnikov, G. S. Narimanov, O. F. Prilutsky, A. S. Pshennikov, V. G. Rodin, V. A. Rudakov, R. Z. Sagdeev, A. I. Savin, Yu. P. Semenov, I. S. Shklovskii, A. G. Sokolov, G. S. Tsarevsky, V. I. Usyukin and M. B. Zakson (1979). An infinitely expandable space telescope. Acta Astron., 6,1-2, 175-201. Pergamon Press.

Gindilis, L. M., N. S. Kardashev, V. A. Soglasnov, E. E. Spangenberg, V. S. Etkin and V. G. Mirovskii (1979). Search for signals from extraterrestrial civilizations by the method of synchronous dispersion reception. Acta Astron., 6,1-2, 95-104. Pergamon Press.

Goldsmith, D., and T. C. Owen (1980). The Search for Life in the Universe. Benjamin/Cummings Pub. Co.

Klein, H. P. (1979). The Viking Mission and the Search for Life on Mars. Rev. Geophys. and Space Physics, 17, 1655-1662.

Morrison, P., J. Billingham and J. Wolfe (1977). The Search for Extraterrestrial Intelligence: SETI. Special Pub. SP-419, NASA Ames Research Center, Moffett Field, Calif.

Papagiannis, M. D. (1980). Strategies for the Search for Life in the Universe. Astrophys. & Space Sci. Library, 83. E. Reidel Pub. Co.

Tarter, J. C. (1981). The Search for Extraterrestrial Intelligence, in Encyclopaedia of the 21st Century, Macmillan Press.

Troitskii, V. S., A. M. Starodubtsev and L. N. Bondar (1979). Search for radio emissions from extraterrestrial civilizations. Acta Astron., 6,1-2, 81-94. Pergamon Press.

\* \* \*

#### AN INTRODUCTORY SETI BIBLIOGRAPHY

##### GENERAL READING:

Billingham, J., (Ed.), Life in the Universe. NASA Conference Publication 2156 (1981); also, MIT Press, Boston, MA 02142, (1981)

Proceedings of the Life in the Universe Conference held at NASA Ames Research Center, June 19th-20th, 1979, "A meeting to explore prospects for research into the nature and distribution of life in the Universe." Thirty-three contributions by specialists from almost as many disciplines, block out much of what is surmised and the little that seems surely known about the history of the Universe, from the Big Bang to the presence of intelligent, technological life. Along the way, major current uncertainties and promising research directions are highlighted in this unique volume.

Bracewell, R.N., The Galactic Club: Intelligent Life in Outer Space. W.H. Freeman and Company (1974)

Highly readable synopsis of major unanswered questions about the nature and distribution of intelligent life, including suggestions for contact by interstellar probes.

Cosmic Search: "A magazine about Space, the Future and the Search for Intelligent Life beyond the Earth, presented in a popular, responsible manner." Published quarterly by Cosmic Quest, Inc., Box 293, Delaware, OH 43015

Goldsmith, D., (editor), The Quest for Extraterrestrial Life: a Book of Readings. University Science Books, 20 Edgehill Road, Mill Valley, CA 94941

A collection of articles dealing with many aspects of SETI and exobiology which spans the centuries from 70 B.C. to the present. The arrangement and the notes by the editor clearly demonstrate what is well understood and what remains controversial or unknowable about these subjects.

Goldsmith, D., and Owen, T., The Search for Life in the Universe. The Benjamin-Cummings Publishing Co., Inc., Menlo Park, CA 94025

A very readable introductory textbook which presents the subject of astro-

nomy from the point of view of what it may tell us about the prospects for finding life elsewhere in the Universe. Extremely entertaining and accurate and it includes more biology and critical humor than is the norm for this type of book. A successful contemporary update to Sagan and Shklovskii.

Kraus, J.D., Big Ear. Cygnus-Quasar Books, Powell, Ohio (1976)

"Big Ear is a personal, behind the scenes account of astronomers, engineers, inventors — humans all — their successes and failures. It is a story about the steel and aluminum structures we have raised to probe the cosmos and of our attempt to answer the question "Are we alone?" (quoted from the Foreword).

Morrison, P., Billingham, J., and Wolfe, J., The Search for Extraterrestrial Intelligence: SETI. NASA SP-419 (1977), available from Ames Research Center, M/S 229-8, Moffett Field, CA 94035, and from the Government Printing Office. Reprinted, with trivial deletions, by Dover Publications, NY (1979)

Report of a two-year series of science workshops on interstellar communication, chaired by Philip Morrison. Discussed are the philosophy, technology, strategy, and timeliness of carrying out an exploration for extraterrestrial intelligence now.

Ridpath, Ian, Messages from the Stars: Communication and Contact with Extraterrestrial Life. Harper Colophon Books, Harper & Row, Publishers, NY (1979)

Well presented account of search for life beyond the Earth, on other planets in the Solar System and in the Galaxy beyond.

Sagan, C., and Drake, F., The Search for Extraterrestrial Intelligence. Scientific American, 232, no. 5, pp. 80-89 (1975)

Condensed account of activities in interstellar communications to date, including description of attempts already made to listen for signals.

Simpson, G.G., The Nonprevalence of Humanoids, in This View of Life: the World of an Evolutionist, A Harbinger Book; Harcourt, Brace & World, Inc., NY (1964), pp. 253-271 and 299-301

Chapter 13 in this graceful and stimulating book by a master evolutionary biologist and systematist, presents a vigorous argument in support of his estimate that humans are likely to be the only intelligent species in our Galaxy.

Sullivan, Walter, We Are Not Alone. McGraw-Hill Book Company, NY (1964)

A gifted and fascinating discussion of the possibility of life throughout the Universe by an outstanding science writer on the staff of The New York Times.

Sullivan, W.T.3rd, Brown, S., and Wetherill, C., Eavesdropping: The Radio Signature of the Earth. Science, 199, 377-388, 27 Jan. (1978)

Besides looking for "beacons", SETI observers should try to eavesdrop. Here is an extensive examination of Earth's microwave signature. Our UHF-TV carriers and powerful aircraft surveillance radars have made us highly "visible" and informative to other species who may exist in the Galaxy outside the Solar System.

Wolfe, J.H., Edelson, R.E., et al., SETI, The Search for Extraterrestrial Intelligence: Plans and Rationale. Life in the Universe, The MIT Press (in press 1981)

Reflects 1979-80 NASA approach to SETI planning; proposes exploration of a well defined volume of multidimensional microwave search space using state-of-the-art ground-based technologies to achieve exceptional sensitivities.

\* \* \* \* \*

#### MORE SPECIALIZED READING

Billingham, J., Oliver, B.M., and Wolfe, J.H., A Review of the Theory of Interstellar Communication. Acta Astronautica, 6, no.1-2, 47-57 (1979)

Summarizes a classic approach to estimating the likely number of transmitting civilizations in the Galaxy.

Billingham, J. and Pesek, R., (Eds.), Communication with Extraterrestrial Intelligence. Acta Astronautica, Special Edition, 6, no.1-2 (1979)

Contains 18 papers covering a wide range of topics, which were presented originally at the 1975-6-7 CETI Review Sessions of the International Academy of Astronautics.

Black, D.C., In Search of Other Planetary Systems. Space Science Reviews, 25, pp. 30-81 (1980)

A thorough review of the history and the modern approaches to the problem of detecting planets in nearby star systems.

Black, D.C., On the Detection of Other Planetary Systems: Detection of Intrinsic Thermal Radiation. ICARUS, 43, pp. 293-301 (1980)

Examines specific aspects of techniques for the detection of planets in the infrared regime where the brightness contrast with the central star is less than in the optical range.

Bracewell, R.N., An Extended Drake's Equation, the Longevity - Separation Relation, Equilibrium, Inhomogeneities, and Chain Formation. Acta

Astronautica, 6, no.1-2, 67-69 (1979)

The author argues the need to refine Drake's simplistic equation for the number of coexisting, transmitting civilizations.

Buyakas, V.I., and 22 other authors (listed alphabetically), An Infinitely Expandable Space Radiotelescope. Acta Astronautica, 6, no.1-2, 175-201 (1979)

Summarizes first design effort for 1-10 km diameter space radiotelescopes. Modules built on Earth are assembled in orbit. Points out the enormous value to general Cosmic exploration of a system of large antennas in space.

Cameron, A.G.W., (Ed.), Interstellar Communication. W.H. Benjamin (1963)

The first collection of original essays on interstellar communication, with stimulating papers by Morrison, Oliver, Bracewell, Cameron, Shklovskii, Towns, Huang, Calvin, Golay, Purcell, and von Hoerner. Also included are two papers by Frank Drake, one of which describes the first definitive attempt to detect ETI radio signals -- Project "Ozma".

Cocconi, G., and Morrison, P., Searching for Interstellar Communications. Nature 184, 844 (1959)

The seminal paper proposing that humans search for signals of extraterrestrial origin at frequencies near the interstellar 21 cm hydrogen emission line.

Dixon, R.S., and Cole, D.M., A Modest All-Sky Search for Narrowband Radio Radiation Near the 21-cm Hydrogen Line. Icarus, 30, 267-273 (1977)

Modest but devoted search running since 1973 at USRO in a 380 kHz band centered on H<sub>1</sub>-line corrected to Galactic Center of Rest. No confirmed ETI signal above  $1.5 \times 10^{-21} \text{ W/m}^2$  in declination range 14-48 deg. N. Uses beam switching.

Drake, F.D., The Radio Search for Intelligent Extraterrestrial Life, in "Current Aspects of Exobiology" (G. Mamiakunian and M.H. Briggs, editors), Ch. IX, pp. 323-345, Pergamon Press, NY (1965)

Early discussion of the logic of a radio search for the existence of extraterrestrial life and an account of Project OZMA, the first post World War II exploration.

Edelson, R.E., An Observational Program to Search for Radio Signals from Extraterrestrial Intelligence Through the Use of Existing Facilities. Acta Astronautica, 6, no.1-2, 129-143 (1979)

Describes a program to search for ETI signals from 1 to 25 GHz over 80% of the sky to flux levels from 190 to 230 dBW per m<sup>2</sup> in a 300 Hz band, and rationale for this strategy.

Edelson, R.E., An Experiment Protocol for a Search for Radio Signals of Extraterrestrial Intelligent Origin in the Presence of Man-made Radio Frequency Sources. Acta Astronautica, 6, no.1-2, 145-162 (1979)

This paper describes the ETI signal detection problem in the face of stochastic noise and man-made interference, and suggests a protocol for their rapid elimination on-line.

Gindilis, L.M., Kardashev, N.S., Soglasnov, V.A., Spangenberg, E.E., and Etkin, V.S., Search for Signals from Extraterrestrial Civilizations by the Method of Synchronous Dispersion Reception. Acta Astron. 6, no.1-2, 95-104 (1979)

A search for very broad band extraterrestrial pulses, using diversity reception and correcting for interstellar dispersion in order to screen out local noise sources.

Greenstein, J.L., Chairman, Astronomy Survey Committee, National Academy of Sciences, National Research Council, Astronomy and Astrophysics for the 1970's. National Academy of Sciences (U.S.A.) (1972)

"One of the most striking conclusions of the report is the high probability it assigns to the existence of intelligent life elsewhere in the universe." (From the letter of transmittal by Harvey Brooks, for the Committee on Science and Public Policy, COSPUP).

Gulkis, S., Olsen, E.T., and Tarter, J., A Bimodal Search Strategy for SETI, pp. 93-105 in Strategies for the Search for Life in the Universe, Papagiannis, M.D., editor. D.Reidel Publishing Company, Boston (1980)

The targeted star observing mode and the sky survey mode. A moderately technical report on the status of the NASA SETI R&D program development being carried out jointly at the NASA Ames Research Center and the Jet Propulsion Laboratory of the California Institute of Technology.

Haldane, J.B.S., The Origins of Life. New Biology, no. 16, p. 12; Penguin, London (1954)

Original early essays on the origin of life and chemical evolution, written before the critical experiments of Miller and Urey.

Harrington, R.S., Planetary Orbits in Binary Stars. Astronomy Journal, 82, 753 (1977)

Study of those binary systems whose separations appear to allow a stable planetary orbit within the habitability zone around the solar type star often present in such systems.

Hart, M.H., N is Very Small, in Papagiannis, M.D., (Ed.), Strategies for the Search for Life in the Universe. D. Reidel Publishing Company, Dordrecht, Holland, (1980)

Summarizes eight years of developing his arguments that we are alone not

only in the Galaxy but probably in the local group of galaxies as well. The antithesis of Sagan's views.

Horowitz, P., A Search for Ultra-Narrowband Signals of Extraterrestrial Origin.  
Science, 201, 733-735, 25 Aug. (1978)

185 Sun-like stars were observed on 1420 MHz  $\rightarrow$  500 Hz in 80 hours with 0.015 Hz bandwidth at N.A.I.C., Arecibo, P.R. Found no signals above  $4 \times 10^{-27}$  Wm<sup>-2</sup> within 500 Hz of the hydrogen rest frequency relative to Solar System barycenter.

ICARUS: International Journal of Solar System Studies, published monthly by Academic Press, Inc., 111 Fifth Avenue, New York, NY 10003

Publishes refereed papers on a wide range of topics related to SETI.

JBIS: for Journal of the British Interplanetary Society, published since 1934 by the British Interplanetary Society, 27/29 Lambeth Road, London, SW8 1SZ, England.

Internationally famous journal. Editorial matter ranges from nontechnical to moderately technical. Besides special issues, "Each issue is now devoted to one of five main subject areas, viz. Space Technology, Space Applications, Astronautics History, Space & Education and Interstellar Studies." Papers in "SETI" area are relatively frequent.

JBIS, 28, No.11, November 1975.

Entire issue devoted to "Development of Intelligence and Civilization in the Universe". Contains papers by S. von Hoerner, K.A. Ehricke, David Viewing, Bob Parkinson, and a translation of USSR Research Programme on the Problems of Communication with Extraterrestrial Civilizations adopted by the Scientific Council on the Problem Area of Radio Astronomy (Academy of Sciences of the USSR).

Kaplan, S.D., (Ed.), Extraterrestrial Civilizations: Problems of Interstellar Communication. NASA TTF-631 (1971)

Interesting collection of Soviet papers from previous years which cover many aspects of interstellar communication.

Kardashev, N.S., Strategy for the Search for Extraterrestrial Intelligence.  
Acta Astronautica, 6, no.1-2, 33-46 (1979)

The author argues the importance of the search to humankind, its self-understanding and its survival. He points out the value of the 10 km space radio telescopes now being designed.

Kreifeldt, J.G., A Formulation for the Number of Communicative Civilizations in the Galaxy. Icarus, 14, 419-430 (1971)

Extends the Drake equation through a linear system formulation which allows for different star generation rates and civilization lifetime and development time distributions. An expression for the variance of the estimate of the number of communicative civilizations is derived.

Lederberg, J., Exobiology: Approaches to Life Beyond the Earth. Science 132, 393-400 (1961)

A distinguished American geneticist explores aspects of life outside the earth; exobiology. Includes suggestions for searching for life in the Solar System.

Machol, R.E., Two Systems Analyses of SETI. Acta Astronautica, 6, no.1-2, 163-173 (1979)

A systems analysis of hyperdimensional search space demonstrates that for the next few years the greatest benefit/cost ratios will arise from development of better spectrum analyzers (up to  $10^9$  channels) and better data processors.

Mallove, E.F., Connors, M.M., Forward, R.L., and Paprotny, Z., A Bibliography on the Search for Extraterrestrial Intelligence. NASA Reference Publication 1021 (1978); available from NASA Scientific and Technical Information Office, or (as long as copies remain on hand) from SETI Program Office, Ames Research Center M/S 229-8, Moffett Field, CA 94035.

An extensive bibliography of interstellar communication and related subject matters. The main divisions are:

1. Life-supporting extrasolar environments
2. Origin and evolution of extrasolar life
3. Methods of searching for extraterrestrial intelligence
4. Decoding signals from extraterrestrial intelligence
5. Philosophical, psychological, and sociological aspects of the search for extraterrestrial intelligence
6. Miscellaneous

Mallove, E.F., and Forward, R.L., Bibliography of Interstellar Travel and Communication. JBIS 33, No.6, pp. 201-248 (1980)

Extensive bibliography of interstellar communication and interstellar travel. Contains 2699 references sorted into thirteen major categories and their subdivisions, and by author(s). Continued development of this bibliography is now in the care of Dr. R.S. Dixon, The OSU Radio Observatory, 2015 Neil Avenue, Columbus, OH 43210.

Miller, S.L., and Orgel, L.E., The Origins of Life on Earth. Prentice-Hall (1974)

Summary of twenty years of laboratory and field work in chemical evolution; written by two of the leading investigators in this field.

Morrison, P., Billingham, J., and Wolfe, J., The Search for Extraterrestrial

Intelligence -- SETI. Acta Astronautica, 6, no.1-2, 11-31 (1979)

After the introduction, this article presents the index, preface, and major section on "CONCENSUS", from NASA SP-419 by the same authors. (See under GENERAL, above.)

Murray, B., Gulkis, S., and Edelson, R.E., Extraterrestrial Intelligence: An Observational Approach. Science, 199, 485-492, 3 Feb. (1978)

Gives arguments for exploring the whole sky for ETI signals in the entire frequency band of the terrestrial microwave window, and suggests a procedure to do it in about 5-years to a flux level of 190 to 210 dBW/m<sup>2</sup> at a resolution of 300 Hz.

Oliver, B.M., Thermal and Quantum Noise. Proc.I.E.E.E., 53, no.5, May (1965)

Tutorial paper; develops theory of black body radiation, thermal and quantum noises from basic physical principles, and shows how to apply results to antenna theory and ideal receivers, radio and optical.

Oliver, B.M., and Billingham, J., Project Cyclops. NASA CR114445 (1972). Revised Edition (1973)

The remarkable, and first, detailed conceptual design study of an expandable system for detecting signals from extraterrestrial civilizations. Proposes an Earth-based system of phased radiotelescopes with a sophisticated data processing system (which is now out of date because of progress in large-scale integrated solid-state digital electronics).

Oliver, B.M., State of the Art in the Detection of Extraterrestrial Intelligent Signals. Acta Astronautica, 18, 431-439 (1973)

A summary of possible engineering systems for the detection of signals from extraterrestrial intelligent life; includes the main features of the Project Cyclops design.

Oliver, B.M., Proximity of Galactic Civilizations. Icarus, 25, 360-367 (1975)

The author examines, on a statistical basis, the number of times in galactic history that two civilizations might have emerged close to each other and at the same time, perhaps stimulating each to search for other civilizations.

Oliver, B.M., Rationale for the Water Hole. Acta Astronautica, 6, no.1-2, 71-79 (1979)

The physical and psychological basis for the thesis that, in the absence of any more cogent reason to prefer another frequency band, the water hole should be considered the primary preferred frequency band for interstellar search.

Oparin, A.I., Life: Its Nature, Origin and Development. Oliver and Boyd, London (1961)

Summary of the writings of the distinguished biochemist; includes some of earliest theories about chemical evolution, as first proposed by the author in the 1920's.

Papagiannis, M.D., (ed.), Strategies for the Search for Life in the Universe. D. Reidel Publishing Company (1980)

Joint session of Commissions 16, 40, and 44, held in Montreal, Canada, during the IAU General Assembly, 15 and 16 August, 1979. 23 papers, many of them concerned with the pros and cons of the popular question "Where are they?"

Pesek, R., Activities of the IAA CETI Committee from 1965-1976 and CETI Outlook. Acta Astronautica, 6, no.1-2, 3-9 (1979)

The Chairman reviews a decade of activity by this committee of the International Academy of Astronautics. He concludes that the committee has done valuable work and should continue to provide IAA support to CETI (SETI in the U.S.A.).

Peterson, A.M., Narasimha, M., and Narayan, S., System Design for a Million Channel Digital Spectrum Analyzer (MCSA). Proc. IEEE 13th Asilomar Conference on Circuits, Systems, and Computers, IEEE Catalog No. 79-CH1468-8C, pp. 14-17

First published account of architectural and hardware approach used in the design of digital megachannel spectrum analyzers for the NASA SETI receiving system. A progress report, in effect, since implementation was delayed, leaving time for still greater sophistication.

Ponnampерuma, C., and Cameron, A.G.W., (Eds.), Interstellar Communication: Scientific Perspectives. Houghton Mifflin, NY (1974)

Collection of essays on interstellar communication presented in a lecture series at NASA Ames Research Center in 1971. A successor to Cameron's original book.

Sagan, C., (Ed.), Communication with Extraterrestrial Intelligence. M.I.T. Press, Cambridge, Mass., (1973)

Detailed and exciting account of the first international meeting on communication with extraterrestrial intelligence, held in 1971 at Byurakan, Armenia, under the joint auspices of the U.S. and Soviet Academies of Science.

Sagan, C., article on Life in the current edition of the Encyclopaedia Britannica, Vol.10, pp.893-911 (1979)

Starts with various definitions of life, describes it as we know or

hypothesize it (pre-Viking Mars Lander experiments), and concludes with a brief discussion of life as we do not know it, i.e., extraterrestrial life.

Schneider, S.H., and Thompson, S.L., Cosmic Conclusions from Climatic Models: Can They be Justified? ICARUS, 41, 456-469 (1980)

"None of this is meant to discourage further ingenious -- or even speculative -- use of climatic models on cosmic questions. But we conclude that cosmic conclusions from climatic models should be accompanied by clear admission of the vast uncertainties in the climatic component of the argument, let alone other parts of the problem."

Seeger, C.L., Strategic Considerations in SETI, and a Microwave Approach. Acta Astronautica, 6, no.102, 105-127 (1979)

Thorough account of search strategy and equipment designs at Ames Research Center circa mid-1979. First description of a million-channel spectrum analyzer design developed by SETI team and A.M.Peterson (Stanford Univ.), and of the consequent data manipulation problem.

Seeger, C.L., The Recognition of Extraterrestrial Artificial Signals. Proc. IEEE, 13th Annual Asilomar Conference on Circuits, Systems, and Computers, Nov. 5-7, 1979; Conference Record, IEEE Catalog No. 79CH1468-8C (1980)

Two basic concerns in designing an efficient SETI receiving system are (1) on-line detection of artificial signals at low signal to noise ratios, and (2), determining if a signal is "natural", or a human artifact, or a sign of extraterrestrial intelligence. One is looking for a steel needle in a scrap iron pile.

Shklovskii, I.S., and Sagan, C., Intelligent Life in the Universe. Holden-Day, New York (1966)

First substantive book on all aspects of theories about the prevalence of life in the Universe. A joint venture over great distance by two distinguished scientists; imaginative, and now a somewhat dated classic.

SPACEFLIGHT: published monthly since 1956 by the British Interplanetary Society, 27/29 Lambeth Road, London, SW8 1SZ, England

"...provides essential reading not to be found in any other publication. Present events and future plans are dealt with in news items and major articles. Extensive participation by readers is developed through correspondence, book reviews, personal accounts and histories...part of a communications network connecting all who have interests in space."

Tipler, F.J., Extraterrestrial Intelligent Beings Do Not Exist. Q. Jl. R. Astr. Soc., 21, pp267-281 (1980)

"Although this argument has been expressed before, its force does not seem to have been appreciated. I shall try to rectify this situation by

showing that an intelligent species with the technology for interstellar communication would necessarily develop the technology for interstellar travel, and this would automatically lead to the exploration and/or colonization of the Galaxy in less than 300 million years." Quite the most complete hypothesizing yet in favor of this particular extremum. More than eighty-one citations. A truly academic paper.

Troitskii, V.S., Starodubtsev, A.M., and Bondar, L.N., Search for Radio Emissions from Extraterrestrial Civilizations. Acta Astronautica, 6, no.1-2, 81-94 (1979)

Extensive account of widely separated synchronous observations since 1970 in centi- and decimeter bands, using dipoles and simple horn antennas. Found evidence for global, sporadic radiation apparently originating with the Earth's magnetosphere.

\* \* \* \* \*

#### APPENDIX I

Special issue of Acta Astronautica  
CETI: Communication with Extraterrestrial Intelligence  
Vol. 6, 1-2 (Jan.-Feb. 1979)

#### CONTENTS

- 1 PREFACE -- J. Billingham and R. Pesek
- 3 ACTIVITIES OF THE IAA CETI COMMITTEE FROM 1965-1976 AND CETI OUTLOOK -- R. Pesek
- 11 THE SEARCH FOR EXTRATERRESTRIAL INTELLIGENCE-SETI -- P. Morrison, J. Billingham and J. Wolfe
- 33 STRATEGY FOR THE SEARCH FOR EXTRATERRESTRIAL INTELLIGENCE - N. S. Kardashev
- 47 A REVIEW OF THE THEORY OF INTERSTELLAR COMMUNICATION -- J. Billingham, B. M. Oliver and J. H. Wolfe
- 59 A REVIEW OF RECENT CONCEPTS OF THE PROBLEM OF THE ORIGIN OF LIFE -- L. M. Mukhin
- 67 AN EXTENDED DRAKE'S EQUATION, THE LONGEVITY-SEPARATION RELATION, EQUILIBRIUM, INHOMOGENEITIES AND CHAIN FORMATION -- R. N. Bracewell
- 71 RATIONALE FOR THE WATER HOLE -- B. M. Oliver
- 81 SEARCH FOR RADIO EMISSIONS FROM EXTRATERRESTRIAL CIVILIZATIONS -- V. S. Troitskii, A. M. Starodubtsev and L. N. Bondar

- 95 SEARCH FOR SIGNALS FROM EXTRATERRESTRIAL CIVILIZATIONS BY THE METHOD OF SYNCHRONOUS DISPERSION RECEPTION — L. M. Gindilis, N. S. Kardashev, V. A. Soglasnow, E. E. Spangenberg, V. S. Etkin and V. G. Mirovskii.
- 105 STRATEGIC CONSIDERATIONS IN SETI, AND A MICROWAVE APPROACH — C. L. Seeger
- 129 AN OBSERVATIONAL PROGRAM TO SEARCH FOR RADIO SIGNALS FROM EXTRATERRESTRIAL INTELLIGENCE THROUGH THE USE OF EXISTING FACILITIES — R. E. Edelson
- 145 AN EXPERIMENT PROTOCOL FOR A SEARCH FOR RADIO SIGNALS FROM EXTRATERRESTRIAL INTELLIGENT ORIGIN IN THE PRESENCE OF MAN-MADE RADIO FREQUENCY SOURCES — R. E. Edelson
- 163 TWO SYSTEMS ANALYSES OF SETI — R. E. Machol
- 175 AN INFINITELY EXPANDABLE SPACE RADIOTELESCOPE — V. I. Buyakas, Yu. I. Danilov, G. A. Dolgopolov, K. P. Feoktistov, L. A. Gorshkov, A. S. Gvamichava, N. S. Kardashev, V. V. Klimashin, V. I. Komarov, N. P. Melnikov, G. S. Narimanov, O. F. Prilutsky, A. S. Pshennikov, V. G. Rodin, V. A. Rudakov, R. Z. Sagdeev, A. I. Savin, Yu. P. Semenov, I. S. Shklovskii, A. G. Sokolov, G. S. Tsarevsky, V. I. Usyukin and M. B. Zakson.
- 203 ON ACTIVE AND PASSIVE CETI FROM AN EARTH SATELLITE ORBIT — M. Subotowicz, J. Usowicz and Z. Paprotny
- 213 INTERSTELLAR COMMUNICATION BY NEUTRINO BEAMS — M. Subotowicz
- 221 MESSAGE THROUGH TIME — G. Marx

---

#### APPENDIX II

A RECOMMENDATION FROM THE 27th CONGRESS OF THE  
INTERNATIONAL ASTRONAUTICAL FEDERATION,  
FOR SUBMISSION TO  
THE INTERNATIONAL TELECOMMUNICATIONS UNION,  
AND THE UNITED NATIONS

Resolution on Radio-Frequency Interference Protection for a Search for  
Extraterrestrial Intelligence (SETI)

In recognition of the rapidly advancing worldwide preparations for the 1979 World Administrative Radio Conference (WARC), the International Astronautical Federation requests that national administrations take note of the following:

There are particular frequency bands which merit priority for the purpose of searching for radio signals from extraterrestrial civilizations. These are:

- (a) 1400 to 1427 MHz
- (b) 1427 to 1727 MHz

The 1400 to 1427 MHz band is important because interstellar transmissions may take place around the hydrogen line, while the 1427 to 1727 MHz band is located between the hydrogen and hydroxyl lines, and lies at the minimum of the galactic noise background. 1400 to 1427 MHz is currently allocated exclusively to the radio astronomy service, and may be shared with it by a SETI service; 1427 to 1727 MHz may be shared with services the transmissions of which will not cause harmful interference to SETI receivers.

Existing radio telescopes are already being used to search for radio signals from extraterrestrial civilizations, and the feasibility of constructing a large ground-based search system has been established. The performance of a ground-based instrument could be seriously degraded by radio frequency interference (RFI), primarily from line-of-sight transmitters such as satellites. The only identified alternatives to an earth-based search system are:

- i) a space-based system, and
- ii) a system on the lunar far side.

Both of these are possible in the future, but at an unknown cost.

Furthermore, a space-based search system, unless shielded at additional expense, is vulnerable to interference from all terrestrial transmitters, and even shielding will not protect it from interference by geosynchronous satellites unless it is located at least at lunar distance. A search system on the far side of the moon is vulnerable to all transmissions originating beyond the lunar orbit. Thus there exists a need for RFI protection for SETI.

Accordingly, the International Astronautical Federation recommends that national administrations at the 1979 WARC approve regulations which include:

- i) creation of a SETI service, and allocation to it as the primary service of the frequencies between 1400 and 1727 MHz,
- ii) allocations for new satellite systems at frequencies outside the protected bands,
- iii) appropriate frequency-sharing criteria for services which can operate in the bands without causing interference to SETI,
- iv) technical criteria for allowable spurious radiation from out-of-band services,
- v) phase-out of services which now operate in the protected bands and which cannot avoid causing interference to SETI.

---

### APPENDIX III

#### A SUMMARY OF SETI

Prepared for the IAF UNISPACE 82 Committee  
by John Billingham

Over the past thirty years advances in the astronomical, planetary and biological

cal sciences have led to the idea that life may be widespread in the universe. In many cases this life may have evolved to the stage of intelligence, and perhaps into civilizations much older than our own. Over the same period of time the science of radio astronomy has emerged as a new and powerful way of exploring the universe, an advance made possible by new techniques in radiotelescope design and new developments in the field of data processing. The possibility that these tools might now be used to search for radio signals transmitted by other civilizations is being examined intensively in a number of countries by groups concerned with the Search for Extraterrestrial Intelligence (SETI).

There is a general consensus that the initial thrust of any thoroughgoing search should be to try to detect signals in the electromagnetic spectrum. Many groups believe also that the microwave window, between 300 MegaHertz and 300 GigaHertz, is a preferred frequency band. And indeed, many small scale searches have already been carried out at different frequencies within this band. The main reason for the selection of this band is that it is a quiet region of the spectrum, a place where faint signals from distant civilizations would be most easily detected.

The major problem for SETI is the large number of unknowns. We do not know where to point the telescope, nor what frequencies and polarizations to use, nor what pattern to look for, nor how sensitive we need to make the SETI system. Various groups have proposed, and some have used, differing values of all these variables for specific search ideas. But the sum of all searches to date has examined only a tiny fraction of the vast volume of "search space" embraced by all the variables.

There are some ways in which the odds against a successful detection can be narrowed. Certain frequencies have been proposed as interstellar communication signposts, for example, the hydrogen, hydroxyl, and water lines. So these and other "spot bands" should certainly be included. Detection is probably easier at the low end of the microwave window. Two polarizations should be used. There are some a priori arguments about search direction: for example, high sensitivity searches of nearby stars like our own Sun are clearly indicated; surveys of the whole sky at lower sensitivities should be carried out to ensure that we are not missing unusually strong signals from more distant stars; certain regions of the sky demand more attention because they contain more stars, for example the galactic plane; and some modest effort should be devoted to looking at nearby galaxies for evidence of supercivilizations.

In tackling the problem of the type of signal to look for, it is soon apparent that there are considerable advantages to the user of very narrow frequency bands, perhaps down to 1 Hertz. There are billions of such channels. One solution, now being actively pursued in the U.S.A., is to design sophisticated spectrum analyzers, capable of examining a large number of separate channels at the same time. Advances in digital electronic technology now permit the design of systems with ten million such channels at extremely low cost. The stream of information emerging from the analyzer at exceptionally high rates must now be fed to an equally sophisticated signal processor. The processor must search automatically for faint signals among the background noise. Three types of signals must be considered: continuous transmissions, fixed or drifting in frequency, pulses, as from a lighthouse, and complex patterns. Many signals will be of astrophysical origin, even more of terrestrial origin. These must be identified and discarded. Any signal which remains at a high level of significance after passing all the automatic tests is then examined in more detail by the observing team.

Radio frequency interference is a major problem for SETI. It is caused by the

large number of transmissions from the surface of the Earth and from Earth-orbiting satellites. The microwave region of the spectrum is rapidly filling up with these transmissions, and the chances of detecting a faint SETI signal are dwindling at a comparable rate. Protection of frequency bands for SETI was explored at the 1979 World Administrative Radio Conference in Geneva, but was not achieved. SETI can use the protected radio astronomy bands, but these are few and far between. If the situation continues to deteriorate for SETI it may eventually be necessary to build large radiotelescopes in space, take them out to lunar distances, shield them from terrestrial interference, and conduct the search in the space environment. There are in fact many other advantages to the use of spaceborne SETI telescopes, and these have been emphasized in papers and proposals in the Soviet Union.

SETI has emerged in recent years as a respectable new discipline, born of a confluence of ideas and developments in science and technology. In many ways it is more exploration than science, though there will be many benefits of SETI systems to radio astronomy and other sciences. It is difficult to estimate the chances of success, but it is clear that the detection of a signal would be an event of great importance for our own civilization. In turn, this emphasizes the importance of considering SETI from the international point of view, and of encouraging the exchange of ideas and results of SETI studies and investigations in an international forum. Since 1971, such a forum has been provided by the International Review Meeting on Communication with Extraterrestrial Intelligence (under the aegis of the International Academy of Astronautics) held annually during the Congress of the International Astronautical Federation. Such exchanges may well lead in the future to cooperative studies and programs in SETI science and technology among various countries, at both informal and formal levels. Such cooperative ventures should be encouraged.

\*\*\*\*\*

IMPACT OF CETI-SETI ON THE  
DEVELOPMENT OF BIOLOGY [B],  
ASTRONOMY [A], RADIOASTRONOMY [R],  
ASTROPHYSICS [AP], AND CULTURE  
[C]—"BARAPC"

M. Subotowicz\*,\*\* and Z. Paprotny\*\*

\*Institute of Physics, M. Curie-Skłodowska University,  
20-031 Lublin, Poland

\*\*Polish Astronautical Society, Warsaw [Warszawa], Poland

ABSTRACT

There exist the possibilities to use future CETI-SETI-hardware to realize the ARAP programs. We suppose the possibility of the situation reversed to the contemporary one when the instruments and telescopes built for ARAP are used from time to time for CETI and SETI. Importance of scientific gains from the search for ETI signals are discussed in the present paper. The difficulties in SETI and CETI are summarized in Drake's idea of the "Cosmic Haystack" that contains about  $10^{29}$  cells of size 1 Hz x 1 Arecibo Beam  $\times 10^{-30} \text{ Wm}^{-2}$ . The ultimate values for some of the parameters, characterizing principal CETI and SETI instruments, are accepted and many of the scientific problems to be solved for ARAP by use of these instruments are listed.

The importance of the search for the extraterrestrial planets for ARAP and CETI is discussed and some of the programs of the detection of these planets are shortly presented: project Orion, speckle interferometry, Space Telescope, spaceborne interferometers and large space radiotelescopes (SRT). The role of the search at natural radio-lines for ARAP and CETI-SETI as well as possibility to make more extensive and accurate catalogues of the stars are discussed. The capabilities of SRT's with huge collecting areas, proposed by Polish (1976, 1979), American (1976) and Russian (1977, 79) groups, are discussed. The technological advances in the development of different instruments for SETI and CETI and their application for ARAP were presented. The impact of SETI-CETI research on the development of biology and on our culture, the better understanding of the nature of our cognition, its universality, knowledge, norms and anthropomorphicity of our thinking are discussed shortly.

**1    INTRODUCTION**

There exist the possibilities to use future CETI-SETI - hardware to realize the ARAP programs. We suppose the possibility of the situation reversed to the contemporary one when the instruments and

telescopes built for ARAP are used from time to time for CETI and SETI.

Importance of scientific gains to be gathered during searches for ETI signals have been stressed by nearly all proponents of large-scale SETI enterprises. The first opinion on this topic we were able to spot was presented by F.D. Drake: "... the constant acquisition of nothing but negative results can be discouraging psychologically. A scientist must have some flow of positive results, or his interest flags. Thus any project aimed at the detection of extra-terrestrial intelligent life should also conduct more conventional research" (Drake, 1965). One thing is worth being noticed here: Drake's opinion was presented in those pioneering years of CETI when our expectations regarding probability of making contact were far too optimistic as we now realize (cf. D.W. Atchley's opinion that ETI should be discovered by 1970 - Atchley, 1960).

Among five goals and objectives of JPL (Jet Propulsion Laboratory), project of 1-25 GHz all-sky survey listed by Edelson et al. (1978), two are directly connected with need of maximizing scientific output of the search:

- simultaneously acquire scientifically interesting radio astronomical data,
- in all cases, maximize the utility of, and provide a focus for, the required technology for broader NASA objectives in communications and information management.

It may turn out that only huge facilities like Cyclops (Morrison et al., 1978) or space-based antennas will give us substantial probability of ETI's detection. In this case the main problem will be with getting financial support for these grandiose project contemporary research already suffers from the lack of money.

In our opinion predictable benefits which introducing such ultimate observing tools like Cyclops could bring for conventional science are so enormous that they would be far more persuasive to all who must vote the funds than any arguments involving possible values of contacting ETI. It is in no case common opinion. Potential benefits which might be gained additionally to SETI with equipment designed specifically for SETI have sometimes been questioned even by astronomers. To cite G. Verschuur (radioastronomer, once involved in SETI): "... I have always felt that if radio astronomers have very good reasons for building a 1-million-channel receiver or a Cyclops antenna, then they are quite capable of rationalizing it on their own terms ..." (Verschuur 1978). A clear distinction between guaranteed and unguaranteed gains of SETI and conventional programs does not exist. It is most evident when astronomical aspects of SETI observations are concerned. SETI is nothing else than observing sky and the only difference between conventional radio astronomy and proposed listenings for ETI signals is that while latter will certainly prove beneficial to former, inverse situation is far less probable. Needless to say, astronomers should be on SETI - side.

## 2 COSMIC HAYSTACK

According to the rather optimistic evaluation of Drake and Sagan there can exist in our Galaxy  $< 10^6$  civilizations. To have substantial probability of finding a civilization one should search  $> 10^5$  stars. A habitable zone can be found around the F, G, K and M-type stars. These stars are stable enough to allow of life's origin and its development in course of biological evolution into

the homo sapiens and technological civilization level. According to Zuckerman and Tarter (1979), 40 MW military radars transmitting with a 91 m antenna would send the signals that could have been detected by similar antenna as far away as the most distant stars observed. F. Drake describes the search for ETI in an eight-dimensional parameters space, the so-called "The Cosmic Haystack". The eight parameters are: 3-spatial, 1 - temporal, 1 - frequency, 2 - polarizations and 1 - transmitter power. The frequency axis covers the spectrum from 300 MHz up to 300 GHz and even more - to 3.000 GHz.

There is  $3 \times 10^{28}$  to  $3 \times 10^{29}$  cells of size 1 Hz x 1 Arecibo Beam  $\times 10^{-30} \text{ Wm}^{-2}$ . Till now a small fraction of the haystack volume ( $\sim 10^{-17}$ - $10^{-18}$ ) have been explored.

Up to date there have been about 20 radio searches for signals from ETI. All the results were negative. Everybody can ask therefore whether investment of money to build SETI-CETI - specific hardware is reasonable to increase sensitivity, resolution and observing time. Another possibility would be to use more efficiently the observing time on existing telescopes. The improvement factor of the order  $10^7$  could be achieved in 1980-ties by use the existing antennas and special SETI instrumentation (mainly MCSA).

### 3    QUALITY OF SETI-CETI INSTRUMENTS AND SOME SCIENTIFIC PROBLEMS TO BE SOLVED

When looking for the CETI and SETI instruments we should accept ultimate values for some of the parameters, characterizing these instruments; they will be placed probably in space. In the case of the interferometric regime we should accept the possible length of the baseline; in one of the projects (Bujakas et al., 1977) it is equal up to 10 AU! At this length of the baseline the authors suppose phantastic resolution of  $10^{-10}$  arc seconds! Let us accept a somewhat more realistic but still very large resolution, namely  $10^{-4}$  to  $10^{-5}$  arc sec. The resolution of the optical space telescope can be taken as equal  $10^{-2}$  to  $10^{-3}$  arc sec. The energy sensitivity of the radiotelescope can be estimated for  $10^{-28}$  up to  $10^{-30} \text{ W/cm}^2$ . Large Space Telescope that will be delivered by the Space Shuttle into the Earth satellite orbit will have the resolution  $5 \times 10^{-2}$  arc sec. that allows to discover the Jovian-like planets from the distance of about 1000 light years. At these ultimate resolutions and sensitivities different effects can be important. For example, at the resolution less than 0.5 arc sec one should take into the account the atmospheric effects, at the resolution less than  $10^{-3}$  arc sec the gravitation effects may be important, a.s.o. CETI and SETI can and will be realized in different spectral ranges: X, UV, visible, IR, mm, microwave and radio-wave. We suppose that in all these spectral regions the proper transmitters CETI and receivers SETI will be built to realize the tasks of the interstellar communication. Together with that there exist possibilities to use the future CETI-SETI hardware to realize the ARAP programs. We suppose the possibility of the situation reversed to the contemporary one when the instruments and telescopes built for ARAP are used from time to time for CETI and SETI. Well known is the idea of "parasitic SETI" realized automatically during the continuous radio-astronomical observations not connected with SETI. This strategy can

lead to serendipitous discovery of ETI. There is already working system called SERENDIP. One can accept similar procedure the powerful hardware built originally for CETI-SETI.

Many of scientific problems important for ARAP development can be solved when observing the Universe in different spectral regions (Sciama, 1971; Sobotowicz, 1974; Kuchowicz, 1979; Shklovski, 1975, and 1976; Klimshin, 1980; Baym, 1977).

- a UV and X-rays region: observation of the neutron stars and pulsars; search for the black holes; looking for the binary pulsars, detection of the gravitational waves and verification of the general theory of relativity,
- b Visible and IR-rays region: Doppler measurements of the redshift also in other spectral regions, determination of the velocities of the outer galaxies, testing of the different models of the Universe (open or closed Universe?), expansion of the Universe, investigation of the quasars, their redshift and distribution, are they local phenomenon in our Galaxy or universal ones,
- c mm- and microwave range: investigation of the distribution of the galaxies in space; testing of the different models of the Universe; looking for the radiosources and quasars; similarity of quasars, N galaxies, radio-galaxies and Seifert-galaxies; intercontinental interferometry with the resolution  $10^{-3}$  arc sec; counting the radiosources and quasars to understand the evolution of some objects in the Universe; measuring the Doppler shift for  $z = \Delta\lambda/\lambda \approx 3$  we can investigate about 80% of the Universe history; measurements of the relict radiation mainly in the microwave region to investigate the Universe history before the formation of the galaxies and the isotropic character of this radiation with the accuracy  $\lesssim 0,1\%$ ; morphology of the galaxies and metagalaxies, mass distribution between the galaxies; homogeneity and isotropy in large scale in the Universe,
- d neutral and ionized gas: observation of the intergalactic neutral gas: H for  $\lambda = 21$  cm and  $L\alpha$ , He - 584 Å and  $He^+ - 304$  Å; the lines in UV can not be observed from the surface of the Earth even at very large Doppler shifts. Looking for the intergalactic ionized gas, thermal bremsstrahlung in X-region, scattering of the electrons, dispersion and absorption of the radio-waves; continuous radiation of the cosmic background depends on the dispersed intergalactic matter and primary cosmic radiation and on the "primary" radiation at the very early stage of the Universe.

#### 4 DETECTING EXTRASOLAR PLANETS

Settling the question whether planets exist around other stars is possibly most challenging task of upcoming astronomical research. Importance of this single question for both SETI and astronomy is unquestionable. For SETI because no-planets stellar systems can be ruled out as targets for searches, while detection of extrasolar planets will strongly support idea of plurality of inhabited worlds,, thus helping to launch big SETI projects. For astronomy because we even do not know whether planetary bodies are common or unique companions to stars. Contemporary theories of stellar evolution suggest that common, but it may well turn out that these single-example theories are worthless as far as planetogenesis is concerned. It is not our intention to discuss problem of detection of nonsolar planets in detail. Let us rather summarize most promising approaches proposed thus far.

#### 4.1. Project Orion

This Earth-based optical imaging interferometer with 50 meters baseline has been discussed extensively in recent NASA publication (Black, 1979). Orion should achieve about  $10^2$  times better accuracy of positional measurements over best ones achievable with classical long-focus astrometry ( $10^{-4}$  to  $10^{-5}$  arc sec for Orion).

#### 4.2 Speckle interferometry with new generation telescope

One of concepts for a new generation telescopes which have been proposed by Kitt Peak National Observatory team under direction of D.N.B. Hall envisions 1000-inch ground-based instrument. Such a MMT-type telescope in conjunction with speckle interferometer would offer resolution sufficiently high to permit direct IR observations of Jupiter-like planets around nearby cool stars (Goldberg, 1977).

#### 4.3 Space Telescope

Ability of the 2,4 meter Space Telescope (ST) to detect extrasolar planetary systems has been discussed recently by Baum (1980). Two of ST instruments are of particular importance in this regard: wide-angle planetary camera and system for precise positioning. Both cameras are fed by rotating, pyramid-shaped mirror placed in optical axis at f/24. Wide angle camera with f/12,9 has field 2,7x2,7 arc min, while planetary with f/30 observes 1,1x1,1 arc min field with maximal attainable resolution. Positional measurements of stars up to 22 magnitude will be carried out with an error of  $\sim 0.001$  arc sec for planetary camera and of  $\sim 0.002$  arc sec for wide-angle one. Periodicities in stars' movements will be sought for, indicating presence of invisible companions.

#### 4.4 Spaceborn interferometers

Independently of Space Telescope and specialized satellites like HIPPARCOS, high precision astrometric measurements can also be carried out with spaceborn interferometers. Four-element (two x two 25 cm mirrors, 2x 10 meters basis line) orthogonal interferometer described recently by Knowles and Thacker (1980) can successfully rival with above mentioned instruments when stars' trigonometric parallaxes and proper motions are concerned. Moreover, observations performed in infrared ( $\lambda \approx 30$   $\mu\text{m}$ ) should enable detection of Jovian planets circling nearby F and G stars, despite differences in optical brightnesses reaching 20 mgt. Similiarly designed spinning interferometer on Earth's orbit has been proposed earlier by Bracewell (1978).

#### 4.5 Apodized telescope

Technique of apodization involving special mask in optical axis allows reducing intensity of the diffraction rings by a factor of  $10^9$  at 1 arc sec from central image on a two meter space telescope, thus permitting direct observation of Jovian planets up to, say, 5 parsecs. However, apodization is on the cutting-edge of contemporary technology and theoretical studies are still in progress (Oliver, 1979).

It is quite evident therefore that in the field of extrasolar planets general goals of CETI and stronymy are indistinguishable and, in case of detection, research in both areas will be equally influenced.

The discovery of the planetary systems around the distant stars would be very stimulating for CETI-SETI as well as for astronomy and cosmogony of the planetary systems. The contemporary astrometric or radial velocity measurements performed from the surface of the Earth are extremely complicated and not very promising. Space Telescope on the Earth satellite orbit may improve the situation but will not change it essentially.

Large spaces radiotelescopes and space interferometers proposed by Polish (Subotowicz, 1976), Russian (Bujakas et al., 1977) and American (Morrison et al., 1977), groups for CETI and space radio-telescopes with the unfilled aperture proposed by M. Subotowicz (1978) for SETI can achieve satisfactory resolution and sensitivity to observe separate planets. This may be of great importance for astronomy and cosmogony.

## 5 SHORT-PULSE SETI

Search for narrowband radio emissions is widely recognized as most prospective method of carrying SETI experiments. It should be remembered however that transmitting civilization could choose another way to facilitate searching for emitted signals. As noticed in early 60's by B. M. Oliver ETI's signals may be rather narrow in time than in frequency (Oliver, 1962; for more detail discussion see f.e. Gindilis, 1979). Assuming that sending civilization is able to generate extremely short signals, emitted energy would in this case be spread over wide band of frequencies, what, in turn, would force us to fundamentally change strategy of SETI experiments. Such a possibility, although marginal may be, suggests carrying complementary searches in time regime, using wide-band receivers and systems able to reject local interferences. This scenario is relevant to recently originated research on astronomical short-time-scale phenomena. Beginning with discovery of pulsars they embrace now such various topics as ionospheric and interplanetary scintillations, microstructure of pulsars' emission, search for radio impulses which are expected to accompany stellar collapses and explosions of black mini-holes, and so on. All this phenomena have time scales ranging from about microseconds to tens of seconds. Radioastronomical experience in this field has been resumed by Phinney and Taylor (1979) and O'Sullivan (1978).

Most widely known SETI-related observations of this kind have been carried out from 1970 in USSR by groups under direction of V. S. Troitskiy (1974) and N. S. Kardashev (1976). According to Russian SETI strategy strong omnidirectional and short impulses are sought for at first with simple Yagis or spiral antennas and wideband receivers. Terrestrial interferences are eliminated by carrying synchronous observations at widely spaced stations up to 8000 kilometers. As far as we know these experiments have failed to detect signals which could be of both intelligent and extra-terrestrial origin. Nevertheless, the "natural" gain of similarly designed observations is considerable (see f.e. Troitskiy, 1979).

More recently, T. W. Cole and R. D. Ekers performed 5 GHz survey for impulsive radio events at CSIRO 64-m Parkes facility. Besides observing globular clusters, galactic centre etc., their programme

included preliminary search for possible ETI's sharply pulsed signals from several nearby G and K-type stars (Cole and Ekers, 1979). During about three days of antenna time the only non-explicable events was recorded while telescope was pointed at 82 Eridani (G5). Four hours of additional observations failed to confirm this single detection.

Experiments like this one clearly demonstrate that both hardware and observing procedures of searches for short-lived natural radio emissions are compatible with needs of SETI and vice versa. As writes Gindilis (1979): "... it is of interest that the search for ETI signals can be conducted concurrently with the detection of sporadic pulsed radio emissions arising during natural bursts..."

## 6 SETI AT NATURAL RADIOLINES

The main difference between conventional and SETI-oriented radio observation at frequencies of neutral hydrogen, hydroxyl radical etc. lies in achieved spectral resolution. While thus far performed mappings of HI distribution or observations of OH sources were done with resolution defined by the fact that narrowest natural emissions are of the order of 1 kHz, search for ETI signals will be carried out with resolution up to circa  $10^2$  Hz and probably much less. This contradiction leads to opinion that results of narrowband SETI, particularly these with bandwidth  $\sim 1$  Hz, will be useless from point of view of conventional radioastronomy since any spectral features will be overresolved (Anonymous, 1977). Different opinion has been presented in Murray's et al. paper (1978), where wide-area spectral-line surveys are envisaged as a by-product of SETI observations in molecular lines with high spectral resolution.

Scientific gains of SETI experiments would be maximized in case of combined programs like JPL's - Jet Propulsion Laboratory s- 1-22 GHz survey, embracing both: high sensitivity and broad-beam surveys of the whole microwave window. Spectral resolution of 300 Hz is at the same time well under thus far reached limits. Cuzzi and Gulkis (1977) enumerate and widely discuss radioastronomical applications of the system proposed by JPL SETI team. Quite novel feature of this program is that it will include surveys at high frequencies, not carried out previously. Scientific importance of observations under

$\lambda \approx 20$  cm has also been noticed by Russian SETI scientists who have carried out surveys on  $\lambda \approx 3,5$  cm, down to about  $10^{-26} \text{ W/m}^2 \text{ Hz}$  (Kardashev, 1976).

To say a few words about astronomical objectives of carrying SETI at frequencies of natural radiolines following may be stressed (for details see paper by Cuzzi and Gulkis, 1977):

- 1.42 GHz search: studies of distribution of HI, observations of absorption features in cold dust clouds, wider than previously performed surveys of high velocity hydrogen clouds, supplementary observations of neutral hydrogen in globular clusters, detection of hydrogen concentrations in extragalactic systems, particularly in Local Group galaxies,
- 1.62+1.72 GHz search: mapping of dense and cold molecular clouds, observations of proto-stellar sites as well as late type stars, measurements of galactic magnetic fields, studies on OH masers in various sources, etc.
- 4,8 and 14.9 GHz search: observations of dense regions of the Galaxy and clouds in spiral arms,

- 22.2 GHz search: broad beam surveys on this frequency could discover many new sources of H<sub>2</sub>O maser emission,
- 23. GHz search: detection of previously unknown sources of ammonia emission from dense interstellar clouds.

At all of above mentioned frequencies recombination lines of hydrogen could be observed as well as radio continuum emission generated by galactic plasma.

## 7 MAKING CATALOGUE OF THE STARS

There were proposed different programs especially for CETI-SETI, e.g. program CYCLOPS built of 2500 radiotelescopes of 100 m diameter, occupying the area of 7 to 20 km<sup>2</sup> on the surface of the Earth. The sensitivity of the CYCLOPS program would be 4 millions times larger than that of the program OZMA. The program CYCLOPS will investigate very precisely about one million stars in the sphere around the Sun of the diameter of about 1000 light years. In this kind of the activity one see the possibility to make very exact catalogue of the stars for astronomy, even in the case of complete failure of SETI. The F, G, K and N type main sequence stars may provide a habitable zones. They are stable for a sufficiently long time to allow life to originate and evolve into a technical civilization. Approximate coordinates of 324, 188 stars contains BD catalogue, 33,342 stars - the second general catalogue of Boss, one of the most fundamental catalogues.

## 8 SPACE-BASED RADIOTELESCOPES (SRT)

Apart from frequency limitations imposed on observations by absorbing properties of terrestrial atmosphere, there are at least three reasons for placing radio astronomical antennas outside the Earth:

- possibility of constructing SRT's with collecting areas for greater than it is allowed on Earth. It is of particular importance since parameters of contemporary receiving systems have virtually reached natural limits except just antennas area. Very large collecting areas would permit us to achieve extremely high sensitivities;
- interferometric space systems would offer baselines measured not in kilometers but in astronomical units, resulting in angular resolution many millions times better than possible with earthbound VLBI systems;
- still-growing contamination of radio spectrum caused by terrestrial interferences can be avoided, provided SRT is placed on sufficiently distant orbit or screened.

SRT were proposed by the Polish (1976, 1979), American (1976) and Russian (1979, 1979) groups.

Most hold project of such an installation presented by Russian scientists (Buyakas et al., 1977, 1979) envisions antenna diameter reaching 10 kilometers. Potential of an interferometer system consisting of two similiar dishes is breathtaking: sensitivity on the order of 10<sup>-10</sup> Jy, angular resolution depending on baseline, equal to about 10<sup>-10</sup> arc seconds for 10 AU baseline. It is quite clear that observing tool like that would revolutionize astronomical research. Holographic imagery of even most distant sources of radiation would be realizable, to mention only one hitherto unthinkable possibility (Kardashev, 1973). It is of particular interest to SETI that

10 km diameter space-based antenna would be able to detect 100 Hz bandwidth 1 MW omnidirectional transmission from 100 ly, with range of detectability growing sharply in case of higher gains on senders side. Astroengineering constructions could be discovered even on intergalactic distances by their own IR radiation or by the effect of screening background emission. These capabilities would make SRT an ultimate tool of both astronomy and SETI. At the same time it should be mentioned that SRT is much cheaper than Cyclops-like ground-based facilities, particularly when equivalent diameters exceeding 2 km are considered (see fig. 12 in Buyakas et al., 1979).

## 9 SETI-CETI TECHNOLOGY ADVANCES

As noticed by many investigators (see f.e. Murray et al., 1978) technology developed specifically for SETI programmes could have important benefits for not only astronomy but also for other endeavors, such as e.g.:

- communication with distant space probes which is important for navigation, celestial mechanics etc. , improved thanks to broadband and low-noise preamplifiers required for SETI. Data on atmospheric degradation of radio signals gathered during SETI searches would also be of importance for deep space communication links;
- data management technology developed for SETI would permit to realize all those projects which are characterized by flood of data, such as spectroscopic monitoring of the environment;
- rich information about radio emissions of anthropogenic origin would be gathered as a result of RFI's identification procedure necessary for SETI observational practice.

In an astronomical aspect two elements of planned SETI systems are of particular importance:

- 300 MHz instantaneous bandwidth receivers operating within terrestrial microwave window, envisioned for JPL (all sky) SETI programme could substantially aid study of molecules in interstellar space;
- introducing mega-channel spectral analyzers will mean  $10^3$  order improvement over analyzers presently available for astronomers.

## 10 SETI-CETI FOR BIOLOGY

The influence of CETI-SETI on the development of the biology can be analysed in two aspects: a) search for the abiogenic organic molecules in space and atmospheres of the stars as the simplest manifestation of the chemical evolution, b) how to understand and to repeat the natural chain of the evolution processes on the Earth of the compound molecules that lead to the living organism with his very complicated metabolism, based on the principal bioelement - carbon and biosolvent - water. The undertaken CETI-SETI program stimulated and will stimulate the fascinating investigations on the origin of life. On the other side, progress in the understanding of the origin of life will stimulate the CETI-SETI efforts.

## 11 CETI-SETI FOR MANKIND

The word "culture" in the title of the present paper contains different social and philosophical problems. It is reasonable to create

new discipline: civilization-science that should deal with the formation of the society, its technical and historical development, the evolution of the different social conditions, creating culture and civilization. It is not excluded that ETI's science is quite different than this on the Earth. There arise the problem of universality and antropomorphicity of our knowledge and of our picture of the outer world. What are the norms of the human thinking and features of the human language? How they arised? Are there possible different norms of thinking and what could be the kind and origin of language of ETI, different from those known on the Earth? What is the life-time of the scientific-technical civilization? What are the laws of the evolution and development of the scientific-technical civilization? Which kind of language should be used in CETI-SETI and what content should be communicated in CETI-SETI?

The answers to these questions might explain us problems of the a) antropomorphicity and universality of our picture of the outer world, b) norms of our thinking, c) nature of our cognition and knowledge (Subotowicz, 1976b).

It would be worth to do the above work even then when our CETI-SETI efforts would turn out to be completely unsuccessful!

The scientific methods of the formulation of the scientific hypothesis and theories should determine to which extent are correct the hypothesis on ETI and do they exhaust all the possibilities of the origin in time and space, development and decline of the scientific-technical civilization (STC). The scientists accept the homogeneous distribution of ETI's in the Galaxy. It is also generally accepted that there is no influence of the life, its evolution and social development on the macroscopic events in the Universe on the scale not larger than that of the planet. It is not true if there exist the Dyson's and Kardashev's - type civilizations. These STC can be more advanced than that on the Earth on thousands or even millions years. How far away in norms, ideas and culture are those STC from that of our? It is very interesting question, to which extent the picture of the outer world determined by our science is connected with the biological and social development of the mankind here on the Earth. The metodologically correct attitude when analysing the origin of any received signal should be: to try to explain it as the natural one. It means that the signals are generated by the inorganic matter and in natural phenomena. This is the general attitude of astronomers, astrophysicists, radioastronomers and of all the physicists. They accept the natural astronomical processes as the incomparably more probable than the artificial ones, realised by the wise beings. It is generally accepted that their rational activity can change neither the laws of the nature nor the direction of the evolution of the matter (in large scale). We should emphasize that our division of the processes as the "natural" and "artificial" is relative one and depends on the level of the contemporary science and knowledge. Some of the people raise the question, is it possible to astronomers to discover the "cosmic miracle" or the artefacts made by the representatives of the STC. The attitude of the astronomers is based on the general methodological supposition that all the observed phenomena are the result of the natural evolution of the matter.

Does this attitude contain the methodological antropomorphism and geocentrism of the terrestrial scientists? Earth and its history and the development of the science and techniques on the Earth can be for the terrestrial scientists the intransgressible model standard. These scientists can underestimate the possible variety of the

realization of life and that of the history of the biological and social development of the living organisms - from the simplest ones to those of the homo sapiens type, able to create STC.

### REFERENCES

- Anonymous (1977), in: The Search for Extraterrestrial Intelligence, NASA SP-419, 207.
- Atchley D.W. (1960). Speculations on communications with other planet civilizations, QST, 44, No. 3, 71.
- Basler R. P., G. L. Johnson, and R. R. Vondrak (1976). Parametric Study of Interstellar Systems, Stanford Research Institute, Menlo Park, Calif., USA, p. 1-193.
- Basler R. P., G. L. Johnson and R. V. Vondrak (1977). Antenna concept for interstellar search system. Radio Science, 12, 845.
- Baum G. (1977). Neutron stars and the properties of matter at high density. Copenhagen, Nordita.
- Baum W. A. (1980). The ability of the Space Telescope to detect extra-solar planetary systems. Celestial Mechanics, 22, 183.
- Black D. C., ed. (1979). Project Orion, NASA SP-436.
- Bracewell R. N. (1978). Detecting nonsolar planets by spinning infrared interferometer. Nature, 274, 780.
- Buyakas and 22 co-authors (1977). Infinitely expandable space radio-telescope, paper presented at an IAF Congress in Prague, Czechoslovakia, (1977) and "Acta Astronautica", 6, 175 (1979).
- Chamberlain, J. W. (1978). The geophysical importance of searching for intelligent life in the Galaxy, FOS-Trans.American Geophys. Union, 59, 900.
- Cole T. W., and R. D. Ekers (1979). A survey for sharply pulsed emissions, Proc. Astron. Soc. Australia, 3, 328.
- Cuzzi J. N., and S. Gulkis (1977), in: The Search for Extraterrestrial Intelligence, NASA SP-419, 147.
- Drake F. D. (1965). The radio search for intelligent extraterrestrial life, in: Current Aspects of Exobiology G. Mamikunian, M. H. Briggs, editors, Pergamon, Oxford, 323.
- Edelson R. E., S. Gulkis, M. A. Janssen, T. B. H. Kuiper, and G. A. Morris (1978). The SETI radio observational project: Strategy, instrumentation and objectives, 1978 IAF Congress, Paper 78-A-43, 2.
- Gindilis L. M., N. S. Kardashev, V. A. Soglasnov, E. E. Spangenberg, V. S. Etkin, and V. G. Mirovskiy (1979). Search for signals from extraterrestrial civilizations by the method of synchronous dispersion reception, Acta Astronautica, 6, 95.
- Goldberg L. (1977). Report on NGT project cited in: Sky and Telescope, 53, 428.
- Kardashev N. S. et al. (1973). Izvestiya SAO, 5, 16 in Russian .
- Kardashev N. S. (1976). Recent CETI investigations in USSR, Space Research Institute Moscow, preprint no. 279 in Russian .
- Klimshin I. A. (1980). Astronomia naszych dniej. Moskwa, Nauka, wyd. II.
- Knowles S. H., and D. L. Thacker (1980). Applications of interferometers in space to astrometry and planetary detection. Celestial Mechanics, 22, 197.
- Kuchowicz B. (1979). Kosmochemia, Warszawa, PWN.
- Morrison P., J. Billingham and J. Wolfe (1977) ed. The Search for Extraterrestrial Intelligence - SETI, NASA SP-419.

- Murray B., S. Gulkis, and R. E. Edelson (1978). Extraterrestrial intelligence: An observational approach. Science, 199, 485.
- Oliver B. M. (1962). Radio search for distant races, International Science and Technology, October 1962, 55.
- Oliver B. M. (1979). Project Orion, NASA SP-436.
- O'Sullivan J. D., R. D. Ekers, and P. A. Shaver (1978). Limits on cosmic radio bursts with microsecond time scales. Nature, 276, 590.
- Phinney S., and J. H. Taylor (1979). A sensitive search for radio pulses from primordial black holes and distant Supernovae, Nature, 277, 117.
- Sciama D. W. (1971). "Modern Cosmology", Cambridge, Univ. Press.
- Subotowicz E. W. (1974). Izotopnaja kosmochimia, Atomizdat, Moskwa.
- Subotowicz M. (1976a). "Ceti from an Earth satellite orbit", paper presented at an IAF Congress in Anaheim, California, 1976, and published in JBIS, Interstellar Studies Issue, 31 (1978), 109-110.
- Subotowicz M. (1976b). Problems of search of extraterrestrial civilizations (in Polish). Postepy Astronautyki, 9, No. 4, 7-22.
- Subotowicz M. (1978). On the new types of space satellite radio-telescopes (SRT) for radio astronomy and CETI. XXIX IAF Congress, Dubrovnik, 1-8.X.1978.
- Subotowicz M., J. Usowicz and Z. Paprotny (1979). On active and passive CETI from an Earth satellite orbit, Astronautica Acta, 6, 203-212.
- Shkłowski I. S. (1975). "Zwiezdy, ich rożdjenije, żizn' i smiert'", Nauka, Moskwa.
- Shkłowski I. S. (1980). Swierchnowyje zwiezdy, Nauka, Moskwa.
- Troitskiy V. S. et al. (1974). Uspekhi Fizicheskikh Nauk, 113, 719.
- Troitskiy V. S., A. M. Starodubtsev, L. N. Bondar (1979). Search for radio emissions from extraterrestrial civilizations. Acta Astronautica, 6, 81.
- Verschuur G. (1978). Sky and Telescope, 56, 551. Review of "The Search for Extraterrestrial Intelligence" NASA SP-419, (1977).
- Zuckerman B., and J. Tarter (1979). "Microwave searches in the USA and Canada", in Proc. of IAU Joint Session on Strategies for the Search for Life in the Universe, Aug. 1979 Symposium, Montreal, Canada, D. Reidel, Publisher.

## THE DEMOGRAPHY OF EXTRATERRESTRIAL CIVILIZATIONS

J. Billingham\* and C. L. Seeger\*\*

\**NASA Ames Research Center, Moffett Field, California, USA*

\*\**San Francisco State University, San Francisco, California, USA*

### ABSTRACT

This paper reviews many of the studies carried out in the last twenty years on the possible nature and distribution of extraterrestrial intelligent life. The significance of the review is the extent to which conclusions of these studies may determine the direction of programs to search for extraterrestrial intelligence (SETI). Extreme views have been presented by some authors.

### KEYWORDS

Life in the Universe; interstellar travel; planetary system formation; emergence of intelligent civilizations; search for extraterrestrial intelligence; SETI.

### INTRODUCTION

A major scientific question of our time concerns the nature and distribution of intelligent life in the Universe. The underlying arguments for the existence of extraterrestrial civilizations have been presented with increasing richness in numerous papers over the last two decades (Billingham and Pesek, 1979). In many instances the arguments have focussed on the equation proposed by Drake as a useful tool for addressing the questions involved in trying to decide on values for N, the number of communicative civilizations currently in the Galaxy.

$$N = R_* f_s f_p n_e f_e f_i f_c L$$

In Drake's formulation, N is related to the average rate of star formation in the Galaxy ( $R_*$ ), the suitability of the stellar system ( $f_s$ ), the fraction of stars with planetary systems ( $f_p$ ), the number of planets in the life zone of each star ( $n_e$ ), the probability of the emergence of life on suitable planets ( $f_e$ ), the fraction of biological systems developing an intelligent species ( $f_i$ ), the number of these attempting communications ( $f_c$ ), and the average lifetime of such a communicative civilization (L). The last factor appears to have the greatest uncertainty.

It is widely recognized that the Drake formula is a simplified description of

what may very well be, in reality, a very complex, probabilistic situation. Attempts have been made to expand his expression to include probability density functions for each of the factors (e.g., Billingham, Oliver and Wolfe, 1979). While the degree of uncertainty increases greatly as one passes from left to right through the equation, what is perhaps more important is that estimated values of the different factors vary enormously according to the views of those making the estimations, even among those who have studied the problem at length. Some conclude the inevitability of communicative life is so high that there may be tens of millions of civilizations coexisting in just our galaxy. Others, assigning extremely low values to one or more factors, conclude that the number is likely to be small; even to be one — our own. And when some factors are held to be vanishingly small, it follows that we may even be alone in the Universe.

Our ignorance of these matters is so very great that a variety of studies are clearly indicated. Among these is the direct exploration of the Universe for electromagnetic signals due to the actions of another intelligent species. However, in developing strategies to carry out searches for other civilizations, it is surely important to estimate anything we can about their demography. Because the matter continues in lively dispute, this paper reviews some of the many discussions which have taken place over the last few years.

#### A VARIETY OF OPINIONS

The Drake equation in simplest form deals with civilizations arising *de novo* on planets of different solar systems. Over the last few years increasing numbers of authors proposed that some civilizations would be bound to develop the capability for interstellar travel, and would necessarily propagate themselves from star system to star system. If this is the case, then the number of inhabited solar systems should increase over estimates derived from the Drake equation. Some authors, continuing with the interstellar travel concept, have suggested that the first civilization or, surely, one among the first to achieve the capability for interstellar flight, would expand to occupy the entire Galaxy in a time which is short on geologic scales.

Two extreme views now emerge. The most prevalent is that a spacefaring civilization of the type described above should have emerged eons ago, and should have occupied the entire galaxy a short time thereafter. Therefore they should have reached the Solar System a long time ago and we should now observe them here, or be descended from them. But we appear, in the dominant view, to have emerged *de novo* on the Earth, and we do not observe others among us. So, the argument goes, we must be the first and only intelligent species in the Galaxy. This argument has recently been summarized extensively by Tipler (1980), and was discussed at length at a meeting in College Park, Maryland, in 1979. The title of the symposium was "Where are they? A symposium on the implications of our failure to observe extraterrestrials" (Hart and Zuckerman, *in press*). This view is similar to that held by Hart and other authors, of which more below.

The other extreme view, held by Papagiannis (1978, and 1980), begins with the same arguments about interstellar travel. However he presents the view that intelligent life may have arisen elsewhere, propagated throughout the Galaxy, and arrived in the Solar System at some undefined time in our past. He suggests that such a spacefaring group would be adapted to living in space settlements (O'Neill, 1978), and would have settled in some convenient location where ample supplies of raw materials and energy exist. Indeed, he proposes that they may have settled in the asteroid belt, preferring such a location to the surfaces of planets or moons; that they are there now and might be located by an appropriate

search. If they are not found in our solar system, Papagiannis argues, then the other scenario should hold, and we must be one of the very few if not the only advanced civilization in the entire Galaxy.

Either of these views rests on a number of major assumptions. A significant one is the notion that interstellar travel can be achieved at relatively high speeds (compared to current human expectations) in very large vehicles and on a continuing basis by one or more parent species and their colonial offspring. Many authors, for example Oliver (1972, 1980) and Goldsmith and Owen (1980), find this concept difficult to accept. They point out that high speed travel with large payloads must consume prodigious amounts of energy, and they are not so confident that future developments will solve the technological and energetic problems, now foreseen, in a fashion that is both attractive and economically acceptable to an intelligent society, unless it must escape some threatening catastrophe -- an incipient supernova close at hand, for example.

Considering the possibility of interstellar travel at speeds on the order of a tenth that of light, or more, suggests the possibility of collisions with bits of interstellar matter. There is, so far, negligible information on the Galactic distribution of matter in the gram-mass range, yet there is reason to suspect it exists in some quantity. Neither does there appear to be an understanding of the collision physics involved. Does a one gram mass at a relative speed of 30,000 km/s merely drill a neat hole? Or, is there a spectacular explosion? At such speeds, collision avoidance is also likely to be a highly energetic matter. So those who assume galaxy-colonizing space traffic speeds below a few hundred km/s seem to be not only mindful of fuel costs, but are also exhibiting admirable caution.

Interstellar travel at rates far below the speed of light does appear to be possible, given our present understanding. Therefore, one may assume it is common in the Galaxy or that it is extraordinarily rare, or anything between these extremes. Any assumption in this range, depending as it clearly does on motivation, would seem to be relatively soft, and a prominent, introductory "If" seems indicated.

Tipler's argument (1980) is now given. Under the declarative title, *Extraterrestrial Intelligent Beings do not Exist*, he writes:

"The basic idea of my argument is straightforward.... if they did exist and possessed the technology for interstellar communication, they would also have developed interstellar travel and thus would already be present in our solar system. Since they are not here (14, 15), it follows that they do not exist. Although this argument has been expressed before, its force does not seem to have been appreciated. I shall try to rectify this situation by showing that an intelligent species with technology for interstellar communication would necessarily develop the technology for interstellar travel, and this would automatically lead to the exploration and/or colonization of the Galaxy in less than 300 million years."

The remainder of this paper is a clear and thorough development of this not uncommon point of view, and should be read in preference to short, often misleading summaries that have appeared in both the scientific and public presses. Here, we comment briefly on what appear to us to be two central ideas in Tipler's paper which, if support for them could be found, would tend to substantiate the paper's title.

"Since they are not here (14, 15), it follows that they do not exist." The two

references in support of this declaration are Klass (1974) and Menzel and Taves (1977). Both these books are discussions of the unidentified flying object syndrome. More generally, the essence of the matter is this. There is, so far, no recognizable physical evidence reasonably acceptable to the scientific community, of the presence of extraterrestrials on or near the Earth or in the Galaxy, now or in the past. This is no more than apparently true, though it may be true in fact, since it is not the result of suitable experimentation and exploration. In such a case, absence of evidence is just that, and no more. It proves nothing, one way or the other, about the possible existence of extraterrestrial intelligent technological species, communicative or not.

Second, there seem to be no prior examples, nor any demonstrable technical and social processes, which "necessarily and automatically lead to the exploration and/or colonization of the Galaxy". Recorded history testifies overwhelmingly in favor of the judgement that predicting humanity's future is not one of humankind's better developed skills. To connect such a prediction to the behavior of hypothetical species is an extrapolation indeed. Because humans may be unable to avoid a significant degree of anthropocentrism generated by their past, questions about future human motivations in this area, let alone extra-terrestrial motivations, are probably largely indeterminate until after the consequences of the actual motives have been observed.

Faced with the apparent energy and travel time problems inherent to space voyaging, Bracewell (1975) proposed that civilizations might send automated probes to explore other star systems. Having made the desired observations, the data would be relayed back to us. This approach appears to be less costly, per probe (and many might be required), than the "manned" interstellar vehicle concept, but it, too, suffers from some possible drawbacks. The further the probes go, the longer it takes to get information back to the home planet. And the energy costs still remain extraordinarily large.

In this connection, Tipler (1980) writes that when we can build a von Neumann multiple selfreplicating universal constructor — and he believes that will be achieved in a hundred years or so — the problem of supplying many probes to explore the Universe will be solved. Would we dare to launch it? It might evolve on its own, and to our detriment. And suppose there is other life in the Galaxy. Can we imagine how it might appreciate our launching a plague of von Neumann machines on the Universe? With less than ultimate knowledge, could we design such a machine to avoid the possibility of its being unwelcome? There is no universal rule that what technology permits must be; and the problem here is philosophical, not scientific, not technological.

Newman and Sagan (1981) use potential theory to reexamine the interstellar diffusion of Galactic civilizations. Numerical and analytical solutions are derived for the nonlinear partial differential and difference equations which specify a range of models. They find that progressive voyages over long galactic distances could take many hundreds of millions of years. They conclude that

"Earth is uncolonized not because interstellar spacefaring societies are rare, but because there are too many worlds to be conquered in the plausible lifetimes of the colonization phase of nearby galactic civilizations. ...There may, however, be abundant groups of  $10^5$  to  $10^6$  worlds linked by common colonial heritage. The radar and television announcement of an emerging technical society on Earth may induce a rapid response by nearby civilizations, thus newly motivated to reach our system directly rather than by diffusion."

Jones (1981), using Monte Carlo calculations, concludes that Galactic filling

times on the order of 60 million years seem probable. He expects that "our descendants will be sufficiently clever to devise a bit of 'magic' so that interstellar voyages can be undertaken at relatively minor cost." And he expects that humanity's drive to expand its living space is likely to continue until it has filled the Galaxy, unless another species achieves this first. He argues from his view of human history that "There is a point that is sometimes missed, however, and that is that only the scale of human activities has increased, not the kind."

Jones is not alone in being attracted to the idea that the first intelligent species to achieve practical space travel capability is fated to expand to the limits of the Galaxy; and, therefore, since they are not here, it is now our unavoidable future. But there are problems with this concept. It requires motivation for continued expansion to be steadfastly maintained as a civilization evolves and changes over long periods of time. It implies a particular and rather narrow interpretation of human history. It assumes that colonizing the Galaxy is no different from colonizing the Western Hemisphere on Earth. It assumes that though the motive for expansion is enduring, the motive for wars between the colonies and their parent society will be suppressed, and so forth. Thus, it seems dangerous to extrapolate from the experiences of human societies which have just reached the stage of space travel within a solar system to the behavior patterns of much more advanced societies.

Turning back from ideas about interstellar travel to the Drake equation itself, some authors have argued that one or more of the terms in the equation may have values that are extremely small. Hence  $N$  is likely to be small.

Kumar (1979) holds that single stars may be uncommon, hence planetary systems similar to our own cannot be a universal phenomenon. This supposition is a subject of much debate. Many astronomers believe that the formation of planetary systems is a normal stage in the process of the condensation of solar nebulae, and that this may be the case for binary or multiple star systems as well as for single stars. Note that we are technically on the threshold of being able to detect the planets of other stars (Black, 1980a,b). So at least this uncertainty should be resolved over the next few decades. A further argument is that conditions on the surface of planets of binary or multiple star systems are likely to be unsuitable for the origin and evolution of life because of the wider variations of environmental variables compared with those on the planets of single stars. The problems with this argument are two-fold. First, reasonably stable environments may be present on the planets of widely separated binaries or multiple star systems (Harrington, 1977). Second, we do not yet know whether biological evolution is accelerated or retarded by fluctuating environmental conditions which may be present in the environments of planets of binary or multiple star systems. To give an example from our own system, some believe that the tidal forces induced on the Earth by the moon may have been an important factor in stimulating evolution on the Earth. Correspondingly, the daily and yearly climatic changes induced by Earth's rotation and revolution about the Sun could be an important factor in biological evolution. Life in the polar regions of the Earth has adapted well to the very considerable environmental changes between winter and summer.

Some have argued that it is most unlikely that planets having the correct initial composition and conditions for the origin of life would be placed precisely at the right distance from their stars so that the very long term climatic stability necessary for the evolution of life might be maintained. Others have disagreed, and think that the "ecoshell" around stars having terrestrial planets could be of considerable thickness. Schneider and Thompson (1981) have emphasized our ignorance:

"But we conclude that cosmic conclusions from climatic models should be accompanied by clear admission of the vast uncertainties in the climatic component of the argument, let alone the other parts of the problem."

(Note that their paper is one of a collection in a volume entitled, "Life in the Universe" (Billingham, 1981), which deals with many of the questions raised here.)

Hart (1975) uses probabilistic calculations of the chances of random associations of molecules coming together to form DNA chains as a basis for saying that the origin of life itself must be rare. According to his calculations, even with "optimistic" assumptions included, he surmises a probability for the origin of life of  $10^{-32}$ . Many exobiologists believe that the probability may be close to one for a suitable planet. The problem with Hart's analysis is that he does not allow certain molecular configurations to be selected, in the evolutionary sense, because of their inherent properties for replication through autocatalysis. The same argument, namely that of Darwinian evolution, holds throughout the entire history of biological advances of ever increasing complexity, culminating in intelligence, cultural evolution, and civilization. Other terrestrial planets, similar to the Earth, orbiting around other F, G or K stars, can be expected to harbor living systems which similarly gradually increase in complexity. Indeed, it may be that the physical and chemical environment of the Earth's biosphere, and the changes in it over time, may be far from optimal for the emergence of complex life forms and intelligence.

We expect that the time for the genesis of intelligence might be much shorter on some planets of some stars. Conversely, in other locations the time may be much longer. We do not yet know which physical, chemical and environmental factors accelerate evolutionary processes. Questions of this type are now being asked. As a research endeavor, this field is likely to be fertile.

Lovelock (1979) has recently put forward the Gaia hypothesis. He proposes that living systems modulate the planetary environment in such a way as to establish conditions favorable for their own stability and evolution. Examples are the evolution of the oxygen atmosphere on Earth, and the regulation of terrestrial CO<sub>2</sub> by increase or decrease in photosynthesis. While there is much debate about the hypothesis, it is clearly a possibility. Those who believe that life is rare for the reason that planetary environments tend to diverge with time in directions increasingly incompatible with life, as may have happened on Mars or Venus, should consider that such tendencies may be modified by life itself in places where life has emerged.

The uncertainty associated with N is clearly large. Our knowledge about the patterns of distribution of civilizations is limited. Certain statements are perhaps reasonable. For example, it is not likely that life, and certainly intelligent life, will emerge on planets of O stars, since their own lifetime is so short. Equally, life will not yet have emerged in any very young solar systems. The galactic center, with high radiation fluxes and intermittent explosions, is not promising as a life site. Life is most unlikely to have evolved very early in the history of a galaxy before carbon, oxygen, nitrogen and other biogenic elements had been produced. Life may not survive massive supernova explosions that are very close. Life is unlikely to evolve to any stage of complexity on gas giants or on small cold planets far out in solar systems. Beyond these reasonable deductions about the demography of extra-terrestrial civilizations, the arguments become more and more difficult, and we return to the Drake equation.

In the light of our present knowledge about the immense number of stars, stellar and planetary formation, the ubiquity of chemical evolution, and the generality of the principles of biological evolution, it would seem most unlikely that we are the only intelligent species. We therefore put forward the hypothesis that civilizations, young and old, are likely to be scattered through the galaxy.

The argument about the number and distribution of intelligent species currently coexisting in our galaxy can only be settled in the foreseeable future by the unequivocal detection of one or more examples.

In our view the likelihood of detection is remote in the absence of a well structured and thoroughgoing search program (Wolfe, et al., 1981).

#### REFERENCES

- Billingham, J., (Ed.) (1981). *Life in the Universe*. NASA Special Publication 456; also, MIT Press, Boston (Nov. 1981).
- Billingham, J., and R. Pesek (1979). *Communication with Extraterrestrial Intelligence*. Acta Astron. 6,1-2. Pergamon Press.
- Billingham, J., B. M. Oliver and J. H. Wolfe (1979). *A Review of the Theory of Interstellar Communication*. Acta Astron. 6,1-2, 47-57. Pergamon Press.
- Black, D. C. (1980a). In *Search of Other Planetary Systems*. Space Science Reviews, 25, pp.30-81.
- Black, D. C. (1980b). On the Detection of Other Planetary Systems: Detection of Intrinsic Thermal Radiation. *ICARUS*, 43, pp.293-301.
- Bracewell, R. N. (1975). The Galactic Club. Freeman Pub. Co.
- Goldsmith, D., and T. C. Owen (1980). The Search for Life in the Universe. Benjamin/Cummings Pub. Co.
- Harrington, R. S. (1977). Planetary Orbits in Binary Stars. *Astronomy Journal*, 82, 753.
- Hart, M. H. (1975). An explanation for the absence of extraterrestrials on Earth. Q.Jl.Roy.Astr.Soc. 16, 128.
- Hart, M. and B. Zuckerman (1981, in press). Extraterrestrials: Where are They? Pergamon Press.
- Jones, E. M. (1981). Discrete Calculation of Interstellar Migration and Settlement. *ICARUS*, 46, 328-336.
- Klass, P. J. (1974). *UFOs Explained*. Random House, New York.
- Kumar, S. (1979). In *Conference Abstracts: Where are They? A Symposium on the Implications of our Failure to Observe Extraterrestrials*, held at College Park, Maryland.
- Lovelock, J. E. (1979). Gaia, A New Look at Life on Earth. Oxford Univ. Press.
- Menzel, D. H., and E. H. Taves (1977). *The UFO Enigma*. Doubleday, Garden City, New York.

Newman, W. I., and C. Sagan (1981). Galactic Civilizations: Population Dynamics and Interstellar Diffusion. *ICARUS*, 46, 293-327.

Oliver, B. M. (1980). Strategies for the Search for Life in the Universe. M. D. Papagiannis (Ed.), p.79-80.

Oliver, B. M., and J. Billingham (1972). Project Cyclops. A Design Study of a System for the Detection of Extraterrestrial Intelligent Life. Report CR 114445, NASA Ames Research Center, Moffett Field, Calif.

O'Neill, G. (1978). The High Frontier. Bantam, New York.

Papagiannis, M. D. (1980). Strategies for the Search for Life in the Universe. Astrophys. & Space Sci. Library, 83. E. Reidel Pub. Co.

Schneider, S. H., and S. L. Thompson (1981). Cosmic Conclusions from Climatic Models: Can They Be Justified?, in Life in the Universe, J. Billingham, ed. NASA Special Publication 456; also, MIT Press, Boston (Nov. 1981); also *ICARUS*, 41, 456-469 (1980).

Tipler, F. J. (1980). Extraterrestrial beings do not exist. O.Jl.R.Astr.Soc. 21, 267-281.

Wolfe, J. H., et al.(1981). SETI, the Search for Extraterrestrial Intelligence: Plans and Rationale, in Life in the Universe, J. Billingham, ed. NASA Special Publication 456; also, MIT Press, Boston (Nov. 1981).

## RECENT PROGRESS IN EXOBIOLOGY AND PLANETARY BIOLOGY

T. H. Jukes

*Space Sciences Laboratory, University of California,  
Berkeley, CA 94720, USA*

### ABSTRACT

Exobiology is the name given to investigations into the possible characteristics of extraterrestrial life. No such forms of life are known to exist. The investigations have included search for life in the solar system beyond our own biosphere and detection of molecules of biological significance in meteorites, comets and interstellar space. Other topics studied include early evolution on the Earth, possibilities for the origin of life, characteristics of organisms that can withstand environmental extremes, and the possible existence and characteristics of life on Earth-like planets beyond the solar system.

### KEYWORDS

Exobiology; planetary biology; extraterrestrial life; evolution

The subject of exobiology is based on the idea that life may exist in other places beyond the Earth and its atmosphere. This idea has interested people for many centuries, and it is a strong stimulus to the exploration of space. Curiosity is an inherent trait in human beings, and probably played a part in our evolutionary success. Some people say that it is hopeless to look for extraterrestrial life, and also that the search is a waste of money when funds are so urgently needed for many other purposes. Against this, we argue that new ideas give rise to new technology that is valuable for unexpected purposes. But the real reason for searching for life elsewhere is to increase our knowledge of the universe and of ourselves.

Many writers have speculated on the possibility of strange forms of life based on unknown biochemistry. Current thinking in exobiology tends to disregard such fanciful hypotheses. For this, we are sometimes criticized as having a parochial or earth-bound viewpoint. I do not agree with this criticism. Spectroscopic examination of distant stars and galaxies, carried out by astronomers, shows that the same chemical elements are present throughout the universe. We know the properties of the elements, because they are present on Earth, and have been subjected to exhaustive experimental manipulation, in their elemental form, and in the form of compounds. Certain elements predominate in living organisms on the Earth, especially carbon, hydrogen, oxygen, nitrogen, phosphorus and sulfur. All

terrestrial organisms contain the same amino acids, and several of these amino acids have been detected in carbonaceous chondrites from outer space. When no living organisms were detected on Mars, this coincided with the fact that the chemistry, temperature and radiation flux of the Martian surface and atmosphere were hostile to life as we know it (Horowitz, 1977; Mazur and co-workers, 1978). A dilemma that confronts us is first, does life become adapted to unearthly conditions, such as the absence of liquid water, on other planets?

Or second, have other planets, beyond the solar system, become adapted to produce the conditions for life as we know it?

Studies of the solar system give us no reason to believe that life exists elsewhere than on the Earth. This finding by space scientists confronts the human species with an overpowering responsibility to work for peace and for the safeguarding of the terrestrial environment.

My next question is: Does evolution inevitably lead to development of intelligent species? Evolution in Gondwanaland led to the development of marsupials. Kangaroos do not search for extraterrestrial intelligence. Neither do the spider monkeys of South America, another of the great continental masses. Only in the third major continent did evolution produce an intelligent species.

Many biologists say that the answer to the above question is "no." However, the multiplicity of opportunities in the Universe says "yes."

One line of speculation in exobiology is concerned with the extent to which extra-terrestrial life could differ from life on Earth with respect to its biochemistry. For example, there are 20 and only 20 amino acids used in synthesis of proteins by living organisms on the Earth. Yet other amino acids are known, and several of these have properties (Dickerson, 1978) which make them entirely suitable for protein synthesis (Table 1). Some of these "non-protein" amino acids have been found in meteorites. There is a possibility that forms of life existing elsewhere might use some amino acids for protein synthesis that are not on the "terrestrial list." This could happen without such organisms being noticeably different from terrestrial organisms.

Table 1 Protein and A Few "Non-Protein" Amino Acids

Protein			Non-Protein
Alanine	Glycine	Proline	$\alpha$ -Aminoisobutyric acid
Asparagine	Histidine	Serine	$\alpha$ -Amino- <i>n</i> -butyric acid
Aspartic acid	Isoleucine	Threonine	Isovaline
Arginine	Leucine	Tryptophan	Norleucine
Cysteine	Lysine	Tyrosine	Norvaline
Glutamic acid	Methionine	Valine	Ornithine
Glutamine	Phenylalanine		Many other simple variations of the protein amino acids are possible

When it comes to the genetic code, discoveries within the past 5 years have shown

that the code is not always the same (Table 2). This throws the question of the genetic code in extraterrestrial organisms wide open. It is no longer necessary to believe that there is only one genetic code, even on the Earth, let alone elsewhere.

Table 2 Differences between Universal Genetic Code and Mitochondrial Genetic Codes

Codons	Amino Acids	
	Universal	Mitochondrial Differences
CUU, CUC, CUA, CUG	Leucine	Threonine, yeast
AUA	Isoleucine	Methionine, human, Neurospora
UGA	Chain termination	Tryptophan, human, yeast, Neurospora
CGG	Arginine	Tryptophan, maize
AGA, AGG	Arginine	Chain termination, human

It seems quite possible that, if life exists on other Earth-like planets in this and other galaxies, its existence would depend on the presence of water and on a range of temperatures similar to those on our own planet. Our best guess is that the organisms would depend on proteins and nucleic acids for their biochemistry and their heredity. The properties of DNA make it certain that evolution will take place in its presence (Jukes, 1980).

DNA has the function of carrying inherited information from generation to generation without making mistakes. Nevertheless, this process is not error-free, and a small number of changes occur. Most of these are harmful, and are discarded. The changes that persist are what make evolution possible.

Proteins can have an almost infinite variety of different compositions and structures. The functions of proteins depend on their structures. Proteins are responsible for the composition of living organisms. All sorts of different biochemical pathways, catalysts, carbohydrates, fatty substances and co-enzymes could occur in life in other worlds.

#### The Search for Life on Mars

In 1976, the Viking landers on Mars carried three sets of experiments designed to detect life in the surface soil, existing as micro-organisms, or resulting from their presence. All three sets of equipment functioned, and returned information to Earth. The consensus is that the responses obtained were nonbiological. The results were discussed by Horowitz (1977) and by Mazur and co-workers (1978). Horowitz, whose experiment on pyrolytic release was carried by the Viking to Mars, appraised the results of all three experiments. He said that "even though some ambiguity remained, there is little doubt about the meaning of the observations of the Viking landers: At least those areas on Mars examined by the two spacecrafts are not habitats of life."

The findings of the Viking experiments are attributed to the high reactivity of

the Martian soil, which perhaps is activated by exposure to solar ultraviolet radiation. As a result, chemical interactions took place between the soil and the reagents used in the Viking tests. At the Mars Colloquium in Pasadena, on September 1, experiments were repeated showing that iron-montmorillonite could catalyze the labeled-release Viking experimental procedure by reacting with formate.

One of the remote hopes for life in the solar system was Titan, a satellite of the planet Saturn, because it has an atmosphere and is large in size. It has a diameter of 5,000 km, and is larger than the planet Mercury. The cameras of Voyager I showed that Titan's surface was totally obscured by a thick layer of atmospheric haze (NASA, 1980).

Data from other sources gave further information on the atmosphere of Titan. It is remarkably similar to that of the Earth in some respects, but radically different in others. Both atmospheres are primarily nitrogen. Both atmospheres exert about the same pressure on their surfaces. Titan's atmospheric pressure is about 50% greater than that of the Earth.

However, Titan's atmosphere contains no free oxygen, and there is a significant amount of methane. The most outstanding difference between the Earth and Titan is that the surface temperature of Titan is about 70° K, or -200° C, on the average. This is too cold for the development or existence of life.

The past decade of exploration of the solar system has not led to the discovery of extraterrestrial life. Instead, we can well conclude that our galaxy probably contains billions of stars, planets and satellites that are completely lifeless, and that the same may be true of countless other galaxies. We speculate that among the vast array of stellar heavenly bodies, there also are a few Earth-like planets that are capable of supporting life.

One aspect of exobiology that has received very little detailed consideration is the question of inorganic trace elements in living systems that might be present in extraterrestrial locations. Life on the Earth, especially in its larger, multicellular forms, is dependent upon the availability of an array of mineral substances, most of them in very small amounts. Excessive quantities of many of the inorganic elements are toxic to large animals, and especially to plants.

Many years ago, biochemists concluded that terrestrial animals had evolved from forms of life that lived in the ocean. This conclusion was based on the mineral composition of blood being something like that of dilute sea water. Sodium, potassium, magnesium, calcium and sulfate are present in both fluids. There is a long list of trace elements that are present in low concentrations in sea water (Table 3). Some of these are toxic at high levels. The simpler living organisms such as molds, have evolved so as to become quite expert at living in high concentrations of toxic elements. However, large animals are less adaptive to such situations, and I wonder if a certain balance of minerals in sea water was essential for the evolution that led eventually to intelligent life on the Earth, and, if so, might this be a deciding factor in the type of extraterrestrial life? One of our most interesting links to a marine origin is our need for iodine in the functioning of the thyroid gland. Human beings and mammals in regions remote from the sea often suffer from iodine deficiency, which leads to congenital deformities and to the young of certain mammals being born dead.

Table 3 Approximate Mineral Composition of Sea Water  
(Partial List)

Element	Concentration (Parts per thousand)	Element	Concentration (Parts per million)	Element	Concentration (Parts per billion)
Chlorine	19	Strontium	13	Arsenic	3 to 24
Sodium	11	Boron	5	Iron	2 to 20
Magnesium	1.3	Silicon	0.2 to 4	Zinc	5 to 14
Sulfur	0.9	Aluminum	0.2 to 2	Copper	1 to 90
Calcium	0.4	Fluorine	1.4	Manganese	1 to 10
Potassium	0.4	Rubidium	0.2	Lead	4 to 5
Bromine	0.06	Lithium	0.1	Selenium	4
Carbon (Inorganic)	0.03	Barium	0.05	Nickel	0.1 to 0.5
		Iodine	0.05	Mercury	0.3

A few years ago, Crick and Orgel made a fanciful but erroneous suggestion (1973) that the need for molybdenum by terrestrial organisms indicated that life had been brought here from a molybdenum star. Their argument was that nickel is much more abundant in the Earth's crust than is molybdenum, and that molybdenum is needed by terrestrial organisms while nickel is not. However, they overlooked the fact that molybdenum is present in sea water at higher concentrations than is nickel (Jukes, 1974). Therefore, the need for molybdenum is an index of the origin of life in the oceans, rather than for its being transported here from a molybdenum star.

With respect to exobiology, my general conclusion is that there is a fairly wide range of concentrations of inorganic elements in sea water that is compatible with the existence of life. Living organisms have produced, by evolution, some remarkable mechanisms for excluding toxic elements, and for selectively concentrating necessary elements in their protoplasm.

Whether life could originate in the waters of an ocean that contained an inhospitable balance of inorganic ions is another question. We call to mind, for example, the existence of sulfuric acid in the clouds of Venus as being suggestive of hell.

The following inorganic elements are needed in the diet of animals (Underwood, 1980): calcium, phosphorus, magnesium, potassium, sodium, chloride, iron, zinc, iodine, copper, manganese, fluoride, chromium, selenium, molybdenum, silicon, vanadium, nickel, tin and arsenic; the last five at very low levels. Boron is needed by plants.

#### Panspermia

The idea was once proposed that life on the Earth could have originated from spores that travelled here through the interstellar medium. Recent knowledge of space

travel has made this unbelievable.

"The combination of high-energy ultraviolet light, X-rays and cosmic rays would probably prove lethal over the length of the trip, a million years or more for random wandering from one planetary system to the next." (Goldsmith and Owen, 1980).

The possible length of time of survival of bacterial spores in space has been investigated experimentally by various procedures. The results were summarized by Horneck (1981). The high vacuum of space inactivates vegetative cells by dehydration and consequent damage to cell membranes. However, bacterial spores of the organism *Bacillus subtilis* survived extended exposure up to 20 days in a high vacuum, but they became more sensitive to solar ultraviolet radiation.

The destructive effect of cosmic heavy ions further decreases the chance that spores can survive for any length of time in space.

The possibility remains that spores might be carried through space on meteorites or dust particles that shelter them from ultraviolet radiation. The small size of the spores would decrease the chance that they would encounter cosmic ion particles. In this way, they might survive long enough to traverse interplanetary distances, according to Horneck, but the possibility of intergalactic survival of spores appears to be vanishingly small. No living organisms have been detected in meteorites.

#### Interstellar molecules

During 1979-1980, there was much interest in the diversity of organic molecules within the interstellar medium. Some of these molecules are identical with those that have been used in laboratory experiments for synthesis of compounds of biological significance. Interstellar molecules of this type include the 12 shown in Table 4 and many others. Zuckerman (1977) lists 39 interstellar molecules composed of carbon, hydrogen, oxygen, nitrogen and sulfur, varying in size from two to nine atoms. A much larger list has been compiled by Mann and Williams (1980).

Table 4      Interstellar Molecules of Biological Interest--  
A Few Examples

---

Hydrogen	Formaldehyde	Acetaldehyde
Ammonia	Formic acid	Cyanoethylene
Methane	Cyanoacetylene	Methylamine
Hydrogen cyanide	Methyl alcohol	Carbon monoxide

---

Laboratory experiments in which amino acids and nucleic acid bases have been produced from small molecules are well known and well established. Miller and Orgel (1974) used mixtures of methane, ammonia, water and hydrogen, subjected to prolonged electric discharge, and found that several amino acids were produced, especially glycine and alanine. Hydrogen cyanide can give rise to adenine, which is a component of DNA. Adenine, guanine and cytosine have also been produced by heating a mixture of carbon monoxide, hydrogen and ammonia in the presence of metal catalysts.

Cyanoacetylene is produced when an electrical discharge is passed through a

mixture of nitrogen and methane. Cyanoacetylene can give rise to aspartic acid and cytosine. Formaldehyde can be polymerized to form sugars.

These small interstellar molecules have no resemblance to living systems, but their presence in outer space reveals the universal nature of compounds of carbon, and suggests that the conditions necessary for the formation of life can exist in various locations throughout the universe.

Such speculations are reinforced by the presence of organic compounds in carbonaceous chondrites, meteorites from outer space, discussed below.

#### Planetary Biology Includes Terrestrial Biology

In 1980 and 1981, the Voyager mission stimulated interest in the chemistry of the outer planets and their satellites. Chemical activities in the atmosphere of Jupiter may provide a model for prebiological chemistry that existed billions of years ago on the Earth.

The term "exobiology" implies the study of life in other worlds beyond the Earth. However, in recent years, the field is turning to studies of life on the Earth itself, particularly with regard to its origin and early evolution. By this means, scientists hope to obtain clues as to the possible existence and nature of life elsewhere.

An interesting development has been the use of the "molecular evolutionary clock." This concept says that nucleotide substitutions in DNA proceed at a certain rate measured in terms of changes per million years (Jukes, 1980). These substitutions occur when one nucleotide is replaced by another in a DNA molecule. Such changes are called "point mutations," and these occur continually and at random. Most point mutations do not persist, either because they are eliminated by an error-correcting mechanism, or because they are deleterious, and hence the organisms that undergo these changes do not survive. However, there are always a few changes that enter the make-up of the species, and are carried along during evolution, either because they are practically harmless, or because, quite rarely, they produce beneficial changes. The process by which mutations in DNA become fixed was first detected in certain proteins such as hemoglobin that are widely distributed in many different species. Mutations in DNA may show up as amino acid replacements in proteins. It is well known that hemoglobins of humans and monkeys are quite similar, but both differ markedly from those of other mammals, such as horses and mice, and the difference gradually widens as we compare our own hemoglobins with those of kangaroos, chickens, frogs, bony fishes and sharks. There is a gene, or region of DNA, that changes much more slowly than the gene that specifies hemoglobin. This is the gene for a large ribonucleic acid (RNA) molecule that occurs in all ribosomes, and is called 16-S ribosomal RNA. It changes so slowly because it is extremely difficult for ribosomes to accept changes.

Nevertheless, the steady pressure of mutations on DNA slowly forces these changes to be accepted. Studies with 16-S ribosomal RNA have recently given us a new idea of the evolutionary separation of life forms on the Earth. This was described by Woese (1981).

All living organisms contain ribosomes, which are particles within the cell on which protein synthesis takes place. They contain RNA and proteins. Woese and his colleagues have compared the ribosomal RNA sequences of many species, including numerous bacteria. The studies show, mathematically, that all living organisms have descended from a common ancestor by divergent evolution. This is a complete substantiation of the basic principle of Darwinism. The results have

supplied a new picture of how bacteria evolved. The studies also show that micro-organisms evolved for about 3 billion years before higher organisms appeared. These higher organisms, including animals, plants and fungi, did not come on the scene until about 600 million years ago. This represents only 20% of the time that micro-organisms have been living on the Earth.

The same research has also enabled the biological "family tree" to be divided into three branches, as compared with the former idea that there were only two main types of life, termed prokaryotes and eukaryotes. The three branches are the true bacteria, the archaeabacteria, and the eukaryotes, the last term referring to organisms whose cells contain a separate nucleus. Two major events have taken place during the evolution of eukaryotes. The first was when certain bacteria entered the cells of eukaryotes and stayed there. The descendants of these bacteria are the mitochondria. These are small packages within all eukaryotic cells, and they carry their own DNA and their own protein coding system. The mitochondria are often termed the "powerhouses" of the cell, where energy is provided for biochemical processes.

The second of these big events took place later. It was another bacterial invasion, and it occurred only in plants. The intruders were cyanobacteria, also known as blue-green algae. They contain chlorophyll, and the descendants of these invading cyanobacteria are called chloroplasts, and, like mitochondria, they carry their own DNA. The chloroplasts are responsible for the maintenance of the oxygen content of our atmosphere, and for the synthesis of food by green plants upon which animals depend. This is familiar, elementary biology.

The invading cyanobacteria brought their DNA with them. Part of the DNA stayed in the chloroplasts, and another part went into the nuclei of the plant cells. The second part has probably enabled plants to produce vitamins and amino acids that animals cannot manufacture (Weeden, 1981). But animals can eat plants. The rise of animal life, as well as plant life, was made possible by an intrusion of ancestral plant cells by blue-green algae.

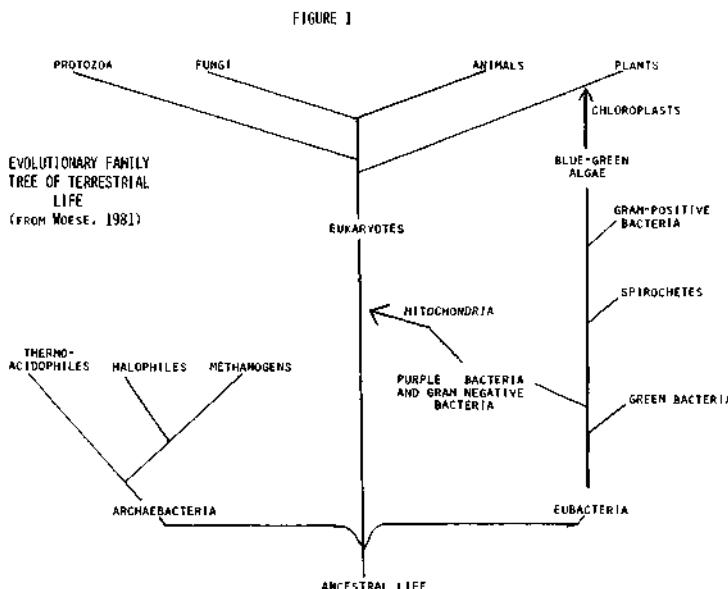
The new family tree of life has encouraged study of the primitive archaeabacteria. Some of these are so sensitive to oxygen that they die as soon as they are exposed to it. However, they live successfully in deep mud and in the stomachs of ruminating animals. Others of them can flourish under very unusual conditions, such as in acid hot springs, or in concentrated brine.

No one knew how to classify micro-organisms except on the basis of visible characteristics until these new findings appeared based on the molecular evolutionary clock. We now have a better idea of the stages through which life must proceed before complexity can be reached. A summary of our family tree is shown in Figure 1, adapted from a diagram published by Woese (1981).

#### Life in Extreme Environments

Sterilization procedures were used for spacecrafts at the Kennedy Space Center in preparation for the trip to Mars, so that Mars would not be contaminated. During these procedures, spores of soil bacteria were detected having an extraordinary resistance to dry heat. The length of time to kill 90% of a sample of these spores at 125° C was 5.8 days. However, the bacteria formed by germination of the spores were killed easily by heat.

The extreme resistance of the spores led to studies of their morphology by scanning electron microscopy. In cross section, the spores had a many-layered coat, each layer with a characteristic fine structure. The spore walls were extraordinarily thick, 87% of the volume of the spore was taken up by various coat structures.



However, no actual explanation of the dry heat resistance of the spores was discovered (Youvan et al., 1977).

Fungi of the common *penicillium* and *aspergillus* groups can be grown in a variety of brine or on moist salt crystals, especially on media rich in potassium chloride, even at temperatures as low as -10° C (BZ Siegel and SM Siegel, 1980).

Certain bacteria can live at temperatures near the boiling point of water. One of these is *Thermus thermophilus*. This contains enzymes that are stable to heat, as well as unstable enzymes. Another species, *Bacillus caldolyticus*, grows at temperatures up to 80° C and produces labile enzymes only at lower temperatures, below 56° C.

In the dry valley desert environments of South Victoria Land, Antarctica, cold and desiccation restrict living organisms severely. Nevertheless, algae and fungi live within the air spores of sedimentary rocks, receiving light through translucent surfaces, warmth from the sun, and a melt water supply. These organisms live as lichens, below the surface of the rocks. They appear to have obtained a niche in which they are protected from the environmental extremes of the general landscape (Siegel, 1980).

Many of these observations were undertaken in order to evaluate the chances of life being found on Mars, under conditions of extreme drought, high ultraviolet flux, and low temperatures. The results showed that micro-organisms could live under some of the most hostile conditions found naturally occurring on the Earth.

#### Organic Compounds in Meteorites and Comets

Studies of the Murchison and Murray meteorites, both of which are carbonaceous condrites, showed the presence of organic compounds (Lawless, 1980). The Murray

meteorite fell in the United States in 1950, and the Murchison meteorite in Australia in 1969. A list of the amino acids found in them is shown in Table 5. More than half of them are not present in proteins. Furthermore, those amino acids with an optically active center were found to contain nearly equal concentrations of the D and L forms, as one would expect if they were not of biological origin.

Table 5 Amino Acids Detected in Carbonaceous Chondrites

Glycine	$\beta$ -Amino- <u>n</u> -butyric acid	Norvaline
Alanine	$\gamma$ -Aminobutyric acid	$\gamma$ -Aminovaleric acid
$\beta$ -Alanine	N-Methylalanine	Proline
N-Methylglycine	N-Ethylglycine	Glutamic acid
$\alpha$ -Aminoisobutyric acid	Aspartic acid	Pipecolic acid
$\alpha$ -Amino- <u>n</u> -butyric acid	Valine	Leucine
$\beta$ -Aminoisobutyric acid	Isovaline	Isoleucine

In a laboratory experiment carried out subsequently, Miller and his colleagues (Ring and co-workers, 1972) showed that all of the primary alpha amino acids found in the Murchison meteorite could be synthesized by the action of electric discharge on a mixture of methane, nitrogen and water, with traces of ammonia.

These results provide a biological foundation for the currently popular theory of the origin of life from simple molecules. Furthermore, it is possible, as Lawless says, "that meteoritic bombardment of the primordial earth surface could have provided large amounts of organic material for life's origin." (Lawless, 1980). Ghosh and Ghosh have suggested (1980) that amino acids and the bases of DNA, such as adenine and guanine, may be formed from smaller molecules in dense interstellar clouds.

#### Comets

Organic compounds have been detected in comets. Oró and co-workers (1980) estimate that during its early history, the Earth captured a mass of cometary material of about  $10^{20}$  kilos. From this, they deduce that about  $10^{19}$  grams of carbon were added by comets to the surface of the prebiotic Earth. This corresponds to about 14% of the total carbon on the surface of the Earth. The capture of comets by the Earth would also have contributed to generating the appropriate aqueous and reducing environmental conditions necessary for generating organic synthesis. The heat and shock waves produced by the impact would probably have decomposed most of the carbon compounds in the comets, but the small molecules produced by the decomposition could have recombined and formed a large variety of organic molecules. Oró and co-workers, as a result of their laboratory experiments, conclude that some of the compounds synthesized in this manner include amino acids, sugars, purines and pyrimidines, which are essential to all living systems.

#### Clays as Catalysts

Nothing is known of the steps that took place to condense the free amino acids into protein molecules. Various proposals have been made. One is that clays or

other minerals provided a surface on which small molecules could be concentrated and subsequently polymerized. Several investigators have carried out experiments to test this hypothesis.

Most clays do not readily absorb amino acids and nucleotides at the pH that is believed to have existed in the primitive ocean, around 7.5. However, Lawless and Edelson (1980) have found that zinc greatly increases the interaction of nucleotides with clays. The authors suggest that divalent metallic ions may have had an important role in the origin of life because of this effect.

#### DISCUSSION

The creationists have invested themselves with the authority to be the voice of the Almighty. Their leader, Dr. Henry Morris, a hydraulic engineer who is head of the Institute for Creation Research in San Diego, has announced that intelligent life cannot exist in other worlds, because this is not described in the Bible. He also has decided and proclaimed that the asteroid belt and the rings of Saturn "reflect some kind of a heavenly catastrophe associated either with Satan's primeval rebellion or his continuing battle against Michael and his angels." Other creationist publications state that Noah took dinosaurs on the ark, and that dragons used to exhale fire, resulting from reaction in a mixture of hydrogen peroxide and hydroquinone. The creationists are not satisfied to believe such fantasies themselves. It is now compulsory by law to teach children such nonsense in the states of Arkansas and Louisiana, whenever the science of evolution is taught. The travesties of the creationists are termed "creation-science," and they claim that there are thousands of scientists who believe that the universe is not more than 10,000 years old, and that the great flood took place as described in the Book of Genesis as a punishment for the sin committed by Adam and Eve, so that all human beings were drowned except for Noah's family of eight. Schoolchildren who do not believe these proposals are threatened with hell fire. The creationist program is being taught in many tax-supported public elementary schools throughout the United States.

It is common among writers of science fiction to portray the inhabitants of other worlds as infinitely wise and benevolent, and in some cases, immortal. These ideas may arise from wishful thinking; a hope that other intelligent beings might not suffer from our frailties.

However, unless intelligent life elsewhere were the product of special creation, I believe we must assume that, if it exists, it has been the result of an evolutionary process. This process on the Earth has taken about 3.5 or 4 billion years, from the first appearance of life to the development of civilization. Perhaps, in another location, the procedure could have been more rapid, but evolutionists are virtually unanimous in regarding the emergence of complexity in living organisms as being a gradual process.

If we believe that extraterrestrial life is based on DNA, or a DNA-like molecule, we should consider some of the biochemical properties of DNA.

Every inherited characteristic of every terrestrial organism must pass through "the eye of the needle" before it appears in the next generation. The thread that goes through the eye of the needle is the filamentous molecule called DNA. This contains the information for each member of each species. The only way the information is carried is in a linear sequence of four variables, the nucleotides of DNA, termed for short, "A, C, G and T." There is much argument about whether human behavior is influenced by DNA, however, it is quite certain that the behavior of many organisms is inherited. This is particularly striking in the case of the complex instinctive behavior of insects. These creatures carry out extraordinary tasks without any parental instruction.

We have probably learned more about DNA in the last 5 years than in all years preceding because of new methods for sequencing it. The new knowledge is bound to influence exobiology.

On the Earth, life depends on the continued existence of DNA, without interruption. If DNA did not change, the Earth would still be inhabited by the first primitive organisms that developed to the point of containing DNA.

Instead, evolution has taken place as a result of changes in DNA. Some of the changes consist in replacing the individual units, the nucleotides, by each other. The other type of change is when the total amount of DNA in an organism increases by a process of duplication. Also, pieces of DNA can be shuffled around within the chromosomes, and short pieces can be added or deleted.

As far as exobiology is concerned, the only example we have of life is on our own planet. Evolution is a process in which organisms survive and persist because they find an environment that suits their needs. Our guess is that the same sort of process takes place elsewhere in the universe on Earth-like planets. To produce intelligent life, we presume there must be evolution, and if there is evolution, there must be some kind of a mechanism of heredity in which changes can take place.

It is tempting to conclude that the terrestrial DNA-protein type of life might exist elsewhere, because there is increasing evidence that the small molecules of organic compounds that are found on Earth are also present in outer space.

Evolution works towards preserving the existence of DNA molecules. Can we, as intelligent beings, shape the course of evolution according to our decisions and desires? Have the inhabitants of other worlds succeeded in so doing?

#### SUMMARY

Exobiology is intended to concern itself with extraterrestrial life, which is not known to exist. Investigations in this field have therefore included the following topics: seeking for life beyond the Earth in the solar system; interstellar molecules of biological significance; carbon compounds in meteorites; origin and early evolution of life; characteristics of primitive organisms; life in environmental extremes. Exobiology has had a stimulating effect on the search for extraterrestrial intelligence.

#### ACKNOWLEDGEMENT

I wish to thank Carol Fegte for cooperation and preparation of the manuscript. Support was received from NASA Grant NGR 05-003-460.

#### REFERENCES

- Crick, F. H. C., and L. E. Orgel (1973) Directed panspermia. Icarus, 19, 341-346.  
Dickerson, R. E. (1978). Chemical evolution and the origin of life. Scient. Am., 238, 70-86.  
Ghosh, K. K., and S. N. Ghosh (1980). A mechanism for the formation of glycine, adenine and guanine in interstellar space. In R. Holmquist (Ed.), Life Sciences and Space Research, Vol. 18, Pergamon Press, Oxford. pp. 37-44.  
Goldsmith, G., and T. Owen (1980). The Search for Life in the Universe. Benjamin Cummings Publ., Menlo Park, Ca. p. 172.  
Horneck, G. (1981). Survival of microorganisms in space: A review. In R. Holmquist (Ed.), Life Sciences and Space Research, Vol. 19, Pergamon Press, Oxford. pp. 39-48.  
Horowitz, N. H. (1977). The search for life on Mars. Scient. Am., 237, 52-61.  
Jukes, T. H. (1974). Sea-water and the origins of life. Icarus, 21, 516-517.

- Jukes, T. H. (1980). Silent nucleotide substitutions and the molecular evolutionary clock. Science, 210, 972-978.
- Lawless, J. G. (1980). Organic compounds in meteorites. In R. Holmquist (Ed.), Life Sciences and Space Research, Vol. 18, Pergamon Press, Oxford. pp. 19-28.
- Lawless, J. G., and E. H. Edelson (1980). The possible role of metal ions and clays in prebiotic chemistry. In R. Holmquist (Ed.), Life Sciences and Space Research, Vol. 18, Pergamon Press, Oxford. pp. 82-89.
- Mann, A. P. C., and D. A. Williams (1980). A list of interstellar molecules. Nature, 283, 721-725.
- Mazur, P., E. S. Barghoorn, H. O. Halvorson, T. H. Jukes, I. R. Kaplan, and L. Margulis (1978). Biological implications of the Viking mission to Mars. Space Science Reviews, 22, 3-34.
- Miller, S. L., and L. E. Orgel (1974). The Origins of Life on the Earth. Prentice-Hall, Englewood Cliffs, New Jersey. p. 85.
- NASA (1980). Voyager 1 Encounters Saturn. U. S. Government Printing Office Publ. #033-000-00817-1. U. S. Government Printing Office, Washington, D. C.
- Oró, J., G. Holzer, and A. Lazcano-Araujo (1980). The contribution of cometary volatiles to the primitive Earth. In R. Holmquist (Ed.), Life Sciences and Space Research, Vol. 18, Pergamon Press, Oxford. pp. 67-82.
- Ring, D., Y. Wolman, N. Friedmann, and S. L. Miller (1972). Prebiotic synthesis of hydrophobic and protein amino acids. Proc. Natl. Acad. Sci. (U. S. A.), 69, 765-768.
- Siegel, B. Z., and S. M. Siegel (1980). Further studies on the environmental capabilities of fungi: Interactions of salinity, ultraviolet irradiation, and temperature in Penicillium. In R. Holmquist (Ed.), Life Sciences and Space Research, Vol. 18, Pergamon Press, Oxford. pp. 59-64.
- Siegel, S. M. (1980). COSPAR Report to the United Nations.
- Underwood, E. J. (1979). New findings with trace elements: Including effects of processing on supplies and availability. In A. Neuberger and T. H. Jukes (Eds.), Biochemistry of Nutrition 1: International Review of Biochemistry, Vol. 27, University Park Press, Baltimore. pp. 207-243.
- Weeden, N. R. (1981). Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. J. Mol. Evol., 17, 133-139.
- Woese, C. R. (1981). Archaeabacteria. Scient. Am., 244, 98-122.
- Youvan, D., M. Watanabe, and R. Holmquist (1977). Morphology of the extremely heat-resistant spores from Bacillus sp. ATCC 27380 by scanning and electron microscopy. In R. Holmquist (Ed.), Life Sciences and Space Research, Vol. 18, Pergamon Press, Oxford. pp. 65-72.
- Zuckerman, B. (1977). Interstellar molecules. Nature, 268, 491-495.

## THE BEGINNINGS OF MAGNETOSPHERIC PHYSICS

J. A. Van Allen

*University of Iowa, Iowa City, Iowa 52242, USA*

### ABSTRACT

A brief account is given of the discovery of the radiation belts of the earth in early 1958, the prompt confirmations of this discovery, and the beginnings of the now flourishing science of magnetospheric physics.

### KEYWORDS

Radiation belts; magnetosphere; history of the magnetosphere; earth's magnetic field; Explorer I.

### INTRODUCTION

As of 1981, magnetospheric physics is a massive and flourishing science which engages the efforts of over a thousand investigators in at least twenty different countries. The current rate of publication in this field is of the order of two or three original research papers per day.

The scientific heritage of the subject lies principally in the areas of geomagnetism, aurorae, geomagnetic storms, and the geophysical aspects of cosmic rays and solar corpuscular streams. This heritage is well represented by the treatises of Chapman and Bartels (1940), Jánossy (1948), Montgomery (1949), Mitra (1952), Alfvén (1950), and Störmer (1955).

The general magnetic field of the earth is essential to the existence of its magnetosphere. Indeed the magnetosphere of the earth or of any other celestial body may be defined as that region surrounding the body within which its magnetic field, however distorted by external currents, controls the motion of electrically charged particles. The magnetosphere of the earth encompasses a huge population of such particles -- electrons, protons, and other ions -- whose source and gross dynamics are traceable principally, but not wholly, to the solar wind, a magnetized, ionized gas emitted by the sun and flowing outward through the solar system.

On the one hand, magnetospheric physics is closely related to laboratory plasma physics; on the other hand, it is presumably applicable to natural phenomena in planetary and astrophysical systems throughout the universe. The theory of ionized

gases in magnetic and electric fields provides the unifying principles of the subject over a vast range of physical scale. The present sophistication of magnetospheric physics stands in stark contrast to its primitive beginnings. The latter are sketched below.

#### PRECURSORY WORK

During the period 1946-1957, a large number of investigations were conducted in the upper atmosphere of the earth with scientific instruments on rockets flown more or less vertically to altitudes of up to 390 kilometers by research workers in the United States, the Soviet Union, England, Australia, France, and Japan (Newell, 1953). Some of these flights were devoted to study of the intensity and nature of the cosmic radiation above the appreciable atmosphere. Such measurements were extended over the latitude range  $77^{\circ}$  N to  $71^{\circ}$  S by the inexpensive balloon-launched rocket technique during the latter part of this period (Van Allen and Gottlieb, 1954; Van Allen, 1959a). This series of flights from ships at sea also resulted in the first direct observations of the primary auroral radiation (Meredith, Gottlieb, and Van Allen, 1955; Van Allen, 1957). Experience in conducting scientific measurements with rocket-borne instruments and the rapid development of high performance rockets following World War II made it realistic to plan scientific work with earth orbiting satellites as part of the 1957-58 International Geophysical Year. The first flight of an artificial satellite of the earth, Sputnik I, was achieved by the Soviet Union on 4 October 1957. The radio signals from Sputnik I at 20.005 and 40.002 MHz were received by stations throughout the world. The distinctive -- beep-beep-beep -- modulation of these signals heralded the beginning of a new scientific epoch. Sputnik II, launched on 3 November 1957, carried the first detectors of energetic charged particles to be flown on a satellite. Two shielded ( $10 \text{ g cm}^{-2}$ ) Geiger tubes (gas-discharge counters) (10 cm length and 1.8 cm diameter) in this payload were intended to provide a geographically comprehensive survey of the intensity of the cosmic ray intensity above the atmosphere. The investigators (Vernov and colleagues, 1958) obtained observations over the Soviet Union for a period of about seven days in the altitude range 225 to 700 km, the latitude range  $40^{\circ}$  to  $65^{\circ}$  N and the longitude range  $25^{\circ}$  to  $143^{\circ}$  E. Reliable data were obtained on the latitude and altitude dependence of cosmic ray intensity. On a particular pass on 7 November a brief series of "bursts" of substantially greater ( $\sim 50\%$ ) than average intensity was observed coherently by the two detectors at latitudes greater than  $58^{\circ}$  N (Fig. 1). This effect was termed a "cosmic-ray burst" but no physical interpretation was suggested. Otherwise the Sputnik II data exhibited a smooth dependence on latitude and longitude at about the level of intensity expected on the basis of previous rocket measurements, at lower altitudes.

#### DISCOVERY OF THE INNER RADIATION BELT OF THE EARTH

In anticipation of the opportunity for extending our sparse geographical survey, with rockets, of the cosmic ray intensity above the atmosphere, I had proposed the flight of a simple radiation detector on an early U. S. satellite (Van Allen, 1956). The instrumentation prepared for this purpose (Ludwig, 1959a) at the University of Iowa had a single, lightly-shielded ( $1.5 \text{ g cm}^{-2}$ ) Geiger tube (10.2 cm length and 2.0 cm diameter) as the basic detector. One such instrument was a part of the payload (Fig. 2) of the first U. S. satellite, Explorer I, launched on 1 February 1958 (UT) (31 January, EST from Cape Canaveral, Florida) into an orbit inclined at  $33^{\circ}$  to the equator with perigee and apogee altitudes of 360 and 2550 km, respectively. A similar instrument was flown in Explorer III, launched on 26 March 1958 into an orbit of the same inclination with perigee and apogee altitudes of 188 and 2800 km, respectively. The Explorer III instrument included

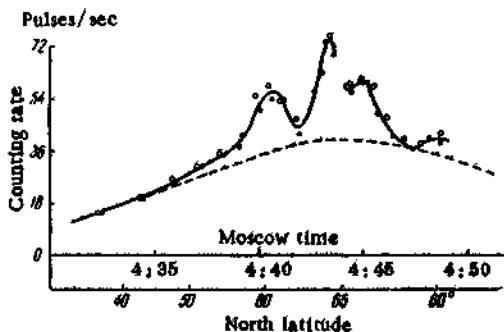


Fig. 1. Time dependence of the counting rate of a Geiger tube on Sputnik II. The dashed curve shows the usual case and the points and solid curve show the occurrence of an unusual "cosmic ray burst" on 7 November 1957 (Vernov and colleagues, 1958).

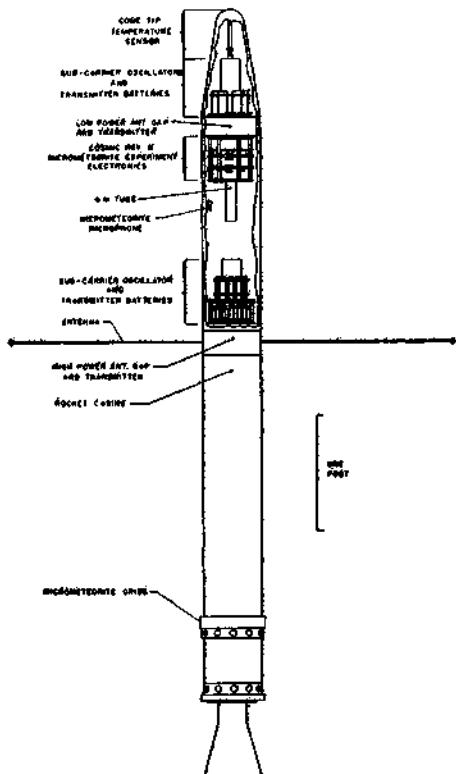


Fig. 2. Outline drawing of the payload of Explorer I (1958 Alpha), as permanently attached to the fourth stage rocket of the propulsion system (Ludwig, 1959a).

the important addition of a miniature magnetic tape recorder which stored the counting rate data for 135 minutes (more than a complete orbital period) and then, on radio command, played back the stored data over the telemetry transmitter in six seconds while within the range of a single receiving station (Ludwig, 1959b).

Within the first few weeks of Explorer I's orbital flight, we had only a rather sparse set of data consisting of segments of the order of one minute duration from many different and somewhat uncertain positions in latitude, longitude, and altitude. During some segments of data, the counting rate of our Geiger tube was of the order of 20 to 100 counts per second, generally within the range that we had expected for cosmic radiation, and the instrument appeared to be operating reliably. In other segments of data there were no counts observed for as long as two minutes. On one noteworthy pass the rate underwent a transition from zero to a reasonable value within about twenty seconds. There was no conceivable way in which the cosmic ray intensity could drop to zero at high altitudes. On the other hand, we had a high level of confidence in the Geiger tube and the associated tuning fork timer and electronic circuitry by virtue of conservative design and the rigorous thermal and mechanical conditions that it had survived in the pre-flight testing program. The puzzle hung over our heads as we tried to find if the strange effect had any systematic dependence on passing through earth shadow, on payload temperature or on altitude, latitude or longitude. Noise-free data accumulated slowly.

Throughout the few weeks following the launch of Explorer I we were heavily occupied in preparing Explorer II (lost in the launch failure of 5 March) and Explorer III; in developing data reduction and analysis techniques, to which we had given relatively little attention before flight; in formulating plans for subsequent flights; and in coping with a steady flow of telephone calls on practical arrangements and on inquiries on our progress. Our original plan was to accumulate a comprehensive body of data on the distribution of cosmic-ray intensity around the earth on a leisurely basis. But the widespread interest in Explorer I had produced an urgent demand for an immediate report of observational results.

Soon after the launching of Explorer III, I flew to Washington, D. C. to confer with Joseph Siry, John Mengel, and others at the Naval Research Laboratory and to pick up preliminary orbital data for Explorer III. Contrary to the implication of some popular accounts, the Vanguard group fully supported the Explorer program in many vital ways. The first successful launch of a Vanguard satellite had occurred on 17 March and the NRL team was operating on an around-the-clock basis, handling now the tracking and data acquisition for three satellites. From NRL I returned to the Vanguard data center on Pennsylvania Avenue and picked up the complete record of a successful playback of data from our Explorer III tape recorder. The playback had been received at the San Diego minitrack station on 28 March. I put the record in my briefcase and returned to my room in the Dupont Plaza Hotel. There with the aid of graph paper and a ruler from a nearby Peoples Drug Store and my slide rule, I worked out the counting rates of our Geiger tube as a function of time for a full 102-minute period and plotted them (Fig. 3). The data provided a beautiful confirmation and extension of the fragmentary data from Explorer I. The counting rate at low altitudes was in the expected range of 15 to 20 counts per second. There was then a very rapid increase to a rate exceeding 128 counts per second (the maximum recordable rate of our on-board storage system). A few minutes later the rate decreased rapidly to zero. Then after about fifteen minutes it rose rapidly from zero to greater than 128 counts per second again and remained high for forty-five minutes, then again decreased rapidly to 18 counts per second as the orbit around the earth was nearly completed. At 3:00 a.m. I packed my work sheets and graph and turned in for the night with the conviction that our instruments on both Explorers I and III were working reliably and giving reproducible results but that we were encountering a mysterious physical effect of a real nature.

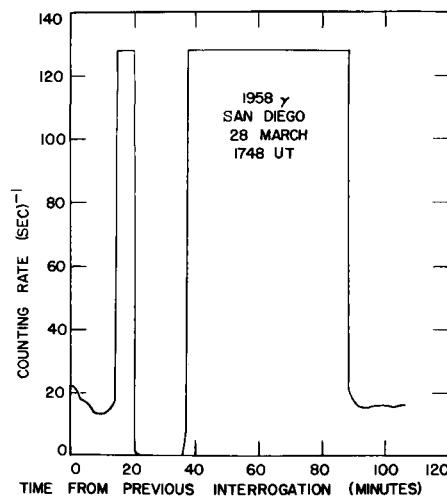


Fig. 3. Apparent counting rate of the Geiger tube on Explorer III (1958 Gamma) as derived from a tape recording for one orbit around the earth. Note that 128 count sec<sup>-1</sup> is the saturation level of the recording system and that a zero apparent counting rate corresponds to a true counting rate greater than 25,000 count sec<sup>-1</sup> (Van Allen, 1 May 1958 and 1961).

Early the following day I flew back to Iowa City and proudly displayed my graph and my conviction to my colleagues, Ernest Ray and Carl McIlwain. It had meanwhile occurred to McIlwain that a sufficiently great intensity of radiation would overload our detection system and drive the apparent counting rate to zero, a fact that we all knew but had temporarily ignored because of our predisposition to reject the possibility of such great intensities and our apprehension about possible malfunctioning of the instrument. During the previous day, Carl had made tests with our prototype Geiger tube system using a small x-ray machine and demonstrated that a true rate exceeding about 25,000 counts per second would indeed result in an apparent telemetered rate of zero. The conclusion was immediate: At altitudes of the order of 1200 km and greater the radiation intensity was at least a thousand times as great as that of the cosmic radiation. Ernie's famous, though consciously inaccurate remark, summarized the situation: "My God, space is radioactive"! Our realization that there was a very great intensity of radiation at high altitudes rationalized our entire body of data.

George Ludwig returned from the Jet Propulsion Laboratory to the University of Iowa on the 11th of April and the four of us continued working feverishly in analyzing the data from Explorers I and III (by primitive hand methods) and organizing them by altitude, latitude, and longitude. The systematic, repetitive run of the data on successive orbits confirmed my earlier conclusion that both instruments were indeed operating reliably.

During mid-April we prepared graphs and a short written statement of our raw findings (Van Allen and colleagues, 1958) and I mulled over their meaning. I entertained two quite different lines of thought: (a) that we might be detecting x rays or gamma rays, possibly from the sun or (b) that the radiation at high

altitudes might be akin to the auroral soft radiation which we had studied during the preceding several years with rockoon flights at high latitudes and most recently by McIlwain's rocket flights from Fort Churchill. We had identified the auroral radiation as being principally electrons having energies of the order of 10's of keV. I quickly rejected hypothesis (a) on the conclusive grounds that the high intensity radiation was present during both daylight and dark conditions, that it exhibited a latitude dependence, and that the extremely rapid increase in intensity with increasing altitude was impossible for any type of electromagnetic radiation. Specifically, the rate of increase of intensity was a factor of ten within an altitude range of less than 100 kilometers at an altitude of the order of 1000 km; the decrease in atmospheric "thickness" within that increment of altitude was totally negligible compared to the some  $1.5 \text{ g cm}^{-2}$  of material in the nose cone and wall of the detector. Such an increase could not be due to decreased absorption of x rays or gamma rays of any wavelength. Hence, I concluded that the effect must be attributed to electrically charged particles, constrained by the earth's external magnetic field from reaching lower altitudes. Such particles could not be arriving directly from infinity because of the concentration of the effect near the geomagnetic equator, in complete defiance of Störmer's theory of the entrance of electrically charged particles from a distant source into the earth's magnetic field. By virtue of my familiarity (a) with an early paper of Störmer (1907) on the bound motion of an electrically charged particle in a dipolar magnetic field, (b) with magnetic field confinement of charged particles in laboratory machines during my 1953-54 work building and operating an early version of a stellarator at Princeton University, and (c) with an early 1958 suggestion of N. C. Christofilos (see later), I further concluded that the causative particles were present in trapped orbits in the geomagnetic field, moving in spiral paths back and forth between the northern and southern hemispheres and drifting slowly around the earth. The intensity of trapped particles whose mirror points were at low altitudes would be diminished by atmospheric absorption. I called this population of particles the geomagnetically trapped corpuscular radiation. Later it was termed the "radiation belt" of the earth because of the toroidal form of the populated region, encircling the earth like a belt.

The foregoing account of observations and interpretation is essentially the one that I gave in a special joint session of the American Physical Society and the National Academy of Sciences in the latter's auditorium on 1 May 1958. A large press conference was held following the lectures by me and other participants in the Explorer I and Explorer III program and the results were reported extensively by news media on the following day. Fortunately and unknown to me until more than a year later, a Voice of America tape recording was made of my lecture and of the ensuing question and answer period. A written transcription of this tape together with copies of my illustrations provided a documented, published record of this lecture, complete with grammatical errors and colloquial language (Van Allen, 1961).

I had adopted at that time the working hypothesis that the trapped radiation consisted of "... electrons and likely protons, energies of the order of 100 keV and down, mean energies probably about 30 keV ..." In this vein of thought, the response of our Geiger tube would be attributed to bremsstrahlung produced as the electrons bombarded the nose cone of the instrument. If this bremsstrahlung interpretation were correct, I estimated that an omnidirectional intensity of  $10^8$  to  $10^9 (\text{cm}^2 \text{ sec})^{-1}$  of 40 keV electrons would be required to account for the counting rates at altitudes of  $\sim 1500$  km over the equator. However, in my 1 May lecture as well as in response to a question at the end I emphasized that we had no definitive identification of particle species and that the particles might be penetrating protons or electrons. I did, however, regard protons and electrons of energies necessary to penetrate the Geiger tube directly, namely  $E_p > 35 \text{ MeV}$  and  $E_e > 3 \text{ MeV}$ , as unlikely in view of our auroral zone measurements with

rocket-borne equipment. Later in 1958 we showed that this opinion was incorrect.

Of the some 1500 real-time recordings of Explorer I telemetry signals in the period 1 February to 9 May 1958, 850 contained readable data.

A much larger body of data was obtained from Explorer III. During its 44 days of useful life, the satellite completed 523 orbits around the earth and 504 tape recorder play-backs were attempted. Of these attempts a total of 408 or 81% were successful. The tape recorder operated perfectly throughout this period and never failed to respond to a command that was electronically successful.

Virtually all of the Explorer I data were reduced and analyzed (Yoshida, Ludwig, and Van Allen, 1960) (Figs. 4 and 5). But only some 10% of the Explorer III data were reduced. Even as of 1981, George Ludwig and I regret not having completed the reduction of the major part of both the real-time and stored data from Explorer III. In early 1958, however, we considered that we had gleaned the principal elements of the data and hardly felt it worthwhile to continue because of the limited dynamic range of the instrument and our preoccupation with follow-on satellite missions carrying detectors specifically designed for the then-known intensities and for particle species identification and other more revealing properties of the radiation.

There is no evidence, so far as I can find, that Soviet scientists were aware of the existence of geomagnetically trapped radiation before my 1 May 1958 report of Explorer I/III observations. It may be noted that the high intensity radiation found by these two satellites was a part of what was later recognized as the inner of two major radiation belts.

#### RADIATION OBSERVATIONS WITH SPUTNIK III

The third Soviet satellite, Sputnik III, was launched on 15 May 1958 into an orbit inclined at  $65^\circ$  to the equator. Its orbit, similar to that of Sputnik II, had perigee and apogee altitudes of 225 and 1880 km, respectively. Among other equipment, Sputnik III carried a scintillation detector for the observation of photons and charged particles. The scintillator was a cylinder of crystalline NaI (Tl) 4.0 by 3.9 cm in size, mounted on a photomultiplier tube whose photocathode had a diameter of 4.0 cm. The effective shielding of the detector was at least  $1 \text{ g cm}^{-2}$ . Three quantities were transmitted: (a) the counting rate of pulses corresponding to energy release in the crystal of 35 keV or more, (b) the anode current of the photomultiplier tube, and (c) the current to an intermediate dynode. A preliminary report of the observations with this instrument was made in a special lecture by S. N. Vernov at the Fifth General Assembly of CSAGI (Special Committee for the International Geophysical Year) in Moscow, 30 July -- 9 August 1958. Among U. S. delegates to this meeting were Ernest Ray of the Iowa group, who gave a preliminary report of Explorer IV results obtained during the first few days of its flight (see later section). The first written report (available in English) of Sputnik III observations that I have found was submitted to Planetary and Space Science on 17 January 1959 and published soon thereafter (Vernov and colleagues, 1959). The stated intention of this investigation was "to obtain data on photons at high altitudes." The choice of detector and associated data handling system does not suggest any prior knowledge of the existence of geomagnetically trapped radiation nor is there any hint of such prior knowledge in the text of the paper. Nonetheless, as with Explorer I, the principal results related to electrically charged particles (in part, via their bremsstrahlung).

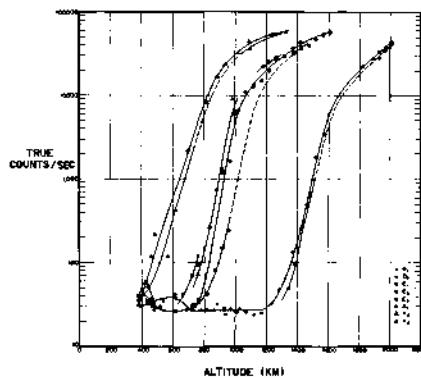


Fig. 4. Altitude dependence of the true counting rate of the Geiger tube on Explorer I as observed within  $19^{\circ}$  of the magnetic dip equator at different geographic longitudes approximately as follows: two curves on the left,  $288^{\circ}$  E; three curves in the center,  $5^{\circ}$  E; and two curves on the right  $105^{\circ}$  E (Yoshida, Ludwig, and Van Allen, 1960).

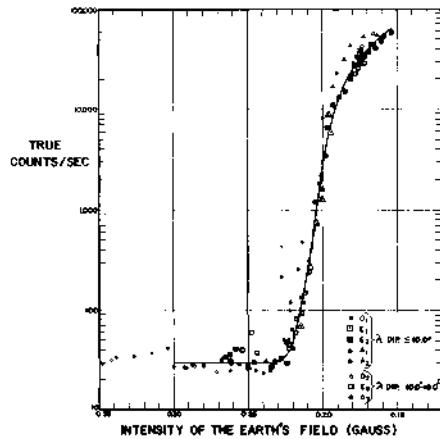


Fig. 5. All of the counting rate data of Fig. 4 replotted as a function of scalar magnetic field strength, illustrating the validity of the geomagnetic trapping hypothesis (Yoshida, Ludwig, and Van Allen, 1960).

Vernov and colleagues distinguish two observed zones of high intensity radiation -- a polar zone (latitudes  $52^{\circ}$  to  $65^{\circ}$  N) and an equatorial zone. In the section of their paper entitled "Electronic Component of Cosmic Rays in the Polar Region", the authors state, in part:

"Comparing this increase of ionization and increase of the counting rate, one can estimate photon energy of the bremsstrahlung radiation and, consequently, the energy of electrons. The photon energy will be of the order of 100 keV or less. The sign 'less' arises due to the fact that in some cases only a low limit of counting rate is known. However, the photon energy cannot be significantly lower than 100 keV, since the threshold is 35 keV. Thus, the most probable value of the energy of electrons responsible for the effect is about 100 keV. ... Thus, the polar region is characterized by a permanent electron flux, the intensity of which varies in a wide range. Taking into account the efficiency of electron registration by the bremsstrahlung radiation, we estimate the electron flux as  $10^3$  to  $10^4$  particles  $(\text{cm}^2 \text{ sec sterad})^{-1}$  ... "

The authors conclude this section of their paper as follows:

"At present it is difficult to give a complete interpretation of the electron component observed. It is not excluded that these electrons are accelerated near the earth by electric fields similar to those assumed to be in the aurora. But it is possible also that the electron component comes from far away, for example, from the sun; and penetrates, despite the small amount of energy of the particles, through the earth's magnetic field due to the difference between the field and that of an ideal dipole."

On a purely observational basis the Sputnik III data represented discovery of the earth's outer radiation belt inasmuch as they were acquired before those of Explorer IV and Pioneer III (see later section). However, the interpretation of Vernov and colleagues did not encompass the idea of geomagnetic trapping for what they called the "electron component of the cosmic rays in the polar region."

The next section of this important paper is entitled "Zone of High Intensity in the Equator Region." Therein, the authors confirmed our Explorers I/III results and added significant detail on latitude and altitude dependence of intensity. They also found the radiation intensity in the equatorial region to have approximate north-south symmetry relative to the geomagnetic equator and to be stable in time. The time stability at low latitudes was contrasted to the marked temporal variability at high latitudes. They estimated that the energy deposition rate in the NaI crystal was of the order of  $5 \text{ erg sec}^{-1}$  at an altitude of 1200 km near the geomagnetic equator.

The authors remark:

"Independently of the mechanism of production of particles in the equatorial zone, it is obvious that the main role in this effect is played by the factor of storing. This fact is proved by the concentration of particles in the equatorial zone where, at a sufficient altitude, they can oscillate for a very long period of time."

At the end of this section of the paper, the authors conclude as follows:

"It is possible now to make a hypothesis that the equatorial zone of high intensity is placed symmetrically to the earth's magnetic field and is characterized by a strong concentration of the density of particle flux in the plane of magnetic equator. Such an equatorial zone provides, apparently, ideal conditions for particle oscillations in a magnetic field, and leakage is determined probably only by ionization losses and losses for radiation. In this case such a long lifetime

of particles is possible (of the order of a year) that even such a weak mechanism of injection as the decay of neutrons leaving the atmosphere under the action of cosmic rays appears to be sufficient to explain the observed intensity."

This conclusion clearly embraces the concept of geomagnetic trapping as appropriate for the equatorial region and suggests that the decay of neutrons produced in the atmosphere by cosmic rays provides a possible source. However, no observational identification of the nature of the trapped particles was achieved nor was any information on intensity vs. range obtained. Also it was not clear whether the authors were thinking of low energy or high energy neutrons or of electrons or protons or both as decay products thereof.

On Sputnik III there were two other detectors relevant to the present discussion. Each detector consisted of a fluorescent film of ZnS(Ag), 5 cm in diameter and  $2 \times 10^{-3}$  g cm $^{-2}$  in thickness, deposited on the face of a photomultiplier tube. Aluminum foils of thickness  $8 \times 10^{-4}$  g cm $^{-2}$  and  $4 \times 10^{-4}$  g cm $^{-2}$  covered the fluorescent films of the respective detectors. The primary purpose of this instrument was to detect and measure corpuscular streams, especially those which cause aurorae and contribute to heating the upper atmosphere. Again the principal results related to geomagnetically trapped particles at relatively low latitudes. V. I. Krasovskii made a preliminary report on his investigation at the Fifth General Assembly of CSAGI in Moscow (Krasovskii and colleagues, 1958), including the concept of geomagnetic trapping, though a full report of findings was not published until some time later (Krasovskii and colleagues, 1961). Excerpts from the latter paper are as follows:

"... large currents of electrons with energies of about 10 keV were detected at 1900 km above the southern part of the Pacific Ocean. The intensity of these currents was very high, and in the majority of cases the apparatus gave off-scale readings, since it was not expected that such high intensities would be present. ... For all the off-scale readings, the energy flux of the electrons under investigation was in excess of 100 erg (cm $^2$  sec) $^{-1}$  up to 1900 km from the earth's surface. If electron currents of this intensity were capable of penetrating into the lower parts of the atmosphere, for example the F-layer of the ionosphere, then they could not have remained unnoticed, since they would appreciably increase the ionization of the upper atmospheres and would give rise to polar aurorae. Since such phenomena were not observed, the recorded currents were explained by an oscillatory motion of the electrons along the magnetic lines of force (July 1958, Fifth CSAGI Meeting, Moscow).

...

"The majority of the recorded electrons move near the plane perpendicular to the magnetic line of force, and the flux of electrons along the line of force in the direction toward the earth (downward direction) is greater than in the opposite direction. Electrons moving toward the earth at small angles to the magnetic line of force were also recorded, and this suggests that particles which penetrate into the lower part of the atmosphere appear as a result of some processes occurring at distances in excess of 1900 km. It was established that the energy flux of electrons with a magnetic moment enabling penetration into the F-layer of the ionosphere can reach values of about 1 erg (cm $^2$  sec) $^{-1}$ ."

FOLLOW ON INVESTIGATIONS WITH EXPLORER IV  
AND PIONEERS I, II, AND III

The next satellite to carry radiation instruments was Explorer IV, under the joint auspices of the U. S. National Committee for the International Geophysical Year and the Department of Defense. Its purpose was twofold: (a) to follow up on the discoveries of Explorers I/III with a system of detectors of adequate dynamic range to cope with the then-known intensity of trapped radiation and of more discriminating properties than those used in Explorers I/III and (b) to observe the effects of a planned series of high altitude atomic bomb tests. The bomb tests had been proposed within the Atomic Energy Commission by Nicholas Christofilos in late 1957 (Christofilos, 1959) for the purpose of producing artificial radiation belts of electrons from the decay of radioactive fission products of the bombs. The University of Iowa group was entrusted with preparing the package of radiation detectors. We settled on an array of four radiation detectors:

Detector A: A circular disc of plastic scintillator, thickness  $0.178 \text{ cm}$ , diameter  $0.762 \text{ cm}$ , cemented on the face of an end-window photomultiplier tube and covered by  $0.14 \text{ g cm}^{-2}$  of aluminum (pulse counting).

Detector B: A circular disc of CsI(Tl) scintillating crystal, thickness  $0.203 \text{ cm}$ , diameter  $0.762 \text{ cm}$ , cemented on the face of an end window photomultiplier tube, with the crystal coated with an evaporated layer of aluminum ( $0.2 \text{ mg cm}^{-2}$ ) and further shielded by a nickel foil of  $0.8 \text{ mg cm}^{-2}$  thickness (quasi-logarithmic electrometer for anode current).

Detector C: A miniature Geiger tube having an omnidirectional geometric factor of  $0.14 \text{ cm}^2$  for  $1.2 \text{ g cm}^{-2}$  shielding and its full value of  $0.705 \text{ cm}^2$  for about  $5 \text{ g cm}^{-2}$  shielding (pulse counting).

Detector D: A nearly identical Geiger tube, enclosed in an additional lead shield of  $1.6 \text{ g cm}^{-2}$  thickness so that its omnidirectional geometric factor was  $0.14 \text{ cm}^2$  for  $2.8 \text{ g cm}^{-2}$  shielding and  $0.823 \text{ cm}^2$  for  $6 \text{ g cm}^{-2}$  shielding (pulse counting).

Detectors A and B were clearly directional by virtue of physical collimators and the geometric shape of the sensitive elements; Detectors C and D were sensitive omnidirectionally but there was an unavoidable variation of the shielding in different directions. Every effort was made to provide the necessary dynamic range to cope with the intensity of the natural radiation and the estimated intensity of that to be artificially injected.

Upgraded high speed stages of the four-stage launching vehicle made it possible to plan an increase in inclination of the satellite orbit from the  $33^\circ$  of the orbits of Explorers I and III to  $50^\circ$  in order to provide improved coverage in latitude. Explorer IV was launched up the east coast of the United States on 26 July 1958 into an orbit of  $50^\circ$  inclination to the equator with perigee and apogee altitudes of 260 and 2200 km, respectively.

All detectors operated properly and there was soon a large amount of fresh observational data available. The findings of Explorers I/III were massively confirmed and major advances were made in knowledge of the distribution of intensity and the nature of the geomagnetically trapped radiation (Van Allen, McIlwain, and Ludwig, 1959a). Energy fluxes as great as  $\sim 100 \text{ erg (cm}^2 \text{ sec sr)}^{-1}$  were measured with Detector B at high altitudes near the geomagnetic equator and excellent angular distributions were obtained. Comparison of the rates of Detectors C and D showed that the radiation near the geomagnetic equator was much more penetrating than that at high latitudes and was not compatible with the bremsstrahlung interpretation that I had suggested earlier. In agreement with the data from

Sputnik III, the distribution of intensity as observed near the earth was found to be bifurcated into a low latitude region and a high latitude region separated by a "slot" of lesser intensity at a geomagnetic latitude of about  $48^\circ$ . The angular distributions of directional intensity and the altitude distributions of omnidirectional intensity in both regions were characteristic of magnetically trapped radiation. This result for the high latitude region contradicted the interpretation of Vernov and colleagues (1958), as quoted above. Contours of constant intensity turned away from the earth at geomagnetic latitudes exceeding about  $60^\circ$ , suggesting that aurorae occur at or beyond the high latitude boundary of trapping. The diversity of detectors on Explorer IV gave a parametric characterization of the radiation at each point along the orbit but we were still unable to reach unique conclusions on the relative proportion of electrons and protons because of ignorance of their respective energy spectra. In a later paper (Van Allen, 1959b) I gave a variety of alternative interpretations.

Meanwhile, the U. S. Department of Defense had produced bursts of two 10-megaton yield atomic bombs, called Teak and Orange, on 1 August and 12 August at approximate altitudes of 75 and 45 km, respectively, above Johnston Atoll in the North Pacific. Three Argus atomic bomb bursts (about 1.4 kiloton yield) were produced over the South Atlantic on 27 August, 30 August, and 6 September at altitudes of about 200, 250, and over 480 km, respectively.

We observed with Explorer IV the effects of all five bursts in populating the geomagnetic field with energetic electrons. Despite the large yields of Teak and Orange the effects on the radiation belts were small and of only a few days' lifetime because of atmospheric absorption corresponding to the relatively low altitudes of ignition.

The three Argus bursts produced clear and well observed effects and gave a great impetus to understanding geomagnetic trapping (Scientific Effects of Artificially Introduced Radiations at High Altitudes, 1959). About three percent of the available electrons were injected into durably trapped orbits. The apparent mean lifetime of the first two artificial radiation belts was about three weeks and of the third about a month. In all three cases a well defined and durable shell of artificially injected electrons was produced. Worldwide study of these shells provided a result of basic importance in giving a full geometrical description of the locus of trapping by "labeled" particles and an upper limit on the diffusion coefficient of the constituent electrons (Van Allen, McIlwain, and Ludwig, 1959b).

During the remainder of the year 1958 there were three U. S. moon shots, Pioneers I, II, and III, launched on 11 October, 8 November, and 6 December, respectively. The most successful of these was Pioneer III, which did not achieve earth escape velocity but had its apogee at  $17 R_E$  (geocentric). Radiation data were obtained along a full radial traversal of the radiation belt region for the first time -- in fact, along two different paths, one outbound and one inbound. It was found that the outer boundary of geomagnetic trapping was at a radial distance of about  $10 R_E$ . In addition, the combination of data from Pioneer III with the low altitude data from Explorer IV revealed for the first time the full geometrical form of the two distinctly different radiation belts, an inner and an outer one (Fig. 6) (Van Allen and Frank, 1959).

Pioneer III completed the discovery phase in the study of the earth's radiation belts, which were subsequently recognized as the distributions of the quite energetic components of the much greater population of (principally) lower energy particles constituting the full plasma physical system, later called the magnetosphere.

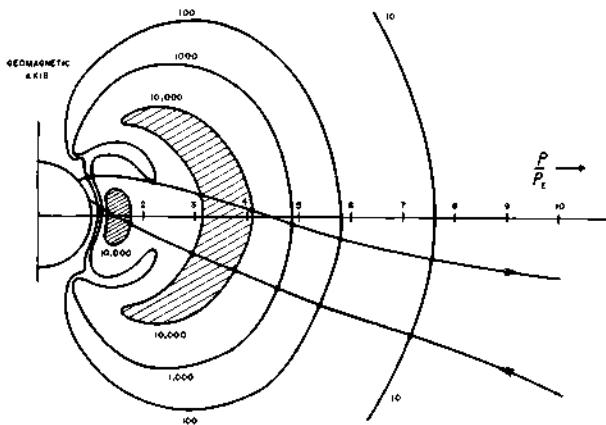


Fig. 6. Intensity structure of the trapped radiation around the earth. The diagram is a geomagnetic meridian section of a three-dimensional figure of revolution around the geomagnetic axis. Contours of constant intensity are labeled with numbers 10, 100, 1000, and 10,000. These numbers are the true counting rates of an Anton 302 Geiger tube carried by Explorer IV and Pioneer III. The linear scale of the diagram is relative to the radius of the earth, 6371 km. The outbound and inbound legs of the trajectory of Pioneer III are shown by the slanting, undulating lines. The intensity structure is a function of detector characteristics (Van Allen and Frank, 1959 as corrected in Van Allen, 1959b).

#### ACKNOWLEDGEMENT

I am very grateful for a Regents' Fellowship of the Smithsonian Institution for the period January-August 1981. During this time I have been in residence at the National Air and Space Museum in Washington, D. C. writing a monograph entitled "Origins of Magnetospheric Physics". The present paper for the XXXII Congress of the International Astronautical Federation in Rome, Italy, 7-12 September 1981 is an abridgment of a portion of that monograph.

#### REFERENCES

- Alfvén, H. (1950). Cosmical Electrodynamics. Clarendon Press, Oxford.
- Chapman, S., and J. Bartels (1940). Geomagnetism, Vols. I and II. Clarendon Press, Oxford.
- Christofilos, N. C. (1959). The Argus experiment. *J. Geophys. Res.*, **64**, 869-875.
- Jánossy, L. (1948). Cosmic Rays. Clarendon Press, Oxford.
- Krasovskii, V. I., Yu. M. Kushnir, G. G. Bordovskii, G. F. Zakharov, and E. M. Svetlitskii (1960). The search for corpuscles with the help of the third artificial satellite. In Artificial Earth Satellites, Vol. 2. Plenum Press, New York, pp. 75-77 (also Uspekhi Fiz. Nauk, USSR, **64**, 425, 1958).

- Krasovskii, V. I., I. S. Shklovskii, Yu. I. Galperin, E. M. Svetlitskii, Yu. M. Kushnir, and G. A. Bordovskii (1961). Discovery of approximately 10-kev electrons in the upper atmosphere. In Artificial Earth Satellites, Vol. 6. Plenum Press, New York, pp. 137-155.
- Ludwig, G. H. (1959a). Cosmic-ray instrumentation in the first U. S. earth satellite. Rev. Sci. Instr., 30, 223-229.
- Ludwig, G. H. (1959b). The instrumentation in earth satellite 1958 gamma. M.S. Thesis, University of Iowa Research Report 59-3. (Also Satellites 1958 Alpha and Gamma: High Intensity Radiation Research and Instrumentation, IGY Satellite Report No. 13, National Academy of Sciences, Washington, D. C., 1961, pp. 31-93.)
- Meredith, L. H., M. B. Gottlieb, and J. A. Van Allen (1955). Direct detection of soft radiation above 50 kilometers in the auroral zone. Phys. Rev., 97, 201-205.
- Mitra, S. K. (1952). The Upper Atmosphere, 2nd ed. The Asiatic Society, Calcutta.
- Montgomery, D. J. X. (1949). Cosmic Ray Physics. Princeton University Press, Princeton, New Jersey.
- Newell, H. E., Jr. (1953). High Altitude Rocket Research. Academic Press, New York.
- Scientific Effects of Artificially Introduced Radiations at High Altitudes -- A Symposium (1959). (Seven observational and theoretical papers by various authors). J. Geophys. Res., 64, 865-938.
- Störmer, C. (1907). Sur les trajectoires des corpuscules électrisés dans l'espace sous l'action du magnétisme terrestre. Ser. IV, Arch. Sci. Phys. et Naturelles, 24, 317-364.
- Störmer, C. (1955). The Polar Aurora. Clarendon Press, Oxford.
- Van Allen, J. A. (1956). Cosmic-ray observations in earth satellites. In J. A. Van Allen (Ed.), Scientific Uses of Earth Satellites, Chapter 20, University of Michigan Press, pp. 171-187.
- Van Allen, J. A. (1957). Direct detection of auroral radiation with rocket equipment. Proc. Natl. Acad. Sci., 43, 57-62.
- Van Allen, J. A. (1959a). Balloon-launched rockets for high-altitude research. In H. E. Newell, Jr. (Ed.), Sounding Rockets. McGraw-Hill, New York. And author's journal of 1957 Antarctic expedition of U.S.S. Glacier.
- Van Allen, J. A. (1959b). The geomagnetically-trapped corpuscular radiation. J. Geophys. Res., 64, 1683-1689.
- Van Allen, J. A. (1961). Observations of high intensity radiation by satellites 1958 alpha and 1958 gamma. Satellites 1958 Alpha and Gamma: High Intensity Radiation Research and Instrumentation, IGY Satellite Report No. 13, National Academy of Sciences, Washington, D. C., pp. 1-22.
- Van Allen, J. A., and L. A. Frank (1959). Radiation around the earth to a radial distance of 107,400 km. Nature, 183, 430-434.
- Van Allen, J. A., and M. B. Gottlieb (1954). The inexpensive attainment of high altitudes with balloon-launched rockets. In R. L. F. Boyd and M. V. Seaton (Eds.), Rocket Exploration of the Upper Atmosphere. Pergamon Press, London.
- Van Allen, J. A., G. H. Ludwig, E. C. Ray, and C. E. McIlwain (1958). Observation of high intensity radiation of satellites 1958 alpha and gamma. Jet Propulsion [September 1958], pp. 588-592.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig (1959a). Radiation observations with satellite 1958 e. J. Geophys. Res., 64, 271-286.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig (1959b). Satellite observations of electrons artificially injected into the geomagnetic field. J. Geophys. Res., 64, 877-891.
- Vernov, S. N., A. E. Chudakov, E. V. Gorchakov, J. L. Logachev, and P. V. Vakulov (1959). Study of the cosmic-ray soft component by the 4rd Soviet earth satellite. Planet. Space Sci., 1, 86-93.

- Vernov, S. N., N. L. Grigorov, Iu. I. Logachev, and A. E. Chudakov (1958). Cosmic radiation measured on the second artificial satellite. Doklady Akademii Nauk, USSR, 120, 1231-1233 (in English, Soviet Physics Doklady, 3, 617-619, 1958).
- Yoshida, S., G. H. Ludwig, and J. A. Van Allen (1960). Distribution of trapped radiation in the geomagnetic field. J. Geophys. Res., 65, 807-813.

# THE SUPERSONIC WIND TUNNEL INSTALLATIONS AT PEENEMUNDE AND KOCHEL AND THEIR CONTRIBUTIONS TO THE AERODYNAMICS OF ROCKET- POWERED VEHICLES

R. Hermann

*Aerospace Science and Engineering, The University of  
Alabama in Huntsville, Huntsville, AL 35899, USA*

## ABSTRACT

The goal in 1937 was to build an aerodynamic-ballistic research institute to provide all aerodynamic, stability and heat transfer data needed for development of rocket-powered vehicles. Design and construction of the supersonic wind tunnel, with 40cm x 40cm test-section, is described, which was operational at Peenemunde in 1939 and moved to Kochel in 1943. Seven wind tunnel Laval nozzles, ranging up to Mach number 5, were in use. Their theory, design and construction is outlined and their wave patterns are illustrated. Detailed measurements of their flow parameters are shown. Results are presented for two groups: for vehicles without wings such as the A4; and for vehicles with wings including delta wings such as the "Glider A9" or the "Wasserfall." The influence of exhaust jet on drag and stability is discussed.

## KEYWORDS

Rocket powered vehicles; supersonic aerodynamics performance; supersonic wind tunnel development.

### A. PURPOSE, DESIGN AND OPERATION OF THE SUPERSONIC WIND TUNNEL FACILITIES

#### Purpose and Goal

The goal in 1937 was to build an aerodynamic-ballistic research institute, capable of furnishing - in a reasonable time frame - all aerodynamic, stability, aerodynamic control, and heat transfer data needed for the development of numerous projects, such as supersonic projectiles (fired from guns), rocket-powered vehicles without wings, stabilized by fins (called missiles) and rocket-powered supersonic vehicles with wings and fin-assemblies or with delta wings. This required the design and construction of supersonic wind tunnels of highest obtainable Mach No. (at least 5 and above), which also should operate in the subsonic range, since all rocket-powered vehicles start with zero velocity.

From 1934-1937, at the Aerodynamic Institute of the University of Aachen, under the direction of Professor Carl Wieselberger, the author had developed, built and operated a supersonic wind tunnel of the blow-down type with a 10 cm x 10 cm test section up

to  $M = 3.3$  with a 3-component balance. Financed by the German Luftwaffe, some of the first models tested were anti-aircraft projectiles.

On January 6, 1936, Dr. Wernher von Braun, of the Germany Army Ordnance, brought the drawing of a pointed, slender body with fins, to be tested for drag and stability up to our maximum Mach No. It was the rocket-powered missile later designated A3. The first proposed fin assembly would have caused an unstable missile. Experiments at the institute produced a stable arrangement (see Section C - A3 Vehicle).

For more detailed measurements larger models and larger tunnel cross-sections were required. More power was needed because of model size and the requirement to test at higher Mach numbers. This showed the necessity to build such a facility with the aid of the Army Ordnance at the Experiment Station just under construction at Peenemunde. Our work started April 1937 with 20 people. Emphasis was placed on combining available scientific basic concepts with a detailed engineering design and a solid construction of all test equipment and modern measuring devices to have a facility for operation for maximum tunnel usage.

Additional assistance for the scientist should be provided by a series of laboratories, working in the area of optics (Schlieren, Interferometer), thermodynamics of high speed gases, heat transfer, electric controls for operation of powerplant and air drying installation, electronic measurement devices, and physics (X-ray densitometer). Their task was the development of original measuring devices, needed for the solution of problems arising during the development of new missile projects and their counterparts in the wind tunnel.

The goals should be easier attained by having a Chief Engineer, with his own staff of qualified designers and with the help of ample design offices, combined with a machine shop, under the direction and control of the design staff. The result was that ideas and proposals of scientists could be designed and fabricated properly, and made ready for testing.

#### Design and Operation

The work in Peenemunde started in April 1937; the tunnel was first operational in the summer 1939 (60 people). After the air raid, the tunnel was moved to Kochel (200 people) in the fall of 1943, rebuilt and put in operation by spring 1944, and step by step improved through the fall (250 people). In Peenemunde, the vacuum reservoir was  $1,000 \text{ m}^3$  (Kochel  $750 \text{ m}^3$ ); 6 vacuum pumps with 800 KW total (Kochel first 4 pumps); 2 test sections of  $40 \text{ cm} \times 40 \text{ cm}$  (Kochel 1 test section); 1 test section of  $18 \text{ cm} \times 18 \text{ cm}$  capable of continuous operation through  $M = 3.3$ . During blow-down operation Peenemunde had a running time of 20-25 sec.; while Kochel had only 15 sec. The range of Mach numbers covered and nozzles are described in Section B.

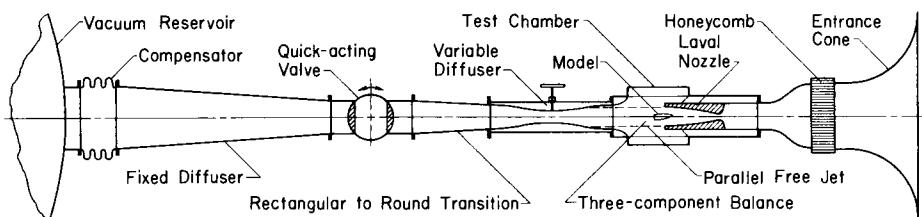


Fig. 1. Supersonic wind tunnel layout ( $40 \text{ cm} \times 40 \text{ cm}$ ) at Peenemunde. (1937)

Figure 1 shows the complete layout. The air enters the entrance cone, in front of

which is the air drying filter, not shown. Air passes through a honeycomb, enters the test chamber with two Laval nozzles, producing a parallel free jet, which is captured by the bell mouth of the variable diffuser. The diffuser had a fast moving throat, to control the pressure in the chamber, needed for a parallel jet. The air flowed through the quick acting valve, which opened and closed at the beginning and end of the usable running time. A fixed diffuser brings the air into the vacuum reservoir.

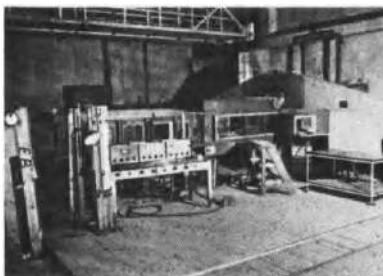


Fig. 2. Test section at Kochel. (Photo 1945)

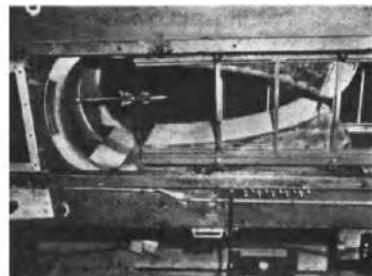


Fig. 3. Test chamber with Wasserfall model. Kochel (1945)

A photo of the test section installed at Kochel is shown in Fig. 2. The movable Schlieren equipment is on top, the variable diffuser and pressure instrumentation is seen. Details of the test chamber with nozzle for  $M = 3.24$  is shown in Fig. 3. A "Wasserfall" model is installed, ready for testing, supported by the mount which varies the angle of attack. This can be done from outside by push-button control without opening the test section; another feature for maximum tunnel usage.

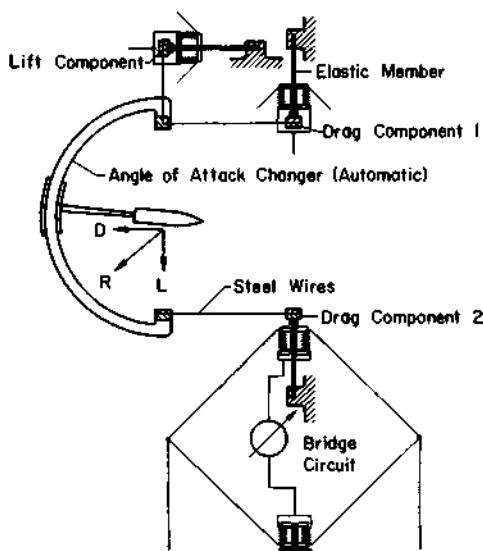


Fig. 4. Schematic of the 3-component balance.

An important measuring equipment is the 3-component balance seen in Fig. 4 for drag, lift and pitching moment. It was built in Peenemunde in 1939 and also used in

Kochel through 1945. It is very ruggedly built, installed in a pressure tight chamber, which is evacuated at each run down to the low static pressure of the respective Mach No. The angle of attack changer is seen here in the schematic, while in a photo in Fig. 3. To determine the C.P., the drag is measured in two components. In customary subsonic wind tunnel balances for aircraft use, the lift is measured in two components. Our system was needed because many of our models near zero angle of attack had no lift, but they always had drag.

The Schlieren-optical system built in 1937 has two nearly parabolic mirrors, 50 cm diameter each, with 10 m distance. It is for observation of flow fields around models; large glass walls on both sides of test chamber (see Fig. 2 and 3) allow full view for Schlieren or Interferometer.

The Schlieren system is mounted on a four-wheel carriage, running along the test section; and also vertically, to move the system fast and safe for observation at any point and also from one test section to the other. The Interferometer (since 1943) has mirror diameters of 20 cm and field of view 9 cm x 10 cm.

The installation of the air drying system was completed in spring 1940. It is based on detailed experiments about condensation shocks caused by humid air done by the author in Aachen in 1935-36. The dryness requirement increases with Mach No. This is a two stage drying system specially developed for intermittent operations, using 600 KW power for heating and airflow. The dryness reached was 0.2 g Water/kg air, or 0.0002 dimensionless. This is a factor 10 smaller than used before (at  $M = 3$ ). This guarantees shock free flow to  $M = 5.2$  in the humid climate.

## B. AERODYNAMIC DATA FOR THE SUPERSONIC TUNNELS AND SOME BASIC RESEARCH CONDUCTED

### 1. Laval Nozzles, Theory and Design.

(a) Basics. The nozzle within a supersonic wind tunnel is of paramount importance for conducting correct and meaningful experiments. Its theory is complex, the mathematical design cumbersome and the mechanical design and construction requires best specialists. The nozzle has two tasks. First to produce the desired supersonic velocity in the test chamber. A Laval type nozzle is applied which produces subsonic velocity in the converging section,  $M = 1$  at the throat, and supersonic velocity in the divergent section. The ratio of exit area to throat area determines the Mach number. For non-movable nozzles this requires a set of nozzles, with decreasing throat area for increasing Mach No., since the exit cross section is kept about equal to accomodate the models.

Second, the flow in the so-called measuring rhombus should be parallel, of constant Mach No. and free of shocks. This can be achieved by forming the walls of the nozzle according to the method of characteristics first given by Prandtl and Busemann, applicable to two dimensional flow. Hence, the nozzles are two dimensional with two equal portions symmetric to the axis. The method is basically graphical, but was improved by the staff of the wind tunnel by substituting calculations.

(b) Nozzles used at Peenemunde and/or Kochel. They have two distinctions. First, they were achieved by the Busemann method, and represent the shortest length of nozzle which can be achieved by any means for a given throat size and shape. Second, their mechanical design and construction is novel, accurate and effective, when boundary layer corrections are needed.

Shortest length is important for high Mach No. 4, 5, ..10 (for future projects) for reasons of weight, handling and for minimizing boundary layer thickness. Each nozzle has a particular radius of curvature at the throat, producing expansion up to the point of inflection, where the angle, made by the profile with nozzle axis,

is denoted by  $\Phi_B$ . If  $v$  is the Prandtl-Meyer angle which provides the desired Mach No. at the exit, then theory shows, the condition:  $\Phi = \frac{1}{2}v$  causes the minimum length of the supersonic portion. This condition is satisfied in all our nozzles. Figure 5 illustrates the discussion and shows wave pattern for the nozzle  $M = 4.4$ . This high Mach No. is selected, to show the complexity of the design (completed in 1938). From a large original drawing it is reduced twice by photographic means. Lower Mach No. are much simpler, for instance  $M = 1.56$  has only eight wavelets.

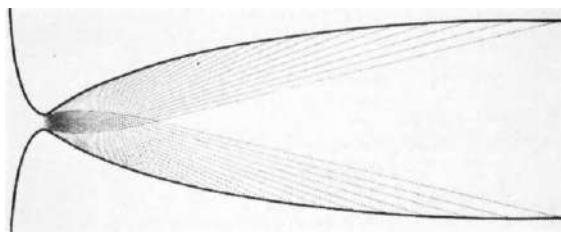


Fig. 5. Wave pattern for nozzle  $M = 4.4$ .

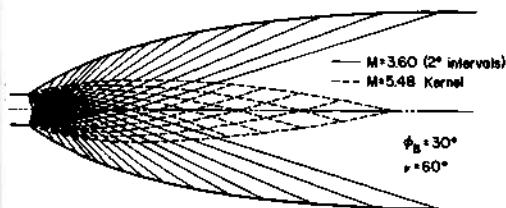


Fig. 6. Wave pattern for sharp throat nozzle.

Table 1 shows characteristic dimensions of six of our nozzles ( $M = 1.22$  to  $4.38$ ) in use for many years, and the last ( $M = 5.18$ ) was completed in Jan. 1945 but not thoroughly tested. The full length of the measuring rhombus gives the allowable maximum model length. The last column is the slope of profile at exit (no dimensions). It is equal to the angle of divergence in radians. The divergence is the consequence of making allowance for the displacement thickness of the boundary layer. The particular value of this angle was obtained by experiments.

TABLE 1 Characteristic Dimensions of Wind Tunnel Nozzles  
(40 cm x 40 cm)

Design Mach No.	$\Phi_{B\max}$ degree	Throat width mm.	Exit width mm.	Length of supersonic part mm.	Length of measuring rhombus mm.	Slope of profile at exit radians.
1.22	2	386.16	399.76	271.0	?	0
1.56	~ 6	325.82	403.76	495.8	480	0.008
1.86	~11	259.20	401.84	593.7	602	0.012
2.48	~20	152.10	404.76	787.4	904	0
3.24	~27	71.52	408.74	920.0	1204	0.004
4.38	35	25.94	400.0	1080.3	1704	0.010
5.18	?	14.30	422.56	1224	?	0.010

(c) Sharp throat nozzles. With higher Mach No. the distance between the throat and the inflection point becomes physically smaller and smaller, hence harder to manufacture (and measure) according to design specifications. This led to the idea of a throat with a sharp edge immediately followed by an expansion around the corner up to  $\Phi_{B\max} = \frac{1}{2}v$ . The inflection point moved up to the throat, reducing the length,  $\ell$ , to an absolute minimum. For  $M = 5$ ,  $\ell/h = 2.56$ ; for  $M = 10$ ,  $\ell/h = 4.98$ . ( $h$  is height of exit cross section.)

Nozzles with sharp throats can be designed with advantage, since all wavelets originate at the sharp edge. Figure 6 illustrates where a nozzle for  $M = 3.60$  is designed from given wavelet field, called Kernel, corresponding to  $M = 5.48$ .

## 2. Nozzle Construction.

The nozzle has a composite construction consisting of a steel base, brass sides or formers, and a cement profile. The cement's name is monolith. The main structure is nothing more than a mold for the cement. The formers, fixed to the sides of the nozzle structure, are of brass, 5 mm thick. They have been shaped to the calculated nozzle profile to an accuracy of about 0.01 mm, final shaping being done by hand. Figure 3 shows part of those details. The cement is poured into the nozzle structure and shaped with the aid of a straight edge. The depth of the monolith layer ranges from 10 to 20 mm. Monolith has several advantages over wood or steel: flexibility to changes in the profile; no special machinery required as for steel; no deformation or cracking of wood. A complete pair of nozzle blocks requires a total of about 2,000 man-hours.

## 3. Experimental Investigation of Nozzle Flow.

Once the nozzle blocks were completed, the flow field was carefully investigated. The procedure varied somewhat with Mach No., but consisted of Schlieren photography, static tube and/or pitot tube measurement. For  $M < 3.3$ , pressure distribution along the axis was measured with a long static tube, which stretched from upstream of the throat to downstream of the measuring section. The tube had nine static holes and the tube was moved lengthwise to cover all spots. Mach No. distribution, measured from Schlieren, was good up to  $M = 3.3$ , above that the definition of the wavelets was poor even when ruled scratches in the nozzle surface were used. Then the static pressure and Mach No. were calculated from pitot pressure reading alone. The various variables were converted into each other, using isentropic flow, an assumption which was carefully proven previously.

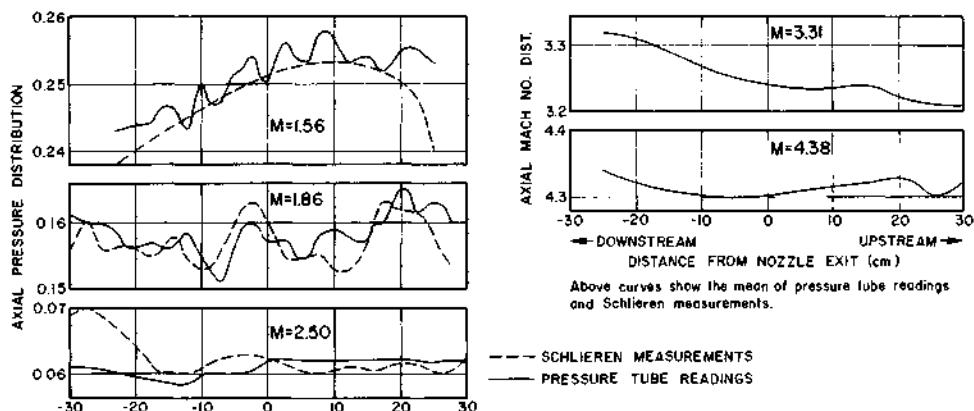


Fig. 7. Measured axial distribution for pressure and/or Mach number for 5 nozzles.

In Fig. 7, the first three show both pressure tube and Schlieren determinations separately. The last two nozzles show the mean of pressure tube and Schlieren. The agreement, on the whole, between pressure and Schlieren measurement is surprisingly good. It follows from the figures, that the maximum deviations of Mach No. from the mean, are about  $\pm 2.5\%$  in the worst case ( $M = 1.22$ ; not shown here) and about  $\pm 0.5\%$  in the best case ( $M = 4.38$ ).

## 4. Some Basic Research Conducted.

When the reader assesses the amount of scientific and engineering accomplishment presented in this historic paper, it is obvious that this could be accomplished only by a great amount of basic research, conducted simultaneously by the staff of

our supersonic wind tunnel installation and by outside consultants, most professors, who worked on contract with us. The author had, in summer 1945, prepared such a summary of basic research in gasdynamics, aerodynamics, thermodynamics and ballistics, a fraction of which is now used.

#### Theoretical gasdynamics.

- (a) Flow in Laval nozzle and the application of the Prandtl-Busemann method with corrections to include the boundary layer for the purpose of producing a homogeneous parallel flow.
- (b) Theoretical work to investigate the stationary, normal shock in the Laval nozzle, including the role of entropy also the "normal condensation shock."
- (c) Professor Tollmien of Dresden developed a method of calculating supersonic flow with rotational symmetry. Pressure distribution along the A4 body was calculated and together with another method developed by Professor Sauer of Aachen compared to experiment.
- (d) The Source-Sink method, first developed by von Karman and Moore, was used to calculate the pressure distribution for axisymmetric, slender bodies and compared with experiment.
- (e) The theory for supersonic flow around an infinite wedge and cone was evaluated with the help of special diagrams and nomograms. The shock waves around cones were examined in a series of Schlieren pictures and compared with the theory.

### C. AERODYNAMIC PERFORMANCE OF ROCKET-POWERED VEHICLES OBTAINED IN THE SUPERSONIC TUNNELS

#### 1. Requirements for Rocket-Powered Vehicles.

Some requirements for such vehicles when under development, are listed here:

- (a) Optimum aerodynamic shapes for minimum drag.
- (b) Lift forces required for certain missions. They are different for a ballistic trajectory, a gliding trajectory or an anti-aircraft missile trajectory.
- (c) Aerodynamic stability in both subsonic and supersonic ranges. Vehicle is required to be definitely stable, but stability must not be too large, since otherwise control requirements of movable rudders with their hinge moments will be excessive.
- (d) Aerodynamic pressure distribution loading and aerodynamic heating on bodies, fins, wings, rudders - needed for the stress analysis in structural design.
- (e) Hinge moments for rudders in the air stream or in the exhaust jet, required for control purposes, for pitch, yaw and roll motion.

Two basic groups of vehicles may be distinguished:

- 1) Rocket-powered vehicles without wings, stabilized by aerodynamic fins; typical: A4, with its forerunners A3 and A5.
- 2) Rocket-powered vehicles with wings stabilized by portion of the wing or by extra fins; typical: "Glider A9" with lift required to stretch the ranges or "Wasserfall" with lift required for maneuverability as anti-aircraft missile.

#### 2. Rocket-Powered Vehicles Group (1).

A3 vehicle. A forerunner of A4 - body shape similar to A4 but with long, narrow fins with a ring at their end, required for containing radio antennas, see Fig. 9. Aerodynamic data (3-components) measured by the author in the supersonic wind tunnel at Aachen with 10 cm x 10 cm cross section in 1936 and early 1937. A3 was the first vehicle, proved by wind tunnel tests to be fin-stabilized in the supersonic range (up to  $M = 3.3$ ). It had a slender body of eight caliber length with sharp pointed nose for minimum resistance in supersonics. It was fired four times in December 1937, reaching only moderate height (400 m) and subsonic speed. A3 was aerodynamically stable around its C.G. (no "somersaulting") and its trajectory went into the wind, as required for aerodynamically stable vehicles. Reason for the short flights was that the gyro control system was too slow and the moments produced by the exhaust

jet rudders too small to counteract heavy winds (about 25 m/sec).

A5 vehicle. The A5 was also a forerunner of A4. Due to the trajectory failures encountered with the A3, another small vehicle had to be developed, with thrust, weight and body size about the same as A3, but with a fast control system to avoid the deficiencies of the A3. Four large air rudders were added at the end of the fins, to produce pitch, yaw and rolling moments to counteract strong natural winds (100 m/sec) at higher altitudes. It was called A5, see Fig. 9, since A4 was already on the drawing board. The fin and rudder assembly was named V12. On purpose, the body of A5 had the same geometry as that of A4. Nearly all aerodynamic data gained for A5-V12 could be used later for the A4-V12. Starting October 1939 through 1941, about 70 to 80 of the A5 vehicles with the completely new control system using a stabilized platform were successfully launched reaching a ceiling of 8 km, range of 15 km, maximum speed about 300 m/sec, close to  $M = 1$ .

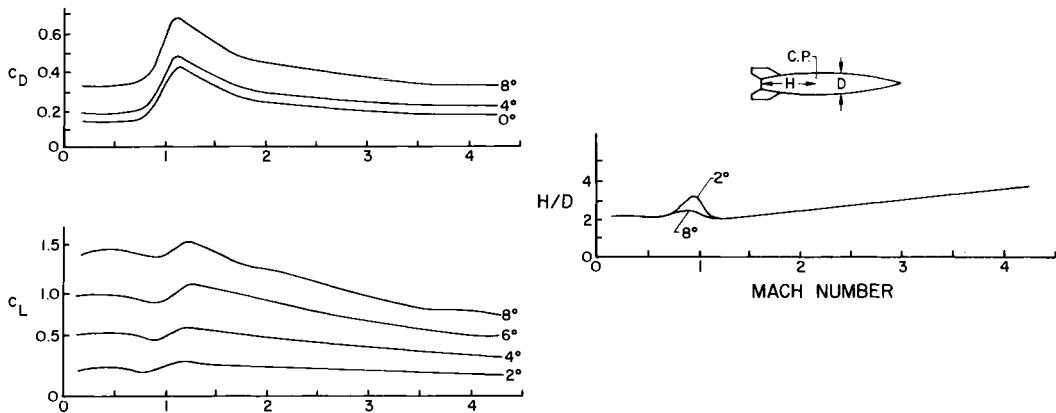


Fig. 8. Drag, lift and center of pressure for A4.

#### A4 vehicle.

Basic aerodynamic data. The political name was V2. Shape of the A4 as a wind tunnel model can be seen in Fig. 9, when disregarding the wings of V12a or V12c. As outlined above, all aerodynamic data, such as drag, lift and center of pressure (C.P.) and others needed for development were measured. They are shown in Fig. 8, obtained in subsonic tunnels up to  $M = 0.85$  and in supersonic tunnels for  $M = 1.2$  and up to  $M = 4.4$ . For  $M = 1$ , it was not possible to test at this time since transonic tunnels, in the present sense of the word, did not exist. Center of Pressure measurements on both sides of  $M = 1$  showed forward motion, indicating possible instability. Hence, the author suggested in July 1938 and initiated drop tests from airplanes from 7,000 m height of models of about 1.5 m length, built from steel. They were carried out in 1939 by the Air Force, but under direction of and evaluation by the staff of the Peenemunde supersonic tunnel. The models reached transition through  $M = 1$  at about 1,000 m altitude, which could be observed with cameras from the ground. The results were satisfactory and laid to rest the concern about instability in the transition region. It is important that the transition itself is of very short duration, since the actual vehicles are under strong acceleration.

Note, that all the curves in Fig. 8 have been drawn solidly through  $M = 1$ , in order to have consistent values for aerodynamic, structural and trajectory calculations. They present the best estimate of the aerodynamics and shall not imply that these values have been obtained.

Stability and control. For aerodynamic stability, the difference  $\Delta = C.G. - C.P.$

(measured from base of the vehicle) must be positive. However, it should be small as possible otherwise the vehicle cannot be controlled easily, since the control moment is proportional to  $\Delta$ . When  $\Delta$  is small, the air rudders can be small and then the hinge moment to turn the rudders are small - that is the power is small, and weight is saved in construction of rudders, motors, and energy storages (batteries). The curves C.P. as function of M (or of time travelled) and C.G. as function of time travelled must be established very exactly, in order to fulfill the requirement for a small  $\Delta$ . For A4,  $\Delta$  was mostly on the order of 0.3 of the diameter (0.3D); however, at M = 5.2 (end of powered trajectory)  $\Delta$  was suspected to be 0.2 to 0.1D.

Note, that the C.P. position can be reliably measured with the 3-component balance only to about  $\alpha = \pm 2^\circ$ . Towards  $\alpha = 0^\circ$ , the value becomes indefinite, 0/0. Due to the importance of stability, several other independent methods had been developed but cannot be listed here.

Pressure distribution over A4. As outlined before, aerodynamic pressure distribution must be measured over the body, fins and air rudders, to obtain local loading, needed for the structural design. One hundred-twenty finely drilled holes were used on the model, connected by 120 small tubes to 120 manometers. In the supersonic range a "half-model" method was developed. The model was bisected along its major axis and was backed by a flat plate, with sharp leading edges. This enabled the attachment of the large number of tubes to the relatively small models used for supersonic testing. Nine velocities, sixteen angles of attach ( $\alpha = \pm 8^\circ$ ) were covered, each point measured 4 or 5 times, resulting in about 120,000 data points, all processed with great care by hand, since printouts or automatic calculators were not available. By integrating the measured pressure distribution along the body, the normal force distribution (per unit length of axis) for each of the above parameters, was obtained, fulfilling the demands of the design engineers.

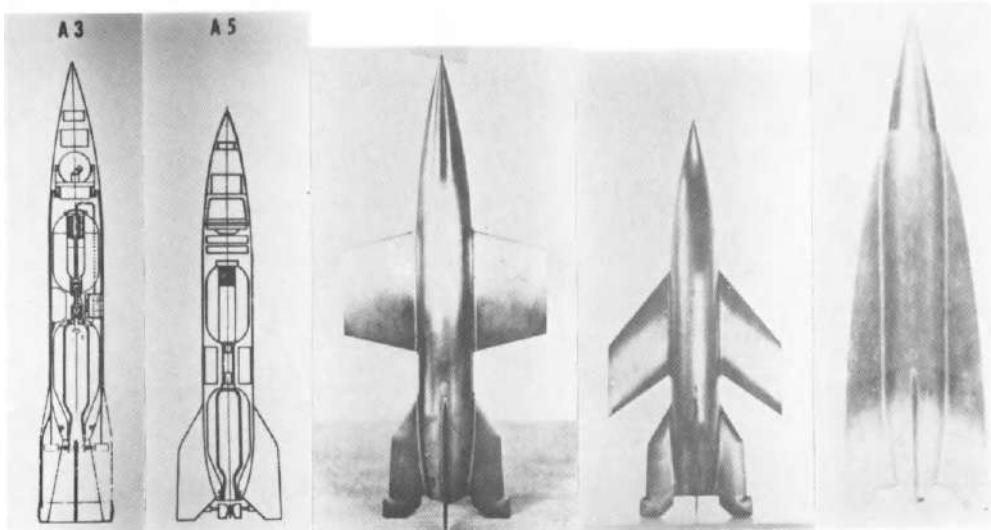


Fig. 9. A3 and A5 (drawings not to scale), A4-V12a, A4-V12c, A4-V13e (photos of wind tunnel models).

### 3. Rocket-Powered Vehicles Group (2).

Drawings and photos of five vehicles of both groups are shown in Fig. 9. Those of the second group have considerable wing area, to produce lift for certain missions, such as long-range gliding or anti-aircraft maneuverability. They are stabilized

either by fin assembly at the very tail (V12a or V12c) or stabilization is taken over by the rear portion of the wing, when a delta wing is used, as in V13e.

"Glider A9" - type vehicles. They had different names at different times for different selling purposes, such as: "Flossengeschoss" (= fin projectile) A4, glider A4, A4b, A9. Development started in 1940. Requirements were stringent: take the body of A4, keep the total weight of propellants and of warhead the same as A4, and attach wings. First wind tunnel tests were carried out in summer 1940 on type A4 V12a, with trapezoidal wings without sweep and on A4 V12c, with  $45^\circ$  swept wings. As an example, V12a had a lift/drag ratio of 4.2 at  $M = 1.86$ . The V12c was 10% better in L/D. Figure 10 shows V12c with two trajectories. Trajectory I has a propellant cut-off angle similar to A4, resulting in a 80 Km ceiling, oscillating trajectory and little gain in range. II has a smaller angle, resulting in 40 Km ceiling with proper gliding trajectory, resulting in maximum range of 450-500 Km. V12c had a good L/D ratio, but a motion of C.P. between subsonic and supersonic range of more than one diameter, almost prohibitive. Also, at certain angles of attack, the stability and controllability with the rudders were weak since the fins were in a downwash of the wings. Hence, other glider-types were investigated. A4 V13e was a delta wing, with curved leading edges, showing L/D ratio 20% smaller than for V12c. A delta wing, broken into three steps, A4 V14f ("Stufen-Gleiter") produced lift on every step, with the third step (behind the C.G.) producing a stabilizing aerodynamic moment. The L/D ratio was a little less than for V12c but the travel of C.P. with Mach No. was much less - an important characteristic.

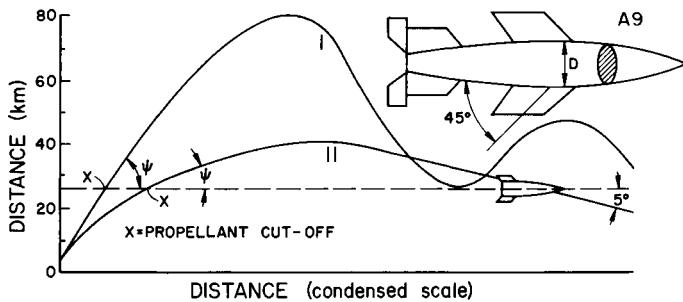


Fig. 10. "Glider A9" with two calculated trajectories.

In order to develop aerodynamic rudders with maximum control effect (maximum lift on the rudder), with minimum hinge moments, pressure distribution measurements along the rudder were conducted with great success, both for the A9-glider and for "Wasserfall."

"Wasserfall" - a short summary. The "Wasserfall" belongs to Group (2), as a rocket-powered vehicle, with considerable wing area and stabilized by an extra fin assembly; see Fig. 3 for a wind tunnel model. "Wasserfall" flies aerodynamically stable with subsonic, transonic and supersonic speeds, developed up to  $M = 3$ . As an anti-aircraft missile, and in order to maneuver in sharp curves in all directions and from all positions, four wings were necessary and large air rudders were placed at the tail. In the subsonic region about  $15^\circ$  angle of attack was necessary; in supersonic region about  $8^\circ$ . Many models with various wings, fin assemblies and rudder forms were tested. The first type used the wing cross twisted  $45^\circ$  against the fin assembly cross to avoid that the fins and rudders would be in the trailing vortex street of the wings. This is correct for small angle of attack, but not for the large ones needed. One type used a concentric ring as a wing, being aerodynamically identical for any angle of attack. Final type was the wing cross and fin cross parallel to each other. The greatest achievement was that the  $\Delta = \text{C.G.} - \text{C.P.}$  for

the total Mach No. range from 0 to 3.0 was almost constant, about  $\Delta = 0.2D$  for angle of attack from  $2^\circ$  to  $8^\circ$ .

Mission requirements in trajectory and flight mechanics, aerodynamic qualities to fulfill those requirements, rudder effectiveness to produce the moments, hinge moments aerodynamically needed, but limited by the servosystem and power available - all relations are so complex, that only a detailed treatment can do justice. Here, this summary must suffice. Important is the fact that after 16 months of intensive work at the wind tunnels first in Peenemunde, then in Kochel, all requirements were fulfilled. (First report was written April 1943; major work was completed in August 1944; some additional details on rudders thru December 1944.)

#### 4. Influence of the Exhaust Jet on the Aerodynamics of the Vehicles.

The influence on the aerodynamics is complicated, different in subsonic and supersonic. It requires theoretical understanding of the jet expansion from the pressure simulating the rocket chamber combustion pressure to the static pressure of the surrounding air at a given flight Mach No. Those expansions are theoretically known for the 2-dimensional case, and expansion into air at rest. Here, the case is 2-dimensional only at the corner, then it is axisymmetric. The expanding jet flow is hit by the outside flow, producing shock waves as on solid cones. An equilibrium is established between the expanding cone flow and the pressure behind the shock wave. Finding out the correct simulation parameters between free flight and wind tunnel test is complex. The experimental difficulties of simulation are so large that only "cold" combustion chamber air was used, not hot combustion gases. Even then, the challenge of measurement of drag and stability was carried out. The mounting of the models was from the top, near the center of gravity, using a hollow pipe, for feeding the simulated exhaust jet. The model was freely movable around the pipe as vertical axis. This arrangement was basically the same for subsonic and supersonic testing, except for the large size of the tunnel and model in subsonic testing and the small size in transonic and supersonic testing.

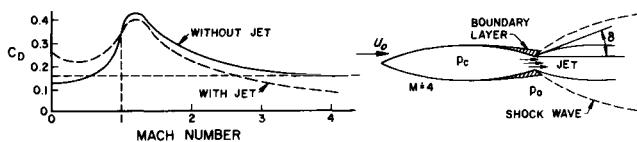


Fig. 11. Drag coefficient of A4 and jet interaction.

Effect on drag. The subsonic region was covered from  $M = 0.13$  thru  $M = 0.85$ . At  $M = 0.13$ , the A5 was used (which is nearly geometrically identical to A4). The exhaust jet had  $M = 1.5$ . Pressure distribution with small holes, attached to manometers, along six meridians were obtained. Without exhaust jet there is overpressure on the nose, some negative pressure along the cylindrical portion and slight positive pressure at the converging tail portion - the first and third contribute to the drag. With exhaust jet the surrounding air is induced like a jet pump, resulting in strong negative pressure at the tail. This results in a drag increase at subsonic speeds, see Fig. 11, which is  $\Delta C_D = + 0.13$ , nearly 100% at  $M = 0.13$ , but nearly disappears at transonic speeds. At supersonic speeds the effect is opposite, and the drag decreased about  $\Delta C_D = - 0.05$  for higher Mach numbers. Note that in this region the drag is decreased nearly 50%. The reason is an interaction of the shock wave, produced by the expanding exhaust jet, with the boundary layer at the convergent tail portion. This results in a pressure increase on that portion, equivalent to a decrease of  $C_D$ .

Effect on stability. The evaluation of the pressure distribution mentioned above showed that the C.P. was displaced about  $0.5D$  towards the tail with the exhaust jet

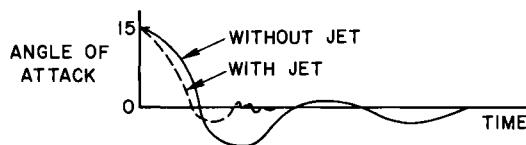


Fig. 12. Stability experiments on A5.

on, causing the A5 (or A4) to become more stable. This is proven by the experiment, permitting free oscillations of the model around its C.G., after elongation to a certain initial angle, here  $15^\circ$ , see Fig. 12. The faster oscillation means larger aerodynamic moments, also larger stability. The larger model velocities are causing faster damping which is also very welcome. Oscillation experiments were also carried out in the supersonic range for the A4, the glider A4 V12c, and the "Wasserfall."

#### ACKNOWLEDGEMENT AND CONCLUSION

Two distinguished persons have been mentioned in the beginning. I avoided mentioning others, since the list would be very long. However, my deepest gratitude goes to Dr. Walter Dornberger, later Major General. From 1937 thru 1945, in all his positions as Chief of a Department of the Army Ordnance, as Commander at Peenemunde, and later as Chief of all rocket development, he was always my direct superior. He was a profound engineer, a man of technical vision, of energy and leadership. This work could never have been accomplished without his constant personal interest, his guidance, sharp criticism, encouragement and constant support. He continued as a friend over the years.

#### REFERENCES

Published reports by the author from 1936 through 1945, several of summary nature. About 200 reports by members of the Wind Tunnel Group 66 are recorded as Peenemunde or Kochel Archive Reports numbered 66/1 (Jan. 1938) through 66/205 (Sept. 1945). Copies of all are in the possession of the author.

## OPTIMAL TRAJECTORIES IN ATMOSPHERIC FLIGHT

N. X. Vinh

*Department of Aerospace Engineering, The University of Michigan,  
Ann Arbor, Michigan 48109, USA*

### ABSTRACT

This paper reviews the main contributions in the general optimization problem in atmospheric flight. The problem is formulated in a general vector form from which the characteristics of the optimal lift, bank and thrust control and the integrals of the motion are established. Using canonical transformation, the appropriate equations for flight in a uniform gravitational field using Cartesian coordinates and in a central Newtonian force field using spherical coordinates are deduced. Performance assessment for aircraft in supersonic flight is reviewed. For planar reentry trajectories, the key equations in the optimization process are proposed. The problem of three-dimensional glide is solved in a reduced-order formulation using the equilibrium glide condition.

### KEYWORDS

Trajectory optimization ; optimal control ; atmospheric flight ; supersonic flight ; reentry trajectories .

### INTRODUCTION

Many problems in science and technology require choosing the best, or the optimal, solution among all the possible solutions. In this second half of our present century, one of the most challenging and fascinating optimization problems is the analysis of the optimal space trajectories. It consists of finding the best trajectory, in some sense, for the motion of a vehicle in a three-dimensional space (Fig. 1). For flight in a vacuum, in which the vehicle is subject only to a gravitational force,  $m \vec{g}$ , either from one or from many sources, and its self-generated propulsive force,  $\vec{T}$ , the theory is very completed and is masterfully presented in a monograph on "Optimal Space Trajectories" by Marec (1979). A space mission, starting from the surface of the earth, naturally begins, and frequently ends, with an atmospheric phase in which an aerodynamic force,  $\vec{A}$ , is added to the total force acting on the vehicle. For a low altitude mission, atmospheric flight covers the entire trajectory. This aerodynamic force depends on the speed, the location in space, the shape and size of the vehicle and the control inputs. These controls, in the form of the angle-of-attack and the bank angle, activated through a

system of aerodynamic devices , together with the thrust control , in an optimal manner , must be selected to bring the vehicle from the initial state , at time  $t_0$  , position vector  $\vec{r}_0$  , velocity  $\vec{V}_0$  and mass  $m_0$  , to the final state  $\vec{r}_f$  ,  $\vec{V}_f$  and  $m_f$  at the final time  $t_f$  such that a certain function of the final state , called performance index , is maximized . An interesting aspect of aerospace maneuver is that the performance index may have a wide variety . Besides the maximization of the final mass for a minimum fuel consumption problem , one may consider the distance achieved for the case of a long-range supersonic transport , the landing coverage of a space shuttle on its return to the earth , or the minimization of the time of flight such as in problems associated with the climb and turn performances of fighter aircraft .

Although the published literature on optimization problems in atmospheric flight is immense , only a few syntheses of the subject exist in the English language . Through the NASA translation program there are the book on "New Variational Methods in Flight Dynamics" (Krotov et al. , 1971) which considers special modes of flight control and the treatise on "Mechanics of Optimum Three-Dimensional Motion of Aircraft in the Atmosphere" (Shkadov et al. , 1975) which deals with reentry trajectories . On the other hand , the excellent book on "Flight Mechanics" by Miele (1962 a) is restricted to performance analysis of aircraft through parametric optimization .

As atmosphere bounded aircraft are extending their operational space beyond the stratosphere and , with the recent maiden flight of the space shuttle Columbia the smooth transition from space flight to atmospheric flight of a winged vehicle has been established , it is time to have a fresh look at the problem of aerospace flight in its entity and consider it as concerning a general dynamical process in order to establish the general optimal characteristics in aerodynamic and thrust controls (Vinh , 1981) . From these general results , a particular problem , when solved , depends on the scale in the physical space considered , and the particular optimal trajectory obtained depends on the end -conditions and the physical constraints to be enforced .

It is customary , when one writes a survey paper , to add a long list of references . However , because of the limitation in the length of this article , we are compelled to assemble a restrictive bibliography . For a more complete listing of the original contributions , we refer the readers to consult the monographs mentioned above .

#### GENERAL FORMULATION

The motion of a vehicle , considered as a point mass with varying mass , flying in a general gravitational field and subject to aerodynamic force and thrusting force , is governed by the equations written with respect to an inertial reference frame (Fig. 1)

$$\frac{d\vec{r}}{dt} = \vec{V} , \quad \frac{d\vec{V}}{dt} = \frac{1}{m} (\vec{T} + \vec{A}) + \vec{g}(\vec{r}, t) , \quad \frac{dm}{dt} = -\frac{c}{g_0} T \quad (1)$$

Besides the usual notation , the parameter  $c$  is the specific fuel consumption which has the dimension of the inverse of the time with the constant  $g_0$  being the gravitational acceleration at a reference level . For a symmetric flight , the aerodynamic force  $A$  is in the plane of symmetry of the vehicle and is conveniently decomposed into a drag force  $D$  opposite to the velocity vector and a

lift force  $\vec{L}$  perpendicular to it. Their magnitudes are given by

$$D = \frac{1}{2} \rho S V^2 C_D, \quad L = \frac{1}{2} \rho S V^2 C_L \quad (2)$$

We assume that the atmospheric density  $\rho$  is a function of the altitude  $h$  and there exists a relation between the drag coefficient  $C_D$  and the lift coefficient  $C_L$  of the form

$$C_D = C_{D_0} + K C_L^2 \quad (3)$$

where  $C_{D_0}$  and  $K$  are functions of the Mach number. In the problem of reentry, at hypersonic speed,  $C_{D_0}$  and  $K$  are constant. On the other hand, in the case where the range in the altitude is small, or more rigorously, in the case of an isothermal atmosphere, we can consider a constant speed of sound and  $C_{D_0}$  and  $K$  are functions of the speed. In high speed powered flight, the angle between the thrust vector and the velocity is small and for all practical purposes we assume the collinearity

$$\vec{T} = T (\vec{V}/V) \quad (4)$$

and use the thrust magnitude as control subject to the constraint

$$0 \leq T \leq T_{\max}(V, h) \quad (5)$$

By Eq. (3), the lift coefficient can be used as control for the magnitude of the aerodynamic force. It is bounded by an upper limit

$$C_L \leq C_{L \max}(V) \quad (6)$$

If physical constraints are added to the problem, the upper limit of  $C_L$  may be function of other state variables. We define the orientation of the aerodynamic force  $\vec{A}$  by the angle  $\sigma$ , called the bank angle. This is the angle between the plane of symmetry of the vehicle and the  $(\vec{r}, \vec{V})$  plane. We assume that

$$|\sigma| \leq \sigma_{\max} \quad (7)$$

where  $\sigma_{\max}$  is function of the state variables as dictated by certain physical constraints.

### Optimal Control

Using the maximum principle, we introduce the adjoint vectors  $(\vec{p}_r, \vec{p}_V, p_m)$  to form the Hamiltonian

$$H = \vec{p}_r \cdot \vec{V} + \vec{p}_V \cdot \left( \frac{\vec{A}}{m} + \vec{g} \right) + \frac{T}{m} \left( \frac{\vec{p}_V \cdot \vec{V}}{V} - \frac{mp_m c}{g_o} \right) \quad (8)$$

To maximize the Hamiltonian with respect to the thrust magnitude  $T$ , we consider the coefficient of  $(T/m)$  called the switching function

$$\Phi = \frac{\vec{p}_V \cdot \vec{V}}{V} - \frac{mp_m c}{g_o} \quad (9)$$

This leads to the following law for the thrust magnitude control

$$\begin{aligned} \Phi &> 0, & T &= T_{\max} \\ \Phi &< 0, & T &= 0 \\ \Phi &\equiv 0, & t \in [t_1, t_2], & T = \text{variable} \end{aligned} \quad (10)$$

The optimal trajectory is a combination of boost arc (B), at  $T = T_{\max}$ , coast arc (C), at  $T = 0$ , and sustaining arc (S), with an intermediate thrust,  $T = \text{variable}$ . Concerning the aerodynamic control, we write

$$\vec{A} = (1/2) \rho S V^2 \vec{a} \quad (11)$$

and consider the Contensou's domain of maneuverability (Contensou, 1969), as described by the control vector  $\vec{a}$  (Fig. 2). It is a paraboloid of revolution about  $\vec{V}$  as defined by the parabolic drag polar (3), bounded by the constraints (6) and (7). Then, the maximization of the Hamiltonian  $H$  leads to the maximization of the dot product  $\vec{p}_V \cdot \vec{A}$ , or  $\vec{p}_V \cdot \vec{a}$ , where the adjoint vector  $\vec{p}_V$ , associate to the velocity vector  $\vec{V}$  is called the primer vector. The optimal bank angle is such that either  $\sigma = \sigma_{\max}$  or that the three vectors  $\vec{p}_V$ ,  $\vec{A}$  and  $\vec{V}$  are coplanar, that is (Fig. 3)

$$(\vec{V} \times \vec{p}_V) \cdot \vec{A} = 0 \quad (12)$$

In the plane of these vectors, the tangent to the drag polar is orthogonal to  $\vec{p}_V$  (Fig. 4). Hence, if  $\epsilon$  is the angle between  $\vec{V}$  and  $\vec{p}_V$ , we have

$$\tan \epsilon = \partial C_D / \partial C_L \quad (13)$$

an optimal relation discovered by Miele (1962 b) for flight in a constant gravitational field. This relation is valid whenever  $C_L \leq C_{L\max}$ . Otherwise, the lift coefficient is on the boundary.

#### Integrals of the Motion.

If the gravitational field is time invariant, the Hamiltonian is constant over the whole trajectory

$$H = C_0 \quad (14)$$

Consider an arc  $M_1 M_2$  between the time  $t_1$  and  $t_2$  of an optimal trajectory. We have the transversality condition (Marec, 1979)

$$[\vec{p} \cdot \delta \vec{x}]_1^2 = [\vec{p}_r \cdot \delta \vec{r} + \vec{p}_V \cdot \delta \vec{V}]_1^2 = 0 \quad (15)$$

for the variation  $\delta$  satisfying the kinematic constraints. If the problem has a rotational symmetry about an axis with unit vector  $\vec{k}$ , a rotation  $\delta \omega = \vec{k} \cdot \delta \omega$  conserves the same performance index. Since  $\delta \vec{r}_i = (\vec{k} \times \vec{r}_i) \delta \omega$ ,  $\delta \vec{V}_i = (\vec{k} \times \vec{V}_i) \delta \omega$ ,  $i = 1, 2$ , we have

$$[\vec{p}_r \cdot (\vec{k} \times \vec{r}) \delta \omega + \vec{p}_V \cdot (\vec{k} \times \vec{V}) \delta \omega]_1^2 = [\vec{r} \times \vec{p}_r + \vec{V} \times \vec{p}_V]_1^2 \cdot \vec{k} \cdot \delta \omega = 0 \quad (16)$$

Let

$$\vec{r} \times \vec{p}_r + \vec{V} \times \vec{p}_V = \vec{C} \quad (17)$$

For a problem with spherical symmetry, e.g. flight about a spherical earth,  $\vec{k}$  is arbitrary and we have  $\vec{C}_1 = \vec{C}_2$ . Since  $t_1$  and  $t_2$  are any two instants of the interval  $[t_1, t_2]$ , the vector  $\vec{C}$  is constant over the whole trajectory and it leads to three scalar constants of integration. For the case of a flat earth,  $\vec{k}$  is along the vertical. Condition (16) implies that the vector  $\vec{C}_2 - \vec{C}_1$  is parallel to the horizontal plane and hence the Z-component of the vector  $\vec{C}$  is constant. But then, because of rotational symmetry about the Z-axis, the horizontal coordinates are ignorable and both the adjoints  $p_x$  and  $p_y$  are constant. Besides these four integrals, valid along the whole trajectory, along a sustaining arc,

we have two additional relations  $\Phi = 0$  and  $d\Phi/dt = 0$  valid for the duration  $[t_1, t_2]$  of the variable thrust. This intermediate thrust appears linearly in the equation  $d^2\Phi/dt^2 = 0$ .

#### Canonical Transformation.

The equations of motion and the optimal control laws for any coordinate system used can be deduced from the general vector equations by using appropriate canonical transformation. It is best to apply the transformation to the Hamiltonian (8) and from it derive the resulting variational equations. For example, in the case of a flat earth, consider the vectors

$$\vec{r} = (X, Y, Z), \quad \vec{V} = (u, v, w), \quad \vec{A} = (A_X, A_Y, A_Z), \quad \vec{g} = (0, 0, -g) \quad (18)$$

in the original inertial Cartesian system and perform the velocity transformation (Fig. 5)

$$u = V \cos \gamma \cos \psi, \quad v = V \cos \gamma \sin \psi, \quad w = V \sin \gamma \quad (19)$$

For the transformation to be canonical while conserving the same Hamiltonian, it is necessary and sufficient that

$$p_u du + p_v dv + p_w dw = p_V dV + p_\psi d\psi + p_\gamma d\gamma \quad (20)$$

Hence, we have for example

$$p_u = \frac{\partial V}{\partial u} p_V + \frac{\partial \psi}{\partial u} p_\psi + \frac{\partial \gamma}{\partial u} p_\gamma \quad (21)$$

and similar relations for  $p_v$  and  $p_w$ . But the transformation (19) is characterized by the orthogonality condition such as  $\sum (\partial u / \partial V)(\partial u / \partial \psi) = 0$ , and for such system we have

$$\frac{\partial V}{\partial u} = \frac{1}{H_1} \frac{\partial u}{\partial V}, \quad \frac{\partial \psi}{\partial u} = \frac{1}{H_2} \frac{\partial u}{\partial \psi}, \quad \frac{\partial \gamma}{\partial u} = \frac{1}{H_3} \frac{\partial u}{\partial \gamma} \quad (22)$$

where the  $H_i$  are the Lame' coefficients

$$H_1 = \sum \left( \frac{\partial u}{\partial V} \right)^2 = 1, \quad H_2 = \sum \left( \frac{\partial u}{\partial \psi} \right)^2 = V^2 \cos^2 \gamma, \quad H_3 = \sum \left( \frac{\partial u}{\partial \gamma} \right)^2 = V^2 \quad (23)$$

with the summations carried over  $u$ ,  $v$ , and  $w$ . Hence, in vector form

$$\begin{bmatrix} p_u \\ p_v \\ p_w \end{bmatrix} = S \begin{bmatrix} p_V \\ p_\psi/V \cos \gamma \\ p_\gamma/V \end{bmatrix}, \quad S = \begin{bmatrix} \cos \gamma \cos \psi & -\sin \psi & -\sin \gamma \cos \psi \\ \cos \gamma \sin \psi & \cos \psi & -\sin \gamma \sin \psi \\ \sin \gamma & 0 & \cos \gamma \end{bmatrix} \quad (24)$$

We notice that the matrix  $S$  is the rotation matrix that brings the  $X$ -axis to coincide with the velocity vector  $\vec{V}$  through a rotation  $\psi$  in the horizontal plane followed by a rotation  $\gamma$  in the vertical plane. In this new axis system, if the bank angle  $\sigma$  is defined as the angle between the vertical plane containing the velocity and the plane of symmetry of the vehicle, the components of the aerodynamic force are  $-D$ ,  $L \sin \sigma$  and  $L \cos \sigma$ , respectively. Therefore, we have for the vector  $\vec{A}$

$$(A_X, A_Y, A_Z)^T = S (-D, L \sin \sigma, L \cos \sigma)^T \quad (25)$$

where superscript T denotes a transposition. On the other hand, for the

velocity vector  $\vec{V}$  we have

$$(u, v, w)^T = S(V, 0, 0)^T \quad (26)$$

Hence, from the last three equations, by writing the optimal condition (12) as  $\det[\vec{V}, \vec{p}_V, \vec{A}] = 0$ , we are led to the control law for the bank angle

$$(\det S) L \left( \frac{p_\psi \cos \sigma}{\cos \gamma} - p_Y \sin \sigma \right) = 0 \quad (27)$$

The determinant is the Jacobian of the velocity transformation and is unity. Hence, the optimal bank angle is given by

$$\tan \sigma = p_\psi / p_Y \cos \gamma \quad (28)$$

For the lift control, by using the fact that if  $(\vec{x}, \vec{y})$  is the dot product of two vectors  $\vec{x}$  and  $\vec{y}$ , then  $(A\vec{x}, B\vec{y}) = (\vec{x}, A^T B\vec{y})$ , we have the following relations

$$\begin{aligned} |\vec{p}_V|^2 &= (\vec{p}_V, \vec{p}_V) = p_V^2 + (p_\psi / V \cos \gamma)^2 + (p_Y / V)^2 \\ \vec{p}_V \cdot \vec{V} &= V p_V = |\vec{p}_V| V \cos \epsilon \end{aligned} \quad (29)$$

Then, using the optimal condition (28)

$$\tan^2 \epsilon = \frac{1}{\cos^2 \epsilon} - 1 = \frac{1}{V^2 p_V^2} \left( \frac{p_\psi^2}{\cos^2 \gamma} + p_Y^2 \right) = \left( \frac{p_Y}{V p_V \cos \sigma} \right)^2 \quad (30)$$

and the general law for the lift control (13) becomes

$$p_Y / V p_V \cos \sigma = \partial C_D / \partial C_L = 2 K C_L \quad (31)$$

with the last equality applied to a parabolic drag polar. Since all vectors, state and adjoint, have been obtained explicitly in terms of the new variables, by forming the different dot products in the Hamiltonian (8) we have the new generating Hamiltonian in the case of a flat earth

$$\begin{aligned} H &= (p_X \cos \psi + p_Y \sin \psi) V \cos \gamma + p_Z V \sin \gamma + p_V \left( -\frac{D}{m} - g \sin \gamma \right) \\ &+ \frac{p_Y}{V} \left( \frac{L \cos \sigma}{m} - g \cos \gamma \right) + \frac{p_\psi L \sin \sigma}{m V \cos \gamma} + \frac{T}{m} \left( p_V - \frac{mp_m c}{g_o} \right) \end{aligned} \quad (32)$$

The following integrals are immediately obtained

$$H = C_0, p_X = C_1, p_Y = C_2, p_\psi = C_1 Y - C_2 X + C_3 \quad (33)$$

The last integral can also be obtained by writing that the Z-component of the vector  $\vec{C}$  is constant.

In the case of a spherical earth, the choice coordinates are the spherical coordinates as shown in Fig. 6. Canonical transformation leads to the same control laws for the lift and bank controls and the following Hamiltonian (Vinh, 1981)

$$\begin{aligned} H &= p_r V \sin \gamma + p_\theta \frac{V \cos \gamma \cos \psi}{r \cos \phi} + p_\phi \frac{V}{r} \cos \gamma \sin \psi + p_V \left( -\frac{D}{m} - g \sin \gamma \right) \\ &+ p_Y \left( \frac{L \cos \sigma}{m V} - \frac{g}{V} \cos \gamma + \frac{V}{r} \cos \gamma \right) + p_\psi \left( \frac{L \sin \sigma}{m V \cos \gamma} \right. \\ &\left. - \frac{V}{r} \cos \gamma \tan \phi \cos \psi \right) + \frac{T}{m} \left( p_V - \frac{mp_m c}{g_o} \right) \end{aligned} \quad (34)$$

In this case, besides the Hamiltonian integral we can deduce from the vector integral (17) the following scalar integrals

$$\begin{aligned} p_\theta &= C_1 \\ p_\phi &= C_2 \sin \theta - C_3 \cos \theta \\ p_\psi &= C_1 \sin \phi + (C_2 \cos \theta + C_3 \sin \theta) \cos \phi \end{aligned} \quad (35)$$

These integrals have been obtained using various schemes in the literature (Moyer, 1973, Shkadov et al., 1975). Finally, we notice that the switching function has the form  $\Phi = p_V - (m p_m c / g_o)$ . In particular, in gliding flight or in short duration flight where the small variation of the mass is neglected,  $p_m = 0$ , and we have  $\Phi = p_V$ .

### SUPERSONIC FLIGHT

For the analysis of aircraft performance, the equations of motion are generated by the Hamiltonian (32). We have, using  $h$  as the altitude variable

$$\frac{dX}{dt} = V \cos \gamma \cos \psi, \quad \frac{dY}{dt} = V \cos \gamma \sin \psi, \quad \frac{dh}{dt} = V \sin \gamma \quad (36)$$

$$\frac{dV}{dt} = \frac{T - D}{m} - g \sin \gamma, \quad \frac{d\gamma}{dt} = \frac{L \cos \sigma}{m V} - \frac{g}{V} \cos \gamma, \quad \frac{d\psi}{dt} = \frac{L \sin \sigma}{m V \cos \gamma}, \quad \frac{dm}{dt} = - \frac{c}{g_o} T$$

Since the aerodynamic and engine characteristics are function of the Mach number it is convenient to rewrite the equations using the following set of dimensionless variables with  $W = mg$

$$\begin{aligned} x &= gX/a^2, \quad y = gY/a^2, \quad z = gh/a^2, \quad M = V/a \\ \omega &= 2W/kp_*S, \quad \theta = gt/a, \quad K_c = ac/g, \quad \eta = 2T/kp_*S \end{aligned} \quad (37)$$

where  $k$  is the ratio of the specific heats and  $p_*$  denotes the pressure at a reference level taken at the tropopause, and furthermore either an isothermal atmosphere or an average value for the speed of sound is used although this assumption can be easily removed. Then, using the relations

$$a^2 = k p / \rho, \quad \delta = p / p_* = \delta(z) \quad (38)$$

we have the equations of motion

$$\begin{aligned} \frac{dx}{d\theta} &= M \cos \gamma \cos \psi, \quad \frac{dy}{d\theta} = M \cos \gamma \sin \psi, \quad \frac{dz}{d\theta} = M \sin \gamma \\ \frac{dM}{d\theta} &= \frac{\eta}{\omega} - \frac{M^2 \delta (C_{D0} + K C_L^2)}{\omega} - \sin \gamma, \quad \frac{d\gamma}{d\theta} = \frac{\delta M C_L \cos \sigma}{\omega} - \frac{\cos \gamma}{M} \\ \frac{d\psi}{d\theta} &= \frac{\delta M C_L \sin \sigma}{\omega \cos \gamma}, \quad \frac{d\omega}{d\theta} = - K_c \eta \end{aligned} \quad (39)$$

It can be shown that for a standard atmosphere, the pressure ratio  $\delta = p/p_* = (\rho/\rho_*)^b$  with the value of the exponent being  $b = 1.235$  in the troposphere and  $b = 1$  in the stratosphere. Hence  $\delta$  is a function, say exponential, of the dimensionless altitude  $z$ . For a smooth application of the variational method,

from aerodynamics and engine data , we can model the characteristics  $C_{D_0}(M)$  ,  $K(M)$  ,  $C_{L_{max}}(M)$  ,  $K_c(M)$  as functions of the Mach number and the dimensionless maximum thrust  $\eta_{max}(M, z)$  as function of the Mach number and the altitude . The equations contain the least number of physical parameters describing real physical phenomena . They can be further simplified such as in the cases :

a / For constant altitude flight , besides the simplification  $\gamma = 0$  , since the ambient pressure  $p$  is constant , we can replace  $p_*$  by  $p$  in the definition (37) of  $\omega$  and  $\eta$  , and in Eqs. (39) we simply have  $\delta = 1$  .

b / In coasting flight or in problems of minimum time maneuver with short duration , we can consider a constant weight . In this case in system (39) , the equation for the mass flow can be deleted . Furthermore , in the remaining equations we can use the new wing loading and the new thrust ratio

$$w = \omega / \delta = 2W/kpS , \quad \tau = \eta / \omega = T/W \quad (40)$$

The variable  $w$  , through the ambient pressure  $p$  is now the altitude variable . In this case , for an isothermal atmosphere ,  $dp/p = d\rho/\rho = -\beta dh$  , and the equation for the altitude is replaced by the equation

$$\frac{dw}{d\theta} = e w M \sin \gamma , \quad e = \frac{a^2 \beta}{g} \quad (41)$$

where  $e$  is a constant specifying the atmosphere .

#### Maximum Range Cruise

At constant altitude , to achieve a maximum range for a given amount of fuel , the optimal trajectory consists of two short arcs at the end-points using  $\eta = \eta_{max}$  or  $\eta = 0$  to adjust the Mach number from the prescribed end-points Mach numbers to the best Mach number for cruise . Nearly the whole trajectory is effected at intermediate thrust such that the following condition is satisfied

$$\frac{2W}{kpS} = w = M^2 \sqrt{\frac{C_{D_0}}{K}} \sqrt{\frac{1 + (C_{D_0} K_c) M + MK_c}{3 - (K K_c) M + MK_c}} \quad (42)$$

In this equation , subscript  $M$  denotes the logarithmic derivative that is  $y_M = d(\log y)/d(\log M) = (M/y)(dy/dM)$  . This requires an accurate modeling of the functions  $C_{D_0}(M)$  ,  $K(M)$  ,  $K_c(M)$  and their derivatives . The optimal cruise Mach number decreases slowly along the flight path , and it is a function of the aerodynamics and engine characteristics and the altitude . Since the optimal Mach number is nearly constant , steady-state cruise can be used as a suboptimal solution to the problem . There exists a best altitude for maximum range and an average optimal Mach number for cruise at that altitude . Using steady-state approximation , the optimal Mach number is obtained by maximizing the function

$$G = M^{1/b} E^*(M) / c(M) C_L^*(M)^{(b-1)/2b} \quad (43)$$

where  $E^* = 1/2 \sqrt{KC_{D_0}}$  is the maximum lift-to-drag ratio with the corresponding lift coefficient  $C_L^* = \sqrt{C_{D_0}/K}$  . Hence , there is a possibility of stratospheric flight ,  $b = 1$  , and tropospheric flight  $b = 1.235$  . Once the cruise Mach number has been established , the best cruise altitude is obtained from

$$2 \sqrt{W_o W_f} / kpS = M^2 C_L^* / d \quad (44)$$

with  $d = 1$  for stratospheric flight and  $d = 1.1$  for tropospheric flight.

#### Turning Maneuver in a Horizontal Plane

At constant altitude,  $\gamma = 0$  and using  $w$  for nearly constant weight

$$\frac{1}{\cos \sigma} = M^2 C_L / w = L/W = n \quad (45)$$

The bank angle and the lift coefficient are related and only one can be used as independent aerodynamic control. Sometimes it is more appropriate to use the load factor  $n = L/W$  as control. It is subject to the constraint

$$1 \leq n \leq \text{inf.} [n_s, M^2 C_{L \max} / w] \quad (46)$$

where  $n$  is a physiological-structural limit in the range 2-6. The general system (39) is reduced to

$$\frac{dx}{d\theta} = M \cos \psi, \quad \frac{dy}{d\theta} = M \sin \psi, \quad \frac{d\psi}{d\theta} = \epsilon \sqrt{\frac{n^2 - 1}{M}} \quad (47)$$

$$\frac{dM}{d\theta} = \zeta \tau_{\max} - \frac{M^2}{w} \left[ C_{D_0}(M) + K(M) \frac{n^2 w^2}{M^4} \right]$$

where  $\epsilon = \pm 1$  and  $\zeta$  is the thrust control parameter varying in the range  $\zeta \in [0, 1]$ . From the constraint (46), in the  $(nw, M)$  space, the domain of flight is bounded by the curves (Fig. 7)

$$nw = M^2 C_{L \max}(M), \quad n = n_s \quad (48)$$

At high altitude, the maximum value of  $M^2 C_{L \max}$ , which is a well determined constant value for any given function  $C_{L \max}$ , is less than  $n_w$  which increases as the altitude increases. Hence, as seen in Fig. 7a, the boundary for turning maneuvers is at  $C_{L \max}$ . However, at low altitude, as seen in Fig. 7b, the maneuver is constrained at  $C_{L \max}$  at low speed and by  $n$  at high speed, with the connection at a Mach number called the corner Mach number,  $M_c$ . The switching function that regulates the thrust control is  $\Phi = p_M^c$ , with  $p_M^c > 0$  for B-arc,  $p_M^c < 0$  for C-arc and  $p_M^c \equiv 0$  for variable thrust arc. The optimal load factor is obtained on the boundary (46) or in the interior of this flight envelope according to

$$\tan \sigma = \epsilon \sqrt{n^2 - 1} = M p_\psi / 2wKp_M \quad (49)$$

This is only possible with B-arc. For C-arc and S-arc, the load factor is on the boundary. With the four integrals of motion, all the adjoint components are known as functions of the state variables and constants of integration. The only difficulty in solving any specific problem consists of finding the sequence of optimal subarcs, B, C and S, and the corresponding load factor and the evaluation of the constants of integration such that the end-conditions are satisfied. The discussion of the switching sequence can be based on the fact that at a junction between subarcs, necessarily on the boundary, if  $dp_M/d\theta < 0$ , the optimal sequence is B → C, and if  $dp_M/d\theta > 0$  the optimal sequence is C → B. There may exist a singular arc at intermediate thrust level in the middle of the turn and in some cases, chattering arc at the end where the bank angle switches rapidly between its bounds.

The basic performance assessment in turning flight can be in terms of minimum

time turn to a heading , to a point or to a line , or minimum turn radius .

### Optimal Maneuvers in a Vertical Plane

With  $\psi = 0$  ,  $\sigma = 0$  and for minimum time maneuver with nearly constant weight we deduce from the general equations (39) together with Eq. (41)

$$\begin{aligned}\frac{dx}{d\theta} &= M \cos \gamma , \quad \frac{dw}{d\theta} = ew M \sin \gamma , \quad \frac{dy}{d\theta} = \frac{MC_L}{w} - \frac{\cos \gamma}{M} \\ \frac{dM}{d\theta} &= \zeta \tau_{\max} - \frac{M^2}{w} (C_{D0} + KC_L^2) - \sin \gamma\end{aligned}\quad (50)$$

The Hamiltonian is

$$\begin{aligned}H &= p_x M \cos \gamma + ew p_w M \sin \gamma + p_M \left[ \zeta \tau_{\max} - \frac{M^2}{w} (C_{D0} + KC_L^2) - \sin \gamma \right] \\ &\quad + p_\gamma \left( \frac{MC_L}{w} - \frac{\cos \gamma}{M} \right)\end{aligned}\quad (51)$$

The two integrals  $H = C_0$  and  $p_\gamma = C_1$  are not sufficient to completely solve the general problem but for all the basic maneuvers , detailed optimal characteristics and performance assessment can be established . We shall consider the following problems :

a / Minimum time to climb . We use the end-conditions

$$\begin{aligned}\theta &= 0 , \quad x = 0 , \quad w = w_o , \quad M = M_o , \quad \gamma = \gamma_o \\ \theta_f &= \text{min.} , \quad x_f = \text{free} , \quad w = w_f , \quad M_f = \text{free} , \quad \gamma_f = \text{free}\end{aligned}\quad (52)$$

Using the Kelley-Contensou condition (Kelley et al. , 1967 , Contensou , 1969) , the singular arc can be ruled out and the climb is performed with  $\tau = \tau_{\max}$  . The optimal lift control is found to be

$$C_L = p_\gamma / 2 K M p_M \quad (53)$$

The maximum thrust can be modeled as

$$\tau_{\max} = T_{\max} / w = k_1 (1 + k_2 M + k_3 M^2) / w \quad (54)$$

with appropriate values for the  $k_i$  selected to match the engine performance as function of the Mach number and the altitude . Then , with  $C_0 = 1$  for minimum time and  $C_1 = 0$  for free range , we deduce from the Hamiltonian the equations for the adjoints

$$\begin{aligned}\frac{d(w p_w)}{d\theta} &= p_M \left[ \zeta \tau_{\max} - \frac{M^2}{w} (C_{D0} - KC_L^2) \right] \\ \frac{d(M p_M)}{d\theta} &= -1 + p_M \left[ \zeta (2 - \tau_{\max} M) \tau_{\max} + \frac{M^2}{w} (C_{D0} C_{D0M} + K K_M C_L^2) \right. \\ &\quad \left. - 4 K C_L \cos \gamma - 2 \sin \gamma \right] \\ \frac{dp_\gamma}{d\theta} &= -ew p_w M \cos \gamma + p_M (\cos \gamma - 2 K C_L \sin \gamma)\end{aligned}\quad (55)$$

where of course  $C_L$  is given by Eq. (53) . An efficient way of computing the trajectory is to use backward integration with the known starting values for the

adjoints,  $p_M(\theta_f) = 0$ ,  $p_\gamma(\theta_f) = 0$  for free final Mach number and flight path angle and from the Hamiltonian integral,  $e(w_{f^*}) = 1/M_f \sin \gamma_f$ . In this way, the parameters for the iteration are the free final values  $M_f$  and  $\gamma_f$  and as state variables, they are easy to estimate. At the end of the integration,  $w = w_f$ , and the prescribed initial values  $M_0$  and  $\gamma_0$  are used to adjust  $M_f$  and  $\gamma_f$ . From Eq. (53), the indetermination of the starting value  $C_L(\theta_f)$  can be removed by applying the 1<sup>st</sup> Hospital rule. This leads to

$$\frac{C_{L_f}}{C_L} / C_L^* = e w p_w M \cos \gamma / 2 K C_L^* = 1/2 K C_L^* \tan \gamma_f = E^* / \tan \gamma_f \quad (56)$$

Since the ratio  $C_{L_f}/C_L$  is small and  $E^*$  is of the order of 10, the exit angle  $\gamma_f$  is large. A typical trajectory starts with a dive to pick up speed, followed by a steep climb. This strategy is to insure that, as the altitude increases, the decreases in  $\tau_{\max}$  due to the effect of the altitude is compensated by the positive effect on maximum thrust due to the high Mach number and also to the high climb rate  $M \sin \gamma$ . This type of zoom climb was described by Kelley (1959) and Bryson and Denham (1962).

b/ Fastest turn in the vertical plane. Three basic aerobatic maneuvers for fighter aircraft in a vertical plane are the loop, Immelman and split-S (Fig. 8). The Immelman is actually the first half of a loop followed by a half roll, and the split-S is the second half of a loop. They are used to effect a  $180^\circ$  rotation of the velocity vector in the vertical plane yielding at the same time a  $180^\circ$  turn for the heading. As seen in the figure, they are characterized by the end-conditions

$$\begin{aligned} \theta = 0, \quad x = 0, \quad w = w_0, \quad M = M_0, \quad \gamma_0 &= \begin{cases} 0^\circ (\text{loop, Immelman}) \\ 180^\circ (\text{split-S}) \end{cases} \\ \theta_f = \min., \quad x_f = \text{free}, \quad w = w_f, \quad M_f = \text{free}, \quad \gamma_f &= \begin{cases} 360^\circ (\text{loop, split-S}) \\ 180^\circ (\text{Immelman}) \end{cases} \end{aligned} \quad (57)$$

Turning is a hard maneuver and the load factor  $n = L/W$  is the choice aerodynamic control as bounded by the condition (46). In this case, the state equations (50) have  $C_L$  replaced by  $n w/M$  and consequently, from Eq. (53), we have the interior load factor control

$$n = M p_\gamma / 2 K w p_M \quad (58)$$

If this control exceeds the bound (46), boundary control must be used. In particular, C-arc and S-arc are necessarily be flown with boundary load factor. Intermediate thrust can be ruled out by the Kelley-Contensou condition and the optimal trajectory is a combination of C-arc with boundary load factor and B-arc with the load factor either in the interior or on the boundary. To solve this problem, a backward integration will remove all the ambiguities in the selection of the optimal thrust and load factor controls. First, the equations for the adjoints are derived for the different cases of load factor control, and in particular for  $p_M$ , which is here the switching function, we have

$$\begin{aligned} M \frac{dp_M}{d\theta} &= -1 + p_M \left[ \zeta \tau_{\max}^{(1-\tau_{\max})} M + \frac{M^2 C_{D0}}{w} (1 + C_{D0} M) \right] \\ &\quad - K \frac{n^2 w}{M^2} (3 - K_M) - \sin \gamma \left[ + 2 p_\gamma \frac{(n - \cos \gamma)}{M} \right] \end{aligned} \quad (59)$$

for the case of interior  $n$  or when  $n = n_s$ , and

$$\begin{aligned} M \frac{dp_M}{d\theta} &= -1 + p_M \left[ \zeta \tau_{max}^{(1-\tau_{max})} + \frac{M^2 C_{Do}}{w} (1 + C_{DoM}) \right] \\ &\quad + \frac{M^2}{w} K C_{Lmax}^2 (1 + K_M + 2C_{Lmax}) - \sin \gamma \left[ -p_\gamma \left[ \frac{M}{w} C_{Lmax} C_{Lmax} + \frac{2 \cos \gamma}{M} \right] \right] \end{aligned} \quad (60)$$

for the case where  $n$  is on the boundary  $C_{Lmax}$ . The equations for  $p_w$  and  $p_\gamma$  are similarly derived from the maximized Hamiltonian. Then, to start the backward integration, for free  $M_f$ ,  $p_M(\theta_f) = 0$ , and the load factor is on the boundary. With a guessed starting  $M$  and a guessed or prescribed  $w_f$ , together with the prescribed  $n_s$ , condition (46) decides the type of boundary load factor. For the starting thrust control, it is dictated by the behavior of the switching function at the final time. Since  $C_1 = 0$  for free range, and  $\sin \gamma_f = 0$  in all cases, from the Hamiltonian integral we have the starting value for  $p_\gamma$  at the final time

$$p_\gamma(n - \cos \gamma_f) = M \quad (61)$$

The starting value for  $p_w$  is either  $p_w(\theta_f) = 0$  for free final altitude or a guessed value  $p_w$  in the case of prescribed final altitude. Then, if  $n = n_s$  on the final arc, from Eq. (59), at the final time,  $dp_M/d\theta = 1/M_f > 0$  and the final arc is a C-arc since  $p_M$  is increasing to the final value zero. On the other hand, if the condition (46) leads to the selection of  $C_{Lmax}$  on the final arc, by using Eq. (61) in Eq. (60) at the final time

$$\frac{dp_M}{d\theta} = -p_\gamma \left[ \frac{C_{Lmax}}{w} (1 + C_{Lmax}) + \frac{\cos \gamma_f}{M^2} \right] \quad (62)$$

Since the maneuver requires  $dy/d\theta = (n - \cos \gamma)/M > 0$ , from Eq. (61),  $p_\gamma > 0$ , and the condition for the final arc to be a C-arc,  $dp_M/d\theta > 0$ , is that at the final time

$$\frac{C_{Lmax}}{w} (1 + C_{Lmax}) + \frac{\cos \gamma_f}{M^2} < 0 \quad (63)$$

with  $\cos \gamma_f = -1$  for the Immelman and  $\cos \gamma_f = 1$  for the loop and split-S. If the inequality reverses, the starting arc is a B-arc. Once the starting thrust control has been decided, the condition  $p_M = 0$  during the integration is used for switching of the thrust. The guessed values  $M_f$  and  $p_w$ , or  $w_f$  in the case of free final altitude, are adjusted by the condition on  $w_o$ ,  $M_o$  at  $\gamma_o = 0^\circ, 180^\circ$ .

#### Energy State Approximation

A general optimization problem in three-dimensional atmospheric flight is a difficult problem to solve. A straight forward application of the maximum principle always leads to a two-point boundary-value problem involving several arbitrary parameters. The difficulty can be alleviated by considering a reduced-order problem. If a problem has a certain variable  $x$  varying slowly, then the approximation  $dx/dt = 0$  will provide an equilibrium relation which can be used to eliminate one component of the state vector, or one component of the control vector. A more sophisticated approximation would involve a combination of different state variables and the elimination of variables that are insensitive in

the optimization process. One such efficient technique is the energy-state approximation. The basic ideas of this promising method are contained in the classical papers by Rutowski (1954), Bryson et al. (1969), Kelley (1971) and Breakwell (1977). On the other hand, the technique of approximating the thrust switching sequence as a periodic control (Gilbert & Parsons, 1976) also leads to an efficient computational technique for aircraft performance assessment.

### OPTIMAL REENTRY TRAJECTORIES

With the space shuttle becoming operational, optimal lifting maneuver at hypersonic speed is now a reality. Using spherical coordinates, with  $T = 0$ , we deduce from the Hamiltonian (34)

$$\begin{aligned}\frac{dr}{dt} &= V \sin \gamma, \quad \frac{d\theta}{dt} = \frac{V \cos \gamma \cos \psi}{r \cos \phi}, \quad \frac{d\phi}{dt} = \frac{V}{r} \cos \gamma \sin \psi \\ \frac{dV}{dt} &= -\frac{D}{m} - g \sin \gamma, \quad \frac{dy}{dt} = \frac{L \cos \sigma}{m V} - \frac{g}{V} \cos \gamma, \quad \frac{d\psi}{dt} = \frac{L \sin \sigma}{m V \cos \gamma} \\ &\quad - \frac{V}{r} \cos \gamma \tan \phi \cos \psi\end{aligned}\quad (64)$$

With a parabolic drag polar independent of the Mach number, a locally exponential atmosphere, an inverse-square force field and the use of the modified Chapman's variables

$$Z = \frac{\rho S C^* L}{2m} \sqrt{\frac{r}{\beta}}, \quad v = \frac{V^2}{gr}, \quad s = \int_0^t \frac{V}{r} \cos \gamma dt, \quad \lambda = \frac{C_L}{C_L^*} \quad (65)$$

we have the equations (Vinh, 1981)

$$\begin{aligned}\frac{dZ}{ds} &= -k^2 Z \tan \gamma, \quad \frac{dv}{ds} = -\frac{k Z v (1 + \lambda^2)}{E \cos \gamma} - (2 - v) \tan \gamma, \quad \frac{dy}{ds} = \frac{k Z \lambda \cos \sigma}{\cos \gamma} + 1 - \frac{1}{v} \\ \frac{d\theta}{ds} &= \frac{\cos \psi}{\cos \phi}, \quad \frac{d\phi}{ds} = \sin \psi, \quad \frac{d\psi}{ds} = \frac{k Z \lambda \sin \sigma}{\cos^2 \gamma} - \cos \psi \tan \phi\end{aligned}\quad (66)$$

The aerodynamic control consists of the bank angle  $\sigma$  and the lift control  $\lambda$ , normalized such that  $\lambda = 1$  corresponds to maximum lift-to-drag ratio. The distinctive characteristic of Chapman's formulation is that the equations are free of the physical characteristics of the vehicle, except for the maximum lift-to-drag ratio  $E$  which is the single most important performance parameter. The altitude variable is the dimensionless density  $Z$ , and as such, for entry trajectory from the vacuum,  $Z \approx 0$ . Because atmospheric flight starts with non-zero value of  $Z$ , we shall take its initial value as  $Z_0 = 0.0005$ . Then, regardless of the type of vehicle, for a strictly exponential atmosphere, the variation in the altitude from entry level is simply  $\Delta h = \log(Z/Z_0)$ . The characteristic of the atmosphere is here represented by the value  $k^2 = \beta r$  which can be approximated as a constant with  $k^2 = 900$  for the atmosphere. The equations provide the Keplerian solution in the limit when  $Z \rightarrow 0$ . Besides the Hamiltonian integral  $H = C_0$ , the integrals (35) for the adjoint variables  $p_\theta$ ,  $p_\phi$  and  $p_\psi$  are valid.

### Optimal Planar Reentry

The relevant variables are the altitude  $Z$ , speed  $v$ , flight path angle  $\gamma$  and the range angle  $\theta$  which is now identical to the variable  $s$ . Their equations are

$$\begin{aligned}\frac{dZ}{d\theta} &= -k^2 Z \tan \gamma, \quad \frac{dv}{d\theta} = -\frac{kZv(1+\lambda^2)}{E^* \cos \gamma} - (2-v) \tan \gamma \\ \frac{dv}{d\theta} &= \frac{kZ\lambda}{\cos \gamma} + 1 - \frac{1}{v}, \quad \frac{d\theta}{d\theta} = 1\end{aligned}\quad (67)$$

Besides the adjoint  $p_\theta = C_1$ , and for free time problems,  $H = C_0 = 0$ , we use the modified adjoint variables

$$P = k^2 Z p_Z, \quad N = v p_v, \quad Q = p_\gamma \cos \gamma, \quad F = P/N \quad (68)$$

with equations

$$\begin{aligned}\frac{dP}{d\theta} &= k^2 \left[ C_1 - \frac{Q(1-v)}{v \cos \gamma} - \frac{[vP + (2-v)N]}{v} \tan \gamma \right] \\ \frac{dN}{d\theta} &= -\frac{2N}{v} \tan \gamma - \frac{Q}{v \cos \gamma} \\ \frac{dQ}{d\theta} &= 2C_1 \sin \gamma - \frac{Q(1-v)}{v} \tan \gamma + \frac{(1-2\sin^2 \gamma)}{v \cos \gamma} [vP + (2-v)N] - \frac{kNZ(1+\lambda^2)}{E^*} \tan \gamma\end{aligned}\quad (69)$$

The optimal lift control is either on the boundary  $\lambda = \lambda_{\max}$  or a modulated  $\lambda$

$$\lambda = E^* p_\gamma / 2 v p_v = E^* Q / 2 N \cos \gamma \quad (70)$$

The Hamiltonian integral for the free time case is

$$\frac{kNZ(1+\lambda^2)}{E^* \cos \gamma} - \frac{kZQ\lambda}{\cos^2 \gamma} + \frac{Q(1-v)}{v \cos \gamma} + \frac{[vP + (2-v)N]}{v} \tan \gamma = C_1 \quad (71)$$

Since the adjoints enter the equations linearly, by normalizing with respect to  $C_1$  and using the Hamiltonian integral, only two adjoint equations need be integrated. To show this, we derive the equation for  $\lambda$  by taking the derivative of Eq. (70), using the adjoint equations (69). We have

$$\frac{d\lambda}{d\theta} = \frac{kZ(1-\lambda^2)\sin \gamma}{2 \cos^2 \gamma} + \frac{2\lambda(\lambda + E^* \tan \gamma)}{E^* v} + \frac{E^*}{2 \cos^2 \gamma} (F - 1 + \frac{2}{v}) \quad (72)$$

For the ratio  $F = P/N$  we have

$$\frac{dF}{d\theta} = \frac{k^3 Z (1-\lambda^2)}{E^* \cos \gamma} + \frac{2F(\lambda + E^* \tan \gamma)}{E^* v} \quad (73)$$

In terms of the state variables and the new adjoints  $F$  and  $\lambda$  which also plays the role of lift control, the Hamiltonian integral becomes

$$\frac{kZ(1-\lambda^2)}{E^* \cos \gamma} + \frac{2(1-v)\lambda}{E^* v} + (F - 1 + \frac{2}{v}) \tan \gamma = \frac{C_1}{N} \quad (74)$$

To solve any free time problem with prescribed initial values  $\theta_0 = 0$ ,  $Z_0$ ,  $v_0$  and  $\gamma_0$ , we integrate the state equations and the equations (72) and (73) for  $\lambda^0$  and  $F^0$ . This requires the evaluation of the initial values  $\lambda_0$  and  $F_0$  such that

the final and transversality conditions are satisfied. The general free-time planar problem is hence a two-parameter problem. In this case, with the present formulation, the Hamiltonian integral (74) is inoperative. When the range is free,  $C_1 = 0$ , this integral is used at the initial time to evaluate  $\lambda^o$  in terms of  $F^o$  or vice versa. The problem is a one-parameter problem.

Typical planar problems are the pull-up maneuver to a final altitude with maximum residual speed, or reaching a maximum altitude with a prescribed energy loss, glide with maximum range, or optimal skip maneuvers. The only difficulty encountered in the computation is in the case of glide with maximum range. The oscillations in the altitude for three cases of maximum lift-to-drag ratio,  $E^* = 1, 2, 3$  are shown in Fig. 9. In this case, the normalized lift coefficient  $\lambda$  is initially oscillating with large amplitude and later on it stabilizes and tends to  $\lambda = 1$  for maximum lift-to-drag ratio glide. For this type of behavior, the initial values  $\lambda^o$  and  $F^o$  have to be evaluated with great accuracy. This sensitivity is alleviated by backward integration.

### Three-Dimensional Equilibrium Glide.

The analysis is manageable in a reduced-order system. In general, for gliding entry, the performance index is the range and this, in turn, leads to trajectories with small and slowly varying flight path angle. The approximation  $\gamma = 0$ ,  $d\gamma/ds = 0$  is called the equilibrium glide condition. Then, using this simplification in the general equations (66), we first have

$$k Z = (1 - v)/v \lambda \cos \sigma \quad (75)$$

and then

$$\begin{aligned} \frac{d\theta}{ds} &= \frac{\cos \psi}{\cos \phi}, \quad \frac{d\phi}{ds} = \sin \psi \\ \frac{dv}{ds} &= -\frac{(1 + \lambda^2)(1-v)}{E^* \lambda \cos \sigma}, \quad \frac{d\psi}{ds} = \frac{(1-v)}{v} \tan \sigma - \cos \psi \tan \phi \end{aligned} \quad (76)$$

The main variables are the longitudinal range  $\theta$ , lateral range  $\phi$ , speed  $v$  and heading  $\psi$ . The Hamiltonian of the reduced problem is

$$H = p_\theta \frac{\cos \psi}{\cos \phi} + p_\phi \sin \psi - p_v \frac{(1 + \lambda^2)(1-v)}{E^* \lambda \cos \sigma} + p_\psi \left[ \frac{(1-v)}{v} \tan \sigma - \cos \psi \tan \phi \right] \quad (77)$$

which is maximized with respect to  $\lambda$  for  $\lambda = \pm 1$ . In general, optimal glide is effected at maximum lift-to-drag ratio,  $\lambda = 1$ . For the bank control, it is either  $\sigma = \sigma_{\max}$  or a variable bank angle such that

$$\sin \sigma = E^* p_\psi / 2 v p_v \quad (78)$$

In the reduced problem, the integrals (35) are conserved and hence, using  $\lambda = 1$  and Eq. (78) in the Hamiltonian integral  $H = 0$ , for free-time problems, we have the explicit variable bank control law

$$\tan \sigma = p_\psi \frac{(1-v)}{v} \times \frac{\cos \phi}{p_\theta \cos \psi + p_\phi \cos \phi \sin \psi - p_\psi \sin \phi \cos \psi} \quad (79)$$

where the adjoints  $p_\theta$ ,  $p_\phi$  and  $p_\psi$  are given by Eqs. (35). By normalizing with respect to one of the three constants  $C_1$ ,  $C_2$ , and  $C_3$ , the general problem becomes a two-parameter problem. As examples, we consider the cases:

a/ Maximum cross range. The longitudinal range is free,  $C_1 = 0$ . The final latitude is maximized,  $J = \phi_f$ ,  $p_\phi(s_f) = 1$ , and the final heading is free,  $p_\psi(s_f) = 0$ . Hence, from Eqs. (35) this gives the solution for  $p_\phi$  and  $p_\psi$

$$p_\phi = \cos(\theta_f - \theta), \quad p_\psi = \cos \phi \sin(\theta_f - \theta) \quad (80)$$

providing the law for the bank control

$$\tan \sigma = \frac{(1-v)}{v} \times \frac{\cos \phi \sin(\theta_f - \theta)}{\sin \psi \cos(\theta_f - \theta) - \sin \phi \cos \psi \sin(\theta_f - \theta)} \quad (81)$$

The unknown parameter is the final longitude  $\theta_f$ , and it is selected such that when  $\theta = \theta_f$ , the final condition imposed on  $v_f$  is identically satisfied.

b/ Footprint of a reentry vehicle. The footprint is defined as the boundary of the reachable domain on the surface of the earth of a reentry vehicle. To find this boundary, we maximize the final latitude for each prescribed longitude. Hence, unlike the previous problem, here  $C_1 \neq 0$ , and it is a two-parameter problem. But one of the two constants involved can be used as a generating, or scanning, parameter and the iteration can be made on the other parameter.

Consider a trajectory from  $M_o$  leading to the final point  $M_f$  on the boundary  $C$  (Fig. 10). If the original coordinate system  $M \theta \phi$  is rotated through an angle  $\psi'$  to  $M' \theta' \phi'$  such that  $M' \theta'$  is parallel to the tangent at  $M_f$  to the footprint  $C$  and if the problem is solved in the new system, we are led to the problem of maximum cross range with free longitudinal range and the control law (81) applies. Hence, we can use  $\psi'$ , which denotes the new initial heading in the new system as arbitrary scanning parameter, and for each value  $\psi'$ , solve the problem as in case a/ using the prime notation on the state variables. This leads to the final optimal values  $\theta'_f, \phi'_f$  and using spherical trigonometry, it can be shown that the transformation to the original coordinates is

$$\tan \theta_f = \tan \theta'_f \cos \psi'_o + \tan \phi'_f \sin \psi'_o \sec \theta'_f \quad (82)$$

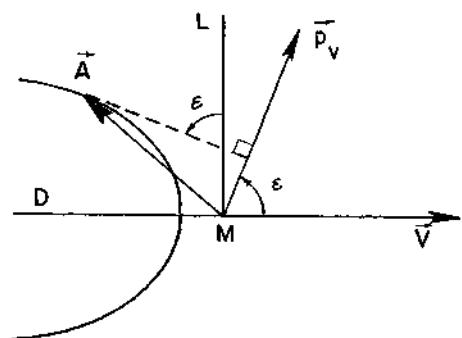
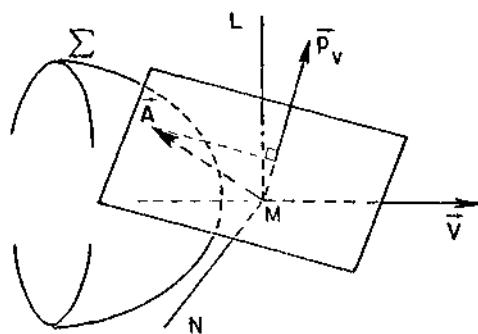
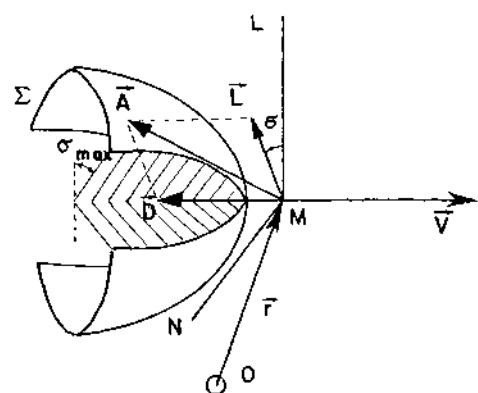
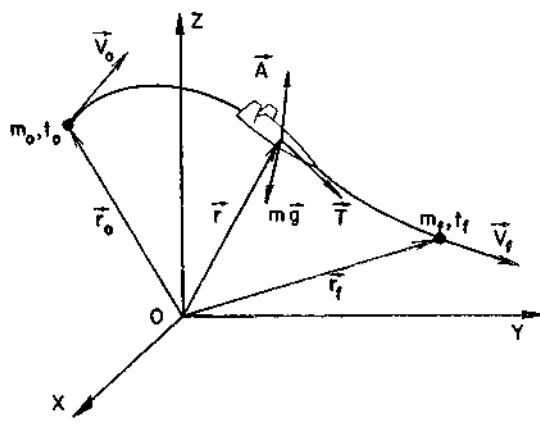
$$\sin \phi_f = \sin \phi'_f \cos \psi'_o - \sin \theta'_f \cos \phi'_f \sin \psi'_o$$

Several related problems in gliding flight can be formulated. For example, we may consider the problem of reaching a point inside the footprint, or a latitude, with maximum residual speed, or when the point is out of reach, the problem of closest approach to that point with a prescribed final speed. For a lifting vehicle with a relatively high maximum lift-to-drag ratio, the problem of closed circuit gliding flight with maximum range angle is also of interest and can be solved by the control law (79).

#### REFERENCES

- Breakwell, J. V. (1977) Optimal flight-path-angle transitions in optimal airplane climb. *J. Aircraft*, 14, 782-786.
- Bryson, A. E. and Denham, W. (1962) A steepest ascent method for solving optimum programming problems. *J. Applied Mechanics*, 29, 247-257.
- Bryson, A. E., Desai, M. N. and Hoffman, W. C. (1969) Energy-state approximation in performance optimization of supersonic aircraft. *J. Aircraft*,

- 6 , 481-488 .
- Contensou , P. (1969) Quelques propriétés générales des domaines de manœuvrabilité à frontière semi-linéaire . Application aux transferts d'orbites . Advanced Problems and Methods for Space Flight Optimization . B. Fraeij s de Veubeke , Ed. , Pergamon , Oxford .
- Gilbert , E. G. and Parsons , M. G. (1976) Periodic control and the optimality of aircraft cruise . J. Aircraft , 13 , 828-830 .
- Kelley , H. J. (1959) An investigation of optimum zoom climb techniques . J.A.S. , 26 , 794-802 .
- Kelley , H. J. , Kopp , R. E. and Moyer , H. G. (1967) Singular Extremals . Topics in Optimization . G. Leitmann , Ed. , Academic Press , New York .
- Kelley , H. J. (1971) Reduced-order modeling in aircraft mission analysis . AIAA J. , 9 , 349-350 .
- Krotov , V. F. , Bukreev , V. Z. and Gurman , V. I. (1971) New Variational Methods in Flight Dynamics . NASA TTF - 657 .
- Marec , J.-P. (1979) Optimal Space Trajectories . Elsevier , Amsterdam .
- Miele , A. (1962 a) Flight Mechanics . Theory of Flight Paths . Addison-Wesley , Reading .
- Miele , A. (1962 b) The calculus of variations in applied aerodynamics and flight mechanics . Optimization Techniques . G. Leitmann , Ed. , Academic Press , New York .
- Moyer , H. G. (1973) Integrals for optimal flight over a spherical earth . AIAA J , 11 , 1441-1443 .
- Rutowski , E. S. (1954) Energy approach to the general aircraft performance problem . J. A. S. , 21 , 187-195 .
- Shkadov , L. M. , Bukhanova , R. S. , Illarionov , V. F. and Plokhikh , V. P. (1975) Mechanics of Optimum Three-Dimensional Motion of Aircraft in the Atmosphere . NASA TTF - 777 .
- Vinh , N. X. (1981) Optimal Trajectories in Atmospheric flight . Elsevier , Amsterdam .



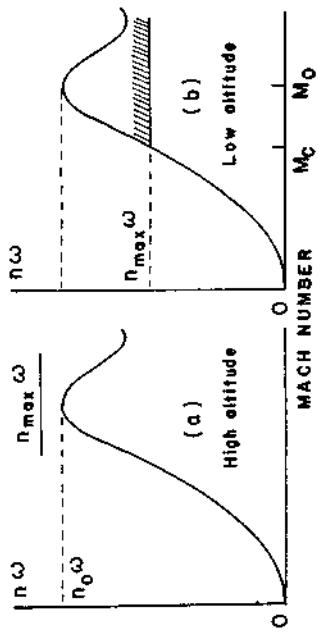


Fig. 5. Velocity transformation

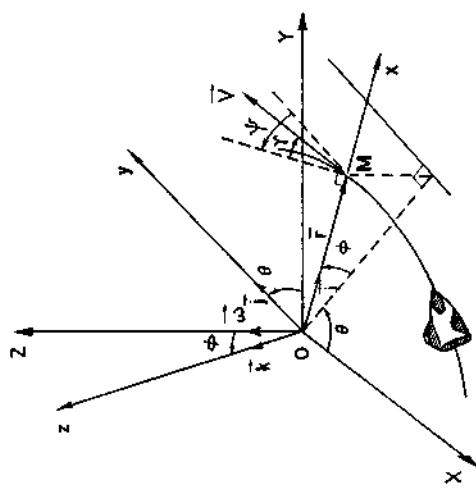


Fig. 6. Spherical coordinates

Fig. 7. Flight envelope

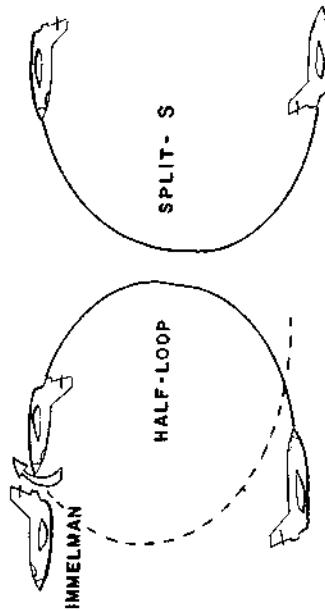


Fig. 8. Loop, Immelman and split-S

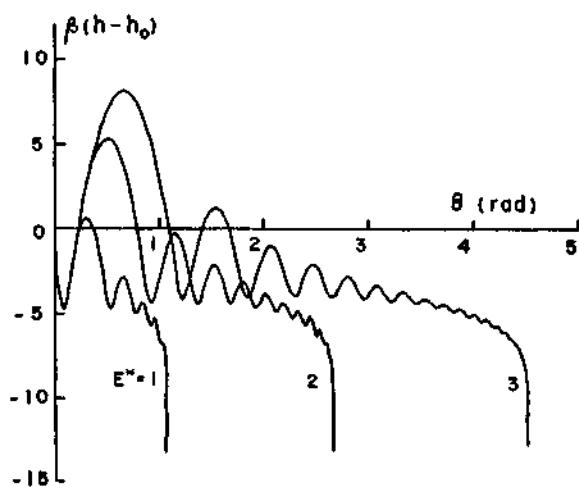


Fig. 9. Altitude oscillation in maximum range  
glide

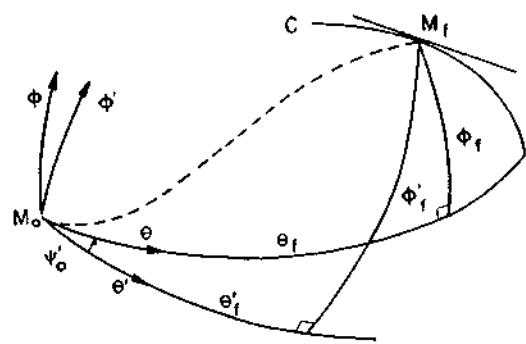


Fig. 10. Rotation of coordinates

## A REVIEW OF THE ATTITUDE CONTROL OF COMMUNICATION SATELLITES

C. A. Markland

*European Space Research and Technology Centre, Noordwijk,  
The Netherlands*

### ABSTRACT

This paper reviews the development of communication satellites and their attitude control systems from the earliest passive vehicles to the present day. Satellite configurations are presented first, including the development in both the Western countries and the Soviet Union. Then a more specialised review is given of the development of attitude control systems from spin stabilisation through dual-spin and momentum bias, to zero momentum. Finally, research problems arising in the control of future missions are discussed.

### KEYWORDS

Communication satellites; stationary orbits; satellite attitude control; spin stabilization; three-axis stabilization; angular momentum; stationkeeping; synchronous satellites; attitude indicators.

### INTRODUCTION

In order to review the development of the attitude control systems of communication satellites, it is necessary first to consider the development of the satellites as a whole and especially their overall configuration. This has been driven by the increasing demands of the communication payload. In particular, the requirement to generate an increasing amount of electrical power on-board the satellite has been the dominant driving factor for satellite configurations and the corresponding attitude control systems.

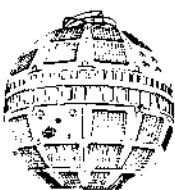
Thus, this paper begins with a review of the development of communication satellite configurations from the general point-of-view of attitude control. This is in two parts, which show the very different approaches adopted by the Western countries and the Soviet Union. This is followed by a more specialised review of the development of the attitude control systems from spin stabilisation through dual-spin and momentum bias, to zero momentum. Particular attention is given to defining the control problems and the solutions used to date. Finally, research problems arising in attitude control for future missions are discussed.

## SATELLITE CONFIGURATION DEVELOPMENT

### The Development in Western Countries

To state that any particular satellite was 'first' may give rise to some contention in some quarter. Nonetheless, defining communication satellites as orbiting vehicles whose main purpose is to facilitate communication, the claimant to this position appears to be the 'Score' satellite launched in December, 1958, just 14 months after the dawn of the Space Age with the Russian Sputnik. In reality, this vehicle was simply a communication payload fitted into the nose cone of an Atlas rocket. It was followed by Echo-1 in August 1960. This was a gas-filled balloon which served as a reflector for bouncing radio signals back to Earth. It was entirely passive except for a battery-powered radio beacon. Telstar-1 was the first active repeater communications satellite. For this satellite, passive attitude control was provided by means of a one-shot spin-up system comprising a tank of high-pressure nitrogen gas and a nozzle. The angular momentum of the spinning satellite then ensures that any attitude change due to external torques is limited to a small nutation rather than a free tumbling motion. The era of active attitude control began with Relay-1 launched in December, 1962. This was spin-stabilised with a fluid-filled toroidal nutation damper, and attitude control was effected by a magnetic torquing coil in conjunction with sun and earth sensors (Wilmotte, 1964).

All of the foregoing satellites were in near-earth orbits. There was considerable discussion at this time as to the desirability of higher altitude orbits. Syncor-2, launched in July, 1962, was the first communication satellite in a 24-hour synchronous orbit. This satellite demonstrated clearly the feasibility and the desirability of such an orbit for most purposes, and the great majority of satellite communication systems in the western countries have since been placed into such orbits. Syncor-2 had an inclination of 33 degrees, but Syncor-3, launched one year later, had nominally zero inclination, i.e. a true geostationary orbit for which the ground antennas do not need to track the satellite. The attitude control of Syncor was achieved by active spin stabilisation with nitrogen gas jets, Vee-beam sun sensors, and RF sensing on the ground using signal strength and polarisation. Hydrogen peroxide was used in a second reaction control system primarily for orbit control (Williams and Cole, 1964). A very similar control system was used in the first commercial communications satellite, Early Bird or Intelsat 1, launched in April, 1965. The system is splendidly simple, and its pointing accuracy of about 1 degree is quite suitable for the toroidal antenna beam of width 11 degrees.



TELSTAR



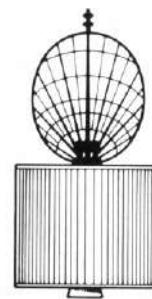
EARLY BIRD

Fig. 1. Spin stabilised communication satellites.

Soon, however, there was demand for higher antenna gain, and correspondingly narrower antenna beams. The state of technology was insufficient to allow a change to three-axis stabilisation (note the Advent Project which was abandoned in 1962) and there was some debate as to whether a pencil beam should be produced by despinning the beam electronically or mechanically. The first of the Lincoln Experimental Satellites (LES-1) and the first NASA Applications Technology Satellite (ATS-1) both produced a pencil beam of some 17 degrees by electrical despin in February 1965 and December, 1966, respectively. This proved to be an interim solution to the problem with mechanical despin being used first on ATS-3 in November, 1967. Since then, this configuration of a simple spinner, with its very simple attitude and orbit control system, together with a mechanically despun antenna section, has developed into being the most used configuration. Some examples are Intelsat III, Skynet, NATO, Anik A, DSCS II, Palapa and Marisat. (Note that many of these look like dual-spin satellites, but there is an important difference from the point-of-view of stability and control.)



INTELSAT III



ANIK A

Fig. 2. Spin stabilised satellites with mechanically despun antennas.

Further extension of this configuration in order to provide more power from solar arrays fixed to the rotating drum of the satellite body and to provide multiple spot beams from large parabolic antenna dishes presented a problem. The foregoing satellites all spin about the axis of maximum inertia, and it can be shown that - for a given amount of angular momentum - spin about the axis of maximum inertia has a lower kinetic energy than spin about any other axis. Hence, if there is a source of energy dissipation on-board the satellite, then spin about any other axis will be damped out, leaving pure spin about the maximum inertia axis, which is the desired state. When the satellite is made longer, by incorporating a longer solar array drum and a taller antenna farm, the spin axis becomes the axis of minimum inertia and energy dissipation makes this spin an unstable condition.

Hughes Aircraft Company found a solution to this problem and they named it the 'Gyrostat' (Iorillo, 1965 and 1967). Now it is generally referred to as the 'dual-spin' configuration. In this configuration, an energy dissipating device (e.g. a damped pendulum) is placed on the despun section, and this has the effect of stabilising the complete prolate (pencil-shaped) satellite in the desired condition of spin. Hughes has used this principle very successfully, first on Tacsat in 1969, and later on the Intelsat IV and IVA series starting in 1971, and recently on SBS.

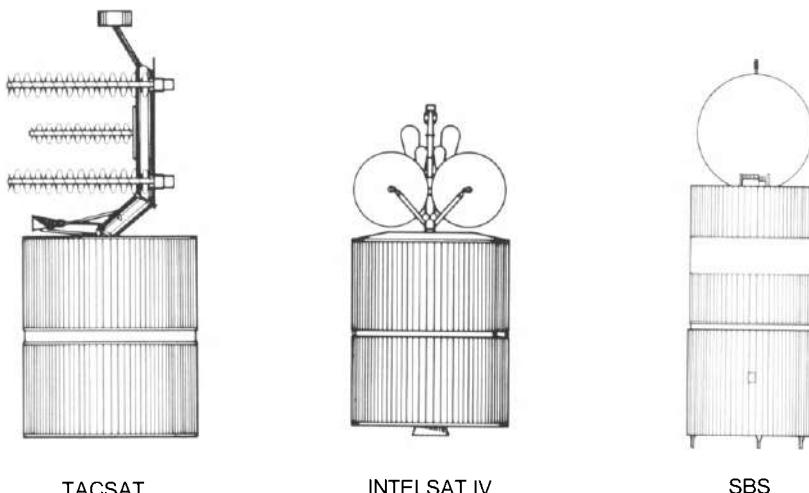
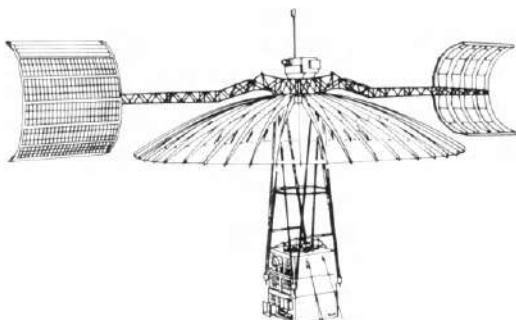


Fig. 3. Dual-spin satellites or 'Gyrostats'.

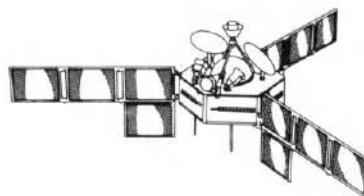
Although this configuration has been used very successfully by its inventors, there was at that time considerable doubt about the theoretical validity of dual-spin stabilisation and about its practical merits in satisfying the foreseen mission requirements. The former led to many theoretical papers (Likins, 1967; Mingori, 1969; Pringle, 1969; Brewer and others, 1970), and the latter led to many comparative mission and configuration analyses (Bakeman and others, 1971). The fact that more electrical power can be obtained in a three-axis stabilised configuration for a given satellite mass has led most satellite designers to favour this configuration.

The first three-axis stabilised communication satellite, ATS-6 launched in May 1974, has also been one of the most complex in configuration and in control. The configuration is shown in Fig. 4 which shows the 9 m diameter unfurlable antenna, fixed solar arrays 16 m tip-to-tip, and the communications and service module supported on trusses some 5 m from the antenna and solar arrays. The attitude control system comprised 3 reaction wheels, hydrazine thrusters, caesium bombardment ion engines, coarse and fine sun sensors, IR earth sensors, a Polaris star-tracker, RF sensing by interferometer and by monopulse, 3-axis rate gyros, yaw rate-integrating gyros, and re-programmable digital control electronics (Redisch, 1975; Freeman, 1976). Whilst this complexity was only conceivable in the context of a two-year experimental satellite, it has been a very successful system.

The second three-axis stabilised communication satellite, and the first communication satellite produced outside the USA and the USSR, was the Franco-German satellite 'Symphonie' launched in December, 1974. As shown in Fig. 4, the solar arrays are again body-fixed rather than sun-tracking. This satellite pioneered the use of bi-propellant for the apogee boost motor, and momentum-bias for three-axis attitude control without measurement of yaw attitude (Metzger, 1976; Pfeiffer and Schroeter, 1978).



ATS-6



SYMPHONIE

Fig. 4. The first three-axis stabilised communication satellites.

The RCA-Satcom launched in December, 1975, and the Canadian satellite CTS or 'Hermes' in January, 1976, established the now-classical configuration for communication satellites with a central box-like body carrying the communication payload, antennas and service module and North-South sun-tracking solar arrays (see Fig. 5). With this configuration the real improvement in electrical power generating capability can be realised. As can be seen from Fig. 6, the array power for a given satellite mass (both at beginning of life) is typically about 1 watt/kg for a spinner and 2 watt/kg for a 3-axis satellite, and in future television satellites the latter figure will be increased to 5 watt/kg. Both of these satellites have momentum bias. Satcom is interesting in that it extends the use of magnetic torquing into geostationary orbit, while CTS used hydrazine and had very large solar arrays.

Since 1978, this overall configuration of communication satellite has been used on many missions, e.g. Fleetsatcom, BSE, OTS, Anik B, and Intelsat V. Many future missions will employ the same satellite configuration, but with variations on the momentum bias configuration and with zero momentum as will be discussed later in this paper.

In terms of numbers, Table 1 shows that some 126 communication satellites have been launched successfully in the period 1958 to 1980. The early move from low earth orbits to the geostationary orbit is shown, and the trend from passive control, through spinners and dual-spin to three-axis stabilisation is clear. The numerical dominance of spinners with despun antennas (which includes large antennas on stable spinners, e.g. Anik A and DSCS II, as well as dual-spin) is remarkable even in recent years. It is evident that significant developments are being made on both types of configuration in order to extend their mission capabilities (e.g. SBS and Leasat in the spinners and Intelsat V, ECS, TDRSS, DSCS III, TVSAT and LSAT in the three-axis stabilised satellites).

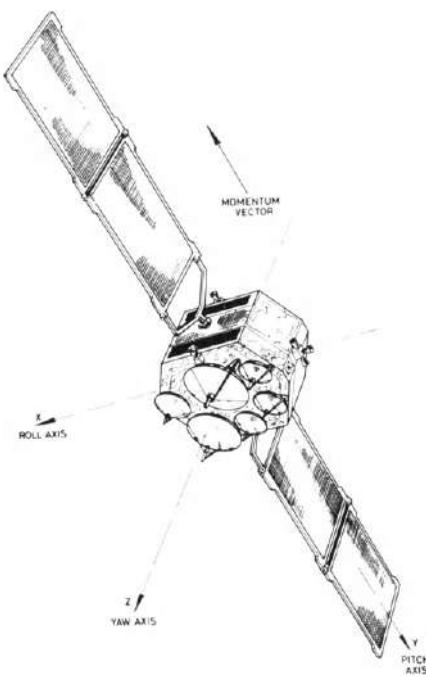


Fig. 5. The 'classical' three-axis stabilised configuration.

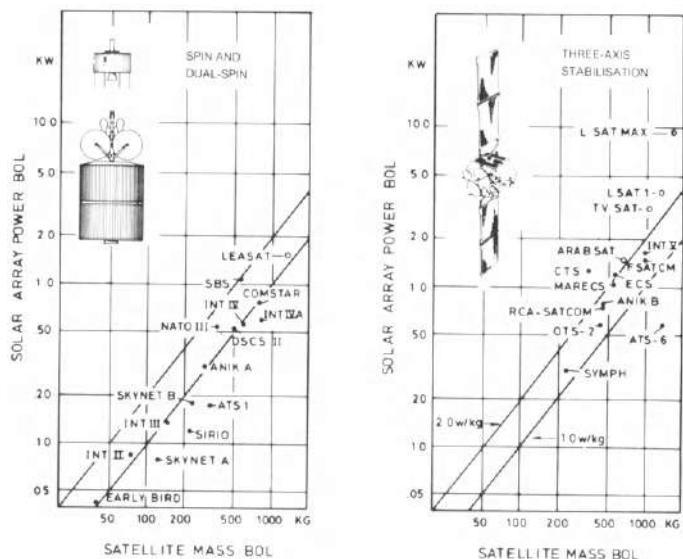


Fig. 6. Satellite power-to-mass ratios.

### The Development in the Soviet Union

The development of communication satellites in the Soviet Union is interesting because it is so very different from the development trend in the Western countries. Nothing has been published about the early experimental satellites, which presumably were flown, but the story begins with Molniya-1A launched in April, 1965. This is a satellite of 1000 kg, three-axis stabilised, in a highly elliptical, inclined orbit. It could hardly contrast more sharply with the configuration and control system designs being developed in the Western countries. The reason lies in mission requirements and also probably in technology. From the point-of-view of mission requirements, the geostationary orbit is not so well suited to the Soviet Union as it is to most Western countries, because much of the Soviet Union is at high latitudes. This means that the elevation angle of ground station antennas pointing to a geostationary satellite will be small, giving difficulties with physical obstructions and atmospherics. Hence, an orbit with a high apogee over the northern hemisphere is desirable, even though the satellite must then be actively tracked by the ground antennas and several satellites are required to give all-day coverage. It is reasonable to suppose that the capability and availability of very powerful launchers also played a role. Further, the high latitude of the launch site makes a high inclination orbit much more accessible than an equatorial orbit, and the technology of hermetically sealed bodies facilitates the use of three-axis stabilisation.

The orbit chosen for Molniya is a repeating orbit with period of one-half of a sidereal day, such that the satellite makes the same path as seen from the ground every day (i.e. every second orbit). This is achieved with the orbit shown in Fig. 7. Because of the high apogee altitude, each satellite can be used for communication for 8 - 10 hours and three satellites are required for all-day communications (Plummer, 1970).

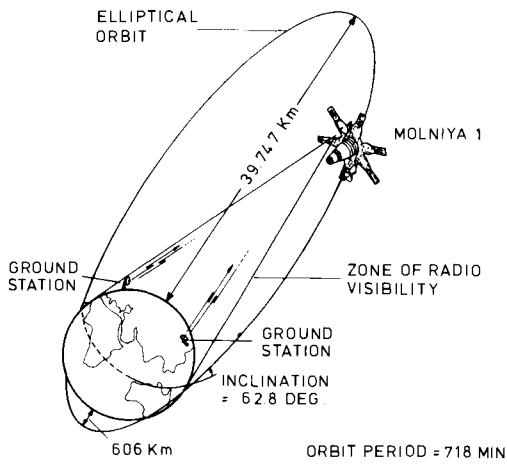


Fig. 7. The orbit of Molniya.

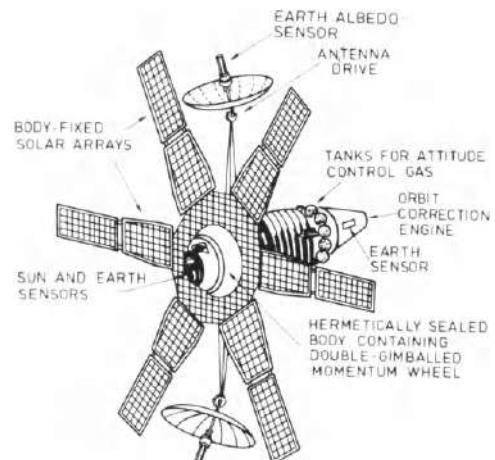


Fig. 8. The Molniya satellite.

The satellite consists of a sealed cylindrical body, with body-fixed solar panels, and two parabolic antennas mounted on booms. The function of attitude control for the body is to point the solar arrays to the sun. This is achieved by means of gas jets for large angles and by a double-gimballed momentum wheel (with its momentum vector directed towards the sun) for damping and for fine control. The requirement to point one of the two antennas to the earth is met by rotating the body about the satellite-sun line by control of the wheel speed and by rotating the antenna about the other axis on a hinge which allows over 180 degrees rotation. The sensor used for this control is an earth albedo sensor (which senses sunlight reflected by the earth). It is fixed to the antenna but with its boresight offset from the antenna boresight because it tracks the illuminated edge of the earth. The pointing accuracy of this system is not known, but since global coverage is the objective, something in the order of 1 degree would be acceptable. There are also two infra-red earth sensors mounted on the main body. These are used to orient the satellite prior to perigee burns for correcting the orbital period (Raushenbakh and Tokar, 1971).

The Molniya satellite has evolved through three stages with mass increasing from 1000 to 1500 kg, but apparently with no change to attitude control. Some 78 Molniyas have been launched since 1965.

In 1974 a Molniya was put into geostationary orbit, and since that date 18 geostationary satellites have been launched. These are known as Statsionar, Ekran, Raduga, and Gorizont. Little has been published about these satellites. They appear to be based on extensions of Molniya technology, and Ekran and Raduga have very large sun-pointing solar arrays.

Another much simpler type of communication satellite has been launched in the Cosmos series starting with Cosmos 939 in 1977. This is a 40 kg spherical satellite launched 8 at a time into 1500 km circular orbit probably with passive spin stabilisation. Satellites with identical parameters were launched earlier, but they were not stated to be communication satellites.

In terms of numbers, some 178 communication satellites have been successfully launched, as shown in Table 2. The small Cosmos satellites and the Molniyas account for the great majority, but the number of large geostationary satellites is growing rapidly.

Year	TYPE OF ATTITUDE CONTROL AND ORBIT										Total	
	passive control		active spin		spin with mech. despun antenna		dual-spin		3-axis stabilised			
	Not GEO	Near GEO	Not GEO	Near GEO	Not GEO	Near GEO	Not GEO	Near GEO	Not GEO	Near GEO		
1958	1										1	
1959	2										0	
1960	1										2	
1961	1										1	
1962	2		1								3	
1963	1		1								2	
1964	1		1								3	
1965	2		4		1						7	
1966		7		1							8	
1967		11		5		1					17	
1968		8		1		1					10	
1969						3					4	
1970	1					4					5	
1971						3		2			5	
1972	1					1		2			4	
1973						3		1			4	
1974	1					3		1		2	7	
1975						1		2		2	5	
1976						5		3		4	12	
1977						6		1			7	
1978	1					3		3		4	11	
1979						3				1	4	
1980								1		3	4	
Total	14	26	6	10		37		17		16	126	

Table 1: Attitude Control and Orbit of Communication Satellites Produced in the Western Countries

Year	TYPE OF ATTITUDE CONTROL AND ORBIT										Total	
	passive control		active spin		spin with mech. despun antenna		dual-spin		3-axis stabilised			
	Not GEO	Near GEO	Not GEO	Near GEO	Not GEO	Near GEO	Not GEO	Near GEO	Not GEO	Near GEO		
1958												
1959												
1960												
1961												
1962												
1963												
1964												
1965									2		2	
1966								2			2	
1967							3			3		
1968							3			3		
1969							2			2		
1970							5			5		
1971							3			3		
1972							6			6		
1973							8			8		
1974							6	1		7		
1975							10	1		11		
1976							7	2		9		
1977	8							6	2	16		
1978	34							6	2	42		
1979	16							5	5	26		
1980	24							4	5	33		
Total	82							78	18	178		

Table 2: Attitude Control and Orbit of Communication Satellites Produced in the Soviet Union

## ATTITUDE CONTROL SYSTEMS

Pure Spin

The early communication satellites were all spin-stabilised, i.e. the stability of the spin axis orientation is assured by the gyroscopic effect of the angular momentum of the satellite body. Thus, in response to an external disturbance, the spin axis makes a limited nutational motion and it does not tumble. The orientation of the other two satellite axes is not controlled and they rotate freely. The antenna beam is directed to the earth by fixing its boresight perpendicular to the spin axis and orienting the spin axis to be perpendicular to the orbit plane. In the simplest satellites this orientation is provided by the launch vehicle and satellite attitude control is passive, e.g. Telstar and IDCSP. The accuracy of pointing depends mainly on the initial conditions provided by the launcher. Normally a few degrees is adequate. Disturbance torques due to satellite interaction with the space environment do change the spin axis orientation, but they are usually negligible (residual magnetic dipoles and solar pressure mainly). More important is the generation of eddy currents that cause the spin rate to decrease (Yu, 1963). Also, particular attention must be given to ensuring that the largest principal axis of inertia coincides with the desired spin axis in order to ensure stability and an acceptable level of wobble. Nutation may be damped by a passive nutation damper.

When satellites are to be put into geostationary orbit, it is usual for the launcher to inject the satellite into an elliptical transfer orbit with apogee near geostationary altitude, and then to use the apogee boost motor on-board the satellite to circularise the orbit. Because this involves satellite re-orientation manoeuvres, an active on-board attitude control system is required. In practice, this involves reaction control by gas jets (magnetic control may be feasible, but it is more appropriate for small attitude trim operations due to the low torque).

The basic attitude control system employs two or more gas jets and two fan beam sun sensors together with some type of earth sensor. Usually only two gas jets are used: one axial jet used in a continuous mode for long duration burns (e.g. initial station acquisition orbit corrections and inclination control or North-South station-keeping) and in a pulsed mode for attitude control by precession of the momentum vector, and one radial jet thrusting through the centre-of-gravity in a pulsed mode for longitude control or East-West station-keeping. The torque due to the axial jet during North-South station-keeping causes a wobble of the spin axis (precession) and this attitude disturbance is either tolerated in the pointing error budget or an extra gas jet must be added to eliminate this torque. (It may be acceptable to use thruster of the redundant gas jet system.) Control of spin rate is not covered in the two jet system, and because the initial spin is usually provided by the launcher and only very small variations are experienced this is either omitted or performed by offsetting the axial and radial thrusters and accepting the resulting cross-coupling.

Passive nutation damping is provided on active spinners in order to reduce the nutation remaining after gas jet manoeuvres in acceptably short periods of time (less than one hour). These are tuned to the nutation frequency and a great variety of energy dissipating devices have been used, e.g. ball-in-tube, fluid-filled toroids and tubes, and pendulums with eddy current dampers. The literature contains many papers on the dynamics of nutation damping (Newkirk and others, 1960; Ayache and Lynch, 1969; Schneider and Likins, 1973; Ancher, 1974; Ancher and others, 1976).

Attitude sensing is achieved by a combination of sun and earth sensing. The sun sensor comprises two slit sensors with fan-shaped fields-of-view, one in the meridian plane and one inclined. By measuring the time interval between the pulses from these two slits, the angle between the spin axis and the satellite-sun line can be determined. In order to determine the complete attitude, an earth reference is needed. On Syncor, this was obtained from the polarisation direction of radio signals received from the satellite, the transmitter polarisation being parallel to the spin axis (Williams and Cole, 1964). More commonly, an IR earth sensor with two pencil beams in a meridian plane is employed, and the attitude of the spin-axis with respect to the satellite-earth line is determined by measurement of the earth chord lengths.

Attitude determination is normally performed on the ground from the telemetered sun and earth sensor data, since the orientation of the spin axis changes only very slowly. Then the attitude control requirements are also calculated on ground and a telecommand to fire thruster pulses at a given phase angle from the sun or earth is issued to the satellite. As can be seen in Fig. 9, the hardware required on-board the satellite is minimal.

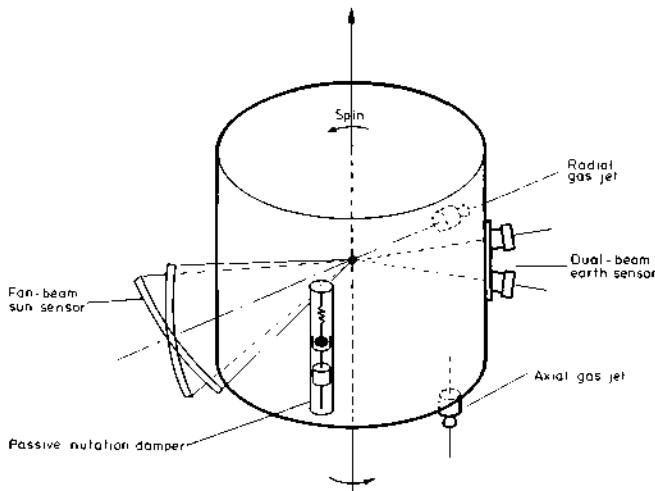


Fig. 9. Control system hardware for active spin stabilisation.

### Spin with Mechanically Despun Antenna

When a mechanically despun antenna is employed, the antenna boresight is pointed to the earth in azimuth (East-West direction) by a despin servo loop while elevation control is performed by the conventional spin-axis controller described in the previous chapter.

The spinning section (termed the rotor) contains an earth sensor so the antenna pointing angle can be calculated by measuring the relative angle between the rotor and the despun antenna. This relative angle can be obtained simply by an optical or magnetic sensor that generates a pulse when (for example) the nominal antenna boresight and the centre of the earth sensor field-of-view coincide. Then by comparing the time of the rising and falling edges of the earth sensor pulse with the time of the optical/magnetic sensor pulse the sign of the torque to be applied to the despin motor can be determined. If linear control or off-nominal pointing is required then the time interval must be measured and multiplied by the measured spin rate in order to determine the actual angle. Additional rate loops may be added to provide the initial despin and rapid acquisition from large angles. In any event, the additional hardware for this loop is small, although bearing friction and sensor sampling and delay effects have to be taken into account in the design (McElvain and Kushi, 1967).

The increase in size of antennas from the simple horn of Intelsat III to the multiple antenna structures as on Intelsat IV leads to problems that are common to both intrinsically stable spinners with despun antennas and intrinsically unstable spinners (i.e. dual-spin configurations which are considered further in the next chapter). These problems are due to the fact that such structures are generally statically and/or dynamically unbalanced with respect to the bearing axes, and so there are product of inertia terms in the dynamics.

The first problem is that the despin control system interacts with the nutational dynamics, and the interaction must be made stable. This can be done fairly readily (Philips, 1971; Brewer, Bushell, and Swift, 1974; Slafer and Marbach, 1975). Further, by employing an accelerometer to measure nutation and feeding this into the despin control system, the requirements on passive nutation dampers can be reduced or eliminated (Slafer and Smay, 1977).

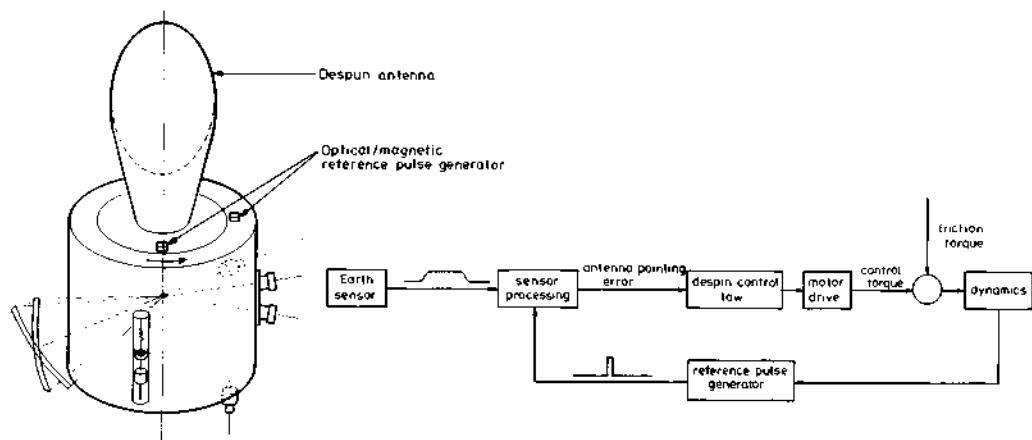


Fig. 10. Control scheme for active spin stabilisation with a mechanically despun antenna.

The second problem is that if the despin motor is not activated (due to a failure in the control loop or in the power supply) then the platform will spin up due to bearing friction and the satellite will move to its state of minimum energy, i.e. a wobbling motion with spin about the axis of overall maximum inertia. It has been shown that, if the rotating section is asymmetric, re-activating the despin control may not bring the system back to the desired attitude, because a certain minimum torque is necessary in order to escape from this minimum energy 'trap'. If the motor cannot provide this torque, then a strategy of torque pulsing may enable an escape. A further problem arises for the same inertia configuration during the despin, which is known as a resonance trap. Here, the periodic mass property variations coincide with a 'nutation frequency' in the system, and again the escape from this trap depends on motor torque magnitude or torque pulsing (Scher and Farrenkopf, 1974; Cochran, 1977).

From the point-of-view of performance, the balance of the rotor and its alignment with the despin bearing will obviously be important in avoiding continuous antenna wobble motion (McIntyre and Gianelli, 1971; Brewer, Bushell, and Swift, 1974). In practice, these effects can be made very small. Also the effect of solar pressure torques has to be considered but, in view of the very large momentum of the rotor, the effect is significant only in the long term. Hence attitude control via the ground as used for simple spinners remains feasible.

#### Dual-Spin

The attitude control systems of dual-spin satellites are basically the same as those of the active spin-stabilised satellite with mechanically despun antennas previously discussed. However, there is a critical distinction with regard to their stability, because the inertia configuration of a dual-spin satellite (i.e. spin inertia less than transverse inertia or 'prolate') implies that it is normally unstable. Landon and Stewart (1964) showed that this configuration is stable if all the energy dissipation is on the despun platform. Unfortunately, this condition is not satisfied in practice because of fuel slosh in the spinning section. Iorillo (1965) derived a mathematical criterion for the energy dissipation rate required on the platform in order to stabilise a satellite having a known energy dissipation rate in the rotor. This result was confirmed by more rigorous analyses (Likins, 1967; Mingori, 1969; Pringle, 1969; Brewer and others, 1970). Thus, the vital distinction in hardware is that for dual-spin satellites the nutation damper must be located on the despun platform.

Obviously, one major difficulty of applying this result is to identify the sources of energy dissipation and their dissipation rates. This is especially difficult for fuel slosh in the rotor, and it is still not fully understood (Martin, 1971; Neer, 1972; Slabinski, 1977). On the first dual-spin satellite, energy dissipation in the despin bearing was found to cause occasional instability for small angles of nutation (Johnson, 1970). On another spinning satellite, the dissipation in heat pipes was found to have an unexpectedly large effect.

The problems of minimum energy and resonance traps states are intrinsically as described previously except that these cause more severe tumbling motion because minimum energy implies a 'flat spin' (Adams, 1980). The destabilising effect of energy dissipation in the rotor could also be important for the prolate, dual-spin configuration during such recovery modes.

### Three-Axis Stabilisation

As noted earlier in this paper, the need for higher satellite power to mass ratios has resulted in the trend to three-axis stabilisation with sun-pointing solar arrays and earth-pointing central body. The classical configuration used for Satcom, CTS, OTS, Intelsat V and others is shown in Fig. 5.

The term 'three-axis stabilisation' means that all three geometric axes of the satellite are controlled to point in a given direction, whereas in 'spin stabilisation' only the spin axis is controlled to have a given direction while the other two axes rotate freely. The conventional axis set is shown in Fig. 5, and the satellite axes are defined with following nominal orientations:

- the yaw axis (Z) is in the plane of the orbit and it points to the centre of the earth,
- the roll axis (X) is also in the plane of the orbit pointing along the orbital velocity vector,
- the pitch axis (Y) completes the right-hand set by being perpendicular to the orbit plane, and if the orbit inclination is zero then it points southward.

Plainly, the objective of attitude control is to maintain this nominal state or some commanded deviation from it.

The external torques that disturb the satellite in its normal operational phase are mainly due to solar radiation pressure on the solar arrays. This creates torques in two ways. First, if the two arrays do not exactly face the sun but they are offset or twisted with respect to each other then they create a torque like that driving a windmill. Second, if the centre-of-pressure of the entire satellite does not pass through its centre-of-gravity, then again there are torques due to this misalignment. The satellite is designed to minimise this misalignment, but unavoidable effects such as fuel consumption cause movement of the centre-of-gravity and array bending causes movement of the centre-of-pressure.

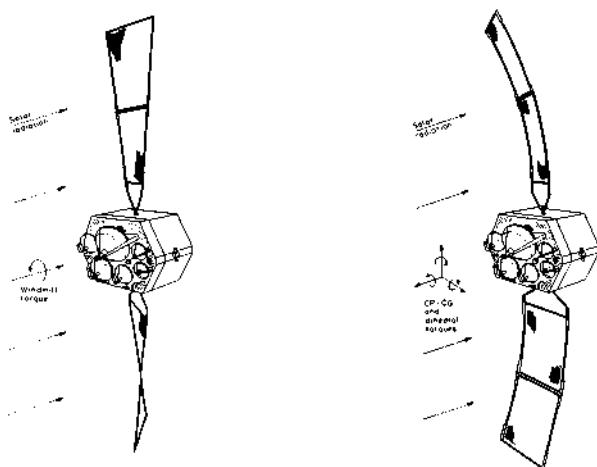


Fig. 11. Satellite disturbance torques due to solar radiation.

As regards attitude sensors, it is obvious to place an earth sensor on the antenna face and this can measure roll and pitch angles by reference to the apparent motion of the earth. However, yaw angles (i.e. rotations about the earth-satellite line) cannot be measured in this way. Yaw can be measured by means of a sun sensor looking along the roll axis, but the sun is not always in the field-of-view of such a sensor. Even if several sensors were used, there are still times when the earth, satellite, and sun are in-line thus preventing continuous yaw measurement. Using a star sensor on the North or South face would provide a continuous measurement, but such sensors are complex and they would be subject to interference from the solar array. This problem of yaw sensing leads to the two classes of three-axis stabilisation that will be discussed later.

All three-axis control systems have used the reaction torque of a wheel as the actuator for controlling pitch attitude. This wheel may be a momentum wheel or a reaction wheel: the distinction being that a momentum wheel operates at a high angular velocity with small variations in velocity around the bias velocity, whereas a reaction wheel operates at positive and negative angular velocities without any bias. In either case, the control of pitch attitude is essentially the same as the despun control of the antenna platform on a spinning rotor.

Thus, the different types of three-axis attitude control systems arise from the different ways of controlling roll-yaw motion. Also, because the satellite configurations are very different from satellites with large spinning rotors which incorporate the solar arrays, the problems of attitude control during station-keeping are more difficult.

#### Momentum-Bias Control Systems

The most frequently adopted solution to the problem of yaw attitude sensing has been to use the fact that the satellite roll and yaw axes rotate with respect to inertially-fixed axes due to the daily orbital motion. Hence, with only pitch control, yaw errors are seen as roll errors after 6 hours. This is illustrated in Fig. 12. In order to strengthen this roll-yaw coupling and to reduce the effect of external torques, a momentum wheel must be installed in the satellite with its angular momentum parallel to the pitch axis. Then, it can be shown that, by combining roll corrections with a torque about yaw, yaw attitude can be controlled without yaw sensing. This may be achieved by offsetting the roll actuator such that it produces both roll and yaw torque, and the technique was termed 'WHECON' by its original authors (Dougherty, Scott, and Rodden, 1968).

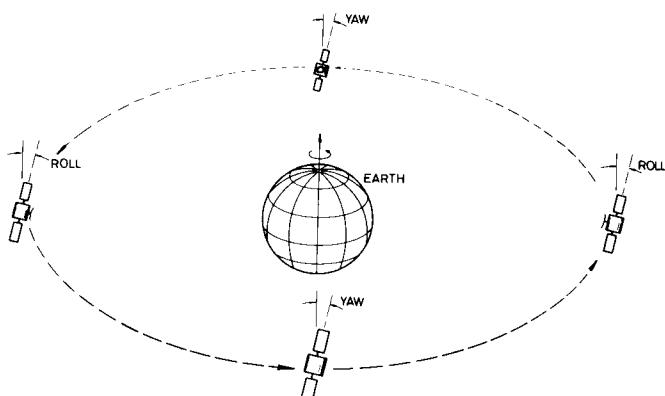


Fig. 12. Interchange of roll and yaw errors by orbital motion.

The validity of this technique may be demonstrated by considering the roll and yaw dynamic equations respectively:

$$I_{xx} \ddot{\phi} + H\dot{\psi} + \omega_o^H \Phi = T_x^d + T_x^c,$$

$$I_{zz} \ddot{\psi} - H\dot{\phi} + \omega_o^H \Psi = T_z^d + T_z^c.$$

Here,  $\phi$  and  $\psi$  are the roll and yaw angles,  $H$  is the angular momentum of the wheel,  $\omega_o$  is the orbital frequency, and the superscripts to the applied torques on the right hand side indicate disturbance and control torque respectively.

Then cross-coupling the control torque by making

$$T_z^c = -kT_x^c,$$

the yaw equation then becomes:

$$I_{zz} \ddot{\psi} - H\dot{\phi} + \omega_o^H \Psi = T_z^d - k(I_{xx} \ddot{\phi} + H\dot{\psi} + \omega_o^H \Phi - T_x^d).$$

Re-arranging gives

$$I_{zz} \ddot{\psi} + kH\dot{\psi} + \omega_o^H \Psi = T_z^d + kT_x^d - k(I_{xx} \ddot{\phi} + \frac{H}{k}\dot{\phi} + \omega_o^H \Phi).$$

Now, the left hand side shows that the control cross-coupling factor  $k$  introduces damping to the yaw dynamics, which is now a classical damped second-order system. The right hand side shows that the yaw motion is driven by external disturbance torques and by the roll motion. The latter effect will be made negligible by the action of a tight roll control loop, and the external disturbances cause a steady-state yaw error that is inversely proportional to the magnitude of the momentum:

$$\Psi_{ss} = \frac{T_z^d + kT_x^d}{\omega_o^H}.$$

This principle is employed in all momentum-bias systems.

The simplest momentum-bias system uses a single fixed momentum wheel aligned with the pitch axis. Both gas jet and magnetic actuators have been proposed for roll-yaw control (Dougherty, Scott and Rodden, 1968; Schmidt, 1978). The system used on FleetSatcom is a low-gain gas jet system, and a similar system is used on OTS and Intelsat V. In this system, damping of the nutational mode is ensured by commanding two pulses from an offset thruster when a roll threshold is encountered. The first pulse drives the roll angle away from the threshold and into a nutational motion, and the second pulse is given after half the nutation cycle is complete which stops the nutation because it makes the satellite momentum vector coincide with the wheel momentum. This is illustrated in Fig. 13 and the control system schematic is given in Fig. 14. In practice, it has been shown that the interval between pulses should be slightly greater than half the nutation period in order to damp any initial nutation (Iwens, Fleming, and Spector, 1974).

This system is often referred to as a 'zero degrees-of-freedom' momentum-bias system because the momentum is not free to move with respect to the main body. It is clear that roll control by gas jets involves a roll deadband, and if the disturbances are large or the deadband is small, then a large number of pulses can be required. Also, if it is desired to point the yaw axis away from the centre of the earth, then it is necessary to use gas jets to precess the momentum around the daily orbit, which is costly in terms of fuel mass.

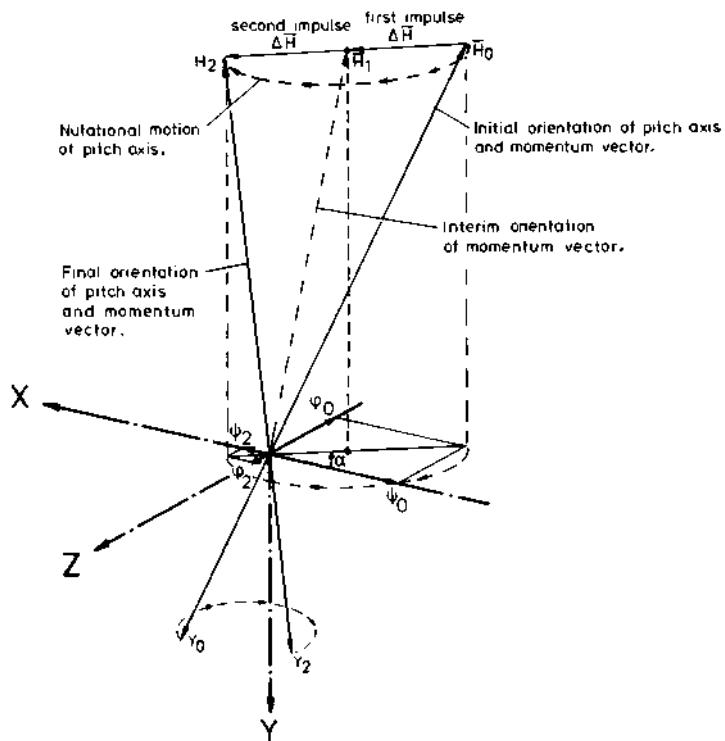


Fig. 13. Roll/yaw control of fixed momentum wheel by the two-pulse technique.

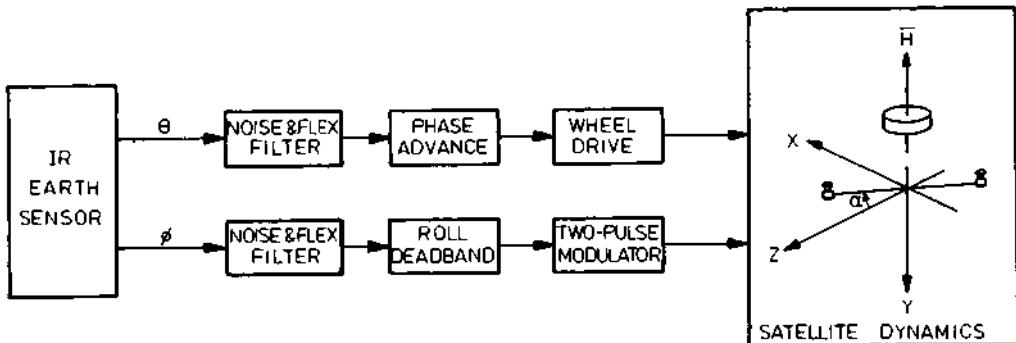


Fig. 14. Fixed momentum wheel attitude control system.

For some missions, including TDRSS and ECS, these factors lead to the adoption of a one degree-of-freedom system in which the momentum vector is steered in one plane. This is usually the pitch-yaw plane, but there may be advantages in skewing this plane. Momentum steering can be achieved by gimbaling a momentum wheel or by arranging two wheels in a Vee-configuration and modulating their speed appropriately. Pitch angle control remains essentially as in the zero degree-of-freedom case, but the roll control loop inherently provides linear control of roll angle, so enabling roll tracking and nutation damping. However,

roll control is slow and rather weak, because it works through the nutation mode and because only a small component of wheel torque acts for roll control. Yaw control is achieved by monitoring the momentum offset angle and commanding skewed gas jets when the yaw deadband is exceeded. Since the permitted yaw error is always several times greater than the permitted roll error, this system inherently uses fewer gas jet pulses and slightly less fuel than the zero degree-of-freedom system. Some improvement in steady-state accuracy is also achieved due to the elimination of the roll deadband (Broquet and Govin, 1977; Lebsok, 1978) but the roll transients due to yaw unloading pulses can be large. Therefore, the operational benefit of the Vee-wheel system is the tracking capability rather than improved accuracy.

The ultimate refinement of momentum bias is the two degree-of-freedom system in which the momentum vector is steered in two axes. This can be achieved by a double-gimballed momentum wheel or by an arrangement of fixed momentum and reaction wheels. This system offers great versatility in roll and yaw tracking. Momentum unloading is required only when the gimbal limits are exceeded, which can often be combined with station-keeping manoeuvres. The design of these systems follows the roll/yaw cross-coupling concept (Scott, 1970; Hammond, 1977). Such systems have not been used for geostationary flight control systems in the West because of their mass and complexity, but they have been used by the Soviet Union and they will be used on the German TVSAT.

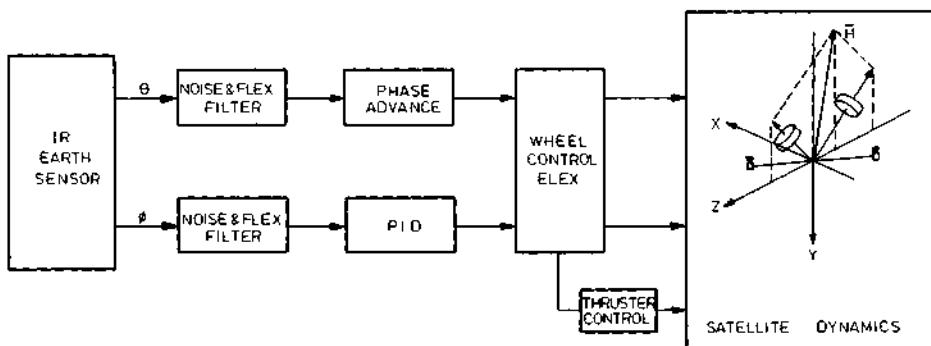


Fig. 15. Vee-wheel attitude control system.

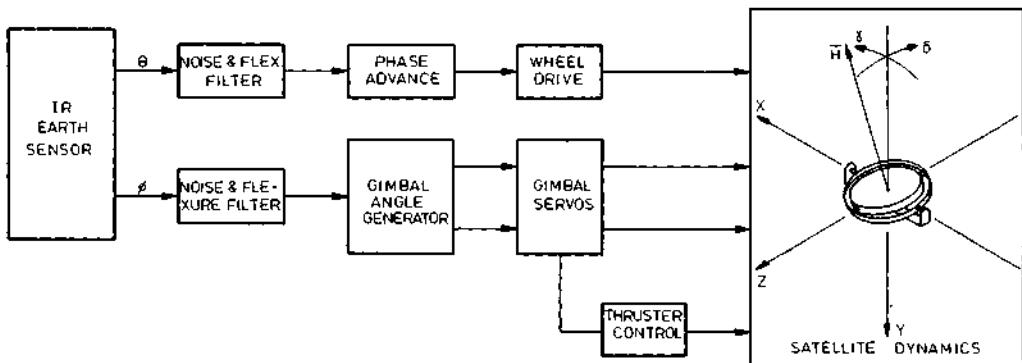


Fig. 16. Double-gimballed momentum wheel attitude control system.

## Zero Momentum Systems

The indirect control of yaw achieved by momentum bias is satisfactory provided the arrays are of a moderate size such that the solar disturbance torques are fairly low. Then with a moderate size of wheel momentum the steady-state yaw error can be made acceptable. For large arrays, the yaw requirements lead to impractical momentum levels. This problem may be alleviated by estimating the disturbance torques and compensating them, but questions arise about the accuracy of the estimation process in tracking random and rapid variations in disturbances (e.g. on exit from eclipse).

Further problems arise with regard to roll control. For the fixed momentum wheel system, the half-cycle nutation damping technique breaks down, because the large array inertias lead to a very long nutation period and roll control requires pulses more frequently than half a nutation period because of the large disturbance torques. Other methods of damping nutation by more frequent use of gas jets involve a heavy fuel mass penalty, and the use of magnetic torquing would involve massive torque rods. The Vee-wheel system has the problem of controlling roll during momentum unloading transients because roll control torque is very limited. Other one degree-of-freedom systems (using a single-gimballed momentum wheel or a reaction wheel) may give better performance, particularly if the momentum is steered in a plane skewed from the ZY plane. Certainly, a two degree-of-freedom system employing a double-gimballed momentum wheel will have no difficulties with roll control, but it is widely considered to be complex.

Hence, at some power level, momentum bias becomes inadequate and one is obliged to measure yaw attitude. The definition of this critical power level requires a detailed trade-off, and the result depends heavily on the assumptions, but the limit would seem to be around 4 kW.

The preferred technique for yaw attitude determination is to employ a digital sun sensor on the roll axis together with a rate-integrating gyro and an estimator. Thus, the sun sensor provides the absolute reference for updating the estimator while the sun is in its field-of-view (about 4 hours per day) and the gyro provides continuous measurements during the entire orbit. Reaction wheels are usually preferred as actuators rather than gas jets, because they are able to absorb cyclic disturbance torques without consuming fuel. Also, they should be more accurate because the deadbands are eliminated.

Having full three-axis attitude measurement and no momentum cross-coupling of the axes, the control loops can be designed simply on a single axis basis. However, sensor noise has to be adequately filtered and - as always with large arrays - great care has to be given to flexible modes of the array. The main hardware problem is in obtaining adequate reliability of the yaw measurement system through a proper level of gyro redundancy.

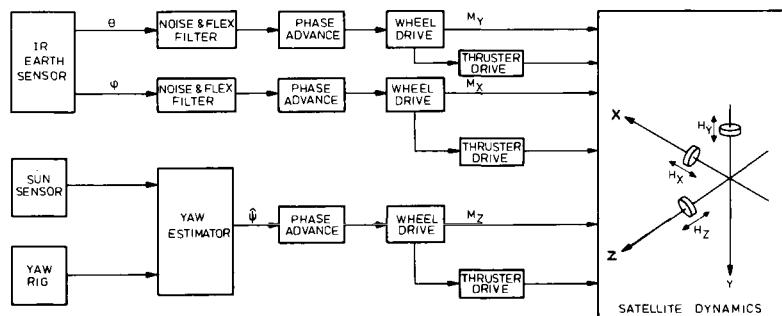


Fig. 17. Zero momentum reaction wheel attitude control system.

## RESEARCH AREAS FOR FUTURE CONTROL SYSTEMS

### Station-Keeping

Obviously, large flexible solar arrays are a major feature of three-axis stabilised configurations, and their flexible modes do interact with the attitude control systems. In the normal operational phase, disturbance torques due to solar pressure are low and so the control system can be designed with low gains and small bandwidths. Hence, interaction with the arrays can be avoided simply by low-pass filtering for arrays up to 2 kW (say 40 micro Nm peak disturbance torque). For higher power arrays, more attention has to be given to these modes by actively including them in the design process, but low-pass filtering of the conventional sensor signals provides modal damping even up to 7 kW (say 200 micro Nm).

During station-keeping a pair of thrusters is fired continuously in the North-South direction in order to compensate inclination drift. The burn time may be ten minutes, and the disturbance torque due to thruster misalignments will be 4 or 5 orders of magnitude greater than solar pressure. Therefore a special control system is required for these periods of the mission, using negative pulse modulation of the gas jets. Furthermore, if it is required to maintain the same pointing accuracy as in normal operation then a high gain, high bandwidth control system is essential, and major difficulties arise from excitation of array flexure. In the past (arrays of about 1 kW) it has been possible to use simple filtering (Beysens, 1976), but this is insufficient for large arrays. An added difficulty is the uncertainty of the array frequencies and damping factors. Designs have been attempted using an observer to accurately filter out rigid body motion, but these are found to be sensitive to modal parameter variations. A promising approach being studied is the use of phase-locked loops to identify the modes and the controller can be designed to damp them (Kopf, Brown and Marsh, 1979). A completely different solution is to perform station-keeping in very short burns such that the attitude transients are acceptable. However, for the moment this remains a problem area for high accuracy, high power missions.

### Reaction Control and Magnetic Torquing

Reaction control gas jets are essential for orbit control and attitude manoeuvres, but there is frequently an incompatibility between the high thrust level required for these functions and the need for accurate pointing in the operational phase. Clearly one could implement a two-level reaction control system, but this would have significant mass and complexity penalties.

The disturbance torques during normal operation are very low on current satellites, say up to 20 micro Nm, and such torque levels can easily be generated by magnetic torque rods even with the low ambient magnetic field at geostationary altitude. In view of the simplicity of generating control torque in this way, the use of magnetic torque rods is an attractive actuator for future satellites, which may be used either for closed-loop control or for open-loop disturbance torque compensation.

### Antenna Pointing

The majority of current satellites point their antennas to the ground receivers only in an indirect way, i.e. the control system sensor detects the earth's infra-red radiation and points the boresight of this sensor to the apparent centre of the earth. Thus, any misalignment between the IR sensor boresight and the antenna boresight produces an unobserved error in antenna pointing. The use of a radio frequency (RF) sensor integrated with the communications antenna and operating with a ground beacon in the coverage zone provides a means of eliminating these misalignments which constitute a major part of the present error budget. The actual errors of the RF sensor itself are about equal to those of an IR sensor (bias = 0.02° and noise = 0.01°/1 sigma) but the noise is of a higher frequency and can be effectively filtered out.

Additional benefits that result from RF sensing are:

- (i) the performance required of the IR earth sensor is lower (accuracy and offset pointing capability) and so a less sophisticated/more reliable device can be used (it is always required for earth acquisition at the beginning of life),
- (ii) by combining information from the IR earth sensor and an RF sensor pointing North or South of the equator, yaw attitude can be determined.

Having RF sensing, a further improvement in pointing performance can be obtained by separating attitude control of the antenna from that of the body by using the RF sensor in closed-loop with an antenna pointing mechanism. This will lead to much faster control, due to the great reduction in inertia, and hence to more accurate control in transient situations. It will also simplify attitude control of the body (momentum steering, manoeuvres, and array modes) especially as regards the roll control loop.

### Transfer Orbit

The preceding discussions have concentrated on attitude control for the operational phase of the communication satellites mission. Naturally this is the driving aspect for the design, and the design of other phases should be organised around this operational system. However, in the past, there has been some incompatibility because the transfer orbit has always been spin-stabilised, even for satellites that are three-axis stabilised in normal operation. The reason is that solid apogee boost motors have generally been used, and their high thrust levels would not be controllable without the averaging effect of a rapid spin. Currently there is a trend to use liquid apogee boost motors because of their higher performance and mission flexibility. This leads to the problem of spin stability for prolate configurations which seem to be the usual case. Active nutation damping has been proposed, but there is fear of violent instability by large masses of fuel due to parametric resonances that are impossible to evaluate by ground testing. Hence the desire to perform transfer orbit in a three-axis stabilised mode, and to reconsideration of the actuator and sensor configuration in order to satisfy this requirement.

Another new aspect of this problem arises from the use of the Space Shuttle as the launcher for geostationary communication satellites. Space Shuttle provides injection into a low circular orbit, and there are a great many potential routes to geostationary orbit, e.g. by low-thrust spiralling or by two or more impulsive transfers. Optimisation in terms of mass and cost is still under study, and it may certainly provide new attitude control problems.

### Very Large Structures

In the more distant future, it has been suggested that very large structures may be built in space for many purposes including communications. These may be built of add-on elements such as the concept shown in Fig. 18 (Preukschat, 1981). Obviously, the relative alignment of all these elements cannot give the level of pointing accuracy currently required for communication antennas. Therefore independent control of the antennas to, say, 0.1 degrees or better and fairly loose control of the overall body to, say, 1.0 degree seems appropriate. Body control by spin or momentum bias is unlikely, but a zero-momentum system together with re-programmable control electronics will enable augmentation of the structure. The requirement for yaw attitude measurement by star sensors is facilitated in this structural configuration.

Clearly many new problem areas will be revealed with such a configuration, e.g. rendez-vous and docking of new modules and adapting to changes in physical configuration. Present problem areas such as external disturbance torques take on new dimensions. Thus, one can be sure that there will be ample subjects for study in attitude control for several years to come.

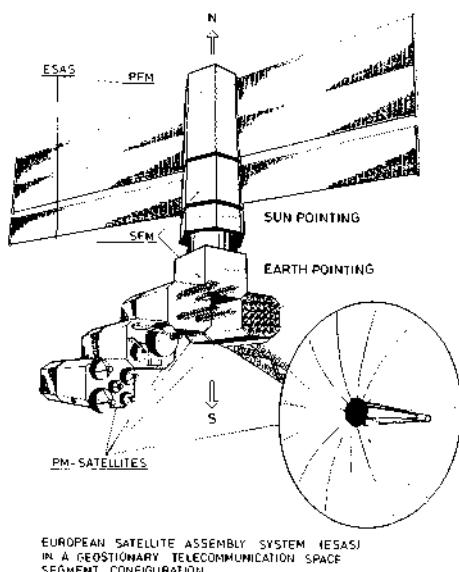


Fig. 18. Concept of a large space vehicle for communications.

### CONCLUSIONS

In the near term, it is plain that both dual-spin and classical three-axis stabilised configurations will be further developed to satisfy mission requirements into the 1990s. Improvements will be effected in antenna pointing performance by employing radio frequency sensing, and the body control system designs will be refined on the basis of known systems. In the longer term, a host of new attitude control problems arise for large, modular communication platforms, and research into these problems is just beginning.

## REFERENCES

- Adams, G. J. (1980). Dual-spin spacecraft dynamics during platform spin-up. J. Guidance and Control, Vol. 3, No. 1, 29-36.
- Ancher, L. J. (1974). Design and testing of a passive nutation damper for the stabilisation of a spinning satellite in the limit of a very small nutation angle. IAF Congress, paper 74-175.
- Ancher, L. J., H. vd Brink, and A. Pouw (1976). Study on Passive Nutation Dampers. ESA-CR(P)-788.
- Ayache, K., and R. Lynch (1969). Analysis of the performance of liquid dampers for nutation in spacecraft. J. Spacecraft and Rockets, Vol. 6, No. 12, 1385-1389.
- Bakeman, D. C., K. W. Kenkins, J. J. Rodden, A. J. Iorillo, T. B. Garber (1971). In M. Summerfield (Ed.), Communications Satellites for the 70's: Systems, Vol. 26, Chap. IX, pp. 607-653.
- Beysens, A. (1976). Design of a control loop with flexibility and in-flight identification of the flexible modes. In Dynamics and Control of Non-Rigid Spacecraft. ESA SP-117.
- Brewer, M. and co-workers (1970). A Study of Dual Spin Stability. ESRO CR-24.
- Brewer, M., R. Bushell, and N. Swift (1974). Dual-Spin Stability Study. ESRO CR-471.
- Broquet, J., and B. Govin (1977). Application of a one degree of freedom system to the attitude control of an operational Ariane large satellite. In Attitude and Orbit Control Systems, ESA SP-128, 29-36.
- Cochran, J. E. (1977). Nonlinear resonances in the attitude motion of dual-spin spacecraft. J. Spacecraft and Rockets Vol. 14, No. 9, 562-572.
- Dougherty, H. J., E. D. Scott, and J. J. Rodden (1968). Analysis and design of WHECON - an attitude control concept. AIAA 2nd Communication Satellite Systems Conference, paper AIAA 68-461.
- Freeman, H. R. (1976). Three-axis attitude control for a geostationary satellite. VII IFAC Symposium Automatic Control in Space, Preprints Vol. 2, pp. 69-83.
- Hammond, M. J. (1977). Aspects of the design and functional testing of a double-gimballed wheel system. In Attitude and Orbit Control Systems, ESA SP-128, 279-290.
- Iorillo, A. J. (1965). Nutation damping dynamics of rotor stabilised satellites. ASME Winter Meeting. Chicago, Ill.
- Iorillo, A. J. (1967). Precision aspects of the Hughes gyrostat system. Symposium on attitude stabilisation and control of dual-spin spacecraft, Aerospace Corporation TR-0158 (3307-01)-16. AD-760154. 257-266.
- Iwens, R. P., A. W. Fleming, and V. A. Spector (1974). Precision attitude control with a single body-fixed momentum wheel. AIAA Mechanics and Control of Flight Conference, paper AIAA 74-894.
- Johnson, C. R. (1970). Tacsat I nutation dynamics. AIAA 3rd Communications Systems Conference, paper AIAA 70-455.
- Kopf, E. H., T. K. Brown, and E. L. Marsh (1979). Flexible stator control on the Galileo spacecraft. AAS/AIAA Astrodynamics Specialist Conference, paper 79-161.
- Landon, V. D., and B. Stewart (1964). Nutational stability of an axisymmetric body containing a rotor. J. Spacecraft and Rockets, Vol. 1, 682-684.
- Lebsock, K. L. (1978). High pointing accuracy with a momentum bias attitude control system. AIAA Guidance and Control Conference, paper AIAA 78-569.
- Likins, P. W. (1967). Attitude stability criteria for dual-spin spacecraft. J. Spacecraft and Rockets, Vol. 4, No. 12, 1638-1643.
- Martin, E. (1971). Experimental investigations on the fuel slosh of dual-spin spacecraft. COMSAT Technical Review, Vol. 1, 1-20.

- Martin, E. (1971). Fuel slosh and dynamic stability of Intelsat IV. AIAA Guidance and Control Conference, paper AIAA 71-954.
- McElvain, R. J., and A. Y. Kushi (1967). Design control system for dual-spin spacecraft configuration. Symposium on attitude stabilisation and control of dual-spin spacecraft. Aerospace Corporation TR-0158(3307-01)-16. AD-670154. 173-184.
- McIntyre, J. E., and M. J. Gianelli (1971). Bearing axis wobble for a dual-spin vehicle. J. Spacecraft and Rockets, Vol. 8, No. 9, 945-951.
- Metzger, R. (1976). Attitude stability of Symphonie - theory and flight results VII IFAC Symposium Automatic Control in Space, Preprints Vol. 2, pp 26-40.
- Mingori, D. L. (1969). Effects of energy dissipation on the attitude stability of dual spin satellites. AIAA Journal, Vol. 7, No. 1, 20-27.
- Neer, J. T. (1972). Intelsat IV nutation dynamics. Progress on Astronautics and Aeronautics, Vol. 33, 57-84.
- Newkirk, H. L., W. R. Haseltine, and A. V. Pratt (1960). Stability of rotating space vehicles. Proc. IRE, April, pp 743-750.
- Pfeiffer, B. R. K., and W. G. Schroeter (1978). The Symphonie satellite system. Acta Astronautica Vol. 5, pp. 327-342.
- Philips, K. J. (1971). Linearisation of the closed-loop dynamics of a dual-spin spacecraft. J. Spacecraft and Rockets, Vol. 8, No. 9, 938-945.
- Plummer, K. L. (1970). Soviet communications satellites. Spaceflight, August, 322-324.
- Preukschat, A. W. (1981). Future space segment concepts for lower cost communication satellite systems. ESA STM-217.
- Pringle, R. (1969). The stability of force-free motions of a dual-spin spacecraft. AIAA Journal, Vol. 7, No. 6, 1054-1063.
- Raushenbakh, B. V., and Je. N. Tokar (1970). Attitude control system of Molniya 1 communication satellite. Automatica, Vol. 7, 7-13.
- Redisch, W. N. (1975). ATS-6 description and performance. IEEE Trans. Aerospace and Electronic Systems. Vol. AES-11, No. 6, 994-1003.
- Scher, M. P., and R. L. Farrenkopf (1974). Dynamic trap states of dual-spin spacecraft. AIAA Journal, Vol. 12, No. 12, 1721-1725.
- Schmidt, G. E. (1978). Magnetic attitude control for geosynchronous spacecraft. AIAA 7th Communication Satellites Systems Conference, paper AIAA 78-570.
- Schneider, C. C., and P. W. Likins (1973). Nutation dampers versus precession dampers for asymmetric spinning spacecraft. J. Spacecraft and Rockets, Vol. 10, No. 3, pp. 218-222.
- Scott, E. D. (1970). Double gimballed reaction wheel control system. Proc. 1st Western Space Congress. Santa Maria, Calif.
- Slabinski, V. J. (1977). Intelsat VI in-orbit liquid slosh tests and problems in theoretical analysis of the data. In Attitude control of space vehicles - technological and dynamical problems associated with the presence of liquids. ESA SP-129. 87-102.
- Slafer, L., and H. Marbach (1975). Active control of the dynamics of a dual-spin spacecraft. J. Spacecraft and Rockets, Vol. 12, No. 5, 287-293.
- Slafer, L. I., and J. W. Smay (1977). Comstar spacecraft product-of-inertia coupled electronic nutation damper. AIAA Guidance and Control Conference, paper AIAA 77-1056.
- Smay, J. W., and L. I. Slafer (1976). Dual-spin spacecraft stabilisation using nutation feedback and inertia coupling. J. Spacecraft and Rockets, Vol. 13, No. 11, 650-659.
- Williams, D. D., and R. W. Cole (1964). In K. W. Gatland (Ed.) Telecommunications Satellites, Iliffe Books, London. Chap. 5, pp. 130-155.
- Willmotte, R. M. (1964). In K. W. Gatland (Ed.). Telecommunications Satellites, Iliffe Books, London, Chap. 4, pp. 84-129.
- Yu, E. Y. (1963). Spin decay, spin-precession damping, and spin-axis drift of the Telstar satellite. Bell System Technical Journal, Vol. XLII, No. 5.

## SOLID ROCKET TECHNOLOGY FOR THE EIGHTIES

C. A. Zimmerman, J. Linsk and G. J. Grunwald

*Lockheed Missiles & Space Company, Inc.,  
Sunnyvale, CA, USA*

### ABSTRACT

Current Solid Rocket Technology is sufficiently mature to offer the propulsion system designer attractive options for space applications. Reliable, high performance solid rocket motors can be developed on low risk basis with high schedule confidence. The paper will address today's state-of-the-art in each of the several component areas that comprise modern solid propellant rocket motors.

Propellant, insulation, chamber and nozzle materials and design approaches are reviewed. Characteristics of each, important to system performance, are identified and discussed in relation to each other and to their interaction with other propulsion or vehicle elements. Design and operational constraints that influence selection and characterization of components are discussed. The paper is not intended to be a detailed treatise on any specific element of the modern rocket motor. It presumes a basic knowledge of both vehicle and propulsion technology. The purpose of the paper is to review and evaluate the state-of-the-industry practice in each of the disciplines pertinent to solid rocket motors with emphasis on interactions.

### KEYWORDS

Solid rocket motors; propellant; insulation; chamber materials; nozzle materials.

### INTRODUCTION

Four decades of progress in the design and development of solid propellant rocket motors has resulted in a state-of-the-art that differs significantly from that in existence in the early forties. The availability and characterization of new materials, combined with advances in analytic and test techniques, have transformed early design practices into more precise applications of scientific and engineering disciplines. The solid rocket motor designer no longer is restricted to materials that just happen to be available. The development of materials satisfying specific needs now provides a useful spectrum of choices. In a similar manner, analytical techniques based upon modern computers have aided the engineer and ballistician in achieving optimum designs. Sophisticated laboratory and subscale testing of mate-

rials and components has allowed the selection and tailoring of each element of the system. Finally, full scale static testing under conditions simulating the expected environment, provides the confidence necessary to certify a rocket motor for unmanned missions and ultimately for manned flight missions.

In the sections that follow, four of the major components that comprise the solid propellant rocket motor are discussed. These are propellant, nozzle, chamber and insulation. Since many applications require that the direction of the thrust vector be controlled, a brief review of several thrust vector control systems is presented in the nozzle section. No attempt will be made to provide comprehensive design criteria, but rather to acquaint the reader with available options and alternatives. Comparisons of properties that characterize the material or approach and often limit its application are presented.

#### PROPELLANT

As the propulsion system designer initiates the process that will ultimately yield a usable rocket motor he invariably starts with the selection of a propellant. Considerations of primary concern which should influence his decision include:

- Performance level - Does the system demand the maximum performance possible within a specified weight or volume, or can performance be traded off for some other attribute?
- Operating environment - What are the environmental conditions under which the system will be tested, stored and eventually used?
- Mechanical properties - What demands will be placed upon the propellant from a mechanical properties standpoint? Are there special stress, strain or modulus requirements?
- Ballistic properties - Are there any burn rate, temperature or pressure coefficient specifications that eliminate potential candidate propellants?

The vast majority of solid propellant motors utilized in U. S. space programs are based upon composite propellants. These will be discussed in the following section.

#### Review of Composite Propellant Technology

Propellant and motor development in the United States has dealt with a very wide range of problems relating to advanced systems, new materials, and extension of solids loading to achieve high performance. Among the work horse binder systems widely used in the past were polybutadiene acrylonitrile (PBAN) - used in the space shuttle boost motors and in 120" TITAN motors, and carboxyl terminated polybutadiene (CTPB). The former encompass solids loadings up to 90%, using epoxide cures. The CTPB system employs epoxide and/or aziridine cures. Because of the regular distribution of cross-links at chain ends, and better control of molecular weight between crosslinks, good low temperature strain capability, important for many rockets and space motors, has been possible with CTPB binders.

During the past decade there has been a marked shift in emphasis from carboxyl to hydroxyl terminated polybutadienes (HTPB) as binder prepolymers; in particular those having a functionality of about 2.3, and crosslinked with various difunc-

tional isocyanates have achieved wide acceptance in the solid propellant industry for the following reasons:

- The HTPB binders afford excellent propellant processability at high solids loading (low end of mix viscosity).
- The HTPB propellants show excellent low temperature strain capability.
- The HTPB propellants appear to be capable of satisfactory long term service life, even after repeated temperature cycling.

The applicability of isocyanate cure for routine solid motor manufacture has been thoroughly demonstrated, and its success is indicated by the increasing number of motors loaded with propellants based on hydroxyl terminated prepolymers.

Hydroxyl and carboxyl terminated polyesters also continue to receive attention for special applications. When polyester binders are employed, attention must be given to long term service life, since it has been shown that hydrolysis of ester linkages by moisture can occur, leading ultimately to propellant degradation. Use of seals in motors containing such systems affords adequate protection. Commercially available hydroxyl terminated polyethers have also been successfully employed in propellant binders and are of continuing interest because of high resistance to moisture.

With the increasing complexity of space motor requirements, the propellant technologist plays a key role in selection of ingredients and formulation optimization to attain the desired combination of performance, physical and processing properties, and service life targets. New compounds and synthetic materials continue to be identified and evaluated, but their utility in solid propellants can only be established after considerable study, based on cost/performance trade-off.

At the highest total solids levels (i.e., 88-92 wt %), it has been shown that small amounts of liquid organic plasticizers improve propellant processing characteristics and high rate mechanical properties. Indicative of the difficulties sometimes encountered, there have been reports of propellants in which mix viscosity increased as temperature was raised during mixing from 120°F. to 160°F., rather than decreasing.

Experience with solid propellants intended for long service life has shown that moisture plays a crucial role not only during cure, but as a result of hydrolytic attack on the ester linkages in propellants based on polyester binders and on the -P-N- linkage present when aziridines are used for crosslinking carboxyl containing polymers.

Moisture effects are pervasive, and the reproducibility of initial mechanical properties requires close attention to its control during processing, and the preparation of samples for tensile properties testing. A number of analytical techniques for moisture measurement are effective. Among these is a modified Karl Fischer test, which has shown accuracy of  $\approx 0.002\%$ , when adequate attention is given to handling and transfer of samples for analysis. Techniques for controlling moisture in propellant processing have been implemented, and moisture levels below 0.02% during the processing of urethane propellants are routinely achievable.

In the formulation of composite propellants, a number of interacting factors must be considered. Selection criteria for ingredients to be used encompass the following:

- Total solids for desired theoretical performance.
- Binder system for required mechanical properties, aging characteristics, and low cost.
- Aluminum and oxidizer particle sizes to achieve selected burning rate, good processability, and optimum mechanical properties.
- Type and amount of burn rate catalyst if needed.
- Curatives and cure stoichiometry; cure catalyst, bonding agent and stabilizer as required.

Iron oxide is a commonly used burn rate catalyst. Its effectiveness is a complex function of ammonium perchlorate size distribution, aluminum content, and solids loading, as well as the particle size of the oxide itself. Complexity arises quickly because changes in any one variable influence not only burn rate, but also burn rate response to changes in the other variables.

Burn rates for aluminized propellants cover the range from less than 0.3 to as high as 2 inches per second at 1,000 psig. High rates are achievable with combination of fine ground oxidizer and burn rate catalysts. In addition to iron oxide, liquid iron containing compounds based on ferrocene have been used. Ferrocene itself has been used, but it has the disadvantage of being subject to loss by sublimation from cured propellant. The chief drawback of the liquid ferrocene derivatives is some volatility and a tendency to migrate into adjoining materials such as liner and insulation. To prevent rate changes due to loss of liquid catalyst, the incorporation of the liquid catalyst into the insulation as a plasticizer has proven necessary in some cases. Small oxidizer particle diameter is uniquely suited to obtain high propellant burn rates. Vibratory energy mills have been used to obtain ammonium perchlorate sizes below one micron diameter, but the technology is expensive, and special propellant processing techniques are required.

The achievement of optimum propellant physical properties involves an empirical balancing of solids particle sizes, cure stoichiometry, cure catalyst, bonding agent, and plasticizer levels. Finally, aging studies over a range of environments is mandatory to assure that the stabilizers (antioxidants) employed are suitable. Considerable effort is required to assure that the formulation selected can meet the intended requirements, as well as the performance characteristics.

The chemistry of bonding agents has been developed by Oberth and others in the U.S. Materials such as polyamines, polyurea, and polyaziridines, have been employed to enhance propellant strain capability. These agents are designed to be preferentially absorbed on oxidizer particles. They can bond chemically to  $\text{NH}_4\text{ClO}_4$ , or form absorbed polymeric shells which can then bond covalently through excess functional groups to the binder during the propellant cure process. They have been used successfully in CTPB and HTPB propellants.

The foregoing considerations apply to all composite propellants. If thermally sensitive energetic plasticizers are employed, special attention to stabilizer selection is mandatory. Propellants containing nitroglycerin and trimethylolethane trinitrate have been successfully stabilized, using aromatic amines and phenols. Theoretical considerations can be useful to some extent in formulation optimization, but the history of solid rocket development suggests that empiricism based on a growing body of knowledge will continue to dominate this activity.

Propellants based on energetic polymers and plasticizers containing both fluoro and nitro groups, difluoramino groups, and azide groups are of interest because they provide high theoretical specific impulse. It has been postulated that fluorine is valuable because a significant amount of the aluminum fuel used in solid propellants is converted to gaseous aluminum fluoride, rather than solid  $\text{Al}_2\text{O}_3$ , so that higher efficiency might be expected. These energetic materials are very expensive and not commercially available. In addition, there are processing hazards, and also age related chemical instability. Research in these areas continues to receive support from a number of government agencies in the United States. Finally, it should be noted that sophistication in the application of newer chemical analysis techniques to the study of propellant cure and degradation reactions is playing an increasingly useful role in dealing with chemical problems, and is reducing the element of 'black art' that has long been associated with propellant development.

#### HMX

One of the noteworthy developments of recent years has been the successful demonstration of the utility and handling safety of cyclotetramethylene tetranitramine (HMX), as well as RDX in the formulation of higher performance propellants. With inert binders, propellants containing up to about 14% HMX have shown negative reaction at zero cards in the card gap test, retaining the 1.3 (Class 2) nondetonator designation. Solids particle size (HMX, AP, Al) determine the outcome in the card gap test, as shown by Oberth at Aerojet Solid Propulsion Company (ASPC), and others.

Class A (150 - 160 $\mu$ ) is shipped in a slurry with water and can be dried readily in a blender. Grinding to particle sizes less than 5 microns is safely accomplished in fluid energy mills. Because of these developments, considerable research has been devoted to formulating high energy systems based on HMX as the major oxidizer. Since the burn rate of propellants is largely insensitive to HMX particle size, controlled particle sizes of  $\text{NH}_4\text{ClO}_4$  are employed to attain required ballistic properties, while HMX and aluminum particle sizes are selected to optimize propellant processability and mechanical properties. Only larger HMX particle sizes (>50 microns diameter) respond well to bonding agents. Propellants containing HMX as the major oxidizer are capable of yielding exceptional elongation with proper selection of particle sizes in urethane binders.

#### Some High Energy Propellant Considerations

The combination of energetic propellants in lightweight composite cases, and high motor pressures, can result in increased damage sensitivity following motor malfunction. High order detonations have in fact been experienced during static test. The major reasons for failure are postulated to result from three scenarios:

- Deflagration to detonation transition (DDT) following chamber rupture creating a bed of damaged propellant in the casebond region.

- Detonation of expelled propellant fragments on nearby objects leading to shock to detonation transition (SDT) by sympathetic reaction.
- DDT following grain failure in the case bond region and implosion of the grain inward, causing fragmentation.

The DDT process has been investigated for some years by Russian workers, and has received significant experimental and theoretical attention with respect to energetic propellants in the U. S. in the past seven years. A number of unique laboratory tools have been employed to study propellant DDT phenomena and to measure a propellant's friability or toughness under simulated malfunction conditions.

The key to formulating "tough" high energy propellants, resistant to granulation, requires close attention to binder chemistry and binder-solids particle size interaction. Tough propellants are resistant to break-up under motor damage conditions. Toughness is defined as the area under the stress-strain curve. To compare toughness, 8 gram samples are fired against a steel target from a shotgun at velocities below ignition threshold. The collected fragments are burned in a closed combustion bomb. Pressure is recorded and the maximum measured  $dP/dT$  is used as a basis for propellant toughness comparison. Similarly, damaged propellant fragments obtained from impact and cold motor blowdown tests can be compared in closed bomb tests to compare the effect of varying blowdown conditions and aging effects.

#### Gas Generators

Gas generator solid propellants for powering spacecraft hydraulic systems, for starting jet engines, and for auxiliary power continue to receive development attention. Propellants based on ammonium nitrate predominate, although ammonium perchlorate and HMX based formulations have been successfully demonstrated. Ammonium nitrate undergoes well known crystalline phase changes, including one at 90° - 130°F. which results in a 4.3% increase in volume, but these phase changes can be accommodated in suitable binders. Propellants with very low elongation have been cycled repeatedly between -65° and +165°F., in configurations which are not case bonded. In end burning configurations it is essential to maintain a very small clearance between the inhibited grain and the insulated case to prevent side burning effects due to the penetration of combustion gases.

Compression molding has been state-of-the-art in ammonium nitrate gas generator production, although burn rate control requires complex processing. In recent years, the formulation of castable ammonium nitrate compositions has been demonstrated, and appears to hold promise. Flame temperatures below 1500°F. can be attained in ammonium nitrate based gas generator propellants, and burn rates as low as 0.02 ips at 500 psig have been attained in very long burn time systems for auxiliary power applications. The achievement of clean-burning, low burn rate gas generants has resulted from the incorporation of a variety of coolants and ballistic modifiers in castable formulations.

#### NOZZLES AND THRUST VECTOR CONTROL

##### Design Requirements and Constraints

Every rocket motor has one or more nozzles at the aft end of the combustion chamber to efficiently convert the energy generated from the burning of the propellant to kinetic energy, thereby producing thrust to the system. This thrust is used to in-

crease the velocity of the vehicle and generally to provide steering forces necessary for control. Approximately seventy percent of the total thrust results from the acceleration of the products of combustion through the throat, and the remainder is developed in the nozzle expansion cone. A typical upper stage partially submerged nozzle configuration is shown in Fig. 1. As is true for all components discussed in this paper, the nozzle is an integral component of a larger system and cannot be optimized independently. The nozzle and the exit cone expansion are controlled in a manner that maximizes performance of the vehicle within the required envelope and vehicle weight and cost constraints.

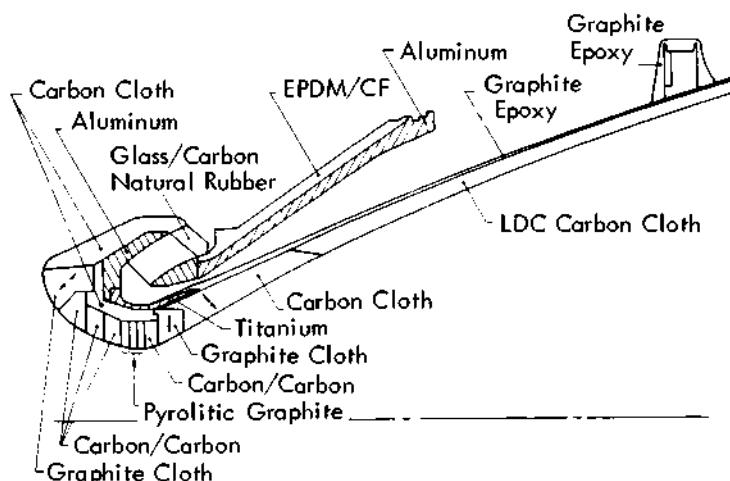


Fig. 1 Upper stage nozzle

Nozzle design itself is an iterative process in which aerodynamic, thermodynamic, and structural considerations are manipulated within given constraints to produce an optimum configuration. Aerodynamics are used to configure the gas-contacting surfaces to minimize losses and to produce the required performance within the envelope limits. In the thermal design phase, liners are selected to maintain the optimum aerodynamic contour, i.e., minimize the erosion, and to select thermal insulators to limit structural components to acceptable temperature levels. Structural elements are then selected to support the thermal components.

Basically, nozzles are either external to the motor or submerged. In the external version, the convergent/divergent nozzle is entirely external to the motor. In the submerged version, the nozzle entry, throat and part or all of the exit cone are submerged in the motor. This concept uses space more efficiently in a volume limited system. The exit cone can be either conical, where the initial divergence angle is the same as the exit angle, or contoured, where the nozzle turns the flow such that the exhaust products exit in a more nearly axial direction, reducing divergence losses.

Most rocket motors require thrust vector control (TVC), thus further complicating the nozzle design. Thrust vector control consists of two basic types:

- Movable nozzle (TVC) - This concept moves the entire nozzle, thus moving the thrust vector to provide side force.

- Fixed nozzle (TVC) - In this concept the nozzle remains stationary, and the thrust vector is deflected by secondary injection or mechanical deflectors to produce side force.

Nozzle design is not only influenced by aerodynamic, thermal, and structural considerations, but also by mission objectives and by system requirements/constraints such as:

- Throat size and acceptable change
- Predicted pressure-time trace
- Production quality and rate
- Design pressure (MEOP)
- Propellant properties
- Diameter of interface with case
- Weight, cost, reliability, and development guideline limits.
- Exit cone configuration
- Envelope limits
- Submergence
- Design TVC angle
- Storage conditions
- Expansion ratio

Fig. 2 is a summary of a typical nozzle design sequence showing major iteration loops and where the above factors influence the design (NASA SP-8115, 1975).

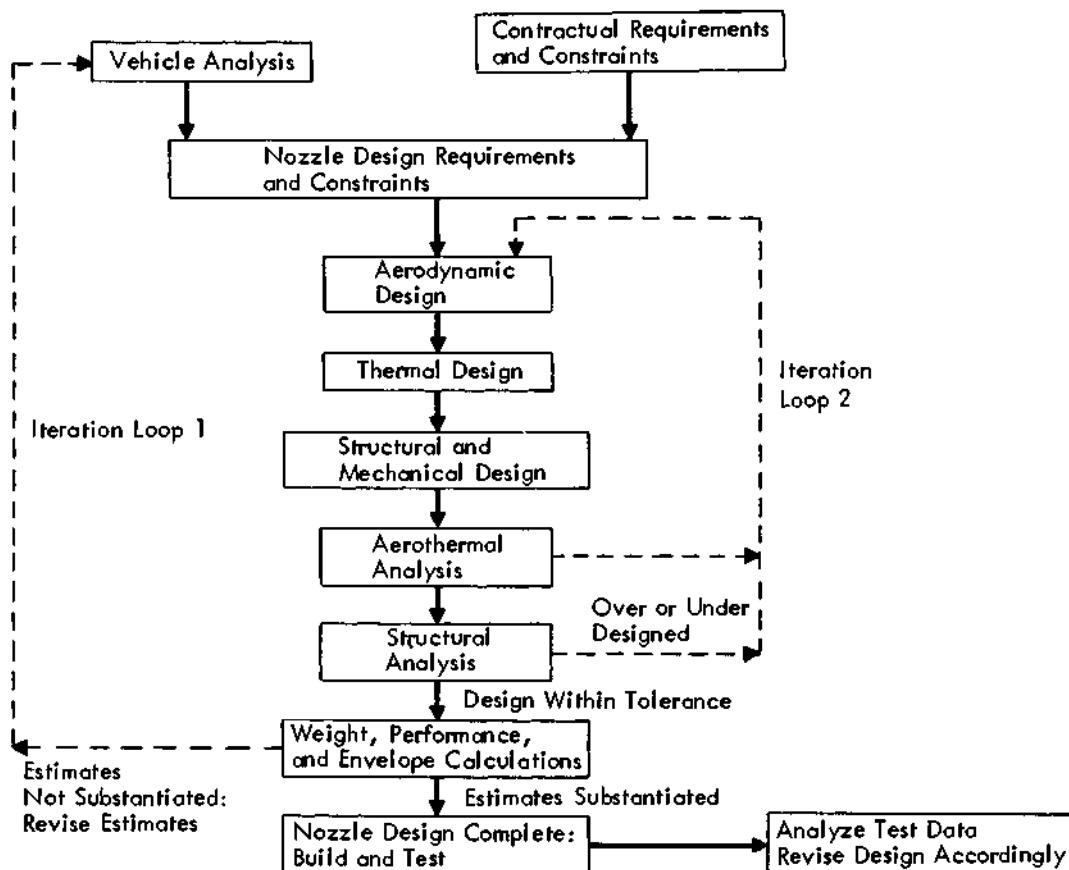


Fig. 2 Nozzle design sequence

Primary emphasis in this paper is placed on nozzles used in larger (>3 feet) motors, utilizing high temperature (> 5500°F) propellants. In the following sections, the discussion focuses on materials, since most of the decisions to be made in the area of nozzle design are intimately related to selection of materials. The hostile environment seen by the nozzle is not found anywhere else in the modern high pressure, high energy solid propellant rocket motor.

#### Current State of the Art

Structural Components. The nozzle must be able to withstand the axial ejection, thrust, and thrust vector control loads as well as the pressure loads on those portions submerged in the chamber. Flight loads, including aerodynamic, inertial, shock and vibration must also be considered. Materials generally available for use are either metals - steel, aluminum, or titanium; or composites of fiberglass, graphite, and Kevlar. These composites can be hand layed up or filament/tape wound, bonded together with either epoxy or phenolic resins. Selection of the best material for each nozzle component must take into consideration the thermal and aerodynamic environments, attach requirements, compatibility with interfacing materials, and the ability to be sealed against gas flow. Further, cost/performance partial effects must be considered especially in boost motors where weight effects on performance are not generally severe. Therefore, metals could be used instead of the more expensive but lighter composites.

Whereas the structural properties of metals used in nozzles are well established, those of composites are less well defined and can vary widely with respect to the type and orientation of reinforcement and binder used. In spite of this limitation and need for further characterization, composite materials are being used whenever possible due to their attractive strength to weight ratio.

Graphite epoxy composites are being successfully used for the nozzle exit cone structural shell at considerable weight savings over aluminum. Glass epoxy and graphite epoxy composites are being used for reinforcements of the flexseal used for TVC in several designs. Considerations other than weight dictated the use of steel reinforcements in the flexseal of the Space Shuttle Solid Rocket Motor. Composite materials will see further use in such areas as the nozzle to motor support structure and the structural support rings for the flexseal.

Thermal Components. Thermal components are those materials that are in contact with the products of combustion for the motor firing duration. Since the pressure/temperature/velocity environment differs along the nozzle surface, generally different materials are used at different stations. These materials are selected on the basis of their capabilities to withstand the environment. In the past, generally ablative materials such as graphite and carbon cloth phenolics were used in the entrance, throat, and exit cone liner. In today's higher performance systems, i.e., motor chamber pressures in excess of 1000 psi, these generally high erosion rate materials are not adequate for use in the throat region because of the inherent performance degradation. The more erosion resistant characteristics of pyrolytic graphite or carbon/carbon are required to nearly maintain the initial expansion ratio of the nozzle in order to minimize performance losses. Graphite and carbon cloth phenolics can and still are being used in the more benign portions of the nozzle.

To simplify the nozzle design, reduce weight, and potentially improve performance, the use of 3D carbon/carbon material in the throat region and free-standing 2D

carbon/carbon exit cones has seen wide application in experimental evaluations of advanced systems (Chase, 1980) (Fig. 3). Typically, a 3D ITE (integral throat exit) is of cylindrical construction made up of layers of axial, radial, and circumferential carbon fiber bundles. This preform is then densified through repeated cycles of impregnation with pitch or other resins followed by carbonization/graphitization cycles. This process is repeated until the desired density is obtained. Different suppliers have used various fibers, preform construction methods, densification resins, and processing cycles resulting in materials with different mechanical/thermal properties.

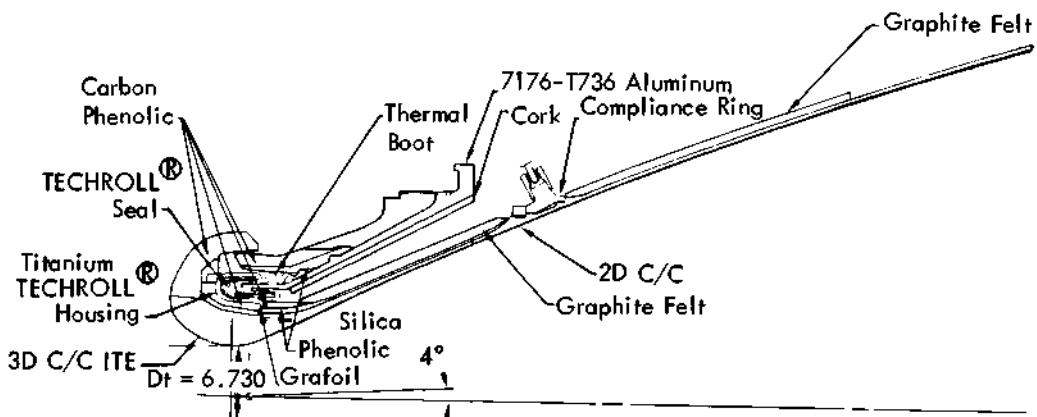


Fig. 3 Inertial upper stage nozzle

The 2D carbon/carbon exit cones used to date have been primarily of involute construction. Once the exit cone is layed up, usually in a female mold using strips of pre-impregnated carbon cloth material, densification proceeds in the same manner as the 3D material. Usually the 2D exit cones are not processed to as high a density as the 3D throats because of the less severe environment seen in the exit cone. The success in the use of 2D carbon/carbon exit cones has not been as dramatic as the 3D ITEs. The primary problem with 2D involute construction is the wrinkling and material distortions that occur during the debulk and cure cycles. These defects can later cause delaminations during densification, resulting in rejection of the part in the manufacturing process. If the parts do survive processing, these defects are normally the failure initiation points during firing. Even though problems have occurred, the advantages of simplicity in design, lighter weight, and potentially higher performance have made the 3D C/C ITE with a free standing 2D C/C exit cone a viable nozzle concept.

Thrust Vector Control Although numerous thrust vector control (TVC) methods such as gimbal ring, ball and socket, fluid bearing, secondary injection (gas or liquid), and mechanical deflectors have been successfully demonstrated and used in the past, the flex seal (Fig. 4) has been the mainstay for the last two decades in large motors. The flexseal requires relatively low actuation forces, can easily be vectored to angles of up to 7°, has demonstrated adequate stiffness requirements for stability, can withstand the motor thermal and pressure effects, and the simplicity of the system has led to low cost. Another advantage of the flexseal is that it can be checked out prior to use and be exposed to all operating environments except temperature. It is the above that has led to its high reliability in motor firings.

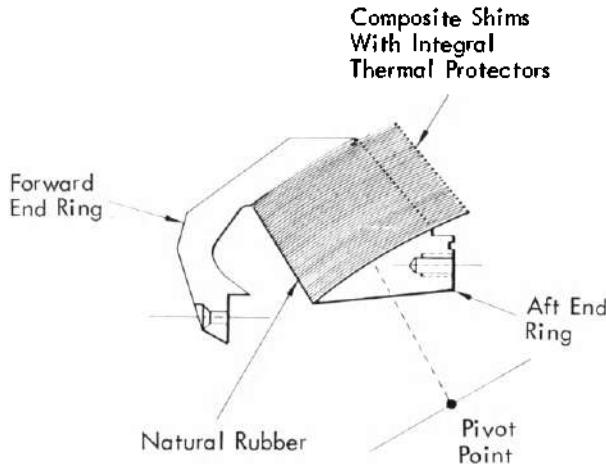


Fig. 4 Typical flexseal

**Extendable Nozzle Exit Cone.** Significant performance improvements, especially on upper stages, can be achieved through the use of the extendable nozzle exit cone (ENEC) concept. This concept basically allows the nozzle to be stored in a shorter length such that the motor can be made longer (adding more propellant) without violating the overall length constraints. It is the addition of propellant that provides the improvement in performance. A typical ENEC translating cone concept (Bruno, McCord, and Myers, 1979) is shown in the stored and deployed configuration in Fig. 5. To date, no ENEC system is in production, although numerous firings have been successfully conducted. Early success was achieved using convoluted metallic cones, but limitations due to temperature have essentially stopped further development. Tests have been conducted with the translating cone concept (both with 2D carbon/carbon cones and conventional materials), demonstrating the feasibility of the concept. Even though an ENEC is not in production, it is considered current state of the art because of its advanced development status.

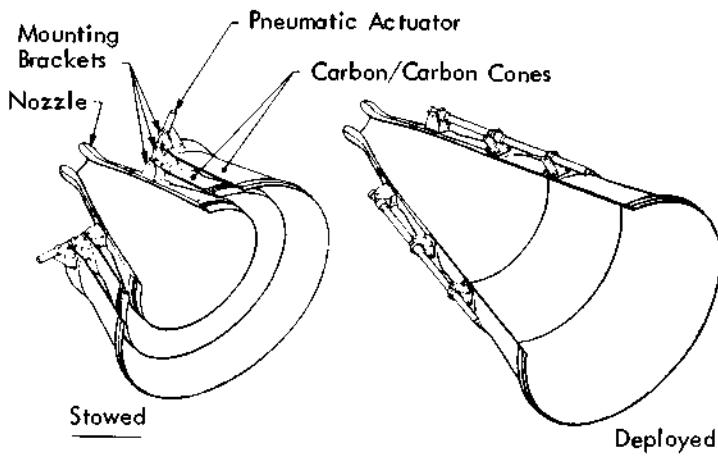


Fig. 5 Hercules ENEC concept

Growth Technology

Future nozzles will see expanded use of graphite epoxy composites replacing the remaining metallic structural components, such as stationary shells, flex seal end rings, compliance rings and other minor elements. Improvements in the 3D carbon/carbon area will come in cost reduction and better understanding of the material through full characterization. Further, structures will be tailored to specific requirements resulting in development of such concepts as 4D/5D construction, 3D contours woven to shape, and 3D/2D hybrids. Significant improvements in 2D carbon/carbon exit cone construction must be made to overcome the manufacturing and performance problems discussed earlier. Alternate methods of fabricating these cones with improved reliability are under investigation and include tape wrapping, braiding, filament winding, and 2D/3D hybrid construction. Some success has been achieved but much more development work will be required.

The performance gain of the ENEC is well recognized and can be further improved through the use of a gas deployed skirt (Carey and Nixon, 1977). This concept uses a folded metallic exit cone extension stored internally near the exit plane of the nozzle or ENEC (Fig. 6). During motor ignition the propellant gases deploy the skirt to the final configuration. Local static pressure retains the shape during motor operation.

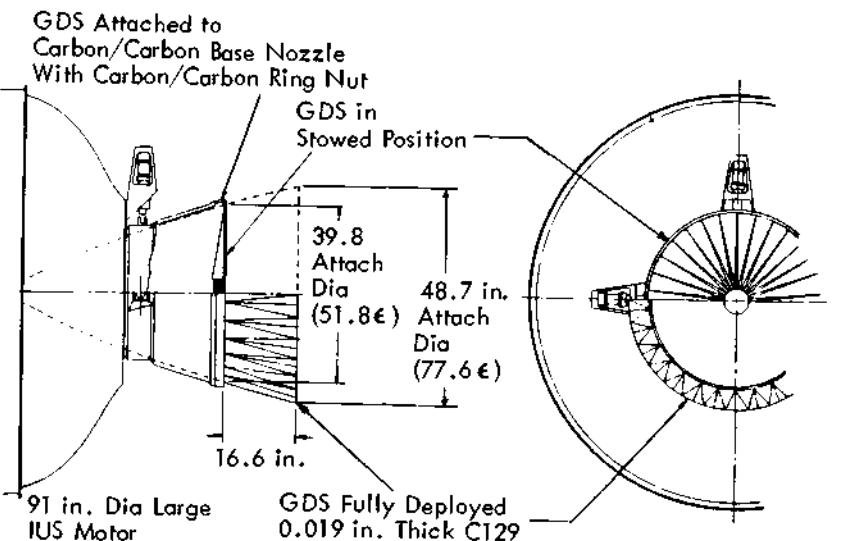


Fig. 6 Gas deployed skirt concept

In the area of TVC, further improvements in the flexseal can be made by elimination of the separate thermal barrier (a manufacturing headache), through the use of elastomers with better controllable properties, better aging characteristics, and lower shear modulus to reduce torque. A new TVC concept, the hot ball and socket (Ellis, Conner, and Kearney, 1979), has been test fired and shows promise for the future. This concept essentially combines the nozzle with the TVC resulting in a simple, light weight configuration (Fig. 7). Extensive use of advanced carbon/carbon materials are required to make this concept viable.

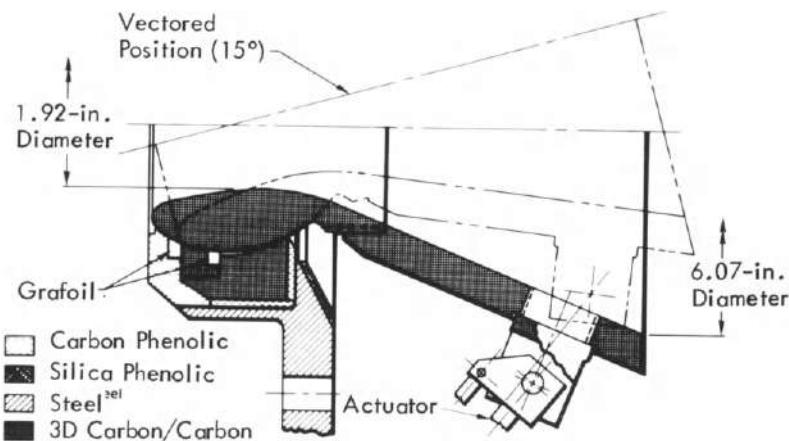


Fig. 7 Hot ball and socket TVC

#### CHAMBERS

Although steel and titanium will continue to be available and utilized, particularly for segmented cases, it is anticipated that the 1980s will witness a dramatic growth in the use of composite materials for motor chambers. Thus, the remainder of this section will be devoted to a discussion of activities in the areas of fibers, resins, design, and manufacturing trends.

#### Fibers

State-of-the-art chambers in the 1980s will likely be constructed of filament-wound composite materials using Kevlar and graphite fibers. Kevlar has been proven to be a tough and reliable fiber in several solid propellant rocket motors. In particular, its specific strength and toughness give Kevlar an edge over most other available fibers. Improvements to Kevlar fiber currently under investigation include fiber surface treatment methods and increased modulus. The surface treatment may help to improve composite shear strengths while an increased modulus would be useful in stiffness critical structures.

Graphite fiber is also finding practical uses in chambers. New high-strain AS4 graphite fibers are now competitive with Kevlar in terms of specific strength (tensile strength/density). Kevlar filament specific strength is  $7.58 \times 10^6$  in. while high strain AS4-graphite is  $7.68 \times 10^6$  in. The new AS4 graphite also recently out performed Kevlar by 9 percent in 6-inch pressure bottle tests and provided an average of PV/W of  $1.44 \times 10^6$  in. for a 3 sample test series versus  $1.3 \times 10^6$  for Kevlar 6" bottles. Currently, high-strain AS4 graphite is used in small motor chambers and is being considered as an alternate fiber for larger rocket motors. Although not as tough as Kevlar, graphite will nevertheless be a viable chamber material because of its high modulus, high specific strength, high shear strength, and competitive cost. If stiffness is important in a given application, graphite could be the choice because of its superior specific modulus (493 for graphite composite versus 351 for Kevlar composite).

RESINS

There are three resin types which will be used for filament winding of chambers: traditional epoxy resins with amine and anhydride cure, vinylesters and low density resins based on hydrocarbon elastomers like polybutadiene. Current practice uses epoxy based resin system. It has been shown that the resin should be tailored to fit the fiber requirements, for example a low modulus resin works best with non-released Kevlar pressure vessels while a higher modulus resin is superior with graphite. The epoxy resins are dependable and easily curable but tend to increase in density as modulus increases. For example, a 20,000 psi modulus resin has a density of 1.09 g/cc and a 400,000 psi resin modulus has a density of 1.23 g/cc. Thus, a substantial weight penalty is paid for the stiffer resin. To overcome the weight penalty, vinylester and polybutadiene resins which give higher moduli with lower densities, are now being studied. Preliminary studies have shown these resins to be very competitive in performance to the low density epoxy resins and will be considered in future large scale systems. In addition, brominated epoxies are likely candidates for use in situations requiring greater strength in high temperature aeroheating and launch environments. The brominated epoxies would be useful in space vehicles which experience significant flame exposure during launch and radiant heating in space.

More durable and higher strength preimpregnated resins will undoubtedly be used in future chambers for hand layup reinforcements. The preimpregnated resins will be better matched to the winding resin. Handling procedures need improvement and storage life should be increased so as to reduce waste and improve composite fabrication.

Design Features

Chamber designs will use improved computer software to achieve more efficient dome shapes and composite lay-ups. Higher order finite elements incorporating rigorous bending, differential stiffness, and composite properties will be in routine use. Work is well underway to adapt Computer Aided Design and Manufacture (CADAM) capability to the construction of finite element grids. Design grids are constructed on a CRT screen and directly transformed into the appropriate finite grid code. A finite element analysis is then conducted. Analytical results in the form of deflection plots, stress/strain plots and isoplots may be displayed on the CRT screen and/or hard copies obtained. Presently, 2D axisymmetric analyses are conducted using the above mentioned system. Research is now underway in which the CADAM grid generator may be linked with a 3D finite element structural program with results being displayed on color graphics. Specific modeling improvements which will be incorporated into the design mode include more accurate modeling of filament winding, hand layup modeling and discontinuity modeling. The end result is a more structurally sound dome and overall chamber design.

New composite design features may result from the improved fibers and resins. These new features may include larger aft port openings due to increased shear and fiber allowables and improved reinforcement methods. An additional area of improvement includes the introduction of composite reinforcements in place of the currently used aluminum.

Manufacturing Improvements

Key manufacturing improvements will center around computer controlled operations and improved winding techniques. In the future, a direct link will be created between the filament winding microprocessor and design computer eliminating the need for separate data dump computer runs and mechanical transfer of information via disks.

Improved winding techniques are under study which may use lighter weight winding mandrels, improved fiber/resin delivery systems, and improved composite machining. Light weight mandrels are easier to handle, require lower torque in winding and are usually easier to disassemble. Both inflatable mandrels and net metal mandrels are under consideration. Improved fiber/resin metering systems are being constructed which minimize fiber damage and increase resin wetting and efficiency. Volumetric metering pumps are being used to blend and mix the resins right at the delivery head of the winding machine. More roller surfaces are also being incorporated into the delivery system to replace friction combs and rubbing surfaces. Improved composite machining is also anticipated. The water jet concept has been tested and is currently being used on a limited basis. Cleaner and more accurate cuts are available with the water jet cutting.

Standard lathe and mill machining improvements are also possible using coolants and lubricants which provide a better cutting surface with carbide or high speed steel cutting tools.

In general, the SOTA will include increased use of computers in the areas of structural design, modeling and analysis with direct linkup to the filament winding microprocessors. The composites in the chamber will utilize Kevlar or graphite fiber and resin low density combinations which will deliver maximum efficiency in the pressure vessel.

#### INSULATION

In considering the goals for improvements in the state of the art, the following appear to be possible in the 1980s.:

- Increased performance - This includes reduced weight and improved low temperature and erosion properties
- Simplification in the manufacturing process

For the first part of this decade, internal insulators for both tactical and strategic rocket motors will continue to be fabricated primarily from filled ethylene propylene diene monomer (EPDM) formulations. This is a result of EPDM's many excellent properties. It has a good low density, yields excellent mechanical properties even when filled, and is processable by conventional rubber fabrication equipment. Development studies are underway to improve processability of EPDM.

While EPDM will likely remain as a key binder polymer in new internal insulator formulations, other polymers such as Hypalon, Styrene/butadiene, and polybutadiene/acrylonitrile will receive increased attention. The interest in new polymers is related to the desire to improve mechanical properties at low temperatures for

motors using composite cases, to reduce the migration of plasticizers and other mobile species, and to simplify the case bond systems.

In addition to new polymers, which will be used both with and without EPDM, new fillers will be developed and incorporated. A major feature of the drive to develop new fillers is the desire to replace asbestos, an established carcinogen. New fillers currently under investigation include Kevlar, graphite, ground phenolic resin (cured), cotton and clay.

The development of improved insulators, for the coming decade, is also taking advantage of materials that suppress combustion such as halogenated hydrocarbons and certain boron based materials, and those that promote char formation and retention such as the phenolic and melamine resins and the polysaccharides. The thrust of these efforts is to improve the insulator's performance to the extent that lighter insulators can be used in future motors.

In addition to developments in new materials, improved methods of insulator manufacture will be actively pursued. The current methods of manufacture (i.e., layup and grind, match metal die compression) are labor intensive, require expensive tooling, and severely lack the flexibility to change a given design rapidly. Significant savings in time and money can be realized if manufacturing methods can be improved. The approaches now being developed are based on advanced winding technology wherein the insulator is wound directly on the chamber winding mandrel. After the insulator is wound and cured, it is machined to the proper dimensions if necessary and the case then wound over it. These approaches will yield considerable savings in insulator cost, fabrication time, and design flexibility. These new manufacturing methods are directly related to the development of new insulator materials. They require materials that are processable, compatible with the case winding resins and fibers, non-carcinogenic, and inert to the propellant.

#### CONCLUSION

Components exist today for the development of reliable high performance solid propellant rocket motors. Design techniques and testing have reached a state of maturity that supports development on predictable schedules with significantly fewer ground tests than previous generations required.

During the 1980s modest improvements in propellant performance are anticipated. Substantial innovations in nozzle materials and the reduction of risks associated with extendable nozzle exit cones will be forthcoming. Filament winding techniques will be extended to largest space motors utilizing improved Kevlar, graphite, other fibers and new low density, flexible resin systems. Insulation processing will be substantially modernized to meet improved reliability and cost reduction goals. Figure 8 summarizes the propellant performance and inert components improvements made in past decades and shows the potential for the 1980s.

The 1980s should continue to see applications of solids in the space programs, where payloads invariably grow during launch vehicle development and additional boost impulse is required. The use of staged solid rocket motors to place payloads in orbit, either from a ground launch or from the Space Shuttle, will expand and establish new challenges to be met by the designer.

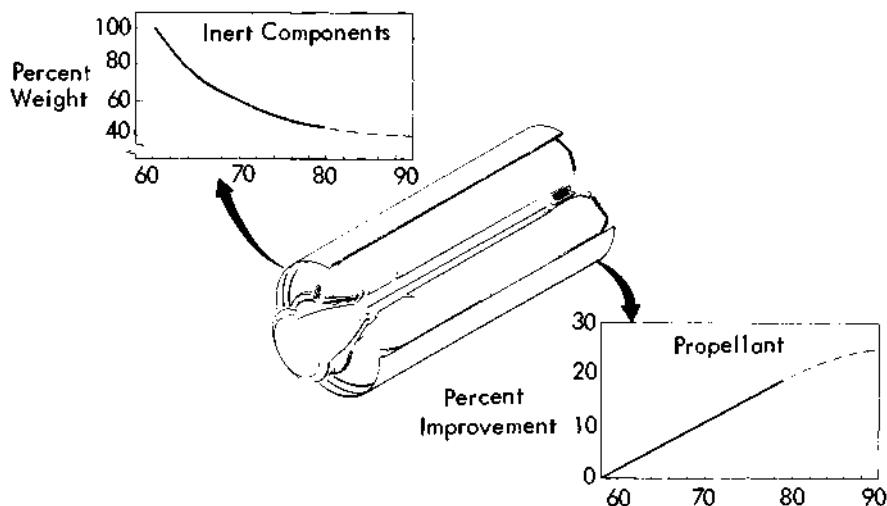


Fig. 8 Propulsion technology trends

#### REFERENCES

- Bruno P. S., J. L. McCord, and L. F. Myers (1979), Hercules Incorporated, Carbon/Carbon, Nested Cone, Extendible Nozzle Development.
- Carey L. and G. Nixon (1977), Gas Deployed Skirt for Rocket Motor Nozzles, AIAA 13th Propulsion Conference.
- Chase C. A. (1980), United Technologies, Chemical Systems Division, Development Status of the Inertial Upper Stage Solid Motors, AIAA Conference.
- Ellis R. A., G. E. Conner, and W. J. Kearney (1979), United Technologies, Chemical Systems Division, and J. W. Wilson, Fiber Materials, Hot Ball and Socket Integral Nozzle and TVC System.
- NASA SP-8115 (1975), Solid Rocket Motor Nozzles, NASA Space Vehicle Design Criteria (Chemical Propulsion).

## LE-5 OXYGEN-HYDROGEN ROCKET ENGINE FOR H-I LAUNCH VEHICLE

H. Nakanishi\*, E. Sogame\*, A. Suzuki\*\*, K. Kamijo\*\*,  
K. Kuratani\*\*\* and N. Tanatsugu\*\*\*

\*National Space Development Agency of Japan (NASDA), Japan

\*\*Kakuda Branch, National Aerospace Laboratory (NAL), Japan

\*\*\*Institute of Space and Astronautical Science (ISAS), Japan

### ABSTRACT

The LE-5 is the first cryogenic engine developed in Japan. It was designed as a second stage of a future launch vehicle named H-I. The H-I vehicle is a three stage rocket capable of placing a payload of 550 kg into a geosynchronous orbit.

The LE-5 engine is a gas generator cycle, medium combustion pressure engine with rated thrust of 10 tons in vacuum. Propellants are liquid oxygen and liquid hydrogen. Propellant mixture ratio can be controlled in three steps by means of two bypass valves. Restart capability is provided by a unique method in which hydrogen gas bled from a combustion chamber is used to spin up turbopumps.

NASDA, in cooperation with NAL, has completed the development of the LE-5 prototype engine. Firing tests were successfully carried out from March to July, 1981. ISAS has developed its own prototype engine to back up NASDA/NAL efforts. ISAS engine tests were also successful.

This paper presents the design and the development status of the LE-5 engine and its components. Also described are the engine system of ISAS and its test results.

### KEY WORDS

launch vehicle, cryogenic engine, liquid hydrogen, liquid oxygen, Japanese space program, H-I launch vehicle, LE-5 engine

### INTRODUCTION

At present, 350 kg class geosynchronous satellites are used in space applications programs in Japan. However in late 1980s, heavier satellites will be required by users. To meet these requirements, the National Space Development Agency of Japan (NASDA) is going to develop a launch vehicle called H-I.

The H-I vehicle is a three stage rocket consisting of a flight proven N-II first stage, a newly developed cryogenic second stage, and a solid third stage (Fig. 1). It is capable of launching a payload of 550 kg into a geosynchronous orbit, in contrast to 350 kg of the present N-II vehicle. The high launch capability of the H-I is attained by a liquid oxygen (LOX)/liquid hydrogen (LH) second-stage engine called LE-5.

NASDA initiated the development of the LE-5 engine in 1977, in cooperation with the National Aerospace Laboratory (NAL) and the Institute of Space and Astronautical Science (ISAS). Starting from virtually no experience in LOX/LH propulsion, NASDA has completed the LE-5 prototype engine in less than five years. NAL has developed some components including a LOX turbopump, and also supplied technical data based on in-house testing. ISAS is working on its own engine system which could be used as an alternative to the NASDA/NAL engine. The turbopump arrangement of the ISAS engine is unique in that LOX and LH turbopump rotors are mounted in a single housing and rotate in opposite directions.

Having completed the prototype engine, NASDA is now engaged in the development of the flight type engine. Now close to completion is a high altitude test stand, with which full-duration simulated-altitude firing tests can be carried out. Qualification tests of the engine will be finished by the end of 1984. The first flight of the H-I vehicle is scheduled for early 1986.

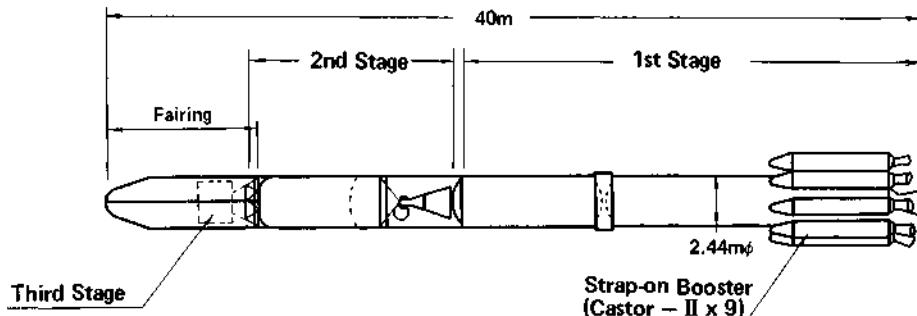


Fig. 1 H-I Configuration

## LE-5 ENGINE DESCRIPTION

### Engine System

The LE-5 is a pump-fed engine with rated thrust of 10 tons in vacuum. Combustion pressure is  $35 \text{ kg/cm}^2 \text{ a}$ . The main characteristics of the engine are shown in Table 1. An external view and a schematic diagram are given in Fig. 2 and 3, respectively. A gas generator is used to drive separate LOX and LH turbopumps. Hot gas from the gas generator first drives the LH turbopump and then the LOX turbopump. Two bypass valves in the hot gas line control propellant mixture ratio in three steps, responding to commands from a guidance computer. The engine thrust is kept at a fixed level.

The main combustion chamber is regeneratively cooled by hydrogen. Part of hydrogen is bypassed and used to pressurize the LH tank and also to cool the nozzle extension, which is attached to the combustion chamber to expand the nozzle area ratio from 9 to 140. A heat exchanger is installed in the turbine exhaust gas manifold. Helium gas from cryogenic spheres in the LH tank is warmed to 90 K by the heat exchanger in order to pressurize the LOX tank. A small turbine is mounted in the hydrogen line to drive a hydraulic pump.

Table 1 LE-5 Characteristics

Thrust (vacuum)	10 ton
Specific Impulse (vacuum)	450 sec
Mixture Ratio	5.5
Weight (dry)	230 kg
Maximum Diameter	1.65 m
Length	2.7 m
Nozzle Area Ratio	140
Combustion Pressure	$35 \text{ kg/cm}^2 \text{ a}$
LH Turbopump	
Shaft Speed	48500 rpm
Discharge Pressure	$56 \text{ kg/cm}^2 \text{ a}$
LOX Turbopump	
Shaft Speed	16500 rpm
Discharge Pressure	$52.5 \text{ kg/cm}^2 \text{ a}$
Gas Generator	
Combustion Pressure	$26 \text{ kg/cm}^2 \text{ a}$
Mixture Ratio	0.9

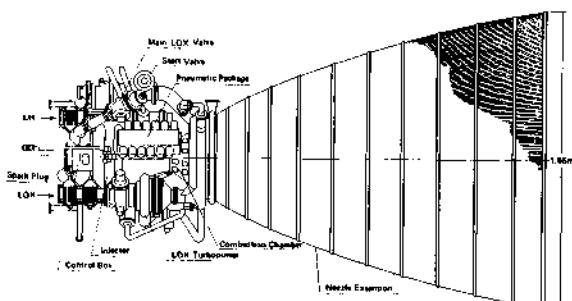


Fig. 2 LE-5 Engine

A unique method is used to start up the engine. Upon receipt of a start signal, the combustion chamber igniter is fired first and then the main propellant valves are opened. Propellants flow into the combustion chamber and burn at low pressure. Liquid hydrogen, which cools the chamber wall, is heated conversely and gasified. Part of gaseous hydrogen is introduced into the two turbopumps through the start valve to drive them. Pump discharge pressure and consequently combustion pressure begin to rise. When adequate pressure level is reached, the start valve is closed and the gas generator is fired. Then the engine bootstraps itself quickly to the nominal operating point. In this method, a start tank or a solid starter is not needed and the system is simplified.

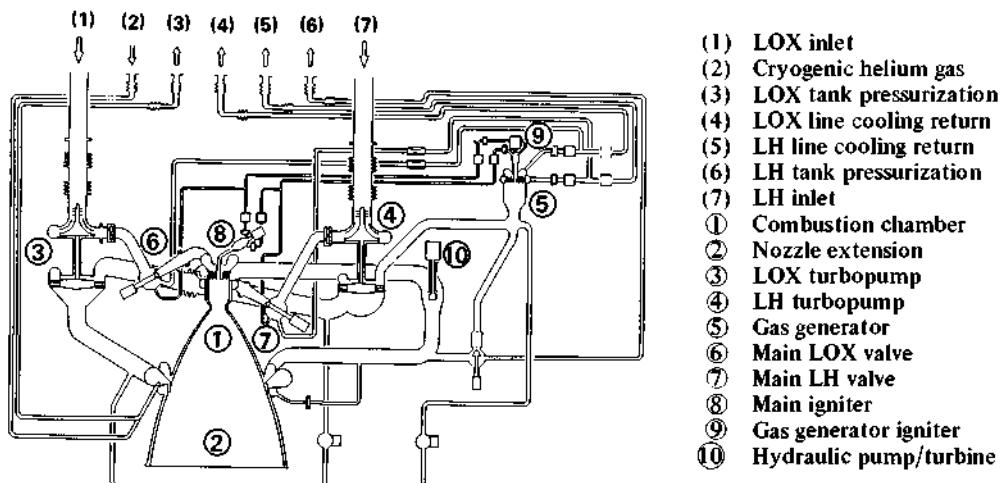


Fig. 3 LE-5 Engine Schematic Diagram

### Major Components

Both the combustion chamber and the nozzle extension are of conventional tubular design, made from 240 double-tapered and 650 single-tapered tubes, respectively. The tapered tubes are formed into their final shape and joined together by means of two-step furnace brazing. The injector consists of 208 conventional coaxial elements. The faceplate of the injector is made of porous material. Next to the injector are cavities in the combustion chamber wall to suppress high frequency oscillation.

The LOX turbopump is a single-stage centrifugal pump with a helical inducer, driven by a two-rotor velocity-compound turbine (Fig. 4). Two ball bearings are cooled by liquid oxygen. Balance holes are drilled in the impeller to reduce the axial force on the rotor. A face-contact metal-bellows seal and circumferential hydrodynamic seals are used to separate liquid oxygen from turbine drive gas. The LH turbopump has a similar configuration with the LOX turbopump (Fig. 5). Two ball bearings are cooled by liquid hydrogen. The axial force on the rotor is compensated by a series-flow balance piston integral with the impeller. The shaft sealing is simpler than that of the LOX turbopump. The gas generator is a cylindrical chamber with a conventional coaxial-element injector.

Each of the main and the gas generator igniters is a torch with a electric spark exciter. All the propellant-line valves are of poppet type and actuated by helium gas. The same valves are used in all the lines leading to the gas generator and the igniters. The two bypass valves are four-position poppet valves. Helium gas for valve actuation is controlled by solenoid valves, which are compactly mounted in a pneumatic package and act in accordance with signals from a micro computer called engine control box.

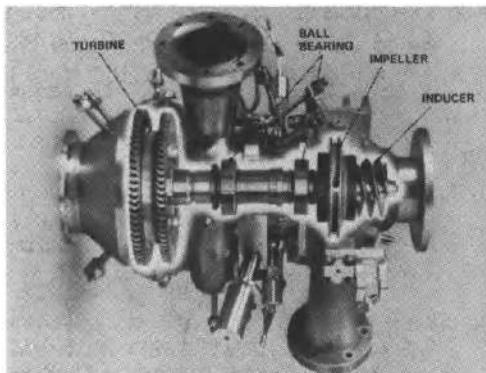


Fig. 4 Liquid Oxygen Turbopump

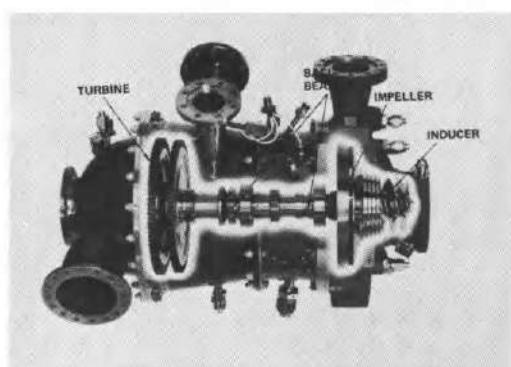


Fig. 5 Liquid Hydrogen Turbopump

### LE-5 ENGINE DEVELOPMENT STATUS

#### Component Tests

Development tests of major components carried out by NASDA or by NASDA/NAL are as follows; water cooled chamber hot firing (1978–79), regeneratively cooled chamber hot firing (1979–81), gas generator hot firing (1978–81), LH pump and turbine testing (1978–79), LH turbopump testing (1979–81), LOX turbopump testing (1980–81), turbopump system tests (1980–81) (Fig. 6). The LOX turbopump was first developed by NAL and then transferred to NASDA. Difficulties met during these tests include shaft oscillation, leakage and outside burning of hydrogen, damage of various parts by combustion gas, misfiring at a start sequence, and ice forming on cold surfaces exposed to combustion gas. These problems were resolved by changing the test procedure or by modifying the design.

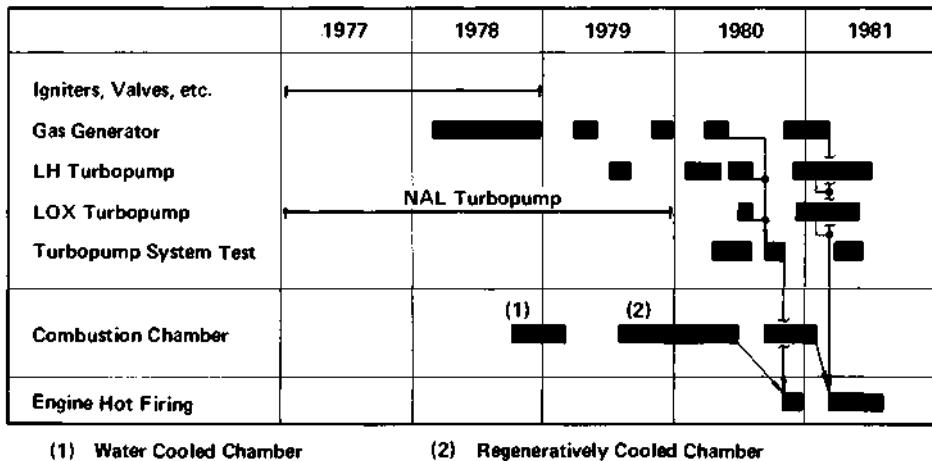


Fig. 6 Prototype Engine Development Tests

### Prototype Engine Tests

Engine system tests were initiated in December, 1980. At first, each component was mounted on separate test beds and connected with pipelines. Continuous firing of 120 seconds was accomplished in this set-up, and the compatibility of the components were verified. Then, prototype engines were fired from March to July, 1981 (Fig. 7, 8). Eighteen tests were carried out for total duration of 1680 seconds, including three tests of continuous 370 seconds. Data are shown in Fig. 9 and 10. All the tests were successful except for a few early ones, where premature cuts off occurred due to minor malfunctions of an engine component or a test equipment. Estimating from the characteristic velocity shown in Fig. 10, specific impulse of over 450 seconds will be attained at flights.

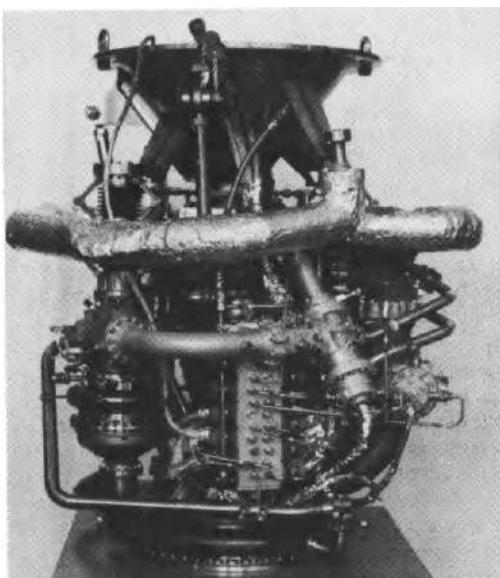


Fig. 7 Prototype Engine

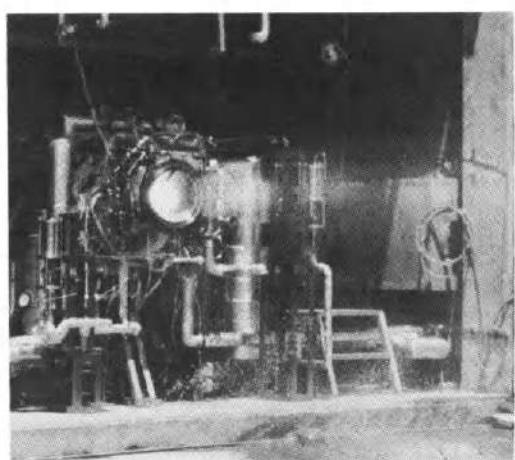


Fig. 8 Prototype Engine Hot Firing

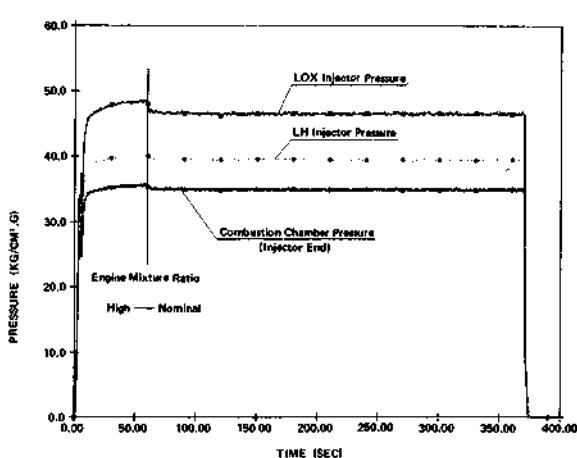


Fig. 9 Typical Test Data

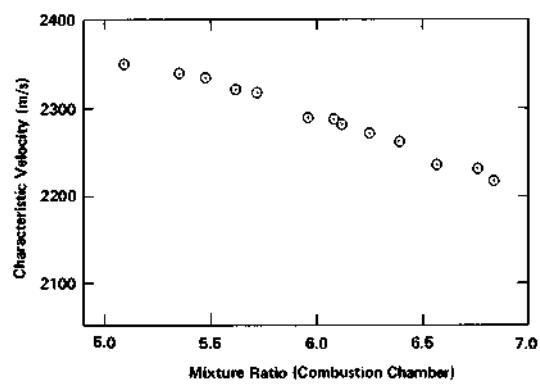


Fig. 10 Characteristic Velocity

### Flight Type Engine Development

The development schedule of the flight type engine is shown in Fig. 11. Now close to completion is a high altitude test stand (Fig. 12), with which full duration simulated altitude firing tests can be carried out. During engine firing, the vacuum cell pressure is maintained at 10 torr by a self-ejecting diffuser and a two-stage steam ejector. The first series of tests are scheduled for April, 1982, using a prototype engine and a nozzle extension. The fabrication of the nozzle extension (Fig. 13) was completed in June, 1981, following the successful development of tapered tubes and brazing alloys. Simulated altitude tests of flight type engines are scheduled from October, 1982.

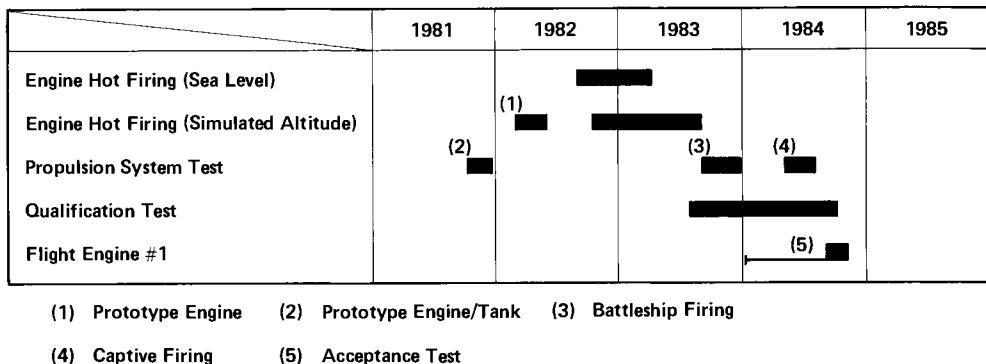


Fig. 11. Flight Type Engine Development Schedule



Fig. 12. High Altitude Test Stand  
 (Under construction)



Fig. 13. Nozzle Extension

Propulsion system tests will be carried out using a newly constructed vertical stand. A prototype engine will be combined with a prototype tank and tested in November, 1981, in order to verify the interface between the engine and the tank at an early stage. Battleship and captive firing tests are scheduled for 1983 and 1984, respectively.

Qualification tests of the engine will be completed in 1984. The first flight unit will be delivered in mid 1984. The first flight of the H-I vehicle is scheduled for early 1986.

## LE-5 RELATED RESEARCH AND DEVELOPMENT AT NAL

### Background

The liquid rocket propulsion research at the National Aerospace Laboratory (NAL) has been conducted at its Kakuda Branch. Since around 1977, the research on oxygen-hydrogen rocket engine at NAL has been directed to support the LE-5 development program. A close cooperation with NASDA, for developing turbopump systems and for acquiring various technical data, has been working very successfully.

The LE-5 related works at NAL are outlined as follows.

### Thrust Chambers

In relation to design and the evaluation of LE-5 thrust chamber, studies on the combustion performance, cooling and high altitude performance were conducted.

As the studies on cooling and combustion performance, the following investigations were performed.

- Experimental investigation of heat transfer characteristics of liquid hydrogen.
- Experimental investigation of the effect of injection parameters on the characteristic velocity efficiency and the heat flux to chamber wall.
- Analytical investigation of the combustion performances.

In order to evaluate many previous heat transfer correlation equations of liquid hydrogen, Joule heated tube experiments were carried out. As the results, Hess & Kunz's equation was considered best for design purpose, and for estimating the effect of surface roughness, Dipprey & Sabersky's equation was sufficient even in the near-critical region. The effect of curvature was estimated through experiments with a curved tube simulating the chamber envelope. A typical example is shown in Fig. 14 which illustrates enhancing or degrading effects of the curvature on the heat transfer rates. In the throat region, an evident delay phenomenon of the enhanced rates toward downstream was observed. These results were verified by firing tests of a subscale, hydrogen-cooled combustor.

A number of firing tests with a 1/20 thrust, subscale combustor were performed investigating the effects of parameters, as the propellant flow rate per injector element, the oxidant-fuel ratio and hydrogen injection temperature, on the characteristic velocity ( $C^*$ ) efficiency or the axial heat flux distribution. Observed  $C^*$  efficiencies were correlated as a function of the hydrogen-oxygen injection velocity ratio as shown in Fig. 15.

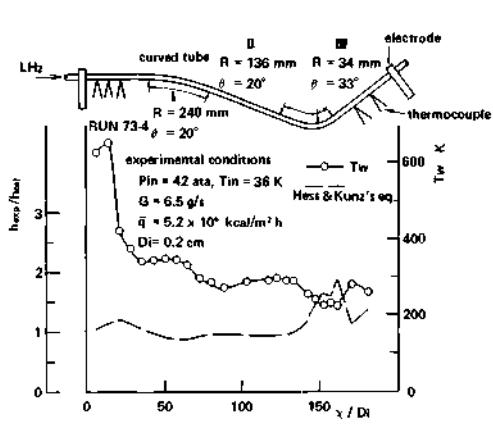


Fig. 14. Heat transfer characteristics along a curved tube

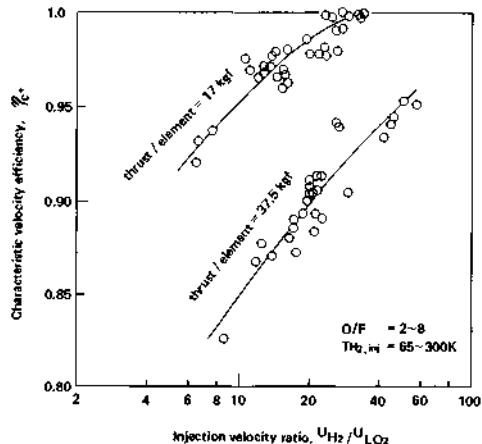


Fig. 15. Characteristic velocity efficiency as a function of injection velocity ratio  
(Combustion chamber; L\* = 0.73m, g = 5.56)

As an analytical approach, a semi-empirical calculation model was developed. The model assumes equations governing the spray dispersion and the droplet vaporization process, and coefficients which meet the measured pressure distributions at subscale combustor tests. Estimated C\* efficiency of LE-5 thrust chamber by this model agreed with experimental values within  $\pm 2\%$ .

As the study on the high altitude performance of the thrust chamber, the following analysis and testing were performed.

- Demonstration of the obtainable specific impulse level by a 1/5 scale nozzle and a prediction of the performance of the full scale nozzle.
- Development of a nozzle flow analysis computer code.
- Analysis and testing on the altitude ignition performance of the prototype torch igniter.
- Analysis and model testing of the exhaust system for the engine altitude performance simulation.

Fig. 16 shows the experimental vacuum specific impulse in comparison with curves calculated by NAL's performance analysis code. The accuracy of the prediction seems to be generally within  $\pm 1\%$ . If conservative values are used for the turbine exhaust loss and the damp cooling loss, the delivered specific impulse of LE-5 is calculated to be 447 sec at the nominal operating condition.

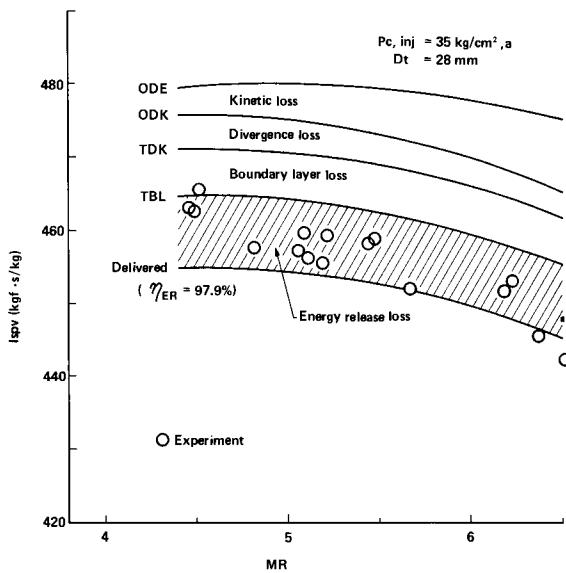


Fig. 16 High altitude performance of a subscale thrust chamber

The main feature of NAL's nozzle analysis computer code is the incorporation of the finite difference method with a concept of shock capturing for the analysis of two dimensional kinetic flow. This facilitates the analysis of real nozzle flows which often involve various kind of waves.

It was shown that the torch igniters for the main chamber and for the gas generator performed excellently over the range of the design conditions. However, an altitude effect, i.e., a non-ignition-region was observed to exist at the reduced environmental pressure.'

A 7.35% scale H<sub>2</sub>/O<sub>2</sub> thrust chamber was used for the model testing of the exhaust system to show the acceptability and the consistency of all the important design values of the supersonic diffuser system, the spray cooling system, and the steam ejector system.

### Turbopumps

The study and development of turbopumps have been conducted for years at NAL. However, the development and testing of LE-5 turbopumps were performed jointly with NASDA. The major contributions of NAL were:

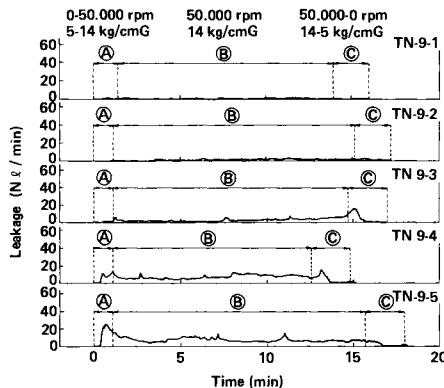
- Study and development of LOX turbopump system for LE-5.
- Study on the suction performance of cryogenic pumps.

- (c) Study and development of self-lubricated ball bearings and shaft seals for liquid hydrogen pumps.
- (d) Testing of the integrated LE-5 turbopump systems.

The prototype, LE-5 LOX turbopump which was developed by NAL is shown in Fig. 4. It consists of single-stage centrifugal pump with three-bladed swept-back inducer and single-stage, two-rotor, velocity-compounded impulse turbine. Main shaft seals are composed of a face contact, metal-bellows seal, a circumferential, turbine hot gas seal and a helium-purged, double circumferential seal. The circumferential seals are of hydrodynamic type with segmented Rayleigh step pads.

In order to investigate the suction performance, pump inducers were tested in water, liquid nitrogen and liquid hydrogen. Test results were applied to design of inducers of LE-5 turbopumps.

A number of tests were performed for self-lubricated ball bearings with glass fabrics-supported Teflon retainers, and face contact, metal-bellows mechanical seals. A mechanical seal, equipped with a newly developed damper and an improved carrier, operated smoothly with very small leakage for fairly long duration as shown in Fig. 17.



**Fig. 17. Seal performance of a mechanical seal**

The full-power, closed-loop operation of the integrated LE-5 turbopump systems was successfully performed in 1980, and the function of power control for a fixed engine thrust and given mixture ratios was validated.

### Engine Systems

A computer program was developed in order to evaluate start transient characteristics of LE-5 which adopted hydrogen bleed cycle as the start mode. A part of this program was used to investigate the turbopump start build-up, in which gas generator ignition troubles due to insufficient pre-conditioning of propellant feed lines were experienced.

## ISAS ENGINE DESIGN AND DEVELOPMENT

### Background

The Institute of Space and Astronautical Science (ISAS) has been engaged in the development of LH/LOX rocket propulsion system and supported the national development program of H-1 launch vehicle.

The development study has started at ISAS in 1975 according to its own program. By 1980 have almost finished the development tests of major components (thrust chamber, turbopump, gas generator and turbine spinner) for the engine assembly. The details on these components were summarized in the paper presented at the 31st IAF congress, Tokyo in 1980.

In early 1980 the engine system has been integrated and then the verification tests have been successfully conducted at Noshiro Testing Center of ISAS. The performance capability and functions of the integrated engine have been verified step by step through three kinds of tests as follows:

- (1) the start up test (in July, 1980)

- (2) the test operated at 60% of rated power (in September, 1980)
- (3) the rated operation test (in December, 1980)

### Engine System

#### General Description

The engine system called ES-702 is designed for an upper stage of H-1 or Mu vehicle which are representative launch vehicles in Japan. A photograph of the engine assembly is shown in Fig. 18. The ES-702 is fed by a turbopump which operates in accordance with the traditional gas generator cycle and generates a rated thrust of 7 ton in vacuum. The hydraulic and pneumatic diagram of the LH/LOX propulsion system to be developed in ISAS is illustrated in Fig. 19. The engine system is shown below a dot dash line in Fig. 19. The engine consists of a thrust chamber, LH and LOX turbopump, a gas generator, a turbine spinner, a heat exchanger, a propellant utilization valve, a pneumatic package, valves and sensors.

The thrust chamber is composed of 200 double tapered stainless steel tubes braised together with outer shell. The injector consists of 90 coaxial elements and a porous material faceplate for transpiration cooling. The turbopump has a very unique configuration to have never been developed for rocket engine. LH and LOX pumps and turbine are mounted on a same axis within a single housing. LH and LOX pumps are arranged at each end and turbine at center of the housing. The turbine consists of single stage nozzle and two rotors. The rotors are mounted on two separate shafts respectively and rotate independently in opposite direction of each other. A centrifugal type of LH pump is mounted on a first stage turbine rotor and a same type of LOX pump on a second stage rotor. The gas generator to drive the turbine is of a spherical reverse-flow type. The turbine spinner for start up is of a solid propellant cartridge type. During the burn period, the LOX tank is pressurized by flowing liquid oxygen through the heat exchanger in turbine exhaust duct. The LH tank is pressurized during the burn period by regeneratively heated hydrogen gas bled directly from the manifold downstream of the thrust chamber cooling passage.

The propellant utilization control (to control the ratio of LH and LOX in tank) is performed by returning liquid oxygen from the discharge side of the LOX pump to the tank through a three position control valve (full open, nominal open and shut off). Nominal open position provides the rated flowrate of liquid oxygen. The engine valves are controlled by a pneumatic system powered by gaseous helium.

#### Engine Integration

The turbopump is mounted by the thrust chamber in parallel to the stage centerline. The LOX pump volute casing is fixed to the thrust chamber injector head with a pin joint. The LH pump volute casing is supported to the thrust chamber throat in order to prevent side way motion but allow thermal expansion in the direction of axis of turbopump. The LH pump is installed in lower side of turbopump assembly and therefore its suction line forms a "U" letter

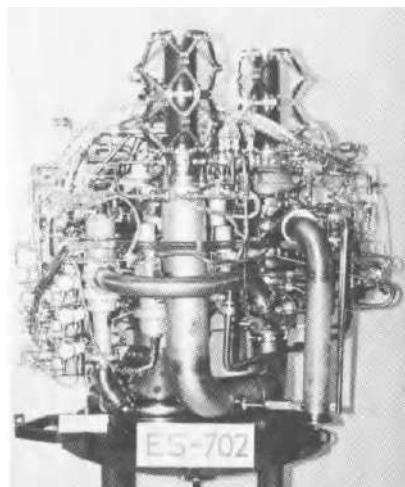


Fig. 18. Engine Assembly Developed in ISAS

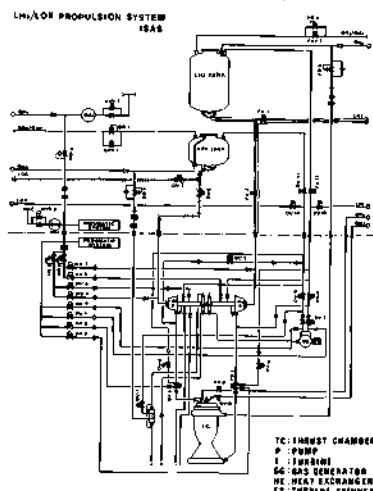


Fig. 19. Hydraulic Diagram of Propulsion System

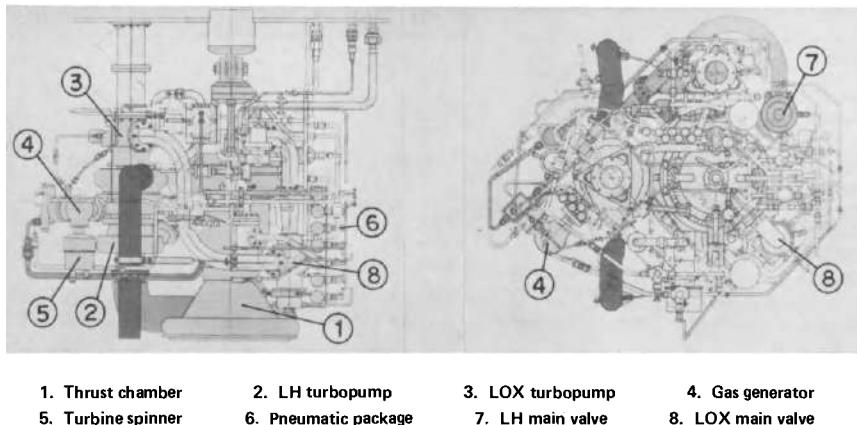


Fig. 20 Schematic of LH/LOX Engine Assembly Developed in ISAS

shape. The gas generator is mounted directly on the turbine nozzle manifold.

The engine assembly is shown schematically in Fig. 20.

#### Engine Start Up System

The turbine starts to rotate by a solid propellant cartridge which also initiates combustion in the gas generator. The start valve of gas generator opens when the discharge pressure of LH and LOX pumps rise adequately beyond the pressure in the gas generator chamber due to the combustion of turbine spinner and then the engine system "bootstraps" itself to steady state conditions. Until the start valve opens, the injectors are purged by gaseous helium in order to prevent contamination of the injectors by the combustion products of solid propellant turbine spinner. The start valve is of a poppet type. The LH and LOX valve elements are mechanically linked and controlled with a single actuator.

#### Test Results

ISAS has conducted the final firing test of ES-702 engine at Noshiro Testing Center of ISAS in December 1980. The feature of the final test is presented below.

The turbine spinner burned for 2 seconds at the beginning of start up transient.

The gas generator start valve opened 0.7 second after start up and then the system "bootstrapped" itself to the steady state conditions, which was attained within 2 seconds after start up. The PU control valve was operated at the steady state condition. The picture of firing test is shown in Fig. 21. In Table 2 are shown the test data obtained in the final test under sea-level conditions.

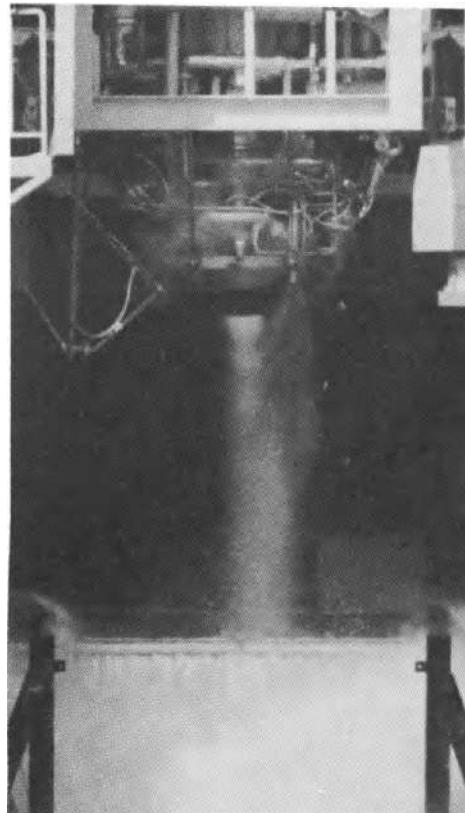


Fig. 21 Picture of Firing Test

PU control valve		full open	nominal open	shut off
<b>Engine</b>				
Propellant flow rate	kg/s	13.0	14.4	15.4
Mixture ratio		4.16	4.59	4.80
Specific impulse	sec	286.8	284.1	285.7
<b>Thrust Chamber</b>				
Thrust	kg	3745	4080	4389
Propellant flow rate	kg/s	12.7	14.0	15.0
Mixture ratio		4.35	4.82	5.04
Chamber pressure	kg/cm <sup>2</sup> a	20.	21.3	22.4
Specific impulse	sec	294.	291.1	292.5
<b>Turbopump</b>				
Rotational speed (LH/LOX)	rpm	36460/12900	37630/13500	38980/13950
Discharge pressure	kg/cm <sup>2</sup> a	33.8/24.5	36.1/26.4	37.9/28.3
<b>Gas Generator</b>				
Propellant flow rate	kg/s	0.238	0.253	0.270
Mixture ratio		0.685	0.729	0.764
Chamber pressure	kg/cm <sup>2</sup> a	20.2	21.7	23.0

Table 2. Test Results Under Seal-level Conditions

## CONCLUDING REMARKS

Having successfully completed the prototype engine tests in early 1981, NASDA, along with NAL and ISAS, cleared a difficult hurdle in the development of the LE-5 engine. Being the first cryogenic engine in Japan, the LE-5 not only contributes to the high launch capability of the H-I rocket, but also opens a door to future engines. Many advances in propulsion technology are being realized through the development of the LE-5. Now under study are such engines like a small light-weight engine for upper stages and a high-pressure high-performance engine for a booster.

## ACKNOWLEDGEMENT

The authors would like to express their gratitude to the staff of NASDA, NAL, ISAS, Mitsubishi Heavy Industries Co., Ltd., Ishikawajima-Harima Heavy Industries Co., Ltd., and other organizations which take part in the development of the LE-5 engine.

## REFERENCES

1. Nakanishi, H., Tomita, B., Sogame, E., "LOX/LH<sub>2</sub> Propulsion System for H-I Launch Vehicle", IAF-80-F-244, 1980.
2. Tanatsugu, N., Nagatomo, M., Akiba, R., Kuratani, K., "Development Study of LOX/LH<sub>2</sub> Propulsion Systems", IAF-80-F-246, 1980.
3. Kamijo, K., Sogame, E., Okayasu, A., "The Development of Small Liquid Oxygen and Liquid Hydrogen Turbopumps of a LE-5 Rocket Engine", AIAA-81-1375, 1981.

## THRUST CHAMBER TECHNOLOGY FOR ADVANCED EARTH-TO-ORBIT ROCKET ENGINES

R. J. LaBotz and R. L. Ewen

*Aerojet Liquid Rocket Co., Sacramento, CA, USA*

### ABSTRACT

Studies of advanced earth-to-orbit vehicles have shown the need for new oxygen-hydrocarbon rocket engines. These advanced engines are characterized by high chamber pressures (to 6000 psia), reusability, and high performance. The propulsion technology necessary to meet these requirements is now being defined. This paper examines thrust chamber cooling and reuse, ignition, performance, and combustion stability in the light of recent engine design studies and the results of tests with oxygen/RP-1 at pressures up to 2500 psia. It is concluded that satisfactory ignition, performance, and stability should be achievable without requiring major technology advancements. The results of engine regenerative cooling studies with oxygen, methane, propane, RP-1, and hydrogen as coolants are presented. The studies indicate that engine operation above approximately 4000 psia chamber pressure will require supplemental mass transfer cooling or the development of improved regenerative cooling techniques.

### KEYWORDS

Oxygen; hydrocarbons; high pressure; ignition; cooling; stability; soot; performance; reusability.

### INTRODUCTION

Studies of advanced launch vehicles have identified LOX/hydrocarbon (RP-1, C<sub>3</sub>H<sub>8</sub>, CH<sub>4</sub>) propellant combinations as particularly attractive by themselves or in combination with hydrogen in a dual-fuel system (Martin, 1979; Beichel, Salkeld, Shulsky, 1980). The features which make the LOX/hydrocarbon propellants attractive are their good performance combined with high bulk density (relative to hydrogen), low cost, and ease of handling. LOX/RP-1, in particular, has been used with success on a number of operational engines (F-1, Titan-1, H-1, etc.). A relatively limited amount of technology work has been performed using methane and propane as fuels.

Although there is a substantial LOX/hydrocarbon technology base as a result of this earlier work, much of it does not apply to the new engines. The new engines will differ from the existing engines in several very important areas. Foremost among these differences are operating pressure and the desire for reusability. The optimum operating pressures for various earth-to-orbit engines was recently presented

by O'Brien (1981) and is summarized in Fig. 1. This figure shows delivered performance as a function of chamber pressure for different hydrocarbon fuels for gas generator and staged-combustion engine cycles. Optimum performance with gas generator cycle engines ranges from 2500 to 4500 psia chamber pressure range whereas that of staged-combustion cycle engines is even higher. These pressure levels are many times greater than those of existing hydrocarbon-fueled engines and make engine cooling and life critical.

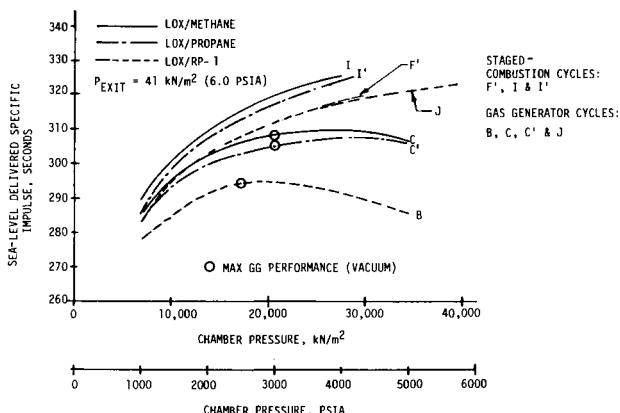


Fig. 1. Comparison of engine cycles as a function of chamber pressure.

Closely related to engine cooling and life considerations is the need for reusability. The vehicle designer desires reusability with low maintenance to keep operational costs low. Reusability with low maintenance generally translates into a more difficult design specification for the propulsion engineer.

The need to extend the LOX/hydrocarbon technology base to meet the requirements of these new engines has been recognized. The Aerojet Liquid Rocket Company (ALRC), under both NASA sponsorship and company funding, has conducted engine design studies and test firings at pressures as high as 2500 psia to begin to build this data base. This paper presents a partial status report on high-pressure hydrocarbon thrust chamber technology. Specific topics addressed are igniter design and engine ignition, combustion stability, performance, engine cooling, and reusability/life.

## DISCUSSION

### Igniters/Ignition

Smooth, reliable ignition is particularly important with oxygen/hydrocarbon propellants. Ignition delays can result in the accumulation of an explosive mixture of unburned propellant which can damage or destroy the thrust chamber when ignition does occur. This is a more severe problem with RP-1 than with  $\text{CH}_4$  or  $\text{C}_3\text{H}_8$  because of the lower volatility of RP-1.

A LOX/RP-1 spark-torch igniter was designed and evaluated firing by itself and when igniting a 2000-psia, 20,000-lbf LOX/RP-1 thrust chamber (LaBotz, Rousar, Valler, 1980). The basic design was similar to that originally developed by ALRC for oxygen-hydrogen use and now being employed on the Space Shuttle main engines. This design uses a spark to ignite an oxidizer-rich mixture which subsequently is mixed with sufficient fuel to provide a near stoichiometric torch. This torch is located

in the center of the injector and is used to provide main chamber ignition. Figure 2 shows the igniter firing through the center of the injector.

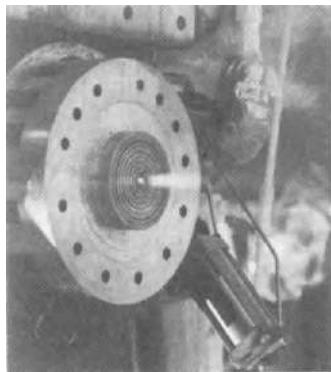


Fig. 2. LOX/RP-1 igniter firing.

The results of the igniter program were very positive, i.e., 100% reliable ignition was achieved. The igniter operated with both an RP-1 lead or a LOX lead and at spark levels down to 10 millijoules. Engine ignition in every instance was reliable and smooth, without ignition delays or chamber pressure spikes. A typical ignition transient is given in Fig. 3. It can be seen that engine ignition occurs as soon as both propellants are present. It was found that the igniter could be shut off early in the engine start transient without any adverse effects. Note that the high pressure spikes in the oxidizer circuit prior to ignition are the result of water hammer in the feed system.

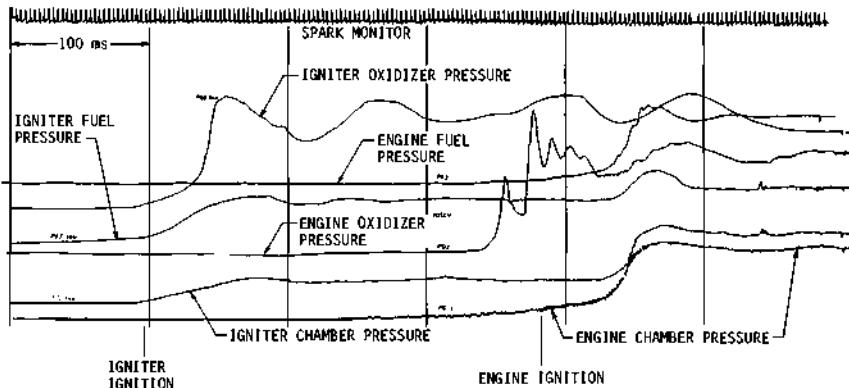


Fig. 3. Ignition transient of 2000-psia LOX/RP-1 engine.

The extension of this ignition technology to  $\text{CH}_4$  and  $\text{C}_3\text{H}_8$  and higher pressure levels should present no significant problems. The greater volatility of these fuels relative to RP-1 makes both igniter ignition and engine ignition easier to accomplish. Also, the higher engine pressures should have no effect on ignition as engine ignition was always found to occur the moment both propellants began to flow, regardless of the ultimate chamber pressure attained. Problems associated with cooling the igniter at high pressure could be avoided by shutting the igniter off early in the ignition transient, well before steady-state chamber pressure is achieved.

### High Frequency Combustion Instability

The development of the first generation of LOX/hydrocarbon engines in the 1950's and 1960's was plagued by occurrences of high-frequency combustion instability. In some cases, engine performance was deliberately downgraded to avoid instability problems with higher-performing injector patterns. The problem of combustion instability was one of the most expensive problems encountered in the development of the early RP-1-fueled engines.

It is difficult to generalize in the area of combustion stability since each injector/chamber design has its own individual characteristics. However, there are several factors which are likely to make combustion instability much less of a problem with the new engines than it was two decades ago. During development of the early engines, instability problems were addressed by modifying the injector pattern and/or baffle configuration. The time required to modify a design was measured in weeks and months so that alternate designs would be carried in parallel, resulting in greatly increased costs. The development of acoustic cavity technology in the last decade has made it possible to alter a cavity configuration in a matter of hours. This alone yields great cost savings. In addition, acoustic cavities provide tuning flexibility not achievable with baffles. The rapid solution of instability problems is also supported by analytical techniques and instrumentation that are superior to those available previously.

One other beneficial aspect of the new engines relative to combustion stability is the high pressure. High pressure leads to smaller chambers. Small chambers are generally easier to stabilize than large chambers as they have fewer acoustic modes in the frequency range which can couple with the combustion process. This permits the designer to focus his damping devices (acoustic cavities or baffles) on a decreased number of modes and bring more damping to bear on suppressing these modes.

At this time, the technology has not progressed to the point where combustion instability problems can be completely avoided during the development of the next generation of LOX/hydrocarbon engines. However, in view of the factors cited above (acoustic cavity technology, improved analytical models, and instrumentation), it is reasonable to anticipate that when instabilities are encountered, they will be resolved more readily than before.

### Combustion Efficiency

High-pressure hydrocarbon engines become attractive from a mission viewpoint only if combustion efficiencies approaching 100% are achieved. This would represent a significant improvement over the efficiencies of the RP-1-fueled engines developed during the 1950's and 1960's. (The combustion efficiencies achieved on these engines ranged from about 94 to 98%.) The problem faced at that time was that high combustion efficiency patterns were also highly unstable. As noted earlier, in several instances, performance was deliberately degraded to achieve stable combustion. With the advances in combustion stability technology, it is unlikely that combustion stability will place as severe a limit on performance as it has in the past.

The achievement of combustion efficiencies between 99 and 100% with LOX and hydrocarbon fuels has been demonstrated by the NASA Lewis Research Center (Pavli, 1979). This work was performed at 600 psia chamber pressure with both liquefied natural gas (LNG) (92% CH<sub>4</sub>) and RP-1 fuels at the 5000-lbF thrust level. Although these data were gathered at lower pressures and thrust levels than planned for the new engines, we believe they indicate there are no inherent limitations in the propellant combinations which will prevent high combustion efficiencies from being

achieved. One effect of high pressure is to increase the propellant vaporization rate. Our experience indicates that this can be a detriment to mixing and can result in reduced performance with a poorly mixing injector. With injectors designed to achieve good mixing, it appears that achieving high combustion efficiency will not present a major problem. The more significant challenge will be to maintain stable combustion during high-efficiency operation.

#### Combustion Chamber Heat Transfer

The cooling of very high-pressure engines is a formidable task. Cooling of engines for multiple reuse is even more difficult due to the need to keep wall temperatures and thermal stresses low. The design of cooled thrust chambers consists of two parts: 1) establishing the hot-gas environment in which the wall must operate; 2) designing a wall cooling system which will provide the required life in this environment.

The Hot-Gas Environment. Until recently, there were no data on the gas-side environment in high-pressure LOX/hydrocarbon engines. In the absence of data, engine studies were based on the extrapolation of turbulent flow heat transfer correlations found applicable at lower pressures. It was also hypothesized that the soot deposits generally found at lower pressures with hydrocarbon-fueled engines would also exist at higher pressures but in reduced thicknesses. These soot deposits acted as a thermal barrier which reduced the heat transfer to the wall, thereby making the soot-coated engines significantly easier to cool than if they had clean walls. Data illustrating this behavior (Pratt & Whitney Aircraft, 1965) are given in Fig. 4.

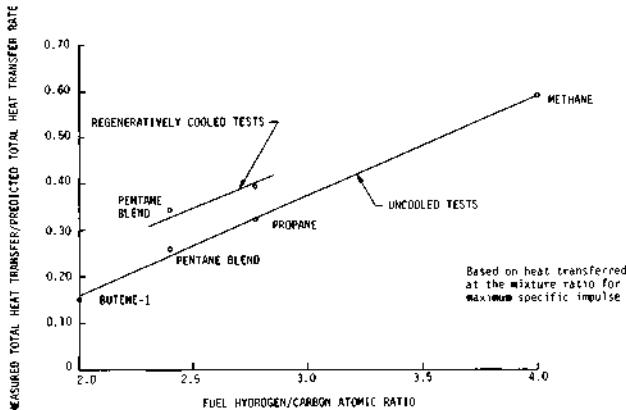


Fig. 4. Effect of fuel hydrogen-to-carbon ratio on total chamber heat transfer.

Within the past year, ALRC completed an experimental heat transfer program for the NASA in which very detailed heat transfer data were obtained with LOX/RP-1 at pressures from 1000 to 2000 psia (LaBottz, Rousar, Valler, 1980). A special water-cooled calorimeter chamber with 32 separately controlled and measured cooling loops was used in this testing. Firing durations were up to 32 seconds. This chamber is shown on the test stand in Fig. 5. The injector provided a very fuel-rich wall environment but did not contain separate film-cooling orifices.

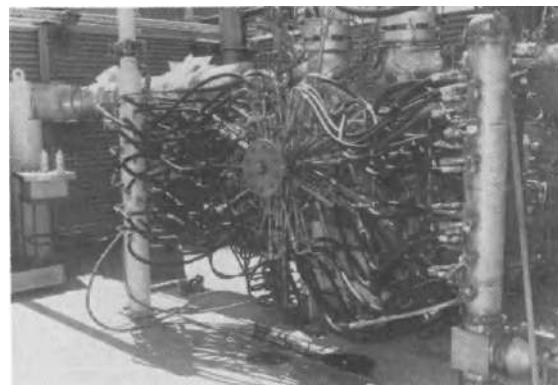


Fig. 5. Calorimeter chamber test setup.

The results of the 2000 psia tests are given in Fig. 6. This figure shows the heat flux as a function of axial position for 3 different engine mixture ratios. The data were surprising in that they gave no evidence of soot deposits forming on the chamber walls. The measured throat heat flux exceeded the predicted "clean wall" heat flux, giving no indication of the existence of a soot thermal barrier. There was also no evidence on the chamber itself of any soot deposits forming. Following firings, the chamber was blackened on the inside, but there were no soot deposits. This condition existed with two different injectors and at O/F ratios as low as 2.0. Following these tests, it was concluded that the soot deposits found on lower-pressure engines were not forming in this 2000-psia thrust chamber.

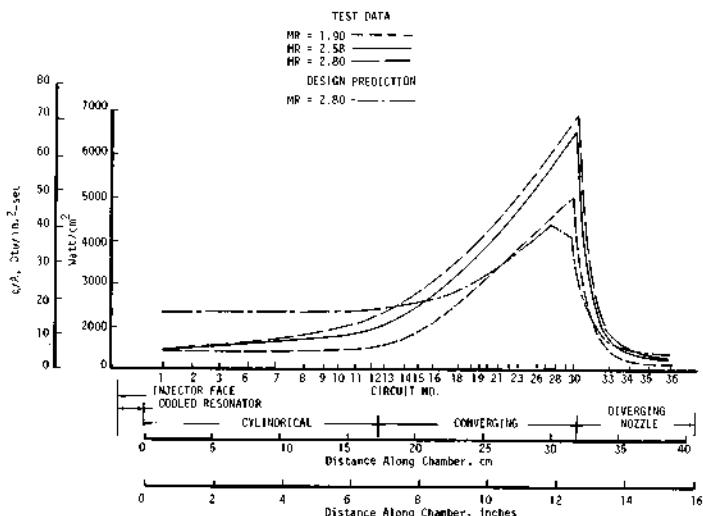


Fig. 6. Effect of engine mixture ratio on heat flux.

The lack of soot at higher pressures was also indicated by the results obtained on other programs conducted by ALRC for the NASA. In one of these programs, high-speed motion pictures were taken of the LOX/RP-1 combustion process through a quartz window in a special combustion chamber (Judd, 1980). The results were classified according to the amount of soot present in the combustion process. Typical photographs illustrating three different soot levels are given in Fig. 7. The effect of chamber pressure on soot formation is illustrated in Fig. 8 which

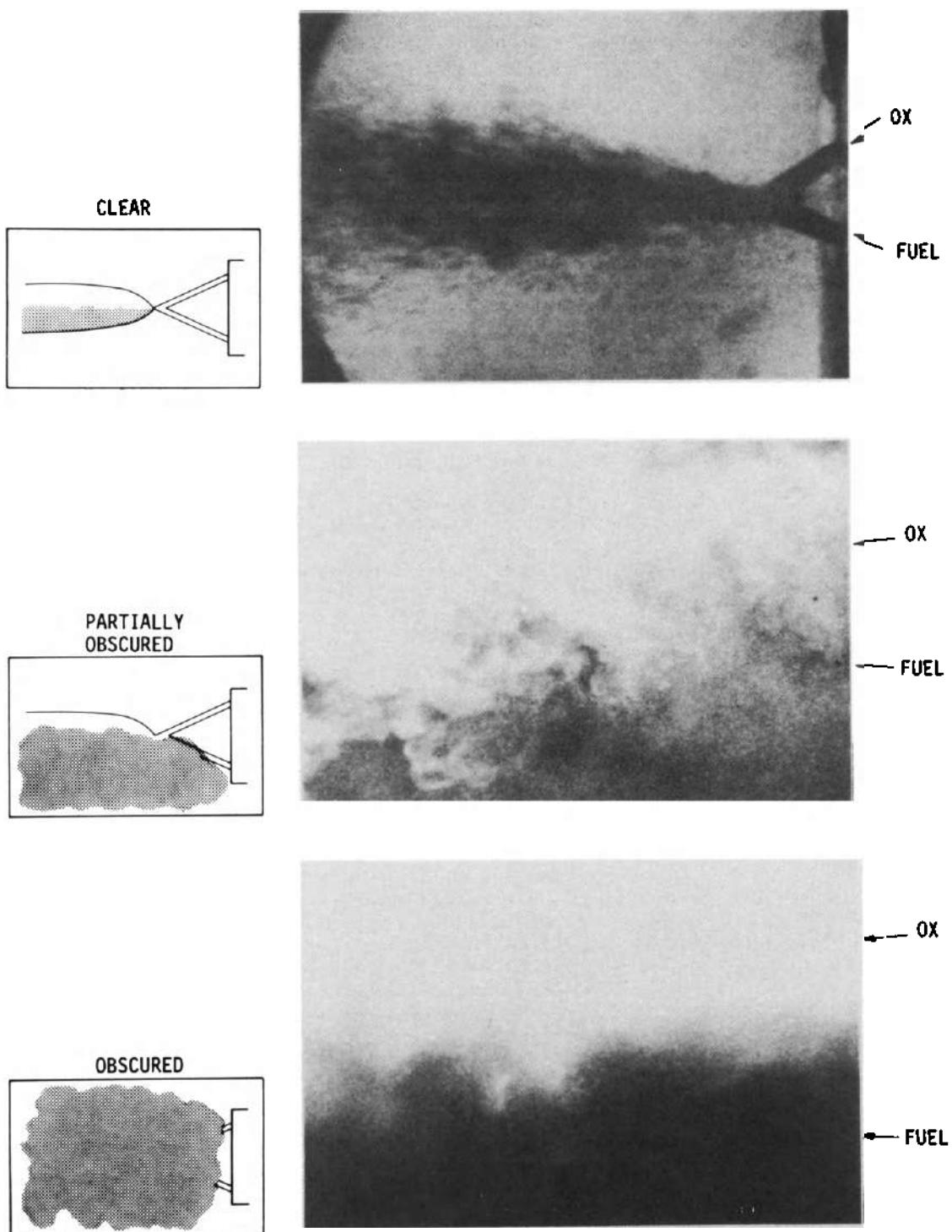


Fig. 7. Modes of carbon formation.

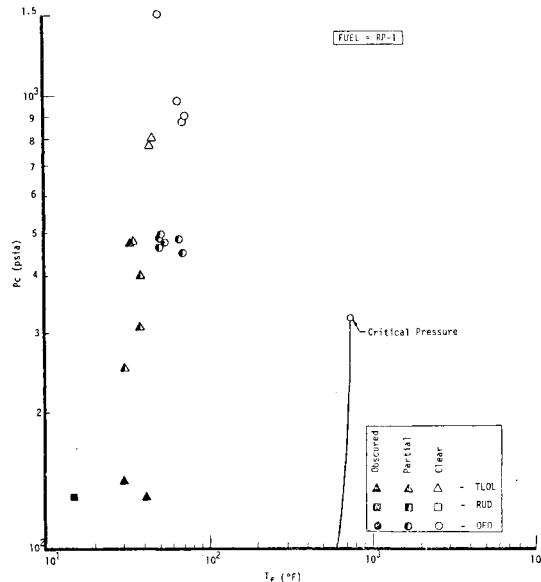


Fig. 8. RP-1 carbon formation correlation.

shows that soot decreases as pressure is increased. A similar effect was seen in recent tests conducted at ALRC at pressures up to 2500 psia with LOX/RP-1 at mixture ratios of 0.3 to 0.4. No appreciable soot deposits were encountered in any of these tests even though large quantities of soot were evident in the exhaust plume.

To date we are not aware of any heat transfer data obtained at high pressure with separately controlled and measured hydrocarbon fuel film-cooling. It is possible that large quantities of fuel film-cooling would result in soot deposits occurring on the chamber wall. However, on the basis of our results with a fuel-rich boundary zone, it appears that no appreciable soot deposits should be anticipated on the hot gas-side of the non-film-cooled high-pressure hydrocarbon-fueled engines. The thermal design of these engines must be capable of withstanding the full clean-wall heat flux without the benefit of a soot barrier.

Wall Cooling System Design. The second portion of the chamber wall cooling analysis and design addresses the ability of the coolant to remove the heat coming into the wall. This problem cannot be addressed in a simple manner due to the many parameters which must be considered. For example, the dual-fuel engines would normally operate the LOX/LH<sub>2</sub> combustor at the same time as the LOX/hydrocarbon combustor. As a result, in a dual-fuel engine, the oxygen, hydrogen, and hydrocarbon propellants would all be available as coolants to cool either part or all of the LOX/hydrocarbon thrust chamber. A similar variety of options is available in selecting the cooling technique to be employed. Although regenerative cooling is preferred due to its minimization of performance losses, it is possible, if necessary, to supplement it with transpiration and film-cooling.

There are other considerations in the design of chambers cooled with either C<sub>3</sub>H<sub>8</sub> or RP-1. Both of these fuels will form "coke" deposits on the inside of cooling passages if the surface temperature exceeds a certain value. These coke deposits degrade the cooling system and lead to accelerated chamber failure. The coking process is not well defined and there is considerable uncertainty as to the

temperature at which coking begins. RP-1 is thought to begin to coke with surface temperatures somewhere in the range of 550 to 800°F, depending on operating conditions and the grade of RP-1 being used. The coking temperature of propane has long been thought to be approximately 800°F. Although the data on decomposition rates are very limited, especially at the high coolant velocities of interest for high-pressure engines, it appears that the coolant-side wall temperatures should be limited to 700°F for commercial propane. This limit was used in the chamber cooling studies described here.

To assess the problems of cooling high-pressure LOX/hydrocarbon engines, a cooling design analysis was conducted for conventionally designed LOX/hydrocarbon engines (O'Brien, Ewen, 1981). This study assumed the use of slotted copper lines and considered RP-1, CH<sub>4</sub>, C<sub>3</sub>H<sub>8</sub>, H<sub>2</sub>, and O<sub>2</sub> as coolants. Hydrogen was included in that it will be available as a coolant on dual-fuel engines. The propellant and coolant combinations evaluated are listed in Table 1. The design analyses covered a thrust range of 200K - 2000K lbF and a chamber pressure range of 1000 - 5000 psia. The chamber geometry evaluated was conventional with a DeLaval nozzle and a 2.3:1 contraction ratio. The chamber length was established using geometric relationships which considered both thrust and pressure. The resulting lengths were typical of those encountered with operational engines. For the purposes of establishing total heat load and coolant pressure drop, the slotted chamber was considered to extend to an area ratio of 8. Although this simple chamber geometry is not necessarily typical of all the high-pressure LOX/hydrocarbon engines, it is believed that the results of cooling analysis can be considered representative.

TABLE 1. Propellant/Coolant Combination

<u>Propellants</u>	<u>Coolants</u>
LOX/RP-1	Oxygen, Hydrogen, RP-1
LOX/CH <sub>4</sub>	Hydrogen, CH <sub>4</sub>
LOX/C <sub>3</sub> H <sub>8</sub>	Hydrogen, C <sub>3</sub> H <sub>8</sub>

The chambers were designed for 100 cycles and 10 hours of operation. The cooling channels were designed for minimum pressure drop consistent with the life requirements. In light of the test results noted previously, it was assumed there would be no carbon deposition on the hot-gas side of the chamber wall with any of the propellant combinations.

#### RP-1 Cooling

The RP-1 cooling analysis was conducted using both 550°F and 800°F as the coking temperature. Cooling with the 550°F coking temperature was found to be impractical due to excessive cooling circuit pressure drop. As an example, at 1000 psia chamber pressure and 200,000 lbF thrust, a pressure drop of 1600 psia was required to meet the life requirements. Use of the higher coking temperature provided a considerably wider range of operation. The results obtained with the 800°F coking temperature are given in Fig. 9 for engines of  $0.2 \times 10^6$  lbF thrust. It is apparent from this figure that the pressure drop is not a very strong function of thrust level and that coolant pressure drop limits the chamber pressure to the 2000 to 2500 psia chamber pressure range over the entire thrust range considered. The somewhat higher pressure drop for the larger engine results from its greater length.

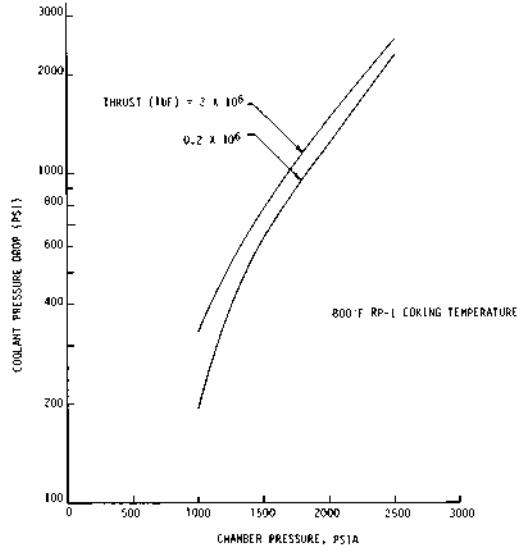


Fig. 9. RP-1-cooled engine pressure drop.

Propane Cooling

The propane cooling studies assumed the use of subcooled ( $200^{\circ}\text{R}$ ) propane as this improved both the bulk density and the cooling capability of the propane. Analyses were run at the  $0.6 \times 10^6$  lbF thrust level with both  $700^{\circ}$  and  $800^{\circ}\text{F}$  coking temperature. The results are given in Fig. 10. The upper limit on chamber pressure is about 4000 psia since the cooling circuit pressure drop becomes excessive beyond that point. The coking temperature ( $700^{\circ}\text{F}$  vs  $800^{\circ}\text{F}$ ) has only a slight impact on the results due to the use of subcooled propane. The subcooled propane provides a very large wall-to-coolant temperature difference. Changing the wall temperature by  $100^{\circ}\text{F}$  does not change this driving temperature difference significantly.

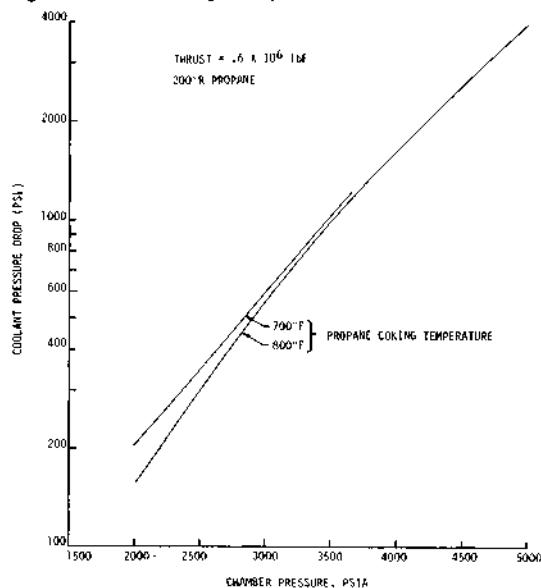


Fig. 10. Propane-cooled engine.

### Methane Cooling

Methane does not undergo thermal decomposition until approximately 1200°F. As a result, the methane cooling designs are not limited by low coking temperatures such as encountered with RP-1 and propane. However, chamber life considerations do place restrictions on the wall design. Methane cooling analyses were run for thrust levels of  $1.2 \times 10^6$ ,  $0.6 \times 10^6$  and  $2 \times 10^6$  lbF. The results of these analyses for the  $0.2 \times 10^6$  and  $2 \times 10^6$  engines are given in Fig. 11. The results show that methane, like propane, has a cooling limit of about 4000 psia beyond which pressure drop becomes excessive. This limit is nearly independent of thrust.

### Oxygen Cooling

The use of oxygen as a regenerative coolant was not given much serious consideration until recently when the heat transfer characteristics of supercritical oxygen were measured at ALRC and an oxygen-cooled chamber was tested by NASA. Historically, there has been reluctance to use oxygen as a regenerative coolant, out of concern that a coolant passage leak would result in ignition of the thrust chamber wall. This concern proved to be unfounded in the NASA testing when leaking oxygen coolant produced no more degradation than would a fuel leak. Although not conclusive, this result does indicate that oxygen cooling is considerably safer than initially believed by some.

Liquid oxygen is common to all of the advanced hydrocarbon engines. Although the cooling analysis was conducted for a LOX/RP-1 engine, the results are generally applicable to all the engines being considered. The results are presented in Fig. 12. The upper limit on LOX cooling is approximately 3500 psia based on coolant pressure drop considerations. The oxygen bulk temperature rise is relatively small due to the large oxygen flowrate.

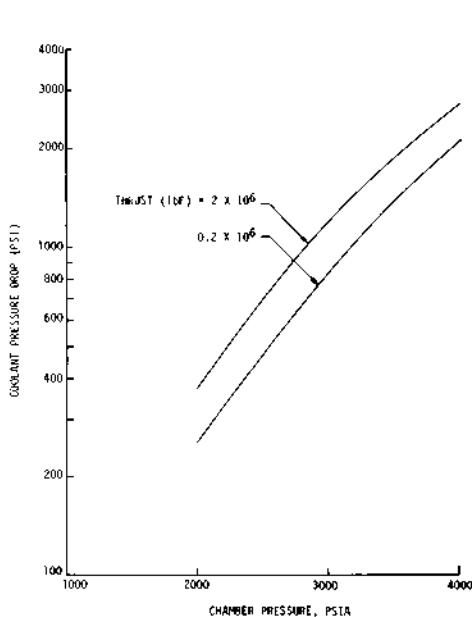


Fig. 11. Methane-cooled engine.

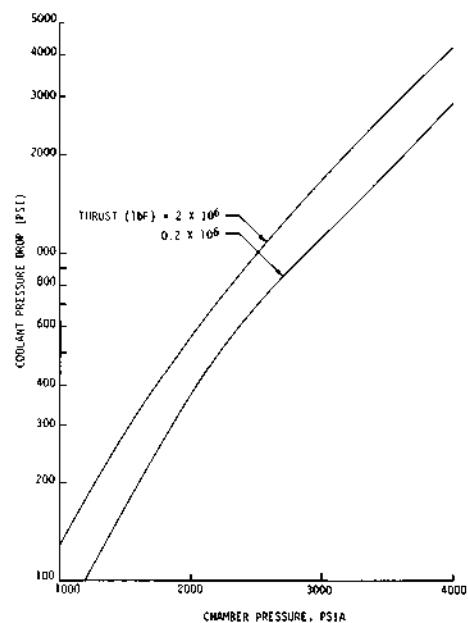


Fig. 12. LOX-cooled LOX/RP-1 engine.

Hydrogen Cooling

The use of hydrogen as a coolant for a high-pressure LOX/hydrocarbon engine was also investigated. Although no hydrogen would be available as a coolant with a simple LOX/hydrocarbon engine, it would be available on a dual-throat or dual-expander engine (Beichel, Salkeld, Shulsky, 1980). By considering hydrogen, it is possible to determine the extent to which the cooling problems of a high-pressure hydrocarbon engine may be alleviated if it is used as part of a dual-fuel engine.

The hydrogen flow required to cool the engine was minimized within the constraints of cycle life and reasonable pressure drop. The results of the hydrogen cooling analyses are given in Tables 2 and 3 and in Fig. 13. It is apparent from the pressure drop results that the use of hydrogen cooling avoids the pressure drop limitations encountered with the other coolants considered. If coolant pressure drop were the only consideration, a hydrogen-cooled chamber could operate over the entire range of thrust and pressure examined. However, there is a limit to the quantity of hydrogen available for cooling the hydrocarbon thruster. The available hydrogen flow depends on the total hydrogen consumption by the O<sub>2</sub>/H<sub>2</sub> thruster and the amount of that total flow required to cool the O<sub>2</sub>/H<sub>2</sub> thrust chamber. The hydrogen flow required to cool the hydrocarbon thruster for one dual-expander configuration is given in Fig. 14. This figure shows that, at high pressure and low thrust, up to 50% of the hydrogen would be needed to cool the RP-1 thrust chamber. This is not feasible as there would be insufficient hydrogen available for cooling the O<sub>2</sub>/H<sub>2</sub> chamber. However, at the higher thrust levels, this is no longer a problem and thus the use of hydrogen cooling becomes feasible. The range of hydrogen cooling feasibility could be extended to lower thrust levels by operating the O<sub>2</sub>/H<sub>2</sub> engine at lower mixture ratios or by increasing the thrust of the O<sub>2</sub>/H<sub>2</sub> engine relative to the O<sub>2</sub>/RP-1 engine.

TABLE 2. Hydrogen Cooling Summary

LOX/RP-1 ENGINE				
THRUST (10 <sup>6</sup> LBF)	P <sub>C</sub> (PSIA)	COOLANT FLOW (LB/SEC)	ΔP (PSI)	ΔT <sub>b</sub> (°F)
0.2	1000	7.5	53	655
	4000	15	449	445
	5000	20	799	355
0.6	1000	15	137	787
	4000	25	405	491
	5000	30	680	429
1.0	1000	20	256	946
	4000	30	429	559
	5000	35	710	498
2.0	1000	40	607	879
	4000	50	715	594
	5000	60	850	473

TABLE 3. Hydrogen Cooling Summary

LOX/CH <sub>4</sub> ENGINE				
THRUST (10 <sup>6</sup> LBF)	P <sub>C</sub> (PSIA)	COOLANT FLOW (LB/SEC)	ΔP (PSI)	ΔT <sub>b</sub> (°F)
0.2	1000	7.5	72	240
	4000	14.6	611	495
	5000	19.4	1105	395
0.6	1000	15.0	199	900
	4000	24.7	563	550
	5000	29.5	1041	484
1.0	1000	25.0	334	352
	4000	29.2	767	642
	5000	34.6	1369	568
2.0	1000	40.0	693	1004
	4000	49.6	963	684
	5000	59.4	1122	551

Cooling Summary

Based on these results, at the present time it appears that the cooling of "conventional" regeneratively cooled single-fuel hydrocarbon engines places a limit of approximately 3500 to 4000 psia on the maximum chamber pressure. To achieve this limit with an RP-1-fueled engine would require the use of oxygen as the coolant. With the propane- and methane-fueled engines, either the oxidizer or fuel could be

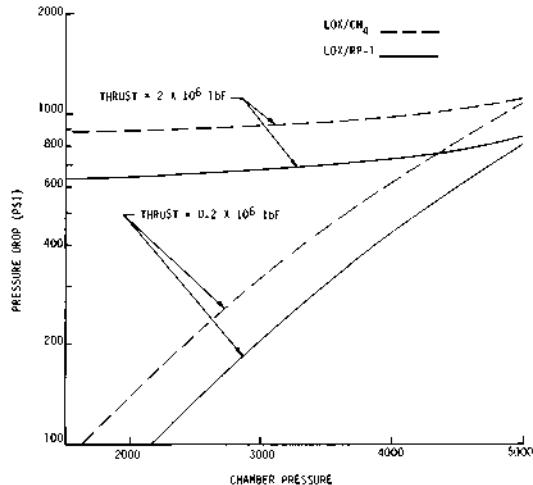


Fig. 13. Coolant pressure drop for H<sub>2</sub>-cooled LUX/RP-1 and LOX/CH<sub>4</sub> engines.

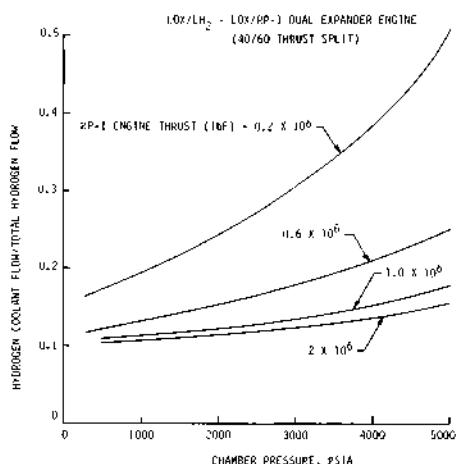


Fig. 14. Hydrogen coolant requirements H<sub>2</sub>-cooled LOX/RP-1 thrust chamber.

used as the coolant. The results of the engine performance studies of O'Brien (1981), given in Fig. 1, show these pressures to be near the peak performance levels for gas generator cycle engines. Increasing the chamber pressure of the gas generator cycle engines above 3500 to 4000 psia produces a performance decrease due to power balance considerations, independent of the cooling results presented here. Thus the gas generator cycle engines cannot be considered to be cooling-limited. This is not the case for the staged-combustion cycle engines where chamber pressures above 5000 psia are desirable from a performance viewpoint but cannot be achieved due to chamber cooling problems. At the present time, these engines must be considered to be cooling-limited unless the engine configuration makes hydrogen available as a coolant.

We do not believe this upper pressure limit of about 3500 to 4000 psia for hydrocarbon- or LOX-cooled engines to be fixed. Design approaches which reduce the heat flux into the coolant could make higher pressure levels feasible. Film cooling, barrier cooling, or transpiration cooling could be employed to extend the operating range to higher pressures, though with an attendant performance loss. It is apparent that cooling is an area in which significant technology work is required to make the higher pressure engines practical.

### Cycle Life

It is generally agreed that multiple reuse is a key element in the development of low-cost space transportation. Normally, the cycle life capability of a thrust chamber is limited by thermal stresses. With an optimized design in which the thermal stresses and coolant pressure drop have been minimized, it is possible to increase cycle life by changing the channel geometry and increasing the coolant velocity. This results in an increase in cycle life, though at the expense of increased cooling circuit pressure drop. The relationship between cycle life and cooling circuit pressure drop was investigated for the case of a LOX-cooled 600,000-lbf LOX/RP-1 engine operating at 3000 psia chamber pressure. The chamber

design employed in this cycle life study, like those used in the cooling studies, was assumed to consist of a standard slotted zirconium copper liner with an electroformed nickel closeout.

The results of this study are given in Fig. 15 as a plot of the relationship between coolant pressure drop and cycle life. This figure shows pressure drop to be almost directly proportional to cycle life over the range considered. Since each of the hydrocarbon-cooled engines considered in the preceding discussion was either pressure-drop-limited or very near to being pressure-drop-limited, it is apparent that the cycle life can be extended only at the expense of chamber pressure capability. This is undesirable since the high operating pressure contributes directly to the delivered performance of these engines. However, it should be noted that the opposite effect is also true. Lowering the life from the nominal design value of 100 cycles used in the cooling studies will produce an increase in the upper pressure limit for these engines. Thus, in vehicle design studies, it should be recognized that a cycle life/chamber pressure relationship exists which can be optimized for a particular vehicle.

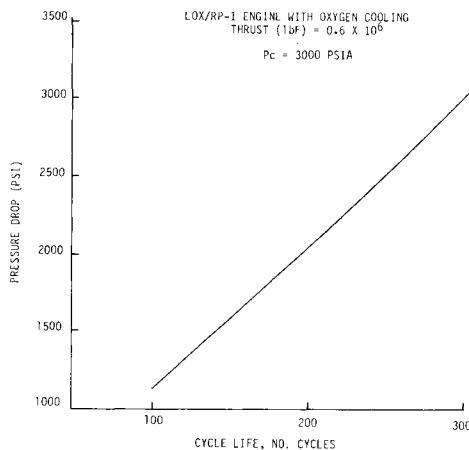


Fig. 15. Effect of cycle life on chamber pressure drop.

### CONCLUSIONS

The successful development of advanced high-pressure LOX/hydrocarbon engines will require thrust chamber assembly technology advances. Based on work accomplished to date, it appears that the required advances in the areas of ignition, combustion stability, and performance can be accomplished with a reasonable amount of effort. However, it is also apparent that the area of thrust chamber cooling will require focused attention. Data obtained during recent tests and thrust chamber cooling studies indicate that simple regenerative cooling of these engines is not feasible for chamber pressures beyond about 3500 to 4000 psia unless hydrogen is available as a coolant. Engine studies indicate these pressures would allow gas generator engines to achieve their optimum performance chamber pressure by operating at the limits of current regenerative cooling technology. Staged-combustion cycle engines, on the other hand, benefit from higher operating pressures and therefore their performance must be considered limited. Cooling aids such as thermal barriers, film or barrier cooling, and transpiration cooling systems all should be addressed in technology development programs in the next few years. Chamber cooling technology must be advanced if we are to take full advantage of the performance benefits of high-pressure hydrocarbon engines for use in future low-cost reusable space transportation systems.

## REFERENCES

- Beichel, R., R. Salkeld, and R. Shulsky (1980). New concept for far-orbit transportation. IAF Paper No. 80-F243.
- Judd, D. C. (1980). Photographic combustion characterization of LOX/hydrocarbon type propellants. Report MA-262T. Contract NAS 9-15724.
- LaBotz, R. J., D. C. Rousar, and H. W. Valler (1980). High-density fuel combustion and cooling investigation. Final Report. NASA CR 165177.
- Martin, J. A. (1979). Dual-fuel propulsion: why it works, possible engines, and results of vehicle studies. AIAA Paper No. 79-879.
- O'Brien, C. J. (1981). Advanced oxygen-hydrocarbon earth-to-orbit propulsion. 1981 JANNAF Propulsion Meeting, Chemical Propulsion Information Agency. Accepted for publication.
- O'Brien, C. J., and R. L. Ewen (1981). Advanced oxygen-hydrocarbon rocket engine study. Report 33452F.
- Pavli, A. J. (1979). Design and evaluation of high-performance rocket engine injectors for use with hydrocarbon fuels. NASA Technical Memorandum 79319.
- Pratt & Whitney Aircraft (1965). Investigation of light hydrocarbon fuels with FLOX mixtures as liquid rocket propellants. Final Report. NASA CR-54445.

## MAGNETIC ATTITUDE CONTROL OF A SPIN-STABILIZED SATELLITE

J.-L. Froeliger and J.-L. Lair

*Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, Centre  
National d'Etudes Spatiales de Toulouse, Toulouse, France*

### ABSTRACT

The Attitude Control System of the satellite ARSENE uses the action of the earth magnetic induction on onboard magnetotorquers to realise the spinning and the spin-axis orientation of the vehicle.

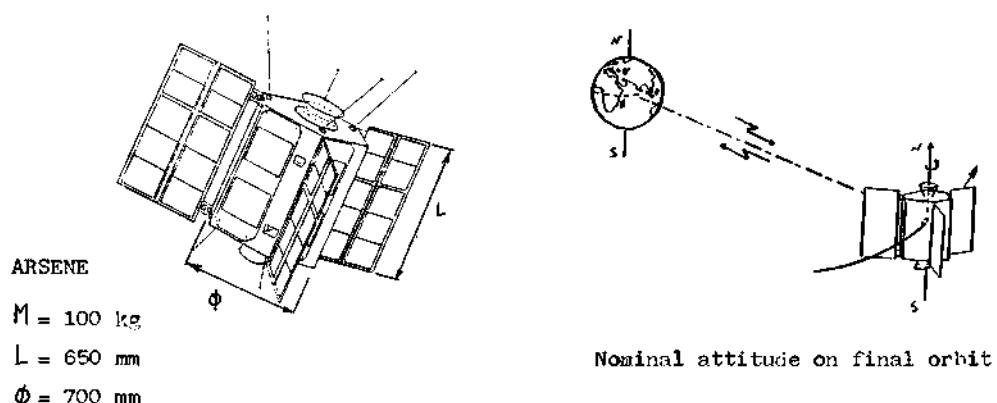
### KEYWORDS

Attitude Control System (A.C.S.) ; magnetic induction ; magnetotorquers; attitude acquisition ; attitude maintenance ; disruptive torques ; dynamic equation, spin-stabilized.

### INTRODUCTION

This paper presents the Attitude Control System (A.C.S.) of the satellite ARSENE. ARSENE is a satellite for the use of radio-amateurs intendent to be launched by the European launcher ARIANE in the year 1985. It will then be loaded in "piggy-back". Its principal task is to give the different radio-amateur stations considerable communication time. Thus, the final orbit is equatorial, elliptic, the apogee is 35600 km, the perigee 20000 km.

The satellite has reached its nominal position when its longitudinal axis is parallel to the North-South Earthaxis. The stabilization choosen is a spin stabilization that suits best the cost and simplicity requirements of the project.



The A.C.S. fonctions are first presented with their specifications. The actuation principle and the application to the different manœuvres are then developped ; finally, the effective piloting is explained. The conclusion insists on the originality of this active piloting law, that may be applied to any spinned stallite and which allows to reorient the vehicle at the same time as it changes its rotational speed.

#### ATTITUDE CONTROL SYSTEM PRESENTATION

##### Attitude Definition

The attitude defines the vehicle orientation in relation to reference axes.

The attitude of a spin-stabilized satellite is characterized by the two following variables :

- orientation of the spin-axis
- rotational speed.

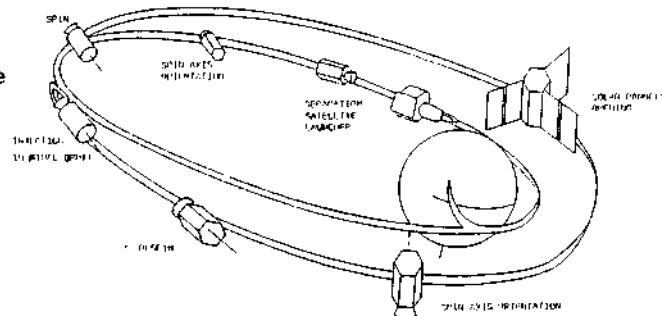
The presence on-board of a nutation damper as well as the dynamic stability of the vehicle ( $I_z/I_x > 1$  in all configurations) allows to neglect the nutation. The spin-axis and the angular momentum  $H$  have the same direction and  $H$  (in modulus and direction) defines completely the satellite attitude.

##### A.C.S. Functions

The Attitude Control System has to :

1. acquire the attitude request for the apogee thrust (on transfer orbit)
2. acquire the attitude request for the mission (on final orbit)
3. maintain the attitude request for the mission (on final orbit).

##### Operations sequence



Specifications

The different manoeuvres for the acquisition and the conservation of the attitude are realised and include the following specifications and requirements :

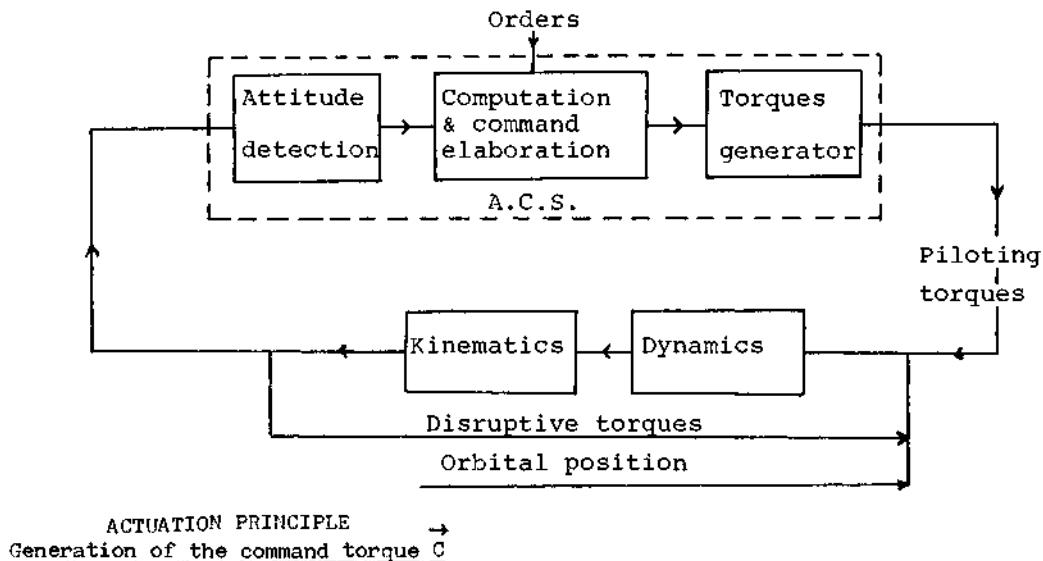
- the satellite must be operational at least a month after the injection on orbit.
- the spacecraft must maintain the spin-axis / Earth North-South axis angle in the range + 5 to - 5 degrees.
- the precision requirements for the final orbit are :
  - perigee altitude : 20000 km  $\pm$  100 km
  - apogee altitude : 35600 km  $\pm$  80 km
    - 100 km
  - inclination :  $0^\circ \pm 5^\circ$
  - weight allowed to the A.C.S. : 8 kg
  - power allowed to the A.C.S. : 5 W mean during the actuations  
1 W mean elsewhere

Functional Diagram

The Attitude Control System includes the following subassemblies :

- subassembly "attitude detection"
- subassembly "command elaboration and computation"
- subassembly "generation of command torques".

The general diagram for the A.C.S. system is represented as follows :

ACTUATION PRINCIPLE →  
Generation of the command torque  $C$ 

Cost, reliability and simplicity made us choose a magnetotorquing actuation. This system uses the action of the earth's magnetic induction on the magnetic moment created on-board by magnetotorquers.

The command torque is then given by the following equation :

$$\vec{C} = \vec{M} \wedge \vec{B}$$

Piloting law

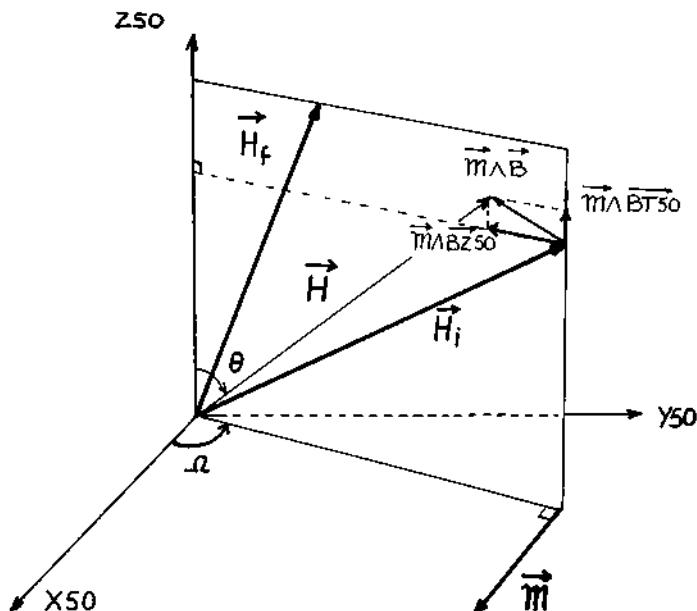
The command torque reacts upon the vehicle attitude as expressed by the dynamic equation, written in the  $\mathcal{Y}_{50}$  inertial frame :

$$\left( \frac{dH}{dt} \right)_{\mathcal{Y}_{50}} = \vec{C} + \vec{C}_p \quad (1)$$

The disruptive torques are taken into account by a recurrent law (see actuation synopsis) and are therefore neglected during the acquisition manoeuvres. Then, the equation (1) becomes :

$$\left( \frac{dH}{dt} \right)_{\mathcal{Y}_{50}} = \vec{C} \quad (2)$$

The piloting modifies the satellite's attitude, by way of the spin axis orientation and/or the rotational speed, to realise the different functions 1., 2. & 3. of the A.C.S. The piloting must realise the transition from the given initial attitude  $H_i$  to the required final attitude  $H_f$ . The command torque must belong to the plane ( $H_i, H_f$ ), so that the magnetic torque created on-board will be normal to the plane ( $H_i, H_f$ ). As the disruptive torques are neglected, the command torque is always in the plane ( $H_i, H_f$ ) and this plane is stationary. (We limit the study to the cases where  $Z_{50}$  lies in the plane ( $H_i, H_f$ )).

Actuation Principle

The knowledge of the earth's magnetic induction allows us to select preferential directions for  $C$  in that plane, by stopping or not the actuation or by reversing the direction of the magnetic moment.

We may have to stop the actuation during magnetic storms for example, when BZ50 reverse itself.

The earth's magnetic induction is modeled by a dipole placed at the earth's mass center and inclined at  $11^{\circ}4$  with respect to the geographical North-South axis. This permits us to simulate numerically the trajectories followed by the angular momentum of the vehicle.

#### Piloting Equations

The dynamics equation is written in the  $\vec{Y}_{50}$  inertial frame (see figure : actuation principle ).

$$\vec{H} = \begin{bmatrix} \sin \theta & \cos \Omega \\ \sin \theta & \sin \Omega \\ \cos \theta & 0 \end{bmatrix} \cdot I \omega$$

$$\vec{B}_{50} = \begin{bmatrix} BX_{50} \\ BY_{50} \\ BZ_{50} \end{bmatrix}$$

$$\vec{M} = M \begin{bmatrix} \sin \Omega \\ -\cos \Omega \\ 0 \end{bmatrix}$$

Then the equation (2) becomes :

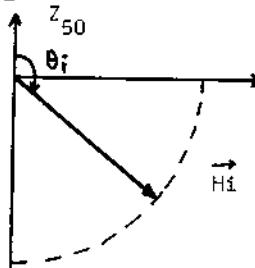
$$(Σ) \quad \left\{ \begin{array}{l} \frac{d}{dt} (\omega \sin \theta) = - \frac{M}{I} BZ_{50} \quad (3) \\ \frac{d}{dt} (\omega \cos \theta) = + \frac{M}{I} (\sin \Omega BY_{50} + \cos \Omega BX_{50}) \quad (4) \end{array} \right.$$

#### APPLICATION TO THE OPERATIONS 1. 2. and 3.

##### Attitude Acquisition for the Apogee Thrust

A spin-axis orientation and a spinning of the satellite is needed to give the vehicle the required direction and adequate gyroscopic rigidity for the apogee thrust.

Injection conditions of the ARIANE launcher. ARSENE is loaded "piggy back" and as a result of time and propulsion restrictions, the launcher will probably not be able to deliver the satellite in the optimal attitude for the subsequent manoeuvres. A rotational speed of 10 RPM is however required to keep the on-board sensors in a good working order. Otherwise, the angular momentum at the injection ( $\vec{H}_i$ ) is supposed to be in the vectorial plane ( $Z_{50}, \vec{H}_i$ ) and all the directions for the  $\vec{H}_i$  vector in the range 90 to 180 degrees will be considered.

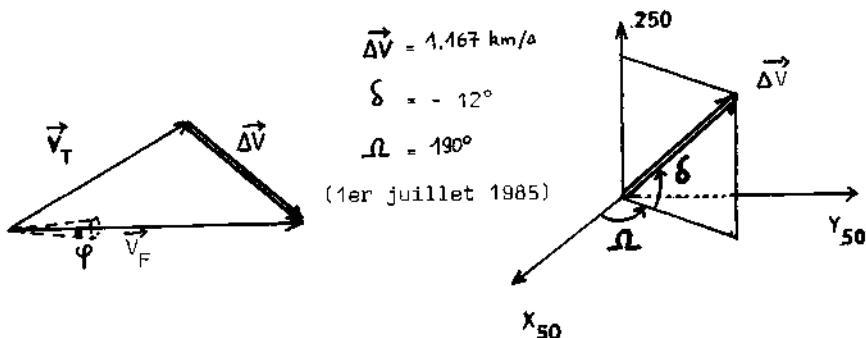


Angular momentum at the injection

Determination of the attitude required for the apogee thrust. ARSENE has to take into account the orbital requirements of the main satellite. The most probable case of a geostationary satellite injected by ARIANE on a transfer orbit with a perigee and an apogee altitude of respectively 200 km and 35786 km and with 8 degrees of

inclination is assumed. This is also the transfer orbit for ARSENE. To reach the final orbit (20000/35600/0°), the modulus and the direction of the speed vector  $\Delta V$  to produce by the apogee thrust, is then determined ; the  $\Delta V$  and  $H_i$  vectors have the same direction.

The launching date does not matter because the injection point (10° W) and the magnetic dipole (69° W) have always the same relative positions.



The imperfections of the apogee thruster - defective alignment of the thrust-axis for example - induce an error in the nominal direction of  $\Delta V$  which leads to a damage of the satellite's final orbit (figure A.). This effect can be shortened by giving the vehicle a high rotational speed : the gyroscopic rigidity must be so that the final orbit squares with the specifications (figure B.).

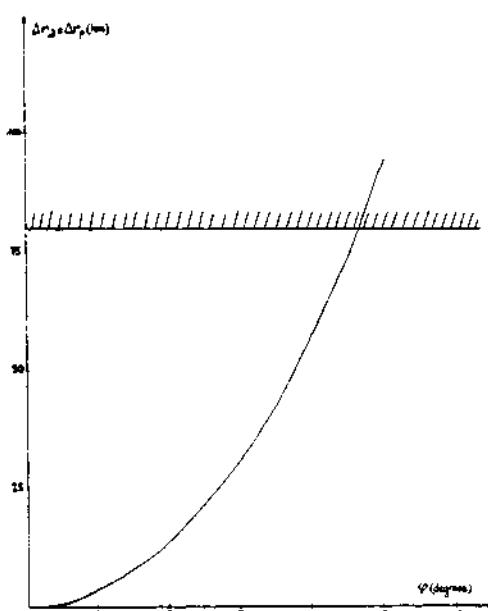


Figure A : Perigee and apogee attitude errors VS wrong orientation of  $\vec{V}_F$ .

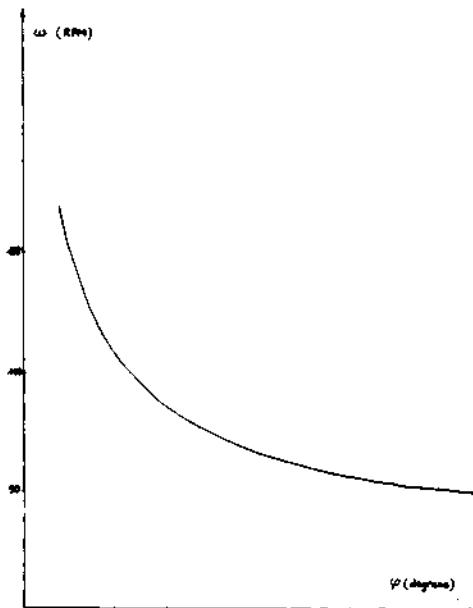


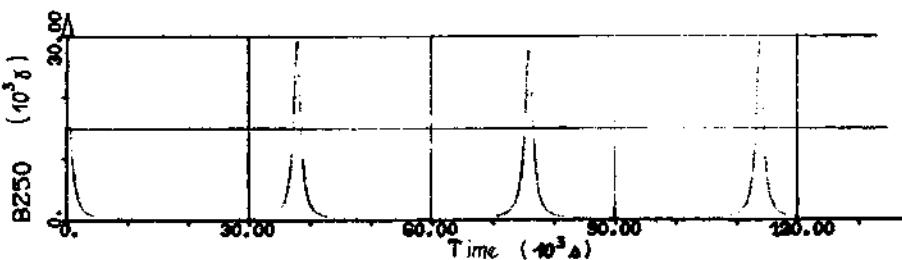
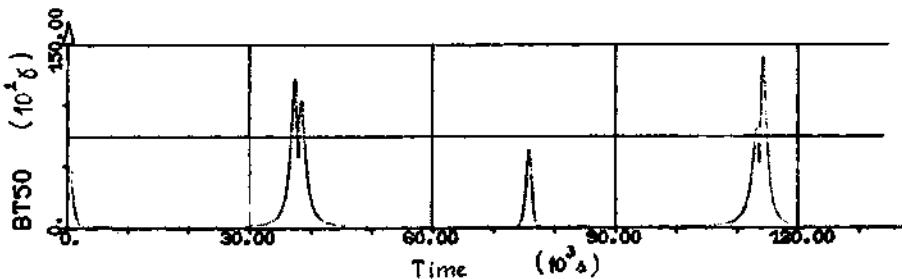
Figure B : Defective alignment of final speed vector VS rotational speed.

This study shows that the minimum spin-rate is about 60 RPM. In so far as possible, (in order to reduce the costs of the project), the TOULOUSE's station only is being used for the remote control and telemetric operations.

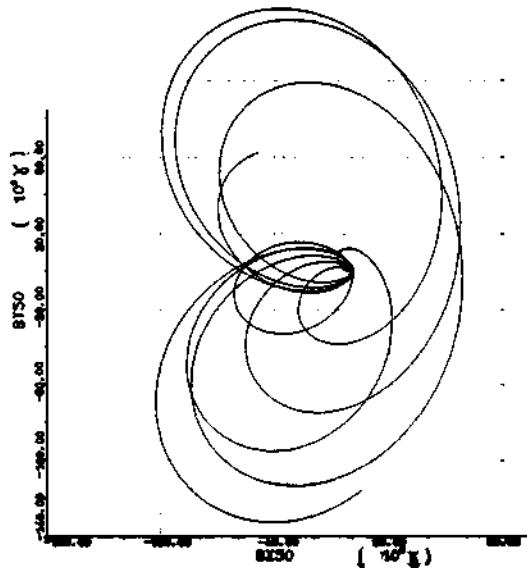
#### Magnetic Induction Characteristics on Transfer Orbit

The evolution of the components of the earth's magnetic field in the J50 inertial frame is given by the following curves :

$$\overrightarrow{BT50} = \overrightarrow{BX50} + \overrightarrow{BY50}; BT50 = \|\overrightarrow{BT50}\|$$



Magnetic field strength as a function of time

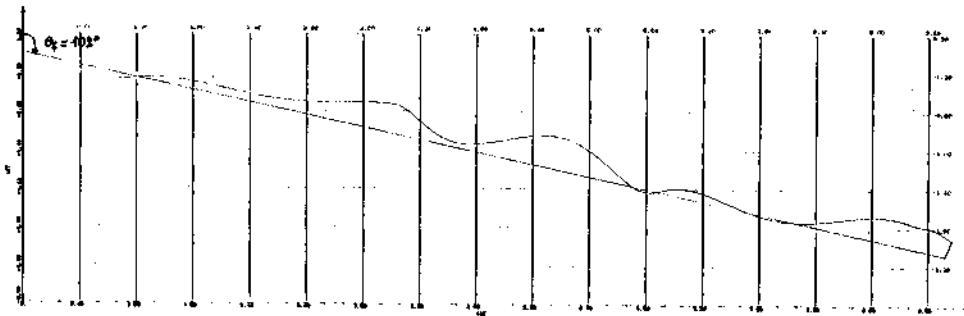


$BY50$  as a function of  $BX50$

The magnetic induction modulus is stronger near the perigee than elsewhere ; BZ50 is always positive and greater than BT50.

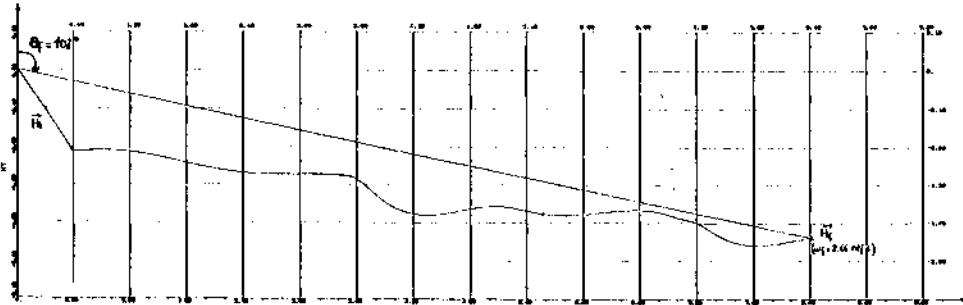
Results.  $\vec{M}$  is normal to the vectorial plane ( $Z50$ ,  $H_1$ ) ; the direction of the  $\vec{M}$  vector, and the actuation periods are chosen so that the angular momentum moves from  $H_1$  to  $H_2$ . The magnetic moment strength on-board is  $250 \text{ Atm}^2$  and the roll moment of inertia is  $6.3 \text{ kg.m}^2$ . The first apogee seen from TOULOUSE is the second apogee of the satellite so that the first localization of the satellite and the manoeuvre starting orders are realised at that time. The actuation strategy depends on the satellite's attitude at injection. Three cases can be distinguished :

1 -  $90^\circ < \theta_i < 128^\circ$ . The most simple one. The actuation begins at  $T_0 = 15 \text{ h } 20 \text{ mm}$  (2nd apogee) and must be stopped between  $T_1 = 41 \text{ h}$  and  $T_2 = 42 \text{ h } 20 \text{ mm}$  according to the required rotational speed. An example is shown on the following figure with  $\theta_i = 105^\circ$ .



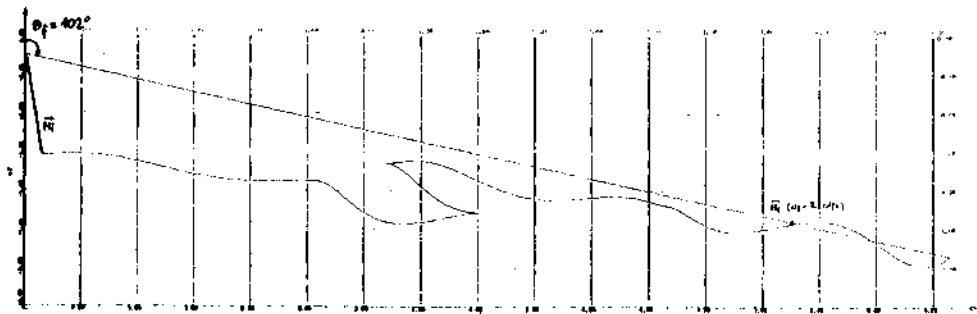
Evolution of the angular momentum -  $\theta_i = 105^\circ$

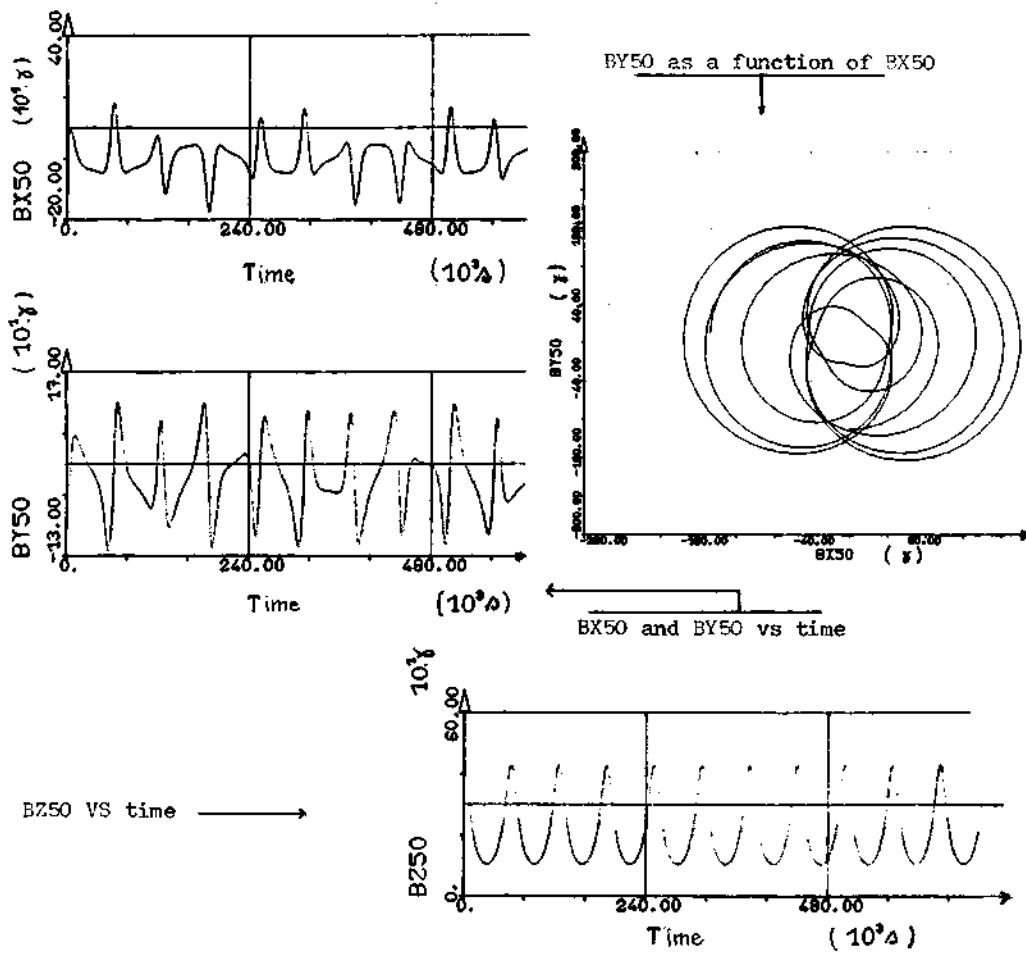
2 -  $128^\circ < \theta_i < 160^\circ$ . The actuation starts at  $T_0$  but must be stopped between  $T_3 = 31 \text{ h } 43 \text{ mm}$  and  $T_4 = 42 \text{ h } 05 \text{ mm}$  to begin again, finally being switched off between  $T_5 = 42 \text{ h } 46 \text{ mm}$  and  $T_6 = 52 \text{ h } 20 \text{ mm}$ . The case of  $\theta_i = 145^\circ$  is presented on the following figure.



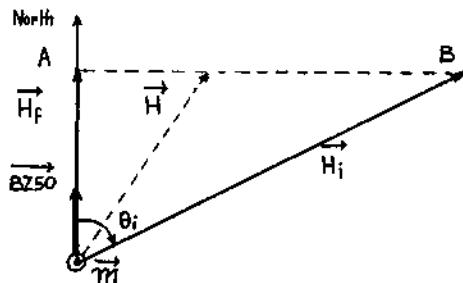
Evolution of the angular momentum -  $\theta_i = 145^\circ$

3 -  $160^\circ < \theta_i < 180^\circ$ . The magnetic moment direction must be reversed between  $T_7 = 31 \text{ h } 40 \text{ mm}$  and  $T_8 = 34 \text{ h } 07 \text{ mm}$  and stopped between  $T_9 = 42 \text{ h } 42 \text{ mm}$  and  $T_{10} = 52 \text{ h } 40 \text{ mm}$ . The example of  $\theta_i = 170^\circ$  is shown.



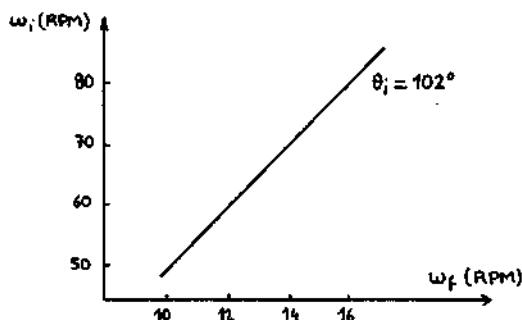


In a first analysis, the components BX50 and BY50 are neglected. The trajectory of the angular momentum  $\mathbf{H}$  is then rectilinear as shown in the figure below.



The final rotational speed is in that case related to the initial one by relation(4) :

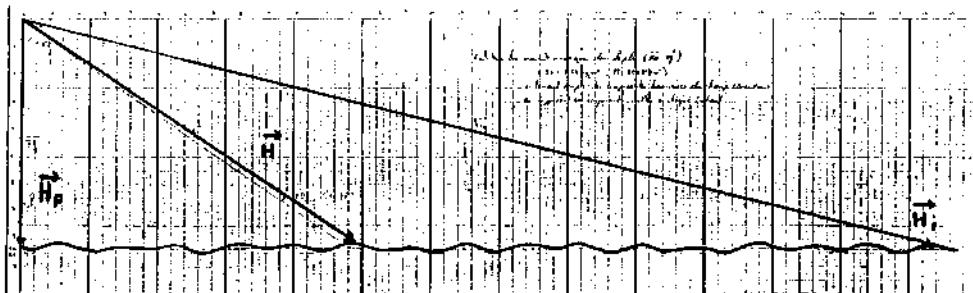
$$(\omega \cos \theta = \text{constant}) \implies (\omega_i \cos \theta_i = \omega_f)$$



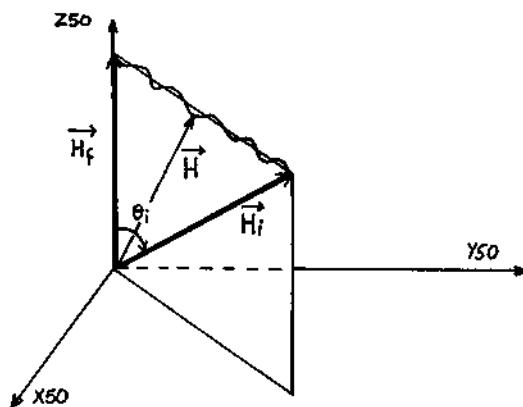
$\omega_i$  as a function of  $\omega_f$

If  $\theta_i = 102^\circ$  (request for the apogee thrust), then  $\omega_f = 0.2\omega_i$ . For example,  $\omega_i = 80$  RPM gives  $\omega_f = 16$  RPM.

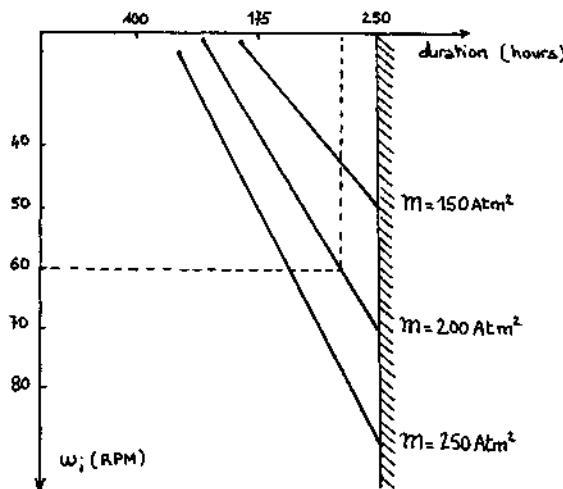
In a second analysis, the components BX50 and BY 50 are taken into account. The direct implication is to introduce an oscillation round the straight line AB. The figure below was drawn after numerical resolution of ( $\sum$ ).



And then the same figure in Space :



The time of operation is also drawn vs the rotational speed for several values of the magnetic moment  $M$ . The upper limit chosen for the duration of the operation 2. is 250 hours.



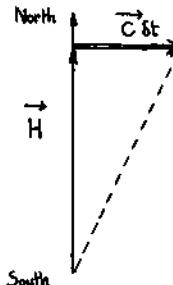
The manoeuvre is then completely realized as follows :

$$\xrightarrow{\quad \left[ \begin{array}{l} \omega_i = 80 \text{ RPM} \\ \theta_i = 102^\circ \end{array} \right] \quad} \xrightarrow{\quad \begin{array}{c} M = 250 \text{ Atm}^2 \\ T = 250 \text{ hours} \end{array} \quad} \xrightarrow{\quad \left[ \begin{array}{l} \omega_f = 16 \text{ RPM} \\ \theta_f = 180^\circ \end{array} \right] \quad}$$

#### Maintenance of the Nominal Attitude

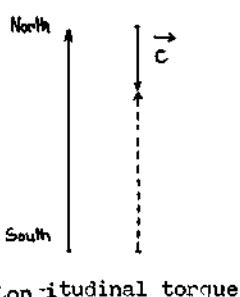
As mentioned previously, the spin-axis is to be kept in a cone of a N-S axis. If no disruptive torque is reacting upon the satellite, attitude motion is described by the Euler-Poincaré laws and, as the nutation is neglected, the spin axis is inertial. In fact, disruptive torques do act on the vehicle on the final orbit. They can be classed in two categories :

- a) The transversal torques, normal to the spin-axis, which effect the motion by letting the spin-axis drift ; they are essentially the torque due to solar radiation pressure and the torque due to gravity gradient.

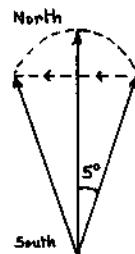


The transversal torque

- b) The longitudinal torques, parallel to the spin-axis, the effect of which is to slow down the rotational speed. They are essentially the torque due to eddycurrents and the torque due to magnetic hysteresis.



Longitudinal torque



Correction strategy

A first estimation leads to neglect the different disruptive torques in relation with the torque due to solar radiation pressure. This torque can be simply modeled by the vector :

$$\vec{C} = C e^{j\Omega_a t}$$

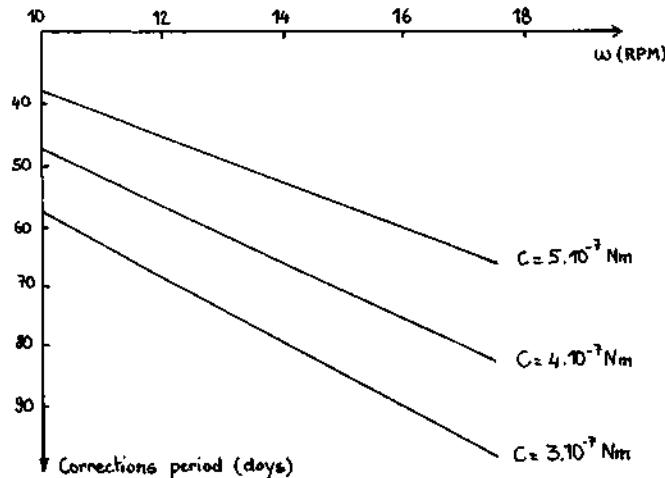
where  $C$  = modulus of the torque;

$\Omega_a$  = apparent angular velocity of the sun around the earth.

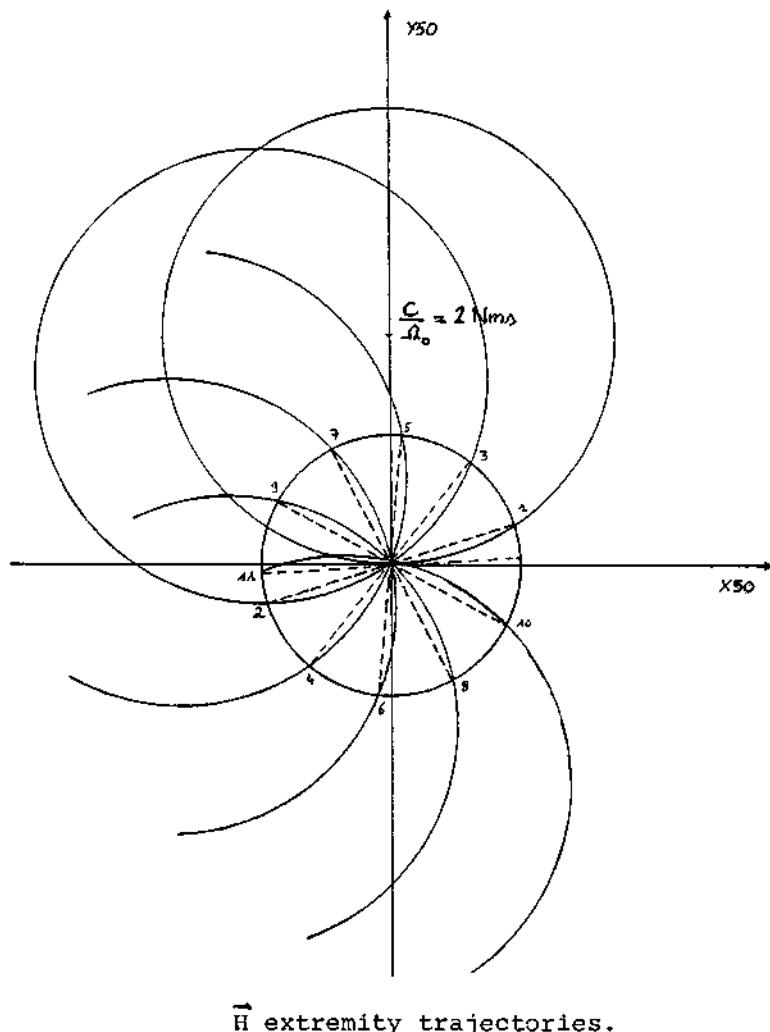
The correction strategy consists in driving the extremity of the disoriented spin-axis to a symmetrical position in relation to the N.S. axis (see figure correction strategy).

The dynamic equation  $(\vec{dH}/dt) = \vec{C}$  is integrated to give the trajectory of the angular momentum :  $H = C_j / \Omega_a (1 - e^{j\Omega_a t})$ . Therefore, on a plane parallel to the equatorial one, the extremity of the spin-axis moves on a circle including the origin (that is the point trace of the N.S. axis.). The complete Figure of the attitude maintenance on final orbit is then drawn on the next page.

The period between two corrections is also given in relation to the rotational speed for several values of the solar pressure torque. Finally, the corrections will take place every two months and need approximatively 18 hours as shown on the following figure.



Correction duration VS Spin rate.



#### EFFECTIVE PILOTING

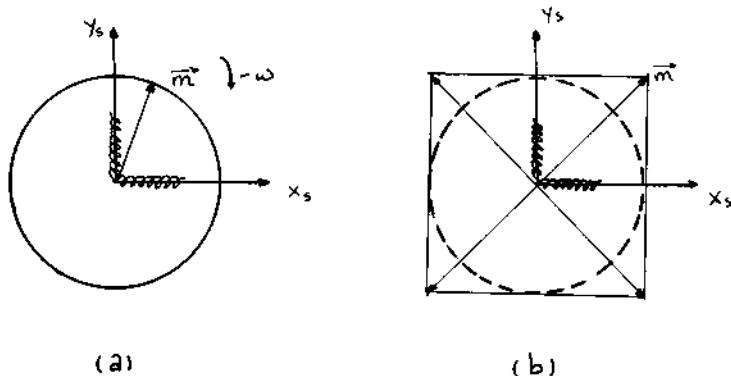
##### Generation of the magnetic current

If the transversal disruptive torques as well as the systematic errors are neglected, the actuation law consists then in generating on-board a magnetic moment with the two following characteristics :

1.  $\vec{M}$  is inertial, i.e. has a fixed direction in the inertial frame.
2.  $M$  is always normal to the plane ( $H_i$ ,  $H_f$ ).

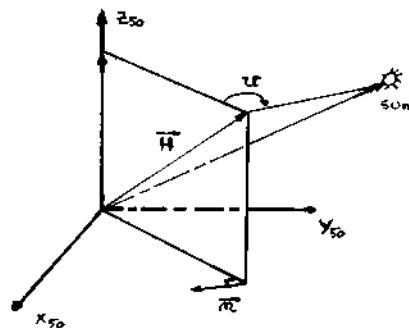
To do that, two magnetotorquers are placed normal to each other in the equatorial plane of the satellite.

The characteristic 1. is established if a sinusoidal magnetic moment with a period equal to the spin period is generated in each torquer ; the two sinusoidal moments are in quadrature. The resultant magnetic moment is then contrarotating in the satellite frame. (figure : resulting magnetic moment (a).).



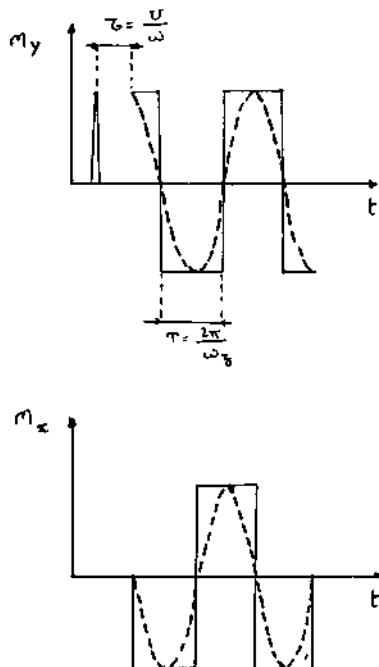
Resulting magnetic moment

The characteristic 2. is established if one of the sinusoidal curve is correctly phase displaced in relation to a sun synchronisation for example. The phaseshift depends then on the dihedral angle  $u$  ( $u = (H_i, H_f), (H, \text{sun})$ ) and the rotational speed.



Dihedral angle  $u$

In fact, as the magnetotorquers used have ferromagnetic cores, the generation of sinusoidal moments in each magnetotorquer suppose an intensity law relatively complex. So the actuation law is a "bang/bang" one. The resulting magnetic moment moves then on the four corners of a square. (figure : resulting magnetic moment (b).).



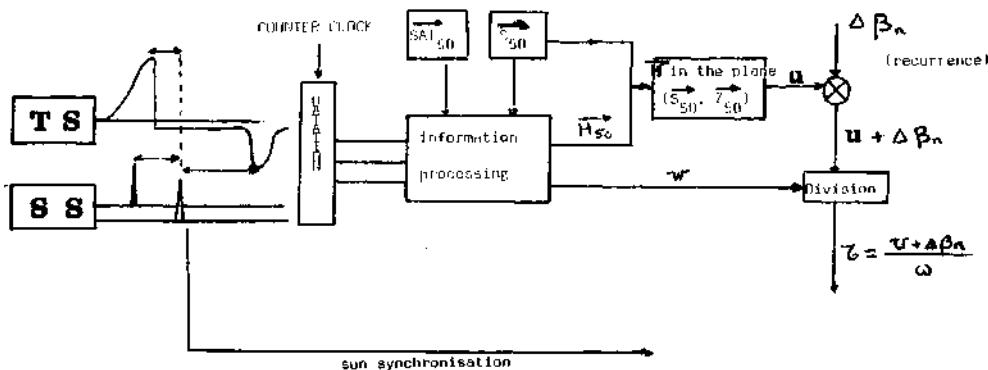
#### Magnetic moments

##### Actuation : Synopsis

To generate the characterised magnetic moment, the Attitude Control System disposes of the following subassemblies :

- subassembly "Attitude detection"
- subassembly "compute and command elaboration"
- subassembly "torque generator".

The synopsis is then given on the next page.



#### Actuation synopsis

The informations delivered by the sensors are first dated and then processed on board by a microprocessor, in relation to commands, remote controled by earth stations. After resolution, the attitude of the vehicle is totally known through its angular momentum  $H$ . The generated periodic curves are then defined by the two variables : - period = spin period  
- phase shift =  $U/\omega$

A recurrent law ( $\Delta\beta_n$ ) takes into account the transversal disruptive torques as well as the systematic errors, the effect of which is to create a drift in the plane ( $H_i$ ,  $H_f$ ). This law gives the error between the computed value for  $H$  and the measured value. The consequence is a phase shift of the magnetic moment  $H$  in relation to the normal to the plane ( $H_i$ ,  $H_f$ ). This phase shift is calculated so that the drifting plane can be driven back to its nominal direction.

#### Subassembly Attitude Detection

It is composed of two crossing IR sensors and a solar sensor . The informations of the terrestrial sensors give the depointing of the spin-axis. The signals of the solar sensor give the rotational speed and the sun colatitude angle in the satellite frame. These signals are sufficient to define completely the satellite attitude.

#### Subassembly Magnetotorquers

Two different designs for magnetotorquers may be envisaged : Air Cored Coil (A.C.C.) and Magnetic Cored Coil (M.C.C.). The magnetic moment to produce is about 250 Atm<sup>2</sup>; A.C.C. implementation should be unrealizable because of weight and size. The M.C.C. of a cylindrical form shall be selected.

The optimisation consists in selecting the most efficient materials, the sizing and the working point of the core which minimizes the weight and the power dissipated. The whole power dissipated is :

$$P_T = P_J + P_H + P_F$$

$P_J$  = power dissipated in the wire

$P_H$  = power dissipated by magnetic hysteresis

$P_F$  = power dissipated by eddy-currents.

The following formulae gives the mass of the magnetotorquer vs magnetic moment, power dissipated in the wire and characteristics of materials :

$$M = M \left( \frac{4\pi\rho_1}{P_J} \cdot \frac{\pi L_0}{\rho_0} \cdot \left( \frac{1+k_1}{k_2} \right)^2 \cdot \frac{B}{N_h^2} + \frac{\rho_0 \mu_0}{B} \right)$$

$\rho_1, \rho_0$  volumic mass and resistivity of the wire

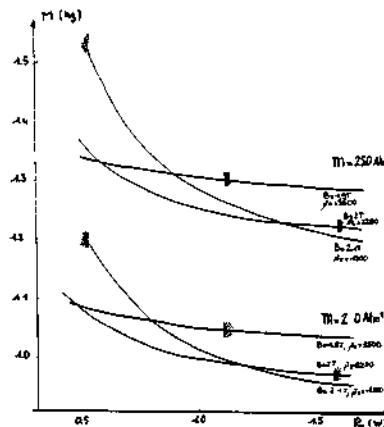
$L_0$  core length

$\rho_0$  volumic mass of the core

$k_1, k_2$  empirical coefficients of Nagooka.

Therefore, aluminium is used for wiring and magnetic materials with relative permeability and high induction saturation values for the core are selected ; this type of materials is mostly ferro-cobalt alloy AFK 502.

Then, the following curves give the magnetotorquer's mass vs dissipated power and magnetic moment for different working points.



#### Balance Sheet

the microprocessing system is not yet completely defined, but its weight should not exceed two kilos, the mass balance sheet will be :

sensors	:	1.5 kg
microprocessing system	:	< 2. kg
magnetotorquers	:	3.2 kg
nutation damper	:	0.5 kg
		<u>&lt; 7.2 Kg</u>

and the mass specification will be respected.

The power dissipated by the sensors and the magnetotorquers is 3,8 W.

#### CONCLUSION

The attitude Control System using magnetotorquers as defined in this study realizes the different functions :

1. orientation and spin of the satellite for the acquisition of attitude for the apogee thrust : the strategy depends on the initial attitude
2. orientation and reduction of the rotationnal speed for the acquisition of the nominal attitude requiring 250 hours of torque actuation
3. maintenance of the nominal attitude : corrections every two months requiring 18 hours of torque actuation.

These functions are fulfilled with respect to the specifications, that is :

- attitude and orbit requirements
- time duration for the operations
- weight and power allowed for the equipments.

This solution, using the action of the earth magnetic field on magnetotorquers achieves the orientation of the spin-axis as well as the spinning (or reduction of speed) of the satellite.

The originality of this active piloting lies in the fact that the two manoeuvres (orientation of the spin-axis and speed modification) are realized simultaneously and that most of the operations are fulfilled on a high orbit (20000/35600). There is no use of a magnetometer to measure the terrestrial magnetic induction ; the magnetotorquer's fiability and lifetime are high in comparison with other stabilisation equipments and, what is quite unusual, the same equipments will be used on the transfer orbit as well as on the final orbit.

#### NOTATIONS

- $\gamma_{50}$  inertial frame
- $B$  earth's magnetic induction
- $M$  magnetic moment
- $C$  command torque
- $H$  angular momentum
- $H_i$  initial attitude
- $H_f$  final attitude
- $C_p$  disruptive torque
- $I$  roll moment of inertia
- $\omega$  rotationnal speed of the satellite
- $u$  diecteal angle ( $\vec{H}_i$ ,  $\vec{H}_f$ ), ( $\vec{H}$ , sun)
- $\tau$  phase shift

## ACKNOWLEDGEMENT

The authors would like to express their gratitude to the Centre National d'Etudes Spatiales de Toulouse where this three-month Study has been done. Special recognition and thanks are due to the Stabilization Department and particulary to Paul DUCHON for his advice.

## REFERENCES

- AUBERTIN, J.M., C.LAZARO (1980). Analyse de mission du Satellite ARSENE. CNES Editions, TOULOUSE.
- DUCHON, P. (1980). Commande d'attitude et d'orbite. Cours de technologie Spatiale. CNES Editions, TOULOUSE.
- FEYNMAN, LEIGHTON and SANDS. Mainly electromagnetism and matter. The Feynman lectures on physics.
- FROELIGER, J.L. , J.L. LAIR (1981). Système de Stabilisation du satellite ARSENE. CNES Editions, TOULOUSE.