

SPACE AND ENERGY

PROCEEDINGS OF THE XXVIth INTERNATIONAL ASTRONAUTICAL CONGRESS
LISBON, 21-27 SEPTEMBER 1975

Edited by

L. G. NAPOLITANO

Istituto di Aerodinamica, Università di Napoli, Naples



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 of Canada Ltd., 75 The East Mall, Toronto, Ontario, Canada
AUSTRALIA	Pergamon Press (Aust.) Pty. Ltd., 19a Boundary Street, Rushcutters Bay, N.S.W. 2011, Australia
FRANCE	Pergamon Press S.A.R.L., 24 rue des Ecoles, 75240 Paris, Cedex 05, France
WEST GERMANY	Pergamon Press GmbH, 6242 Kronberg-Taunus, Pferdstrasse 1, Frankfurt-am-Main, West Germany

Copyright © 1977 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 1977

Library of Congress Catalog Card No. 76-46135

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

Printed in Great Britain by A. Wheaton & Co., Exeter

ISBN 0 08 021053 8

PREFACE

This volume continues the series "Astronautical Research" which began in 1970. It contains a number of survey papers, invited lectures and papers of general interest presented at the 26th Congress of the International Astronautical Federation. Together with a special issue of the Journal *Acta Astronautica* it constitutes the Proceedings of the Congress itself. The two companion publications complement and integrate each other.

The annual Congresses of the Federation are structured so as to develop a central theme and at the same time provide an international forum for the presentation of original ideas and results in different fields involved in the space sciences and technologies. The Congresses also serve to monitor the developments and trends of interest and research evolving within the space community which the Federation considers as made up of both the developers and the users of space science and technology.

The editing of the present volume aimed at properly reflecting the structure of the Congress and appropriately witnessing the essential trends that emerged from it. The volume is thus made up of four parts.

The first part deals with the central theme of the 1975 Congress: "Space and Energy". It begins with the sixth I.A.F. Invited Lecture given by Prof. A. Jaumotte, the title of which aptly summarizes the thematic aims and scopes of the Congress: "The Earth, an open system. How to utilize Solar Energy".

The second part, devoted to the "Engineering and Life Sciences", contains papers giving surveys and/or assessments of a number of basic topics in Astrodynamics and Bioastronautics.

The third part is entitled "Space Systems and Space Technology" and contains overall views on propulsion system concepts for a single stage Shuttle, on the reliability of outer planet spacecraft, and on specific systems such as an experimental hybrid rocket, the Large Space Telescope and the 1980 Pioneer Jupiter Orbiter.

The last part is devoted to the application satellites, with particular reference to Communication and Earth Resource Satellites. A first paper surveys progress and trends in communication satellites and a companion paper reviews current and future technology of Communication Satellites. The overall picture is completed by a group of papers dealing with National and International Space Communication Projects, namely: the Italian Sirio Program, the Japanese and the Norwegian Domestic Communication Satellite Systems, the ESA Marisat System. The last group of papers deals with a number of problems connected with the technology and applications of remote sensing and observations from Satellites. A comprehensive survey on Earth and Ocean Dynamics satellites and systems is also included.

In closing, I wish to express my appreciation to Dr. J. Bouttes, Chairman of the International Program Committee, for his efficient and invaluable help in collecting the manuscripts from authors. The cooperation of the I.A.F. Secretariat is gratefully acknowledged. Many thanks are also due to all authors who have graciously complied with the sometimes very strict requirements I had to exact from them. Last but not least I would like to thank Mr. Strange and the staff of Pergamon Press for their cooperation in making this volume available as economically and rapidly as possible from the authors original typescripts. In order to achieve this it was not always possible for them to obtain larger or clearer illustrations for some of the contributions, but it is hoped that the reproductions printed are sufficient for the reader.

EARTH, AN OPEN SYSTEM

The Use of Solar Energy

André L. Jaumotte
President of the University of Brussels

1. INTRODUCTION

I would like to start this paper by making two observations :

- a) since the reports of the Club of Rome on growth [1, 2] and the letter from M. Mansholt to President Malfatti, there has been too much of a tendency to view the earth as a closed system living in a state of autarchy on its non-renewable resources, whereas it is an open system nurtured by the enormous amount of energy which is sent out to it from the sun;
- b) my second remark refers to man in particular and living beings in general. They are - and the definition comes from Christian de Duve, the Nobel Prizewinner in Medicine - self-programmed, self-manufactured, self-maintaining and self-decaying chemical machines, running either directly or indirectly with the help of solar energy.

These two observations on the earth and on living beings characterise the essential part played by this energy, solar energy, which comes to us from space.

Edmond Rostand says in Chantecler, "O Sun, without which things would be but what they are". In fact, it is much more a case of, "O Sun, thanks to which things are what they are".

2. THE EARTH'S THERMAL BALANCE-SHEET [3 , 4 , 5 , 6]

The solar constant or the flow of solar radiation received per unit of air at the surface of a sphere concentric with the sun and having a radius equal to the distance separating the earth and the sun is 1.395 kW m^{-2} .

The energy intercepted by the globe is enormous : $1.8 \cdot 10^{11} \text{ MW}$ (180 milliard MW), whilst the total thermal energy consumed on earth in 1970 was $8 \times 10^6 \text{ MW}$ (calculated on the basis of 2 thermal kW per inhabitant, and 4×10^9 inhabitants). Thus, the energy received from the sun is 22,500 times as much as all the forms of energy consumed by man every where on earth. If we confine ourselves to the European Economic Community, solar energy received represents about 150 times the total consumption of primary energy.

It is interesting to note that the product of 1 p.c. of the photosynthesis

*

The numbers given in brackets refer to the attached bibliography

(forests, agriculture) of about one tenth of the land surface of the globe corresponds to an energy value of some millions of MW, that is, of the same magnitude as the present world consumption of energy.

Assuming that it is uniformly spread over the whole earth, the sun's energy contribution is 0.348 kW/m^2 .

How can the earth's thermal balance-sheet be drawn up ?

35 p.c. of the energy is reflected outwards by the fact of selective and diffused reflection from the atmosphere and the surface of the earth. The other 65 p.c. is absorbed by the earth-atmosphere system. What becomes of this energy ?

The answer is that it is redistributed as a result of multiple transformations which take place in the atmosphere, the hydrosphere and the lithosphere. These transformations produce the evaporation of water, the winds, vegetation and all forms of animal life.

Fig. 1 schematises the annual average radiative balance for the whole earth.

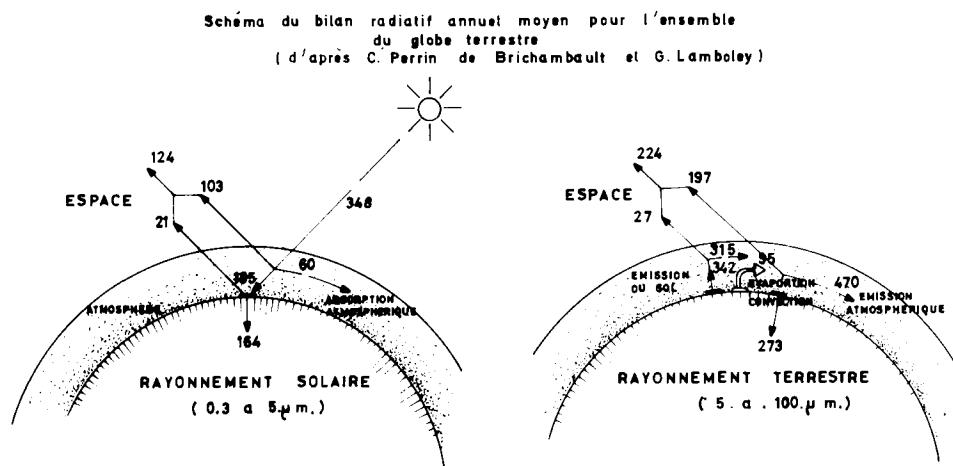


Fig. 1. Composantes moyennes annuelles du bilan énergétique pour le système Terre - Atmosphère
(Watts par m^2)

Does the activity of man provoke any significant change in the thermal balance sheet of the earth ?

Taking 1970 as a year of reference, heat losses corresponding to the production and consumption of energy in the world amounted to :

- 0.015 W m^{-2} over the whole of the earth's surface ($510 \times 10^6 \text{ km}^2$);
- 0.054 W m^{-2} over the continents only ($143 \times 10^6 \text{ km}^2$).

This last figure should be compared with the average value of the flow of geothermal heat at the surface of the earth : 0.063 W m^{-2} .

The movement of heat-loss connected with human activity is towards the atmosphere, in the form of perceptible and latent heat. The loss should be compared with the 95 W m^{-2} of the net energy which is transferred from the earth to the atmosphere and which conditions the world's weather. It only represents 0.017 p.c. of this. Its influence is therefore negligible.

The situation is different at local level, where repercussions of human activity are observable in the microclimate. Let us take two figures, both relative to the Federal Republic of Germany. For the whole of the Federal Republic the global consumption of energy is 1 W m^{-2} , and for the Ruhr, 17 W m^{-2} . For the USA, the figure is 0.235 W m^{-2} .

In the temperate zone the study of the climates of built-up areas with population figures in excess of a million has revealed an increase in temperature of about $0.7 - 2^\circ\text{C}$ (the annual average). In conjunction with atmospheric pollution, this phenomenon may bring about localized microclimatic changes (convective phenomena, frequency of fogs which develop into low strati, etc.). On the other hand, there is no proof that the thermal structure of air-masses is significantly disturbed by the fact of their passing over urban and industrialized areas despite the increase in localized harmful effects.

The increasing concentration in the atmosphere of CO_2 and aerosols is much more significant. These are the left-overs from combustion, from the fact that human activity has taken place.

The effects of the two types of action are incompatible, but the balance tends towards a cooling-down which will become more marked unless nuclear energy more or less takes over from fossil fuels.

A more sophisticated analysis of the meteorological effects of localized releases of large quantities of heat should be the subject of an integrated programme of temperature-monitoring. Temporal follow-up should be provided by a specially designed satellite equipped, according to the suggestion of Penner and Wolfhard [6] with special detectors [7].

3. HOW CAN SOLAR ENERGY BE PUT TO USE ?

Application of Low-Temperature Thermal Energy [8, 9, 10, 11, 12]

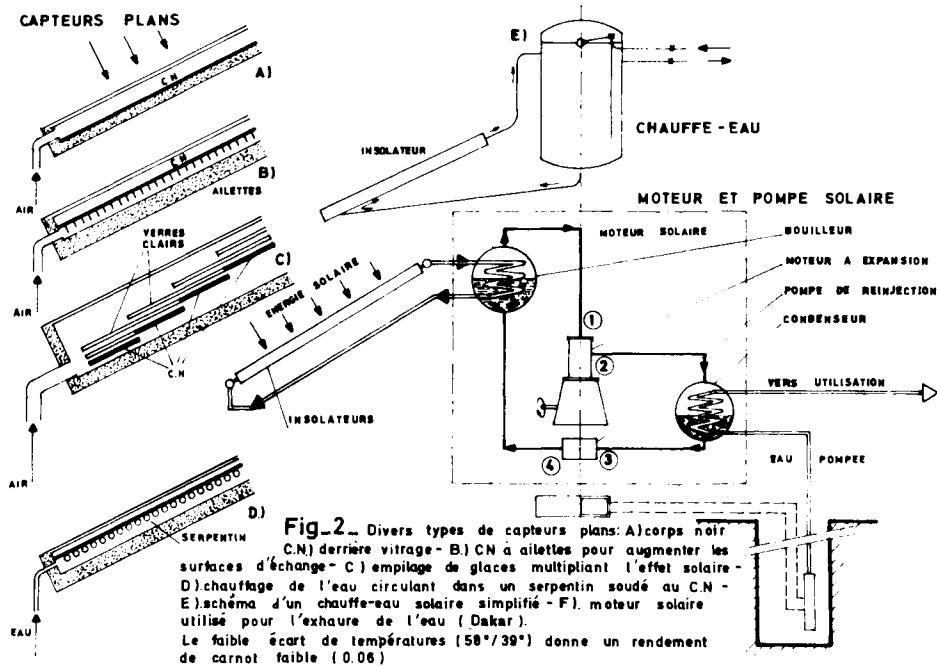
The sun radiates very much like a black body with a temperature of 5.800°K . As far as the short waveband is concerned, this radiation is precisely limited to 0.2μ , reaches its maximum at 0.5μ and becomes insignificant over 5μ .

Where loss through conduction or convection does not occur, a blackened disc exposed under normal circumstances to solar radiation in an environment of 300°K will, in theory, assume a temperature of just over 100°C . Natural convection reduces this temperature to about 60°C .

It should be noted that, when incident radiation is directed, the radiation reflected from the heated disc emanates in all directions. If the back of the disc is polished in order to limit its radiation, the temperature of the disc in a vacuum will reach 150°C instead of 100°C .

If a sheet of glass is placed in front of the disc at a certain distance from its incident side, the radiation will be trapped (the greenhouse effect). When atmospheric absorption is taken into account, incident radiation at ground level lies between 0.3 and 2.5μ . Glass does not absorb this range. On the other hand, it is practically impenetrable to infra-red radiation reflected from afar (between 4 and 30μ). A black disc which reaches 100°C in a vacuum will have its temperature increased to 150°C behind a single sheet of glass, to 185°C behind a double sheet and to 210°C behind a triple sheet. With 0.86 as the factor of transmission of a sheet of glass, the temperature is brought

down to 190°C . It should be added that these glass covers placed close to the receiving surface form an efficient anticonvective device. Fig. 2 schematises the positioning of flat-plate collectors based on the greenhouse effect.



A propos, the use of greenhouse in agriculture is a development of this kind of collector.

Another stratagem - that of the non-radiating structure - is possible. The amount of radiation reflected back towards its source is reduced by adding a cellular element to the heat-collector. Let us imagine that a honeycomb structure is placed above the surface of the heat-collector (the black body). This honeycomb structure is made up of, let us say, hexagonal aluminium tubes, the height of which is about ten times their width. The rays reflected by the black body which forms the base of the structure are not parallel to the axes of the tubes. They reflect off their inside faces and emerge at a narrow solid angle. Loss by radiation is thus limited. (Fig. 3.).

All non-concentrating flat-plate collectors use one or the other of these principles.

We will take two examples - the solar dwelling and the solar irrigation pump.

3.1. Dwellings [13].

One straightforward use of a heat-collector based on the greenhouse effect is to heat water for domestic purposes. This use is widespread in regions with a high rate of sunshine. The daily heat supply is stored in a tank containing hot water, the temperature of which can be as high as 95°C . Circulation is effected by means of a thermal siphon. By way of an example, in a hot climate

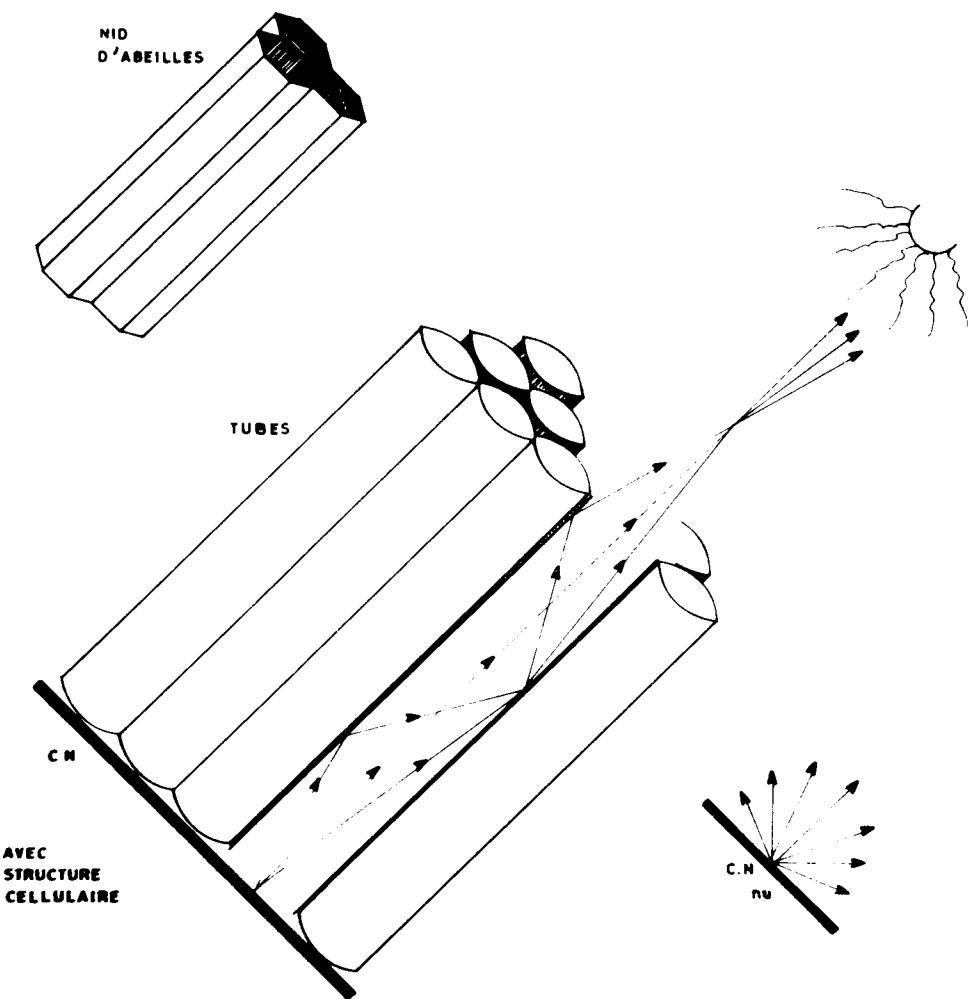


Fig. 3 - Rayonnement du corps noir (C.N), nu, et avec une structure cellulaire.

(as in Israel), a collecting surface of 2 m^2 is sufficient to bring the water in a 200 litre tank up to 65°C . This type of water-heater can also be used in a temperate climate provided that a reserve supply of heat is available. Solar water-heaters are in commercial production in Israel, Japan, the USSR and the USA. (Fig. 4).

The distillation of sea-water and brackish water is also worthy of mention. Numerous attempts have been made in different countries to heat houses by

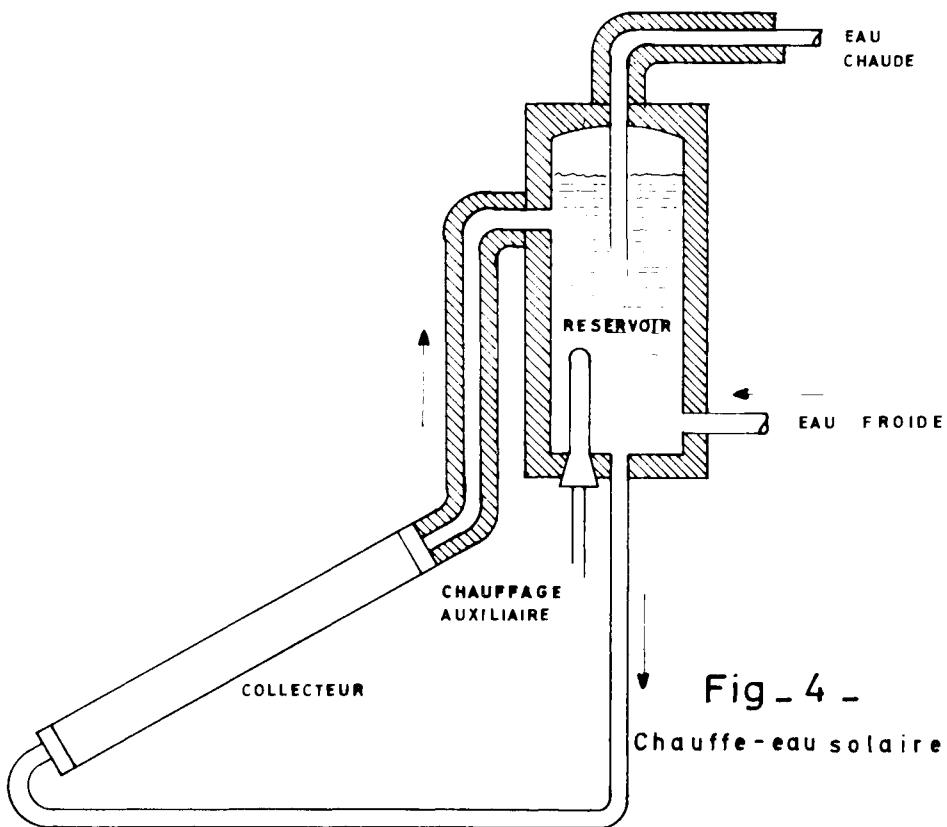


Fig. 4 -
Chauffe-eau solaire

means of solar energy.

Prototypes exist in the USA (California and Florida), in France [Chauvency-le Château (Meuse) and Font-Romeu-Odeillo (Pyrénées Orientales)], and in Germany (Aachen)].

Two systems are possible, one using hot water and the other hot air. Both are backed up by a reserve supply of energy.

The system using hot water is a development of the water-heater. The heat-collector of the hot-air system may either be mounted on the roof or fixed along a vertical wall. In our hemisphere it should face south.

Because it is lighter than the air inside the dwelling, the hot air contained in the glass suntrap of a roof-mounted unit will have to be transferred by mechanical means to the rooms to be heated.

Where the glass suntraps are mounted on vertical outside walls, it becomes possible to effect the processes of collection, storage and the circulation of air by natural means (thermal circulation). The sun's rays pass through the glass panels (1) and are absorbed by the collecting surface (2) which consists of concrete slabs or a quantity of water. Natural processes set up a permanent thermal current which moves in a clockwise direction. A system of this type functions without mechanical assistance and will continue to work after sunset thanks to the partial storage of heat received (for this purpose a 35 cm thick

wall which stores about half the heat received is necessary). (Fig. 5).

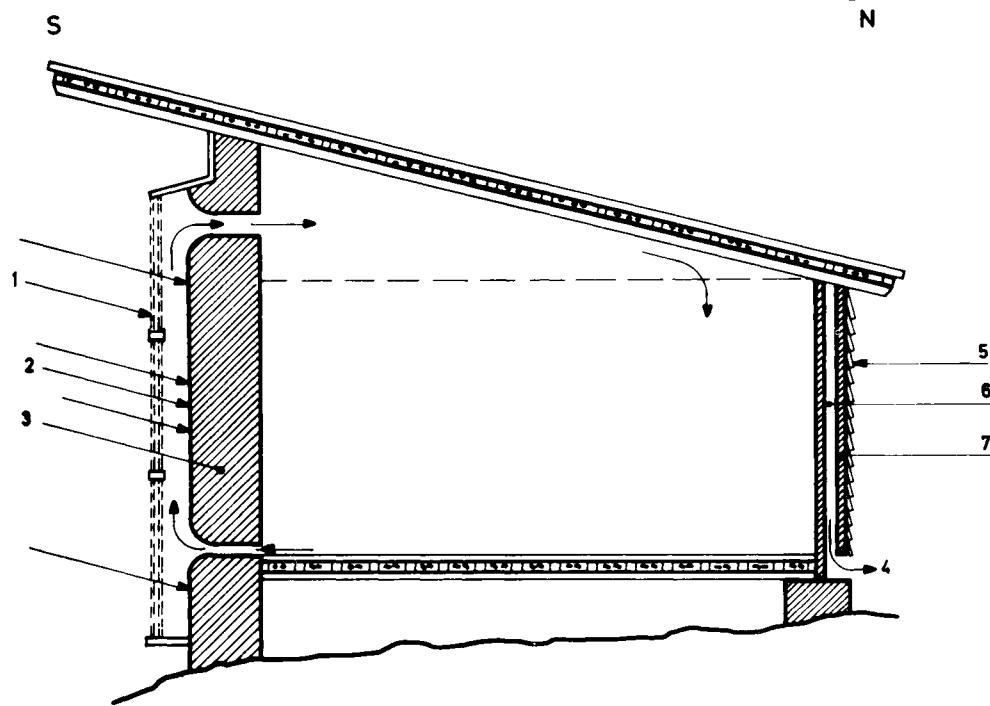


Fig. 5 - Chauffage solaire par serre sur facade verticale

The heat from the suntrap can be used both to ventilate and to heat rooms. Fig. 6 and 7 illustrate a recent patent CNRS - ANVAR (TROMBE and MICHEL). As J. MICHEL [13] explains, "Fig. 6 and 6b represent the heating circuit. Hot air, circulating as in Fig. 5, enables the rooms to the rear of the wall to be heated (see the legend to Fig. 6a and 6b). In the models given in Fig. 7a and 7b (see legend), the outside wall acts as a suction device and air is expelled at the top of the suntrap. Pressure decreases inside the dwelling and cold air is able to enter from a north-facing outside wall or an air-conditioning unit. With the abovementioned improvement, solar suntraps have a much wider range of application than when they are only used for heating purposes".

It will be noticed in Fig. 6a and 6b, and in Fig. 7a and 7b, that the receiving surface of the suntrap is surmounted by laterally fixed pipes which frame its illuminant sections. These pipes increase natural circulation in the same way as the chimney over a traditional fireplace.

In the majority of cases solar heating has to be backed up by a secondary, i.e. electric, source of heating.

In the basic solar house at Chauvency-le-Château (Meuse), solar energy provides 75 p.c. of the heat necessary to maintain a temperature of 20°C. The house (106 m²) has been occupied since 1972.

Some solar houses are under construction at Odeillo-Font-Romeu (Pyrénées Orientales). The secondary supply heat input is electric. Another example is the

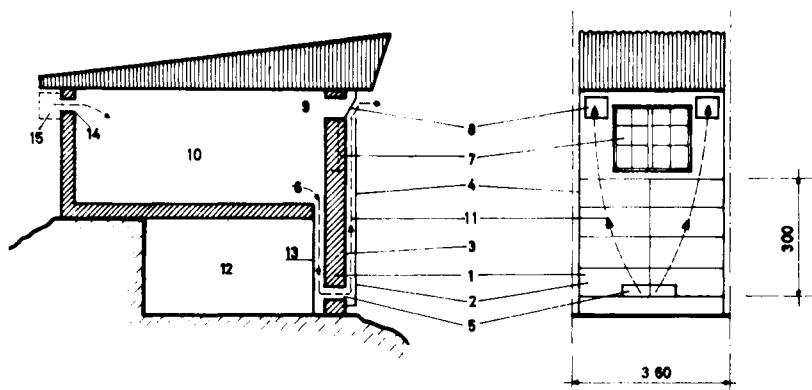


Fig. 6^a et 6^b. Dispositif du chauffage avec cheminées latérales d'appel d'air (1) mur accumulateur ; (2) vitrages ; (3) surface réceptrice du rayonnement ; (4) circulation de l'air ; (5) ouverture inférieure ; (6) passage de l'air aspiré ; (7) partie éclairante ; (8) passage de ventilation fermé ; (9) entrée de l'air chaud dans la pièce ; (10) pièce à chauffer ; (11) circulation de l'air ; (12) étage inférieur garage ; (13) cloison ; (14) passage de l'air de climatisation fermé ; (15) climatiseur.

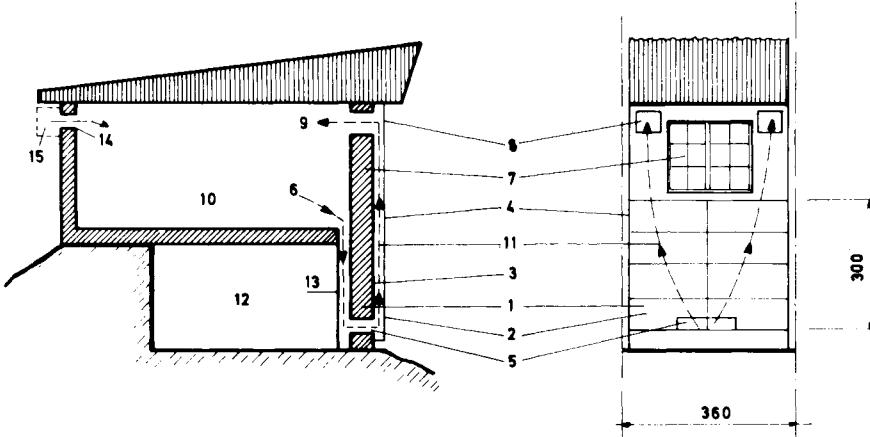


Fig. 7^a et 7^b (9) Sortie de l'air chaud vers l'extérieur ; (6) aspiration d'air de la pièce ; (14) entrée de l'air climatisé.

Thomason house, several examples of which have been built near Washington.

Heating can be combined with air-conditioning by using an absorption refrigerator which requires a heat source of about 90°C. Happy coincidence - the need for cooling coincides with periods of intense sunshine. Fig. 8 outlines a dual-purpose heating and air-conditioning unit.

The only problems which remain to be solved are those of reducing the price by 50 p.c. through mass-production, and of persuading builders to employ this type of heating. The building industry is traditionally slow to adopt new technologies.

3.2. Irrigation [14, 15].

Cyclic needs are best satisfied by solar energy which itself is cyclic. This

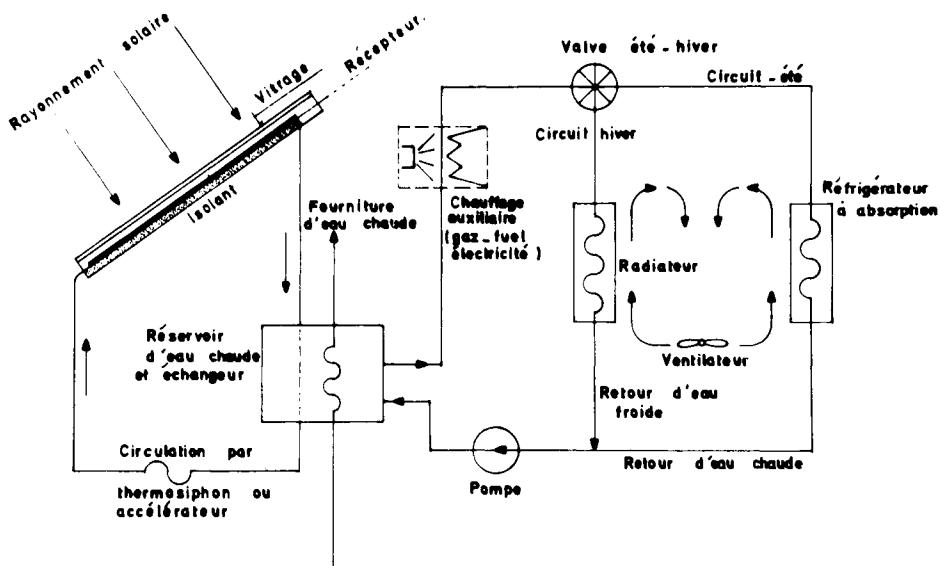


Fig. 8 -

is the case with water-pumping for irrigation and stock-watering in arid regions with a high sunshine level (e.g. the Sahara, the Middle East and the High Plateaux of Brazil).

A financially interesting method has been developed by using flat-plane heat-collectors and a low-temperature thermal cycle operating between a temperature of 60 - 70°C obtained at the exit of the collectors, and the temperature of the water being pumped. This acts as the cold-water source. A diagram will be found in Fig. 2. Developments are based on research carried out over 10 years or so at the Institut de Physique Météorologique in Dakar (Professor Masson).

The heat-collector consists of a cover made of self-supporting sheets of asbestos 5.50 m long (Fig. 9). The bottom of each unit includes a thin, blackened metal sheet, three water circulation pipes and a sheet of glass. The water is heated to 60 - 70°C. Circulation between the hot and cold water sources is effected by means of a thermal siphon. The motor (pistons or turbine) uses butane, propane or freon between 60 - 70°C and 35 - 40°C as fluid. The pressure level depends on the nature of the fluid.

The water from the pump is stored in fibreglass tanks. The following table lists developments of the Société Française d'Etudes Thermiques et d'Energie Solaire (SOFRETES), both in service and under construction. (see page 11).

Figure 10 illustrates one of these developments.

Units equipped with an alternator and producing electricity are obviously possible.

A 25 kW unit is under construction for Mexico (San Luis de la Paz).

The development of this kind of unit is related to its economic viability. Its competitor is the diesel motor which requires fuel-oil, maintenance and trained personnel.

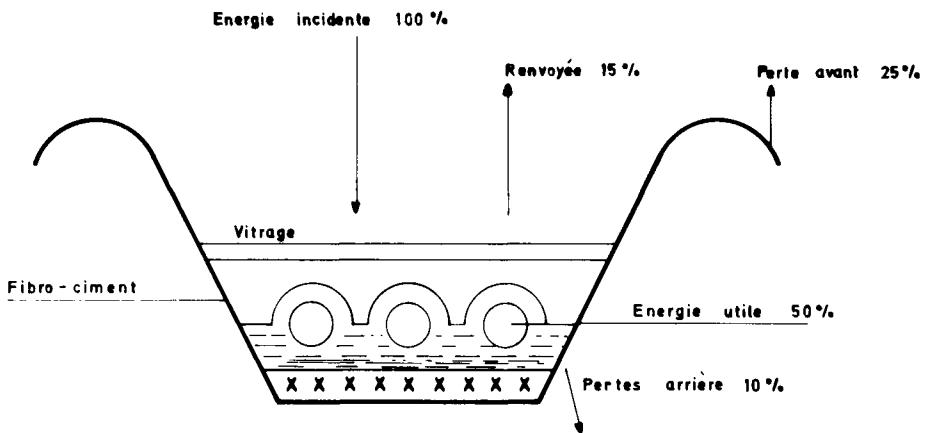


Fig. 9 - Éléments de toiture " CANALETAS ..



Fig. 10

Figure 11 gives a comparison in terms of the rate of interest on the sum invested for a 50 kW motor that runs 6 hours a day for 350 days a year i.e. 2.100 hours [15].

It has been established that the solar motor is heavily penalised by current high interest rates. Its price should be brought down to half its present level by mass-production. This can be done.

Figure 12 gives a comparison with the energy output of an oil-fired urban thermal power-station with a medium-voltage power-transmission line.

Solar installations will be profitable from now on if they are located more than 60 km from an urban power-station.

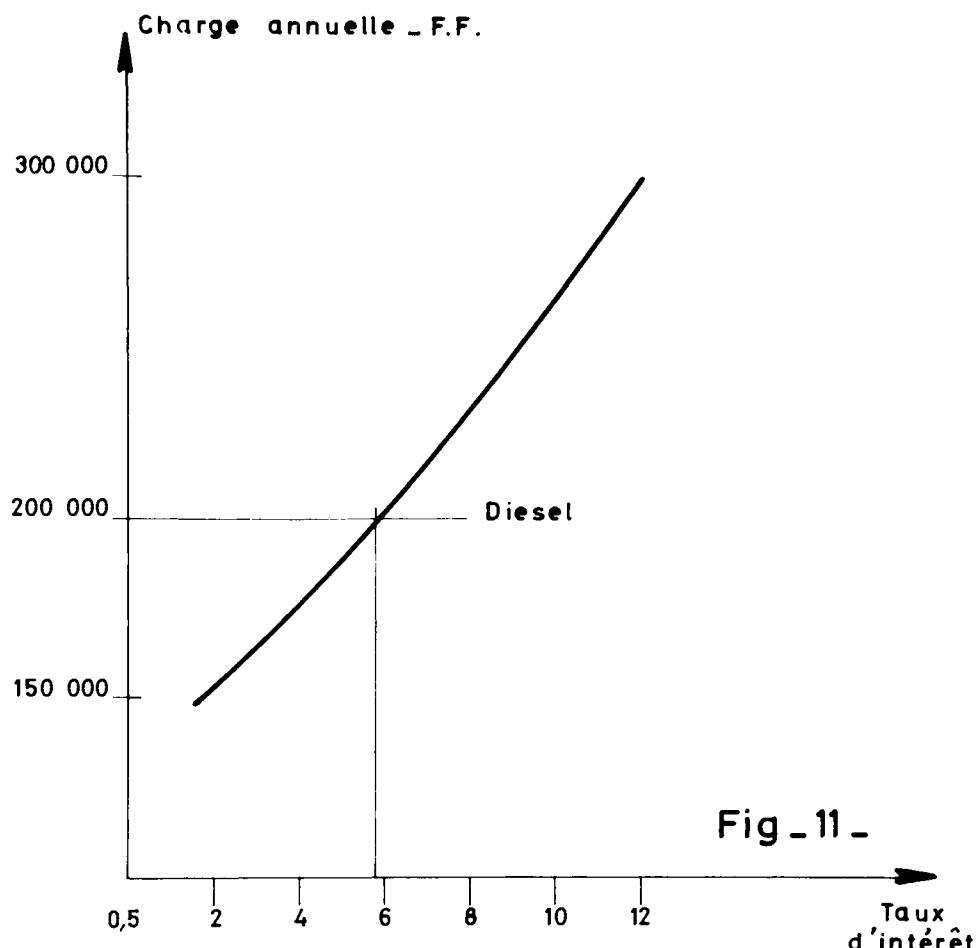
It seems that this type of unit has an immediate future where the provision of water for village, stock-watering and agricultural purposes is concerned.

IN SERVICE

Location and date of going into service.	Surface area of the collector.	Yield	Manometric height	Length of operational period
Institut de Physique Meteorologique Dakar (Senegal) 1966	12 m ²	1200 l/hr	15 m	5 to 6 hr
" 1968	88 m ²	6 m ³ /hr	25 m	5 to 6 hr
Ecole Inter-Etat d'Ingenieurs de l'Equipement Rural Ougadoudou (Upper Volta) 1971	30 m ²	2 m ³ /hr	20 m	5 to 6 hr
Office de l'Energie Solaire (Niger) 1973	60 m ²	6 to 7 m ³ /hr		5 to 6 hr
Chinguetti (Mauritania) 1973	72 m ²	8 to 10 m ³ /hr		5 to 7 hr

UNDER CONSTRUCTION OR PROJECTED

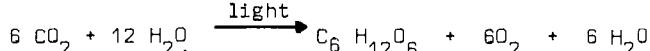
Location	Surface area of the collector	Yield	Manometric height	Length of operational period
Upper Volta pump for stock-watering	50 m ²	5 m ³ /hr	30 m	6 hr
Senegal 5 pumps	85 m ²	8 to 10 m ³ /hr	20 m	5 to 6 hr
Algeria Wilaya des Oasis	350 m ²	21.5 m ³ /hr (6 l/sec)	30 m	5 to 6 hr
Chad 5 village water supply pumps	70 to 200 m ²	5 to 10 m ³ /hr	15 to 60 m	5 to 7 hr



4. APPLICATIONS INVOLVING PHOTOSYNTHESIS [16].

It was Robert Mayer who, after discovering the principle of the conservation of energy, discovered photosynthesis. He wrote, "Plants take one form of energy - light, and convert it into another - chemical energy (1845).

Photosynthesis is a fixation of the carbon present in the atmosphere in the form of carbon dioxide. The reaction is the reduction by water of this carbon dioxide to form a glucide. The reaction is made possible by the energy contribution of the light absorbed by chlorophyllian pigments. The general formula is :



There is a store of energy in the molecule of glucide formed.

The energy generated by the combustion of an atom of carbon in the form of a glucide is equal to about half of what can be obtained by the combustion of a monocarbonated fraction of methane.

If the product of photosynthesis is defined by the ratio
stored chemical energy
light energy provided

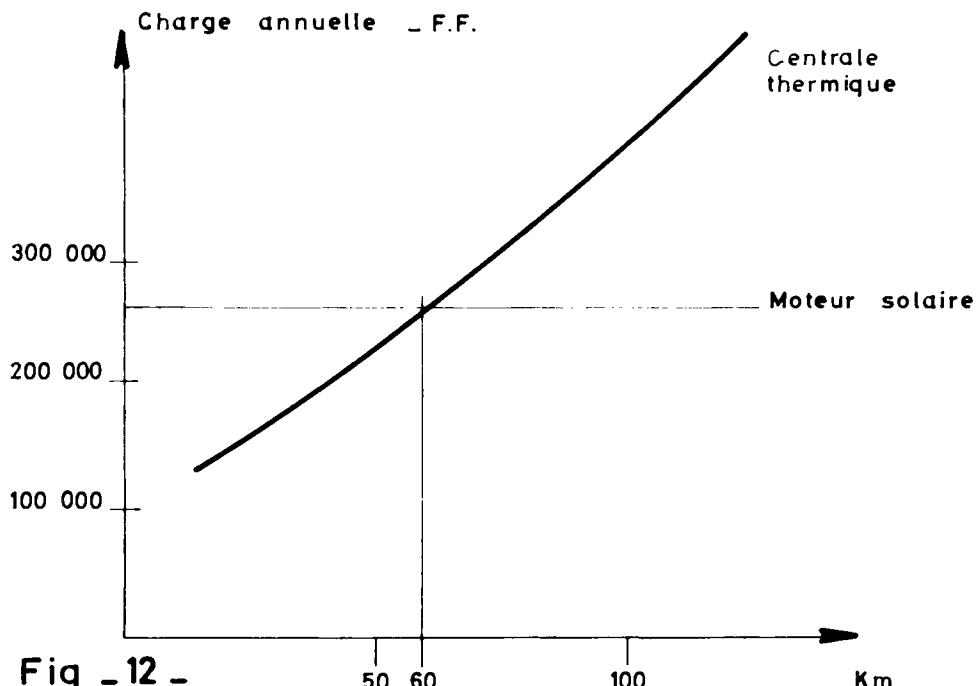


Fig - 12 -

its average value will be found to be of the order of 2 p.c. for forests, 1 p.c. for most of the flora in the temperate zone from May to June, 0.6 p.c. for the annual average and some thousandths where vegetation is sparse.

The average photosynthetic capacity of the foliage of a given area in our climate varies from 150 kg of carbon per hectare per day in June to 27 in December. Under average conditions these values should be reduced by 20 - 25 p.c.

Day and night metabolism re-processes some of the primary products of photosynthesis to form various substances : protides and lipids, and secondary substances such as polyphenols, alkaloids etc. This reprocessing is characterised by a loss of energy.

Thus, the energy yield corresponding to plant growth is from 10 to 20 p.c. lower than the original photosynthetic product.

The systematic study of possible improvements is one of the essential aims of modern agricultural science. The technology brought into existence by the green revolution has not exhausted the possibilities of the vegetable kingdom.

When everything has been taken into account, non-marine plants and the vegetable plankton of the oceans fix about 40 milliard metric tons of carbon between them, i.e. 15 - 20 times more than the carbon in the coal and oil consumed during the same period. Nor must it be forgotten that coal and oil are only the fossils of past photosynthetic processes.

An interesting use of waste vegetable matter is the production of fuel gases - so-called manure gases - by controlled anaerobic fermentation [17].

Manure-gas is made up of methane (55 - 60 p.c.) and carbonic gas. Its upper calorific limit is in the region of 5.500 kcal/m³. If the CO₂ is washed out with water, almost pure methane is obtained.

The technology is closely related to that used in decomposition by mud. Under economically viable conditions a metric ton of straw can yield 200 - 250 m³ of gas. The production of gas. The production of straw may be as high as 6 metric per hectare.

The gas can be stored in gasometers or under pressure in bottles.

Various developments exist, particularly at the Institut Agricole d'Algérie at Maison-Carrée (Algiers).

5. APPLICATIONS OF SOLAR ENERGY INVOLVING OPTICAL CONCENTRATION [8, 9, 10, 11].

5.1. Concentration.

Everybody has ignited a piece of paper place at the focal point of a magnifying glass. In the case of large surfaces, mirrors must be used.

The ideal is a one-piece parabolic mirror made of metal or glass. After the war, searchlight mirrors provided the wherewithall for numerous small solar ovens.

A reflector in the shape of a truncated cone can be used for small cookers. The Francia boiler (Fig. 13), which has a non-radiating cellular structure, uses this principle. These conical reflectors have one strange property; because the angle of incidence of the rays increases as they reflect back and forth, they are likely to retrogress before reaching their target. The field of optimization is narrow and must be carefully calculated. The size of one-piece parabolic mirrors cannot exceed a few metres.

Where large sizes are involved, it is necessary to make use of a mosaic of mirrors which are positioned according to the shape of the parabolic frame which holds them. If it is small enough, each element can be flat. The compromise solution of bending each elemental mirror is preferred. Mechanical shaping is effected by means of push-and pull-rods.

In the case of the large 1.000 kW solar oven at Odeillo (Pyrénées Orientales), constructed by Professor Trombe for the French CNRS (Fig. 14), the pull-rods are arranged on the back and act on pierced mirrors. Their individual parabolic form is approximate. The object of the exercise is to use a rule-of-thumb approach to try and obtain the smallest possible focal area of each mirror.

The Odeillo parabolic mirror consists of 9.000 silvered mirrors occupying a surface area of 2.500 m². With a focal length of 18 m , the gaussian image point should have a diameter of 17 cm . In fact, its energy is distributed in a gaussian curve. The maximum temperature at the centre of the focal area goes up to 3.800°C . The nominal energy of 1.000° kW is obtained in a zone 50 cm in diameter where the average temperature is 2.500°C . The levels of useful energy are 750 kW at 3.000°C and 300 kW at 3.500°C .

Such high temperatures are of no practical interest except in certain areas of high-temperature chemistry ; the study of hyper-refractories and thermal shocks etc.

The aim of the process of concentration is not to obtain an ideal focal point. The battery of mirrors, curved or flat, which is placed on a hillside forms an optimized focal area projecting into a collector [the MARTIN & MARRIET project 4]. Collectors have undergone considerable improvement with the development of space propulsion systems using solar energy to heat a liquid propellant.

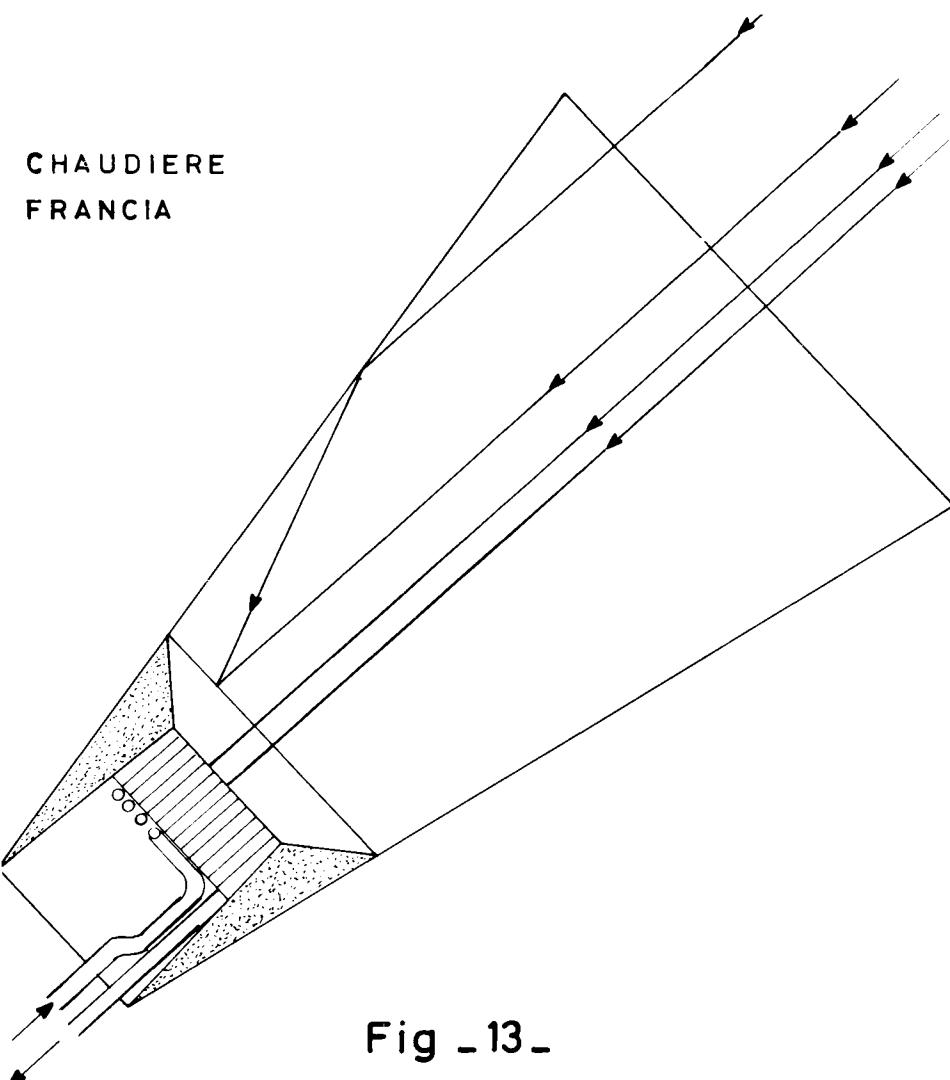


Fig. -13-



Fig. 14

With the authorization of the Laboratoire de l'énergie Solaire, Centre National de la Recherche Scientifique à Odeillo-Fontromeu, France.

5.2. Adjusting and Tracking.

Concentrating mirrors must obviously be turned towards the sun. The use of equatorial mounts of the type employed in astronomy is only possible in conjunction with mirrors of comparatively small size. Large mirrors are fixed and an intermediate collector is introduced which acts as a heliostat.

63 orientator mirrors measuring 45 m^2 each are arranged in tiers and reflect on to different areas of the Odeillo parabolic mirror. Fig. 15 shows the self-equipped automatic tracking device which carries out azimuthal and zenithal adjustments by means of a hydraulic lever.

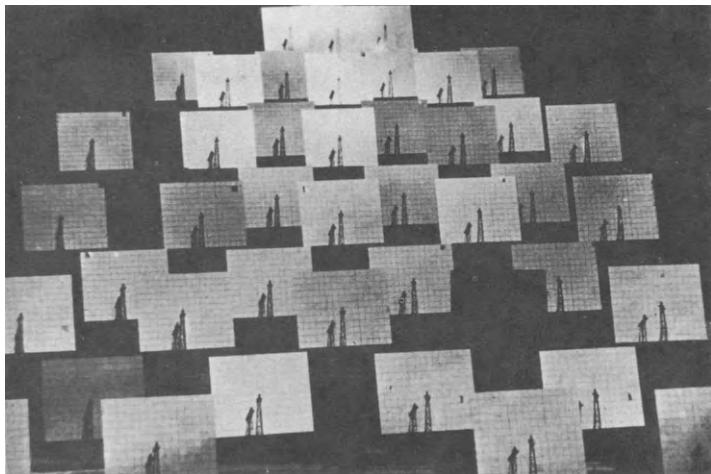


Fig. 15

With the authorization of the Laboratoire de l'énergie Solaire, Centre National de la Recherche Scientifique à Odeillo-Font-Romeu, France.

In Tashkent (Armenia), the Russians use mirrors mounted on trolleys which move around a circular track in accordance with the movement of the sun.

It is possible to dispense with the need for continual adjustment by using multi-directional collectors; these are East-West orientated cylindro-parabolic reflectors with a focal line in place of a precise focal point. This system lends itself to heating a fluid circulating in a channel placed on the focal line.

Finally, the concentration of solar energy raises few problems in the field of basic physics. The questions to be solved are of a technical, technological and financial nature.

5.3. Plans for Solar Power-stations [10, 19, 20].

Two solutions are possible depending on the degree of perfection of the concentrator.

Where cylindro-parabolic concentrating mirrors are used, the heat-carrying liquid can be raised to an average temperature. Fig. 16 outlines a double-storage unit where an exchanger produces the steam to drive the turbine.

Storage may either be passive (large insulated tanks), or effected by a change of state (eutectics). As part of an international programme organised by Honeywell, more than 300 binary, ternary or quaternary mixtures were studied

as means of storage in the 260 - 320°C range, and near to 600 in the 450 - 550°C range. The cheapest mixtures are :

KCl - K NO₃ ; NaCl - Na NO₃ ; NaCl - Na NO₃ - Na₂SO₄ .

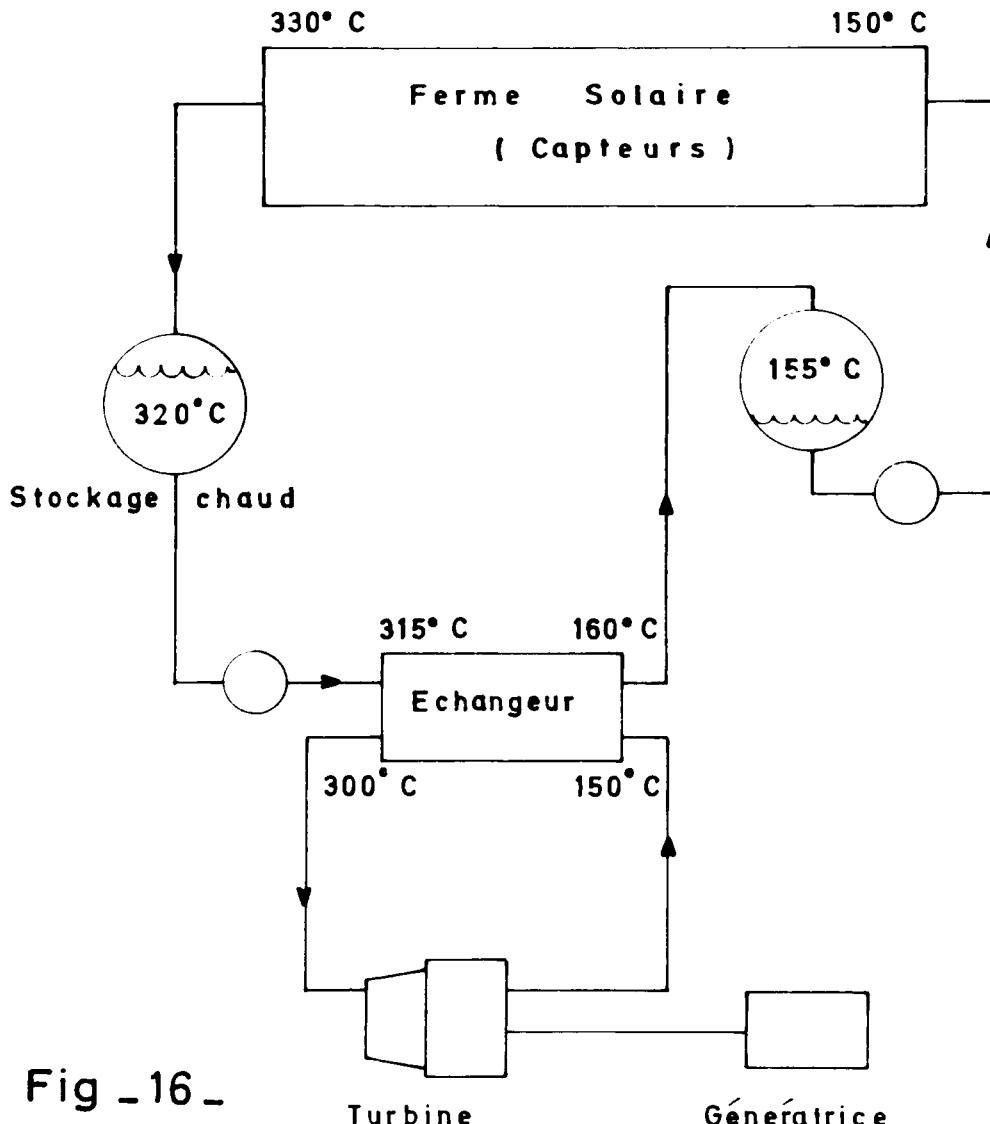


Fig. 16 -

Turbine

Génératrice

These are suitable for effecting storage at 300°C. A process which produces higher temperatures uses a device similar to that of the Odeillo oven, but simplified (Fig. 17). The heliostat is made of flat-plate mirrors, and the concentrator is a cylindro-parabolic mirror. Heat is concentrated on a tube placed along the focal line of the concentrator. It is drawn off by a heat-carrying liquid and stored in the form of perceptible or latent heat. It is obvious that such equipment is only viable in regions with a high sunshine

rate (at least 2.100 hours per annum). In MEINEL and MEINEL's project at the University of Arizona, the storage problem is avoided by linking the solar power-station to a hydro-electric unit. The economic viability of solar power-stations rests entirely on the hypotheses made on the investment in heliostat and collector equipment, and on the price of maintaining the installation. Given a quantity of development aid and a need for a sufficiently large number of units, solar power-stations could be competitive about 1985 and make a significant contribution to the production of energy about 2000 AD.

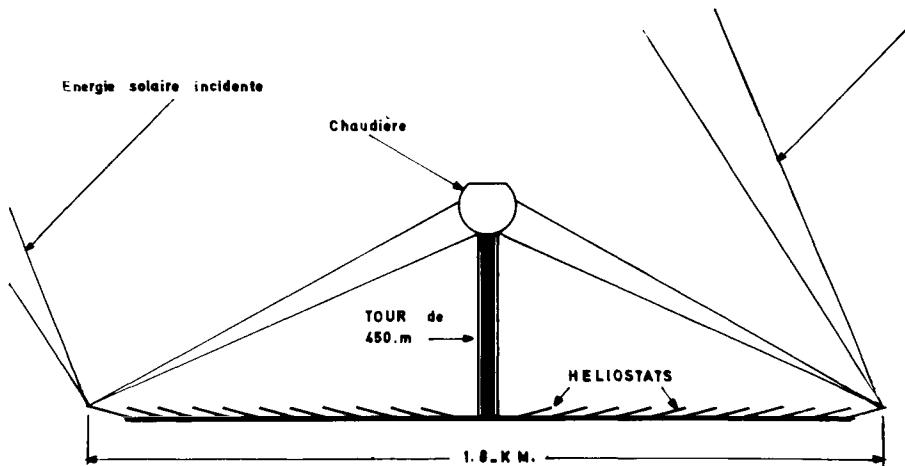


Fig. 17. Coupe d'une tour - collectrice d'énergie solaire

6. COLLECTION BY MEANS OF PHOTOCHEMICAL CELLS [21].

As a result of its employment in space, direct photon-electron conversion has made tremendous strides forward. The first unit was taken up by the Vanguard rocket in about 1958 and had a capacity of 0.1 W. The largest unit is that of Skylab-Apollo which has a capacity of 21 W. The great advantages of the photochemical cell are that it does not use a cooling liquid and that it responds to diffused radiation. Collection by monocrystalline silicon cells (p-n junction) - the only ones manufactured on a sufficient scale - produces a yield of 11 - 14 p.c. Progress still remains to be made as far as protection against corrosion, ageing and, above all, cost reduction are concerned. The cost of monocrystal silicon solar cells produced by hand in small numbers should be reduced by a factor of 200 - 300 to allow for competitive development (their yield is some 11 - 14 p.c. and their cost 7.000 dollars per m²). It should be borne in mind that between 1950 and 1971 the price of silicon diodes was brought down by a factor of 100 as a result of the extraordinarily rapid growth of the market. The present-day total annual output of solar cells in the USA is about 50 kW.

Current research on cells made of thin slices of polycrystalline material such as cadmium sulphide and gallium arsenide gives rise to the hope that their life will be extended and that mass production will considerably reduce costs. It should be noted that it is possible to combine a moderate concentration with photochemical cells to reduce the required surface. Even when conversion

by photochemical cell has been put on an economic footing, it does not solve the problem of storing the electricity produced. In certain geographical environments storage by hydro-electric pumping stations is possible. As far as the distant future is concerned, the production of hydrogen by electrolysis and its storage in the form of liquid hydrogen is a possibility. Present-day applications remain limited to space and to certain special fields here on earth.

On the Paris-Bordeaux air-route there is a radio-beacon with a range of 100 kms which has been running on photochemical cells since 1968 (Saint-Girons). Apart from these cases, a link in the form of a thermal cycle is required. A wide field of research remains open relative to photochemical cells and their use in industry.

7. SOLAR POWER-STATIONS IN SPACE : THE SATELLITE GENERATOR SSPS : SATELLITE SOLAR POWER STATION [22, 23].

The drawbacks of solar energy on earth are overcome by the use of a synchronous orbit in space : energy is available without interruption and erosive effects are avoided.

The development of the space shuttle (which will be available at the end of the present decade) enables us to imagine a satellite power-station on a synchronous orbit (Fig. 18).

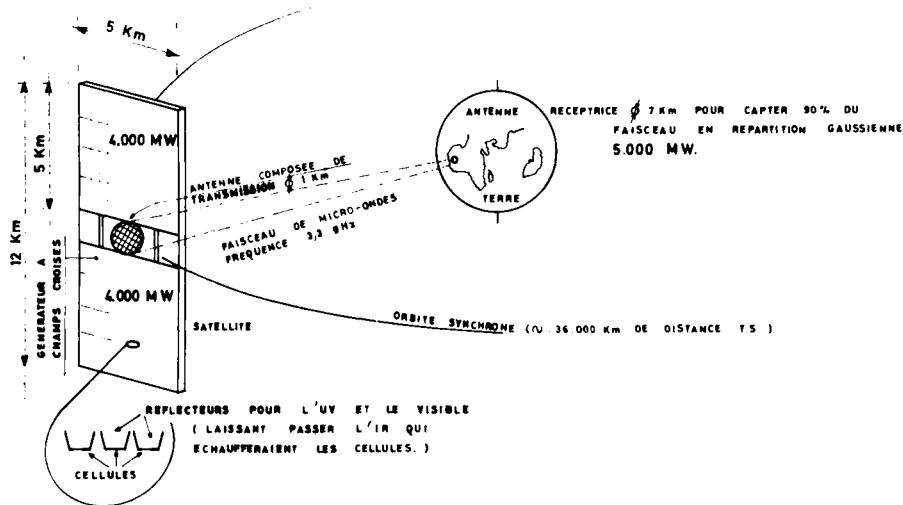


Fig. 18 - SCHEMA D'UNE CENTRALE SOLAIRE SPACIALE

The project is the work of Peter P. Glaser, Vice-President and Head of the Engineering Department of Arthur D. Little Inc.

The power-station would consist of two symmetrically-fixed collectors which would use photochemical cells to convert solar energy directly into electrical energy. This energy would be passed to microwave generators built into a transmitting aerial located between the two collectors. The position of the satellite would be such that the two collectors would always be facing the sun. The transmitting aerial would send microwave beams to a receiving aerial on earth in which the microwave energy would be converted into electricity. Microwave beams enable transmissions to take place regardless of atmospheric

conditions. To all intents and purposes, a 24-hour service is possible. The high yield resulting from the conversion of the microwave beam into direct current in the receiving aerial (0.75 already attained - 0.90 in the future) reduces loss to a minute fraction of the loss arising from thermodynamic conversion. Precautions should be taken to ensure that the microwave beam is not accidentally directed towards a target other than the receiving aerial.

To produce 5.000 MW on earth, the collectors would have to produce 2 x 4.000 MW. Fig. 18 gives the approximate size of such a Satellite Solar Power-station (SSPS).

Will this particular piece of science-fiction ever become science-fact? All that is needed is progress and the reduction of the cost of existing technologies. No innovations are necessary.

In 20 years, from 1958 to 1978, we will have moved from the Vanguard to the Apollo rocket, from the first little satellites to Skylab and the Space Shuttle. The challenge is the same. Financial estimates have been drawn up, but they appear to us to be too tenuous to merit reproduction.

8. CONCLUSIONS.

The use of solar energy has been arousing interest since the beginning of the industrial era.

A steam-engine using solar energy as fuel was the attraction of the 1878 Paris World Exhibition. A mirror concentrated the sun's rays on a boiler, and the power of the steam engine was used to print a newspaper called "The Sun". Machines using the same principle were built in California in 1901 and in Egypt in 1913. Built by Shuman and Boys, the engine at Meadi (Egypt) made use of cylindro-parabolic concentrators. Its 40 kW output was used in the irrigation of cotton fields.

The energy crisis has re-focussed attention on the possibilities of this source of clean energy.

Fig. 19 summarises the position of solar energy as a means of heating in relation to other sources of energy. It shows clearly that the threshold of competitiveness has already been reached.

Even now, variously ambitious programmes are being devoted to it, principally in the USA, the USSR, Japan, Australia, France, Germany etc.

The immediate target is the heating and air-conditioning of buildings. This is followed by thermal and photochemical conversion. A summary of the programmes is to be found in a recent O.C.D.E. publication [24].

Nonetheless, the programmes remain limited and there is no sign of any of those great waves of collective enthusiasm which lead to innovation.

The uses of the coming decade remain precisely defined: domestic water-heating, the heating and air-conditioning of buildings, the improvement of photosynthesis and the use of rejected matter (manure gas).

Solar power-stations, whether on earth or in space, remain a prospect which is tied in with the gradual evolution of other sources of energy: the breeder reactor, geothermal energy and nuclear fission.

With the implementation of a large-scale research and development programme, 2000 AD could see solar energy providing the USA with 35 p.c. of its domestic energy, 30 p.c. of its gaseous fuels, 10 p.c. of its liquid fuel and 20 p.c. of its electricity [25]. We consider this estimate to be very optimistic.

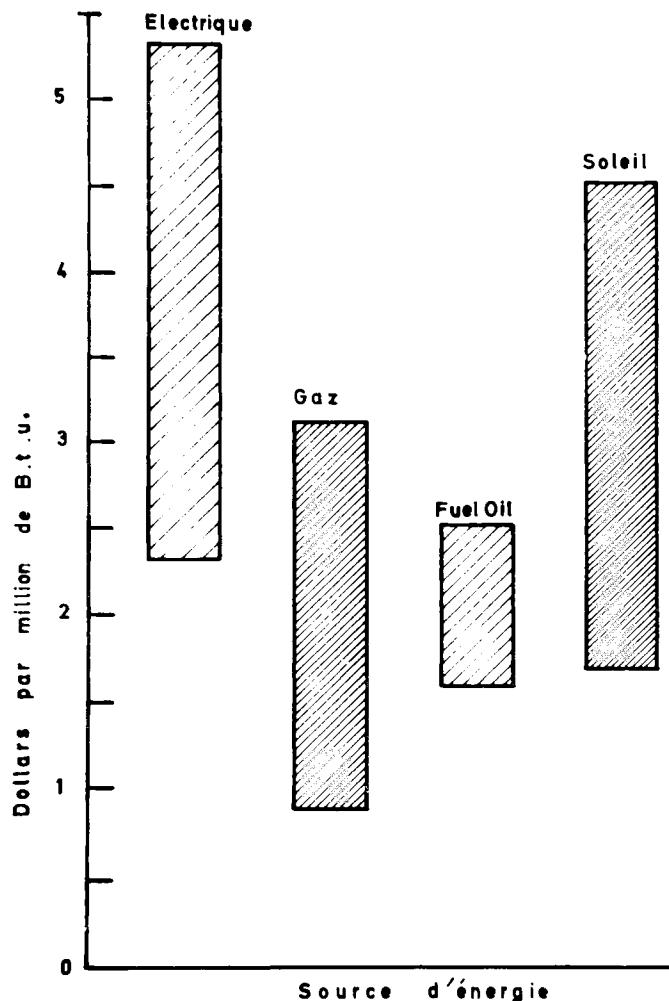


Fig. 19. Comparaison du prix du chauffage solaire par rapport à d'autres sources d'énergie

(Source: From Solar Energy as a National Resource, Department of Mechanical Engineering, University of Maryland, 1973.)

A word about pollution to finish. It is often stated that solar energy is non-polluting. This is not the case because the albedo i.e. the coefficient of reflection of the earth, is affected. In fact, useful energy is only derived from its reflected portion. As long as solar energy is employed in small amounts, the effect will remain minimal. In the case of large amounts, the modification to the albedo would be reflected in the climate.

Solar energy will slowly but surely become part of our as yet hazy future

development. At all events, whatever else may happen, humanity can rest assured that it will not die of starvation due to a lack of energy. Solar energy is there, free at source but expensive to exploit. It should not be forgotten that fossil energy is just as free at source, and that the cost is that of making use of it - mining, transport, transformation.

The energy-problem is straightforward. There is only one set of alternatives, i.e.

1. The use of solar energy in different forms
 - either in real time,
 - or in deferred time.
 2. The imitation of the sun by attempting to set up and control reactions which operate on the sun's surface.
- As far as the use of solar energy is concerned, space has made an important contribution on three levels :
1. A better knowledge of the balance of energy and resources available, and of the evolution of temperature at the surface of the earth.
 2. The know-how contributed by space technology towards the better use of solar energy, in particular through the development of high-yield photochemical cells and the idea of apparatus for collecting and focussing solar energy.
 3. The direct application of space technologies which makes projects like those of satellite generators possible.

BIBLIOGRAPHY.

- [1] Halte à la croissance? Fayard, Paris (1972).
- [2] Mihajlo Mesarovic et Eduard Pestel. Stratégie pour demain? 2e Rapport au Club de Rome, Seuil, Paris (1975).
- [3] J. Jacquet. L'influence de l'homme sur le bilan thermique de la planète. Électricité de France. Bulletin de la Direction des Etudes et Recherches. Série A, 3/4, 5-21 (1974).
- [4] Influence des activités de l'homme sur le cycle hydrométéorologique - Compte-rendu des troisièmes Journées de l'hydraulique, Paris, 14-18 sept. 1974, 2 volumes, 604 pages, Société Hydrotechnique de France.
- [5] S.S. Penner et L. Icerlan. Energy - Volume I - Demands, Resources, Impact, Technology and Policy. Addison-Wesley Publishing Company Inc. Reading Massachussets (1974).
- [6] S.S. Penner et H.G. Wolhard. Monitoring of the Climatological Impact of World-Wide Thermal Loading associated with Energy Consumption - 22 pages.
- [7] K. Ya. Kontratyev, O.B. Vasilijev, O.M. Pokrovsky et G.A. Ivanyan. Remote Sensing of Natural Formation from Measurements of Radiance Coefficients. Acta Astronautica, 1, 1415-1426 (1974).
- [8] 2000. Le Soleil. La documentation française, Paris (1975).
- [9] Ivan Peyches. L'énergie solaire : Mythe ? ou salut de demain ? Sciences et Techniques, 3, 5-12 (1973).
- [10] Ivan Peyches. Du nouveau dans la conquête de l'énergie solaire. Sciences et Techniques, 7, 5-14 (1973).
- [11] Le soleil au service de l'homme. Colloque UNESCO, 2-6 juillet 1973 - 300 communications - collection des pre-prints.
- [12] Rainer Köhne. Der Flache Sonnenkollektor. Energiebilanz und Wirkungsgrad. Deutsche Luft und Raumfahrt Forshungsbericht, 33, (1975).
- [13] J. Michel. Le chauffage par rayonnement solaire. Sciences et Techniques, 13, 15-22 (1974).
- [14] Jean-Jacques Beauventre. Une pompe solaire à la disposition des pays sahariens. Sciences et Techniques, 13, 48-52 (1974).
- [15] Documentation SOFRETES, Montargis, France.
- [16] Alexis Moyse. La photosynthèse : l'énergie solaire au service de l'agriculture. Sciences et Techniques, 13, 15-22 (1974).
- [17] Marcel Isman. Les problèmes de l'énergie. Energie solaire et ... gaz de fumier. Sciences et Techniques, 18, 17-22 (1974).
- [18] William R. Cherry, Harnessing Solar Energy : The Potential - Astronautics and Aeronautics, 8, 30-36 (1973).

- [19] A.F. Hildebrandt et L.L. Vant-Hull. A tower-top focus solar energy collector, Mechanical Engineering, 9, 23-27 (1974).
- [20] A.B. Meinel et M.P. Meinel, Physics Look of Solar Energy, Physics Today, 2, 44-50 (1972).
- [21] J.J. Loferski - Large-scale solar power via the photovoltaic effect. Mechanical Engineering, 12, 29-32 (1973).
- [22] Peter E. Glaser, Satellite Solar Power Stations to Meet Future Energy Demands, Industries Atomiques et Spatiales, 4, 77-95 (1973).
- [23] Peter E. Glaser. To-morrow Solar Energy ? Industries Atomiques et Spatiales, 1, 26-35 (1974).
- [24] Science et Technologie pour l'Energie. OCDE, 173 p., Paris (1975).
- [25] Solar Energy as a National Energy Resource. NSF/NASA Solar Energy Panel, December 1972.

ADDITIONAL ITEMS.

NSF/NASA - Solar Energy Panel Report. An Assessment of Solar Energy as a National Energy Resource, Dec. 1972, Mechanical Engr. Dept. Univ. of Maryland

ALICH, J.A., INMAN, R.E. : Effective utilization of solar energy to produce clean fuel, 164 p., NTIS (1973).

ALLEN, R : Proceedings of the solar heating and cooling for buildings workshop held in Washington on March 21-22nd, 1973, Part. A : Technical Sessions, 231 p., NTIS (1973).

ARGONNE NATIONAL LAB. : Solar Energy Evaluation Group Report. 50 p. NTIS (1973).

CLARK, A.F., DAY, J.A. et al. : The shallow solar pond energy conversion system, an analysis of a conceptual 10-MWe plant, 32 p. NTIS (1974).

DANIELS, R : Direct use of the sun's energy. Yale U.P. (1975)

DUFFIE, J.A., BECKMAN, W.A. : Solar energy thermal processes. 416 p. John Wiley, (1974).

FRY, L.J. : Practical building of methane power plants for power and electricity, natural fertilizer, labor saving and rural energy independence. Diags, photos, D.A. Know (1974).

INSTITUTE OF ENVIRONMENTAL SCIENCES 1974 proceedings, 2 vols. : Cost effectiveness in the environmental sciences. The Energy crises and energy from the sun. 668 p., Instrument Soc. of America (Distr. by Wiley) (1974).

JONES, B.W. : Ed.: Energy and housing. A symposium held at the Open University, Milton Keynes, 160 p., Pergamon (1975).

KINSELL, L.L. COULSON : Solar and Terrestrial Radiation : Methods and Measurements. 336 p. Academic Press (1975).

PEYTURAUX, R : L'Energie solaire, PUF, Paris.

RINGE, A.C. : Solar energy and Wind power : a selected bibliography, 80 p. NTIS (1974).

- SAYED, M, PARTAIN, L. : Effect of shading on CdS/Cu(x) Solar Cells and optimal solar array design. 52 p., NTIS (1973).
- SOLAR ENERGY; A BIBLIOGRAPHY, 360 p., NTIS (1974).
- SOLAR THERMAL ELECTRIC SYSTEMS. 288 p., NTIS (1974).
- STEOVICH, V.A. : Solar Energy. 478 p. NTIS (1974).
- VALE, B., VALE, R. : The Autonomous house; planning for self-sufficiency in energy, 111 illus., Thames & Hudson, (1975).
- WILLIAMS, J.R. : Solar energy : technology and applications. 144 p., Wiley (1974).
- ZENER, C., LAVI, A. ROTHFUS, F. et al.: Solar Sea Power, 113 p., NTIS, (1974).

THE ELECTRICAL POWER SYSTEM FOR SPACELAB

B.Gohrbandt, E. F. Schmidt
AEG TELEFUNKEN, Hamburg/Wedel

1. INTRODUCTION

As part of NASA's SPACE SHUTTLE within the scope of the American Post Apollo Program, the European SPACELAB is one of the most important payloads of the SPACE SHUTTLE Orbiter. Simultaneously, it represents the European entry into manned space flight. As such, it is designed to accommodate a variety of different experiments under space conditions. In contrast to the previous practice of conducting experiments in space by automatic operations, up to four scientists will themselves directly carry out their experiments in SPACELAB. To accomplish this, environmental conditions for the scientists during launch and orbit of SPACELAB will be such that more or less untrained men can work in SPACELAB under "shirt-sleeve" conditions.

Based on a Memorandum of Understanding between NASA and ESA, the European Space Agency, ESA is the customer of a consortium of European industry, which is located in ten European countries. These countries contribute to the overall financing of SPACELAB in accordance with figure 1. As can be seen

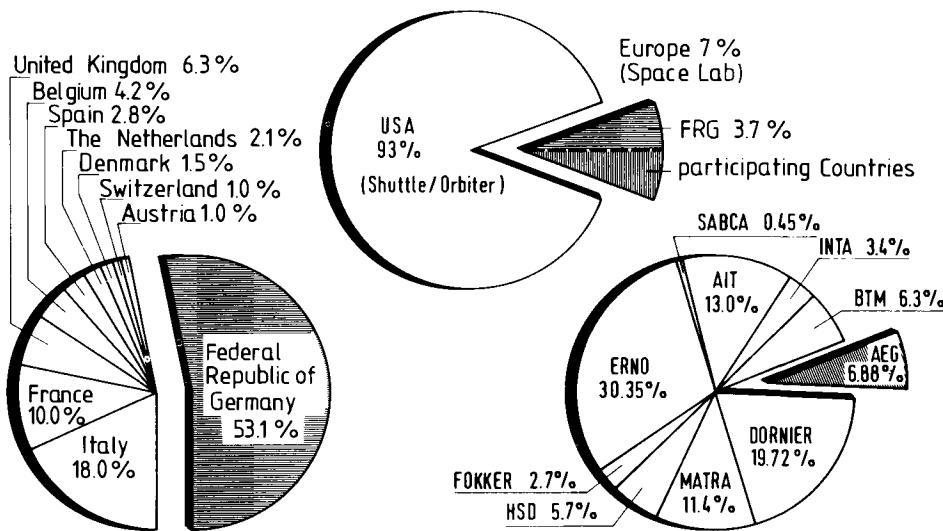


Fig. 1 Cost sharing for the SHUTTLE/SPACELAB program

from the figure, Germany contributes more than 50 % of the total demand. Therefore, a German firm, VFW-Fokker/ERNO, became the prime contractor for this important program. As a co-contractor, AEG-Telefunken of Hamburg/Wedel has been entrusted with the design, manufacture, and test of the Electrical Power and Distribution Subsystem (EPDS).

2. MISSION REQUIREMENTS FOR SPACELAB

SPACELAB is to be designed such that it meets a multitude of requirements which are primarily determined by the experiment needs within the general context of weight, power consumption and heat dissipation. The limits are given by the SHUTTLE Orbiter provisions for SPACELAB. They amount to

- 14.5 tons for SPACELAB including its experiments
- 7 kW average, 12 kW peak electrical power
- 930 kWh max. electrical energy
- 8.5 kW heat rejection capability

Within these limits, several configurations of SPACELAB can be composed from individual modules: Pressurized Module only of variable length, a combination of pressurized module and vacuum-exposed experiments on pallets or pallets only, see fig. 2. This modular concept for SPACELAB has been identified as the most cost -effective approach to accomodate the entire range of experiment categories listed in figure 2.

For all installations, a further major design requirement is a turn-around time of some 13 days between two SPACELAB flights. This means a short Mean Time to Repair (MTTR) of only 1 hour in the field of Maintanability. The designer of SPACELAB equipment, therefore, must design for rapid access to and replacement of defective installed units or harnesses.

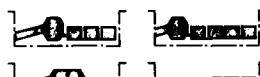
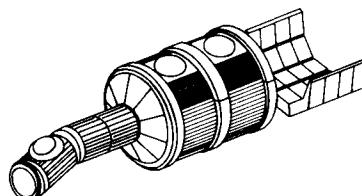
Finally, to remain attractive for experimenters, SPACELAB is to be designed for maximum SPACELAB-payload (experiment) accomodation. The design goal for this area is about 9 tons for experiment masses and of 3 to 4 kW of electrical power capability. A mission success of 0.95 per mission, the nominal duration of which is 7 days, extendable to 30 days maximum, is required.

To accomplish the different flight configurations of SPACELAB, the pressurized module for the scientists as well as the pallet is modularized. The pressurized module consists of two Segments, each 2.7 m long. The Core Segment contains the installed subsystems:

SPACELAB Requirements
 (336 planned NASA SPACELAB Missions)
 (1980 - 1991)

Earth Observations	10.7 %
Material Sciences	12.8 %
Technology	13.7 %
Astronomy	18.2 %
Biology	8.3 %
Navigation/Communication	3.3 %
Physics	<u>33.0 %</u> 100.0 %

SPACELAB Configurations



Payload 5.5-91 tons
 Length(max) 15.00m
 Diameter 4.06m

Fig. 2 Missions and configurations of SPACELAB

- Life Support
- Thermal Control
- Data Management
- Electrical Power Conditioning, etc.

which occupy about 1 m on both sides inside the Core Module. The rest of 1.7 m is dedicated to experiments. The second segment is fully dedicated to experiments.

Configurations consisting of the Core Segment only or of both Core Segment plus Experiment Segment are foreseen, depending on the volume of Experiments to be flown under normal atmospheric conditions. In case that experiments are to be vacuum-exposed, these experiments must be mounted on the pallets: Telescopes for astronomy, antennas for communication etc. For this category, up to 5 pallets, each 3 m long, may be installed. In case that 5 pallets are needed, no pressurized module can be used since the length of the cargo bay of the SHUTTLE Orbiter is limited to some 18 m. However, a combination of a pressurized module with up to 3 pallets will be possible.

The varying geometry of SPACELAB caused by this modular concept has a most significant impact to the harnesses: Each segment or pallet has to be equipped with a permanently installed power and signal harness. A flexible concept for these permanently installed harnesses is to be developed such that a highest possible portion of the cabling can already be interconnected by these harnesses and that completion of overall cabling requires only minor extra wiring.

3. ELECTRICAL POWER AND DISTRIBUTION SUBSYSTEM (EPDS)

3.1 Subsystem Design Requirements

EPDS has to condition, control, distribute and protect the raw fuel-cell power of the SHUTTLE to all subsystems and payload equipment of SPACELAB. Therefore the special features of EPDS are:

- Use of unregulated power delivery generated by orbiter fuel-cells, which are limited to 7 kW average and 12 kW peak.
- Conditioning of unregulated DC power into AC and regulated DC power by means of
DC/AC inverters for 115 V/ 220 V, 400 Hz
DC/DC converters for 28 V \pm 2 %
- Control and monitoring of power conditioning functions
- Emergency power handling independent of primary power
- Protection of system against circuit failures
- Adaptation of power distribution concept to modular system concept
- Separation of subsystems and experiment power buses
- Regulated AC and unregulated DC distribution
- Separation between power and signal harness
- Direct control of subsystem and experiment subgroup loads with respect to failure propagation.

The amount of power and energy to be used by experiments depends on the flight configuration. The design goal is 4 to 6 kW max. power to experiments, whilst 600 - 730 kWh energy will be available to experiments during 7 days mission. The experiments receive their power from experiment distribution boxes on the pallet or from experiment power switching panels in the module. Satisfaction of additional power demands is covered by a peak storage battery system. A summary of the specific power requirements for SPACELAB is shown in figure 3.

The main design criteria for SPACELAB-EPDS are:

- Design-to-safety overrides design-to-cost
- Max. power and energy to the experiments
- Meeting experiment needs for power type and quality
- Use of known techniques shall assure low risk and low cost
- Provision of flexibility in accommodation of different experiment payload changing from mission to mission
- Providing compatible interfaces with the orbiter electrical system and with the payloads.

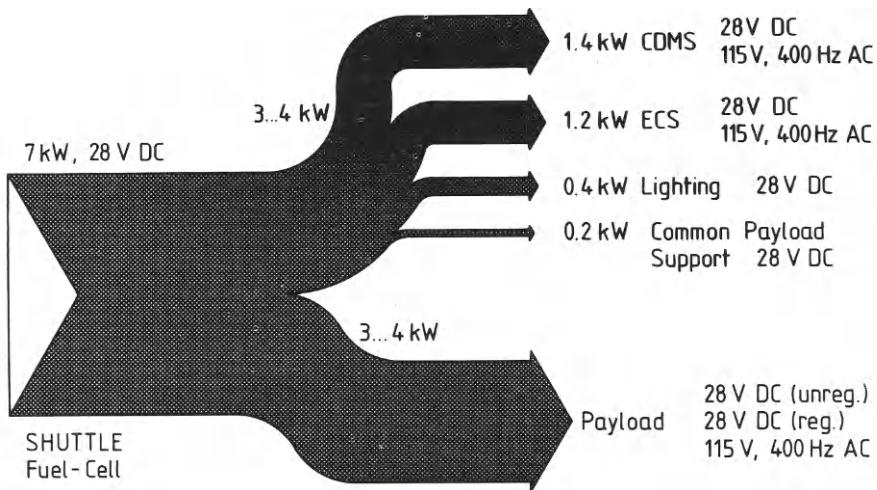


Fig. 3 SPACELAB power requirements

- Meeting low development, production, operation and maintenance costs
- Design based on existing and available European technologies as far as possible
- Modular design of power modules, distribution boxes and harnesses as far as applicable
- Low weight as possible
- Power distribution equipment to be located partly under the floor, and partly in Standard 19 inch racks
- Equipment with use on pallets shall be designed for vacuum operation
- Equipment will be cold-plate cooled and aircooled , depending on location of installation.

Finally, the strong interface with the orbiter imposes for electromagnetic compatibility constraints. Therefore, special measure are necessary to isolate the primary lines from the structure for all onboard equipment.

3.2 EPDS- Block Diagram

Three main power conditioning and distribution concepts have been considered more thoroughly to meet the requirements:

- AC-Distribution with decentralized transformer-rectifier-units
- DC-Distribution with decentralized converters
- Hybrid AC- und DC-distribution.

The trade-offs for the three concepts considered general subsystem aspects, e.g. weight, reliability, EMI, flexibility and losses. The SPACELAB related System criteria, overall cost aspects, maintainability, safety, influence also the choice of the EPDS-concept.

The greatest advantages offered by the hybrid DC-AC-Distribution are the low overall system costs and the simplicity of the electrical interfaces to the loads. For these reasons this concept has been selected and designed in more detail. To meet the payload requirements, a design has been chosen, which allows the experiment to use unregulated DC directly from the orbiter fuel-cell or SPACELAB provided AC and regulated DC. The power types and qualities 28 V DC, 115/200 V AC, 3 Phases - represent reasonable compromises between possible power conditioning alternatives. Of the different AC powers, 400 Hz, 115 V is chosen on the basis of application of off-the shelf avionic equipment. The distribution is based on 28 V DC or 115 V AC, the choice between these two is governed by the requirement of minimum conversion losses. If a load can use 28 V DC, one should not convert it to 115 V AC, because the lower harness weight does not compensate for the higher inverter weight, and the higher energy losses.

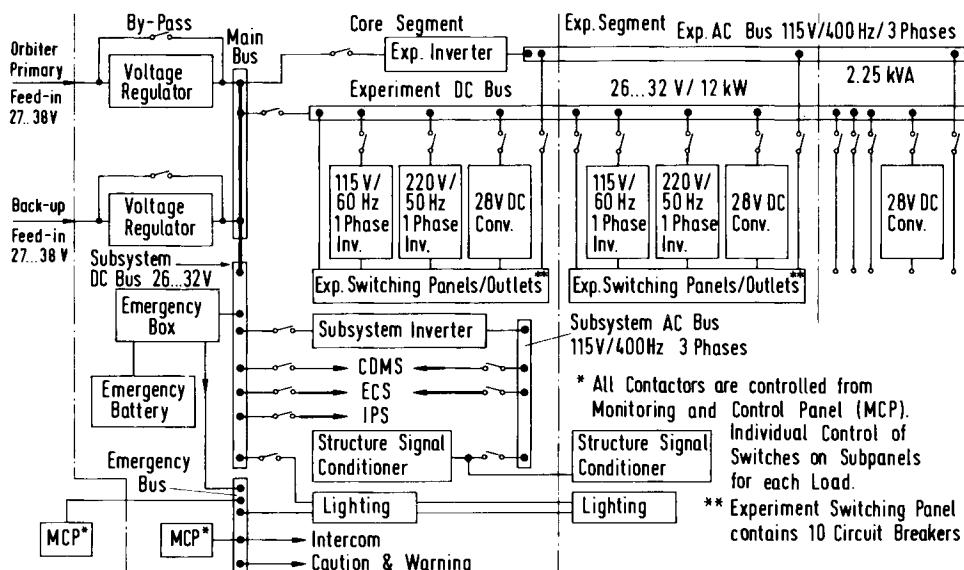


Fig. 4 SPACELAB EPDS-blockdiagram

The wiring harnesses and interface units are arranged in such a way that rapid disconnection for short ground turn-around is possible. The Subsystem wiring is permanently installed and has not to be changed significantly between missions. On the other hand mission-dependent harnesses can be installed independently on the structure floor/rack assembly when separated from the module pressure shell.

The conversion losses for subsystems supply of the chosen DC-AC distribution are shifted mainly to subsystems. Losses within EPDS are about 400 W, if payloads are supplied with unregulated DC. Losses of 750 W will occur if payloads are supplied with regulated DC. The highest losses of 920 W are estimated, if payload are supplied with AC.

Due to EMI-requirements the grounding concept has to be observed by all subsystems and payloads. The reliability for EPDS itself is rather high, but redundancy for Subsystem-Inverter is required. Concerning costs this is the best concept with respect to overall system costs. EPDS costs itself are also relatively low. Costs for generation of special voltages are shifted to subsystem and payloads, as mainly CAM-equipment can be used operating with either unregulated 28 V DC or 115 V/400 Hz. This concept is rather insensitive to changes within other subsystems and interfaces are relatively simple.

3.3 Power Sources

Electrical power for the orbiter and SPACELAB will be generated by hydrogen/oxygen fuel-cells. Three identical fuel-cells are foreseen for the orbiter, none is foreseen for SL itself. Any one of the 3 fuel-cells in the orbiter is able to guarantee safe return of the orbiter as far as electrical power is concerned. Any two fuel-cells can cover all orbiter power demands; under normal operating conditions, the third cell will be dedicated to SPACELAB.

The power given by SHUTTLE fuel-cell is 7 kW average up to 12 kW peak for 15 minutes. The fuel-cell itself is capable of delivery 12 kW continuous, but the limit is set by the heat rejection capability of the orbiter. At the interface with SPACELAB the voltage may vary between 27 and 34 V, see figure 5. In the basic design, fuel for about 950 kWh is available; additional energy can be accommodated by extra fuel kits, SPACELAB chargeable, however.

The peak power battery for peak power experiment loads is a facility offered to experiments and flown only on request. It is intended for experiments with high power demands - 10 to 20 kW - for short periods. Since these batteries are flown on a mission by mission basis, AgZn as well as NiCd batteries, dependent on mission constraints, may be used. A modular approach, fea-

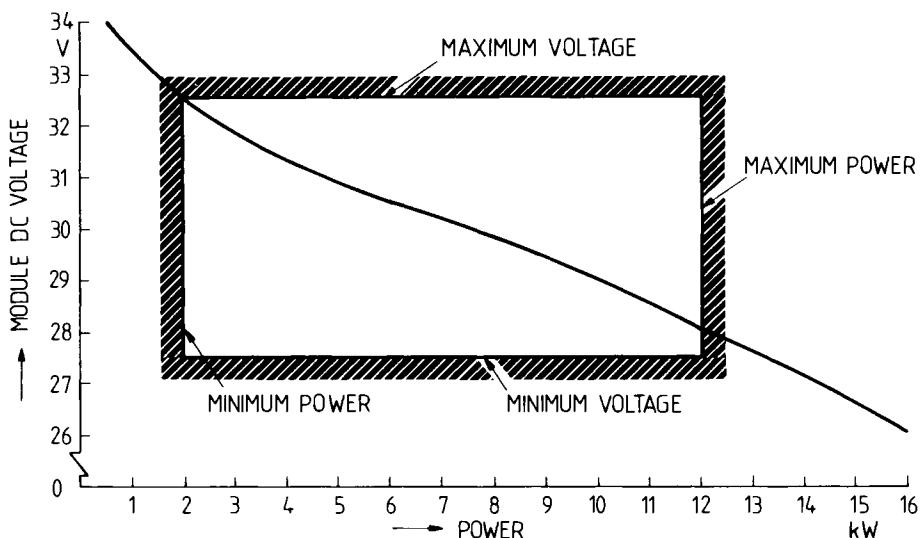


Fig. 5 Shuttle fuel-cell module power curve

ring up to 4 battery modules connected in parallel, is being studied. The peak power system shall be able to provide an output current of 500 Amps. max. With the peak power system, power demands exceeding the 7 kW level by 5 kW for more than 15 min. can be satisfied. Recharging of the batteries is accomplished by battery module-dedicated charge regulators of the buck/boost type with a charge rate varying between C/7 and C/15 depending on overall power profile possibilities.

3.4 Power Conditioning Equipment

Based on the experience of previous space projects on the one hand and on research and development contracts with the German space research authority, the Gesellschaft für Weltraumforschung on the other hand, a spectrum of space-configured hardware for modularized kW-power conditioning systems as used in SPACELAB has already been developed. Therefore, in some cases only relatively small adaptation work is required to manufacture the related SPACELAB hardware.

Inverters

Two identical inverters are provided for the baseline concept, one feeding the experiment AC-bus, the other the subsystem AC-bus. Each inverter of a mo-

dular type has a maximum power capability of 2.25 kVA, an output voltage of 115/200 V, 3 phases, with starpoint lead at 400 Hz.

Each inverter consists of two DC/DC-converters providing a ± 190 V DC-output. Three inverter bridges, controlled in a 120° phase-staggered mode, are supplied by these voltages. The AC-voltage obtained this way is then filtered to provide the specified sinusoidal output voltage. Regulation is achieved within the DC-conversion parts, where a DC-voltage equivalent to the peak AC-voltage is produced using pulsedwidth modulation techniques for regulation.

Electrical performance data for inverters:

operational input voltage	:	24 - 32 V
abnormal input voltage range	:	20 - 36 V
output voltage	:	115/200 V
output configuration	:	3 phases with neutral lead
output voltage tolerance	:	$\pm 5\%$
output frequency	:	400 Hz
output frequency tolerance	:	$\pm 1\%$
total distortion	:	2 %
nominal output power	:	2.25 kVA
efficiency at nom. load	:	80 %
tolerable power factor	:	0.85 lag to 0.85 lead.

Input-output isolation of the inverter is provided by using transformer signal coupling at the bases of the power switching transistors, within both DC/DC-converter and inverter circuits, and transformer power coupling within the DC/DC-converter. The regulation logic is separately supplied by an auxiliary power supply with floating outputs. High frequency switching techniques are employed in order to considerably reduce dimensions and weight of transformers and filters within both the converter and the inverter.

Converters for experiments

Grounding concept design principles and requirements for regulated DC justified the use of converters. The converters provide regulated 28 V floating output. They are designed for delivery of 500 W power which is conditioned in two parallel phase-staggered power modules. The DC/DC-converter concept is based on the principle of the transformer-buck/boost regulator. The regulator can operate at large input voltage variations with good efficiency, has input-output isolation, and excellent overload behaviour while the circuit simplicity helps to keep a high reliability figure. These power modules operate at a high switching frequency. This minimizes dimensions and weights considerably.

Two transformer-coupled buck/boost regulators are operating in parallel on a common output filter. The output filter operates roughly with twice the carrier

frequency resulting in a reduction of filter size. The common input filter is designed for filtering the two phase-staggered sawtooth currents. This technique helps effectively to overcome the disadvantages of a single regulator of this type with regard to high voltage and current ripples respectively. The power modules comprise energy storage transformers, power switching transistors, driving circuits, input and output filtering, and output voltage sensing. AEG has already designed such converters.

The DC/DC-converters are protected against excessive overload that could affect the output circuit. A current sensor is provided sensing the output current. At a defined current level the overload protection sends a signal to the control logic that reduces the pulse widths by overriding the error signal. Thus, the output current is limited to the defined maximum value.

Electrical Performance Data of Converters:

Normal input voltage range	26.5 - 32 V
Abn. input voltage range	20 - 40 V
Output voltage	28 V ± 2 %
Maximum output power	500 W
Efficiency	appr. 85 %

The only EPDS equipment mounted on cold plates are the inverters and the converters because of their relative high heat dissipation. Converters in the segments could be forced-air cooled. But, because of the sophisticated thermal lay-out for two types of cooling, inverters on pallets are located on cold plates only.

3.5 Power distribution, monitoring and protection

The power distribution system features DC as well as 400 Hz, 3 phase AC-distribution for "basic SPACELAB" and for SPACELAB's payload. The technological concept is putting emphasis on minimum weight penalty. An integral part of the distribution system are the provisions for monitoring and protection of the power system.

The power generated by the fuel-cells of the SHUTTLE is fed via a connector plate comprising primary and back-up feeds into the main power distribution box, which is located in the core module, see figure 6. From this mainbus the experiment DC-bus and subsystem (SS) DC-bus are derived. They are switchable by remote-controlled circuit breakage. The mainbus supplies also the experiment-dedicated inverter and the subsystem inverter. Both inverters are switched by remote-controlled circuit breakers located in the main power distribution box.

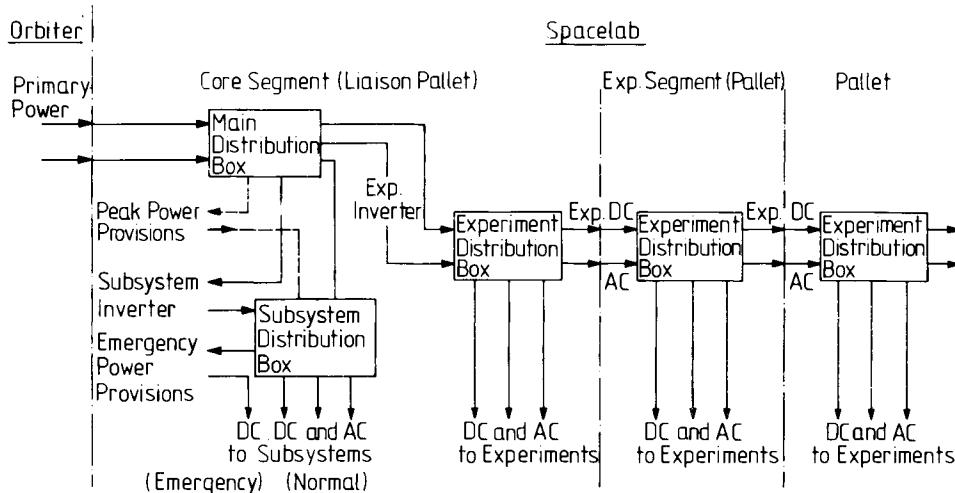


Fig. 6 SPACELAB power distribution scheme

The SS-DC-bus and the SS-inverter output are routed into the SS-Power Distribution Box. From this box all subsystem DC-busses are supplied. Subsystem units are controlled from dedicated subsystem switching panels. Similarly, the SS-AC-bus is routed into the SS-Power Distribution Box in the same way as the DC-bus.

The Experiment DC-bus is routed into the Experiment-Power Distribution boxes provided on each module or on each pallet. The busses are distributed via remote-controlled circuit breakers (RCCB) to Experiment Switching Panels installed within each rack section located in the secondary structure. From the switching panels the different experiment loads are switched manually by circuit breakers. Manually switchable power outlets are provided from these panels, too. The RCCB are controlled via an OR-gate allowing activities from the Monitoring & Control Panel for EPDS (MCP) or from the discrete output of the Remote Acquisition Units of the Data Management Subsystem. The latter mode being used on ground for speeding up of switching operations. At the interface between main-DC-bus and experiment DC-bus the RCCB, located in the Power Control Box, allows to switch-off all experiments manually from MCP and automatically by an overcurrent sensor.

In order to prevent failure propagation - both overload and short circuit - from a defective experiment to other loads, load-dedicated circuit breakers are foreseen within the Experiment Switching Panels. Groups of experiments as well as subsystem loads are protected by the RCCBs within the Subsystem or the Experiment Distribution boxes, respectively. These RCCBs are acting as normal switches and overload protection. Short circuit protection within

EPDS is made both by careful selection of response time and current switching capability of RCCBs.

All relays and power contractors that are related to power conditioning are remote-controlled by the MCP. Also, the AC-and DC-power contractors within the experiment distribution boxes are controlled from the MCP. The most important analog parameters of the EPDS are indicated by panel meters, allowing status monitoring of the EPDS independently of the CDMS data processing. The supply of the MCP by the emergency bus shall guarantee the switching and monitoring capability during critical situations in order to retain normal operation of the EPDS.

Subsystem wiring is permanently installed and does not have to be significantly changed between missions. Convenient interface units are located throughout SPACELAB to enable interconnection of mission-depending wiring. Wiring harnesses and interface units are arranged in such a way that rapid disconnection for short ground turn-around is possible.

Mission-dependent harnesses can be installed independently on the structure floor/rack assembly, when separated from the module pressure shell. Interconnection to the SPACELAB subsystems is easily accomplished after the floor/rack assembly is inserted into the module. Due to the selected mission-dependent distribution concept and to the high number of available outlets of Experiment Distribution Boxes and Switching Panels, a highly flexible distribution concept is obtained. Thus mission modifications are accommodated by an exchange of some cabling between Experiment Switching Panels and the experiments only.

The mission-independent harness is subdivided into:

- tunnel harness
- Core Segment harness
- Experiment Segment harness
- Igloo harness
- Standard Pallet harness
- Utility Bridge cabling

Separation of wire bundles with regard to power, signal and emergency cabling is foreseen.

The present EPDS-concept fully meets the EMC-requirements. If the EMC analysis shows too small margins, additional design changes are easy to introduce by synchronization of inverters and/or converters or by additional shielding, twisting and bonding.

3.6 Emergency Power System

In order to assure a safe and orderly evacuation of the crew in the event of an emergency, an emergency power-system has to be provided. The power capability of this emergency power-system is rated such that sufficient power for operation of essential loads is provided for the period of evacuation of the crew. In case of a failure on the primary power feeder, 2×2.5 kW as back-up power can be delivered via extra power feeders from the SHUTTLE. For safety reasons, a separate emergency power harness is installed. This assures a reliable supply of emergency loads in case of a power breakdown in the normal harness.

An Emergency Power Box, manufactured according to the highest-class-specifications for highest quality contains all monitoring and command electronics to route emergency power to the most essential loads and to shut down normal power in case of safety-critical abnormalities in the normal power system. Simultaneously, alert signals are generated and routed to the caution and warning System.

3.7 Lighting System

Finally, also the lighting system (Fig. 7) is part of the EPDS of SPACELAB. It consists of different types of lights for several lighting purposes, lighting for "basic SPACELAB" and for experiments. 27 to 32 V DC unregulated main bus voltage is applied for the complete lighting system.

For the overall illumination 6 lights at the ceiling of each segment are installed in two rows of 3 lights each. One of the 6 lamps in each segment shall serve as emergency illumination since it is supplied via the emergency bus. For illumination of the general-purpose work bench as well as the display and control panel in the core module two lights are installed on each place in the upper corner. The type of lamps used are 20 W fluorescent lamps as specified by US Department of Defense specifications MIL-L-116/17.

The top scientific airlock has to be illuminated by two lights which are installed inside the airlock cylinder near the two viewing lenses. The after scientific airlock is illuminated by three lights mounted on the walls inside the cylinder. One of them shall serve as emergency illumination by supply via the emergency bus. 25 W incandescent lamps are used as specified by US Department of Defense specifications MIL-L-6363.

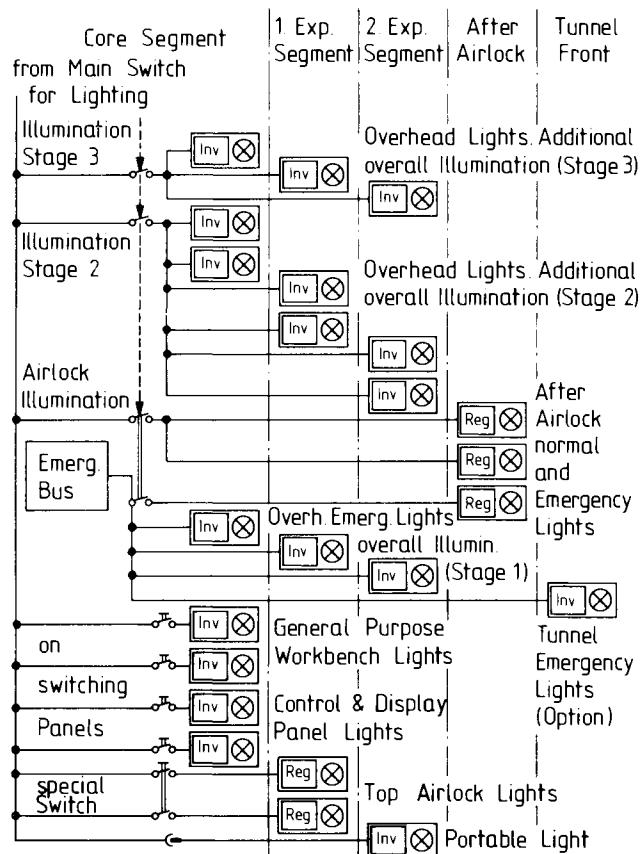


Fig. 7 SPACELAB Lighting system

Normal Power Consumption:

Tunnel illumination	24 W
Overall illumination	96 W/per module
Illumination of general-purpose work bench and display and control panel	96 W total
Top airlock illumination	60 W total
Aft airlock illumination	83 W total

Emergency Illumination Power Consumption:

Tunnel illumination	24 W
Overall illumination	24 W/per module
Aft airlock illumination	30

3.8 Signal Conditioning

For monitoring and operation purposes, pressure, temperatures and hatch positions of the scientific airlock of SPACELAB's Experiment Module must be measured and displayed. Furthermore, automatic heater control for maintaining temperatures inside the airlock is required.

To accomplish this, several electronic circuits must be developed and manufactured as an integral part of the Airlock Control Panel:

- Several physical stimuli are to be converted into electrical signals by signal transducers
- The output signals of the transducers must be amplified and level-detected in order to meet the requirements of the electrical interface, e.g. to the Data Management System.
- In order to avoid ground loops, all conditioned signals must be floating.

4. SUMMARY

The Electrical Power and Distribution Subsystem is being designed to provide the flexibility necessary to support a wide range of payloads for different SPACELAB configurations. In supporting the SPACELAB requirements, the EPDS will use state-of-the-art hardware as far as possible to minimize development costs. The power conditioning equipment is based to a large extent on the technology advancement program of the German Ministry of Research and Technology, which demonstrated the feasibility of high power systems for space crafts. In addition, other hardware items are based on previous experience of AEG-Telefunken in the area of power systems for aerospace applications.

LASERS FROM FISSION (GASEOUS CORE REACTORS AND NUCLEAR PUMPED LASERS FOR SPACE POWER GENERATION AND TRANSMISSION)

Richard T. Schneider*, Karlheinz Thom** and Herbert H. Helmick

*U.S. Office of Naval Research, London Branch Office**NASA Headquarters, Washington, DC, USA

***Los Alamos Scientific Laboratory, Los Alamos, NM, USA

ABSTRACT

Gaseous core reactors and nuclear pumped lasers have the potentials of space propulsion at exhaust jet velocities up to 50,000 m/sec, space power at 1 kg/kW specific mass, power transmission via laser beams over astronomical distances and beneficial applications on Earth.

The energization of lasers directly by nuclear reactions has recently been achieved. In experiments conducted jointly by the University of Florida and the Los Alamos Scientific Laboratory, New Mexico, a helium-xenon laser was directly pumped by fission fragments. The obtained laser wavelength was 3.5 μm . A group of researchers at the Sandia Corporation in Albuquerque, New Mexico, was successful in energizing a carbon monoxide laser by fission fragments at wavelengths in the 5- μm band. At the University of Illinois lasing was achieved at wavelengths of 8629 Å and 9393 Å in a neon-nitrogen mixture. A program of gaseous core reactor research is underway with experiments being conducted at the Los Alamos Scientific Laboratory in New Mexico, USA. The program utilizes a beryllium moderator-reflector, forming a cylindrical cavity of 1 m diameter and 1 m length. This system and associated control system hardware, uses components from the previous ROVER nuclear rocket program. Various configurations of canisters containing enriched gaseous uraniumhexafluoride fuel are inserted into the reactor cavity for research on neutronics and nuclear induced optical radiation.

Critical mass, control swing and the effects of poison were measured by simulating enriched uranium hexafluoride fuel with uranium foils, which were placed in homogenous and inhomogeneous distributions in the cavity. Critical mass was determined at about 6 kg 93% enriched 235 uranium. A uranium hexafluoride canister system was built for safe operation in the reactor cavity and for physics measurements and observations at nuclear criticality.

It is anticipated that this work will result in the demonstration of principles of a new type of nuclear power reactor, and of laser power output from such a reactor.

INTRODUCTION

In a previous paper (Ref. 1) entitled "Physics and Potentials of Fissioning Plasmas for Space Power and Propulsion," research on/and applications of

conceptual gaseous core nuclear reactors and nuclear energized lasers were described at the XXVth International Astronautical Congress, held in September 1974, in Amsterdam. Such research is mainly supported by the U.S. National Aeronautics and Space Administration. Originally motivated by advanced propulsion in space, the early work on gaseous core reactors was centered on problems germane to very high temperature operation (Ref. 2).

In the course of further research it became clear that in a fissioning uranium plasma, or, more generally, in fissioning gas, the distribution of energy levels of the plasma or gas could deviate from thermal equilibrium. The consequence of non-equilibrium distribution of energy levels is that radiative heat transfer in rocket application poses different problems than those germane to black body radiation, which was assumed to exist in gaseous core reactors for propulsion.

In the paper to the XXVth IAC, the possibility of nuclear pumping of lasers was described along with results of basic research that showed non-equilibrium distributions of excited states to exist in gases exposed to a fission fragment flux. However, at that time, the nuclear pumping of lasers had not yet been demonstrated by experiment. In the meantime this has occurred.

In this context nuclear pumping of lasers is understood to be the excitation of a laser by the kinetic energy of the fission fragments (nuclear energy) only. While the population of the upper laser level may actually be achieved by electron collisions, similar to E-beam pumped lasers, the energy of these electrons is obtained from the kinetic energy of the fission fragments via collisions during the slowing down process of these highly energetic, multiple charged ions.

It should be realized that for an efficient nuclear-energized laser all, or an appreciable part, of the fission energy should be released into the laser gas or gas mixture and not be converted to heat in a part of the reactor having solid fuel elements. Consequently, the laser gas, or gas mixture, must undergo nuclear fission itself. Thus, either a gaseous core reactor system or a hybrid system, in which at least a substantial fraction of the total reactor power is generated by a fissioning uranium gas or plasma, must be employed. Because of this necessary combination, nuclear pumped laser research is discussed here along with progress of research on gaseous core reactors.

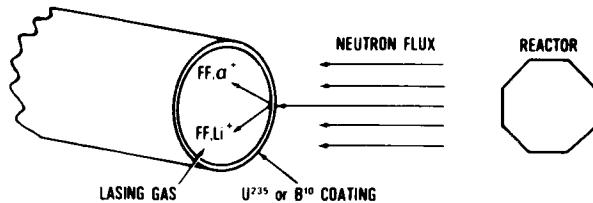
METHODS OF NUCLEAR PUMPED LASERS RESEARCH

Concepts of using nuclear energy directly to excite lasers have been described in the literature since about 1961 (Ref. 3). A major distinction exists between the excitation of atomic and molecular states and the excitation of nuclear states. Only the former is treated in this paper. Furthermore, the described research is restricted to gas lasers. Attempts to pump crystalline materials or liquids by nuclear radiation have been made by several researchers with varying success (Ref. 4).

The major difference of nuclear pumping of a gas laser from other methods of pumping is the very high energy of the energizing particles (fission fragments). Although gamma rays from nuclear reactions have the potential of inducing population inversions in gases, they are not considered here for nuclear pumping, since in a nuclear reactor most of the energy released by the fission reaction appears in the form of kinetic energy of fission fragments.

For research purposes, several methods are available for providing a flux of fission fragments or other high energy nuclear particles and for study of their interactions with a test gas as depicted in Fig. 1. All depend on an

I. $B^{10}(n,\alpha) Li$ OR $U^{235}(n,f)FF$



II. $He^3(n,p)H^3$

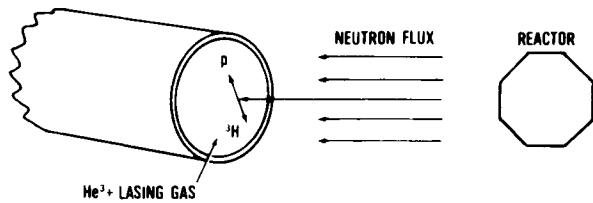
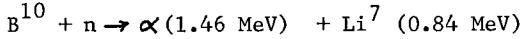
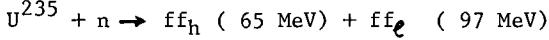


Fig. 1. Methods of nuclear pumped laser experiments.

external source of moderated neutrons. One group of experimental setups requires that the inside of the laser tube be coated with boron 10 or enriched uranium (U_3O_8). The following reactions produce energetic particles:



and

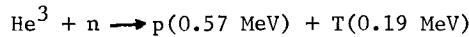


where n = neutron, α = alpha particle, ff_h = heavy fission fragment, and ff_e = light fission fragment.

Only the latter reaction provides a fission fragment flux, while the α -particle flux of the first reaction can be considered only for simulation of nuclear pumping. In both cases the particle flux has to emerge from the surface of the solid material which forms the coating in order to interact with the laser gas. Therefore, the coating needs to have the thickness of only one mean free path of the fission fragment (or α -particle) in the solid. Any increased thickness of the material will have the effect that additional fissions occurring in the coating will heat the solid material without contributing to the fission fragment flux emanating from the surface. This fact limits the achievable nuclear power input into the gas. This limitation disappears for a fissioning gas; as a matter of fact, the achievable power density should far exceed that density obtainable by electrical or chemical power input. The coated tube technique, however, has the advantage of

relative experimental ease and of allowing the testing of various undiluted gas mixtures.

A volume source of energizing particles is achieved when, instead of coating the walls, the laser tube is filled with He^3 and an additional (laser) gas mixture. The He^3 isotope has a large cross section for the capture of neutrons. The reaction



provides a particle flux of fast protons (p) and fast tritons (T). Such a volume source has the advantage that at constant neutron flux the nuclear power in the laser tube increases with He^3 partial pressure. Of course, this reaction, again, can only be considered as a simulation, albeit a very convenient one, of a fissioning gas.

All these methods have been tried. The coated tube technique has been successfully employed in experiments independently conducted by three different groups of researchers. The He^3 volume source method has so far yielded only partial success in experiments where nuclear excitation doubled the efficiency of an electrically pumped CO_2 laser (Ref. 5). It is now believed that at pressures of 10 atm and higher, excimer laser systems are promising candidates for the He^3 volume source excitation.

EXPERIMENTALLY REALIZED NUCLEAR PUMPED LASERS

General Remarks

Nuclear pumping of a laser is now understood to be the direct conversion of the kinetic energy of fission fragments into coherent electromagnetic radiation. It is, therefore, a process of energy conversion that does not involve thermodynamical efficiency limitations. In principle, nuclear pumped lasers could become a means of utilizing fission power at energy levels (and at equivalent temperatures) greatly exceeding those of conventional nuclear power generation. This may result in significantly improved efficiency and power density, apart from other benefits that can be derived from the monochromaticity and coherence of laser energy. In the recent past a major motivation for relevant research has been the needs of the space program. The energy sources for the nuclear pumping of lasers are conceived to be gaseous core reactors, similar to those which for many years have been investigated under NASA programs. The possibility of non-equilibrium radiation from such reactors with its concomitant fission induced population inversion, is an area of major concern in this research. This has been intensively investigated primarily by use of the previously described coated-tube technique in combination with research reactors as neutron sources. While the main objective was to demonstrate non-equilibrium distributions of ionized and excited states in fission-fragment energized gases and plasmas, nuclear pumping lasers was also attempted. The experimental work was supplemented by theoretical research, which, however, is still greatly hampered by the lack of data, such as cross sections and transition probabilities. In this earlier work, the existence of non-equilibrium radiation from fission fragment excited gases was clearly demonstrated by spectroscopically measured deviations of energy levels from Boltzmann distributions (Schneider, 1973).

For initial demonstration of nuclear pumping of lasers, owing to the incidental limitations of the coated tube technique, extremely high neutron flux densities were needed in order to deliver sufficient fission power for the laser gas to surpass threshold conditions.

A Nuclear Pumped He-Xe Laser

The largest neutron flux density from an experimental reactor can be obtained from a fast burst reactor, such as was developed for example at the Los Alamos Scientific Laboratory (Ref. 6). Such a reactor consists of a cylinder manufactured of highly enriched uranium 235 metal, control and safety rods totaling a weight of 65 kg. It is unshielded and unmoderated, so to speak bare, and because of this fact is named GODIVA. By pneumatic means the control rod is quickly inserted into a position within the cylinder to form a supercritical mass and a pulse in excess of 100,000 MW of power is obtained for 100 usec, during which time a neutron flux of a peak value of about $10^{18}/\text{cm}^2 \cdot \text{sec}$ is emitted.

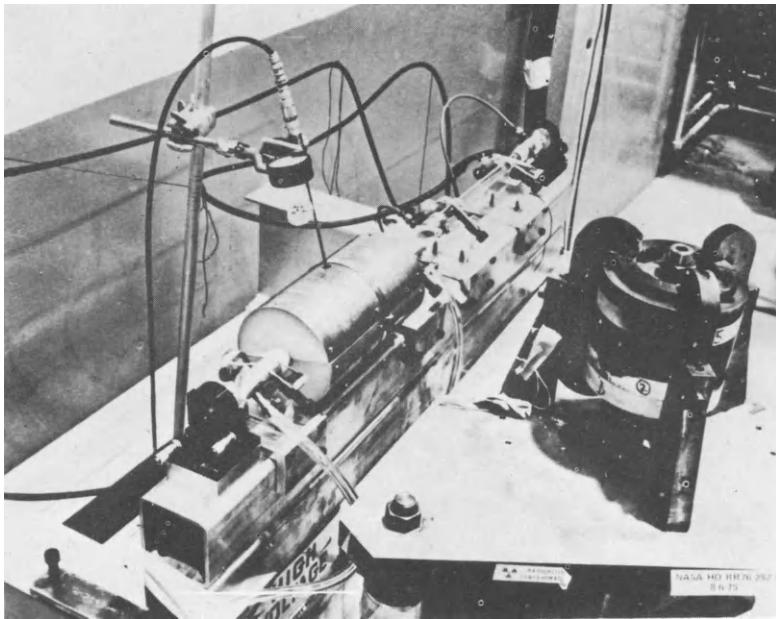


Fig. 2. He-Xe laser assembly with Godiva Reactor.

Figure 2 shows the GODIVA fast burst reactor in the right foreground, close to a laser assembly incorporating the coated tube technique.

The laser consisted of a 2.2-cm-diameter tube lined with a ^{235}U foil of 33 cm length. The tube, fitted with potassium chloride Brewster windows, was sited within a resonant optical cavity. The optical path led through a 0.2-cm-diameter output aperture in one of the mirrors to a heavily shielded, liquid-nitrogen-cooled, gold-doped germanium detector. The detector was

located in an adjacent laboratory room 18 m from the laser. The light path had four 90° bends for adequate shielding of the detector against streaming and neutron radiation from the reactor. The laser tube was surrounded by a 8.5-cm thick polyethelyne sleeve, serving as a neutron moderator. An ion chamber probe inserted in the cadmium covered polyethelene moderator monitored the fission rate near the uranium foil to determine the fission power input into the laser.

With this arrangement, nuclear pumping of a He-Xe laser was achieved (Ref. 7). By use of three Corning glass color-filters the observed wavelength was established to be around 3.5 μm . It is assumed that the strong Xe 5d $5/2^+$ transition was produced.

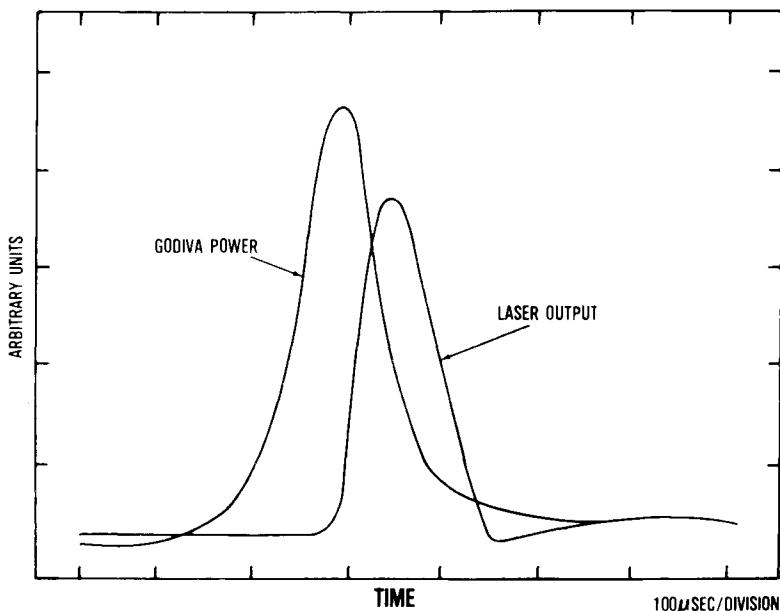


Fig. 3. Oscilloscope traces of neutron flux and laser output.

Figure 3 shows typical signals of the laser output and the fast neutron pulse versus time. The laser pulse was delayed by approximately 50 μsec in respect to the fast neutron pulse. This effect was caused by the neutron thermalization time which was required for moderation of the fast neutrons since the fission reactions in the U₃O₈ coating have to be induced by thermal neutrons. An analysis of the curves also exhibits a threshold for onset and termination of laser action, which, in a different way, was confirmed by the variation of the GODIVA reactor power. Best results were obtained from a Xe/He mixture ratio of 1:4 at 350-Torr pressure. The measured laser intensity was at least 0.1 W. Fission fragments from the coating of the tube were found to be the only source of laser pumping, and actual lasing was confirmed by various conventional methods.

The work was conducted jointly by the University of Florida and the Los Alamos Scientific Laboratory personnel.

A Nuclear Pumped CO-Laser

About the same time period as the He-Xe laser experiments were conducted at Los Alamos, a research team at the Sandia Corporation in Albuquerque, New Mexico, successfully energized a carbon monoxide laser solely by fission fragment irradiation (Ref. 8). In need of high neutron flux densities for achieving threshold, Sandia's fast burst reactor, SPR II, a device similar to the Los Alamos GODIVA reactor, was employed. Beyond considerations to achieve threshold, it was figured that the SPR II in a single burst produces enough neutrons external to its core to deposit 2×10^6 J in a surrounding medium containing ^{235}U nuclei. This medium could be part of a large laser system and, according to the mentioned citation, energy at the megajoule level could be dumped into a lasant. Such a large nuclear pumped laser would not require the complex electrical power supplies now being developed for large conventionally excited lasers. The research appeared, therefore, to have been motivated by the possibility of producing high power laser pulses. Overall-efficiency considerations similar to those which led to the NASA concept of a reactor-laser synthesis in form of a self-critical gaseous laser system were apparently not made, and the course of research seems to have been directed predominantly toward achieving the above mentioned goal.

The CO vibrational laser is well suited for nuclear excitation utilizing pulsed reactors and was therefore selected for tests with the SPR II. The characteristics of CO lasers as known from electrical pumping indicate operation at high pressure, excitation of a wide range of vibrational levels, high efficiency and a long radiative lifetime, which appeared to be advantageous, for an appropriate match with the neutron pulse from the reactor.

In the experimental arrangement the vertical axis of the laser was parallel to the axis of the SPR II reactor which, from a distance of 20 cm produced a 60- μsec pulse of fast neutrons at a flux density of $5 \times 10^{17}/\text{sec. cm}^2$ near the edge of the laser apparatus. The cylindrical laser chamber was divided into two parts by a CaF_2 Brewster window. The lower part was immersed in liquid nitrogen (77 K) and contained pure CO at about 100 Torr pressure. At the bottom there was a totally reflecting mirror. The lower chamber was surrounded by a 20 cm long cylinder coated with enriched uranium oxide. The scheme used also a polyethelene moderator sleeve. The upper part of the test chamber was evacuated to provide for thermal insulation to protect the output mirror. The optical detection apparatus was located 15 m away from the laser. It consisted of a calorimeter and a CaF_2 take-off plate which deflected a small part of the laser beam into a Ge:Au detector. The laser signal passed through either an uncoated Ge or a narrow-band interference filter centered at $5.24 \mu\text{m}$, which indicated lasing in the 5.1 to $5.6-\mu\text{m}$ vibrational bands of CO.

Oscilloscope traces of neutron and laser pulses showed similar patterns as those of the Los Alamos Experiments, indicating likewise a threshold behavior of the system. The laser pulse energy was measured at $1 - 3 \times 10^{-4}\text{J}$ which corresponds to a peak power of $2 - 6\text{ W}$. The neutron flux penetrating the uranium oxide coating was calculated using a three-dimensional Monte Carlo method. From this the average energy deposition in the laser gas was estimated to be of the order of 200 J/liter , which translates into a laser efficiency of $0.1 - 0.3\%$. Although the efficiency of electrically excited

CO lasers is much higher, this result must be considered quite encouraging, much higher efficiencies pending upon optimization of the system.

Ne-N₂ Laser Pumped by α -Particles

Nuclear pumping of a neon-nitrogen gas laser was recently achieved by a group of researchers at the University of Illinois (Ref. 9), who appear as a third party now joining a rapidly growing club of nuclear pumped laser operators. However, such a view would be misleading in respect to an evaluation of the history of the many contributions that eventually have led to the successful energization of lasers directly from nuclear reactions. In such a perspective, it is recognized that the University of Illinois group has been an outstanding contributor to the basics of nuclear pumped lasers for many years. Their achievement of lasing solely by nuclear particle irradiation of a gas seems to be delayed in respect to the progress made by others mainly because of the lack of a neutron source as powerful as the GODIVA or the SPR II fast burst reactors.

Instead, at the University of Illinois, a TRIGA research reactor was used, a heterogeneous moderated reactor that in a through-port close to the core can produce pulses of thermal neutrons at flux densities a few times $10^{15}/\text{cm}^2 \cdot \text{sec}$.

Over the years, various laser gas mixtures have been explored in this reactor, employing the coated tube technique. Gain measurements have been conducted, indicating nuclear induced population inversion. However, sufficient nuclear power deposition for passing threshold conditions has not occurred in the test gases until the recent pumping of the Ne-N₂ laser system.

However, during all these attempts, other significant research was conducted on the fundamental phenomena of high-energy particle-induced excitation and ionization in gases. Highlights of this work (Ref. 10) include the finding that secondary electrons, sometimes called δ -rays, are in most cases the more effective agents for ionization and excitation than, for example, fission fragments. The δ -rays are, of course, generated and energized by the nuclear reactants. Other important results were generated by computer analysis of gaseous systems irradiated by particles or fission fragments, in the search for electron distribution functions that are important for the prediction of ionization and excitation rates, and finally, for the prediction of population inversion. Characteristically, such distribution functions were Maxwellian peaking at low temperatures but with pronounced spikes on the higher energy tails, indicating a high potential for nuclear induced population inversion in gases.

This and other information (Ref. 11) was exchanged at NASA-sponsored meetings on fissioning plasma research prior to any achievement of nuclear pumping of lasers, and it is therefore fair to state that among the participants of these meetings there exists no claim to priority in the endeavor of nuclear pumping of lasers.

In the University of Illinois nuclear driven laser experiment the coated tube technique was employed using boron 10. This material produces less energy at neutron irradiation than enriched uranium. However, this nuclear reaction does not produce neutrons and is therefore more easily acceptable in reactor tests. A 2.5-cm-i.d. quartz laser tube was used, containing a 60-cm section of aluminum tubing coated inside with the boron-10. The cell was inserted into the TRIGA reactor. Pressure and concentration was varied by means of a

vacuum system attached to the portion of the laser extending out of the reactor. Lasing occurred at pressures above 75 Torr Ne and a very low partial pressure of N₂, at wavelengths of 8629 Å and 9393 Å.

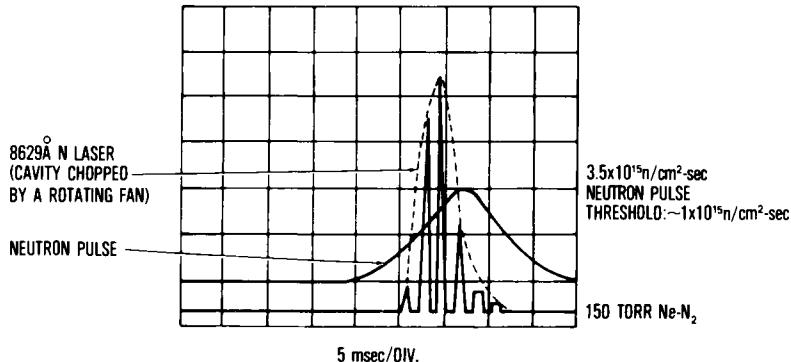


Fig. 4. Ne-N₂ nuclear pumped laser.

The traces in Fig. 4 (courtesy of Dr. G. Miley, Univ. of Illinois) show the laser output superimposed on the neutron pulse history. The laser cavity was chopped by means of rotating blades that periodically blocked the back mirror. From these traces, again, threshold behavior can be deduced, in addition to the proof of lasing.

Other significant results are best appreciated from a comparison of the nuclear pumped laser research results obtained so far by all groups involved in this work (Fig. 5, compiled at the University of Illinois). In the figure, it appears that progress is shown by looking down the columns, for example: decreasing wavelength, increasing pressure, decreasing required neutron flux densities, and length of laser pulse. Other characteristics do not fit into such a scheme, naturally so, because what has been described in the preceding pages are the first confirmed measurements of a novel physics effect: the nuclear pumping of lasers. Potential applications are great and the research devices described here will not resemble the final product at all. The planners of space technology may want to watch progress in this research with great interest; a request for engineering data would, however, be too early.

	WAVELENGTH	PRESSURE (TORR)	TH. THRESHOLD $\theta(n/cm^2\sec)$	ENERGY DEPOSITED (J.)	LENGTH OF LASER PULSE	PEAK POWER DEPOSITED (WATTS/.)	PEAK LASER POWER OUTPUT (WATTS)
FISSION FRAGMENT CO LASER SANDIA LABS	5.1-5.6 μ m	100	$\sim 5 \times 10^{16}$	200 150 μ sec	50 μ sec	1.3×10^6	2.6
FISSION FRAGMENT He-He ⁺ LASER LOS ALAMOS & UNIV. OF FLA	3.0-4.2 μ m	200	3×10^{16}	50 150 μ sec	135 μ sec	3.3×10^5	> 1
$B^{10}(n,\alpha)L^1$ Ne/N ₂ LASER UNIV. OF ILLINOIS	8629 \AA AN0 9393 \AA	75-400	1×10^{16}	313-1000 10 msec	6 msec	3×10^4 1×10^5	
He ³ (n, T) p He ³ -He ⁺ LASER IPL-LASL	EXPERIMENTS IN PROGRESS. USING VOLUME EXCITATION LASING EXPECTED AT 10 ATM PRESSURE. $\approx 6000 \AA$						

Fig. 5. Nuclear pumped lasers (May 1975).

UF₆-Admixtures to Laser Media

The energy released at nuclear fission is almost entirely invested in kinetic energy of the fission fragments. A small fraction appears as gamma rays, and very little energy is in the neutron flux. Therefore, concepts for nuclear pumped lasers that are based on the neutron fluxes from external driver reactors are intrinsically inefficient and should be considered as research devices only. For example, in the case of the He-He⁺ laser that was energized by the neutron flux from the GODIVA fast burst reactor, energy on the order of 50 J was emitted from the uranium foil inside the laser tube, while the energy of the reactor pulse was in the megajoule range. Of course, the fission energy generated in the reactor stays there. The laser system is coupled to the reactor power solely by the escaping neutron flux, and only a small fraction of GODIVA's neutrons actually reaches the laser tube. Consequently, as mentioned earlier in this paper, it is desirable that the lasant should fission itself or be mixed with a gaseous fissioning material inside a reactor-laser cavity.

The most likely fuel to be used in such a reactor-laser combination is UF₆. Therefore research concerned with the optical properties of fission-fragment excited UF₆ and UF₆ mixtures has been initiated. Although the structure of the UF₆ molecule is not yet completely understood, theory indicates that optical emission will be weak. So far, no success has been achieved in exciting the emission spectrum of UF₆ by either microwaves, electric discharges or fission fragments.

However, by using fission fragment excitation, some progress was made in obtaining emission spectra of other gases mixed into UF₆. In these

experiments, for convenience natural UF₆ was used in connection with the coated tube technique. This, of course, is only a simulation of a fissioning gas. The physical differences are mainly that the kinetic energy of the fission fragment flux coming from the solid surface has an energy distribution function different from a flux generated within the gas. And again, the energy input into the gas mixture is pegged low by the limitations peculiar to the coated tube technique.

Nevertheless, these preliminary results (Ref. 12) show that optical emission of gases mixed with UF₆ can indeed be obtained. Fig. 6 shows the relative

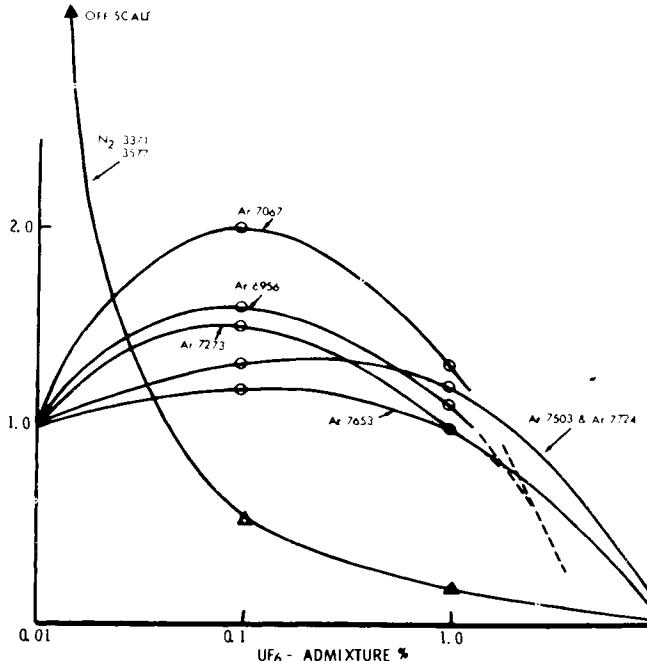


Fig. 6. Relative intensity of Ar and N₂ in UF₆ mixture.

intensity of certain N₂ and Ar lines. After adding a certain amount of UF₆ to these gases the UF₆ drastically quenches the N₂ lines, while the normal emission of fission-fragment excited Ar is actually increased by small amounts of UF₆. Some argon lines (Ar 7503 and Ar 7724) are quenched at the 10% level, others at a lower level. Although these experiments are of a too preliminary nature to draw far reaching conclusions, one fact can be considered to be established, namely, that it is indeed possible to obtain a light emitting fissioning gas.

RESEARCH ON UF₆-FUELED REACTORS

With the requirement to employ a fissioning gas, the nuclear pumped laser research merges with research on gaseous fueled reactors, as is currently being conducted under NASA programs. The final section of this paper describes the scope and progress of the gaseous fueled reactor work.

The distinguishing feature of gaseous-fuel reactors as far as neutron physics is concerned, is that the fuel is placed in a cavity and is surrounded by the moderator - the only moderator - in the system. Such systems, known as cavity reactors, have posed interesting problems in criticality analyses and experiments. Research in this area dates back to the early Fifties. More recent work on criticality (Ref. 13) used experimental facilities at the National Reactor Testing Station, Idaho. Important findings of this research include the requirements that fuel volumes be greater than 20% of the cavity volume, and that the cavity reactor system be designed for a minimum of nonessential reactivity penalties in order to maintain a low critical mass. A significant lesson from this research was the difficulty experienced in calculating the reactor physics parameters, such as multiplication factors and fuel worths, and therefore the value of benchmark experiments for verification of cavity reactor calculations.

A new program to demonstrate gaseous $^{235}\text{UF}_6$ fueled reactor operation and to probe non-equilibrium radiation from the fissioning gas or from a gaseous admixture was started in 1974 at the Los Alamos Scientific Laboratory, utilizing reactor components from the ROVER nuclear rocket program. Fig. 7 shows a projected program of stepwise increased performances of a reactor experiment that utilizes the same reflector-moderator cavity and control systems, but provides for various insertions into the cavity. In this fashion the anticipated progression of reactor operation and research will be accomplished most economically and in the shortest time.

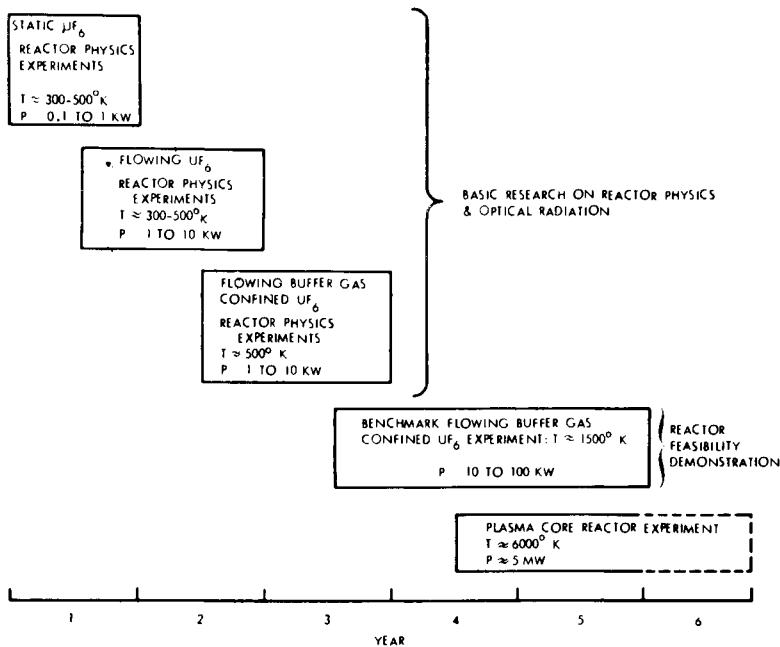


Fig. 7. Gaseous fuel nuclear reactor experiment series.

The first experiment is essential to establish a baseline for the specific configuration on material worths, control effectiveness, and neutron physics.

After this is accomplished, a logical progression involves stepwise increases of power until a level is obtained at which the feasibility of such a reactor concept can be judged with a reasonable degree of confidence. Obviously, methods for heat transport and, eventually, separation of the high temperature fissioning UF₆ gaseous fuel from the cavity surfaces have to be incorporated.

When changing to conditions where the nuclear fuel is in motion, where it is compressible, and where it has to share the cavity volume with confining buffer gases, reactor physics becomes quite complicated, particularly with respect to transient phenomena. A complete analysis of reactor dynamics for this situation has not yet been accomplished. It is expected that a delicate balance of positive and negative reactivity feedbacks must be maintained to keep criticality and to prevent power excursions. Obviously, the stability of the reactor depends heavily on the effectiveness of response of the control elements.

At the present the status of the program may be located in the middle of the first box in the chart.

In the experimental program at the Los Alamos Scientific Laboratory, four core configurations were investigated in the preparation of ²³⁵UF₆ loadings to be tested in the near future (Ref. 14). All of the ROVER beryllium reflector components that enclose a cylindrical cavity of roughly 1 m diameter and 1 m length have holes required for coolant flow, tie rod access, instrumentation lead channels, reflector density adjustment, and control drum containment. These four core configurations are described as follows:

Core Number 1: In this experiment, a homogeneous distribution of a static ²³⁵UF₆ loading was simulated with uranium foils, placed on a set of ten equally spaced aluminum discs, as seen in the photograph, Fig. 8. The critical mass was 19 kg of 93.2% enriched uranium 235. The reactivity swing for the 18 control drums was \$6.10. This amount of control span is approx-

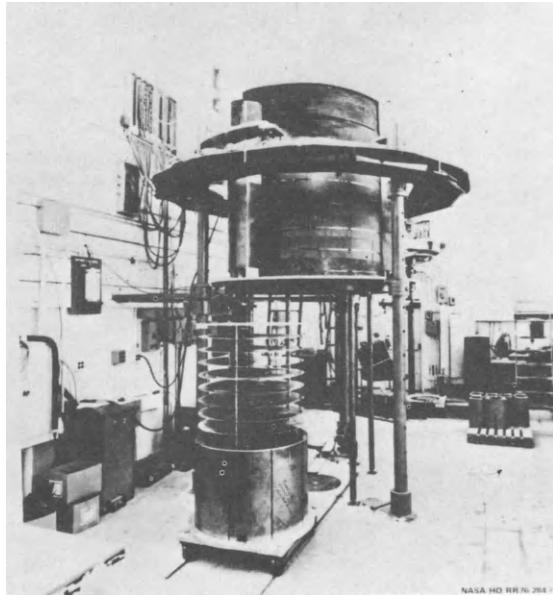


Fig. 8. UF₆ simulation core I.

POLAROID

NASA HQ REEN 2012.3

imately five times the quantity estimated to be necessary for control of reactor transients that could be reasonably expected for the first experiments with UF₆.

Core Number 2: This experiment was a mock-up of an early UF₆ experiment, in which a central zone of UF₆ would be surrounded by a solid fuel driver section. Such a situation was simulated by an 0.94-m diameter aluminum cylinder carrying a layer of uranium foil, centered in the cavity. A critical mass of about 19 Kg was indicated.

Core Number 3: The critical mass of 19 kg in both of the previous experiments was by a factor of 2 higher than calculated for uniformly distributed fuel in an idealized reflector. The cause for the large critical mass was a combination of the effects of the holes in the moderator and of contaminants in the beryllium. A solution to this neutron undermoderation problem was to add an annulus of beryllium to the cavity. This reduced the initial critical mass to 6.84 kg.

Core Number 4: This experiment provided for the transition to tests with ²³⁵UF₆. The beryllium flux trap from the previous experiment was employed, and the uranium foils were replaced by uranium-graphite ROVER fuel elements in an outer driver zone. In the inner experimental area consisted of an aluminum container similar to the ones to be used for future UF₆ experiments. During a final phase of simulation tests, it will house low density uranium-graphite fuel elements.

After the conclusion of these simulation tests, an aluminum canister filled with UF₆ will be located within the Be-reflector of the cavity reactor critical assembly. A UF₆ handling equipment is designed to provide a controlled supply of gaseous UF₆ to the canister. Provisions are made for remote operation of the entire flow system. Gaseous UF₆ is introduced into the canister by control of the UF₆ supply oven temperature. The UF₆ is removed from the cavity and returned to the supply cylinders by cryopumping.

The equipment is designed for operation at pressures up to 4 atm. Static pressure tests of the supply at 10 atm pressure have been completed.

The core canister is composed of two concentric aluminum canisters. During the criticality tests the inner canister will be charged with UF₆. The temperature in the canister will be maintained constant by controlling the flow of N₂ in the region between the inner and outer canister shells. The canister has provision for positioning fission wires and for attachment of a window for detection of possible fission-fragment-induced electromagnetic radiation emission produced during the criticality tests.

CONCLUSION

The feasibility of the nuclear pumping of lasers was demonstrated in three experiments conducted independently at three different laboratories. The efficiency in terms of laser output versus energy deposition into the laser gas is still low. However, reasonable judgment supports confidence that this efficiency can be improved significantly. Depending on the requirements of future applications, a synthesis of a nuclear pumped laser and a gaseous fueled nuclear reactor may be the ultimate goal of nuclear pumped laser research. NASA-sponsored work is progressing in this direction. Current work at the Los Alamos Scientific Laboratory is approaching a phase in which

an enriched gaseous 235 uranium hexafluoride fuel will fission in a cavity reactor, at conditions which allow for various kinds of neutronics measurements and measurement of optical radiation. The ultimate goal of a "self-critical" nuclear pumped laser, or a lasing gaseous core reactor has significant potential benefits. A part of the basic research needed is being conducted under NASA programs. Appreciable additional contributions must come from the scientific community.

REFERENCES

- (1) K. Thom, R.T. Schneider, F.C. Schwenk, Physics and Potentials of Fissioning Plasmas for Space Power and Propulsion, XXVth International Astronautics Congress, Amsterdam, The Netherlands (1974).
- (2) K. Thom, and R.T. Schneider, editors, Proceedings of a Symposium, Research on Uranium Plasmas and their Technological Applications, NASA SP-236 (1971), and R.G. Ragsdale, editor, proceedings 2nd Symposium on Uranium Plasmas: Research and Applications, AIAA (1971).
- (3) K. Thom, and R.T. Schneider, Nuclear Pumped Gas Lasers, AIAA J., 10 (1972).
- (4) R.T. Schneider, On the Feasibility of Nuclear Pumping of Lasers, Third Workshop on Laser Interaction and Related Plasma Phenomena, Rensselaer Polytechnic Institute, Troy, N.Y. (1973).
- (5) H. Rhodes, Direct Nuclear Excitation of a CO₂ Laser, PhD Dissertation, Univ. of Florida (1972).
- (6) T.F. Wimmet, R.H. White, and R.D. Wagner, Fast Burst Reactors, Proceedings of a National Topical Meeting, pp. 81-105, University of New Mexico, Albuquerque, NM (1969).
- (7) H.H. Helmick, J.L. Fuller, and R.T. Schneider, Direct Pumping of a Helium Xenon Laser, Appl. Phys. Lett., 26, 6, 327 (1975).
- (8) D.A. McArthur, and P.B. Tollefson, Observation of Laser Action in CO Gas Excited Only with a Nuclear Reactor, Appl. Phys. Lett. 26 (1975).
- (9) R.J. DeYoung, W.E. Wells, J.T. Verheyen, and G.H. Miley, A Direct Nuclear Pumped Ne-N₂ Laser, IEEE/OSA Conference on Laser Engineering and Applications, Washington, D. C. (1975).
- (10) W.E. Wells, Gaseous Quantum Electronics in Nuclear Engineering Education, NUCLEAR TECHNOLOGY (1975).
- (11) R.A. Walters, R. Paternoster, and R.T. Schneider, Recent Experimental Results Concerning Nuclear Pumped Gas Lasers, Proc. SPIE, Developments in Lasers Technology (1973).
- (12) G. Albrecht, Light Emission of a Ar-N₂-UF₆ Mixture, Special Project, Univ. of Florida, Dept. Nucl. Engrg. Sciences, Gainesville, FL (1975).
- (13) J.G. Kunze, G.D. Pincock, and R.E. Hyland, Cavity Reactor Critical Experiments, Nuclear Applications, 6, 2, 104-115 (1969).
- (14) G.A. Jarvis, W. Bernard, H.H. Helmick, and R. White, Beryllium Reflected Cavity Reactor for UF₆ Critical Experiments, AIAA/ASE 11th Propulsion Conference, Anaheim, Calif. (1975).

SATELLITE SOLAR POWER STATION CONSTRUCTION IN SPACE

Gerard K. O'Neill

Physics Dept., Box 708, Princeton University, Princeton, New Jersey 08540

Within this century it may be feasible to establish manufacturing facilities in space, possibly in the vicinity of one of the Lagrange libration points of the earth-moon system. A space manufacturing facility, community or colony would be a self-sustaining habitat for a large number of people (of the order of 10^4). Its energy needs would be met by solar power, used directly as sunlight for agriculture, as process heat for industry when concentrated by mirrors, or indirectly as electricity.

A space colony could be economically useful because:

1) Nearly all of the material for construction of the colony, and all the raw material for the products it would export, would be obtained from the lunar surface by a ground-based launching device. Economy in material supply would therefore exploit the factor of twenty in energy by which it is easier to bring materials to escape velocity from the moon than from the earth.

2) The "bootstrap principle": the first colony would be used not only to construct manufactured products, but to build more colonies, so that the productive output would be exponentially rather than linearly dependent on time.

The utility of a space manufacturing facility would therefore be maximum for those products whose end-use would be in free space or in high orbit rather than on the earth.

We have examined the production of satellite solar power stations at a space colony. Unselected lunar soils contain typically 40% oxygen, 20% to 30% metals, and 20% silicon by weight. They are therefore suitable as raw materials for satellite power stations either of the turbogenerator or photovoltaic types. Oxygen produced at the colonies as a waste-product in chemical processing would be available in abundance as reaction mass, for the transfer of a finished power station from the libration point to geosynchronous orbit. There the station could beam down microwave power to the earth's surface. In that approach, the relatively light and fragile components of a satellite power station would be produced in zero gravity, and would never have to stand the stress of launch from the earth. Correspondingly, power plant designs could be optimized for efficiency without launch-vehicle restrictions.

Cost-benefit analyses for the entire colony/power-station project have been made. Under a wide variety of assumptions as to the overall cost of the project, in all cases assuming discounted economics with a 10% discount rate, the benefit/cost ratio is much higher than unity. Because of the exponential nature of the bootstrap process, the amount of generator capacity produced within a few years is substantial *within* the first ten years after completion of the first colony, typically the total integral of energy supplied to the earth is greater than the total energy now stored in the form of oil in the Alaska North Slope).

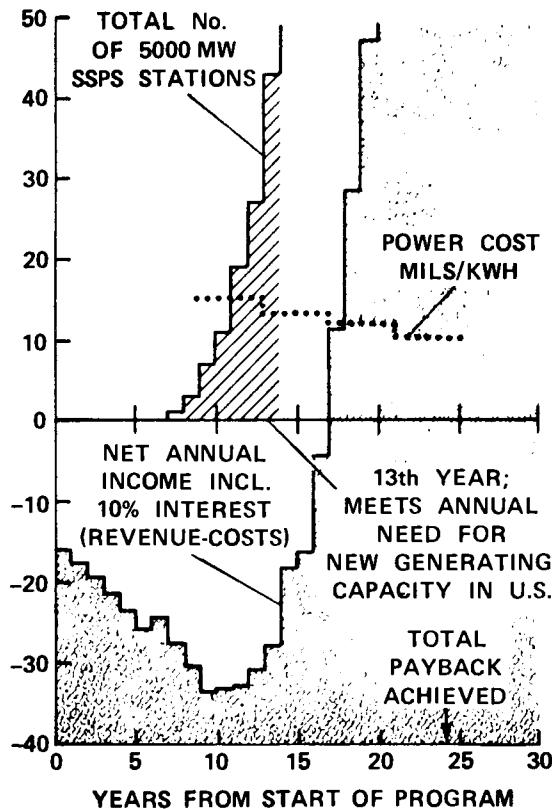


Figure 1 shows costs, revenues and generator capacity for a sequence in which the first colony is capitalized at 96 Billion Dollars before interest charges, and in which an additional 82 Billion Dollars is assumed to be spent in furnishing additional colonies with the elements C, N and H, rare on the moon. The capitalization is based on a launch cost of \$950/kg to the libration point, on a power plant mass of 10 tons/megawatt, and on human productivity figures for processing and construction (in zero or low gravity, in a shirt-sleeve environment) no higher than is now customary on the earth for heavy industry. The busbar price of energy on the earth is taken initially as 15 mils, competitive with nuclear power, and is dropped in steps to 10 mils, consistent with generator facilities in base load service, achieving strong market penetration. The benefit/cost ratio exceeds 3.

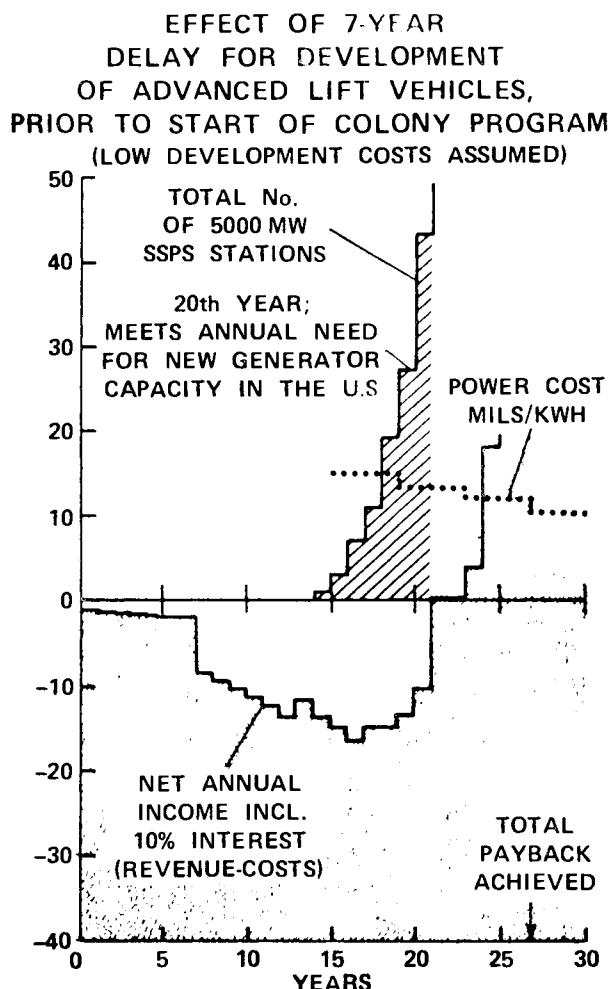


Figure 2 is a time-line in which, instead of building the first colony on the basis of shuttle-derived lift-vehicles, a delay is inserted for the development of advanced lift-vehicles with assumed low launch costs (\$75/kg to the libration point, similar to what has been assumed in Boeing studies of earth-launched satellite power stations). In that case the capitalization drops by a factor two, but all benefits are delayed for the seven-year development time of the new launch vehicles.

METHODES QUALITATIVES EN MECANIQUE CELESTE

Christian Marchal

Office National d'Etudes et de Recherches Aérospatiales, 92320 Chatillon (France)

RESUME

L'étude qualitative d'un problème vise seulement à établir rigoureusement quelques propriétés de ses solutions. Ces propriétés se rapportent en général à l'existence, l'unicité, la périodicité, la stabilité des solutions ou encore à leur structure ou à leur évolution finale.

Quatre exemples très simples de méthode qualitative sont successivement présentés, ils concernent les solutions périodiques, l'analyse du voisinage d'une collision triple, les évolutions finales et les solutions oscillatoires de Sitnikov.

SUMMARY

The only purpose of the qualitative analysis of a problem is to establish rigorously some properties of its solutions. These properties generally concern the existence, the uniqueness, the periodicity, the stability of solutions or their structure or their final evolution.

Four very simple examples of qualitative method are successively presented, they concern the periodic orbits, the analysis of the vicinity of a triple collision, the final evolutions and the Sitnikov oscillatory solutions.

1 - INTRODUCTION

Les systèmes d'équations différentielles non intégrables se rencontrent dans de très nombreux problèmes et leur étude se fait généralement par des méthodes quantitatives approchées (théorie des perturbations, intégrations numériques etc...). Cependant, hormis des cas particuliers comme ceux des solutions périodiques, la précision de ces méthodes se dégrade peu à peu et un moment vient où la solution obtenue n'a plus rien à voir avec la solution réelle.

Les méthodes qualitatives ont été développées pour remédier à ces inconvénients, elles conduisent à des résultats certes partiels mais aussi rigoureusement démontrés et généralement indéfiniment valables comme par exemple ceux liés aux intégrales premières, elles ont de plus l'avantage de donner des idées générales de l'ensemble des solutions ce qui est très difficile à obtenir autrement, elles couvrent des domaines très variés et sont dues principalement à Poincaré, Sundman, Chazy, Khilmy, Merman, Sitnikov, Alexeev,... Nous allons analyser les exemples les plus caractéristiques.

2 - LES SOLUTIONS PERIODIQUES

Un exemple simple de méthode qualitative est la recherche et l'étude des so-

lutions périodiques de la manière mise au point par Poincaré (Poincaré 1892, 1893, 1899).

Poincaré utilise les symétries naturelles du problème des n corps ponctuels et opère par continuité à partir du "cas planétaire" pour lequel l'une des n masses (soleil) est beaucoup plus lourde que toutes les autres (planètes) qui ont donc des orbites presque képlériennes par rapport à la masse principale.

La figure 1 représente un cas de "symétrie passé-futur" : à un instant donné, les n corps sont alignés et les vitesses sont normales à la droite d'alignement, il est alors aisé de vérifier que les distances mutuelles sont toutes les fonctions paires du temps par rapport à l'instant d'alignement. Si une telle symétrie se reproduit les distances mutuelles deviendront des fonctions paires du temps par rapport à deux instants différents t_1 et t_2 , elles seront donc nécessairement périodiques avec une période égale à $2(t_2 - t_1)$ et la solution étudiée sera elle-même périodique (avec une rotation d'angle constant à chaque période).

Essayons donc d'obtenir une solution comportant deux symétries successives dans un "cas planétaire".

Utilisons les axes du centre du Soleil et considérons un cas plan avec $n = 4$.

Les quatre corps sont à l'instant t_1 sur l'axe Ox_1 et l'on veut qu'ils soient à l'instant t_2 sur l'axe Ox_2 , les vitesses radiales étant nulles dans chaque cas.

Posons $\widehat{x_1 O x_2} = \alpha$ et faisons, par exemple, effectuer à la masse m_1 un angle α , à la masse m_2 un angle $\pi + \alpha$ et à la masse m_3 un angle $2\pi + \alpha$. On dispose des positions initiales et des vitesses circonférentielles initiales de m_1 , m_2 , et m_3 pour obtenir le résultat cherché ce qui est aisé si ces 3 masses sont infinitésimales (cas képlérien à orbites circulaires) et ce qui comporte nécessairement une solution voisine de cette première solution si $\sin \alpha \neq 0$ et si les masses m_1 , m_2 et m_3 sont suffisamment petites ainsi que l'on peut s'en rendre compte en étudiant les orbites képlériennes voisines et en recherchant la solution par approxima-

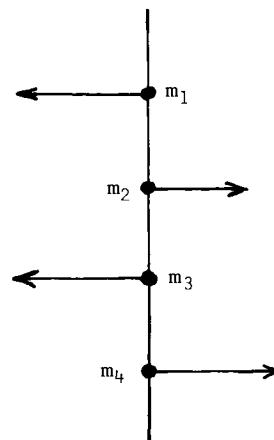


Fig. 1. Cas de symétrie passé-futur
(Case of past-future symmetry)

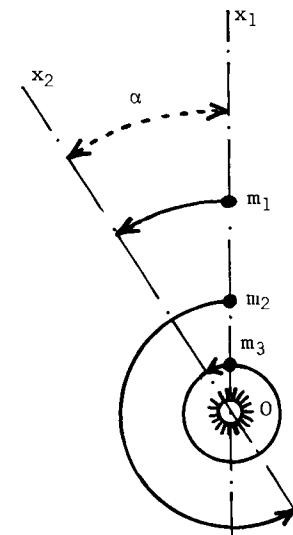


Fig. 2. Exemple de solution périodique
(Example of a periodical n-body solution)

tions successives.

On peut aisément étendre ce résultat à un nombre quelconque de planètes et à des mouvements de sens rétrogrades pourvu que les valeurs absolues des angles parcourus par chaque planète de t_1 à t_2 soient toutes différentes, on peut même ajouter des satellites à chaque planète pourvu que le mouvement reste du type "hiérarchie de mouvement quasi-képlériens" et que les symétries initiales et finales soient respectées.

Les solutions périodiques ainsi définies sont planes et ont des orbites quasi-circulaires, elles peuvent être étendues par continuité à des rapports de masse plus proches de 1 et bien d'autres solutions périodiques peuvent être obtenues par divers moyens surtout dans le cas du problème des 3 corps. Poincaré conjectura que l'ensemble des solutions périodiques est dense dans l'ensemble des solutions bornées (Pars, 1965) et il semble bien que ce soit vrai.

3 - VOISINAGE D'UNE COLLISION TRIPLE

Les solutions du problème des deux corps ponctuels sont bien connues et le voisinage d'une collision double peut être régularisé et ne présente pas de difficultés analytiques particulières, il n'en est pas de même du voisinage d'une collision triple : des solutions infinitement voisines de celle conduisant à la collision triple s'écartent l'une de l'autre d'une quantité finie voire même infinitiment grande après l'instant de la collision triple ! Ces collisions ou quasi-collisions représentent donc des événements majeurs dans l'évolution d'un système de n-corps et sont particulièrement difficiles à étudier numériquement, on peut s'en convaincre aisément de la façon suivante :

3.1 - La transformation par homothétie

Si $\vec{r}_i(t); i = \{1, 2, \dots, n\}$ est une solution du problème des n corps ponctuels, $\vec{r}_i(t) = k^2 \vec{r}_i(t/k^3); i = \{1, 2, \dots, n\}$ en est une autre ainsi qu'il est aisé de le vérifier (k étant un scalaire constant non nul).

Dans la nouvelle solution les rayons vecteurs sont donc multipliés par k^2 et les vecteurs vitesses par k.

3.2 - Application de la transformation par homothétie

Considérons 2 solutions particulières du problème des 3 corps (Fig. 3), tout d'abord la solution en trait plein : c'est une solution de Lagrange où les 3 corps sont constamment alignés et décrivent des orbites paraboliques (dans le cas de la figure 3 les masses m_1 et m_2 sont égales, m_3 reste donc à l'origine), enfin la solution en pointillé : elle est asymptotique à la solution parabolique de Lagrange pour le passé mais dans le futur les masses m_1 et m_3 forment une binaire elliptique tandis que m_2 s'échappe avec une vitesse hyperbolique finale \vec{V}_f .

Appliquons la transformation par homothétie à ces deux solutions avec un rapport k tendant vers zéro : la solution de Lagrange tend vers une solution à collision triple pour laquelle $\vec{r}_i = A_i |t|^2/3; i = \{1, 2, 3\}$ les 3 vecteurs A_i étant fixes, la seconde solution à la même limite pour $t < 0$ mais pour $t > 0$ la taille de la binaire m_1, m_3 est proportionnelle à k^2 et tend vers zéro tandis que la vitesse finale de m_2 est \vec{V}_f/k et tend vers l'infini... ce qui montre bien la complexité du voisinage d'une collision triple.

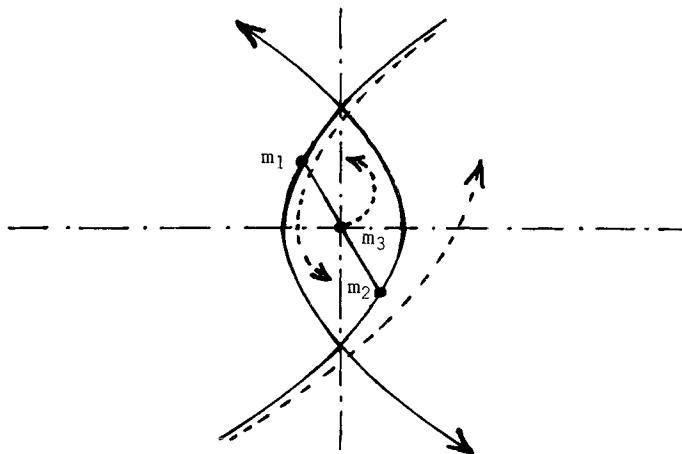


Fig. 3. Etude du voisinage d'une collision triple
(Study of the vicinity of a triple collision)

4 - CRITERES D'EVOLUTION FINALE

La notion de critère d'évolution finale est très générale et s'applique à la plupart des systèmes d'équations différentielles, elle consiste à partager l'espace des phases en deux ou plusieurs morceaux entre lesquels le mouvement ne peut se faire que dans un seul sens ; les solutions sont ainsi classées en : A) Solutions indéfiniment dans le premier morceau, B) Solutions passant (une seule fois) du premier morceau au second, C) Solutions indéfiniment dans le second morceau, etc... Ainsi, par exemple, si $G(\bar{X}, t)$ est une fonction continue de l'état \bar{X} et du temps t et si le long de toutes les solutions la condition $G_1 < G < G_2$ entraîne $dG/dt \geq 0$ on a alors nécessairement lorsque t est croissant (et aussi lorsque t est décroissant), soit $\limsup G \leq G_1$, soit $G_1 < \lim G < G_2$, soit $\liminf G \geq G_2$ et la valeur actuelle de G donne des indications sur les valeurs passées et futures de G .

Un exemple simple d'un tel critère dans le problème des n -corps est donné par l'évolution de la longueur λ (Marchal et Saari, 1976 ; Marchal, 1975), longueur définie comme suit :

- A) soient m_1, m_2, \dots, m_n les n masses étudiées classées par valeurs croissantes,
- B) soit $\varphi = \{i, j, k, \dots\}$ un sous-ensemble propre non vide de l'ensemble $\{1, 2, \dots, n\}$
- C) soit P l'ensemble des sous-ensembles φ ; P a donc $2^n - 2$ éléments,
- D) soit λ_φ la distance séparant le centre des masses m_i dont l'indice i appartient à φ du centre des autres masses.

La longueur λ (Fig. 4) est alors définie par :

$$\lambda = \sup_{\varphi \in P} \lambda_\varphi \quad (1)$$

λ correspond donc à une certaine partition Φ des n masses et à cause de sa définition, on obtient, en apposant M la masse totale et r_{ij} la distance des masses m_i et m_j :

$$\lambda \leq R < 2\lambda; \text{ avec } R = \sup r_{ij} \quad (2)$$

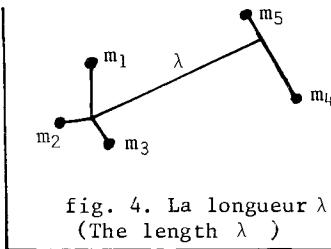


fig. 4. La longueur λ
(The length λ)

$$\left\{ i \in \Phi, j \notin \Phi \right\} \Rightarrow r_{ij} \geq \lambda (m_i + m_j)/M \geq \lambda (m_1 + m_2)/M \quad (3)$$

Les deux moitiés de la partition correspondant à λ et à Φ sont donc séparées par une distance supérieure ou égale à $\lambda (m_1 + m_2)/M$ et la force d'attraction qui les relie est donc bornée, on en déduit une limite sur l'accélération de λ (G étant la constante de la loi de Newton) :

$$\lambda'' = d^2\lambda/dt^2 \geq -K/\lambda^2; \text{ avec } K = GM^3/(m_1 + m_2)^2 \quad (4)$$

Remarquons que cette condition est conservée même quand il y a modification de la partition Φ correspondant à λ : λ est alors continue et λ' a une discontinuité positive ou nulle, on peut donc alors considérer que λ'' a un "dirac" positif ou nul ce qui s'accorde avec $\lambda'' \geq -K/\lambda^2$ (cf pour plus de détails les deux références citées ci-dessus).

Dans ces conditions, on peut utiliser pour fonction $G(\vec{x}, t)$ définie au début de cette section :

$$G = \lambda' |\lambda'| - 2K/\lambda; \text{ avec } G_1 = 0; G_2 = +\infty \quad (5)$$

Notons que G n'est pas une fonction continue à cause des discontinuités de λ' , mais quand G est positif ses discontinuités sont toujours dans le sens positif ce qui ne modifie pas les conclusions.

Ceci conduit à 5 évolutions finales possibles de la longueur λ et du système de n -corps ponctuels étudié.

A) La collision triple ou multiple à un instant fini

Contrairement aux collisions binaires ces collisions ne peuvent être régularisées en général et la solution ne peut être définie de manière unique après la collision. Ce premier type d'évolution finale correspond à un ensemble de mesure nulle de l'espace des phases (Saari, 1973).

B) L'expansion infinie en un temps fini

Cette évolution finale très particulière ne se présente que si $n \geq 4$ et exige de très fortes oscillations : pendant que les longueurs λ et R vont à l'infini il y a au moins deux distances mutuelles r_{ij} pour lesquelles $\limsup r_{ij} = +\infty$ et $\liminf r_{ij} = 0$. Un exemple de ce type d'évolution peut être trouvé dans (Mather et McGehee, 1974).

C) L'expansion super-hyperbolique

Ce cas très lié au précédent ne se rencontre aussi que pour $n \geq 4$ la longueur λ et le diamètre R du système y croissent plus vite que le temps t ($t \rightarrow \infty$ entraîne $G \rightarrow \infty$; $\lambda' \rightarrow \infty$, $\lambda'/t \rightarrow \infty$ et $R/t \rightarrow \infty$) et il y a au moins deux

distances mutuelles r_{ij} pour lesquelles : $\limsup_{t \rightarrow \infty} (r_{ij}/t) = +\infty$ et $\liminf_{t \rightarrow \infty} (r_{ij}/t) = 0$

D) L'expansion hyperbolique ($t \rightarrow \infty$ entraîne $G \rightarrow G_f$ fini et positif, $\lambda' \rightarrow G_f^{\frac{1}{2}}$
 $\lambda \hookrightarrow t G_f^{\frac{1}{2}}$).

Ce cas est très usuel : lorsque $t \rightarrow \infty$ le système se divise en sous-systèmes s'écartant linéairement les uns des autres et à l'intérieur desquels les distances mutuelles sont au plus d'ordre $t^2/3$.

Dans chaque sous-système l'énergie et le moment cinétique tendent vers une limite finie.

E) L'expansion parabolique ou sub-parabolique (pour tout t : $G \leq 0$ donc $\lambda' \leq (2K/\lambda)^{\frac{1}{2}}$, et $\lambda = 0$ ($t^2/3$)).

La longueur λ et toutes les distances mutuelles sont au plus d'ordre $t^2/3$ et, dans les axes du centre de masse, l'énergie est négative ou nulle.

Ce dernier type d'évolution finale présente des solutions d'une grande variété non encore complètement investiguée : solutions avec des évasions paraboliques, solutions bornées, solutions oscillatoires ($\limsup R = +\infty$; $\liminf R < +\infty$). etc...

L'utilisation des théories de la mesure permet de préciser la structure des solutions bornées et des solutions oscillatoires et de souligner leurs caractères de stabilité, en particulier elles ont "presque toutes" un type identique dans le passé et dans le futur.

5 - MOUVEMENTS OSCILLATOIRES DE SITNIKOV

Comme dernier exemple de méthode qualitative simple, présentons les mouvements oscillatoires de Sitnikov (Sitnikov, 1961), c'est-à-dire des mouvements non bornés mais revenant une infinité de fois à des dimensions bornées.

Sitnikov utilise deux masses égales m_1 et m_2 tournant sur deux ellipses képlériennes symétriques dans le plan Oxy (Fig. 5). La masse m_3 est infinitésimale et est lancée le long de l'axe Oz, elle y reste indéfiniment à cause de la symétrie du problème.

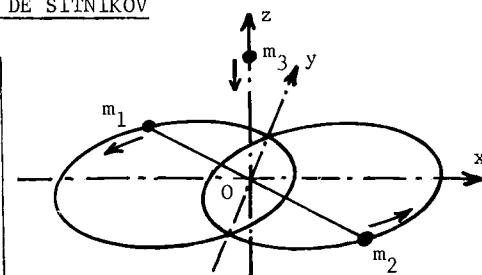


fig. 5. Mouvements oscillatoires de Sitnikov
(Sitnikov oscillatory motions)

Etudions donc le mouvement de m_3 le long de Oz.

Si m_3 passe en O à un moment où m_1 et m_2 sont soit au périhélie, soit à l'apogée, la situation est celle d'une symétrie passé-futur et le mouvement de m_3 est une fonction impaire du temps autour de l'instant de passage en O

Si m_3 passe en O pendant que m_1 et m_2 s'écartent, l'énergie de m_3 augmente au cours du passage et c'est l'inverse qui se produit si m_1 et m_2 se rapprochent.

Launchons donc m_3 depuis O à un moment déterminé (par exemple lorsque m_1 et m_2

sont au périgée) et analysons le mouvement de m_3 le long de Oz pour différentes vitesses initiales de lancement z'_0 .

Si z'_0 est petit m_3 n'ira pas loin et retombera vers 0, si au contraire z'_0 est grand m_3 aura une évasion hyperbolique et le premier maximum z_1 de z en fonction de z'_0 est donné par la courbe de la figure 6 avec une certaine vitesse de libération L correspondant à une évasion parabolique.

Si $z'_0 < L$ la masse m_3 retombe vers l'origine 0 la dépasse et soit s'en va à l'infini, soit atteint un minimum z_2 avant de revenir vers 0; quelle est la fonction $z_2(z'_0)$?

Si m_3 est repassé en 0 a un instant où m_1 et m_2 étaient soit au périgée,

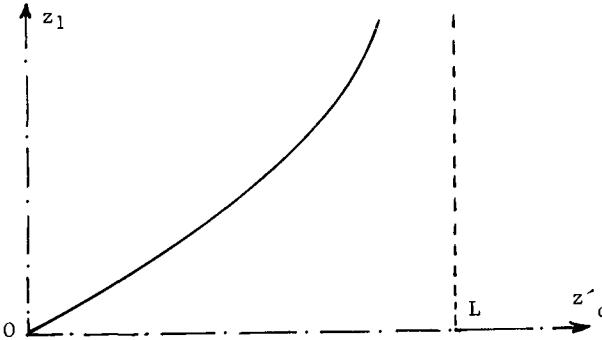


Fig. 6. Le premier maximum z_1 de z en fonction de la vitesse initiale z'_0
(The first maximum z_1 of z in terms of the initial velocity z'_0)

soit à l'apogée on a $z_2 = -z_1$ à cause de la symétrie passé-futur, si m_1 et m_2 étaient en train de se rapprocher on a $0 > z_2 > -z_1$ si au contraire m_1 et m_2 s'écartaient on a $z_2 < -z_1$, on peut même avoir une évasion parabolique ou hyperbolique de m_3 si le gain d'énergie au cours du passage près de 0 est suffisamment élevé et si z_1 était déjà suffisamment grand. D'où la fonction $z_2(z'_0)$ donnée sur la figure 7, elle comporte une infinité d'asymptotes verticales (évasions paraboliques) avec des abscisses telles que L_{A4} , L_{P4} ayant un point d'accumulation à la vitesse L .

Examinons maintenant z_3 , le troisième extrémum de z , la discussion est tout à fait analogue et si l'on fait un agrandissement de l'une des branches de la figure 7, par exemple la branche A_4 , B , P_4 , on obtient la fonction z_3 de z'_0 de la figure 8 avec une infinité d'asymptotes verticales s'accumulant aux vitesses L_{A4} et L_{P4} .

Pour z_4 , z_5 ,... le même schéma se reproduit indéfiniment avec une complexité croissante et l'on voit que si l'on considère une suite arbitraire indéfi-

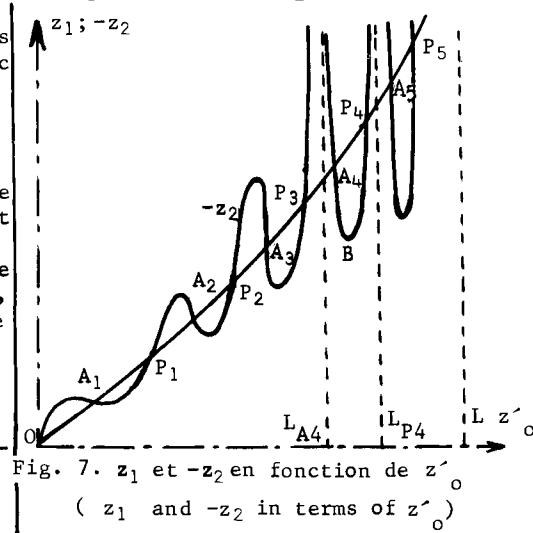
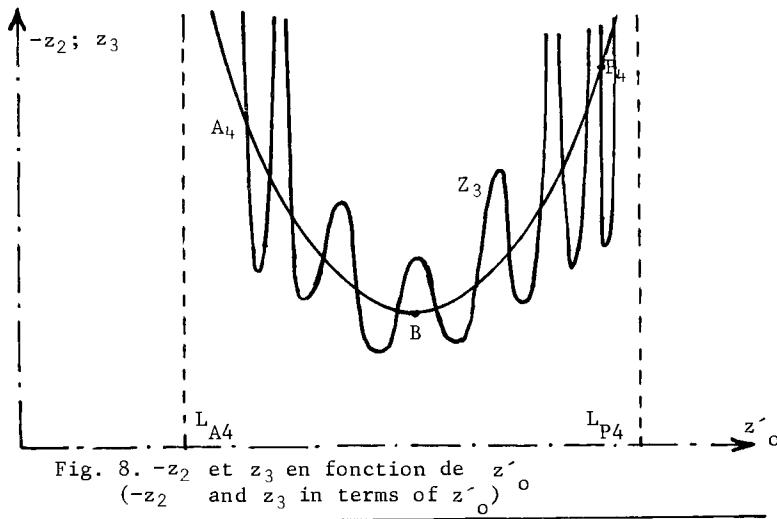


Fig. 7. z_1 et $-z_2$ en fonction de z'_0
(z_1 and $-z_2$ in terms of z'_0)

niment croissante $a_1, a_2 \dots a_p$ il existe au moins une vitesse initiale z'_0 que l'on peut obtenir par élimination, telle que pour tout P on ait $|zp| > ap$; cette vitesse initiale conduit donc à un mouvement de type "oscillatoire".



L'étude de ces mouvements a été développée en (Moser, 1973) et en (Raulic et Hurand, 1975), ils existent certainement pour tous les systèmes de n-corps, dès que $n \geq 3$, quels que soient les rapports des masses mais ils ne représentent sans doute qu'un ensemble de mesure nulle de l'espace des phases.

6 - CONCLUSION

Les systèmes d'équations différentielles, et en particulier les systèmes de n-corps de la Mécanique Céleste peuvent être étudiés par des méthodes qualitatives très diverses et aboutissant à des résultats d'une grande variété qui permettent de préciser beaucoup l'image de l'ensemble des solutions. En dehors des exemples présentés ci-dessus les résultats les plus intéressants concernent les évolutions originelles et finales dans le problème des 3 corps (Chazy 1922, 1929, 1932 ; Schmidt 1959 ; Khilmy 1961 ; Alexeev 1963 ; Sagnier 1969), l'étude du mouvement de 3 corps le long d'une droite (Nahon 1973 ; Irigoyen 1974) et surtout l'existence de solutions quasi-périodiques au voisinage des solutions périodiques et stables dans l'approximation linéaire (Kolmogorov 1957 ; Arnold et Avez 1967 ; Moser 1968).

Les perfectionnements des ordinateurs ont conduit à de grands progrès des méthodes quantitatives et surtout des intégrations numériques, il en est résulté des idées qualitatives nouvelles comme celle de la semi-ergodicité (Hénon 1966) mais ces nouveautés restent très discutées et il conviendrait de déterminer au moins qualitativement l'importance, la structure et le comportement de ces solutions semi-ergodiques qui, en particulier, jouent un rôle essentiel dans les questions de stabilité.

7 - REFERENCES

- Alexeev, V.M. (1963) Problemy dvizhenija isk.neb.tel. Izd-vo Akademii Nauk, SSSR, p 50.
- Arnold, V.I. ; Avez A. (1967) Problèmes ergodiques de la mécanique classique Gauthier Villars, Paris.
- Chazy, J. (1922) Sur l'allure finale du mouvement dans le problème des 3 corps dans le temps croît indéfiniment, Annales de l'Ecole Normale Supérieure, 3ème série, 39, p 29-130.
- Chazy, J. (1929) Journal de Mathématique pure et appliquée, vol 8, p 353.
- Chazy, J. (1932) Bulletin Astronomique, 8, p 403.
- Hénon, M. (1966) Exploration numérique du problème restreint III, Bulletin Astronomique, Paris, Vol 1, fasc 1, p 57.
- Irigoyen, M. (1974) Etude du choc de deux petites masses en présence d'un troisième corps de masse prépondérante, Celestial Mechanics, Vol 9, p 491-506.
- Khilmy, G.F (1961) Qualitative methods in the many-body problem, Russian Tracts on Advanced Mathematics and Physics, Vol 3, Gordon and Breach Science Publishers, New-York.
- Kolmogorov, A.N. (1957) Théorie générale des systèmes dynamiques et mécaniques classiques Proc. Int. Congress of Math. Amsterdam.
- Marchal, C. (1975) Qualitative Methods and Results in Celestial Mechanics, Survey paper, ONERA, tiré à part n° 1975-77.
- Marchal, C. et Saari D.G. (1976) On the final evolution of the n-body problem, Journal of Differential Equations.
- Mather, J.N. et Mc Gehee, R. (1974) Solutions of the collinear four-body problem which become unbounded in finite time, National Science Foundation Grant N° GP 4313X and 38955.
- Moser, J. (1968) Quasi-periodic motion in the 3-body problem, Bulletin Astronomique 3, p 53-59.
- Moser, J. (1973) Stable and Random Motion in Dynamical Systems, Princeton University Press, Princeton.
- Nahon, F. (1973) Trajectoires rectilignes du problème des 3 corps lorsque la constante de la force vive est nulle, Celestial Mechanics, Vol 8, p 169-175.
- Pars, L.A. (1965) A treatise on analytical dynamics, Heinemann London p 435.
- Poincaré, H. (1892, 1893, 1899) Les méthodes nouvelles de la Mécanique Céleste, Gauthier-Villars, Paris.

- Raulic, P. et Hurand, P. (1975) Quelques mouvements du problème des 3 corps,
Rapport de stage - Université Pierre et Marie Curie - Laboratoire de Dy-
namique et Statistique Stellaire - Institut Henri Poincaré - Paris.
- Saari, D.G. (1973) Improbability of Collision in Newtonian gravitational
systems II, Transactions of the American Mathematical Society, 181, p 351.
- Sagnier, J.L. (1969) Aspects qualitatifs du problème des trois corps,
Seminarios do Observatorio Astronomico, N.T. 07/69, 10-17 juin.
- Schmidt, O.J. (1959) Quatre leçons sur la théorie de l'origine de la terre,
Editions en langue étrangère, Moscou.
- Sitnikov, K.A. (1961) The existence of oscillatory motions in the 3 body
problem, Soviet Physics Dokl. 5 p 647-650.

BAHNSTÖRUNGEN DURCH SCHWEREWELLEN

Oskar Bschorr

*Deutsche Forschungs — und Versuchsanstalt für Luft— und Raumfahrt E. V.,
Oberpfaffenhofen, Federal Republic of Germany*

ÜBERSICHT

Aufgabe der Untersuchung ist, Beziehungen über die Strahlungseigenschaften tieffrequenter Schwerewellen abzuleiten, die dadurch verursachten Störungen interplanetarer Bahnen zu berechnen und Bahneigenschaften auf deren Eignung als Schwerewellendetektor zu untersuchen.

Die bekannt gewordenen Schwerewellendetektoren sind vorwiegend zur Registrierung von kurzzeitigen Impulsen, hervorgerufen durch Sternkatastrophen, ausgelegt. Es bereitet erheblichen apparativen Aufwand, deren Empfangsfrequenzen in den unteren Bereich zu legen. Eine Weber-Antenne mit einer Cut-off-Frequenz von z.B. 1 Hz hätte eine Länge von 2,5 km. Andererseits sind nahe Flyby-Vorgänge von Sternen wirkungsvolle Gravitationsstrahler in einem Bereich weit unter 1 Hz.

Um Beziehungen über die Gravitationsstrahlung - hervorgerufen durch Flybys - abzuleiten, wird folgendes Modell zugrunde gelegt: Ein abgeschlossener Sternhaufen wird als kanonische Gesamtheit aufgefaßt. Damit gelingt es, die Lösungstechniken der statistischen Thermodynamik zu übertragen. Für den linearisierten Fall ergibt sich, daß die Gesamtstrahlung mit der achten Potenz und das Maximum des Spektrums mit dem Quadrat der charakteristischen Haufengeschwindigkeit anwächst. Diese Ergebnisse sind analog dem schwarzen Strahler und auch dem Turbulenzlärmstrahler.

In Anlehnung an das Kirchhoffsche Emissions/Absorptionsgesetz kann auch gezeigt werden, daß ein sich bewegendes Massensystem sowohl Gravitationsenergie emittieren als auch absorbieren kann. Bei Absorption nimmt eine Keplerbahn Gravitationsenergie auf, so daß sich die Bahnenergie und damit die Halbachse vergrößert.

Mittels der Gaußschen Methode wurden die Störungen auf die Bahn-elemente, insbesondere der großen Halbachse, untersucht. Bei Zugrundelegung plausibler Schwerewellenintensitäten sind - astronomisch gesehen - die zu erwartenden Bahnstörungen der Planeten gering. Bei sonnenfernen Körpern, wie den Oortschen Kometen, ist der Schwerewellenfluß am größten. Zur meßtechnischen Erfassung tieffrequenter Schwerewellen erscheinen Umlaufzeitmessungen

und Librationspunktsystem in der Art eines Heterodyn-Detektors am empfindlichsten.

EINLEITUNG

Die von Weber (1) - (2) vorgeschlagenen Antennen zum Nachweis von Schwerewellen sind von mehreren Gruppen nachgebaut und verbessert worden (4) - (6). Im wesentlichen handelt es sich hierbei um einen schwingungsisoliert gelagerten Aluminium-Zylinder geringer Eigendämpfung. Erregt durch einen Schwerkraftimpuls wird das Nachschwingen der Eigenmoden beobachtet. Die Grenze der Meßempfindlichkeit bildet das thermische Rauschen. Zur Verbesserung der Empfindlichkeit wurde kryogene Kühlung von Antenne und Nachfolgeelektronik vorgeschlagen, um den Rauschpegel herabzusetzen (7). Es sind auch eine Reihe andere, vom Weber-Typ abweichende Laborexperimente in Erwägung gezogen worden. Einen elektromagnetischen Detektor hat Braginski (8) untersucht. Von Bschorr (21) wurde eine Hornantenne vorgeschlagen, die ab der Cut-off-Frequenz lineare Empfindlichkeit aufweist. Neben den Laboruntersuchungen sind auch geophysikalische und bahnmechanische Effekte in Betracht gezogen worden. Messungen des Abstands von Erde-Mond zeigten eine zu geringe Empfindlichkeit als Schwerewellendetektor. Quadrupolschwingungen der Erde oder des Mondes, angeregt durch Schwerewellen, sind wegen des seismischen Rauschens ebenfalls sehr unempfindlich und auf die Eigenperioden von 54 Minuten bei der Erde und 15 Minuten beim Mond beschränkt (9). Dyson (10) hat auch die Einwirkung von Gravitationswellen auf die Erdkruste untersucht.

Die Arbeitsfrequenzen der bekannten Schwerewellendetektoren liegen vergleichsweise hoch. Die Eigenfrequenz der Weber-Antenne liegt bei 1,6 kHz. Zur Aufnahme kurzzeitiger Impulse, herrührend von Gravitationskollapsen, ist eine solche Cut-off-Frequenz zwar ausreichend, auf der anderen Seite gibt es Vorgänge - z.B. Flyby von Schwarzen Löchern - die sehr viel tiefer frequente Schwerewellen aussenden. Nach Braginski (15) wird bei einem Flyby von 2 Schwarzen Löchern größtenteils 10% der Eigenmasse im Frequenzbereich von 10^{-4} - 10^{-9} Hz abgestrahlt. Durch entsprechende Dimensionierung ist es zwar möglich, die Ansprechfrequenzen herabzusetzen, jedoch der Frequenzbereich mit Schwingungsdauern von Minuten, Stunden, Tagen und Jahren erscheint damit nicht erschließbar.

Es wird auch die Aufgabe gestellt, Aussagen über Intensität und Spektralverteilung von Gravitationsrauschen, hervorgerufen durch die Massenbewegungen (Flybys) in Sternhaufen, zu gewinnen. Im weiteren soll untersucht werden, inwieweit sich interplanetare Bahnen als Indikatoren zum Nachweis tieffrequenter Schwerkraftstrahlung eignen.

THERMISCHES SCHWERKRAFTRAUSCHEN

Modell

In Anlehnung an die elektromagnetische Strahlung soll zwischen einem *thermischen* und einem *nichtthermischen* Anteil an Gravita-

tionsstrahlung unterschieden werden. Der *nichtthermische* Anteil resultiert vor allem aus katastrophalen Sternprozessen wie z.B. Gravitationskollapsen und Supernova-Ausbrüchen. Der *thermische* Teil hat seine Ursache in Wechselwirkungen der Sterne untereinander. Für diesen Entstehungsmechanismus sollen in Analogie zur Theorie des Schwarzen Strahlers die Strahlungseigenschaften angegeschrieben werden.

Als Modell wird ein abgeschlossener Sternhaufen mit einer für statistische Belange ausreichenden Zahl von Mitgliedern (Sterne, Schwarze Löcher), zugrunde gelegt. Diese stehen über Gravitationskräfte in Wechselbeziehung miteinander und können Energien übertragen. Weiter wird unterstellt, daß ein solcher Sternhaufen (quasi)stationär, homogen und isotrop ist.

Mit diesen Voraussetzungen kann der Sternhaufen als kanonische Gesamtheit aufgefaßt werden. Damit lassen sich die bekannten Beziehungen der statistischen Thermodynamik analog auf einen Sternhaufen übertragen.

Aufgrund der Energieübertragung durch Gravitationskräfte stellt sich im Sternhaufen eine Verteilung der Geschwindigkeiten und damit der kinetischen Energie der Sterne ein, die die größte Wahrscheinlichkeit hat. Im besonderen wird diese Verteilung durch einen sogenannten Verteilungsmodul beschrieben. Dieser hat die gleichen Eigenschaften wie die absolute Temperaturskala der Thermodynamik. Im linearisierten Fall ist der Verteilungsmodul θ proportional zum Quadrat der mittleren Geschwindigkeit v der Elemente der kanonischen Gesamtheit.

$$\theta \sim v^2.$$

Für die weiteren Ableitungen wird ferner die Entropie Σ des Systems benutzt. Diese lässt sich auf die Zustandswahrscheinlichkeit zurückführen

$$\Sigma \sim \ln W.$$

Außerdem impliziert das zugrundegelegte Modell neben der Energiekonstanz (1. Hauptsatz) auch die analoge Entropieaussage des 2. Hauptsatzes der Thermodynamik.

Thermische Schwerkraftstrahlung

Für das definierte Modell des Sternhaufens lassen sich mit den Mitteln der statistischen Thermodynamik Aussagen über die abgestrahlte Schwerkraftstrahlung ableiten. Für die Intensität der Gesamtstrahlung ergibt sich analog zum Stefan-Boltzmannschen Strahlungsgesetz (und zum Debye-Gesetz für Festkörperenergien)

$$I \sim \theta^4 \sim v^8.$$

Für die spektrale Intensität $i(v, \theta)$ ergibt sich

$$i(v, \theta) = \theta^3 F(v/\theta).$$

Die zugrundegelegten Modelleigenschaften reichen nicht aus, die volle Spektralfunktion $F(v/\theta)$ abzuleiten, doch können damit Aussagen über das Maximum der spektralen Intensität i_M und v_M abge-

leitet werden, was dem Wienschen Verschiebungsgesetz entspricht.

$$\begin{aligned} i_M &\sim \theta^3 \sim v^6, \\ v_M &\sim \theta \sim v^2. \end{aligned}$$

Es ist bemerkenswert, daß es mit denselben Lösungsmethoden gelang, auch die Eigenschaften des Turbulenzlärms abzuleiten (Lit. 11) und damit das Lighthillsche v^8 -Gesetz zu bestätigen. Dieses besagt, daß die von einer Turbulenzmenge abgestrahlte Schallstrahlung durch Quadrupole hervorgerufen wird und daß deren Gesamtintensität mit der achten Potenz der mittleren Turbulenzgeschwindigkeit anwächst.

Kirchhoff'sche Emissions/Absorptionsbeziehung

Auf die Emissions/Absorptionsbeziehung von Gravitationsstrahlung soll detaillierter eingegangen werden, da sich hierauf besonders die Aussage über säkulare Änderungen der Energien von Raumbahnen stützt. Diese basiert wieder auf dem im vorhergehenden Abschnitt beschriebenen Modellkonzept.

Die Energie des zugrundegelegten Sternhaufens setzt sich aus 2 Komponenten zusammen: Die mechanische Energie (potentielle und kinetische) der Sterne und die Gravitationsstrahlung (daß die Gravitationsstrahlung hier einen kleinen Bruchteil ausmacht, ist für die Gültigkeit der Ableitung ohne Einschränkung). Diese Energiearten werden ständig ineinander umgewandelt. Dieser Mechanismus ist gerade für die Gravitationsstrahlung von Misner (12) abgeleitet worden. Im dynamischen Gleichgewicht wird genauso viel mechanische Energie in Gravitationsstrahlung umgewandelt wie umgekehrt Gravitationsstrahlung in mechanische Energie. Es wird also genauso viel Gravitationsstrahlung emittiert wie absorbiert. Emission bedeutet die Umwandlung von mechanischer Energie in Gravitationsstrahlung, während Absorption den umgekehrten Vorgang darstellt.

In Analogie zu der Kirchhoff'schen Ableitung kann eine Beziehung zwischen dem spektralen Gravitationsemissionskoeffizienten $\epsilon(\theta, v)$ und dem spektralen Gravitationsabsorptionskoeffizienten $\alpha(\theta, v)$ abgeleitet werden. Bei der Herleitung wird auf die Thermodynamikliteratur verwiesen. Die vergleichbare Ableitung der Emission und Absorption des Turbulenzlärms ist in (Lit. 11) dargestellt.

Analog ergibt sich

$$i(\theta, v) = \frac{\epsilon(\theta, v)}{\alpha(\theta, v)} .$$

Im dynamischen Gleichgewicht ist in einem (quasi)stationären Sternhaufen mit dem Verteilungsmodul θ das Verhältnis von spektralem Gravitationskoeffizienten $\epsilon(\theta, v)$ zum spektralen Gravitationsabsorptionskoeffizienten $\alpha(\theta, v)$ gleich der spektralen Intensität der Gravitationsstrahlung $i(\theta, v)$.

Emissionskoeffizient $\epsilon(\theta, v)$ und Absorptionskoeffizient $\alpha(\theta, v)$

beziehen sich auf den Umsatz pro Volumeneinheit. (Auf die Oberfläche des Sternhaufens integriert, ergibt sich dieselbe auf die Oberflächeneinheit bezogene Emissions/Absorptionsbeziehung.)

Die Kirchhoffsche Formel sagt aus, daß ein guter Schwerewellenemitter auch ein guter Schwerewellenabsorber ist und umgekehrt. Im weiteren folgt, daß sich bei 2 Systemen mit unterschiedlichen Verteilungsmoduln θ_1 und θ_2 durch Strahlungsübertragung ein dynamisches Gleichgewicht bei einem mittleren Verteilungsmodul einstellt. Wenn $\theta_2 > \theta_1$, absorbiert das System 1 auf Kosten von 2 Gravitationsstrahlung.

BAHNSTÖRUNGEN DURCH SCHWEREWELLEN

Ausgangsgleichungen

Eine Schwerewelle in Richtung der z-Achse verursacht folgende Störbeschleunigungen:

$$x'' = \frac{c}{2} A'' x, \quad y'' = \frac{c}{2} A'' y, \quad z'' = 0.$$

Für eine monochromatische Schwerewelle ergibt sich der Feldverlauf

$$A = A_0 e^{i\omega(t - z/c)}.$$

Die Intensität I dieser Welle ist

$$I = \frac{c^5 \omega^2 A_0^2}{32 \pi G}.$$

Zur allgemeinen Berechnung der Bahnstörung durch eine Schwerewelle ist es zweckmäßiger, die Bahn in die x,y-Ebene mit der x-Richtung zum Perihel zu legen. Trifft eine Schwerewelle unter dem Winkel ϑ auf die Bahnebene, so werden die Störbeschleunigungen auf den Bahnkörper

$$R = \frac{c}{2} A'' r \cos^2 \vartheta \cos 2f,$$

$$S = \frac{c}{2} A'' r \cos^2 \vartheta \sin 2f,$$

$$W = \frac{c}{2} A'' r \cos \vartheta \sin \vartheta.$$

R, S, W sind die Komponenten der Störbeschleunigung in radialer, tangentialer Richtung und senkrecht zur Bahnebene. Das Quadrat des Einfallscosinus bei R und S kommt einmal durch die Projektionsverkürzung von r und zum anderen durch die Projektion der Beschleunigung in die Bahnebene. Der Faktor 2 bei der wahren Anomalie f kommt durch den Quadrupolcharakter der Schwerewelle.

In der Gaußschen Form sind die momentanen Störungen der Elemente einer Keplerbahn, hervorgerufen durch Störbeschleunigungen R, S, W (Lit. 14).

$$a' = \frac{2}{n\sqrt{1 - e^2}} (R e \sin f + S e \cos f + W),$$

$$e' = \frac{\sqrt{1 - e^2}}{n a} (R \sin f + S(\cos u + \cos f)) .$$

Die Auswertung für den allgemeinen Fall führt auf umfangreiche Reihenentwicklungen, und zwar ergeben sich die Koeffizienten der Reihe wieder durch Reihenentwicklungen. Die Konvergenz verschlechtert sich mit zunehmender Bahnexzentrizität e . Dieser Sachverhalt ergibt aber auch, daß solche Bahnen auch von den Harmonischen der Kreisfrequenz n angeregt werden und nicht wie im linearen Fall nur von der Grundfrequenz allein.

Die Effektivität der Schwerewellenstörung wird auch dadurch vergrößert, daß im oberen Frequenzbereich auch eine höhere Wellenintensität zu erwarten ist.

Impulsstörungen

Am einfachsten sind die Störungen, hervorgerufen durch impulsartige Schwerewellenereignisse, anzugeben. In diesem Sonderfall wirkt der Bahnkörper vergleichbar einem ballistischen Pendel. Die Impulsdauer muß dabei kleiner als die Umlaufzeit sein. Für den Mond als Antennenkörper ergibt dies Impulsdauern von Tagen, bei der Erde solche von Monaten und bei Jupiter von Jahren. Bei Erdsatelliten liegen die entsprechenden Impulse bei Minuten bis Stunden. In solchen Bereichen sind gerade Gravitationsimpulse von Flyby-Vorgängen zu erwarten. Höher abgestimmte Antennen sind nicht in der Lage, derart tieffrequente Vorgänge aufzunehmen.

Die Störung Δa der Halbachse durch einen Impuls der Dauer Δt ergibt sich

$$\Delta a = a' \Delta t = \frac{2 \Delta t}{n \sqrt{1 - e^2}} (R e \sin f + S e \cos f + S) .$$

Für die kreisförmige Bahn, wie sie bei den Großplaneten und bei vielen Morden gegeben ist, vereinfacht sich

$$\Delta a = \frac{d A''}{n} a \Delta t \cos^2 \vartheta \sin^2 f .$$

Die dadurch bedingte Änderung T der Umlaufzeit ergibt sich nach dem 3. Keplerschen Gesetz und mit $n = 2\pi/T$:

$$\Delta T = \frac{3}{4\pi} c A'' \Delta t T^2 \cos^2 \vartheta \sin^2 f .$$

Periodische Störungen

Periodische Störkräfte sind bei rotierenden Sternen mit Quadrupolanteil oder bei engen Doppelsternen zu erwarten.

Mit Beschränkung auf kreisnahe Bahnen und senkrechtem Welleneinfall wird die Störung der Halbachse

$$a' = \frac{1}{n} \frac{c A(\omega) \omega^2}{2} a \sin(\omega t - 2f + f_0) .$$

Die maximale Änderungsgeschwindigkeit ist

$$(a^*)_{\max} = \frac{c A(\omega) \omega^2 a}{n}$$

Der maximale Hub a_{\max} der großen Halbachse ist näherungsweise

$$a_{\max} = \frac{c A(\omega) \omega^2 a}{n(\omega - 2n)} .$$

Es handelt sich hierbei um eine periodische Störung mit der Periode $T_{\text{stör}}$

$$T_{\text{stör}} = \frac{2\pi}{\omega - 2n} .$$

Säkulare Störungen

Die durch Schwerewellenrauschen verursachten säkularen Bahnstörungen sollen unter Umgehung einer aufwendigen Störungsrechnung der Größenordnung nach abgeschätzt werden. Hierbei wird auf die von Planck abgeleitete Analogie zurückgegriffen.

In einem elektromagnetischen Strahlungsfeld befindet sich ein linearer Oszillatator, ein Hertzscher Dipol mit einer bestimmten Eigenfrequenz ω_0 . Ein solcher Oszillatator entnimmt dem elektromagnetischen Wechselfeld Energie und wird so zu Eigenschwingungen angeregt. Die Energieaufnahme geht so weit, bis sich ein Gleichgewicht zwischen Energieaufnahme und Energieverbrauch durch die Oszillatordämpfung einstellt.

Die Schwingungsgleichung eines solchen Oszillators lautet

$$x'' + 2\rho x' + \omega_0^2 x = b(t) .$$

$b(t)$ ist die auf den Oszillatator wirkende Störbeschleunigung.
Die Spektralverteilung von $b(t)$ ist $b(\omega)$:

$$b(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} b(t) e^{i\omega t} dt .$$

Bei kleiner Dämpfung ρ tritt bei ω_0 eine ausgeprägte Resonanzspitze auf, so daß praktisch nur in diesem Frequenzbereich ein Energiumsatz stattfindet. Unter der Voraussetzung, daß $b(t)$ inkohärentes Rauschen darstellt, lässt sich die Gesamtenergie E des Oszillators aufsummieren:

$$\frac{E}{m} = \frac{\pi}{8} \frac{b^2(\omega_0)}{\rho} .$$

Der Leistungsumsatz N des Oszillators mit der Umgebung ist

$$\frac{N}{m} = \frac{\pi}{4} b^2(\omega_0) .$$

Diese Leistung nimmt der Oszillatator aus dem Strahlungsfeld auf und gibt sie über Strahlung (oder anderen Mechanismus) ab.

Zur Abschätzung der Energiebilanz eines auf einer Keplerbahn um-

laufenden Satelliten mit einem Gravitationswellenfeld soll dieses Modell herangezogen werden.

Die Gleichung einer Bahn in der x,y-Ebene lautet

$$x'' + \frac{\mu}{r^3} x = b_x(t) ,$$

$$y'' + \frac{\mu}{r^3} x = b_y(t) .$$

Ohne Einschränkung der Allgemeinheit soll eine kreisförmige Bahn mit $r = a$ ($a =$ große Halbachse) weiter verfolgt werden. Nach dem 3. Keplerschen Gesetz wird die Kreisfrequenz

$$\frac{\mu}{r^3} = \frac{\mu}{a^3} = n^2 \equiv \omega_0^2 .$$

Diese Beziehung, in die Bahngleichung eingesetzt und eine (wenn auch vernachlässigbar kleine) Dämpfung durch Gravitationseigenstrahlung angenommen, ergibt formal dieselbe Ausgangsgleichung wie bei Planck. Ein Unterschied - was vor allem die mathematische Kompliziertheit ausmacht - liegt aber darin, daß der Plancksche Oszillator linear ist, während die Keplerbahn als eine nichtlineare Schwingung mit abnehmender Eigenfrequenz n bei zunehmendem Bahnradius

$$r = \sqrt{x^2 + y^2} ,$$

d.h. bei zunehmenden Ausschlägen x und y , anzusehen ist. Unter Negierung der Nichtlinearität würde ein solcher Kepler-Oszillator aus einer inkohärenten Gravitationsstrahlung senkrecht zur Bahnebene die Leistung N aufnehmen:

$$\frac{N}{m} = 2 \frac{\pi}{4} b^2(n) .$$

Der Faktor 2 berücksichtigt, daß es sich bei der Kreisbahn um 2 Oszillatoren (x - und y -Richtung) handelt.

Es ist zu vermuten, daß die Keplerbahn nicht im Gravitationsstrahlungsgleichgewicht steht, sondern daß die Umgebungsstrahlung sehr viel intensiver ist als die Eigenstrahlung. Jedenfalls ist die Gravitationsstrahlung von Keplerbahnen, selbst über kosmologische Zeiträume gemessen, vernachlässigbar klein. Für diesen Fall wird die Strahlungsleistung N absorbiert, d.h. in mechanische Energie (= Bahnenergie) umgewandelt. Die Bahnenergie E ist

$$\frac{E}{m} = - \frac{\mu}{2a} .$$

Durch Differentiation nach der Zeit t wird

$$\frac{\dot{E}}{m} = \frac{\mu}{2a^2} \dot{a} .$$

N und E gleichgesetzt, ergibt sich eine Beziehung zwischen der

Intensität der Umgebungsstrahlung und der Zunahme der Bahnachse a.

$$a^* = \frac{\pi a^2 b^2 (n)}{\mu} .$$

Da die Umdrehungsdauer T genauer meßbar ist, wird a^* durch T^* ersetzt:

$$T^* = \frac{3}{8\pi} \frac{T^3}{a^2} b^2 (\omega_0) .$$

BAHNSTÖRUNGEN ALS SCHWEREWELLENDETEKTOREN

Satellitensystem

Es ist bereits vorgeschlagen worden, das Erde/Mond-System als Schwerewellenantenne zu verwenden. Dabei war vorgesehen, den Abstand Erde/Mond mittels Laser-Interferenz zu ver messen. Bei einer erreichbaren Meßgenauigkeit von ~ 10 cm ist dieses Verfahren einem Weber-Detektor um Größenordnungen unterlegen. Weber vermag bei einer Antennenlänge von 150 cm Anschläge von 10^{-14} cm zu registrieren. Um dieselbe Genauigkeit zu erreichen, müßte der Erde/Mond-Abstand auf 10^{-6} cm genau vermessen werden.

Genauer als Entfernungsmessungen sind Zeitmessungen. Michelson und Sadler und neuerdings Flandern (16) haben in einem anderen Zusammenhang Messungen der Mondumlaufzeit durchgeführt. Sie benützen dabei die Mondscheibe selbst als Kameraverschluß bei der Beobachtung von Fixsternen. Dabei werden Genauigkeiten von $\Delta T/T$ von $\sim 10^{-11}$ pro Jahr erreicht. Bezogen auf die Messung der Mondentfernung, entspricht dies einer Genauigkeit von $\sim 0,5$ cm.

Die von Flandern (16) festgestellte Abnahme der Mondumlaufzeit von 10^{-10} pro Jahr, auf Schwerewelleinfluß zurück geführt, ergäbe unerklärlich hohe Intensitäten. Jordan (17) deutet diese Abnahme als Beweis für die Diracsche Hypothese der abnehmenden Gravitationskonstanten. Einer solchen Theorie, die eine Zunahme des Erdradius $\Delta R/R \sim 10^{-11}$ pro Jahr bewirken würde, wird aber von Geologen (18) widersprochen.

Librationspunkte (Bild 1)

In Mehrkörpersystemen treten jeweils Librationspunkte auf. Diese sind durch Potentialsenken gekennzeichnet. Als Maß für die Stabilität eines Librationspunktes kann die Fluchtgeschwindigkeit angesehen werden, die notwendig ist, um den Librationspunkt zu verlassen. Die stabilsten Librationspunkte weist das Sonne/Jupiter-System auf. In diesem befinden sich Planetoiden, die sog. Trojaner. Die Librationspunkte liegen auf der Planetenbahn jeweils um 60° versetzt und haben dieselbe Umlaufzeit wie der Planet.

Ein solches Librationspunkt/Planet-System entspricht sinngemäß einem von Braginski et al. (19) vorgeschlagenen Schwerewelldetektor, dem sogenannten Heterodyn-Detektor. Dieser besteht aus 2 unter 90° gekreuzten Balken, die um ihren gemeinsamen Mittelpunkt rotieren. Die gleichsinnig zirkular polarisierte Kompo-

nente einer Schwerewelle regt ein solches System zu Drehschwingungen an. Für monochromatische Wellenzüge, wie sie z.B. von Pulsaren zu erwarten sind, errechnet Braginski eine hohe Empfindlichkeit. Auch bei Impulserregung ist ein solches System empfindlicher als eine Weber-Antenne. Da Planet und Librationspunkt um 60° gekreuzt sind, ergibt sich eine geringföge Ver schlechterung ($\sim 13\%$) gegenüber einer 90° -Anordnung.

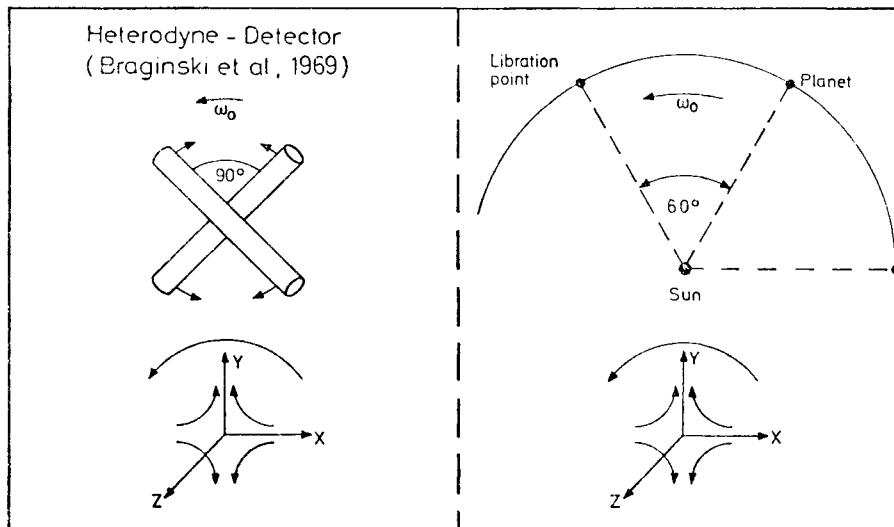


Bild 1. Heterodyne Detector.

Aus der Größe der Fluchtgeschwindigkeit und eines besetzten Librationspunktes kann ein Gravitationsereignis nach oben abgegrenzt werden. Bezug auf die Librationspunkte des Jupiter, liegt diese Grenze sehr hoch. Die Librationspunkte aber der kleineren Planeten bzw. der Monde, die eine tiefere Empfindlichkeitsschwelle haben, unterliegen andererseits den Nahfeldstörungen der Großplaneten, so daß eine Aussage unsicher wird.

Eine Fortentwicklung dieser Methode wären in den Librationspunkten, z.B. des Erde/Mond-Systems, positionierte Dragfree-Satelliten, deren Distanz untereinander mit Laser vermessen wird.

Doppelsatellitensysteme

Der von Press (20) vorgeschlagene Mechanismus zur Aufnahme von Schwerewellen besteht aus einem geschlossenen, z.B. ringförmigen Rohr. Im Rohr zirkuliert eine Flüssigkeit. Der Ring wird gleichzeitig in Rotation versetzt. In einem solchen System ergibt sich unter Einwirkung von Schwerewellen eine Beschleunigung der Flüssigkeitsströmung (Schwerewellen-Zyklotron). Wegen der Flüssigkeitsreibung ist dieser Effekt auch bei Verwendung von Hyperflüssigkeit belanglos.

Sind Planetenbahn und Mondbahn nicht koplanar, so führt der Mond eine vergleichbare Bewegung aus, wie ein Flüssigkeitsteilchen im rotierenden Ringrohr. Analog zu dem Ursprungsmodell ergibt sich eine Beschleunigung des Mondes.

Oortsche Kometenwolke

Während die sonnennahen Planeten- und Mondbahnen relativ stabil und damit weniger durch Schwerewellen beeinflußbar sind, nimmt die Bahnstabilität nach außen ab. Gleichzeitig nimmt die Größe der Störbeschleunigung durch Gravitationswolken mit dem Sonnenstand zu. Kometen in dem sonnenfernen Oortschen Gürtel reagieren so auf Schwerewellen empfindlicher. Eine Störbeschleunigung kann den Kometen auf Fluchtgeschwindigkeit bringen, während eine Verzögerung zu sonnennahen Tauchbahnen führt. Eine direkte Bahnbeobachtung in dieser Entfernung ist zwar nicht möglich, doch indirekt vermag die Häufigkeit von parabelnahen Kometen Indizien zu liefern.

FORMELZEICHEN

A	Schwerewellenamplitude
a	Große Halbachse
c	Lichtgeschwindigkeit
E	Energie
e	Exzentrizität
f	Wahre Anomalie
F	Spektralfunktion
G	Gravitationskonstante
I	Gesamtintensität
i	Spektrale Intensität
J	Besselfunktion
m	Masse
N	Leistung
n	Umlaufkreisfrequenz
p	Ellipsenparameter
R	Radiale Störbeschleunigung
r	Bahnradius
S	Tangentielle Störbeschleunigung
T	Umlaufperiode
t	Zeit
u	Exzentrische Anomalie
v	Geschwindigkeit
W	Normale Störbeschleunigung
W	Wahrscheinlichkeit
x,y,z	Karthesische Koordinaten
α	Absorptionskoeffizient
ε	Emissionskoeffizient
ν	Frequenz
μ	Gm_0
θ	Verteilungsmodul
Σ	Entropie
ρ	Dämpfung
ω	Kreisfrequenz

LITERATUR

- (1) J. Weber, Detection and Generation of Gravitational Waves, Phys. Rev. 117 (1960).
- (2) J. Weber, Gravitational Radiation Experiments, Phys. Rev. Lett. 24 (1970).
- (3) Gravitational Radiation and Gravitational Collapse, Dr. Reichel Publishing Comp., Dordrecht/Boston, 1974.
- (4) J.A. Tyson, Detection of Gravitational Radiation, In: (3).
- (5) P. Kafka, On the Munich-Frascati Weber-Type Experiment, In: (3).
- (6) S. Bonazzola et al., Newton Gravitational Radiation Detection Experiment, In: (3).
- (7) S.P. Boughn et al., The Use of Cryogenic Techniques to Achieve High Intensity, in Gravitational Wave Detectors, In: (3).
- (8) V.B. Braginski, Electro-Magnetic Detectors of Gravitational Waves, In: (3).
- (9) J. Weber, Gravitational Radiation, Phys. Rev. Lett. 18 (1967).
- (10) F.J. Dyson, Seismic Response of the Earth to a Gravitational Wave in the 1-Hz Band, Astrophysics J. 156 (1969).
- (11) O. Bschorr, Untersuchungen über den Turbulenzlärm. DGLR-Jahrbuch 1968.
- (12) C.W. Misner, Mechanisms for the Emission and Absorption of Gravitational Radiation, In: (3).
- (13) C.W. Misner, K.S. Thorne, J.A. Wheeler, Gravitation, W.H. Freeman and Comp., San Francisco, 1971.
- (14) D. Brouwer, G.M. Clemence, Methods of Celestial Mechanics, Academic Press, New York/London, 1961.
- (15) V.B. Braginski, Signal to Noise Ratio in the Satellite Gravitational Wave, XXIV IAF, Baku, USSR, 1973.
- (16) P. Jordan, Die Diracsche Hypothese und die Theorie der Erdexpansion, VDI-Nachrichten 32 (1974).
- (17) P. Jordan, The Expanding Earth, Pergamon Press, Oxford, 1971.
- (18) M.G. Wunderlich, Das neue Bild der Erde, Hoffmann und Campe Verlag, Hamburg, 1975.

- (19) V.B. Braginski, Y.B. Zel'dovich, V.N. Rudenko,
Reception of Gravitational Radiation of Extraterrestrial
Origin,
Zh. Eksp. 9, Theor. Fiz. Pis'ma 10, 437-441 (1969).
- (20) W.M. Press, K.S. Thorne, Gravitational Waves Astronomy.
Ann. Rev. Astron. Astrophys. 10, 335-374 (1974).
- (21) O. Bschorr, Graviskop.
MBB TN BT22-1/75 (1975).

PLAN FOR FLIGHT EVALUATION OF ATTITUDE STABILIZATION AND FLEXIBLE SOLAR ARRAY DYNAMICS FOR THE COMMUNICATIONS TECHNOLOGY SATELLITE

Frank R. Vigneron and R. A. Millar

*Department of Communications, Communications Research Centre, P.O. Box 11490
Station 'H', Ottawa, Ontario, Canada K2H-8S2*

ABSTRACT

The Communications Technology Satellite (CTS) will be launched into synchronous orbit with a Thor Delta 2914 vehicle in early 1976. CTS is a joint program of the Canadian Department of Communications (DOC), the National Aeronautics and Space Administration (NASA), and the European Space Agency (ESA). The agreements between DOC and NASA formally identify several communications satellite systems technologies to be developed and flight tested in the course of the program, among which are a 3-axis stabilization system to maintain accurate antenna boresight pointing on a spacecraft with flexible appendages, and an extendible lightweight solar array with an initial power output greater than one kilowatt. The present paper describes the corresponding plan and technical activity which is presently being implemented by the Department of Communications in conjunction with SPAR Aerospace Products Limited, University of Toronto Institute for Aerospace Studies, and Telesat Canada.

The 3-axis attitude stabilization system for CTS utilizes the 'momentum wheel with offset thrusters' principle for on-orbit stabilization. The accuracy of stabilization is projected to be $\pm 0.1^\circ$ in roll and pitch, and 1.1° in yaw. The solar array for CTS has a beginning-of-life power of 1.2 kilowatt. It employs two solar panels, each 653 cm long and 130 cm wide. The array folds up concertina fashion for stowage, and in deploying and deployed states has significant structural flexibility. CTS will carry instruments to enable flight test, among which are array accelerometers, an array deflection sensing system, array load cells, and temperature sensors. A number of passive 'monitoring' and active 'special excitation' events will be performed during the two-year mission, to flight test attitude stabilization and solar array design, and to explore structural flexibility and the flexibility and control interaction.

In the paper, the stabilization system and solar array are briefly described. Pre-launch analysis, ground test, and parameter determination are discussed. Basic uncertainties which require flight test validation are discussed, and corresponding flight test objectives are outlined. The flight instrumentation to be carried on CTS is described. A computer-based data processing system to be used to handle, distribute, and analyse the flight data, is described. The passive and active events to be monitored during the CTS mission are outlined and discussed.

INTRODUCTION

An upper bound on spacecraft power for the flight-proven versions of current communications satellites, which are typified by Anik or Intelsat IV, is dictated by the choice of the "dual spin" configuration (in which the maximum spacecraft power is governed by the number of cells which can be mounted on the rotating drum of the structure) and the choice of launch vehicle. The bounds appear to be about 500 watts and 1.2 kw for dual spin satellites launched by the Thor Delta and Atlas Centaur vehicles, respectively. 3-axis stabilized spacecraft (i.e., spacecraft which employ a configuration in which a central body is stabilized with respect to the earth by active autonomous means and has deployed from it a lightweight sun-tracking solar array in the arrangement shown in Fig. 1) enable a power level of 1.2 kw to be achieved within the limitations imposed by a Thor Delta launch. 3-axis stabilized

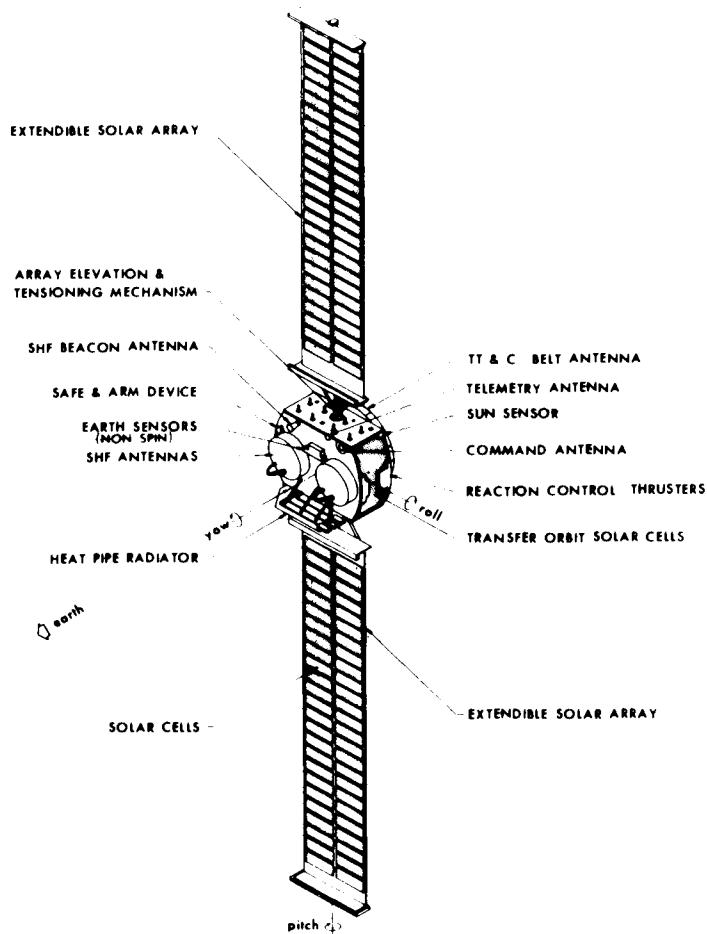


Fig. 1 CTS Configuration, On Station, Array fully deployed.

spacecraft hold the promise of being more cost effective for many communications missions than dual-spinners, but at present do not have the flight qualification status (in communications missions) associated with the latter type. Of particular concern with regard to 3-axis stabilized spacecraft are factors associated with the attitude control, and related structural flexibility implications of the large lightweight solar arrays.

Active 3-axis stabilization has been successfully utilized for a number of missions in orbits other than synchronous, for example, Nimbus, OGO, OAO, Ranger, Mariner, and Apollo. The problem of adapting 3-axis stabilization to commercial communications missions lies mainly with the requirement for operation in synchronous orbit with long life (in excess of five years), high reliability, low weight, and high boresight pointing accuracy, to the same degree as required for dual-spin stabilized spacecraft. However, the current consensus is that the development of 3-axis control for communications satellite systems does not necessitate a major technological breakthrough, but rather demonstration and accumulation of a flight record with one or more of many promising stabilization systems appropriate for this type of application.

Phenomena relating to structural flexibility of spacecraft contribute a major operational uncertainty to mission planning. The available design tools for large flexible vehicles are based on analytical modelling, related computer simulation and test data on components. Ground-based confirmation of the design at the systems level is not possible because the configurations are not generally structurally self-supporting in the earth's one-g environment (system tests based on air bearing which are of value in design of small relatively rigid spacecraft, will play almost no role in design of large flexible spacecraft). Of a number of flexible deployable spacecraft arrays of the type appropriate for communications missions, the FRUSA is the only one upon which flight data has been reported (1); thus for the 'Deployable Solar Array' class of appendage, a data base essential to confident spacecraft design is not yet available. Qualification of higher power 3-axis stabilized spacecraft with regard to spacecraft flexibility must necessarily involve future flight demonstration and acquisition of flight-derived structural dynamics data.

The Communications Technology Satellite (CTS) will be launched into synchronous orbit with a Thor Delta 2914 vehicle in early 1976. CTS is a joint program between Canada, Department of Communications (DOC), and the United States, National Aeronautics and Space Administration (NASA). The European Space Agency (ESA) is also participating by supplying certain communications components and by supporting the development of the solar cells and flexible substrates for the deployable solar array. The agreements between DOC and NASA formally identify several communications satellite systems technologies to be developed and flight tested in the course of the program (2) among which are a deployable solar array with an initial power output greater than one kilowatt, and a 3-axis stabilization system to maintain accurate antenna boresight pointing on a spacecraft with flexible appendages. The following paper describes the plan for flight evaluation in attitude stabilization and control, and flexible spacecraft dynamics for CTS. Particular attention is given to the technical objectives, the ground evaluations and test program, various mission events planned, the baseline of anticipated results, the complement of spacecraft instrumentation, and the data handling system.

GENERAL GOALS AND RATIONALE ASSOCIATED WITH THE CTS TECHNOLOGY INVESTIGATIONS

The general goals in the CTS program in solar array and attitude stabilization technology are: (i) design and develop appropriate subsystems (hardware) for CTS, and qualify them by flight; (ii) establish a data base sufficient to enable confident extrapolations of these technologies to meet design requirements of future communications satellites of the CTS type.

The above goals cannot be met solely by analysis and ground-based testing, because these spacecraft are too large and flexible to be tested adequately in simulated zero-g laboratory conditions. Also it must be recognized that the complement of spacecraft instrumentation which it is practical to fly (i.e. in terms of weight, feasibility, etc.,) falls short of that normally associated with conclusive deductive experimentation, particularly in the structural flexibility area. For these reasons the approach adopted in the CTS program involves:

- (a) a heavy emphasis on the development of mathematical models which can be used to unify the ground-based test data and then extrapolate to predict on-orbit performance;
- (b) extensive ground test, particularly at the component level, to establish the hardware characteristics;
- (c) acquisition of flight data, with a small (but hopefully well designed) set of instruments, over a set of planned passive and active mission events;
- (d) quantitative comparison of flight measurement with predictions (referred to above), appropriate iterations in analysis and ground test, to finally consolidate a sound technical base.

THE DEPLOYABLE SOLAR ARRAY

The solar array for CTS includes two fold-up deployable solar array units (DSA's) each of which, when deployed, is approximately 721 cm long by 130 cm wide and supplies 630 watts (beginning-of-life) of electrical power. Each DSA consists of a flexible blanket (upon which are mounted solar cells and associated bus wiring)*, a single BISTEM stainless steel book, an inboard pallet, an outboard pressure plate, and related substructure**, configured as shown in Fig. 2. Internal to the spacecraft is a drive and track mechanism for maintaining the DSA normal to the sun, a deployment actuator, a slip ring assembly, and other associated equipment. The portion of a DSA which protrudes from the spacecraft is structurally flexible, and had a total mass of approximately 19 kg. The blanket substrate is a composite fabric 76 μm thick, made up of a kapton layer bonded to a glass fibre reinforcing layer. Solar cells, coverglasses, and adhesives, having a total thickness of 440 μm , are bonded to the fiberglass side of the substrate. Electrical power and instrumentation flexible bus lines are located at the edges and at the centre-line of the blanket. At the inboard end, these bus lines connect via a diode board to a flexible connector cable which in turn connects to the spacecraft electrical subsystem via a slip ring assembly. The blanket is structurally supported by a 3.40 cm diameter deployable BISTEM which is offset from the blanket approximately 6.3 cm as shown. The ends of the blanket are supported by the inboard pallet and the outboard pressure plate. During launch the pallet and plate serve as a protective sandwich structure for the stowed blanket. The blankets are deployed concertina fashion during the early phases of the mission (Fig. 3). The deployed blanket is maintained at a nominal tension of 35.6 N via a constant force spring mechanism located at the

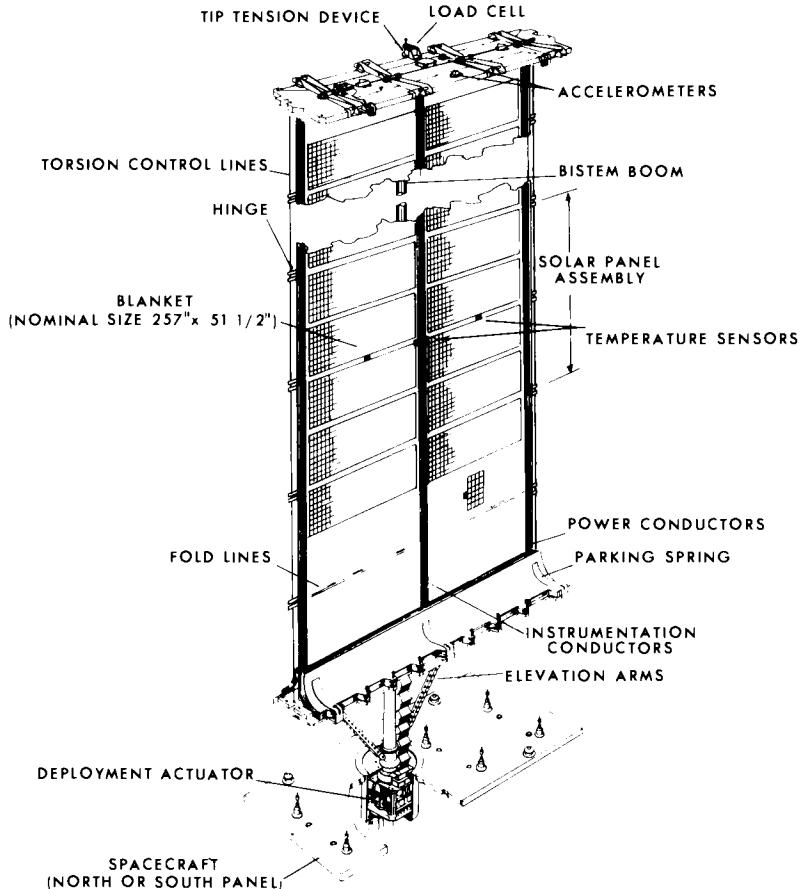


Fig. 2 CTS Deployable Solar Array (DSA)

outboard end of the boom. To provide torsional stiffness, at each edge of the blanket there is a thin stainless steel line tensioned to a constant force of 1.3 N. The outboard parts rotate relative to the spacecraft, via a drive and track mechanism which is energized by a stepper motor. The DSA is normally rotated at a rate of one revolution per day. In addition, a capability exists to rotate at a rate of 35.7°/min. An autotrack feedback control system which

*The blanket was designed and built by AEG-Telefunken, Hamburg, Germany.

**The mechanical portions of the DSA were designed and built by SPAR Aerospace Products Limited, Toronto, Canada.

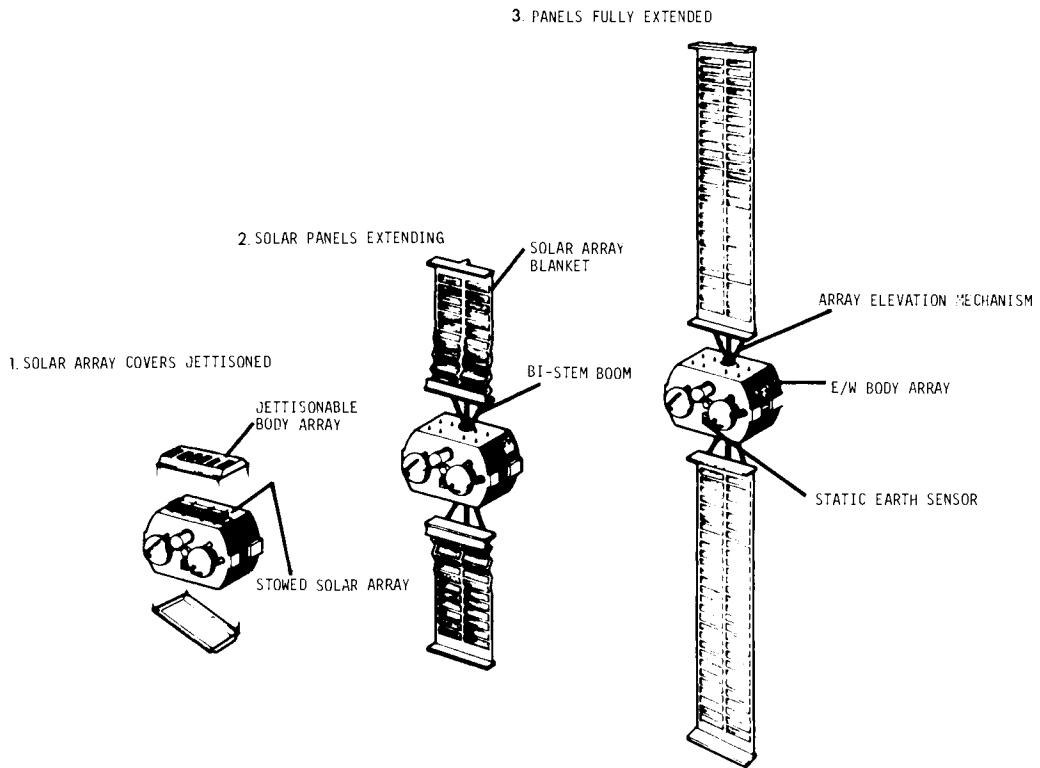


Fig. 3 Array Deployment

utilizes an analog sun sensor mounted on the inboard pallets can be used to automatically step the DSA. Instrumentation will be discussed in a later section. A more complete description of the DSA is given in (3) and (4).

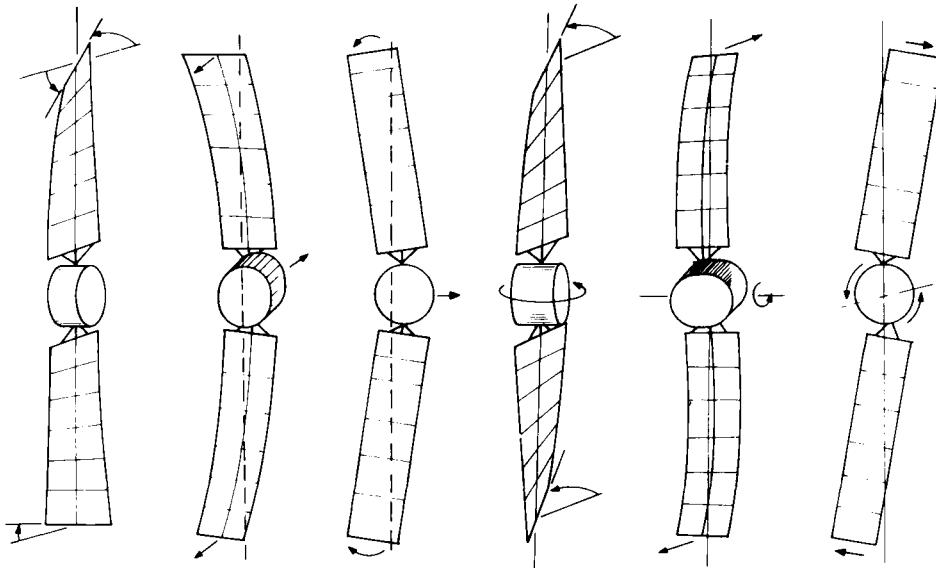
Analysis and Ground Testing. Each of the components of the array (the BISTEM, elevator arms, blanket, torsion control devices, tip tension device, drive and tracking assembly) have been tested individually to establish elastic properties, mechanical deadzones and hysteresis, thermal characteristics, and mass properties. Tests of this level are relatively easy to perform successfully in a laboratory (one-g) environment, and provide basic parameter values.

Vibration tests have been conducted in vacuum and thermal environment under one-g conditions (5, 6). The work has established structural modes and associated frequencies and damping factors, the influence of temperature environment on the modes, and the existence of minor non-linear superharmonic

vibration characteristics. A corresponding analysis has been done to establish a mathematical model which enables confident calculation of the major constrained (fixed-base) modes and associated frequencies under one-g conditions, and extrapolation of the values to the corresponding zero-g on-orbit state (7). The modeling techniques involve linearization of motion equations, and fall short of describing the latter-noted superharmonic vibration characteristics. Further, the torsional (twist) frequencies of vibration of the DSA are found to be very dependent on tension distribution within the blanket, which in turn, is dependent on temperature, twist configuration, and factors associated with the manufacturing and assembly of the units.

The modal frequencies of the flight spacecraft (two DSA's and a central body) have been calculated on the basis of the knowledge described in the preceding paragraph (8, 9). Confident prediction may be made in a quantitative sense, but is hindered in accuracy in a minor way due to a number of manufacture and assembly related factors which are not directly controllable or measurable, among which are: the above noted sensitivity of torsional frequencies to blanket tension distribution; uncertainty related to asymmetries in final on-orbit mass distribution and DSA tension; uncertainty associated with certain stiffnesses in the BISTEM and drive and track assembly bearings. Prediction of corresponding modal damping may be done only with low confidence (i.e., within an order of magnitude) because of similar factors. Predictions are illustrated in Fig. 4.

MAJOR NATURAL STRUCTURAL MODES OF VIBRATION



SYMMETRIC (EVEN) RAD/SEC			ANTISYMMETRIC (ODD) RAD/SEC		
PREDOMINANT TWIST	PREDOMINANT OUT-OF-PLANE	PREDOMINANT IN-PLANE	PREDOMINANT TWIST	PREDOMINANT OUT-OF-PLANE	PREDOMINANT IN-PLANE
1.06	0.79	1.62	1.08	2.37	4.71
3.18	3.13		3.19	3.16	
5.80	5.78		5.81	5.84	

Fig. 4 Major Natural (unconstrained) Structural Modes of Vibration of CTS

An analysis and test program to establish the thermal behaviour of the blanket and boom configuration has been conducted (10). In addition thermal flutter possibilities have been assessed with the conclusion that flutter is not expected to occur. The nominal temperature profile of boom and blanket is depicted in Fig. 5.

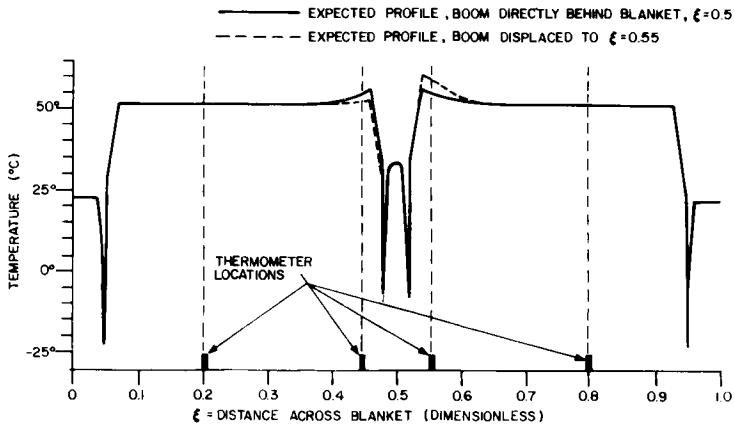


Fig. 5 Temperature Profile of Boom and Blanket

An extensive program has been conducted involving in excess of 70 deployments of the DSA on a specially designed horizontal deployment rig, both in ambient and in thermal vacuum conditions (11). The program serves to qualify the mechanical and electrical aspects, and to some extent gives a qualitative confidence in the structural dynamics aspects. Unfortunately it does not appear possible to simulate the on-orbit "zero-g" state of deployment in ground testing. Quantitative knowledge of structural dynamics during deployment is based solely on analytics and component test information.

THE ATTITUDE STABILIZATION AND CONTROL SYSTEM

The stabilization and control system for CTS is designed to serve three distinct stages during the CTS mission: first it provides control of the orientation of the spacecraft in its spinning configuration during the latter phases of launch (transfer orbit, injection into synchronous orbit, and station acquisition); second it provides attitude stabilization and control during a transition phase from a spinning non-deployed configuration to an on-orbit operational state in which the solar array is deployed; and third, it provides 3-axis stabilization to high accuracy during a two year operational mission. Attitude control during the first stage (transfer orbit through to station acquisition) for a Thor Delta vehicle launch into synchronous orbit is regarded as a developed and mature technology, and will not be elaborated on further in this paper.

Design, parameters, performance characteristics, and other information pertaining to CTS are conceptually as explained in (12), with the following revision: the secondary mode controller has been deleted; the sun sensor

complement has been changed to incorporate digital sensors with 4π steradian field of view). The sensor, control components, and actuators of CTS are listed and described briefly in Table 1 and illustrated in Fig. 6.

TABLE 1 Stabilization and Control Components

Component	Weight(lb)	Major Parameters
Earth Sensor (non-spin)	11.2	Field of View = $\pm 21.7^\circ$ Linear range = $\pm 2.8^\circ$ Accuracy = $\pm 0.05^\circ$
<u>Components</u>		
Digital Sun Sensor	5.9	$\pm 1^\circ$ acc. over 4π steradian FOV $\pm 275^\circ$ acc. in $0.1^\circ \times 64^\circ$ FOV
Attitude Control Electronics	11.9	
Momentum Wheel	17.3	Momentum 15. ft.lb.sec at 3750 rpm
Earth Sensor (spin)	3.0	Acc. $\pm 0.1^\circ$ at 60 rpm
Nutation Damper	0.85	Time Constant of 30 min.; threshold 0.05°
3-Axis Rate Gyro	2.5	Null Acc. $\pm .12^\circ/\text{Sec}$; Dynamic Range $\pm 10^\circ/\text{Sec}$.
Miscellaneous hardware & wiring harness	1.9	
Total Weight of System	54.4	
<u>Other</u>		
Fuel Weight for On-Orbit Stabilization	8.9	Two years in normal mode
Fuel Weight for Transfer & drift orbit precessions	3.4	
Fuel Weight for Station Acquisition	27.0	
Fuel Weight for Attitude Acquisition	2.1	
Reaction Control System tanks & hardware	39.5	Two 5# thrust engines sixteen 0.15# thrust engines

A catalytic blowdown hydrazine reaction control system with sixteen low thrust engines, each of nominally 0.2 lb. thrust level, will provide stabilization and control torquing for the second and third stages of the mission (13).

Attitude Acquisition. The second stage, referred as the 'attitude acquisition phase'*, commences from an initial state with the angular momentum vector of the (spinning) spacecraft normal to the synchronous orbit plane, and progresses through a sequence of elementary dynamic manoeuvres listed below:-

- a) Passive spin stabilization (of up to two days);
- b) Despin from 60 rpm to about 2 rpm via yaw thrusters and a yaw rate gyro (spin about maximum moment of inertia axis);
- c) Nutation monitoring (rate gyros and non-spinning sun sensor) and active nutation damping via roll and pitch thrusters if necessary;
- d) Further reduction of spin rate using yaw thrusters;
- e) Acquire and maintain sun line of sight lock along the negative roll axis using pitch and yaw angle information from a non-spinning sun sensor;

* Designed and implemented by SED Systems Ltd., Saskatoon, Canada.

- f) 90° rotation about pitch to transfer sun line lock from the negative roll axis to the positive yaw axis;
- g) Solar array deployment;
- h) Momentum wheel spin up, enable on-board constant wheel speed hold mode;
- i) Precess about sun line to place the positive pitch axis southerly and positive roll axis approximately tangential to the orbit trajectory;
- j) Rotate about pitch to acquire earth in the non-spinning earth sensor field of view;
- k) Activate the on-board attitude control system to enable capture of the final on-orbit pointing mode.

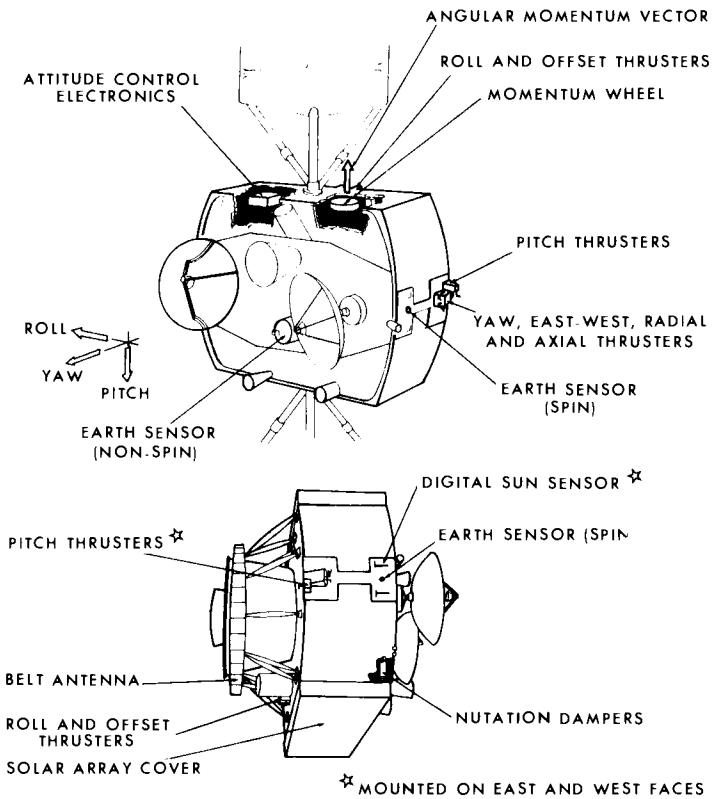


Fig. 6 Sensors, Control Components, and Actuators for Attitude Control

Maneuvers (a-d), (g), and (k) are, from a controls viewpoint, 'man-in-the-loop' maneuvers (initiations, corrections, and trims are made by a human operator on the basis of off-line analysis of transmitted spacecraft data). Maneuvres (e-f) and (h-j) involve some extent of 'ground-based computer in the loop' operation: angular or angular rate information is sensed by the spacecraft, then telemetered to ground; the signals are used as inputs to a

control algorithm which is resident in a triply-redundant HP2100 minicomputer which, in turn, will compute commands for an appropriate thruster firing sequence; the sequence is then transmitted to the spacecraft and executed.

The above sequence minimizes spacecraft weight for this function, as minimal on-board components and power (battery storage) are required. The sequence is comprised of non-complex functions from a controls viewpoint, although it is rather complex from an operations viewpoint. The controls side is quite tractable, since: (a) both rate and position are sensed directly; (b) a high pointing accuracy is not required (± 2 degrees at most); (c) the controllers of this phase must control or stabilize only for a matter of hours during the total mission; (d) maximum flexibility in algorithm implementation is available as a result of using ground based minicomputer in the control loop. A distinct disadvantage is that the system is dependent on RF links to a ground station with which are associated time delays and overall operational and implementation complications. Further details on the controls implementation is given in (14).

On-orbit 3-axis Stabilization. The CTS satellite utilizes a 'momentum wheel with offset thrusters' stabilization system for 3-axis on-orbit stabilization.* The spacecraft is a 'tri-spin' satellite in this mode, in that the momentum wheel, the spacecraft body, and the solar array each rotate at different rates about a common minimum of inertia axis (the pitch axis).

The earth sensor provides pitch and roll error signals to separate channels of the controller electronics. The pitch controller is of the linear proportional-integral-derivative (PID) type and utilizes the pulse width modulator driven momentum wheel as its torque actuator. Tachometer feedback around the pulse width modulator and wheel (with the PID controller disconnected) provides the constant wheel speed hold mode referred to earlier. The pitch thrusters can be activated for momentum dumping via ground command.

The roll/yaw controller (or offset controller) is basically an augmented dual time constant pseudorate controller which activates, via response to earth sensor derived roll error signals, a set of thrusters whose torque vector is offset from the roll axis in the roll/yaw plane. To minimize fuel usage over the mission a roll error signal noise filter can be switched in or out, and any one of four pseudorate parameter sets can be selected via ground command. It is planned that earth capture will utilize the higher gain parameter set four with the noise filter out and that the nominal on-orbit pointing mode will use parameter set one with the noise filter in. More detail on the CTS implementation of the system is given in (12), (15) and (16).

Simulation and Ground Test. Each of the components of the attitude control system (the earth sensor, the sun sensor, the controller electronics, the thruster and the pulse width modulator and wheel) have been tested individually and to some extent collectively (earth sensor, control electronics, wheel, and sun sensor, rate gyros, ground station) to establish accurate models for subsequent incorporation in both digital and hybrid computer computer simulations of the complete control loops including flexible body dynamics. In addition, the hybrid computer simulation was used in a real time mode to close the loop around the controller electronics to verify its dynamic integrity.

* Designed and manufactured by SPAR Aerospace Products Ltd., Toronto, Canada.

Of necessity, the on-orbit controller design (hardwired as opposed to programmable) had to be fixed before accurate models for the array dynamics based on test results of flight representative hardware could be established. Simulation results indicate that stability and pointing accuracy, particularly in roll and yaw, depends on the loop time delay and the somewhat elusive array structural frequencies and damping factors. Recent analysis and simulation studies indicate that large oscillations of the fundamental flexible mode, if excited to sufficiently large amplitude, (for modal frequencies and damping slightly lower than those expected for CTS) could diverge in time (unstable behaviour) due to non-linearities associated with the pseudorate controller. In this connection, DSA stepping and slewing intervals, momentum dumping and east-west station-keeping thruster sequences have been designed so as not to excite the flexible modes of vibration.

TECHNICAL QUERIES AND CORRESPONDING OBJECTIVES

The CTS type of configuration (i.e., two flexible sun-tracking DSA's deployed from an earth-tracking main body configured as per Fig. 1) to date has not been flight tested, but is the most popular of several possible concepts of the coming generation of 3-axis stabilized communications satellites.

The following CTS objectives in solar array dynamics and attitude control illustrate the nature of queries that are, at present, outstanding with regard to technology development for CTS and other spacecraft of this class:

1. Establish by flight measurement, and correlate with theory, the solar and magnetic torque environment for the CTS configuration.
2. Demonstrate that flight measurements and theory are in accord for on-orbit 3-axis stabilization as implemented by the momentum wheel with offset thrusters principle (on-station mode).
3. Demonstrate that flight performance is in accord with theory for stabilization and control during attitude acquisition with particular attention to: (a) active nutation damping; (b) two axis stabilization about the sunline; (c) stabilization during solar array deployment (d) intentional precession maneuvers.
4. Demonstrate that measurement and theory are in accord for the static and quasi-static flexible solar array mechanics, with particular attention to: (a) deployed static shape in sun-light and eclipse; (b) deformation history while traversing in the shadow-sun boundary.
5. Demonstrate that measurement and theory are in accord for dynamic solar array mechanics, with particular attention to: (a) solar array deployment dynamics; (b) estimation of internal array forces and bending moments; (c) modes of vibration, modal frequencies, and identification of corresponding modes and frequencies; (d) damping associated with dynamic motion; (e) magnitude of deformation response to a known force input.
6. Explore by flight experiment, the effect of spacecraft structural flexibility on stabilization system operation.
7. Acquire flight performance data on on-orbit 3-axis stabilization, implemented by generic derivatives as are feasible within the scope of the hardware and ground based control capability.

With regard to Objective 1, the popularity of the CTS type of configuration

stems in good part from the fact that it is nominally symmetric about the roll-yaw plane, so that without a great deal of inconvenience, the spacecraft layout can be arranged with respect to spacecraft center of mass and solar pressure location, so that the secular part of the solar torque is nominally zero. This is a tremendous advantage from an attitude control point-of-view because it enables the torque actuators and fuel consumption for stabilization to be minimal. The realities that complicate this concept are that the configuration acquires asymmetries due to sun-induced thermal deformation, manufacturing tolerances, and uncertainties associated with the on-orbit deployment process. The calculation of these torques is not a deterministic process, and there is at least an order of magnitude of error uncertainty associated with present estimates.

With regard to Objective 2, a major index to be determined is the stabilization accuracy in roll, pitch, and yaw. Calculations of this index involve a rather complex error analysis, and associated statistical addition. This question can be answered in a non-statistical way only through acquisition of flight data.

With regard to Objective 3, experience will be gained in order to better judge to what extent 'ground based control' modes of operation are feasible and desirable in the acquisition and on-orbit phases of the mission.

Objectives 4-6 are aimed at establishing whether or not the theory regarding structural flexibility and stabilization (referred to in a previous section) is sufficiently complete and accurate, a prime question in light of experience documented in (17).

Objective 7 stems from a basic query, "which implementation of 3-axis stabilization is best suited to the CTS class of spacecraft?". CTS will enable the acquisition of flight data on the "momentum wheel with offset thrusters implementation", and will show its advantages and weaknesses. It will also be possible to acquire flight experience with a number of variants (generic derivatives), for example: (i) 'momentum wheel with offset thrusters' augmented by sun sensor derived yaw information, and (ii) momentum wheel with open loop roll and yaw thruster control.

INSTRUMENTATION ON CTS

The following complement of instruments (depicted in Figs. 2 and 7) will be flown on CTS, specifically with a view to making flight measurements in support of solar array and attitude stabilization technology.

- (a) Array Accelerometers. Six accelerometers, three each on the outboard pallet of each DSA, will enable inflight measurement of pallet accelerations normal to and in the plane of the blanket. The accelerometers are Systron-Donner Servo type, with range ± 100 milli-g's and resolution of about 0.7 milli-g's. Various accelerometers can be sampled at one of 3, 4, or 5 times per second, as noted in Fig. 7. The accelerometer signals will be filtered on the spacecraft to eliminate electromagnetic and related spacecraft noise at above 1 Hz. The full scale range is interfaced with spacecraft telemetry in a non-linear manner in order to expand ± 30 milli-g's over about one-half the available telemetry range, at the expense of a compression of $\pm (70-100)$ milli-g's over the remaining half of the range.

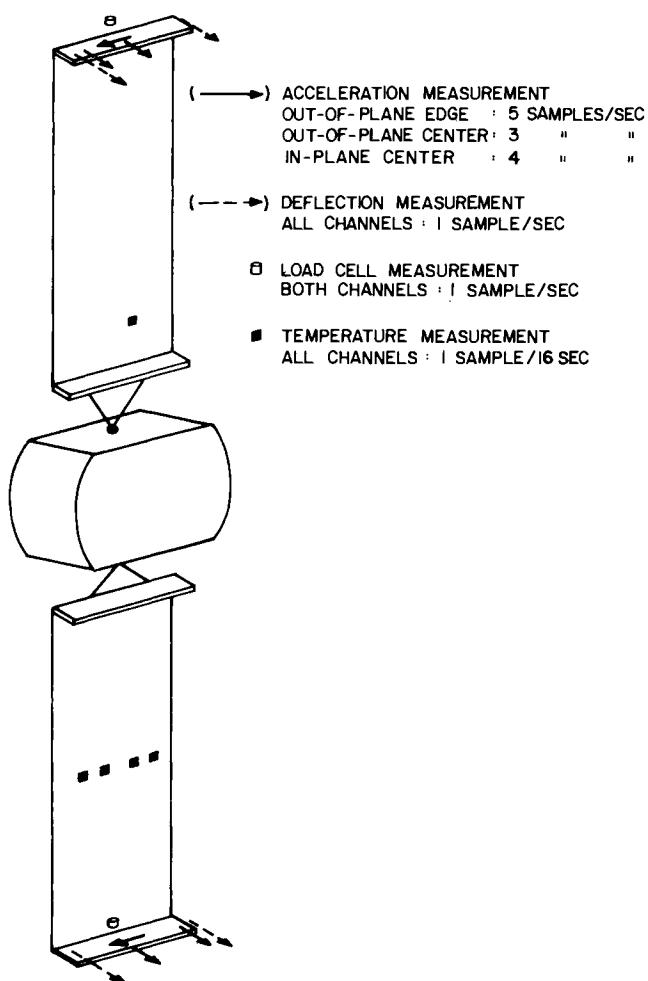


Fig. 7 Schematic Representation of Instrumentation

- (b) Array Deflection Sensing System. The spacecraft will carry four deflection sensors which will enable measurement of components of static and quasi-static deflection of the array tip relative to the spacecraft. Each unit consists of an emitter and detector mounted on the DSA inboard pallet, and a retroreflector mounted on the outboard pallet. The emitter, a light emitting diode, emits a cone of light, a portion of which is returned by the retroreflector to the detector. The detector, which is made with appropriately masked silicon cells, (similar to an analog sun sensor), senses the two components (in-plane and out-of-plane) of the angle of the returning light; out-of-plane or in-plane deflection is then equal to the appropriate angle multiplied by the array length. The information can be sampled and transmitted to the ground once per second. The

sensors will operate in sunlight or eclipse, during deployment, and for the 2 year life of the spacecraft. The field of view and resolution of each sensor is ± 30.5 cm \times ± 25.4 cm, and approximately ± 3.8 cm of tip deflection respectively.

- (c) Array Temperature Sensors. Five platinum resistance thermometer temperature sensors with measurement range of -200°C to 120°C will enable temperature measurement on the blanket of the DSA. The temperatures can be sampled and transmitted to ground with a resolution of about 1° once every 16 seconds.
- (d) Array Tension Load Cell. A load cell with a specially designed strain gated beam is incorporated into each DSA at the outboard pallet. The devices enable measurement of sail tension over the range 0 to 110 N, and can be sampled by telemetry once per second.
- (e) Solar Cell Test Group. A solar cell test group will enable measurement of eleven reference points of power-related data (I-V curve, etc.) on the DSA. Each measurement can be sampled and transmitted to ground once every 16 seconds.

In addition, a number of sensors which are provided to support spacecraft operations are sampled by telemetry once per second and hence can be used as instruments for flight evaluations. These are boom length potentiometers, potentiometers in the tip tension spring mechanisms, digital angular position encoders, DSA mounted sun sensors, extension arm microswitches, earth sensor, sun sensor, rate gyros, pulse width modulator, pulse width counter, wheel speed counter, offset pulse width counter, offset pulse rate counter, and thruster fired indicator. These subassemblies are sampled once per second by spacecraft telemetry.

DATA PROCESSING

The flow of flight data is illustrated in Fig. 8. Data taken by the spacecraft is transmitted to the CRC ground station in binary form, where it is transferred to Digital Original Data Record tapes (2 per day) for permanent retention. The tapes are then processed (synchronization, decommutation, and conversion from binary to engineering units) on a daily basis in CRC's Sigma 9 computer, and separate tapes containing appropriate subsets of the data are created for transmittal to a number of external participating agencies (note next section). Data may be further processed at CRC via an additional Sigma 9 computer program which executes special calculations (combinations of data, averaging, fast fourier transformtions, etc.), tabulates output, and generates graphs*.

MISSION EVENTS AND EXPERIMENTAL APPROACH

Mission events as presently planned (18) are listed and discussed below, and are shown vs. a time line in Fig. 9.

Attitude Acquisition Monitoring. During control maneuvers (despin, active nutation damping, sun acquisition, 90° pitch rotation, earth acquisition, precessions about yaw, and capture of the normal mode), data from all instruments will be recorded. Data will be processed off-line and compared to expected

* The data processing software was designed and implemented by HiTech Canada Ltd., Ottawa, Canada.

performance, and deviations will be analyzed.

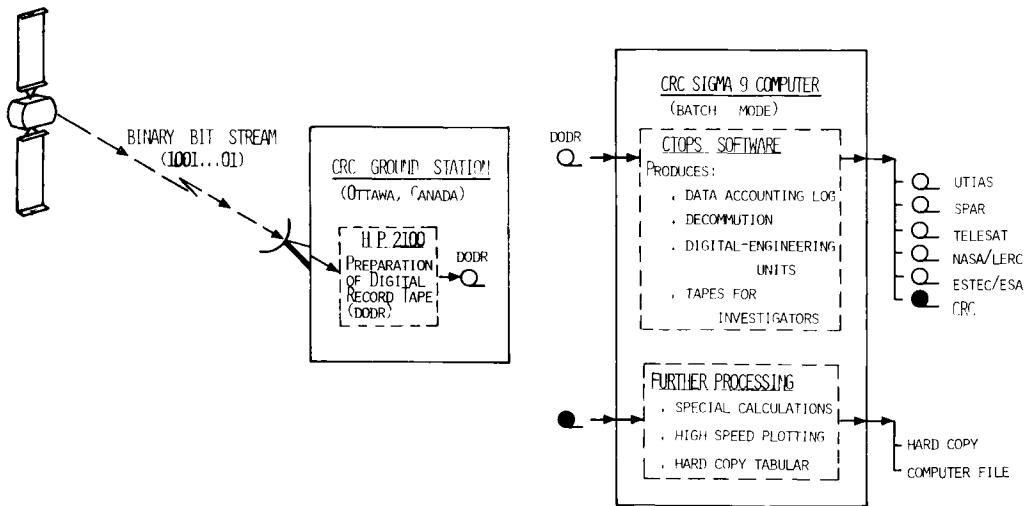


Fig. 8 Processing of Flight Data for CTS

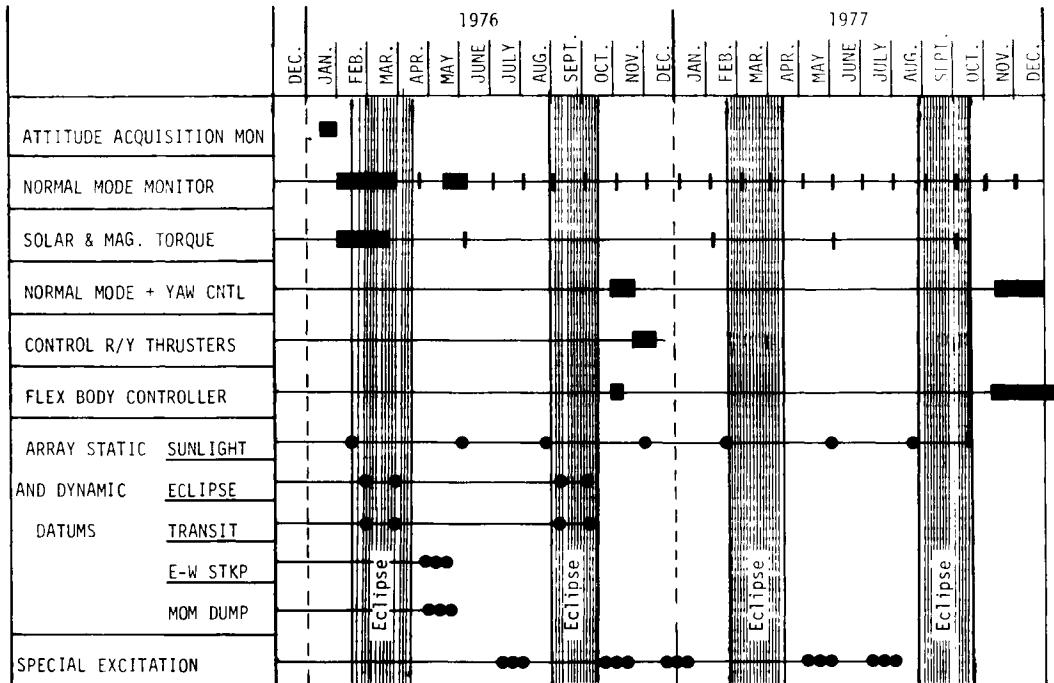


Fig. 9 Mission Events vs Time

The deployment of the DSA will be evaluated in relation to design specifications. The jettisonable covers flag and elevation arm lock flags will be monitored prior to and at the initiation of deployment. The boom extension potentiometer, the deflection sensing system, the six accelerometers, and the load cell (blanket tension) will be monitored as a function of time to record history and characteristics during deployment and immediately after final tensioning. An anticipated hesitation during boom extension will be observed from deployment motor current and boom extension potentiometer measurements. The tip extension potentiometer and on-off microswitch will be monitored to determine when extension ceases. Performance of the DSA orientation mechanism will be evaluated from DSA sun sensors, spacecraft sun sensor, and encoder disc measurements vs. time.

Normal Mode Monitoring. A datum for normal mode pointing accuracy over $2\frac{1}{2}$ days of undisturbed operation (no momentum dumping or east-west stationkeeping) will be determined within the first month of operation, and monitored on a monthly basis thereafter. Pointing accuracy will also be noted during momentum dumping, east-west stationkeeping, DSA stepping and slewing, and eclipse. Pitch and roll pointing accuracy will be determined directly from earth sensor information. Flight measurements of offset thruster pulse rate, pulse width, and roll angle will be compared with simulation results to infer what yaw pointing accuracy is being achieved. In addition, the practicalities of determining yaw from telemetered roll information using modern estimation techniques will be assessed. These data will be compared and combined with pitch, roll, and yaw information extracted as is possible from sun sensor information. ACS performance response to parameter changes, changes in noise filter configuration, and perturbative disturbances will be monitored and compared with expectations from simulation results.

Solar and Magnetic Torque Determination. With the spacecraft operating under nominal conditions, the array dynamics and gravity torques will be negligible. The attitude control system will be appropriately cancelling the external torques due to solar pressure and the earth's magnetic field. Monitoring the control system torquing history (pulse width modulator, pulse width, and wheel speed in pitch, and offset thruster pulse width and rate counter in roll/yaw) thus will permit an evaluation of the external torque. The torque will be determined from $2\frac{1}{2}$ day operation periods in sunlight (during the first month and every four months thereafter), and operation periods in eclipse. In eclipse when the solar torque is zero, an attempt to determine magnetic torque from control system torquing action will be made. The difference between magnetic torque level and the torque level in sunlight will then be the solar torque. In the second year of the mission, there will be an opportunity to inhibit the thruster system completely, and estimate torque on the basis of attitude drift.

In addition an independent calculation of solar torque can be obtained from information on the solar array static shape provided by the deflection sensing system. The solar torque obtained in this manner can then be compared with that obtained from observations of control system performance.

Array Static Shape Datum and Monitoring. The static shape of the solar array is governed by mechanical loads on the structure, initial shape of the boom as manufactured, and thermal effects which vary throughout the mission. For normal operation, it is expected that the boom will assume a final shape after deployment which will change only in eclipse. A datum for array static shape will be determined within two days of attitude acquisition. Measurements will

be repeated 18 times during stable on-orbit conditions and during eclipse entry, transit, and exit. The static shape and quasi-static variations will be deduced from measurements made with the deflection sensing system, load cells and temperature sensors. Accelerometers will be monitored to verify that the behaviour is static.

Deflection measurements will be correlated with theory and ground test results. In addition the torque as determined by RCS thruster operation (see Solar Torque Determination) will be cross correlated with solar torque as deduced from deflection measurements. The lateral temperature distributions across the blanket will be determined with the temperature sensors and compared to theory and ground test measurements. The location of the boom centerline relative to the longitudinal centerline of the blanket influences the lateral blanket temperature distribution (Fig. 5), and thus temperature measurements provide a means of deducing relative positions of boom and blanket. Changes in substrate dimension as a function of temperature and time will be examined on the basis of data from the tip tension potentiometer and deflection sensing system.

Array Dynamics Datum and Monitoring. A datum for array dynamics will be determined via the six accelerometers during stable on-orbit conditions. Array dynamics will be followed also through standard housekeeping events (momentum dumping, east-west stationkeeping, and array fast slew) and during eclipse entry, transit, and exit.

It is anticipated that for normal operation the dynamic motions will be at or below the threshold of the accelerometers. The objective is to show that these motions are below a certain level and thereby validate this aspect of the theory. In event that this is the case, Special Excitation of the type discussed below will be implemented to induce measurable dynamic states.

Special Excitation. A number of special excitation events will be performed in which spacecraft vibration modes will be excited by carefully controlled sequences of thruster firings. The experiment design will be based on extensive ground testing of engineering model and flight arrays; the events have been 'qualified' by ground test (6) to demonstrate that no damage to blankets and no severe spacecraft motion results from them.

Modal frequencies will be deduced from accelerometer data, and the deflection sensing system, the earth sensor and accelerometer phase differences and magnitudes will be used to sort and identify the various classes and modes. Information from theoretical models and ground based test will be used to supplement the measured data. Damping will be determined by observing the decay of modal amplitudes of motion following special excitation operations. The accelerometer response to a commanded pulsing sequence will be obtained; the gains (i.e., output/input) deduced from the measured data will be compared with those predicted by theoretical model and ground test program results.

Hybrid computer simulation of a typical excitation and corresponding acceleration response is shown in Fig. 10 (extracted from (19)).

Normal Mode Supplemented with Yaw Control. Yaw information may be measured directly by sun sensor over portions of the orbit. It is planned, as an experiment over a 2 day period, to supplement the normal mode operation with appropriate thruster firings based on sun sensor readings processed on the ground in real-time or possibly off-line.

Ground Control of Roll/Yaw Thrusters. It has been postulated that the onboard real-time control of roll and yaw attitude could be replaced or improved by ground commanded trim operations done at periodic (e.g., weekly) intervals in a manner similar to that now used for dual-spin spacecraft. It is planned to evaluate this concept with CTS in a 2 day event in the second year of the mission.

Flexible Body Controller. In the latter stages of the CTS mission, a 2 day event is planned in which an alternate ground based control law (wide band-width controller and control with accelerometer feed-back) which will account for flexibility effects will be implemented.

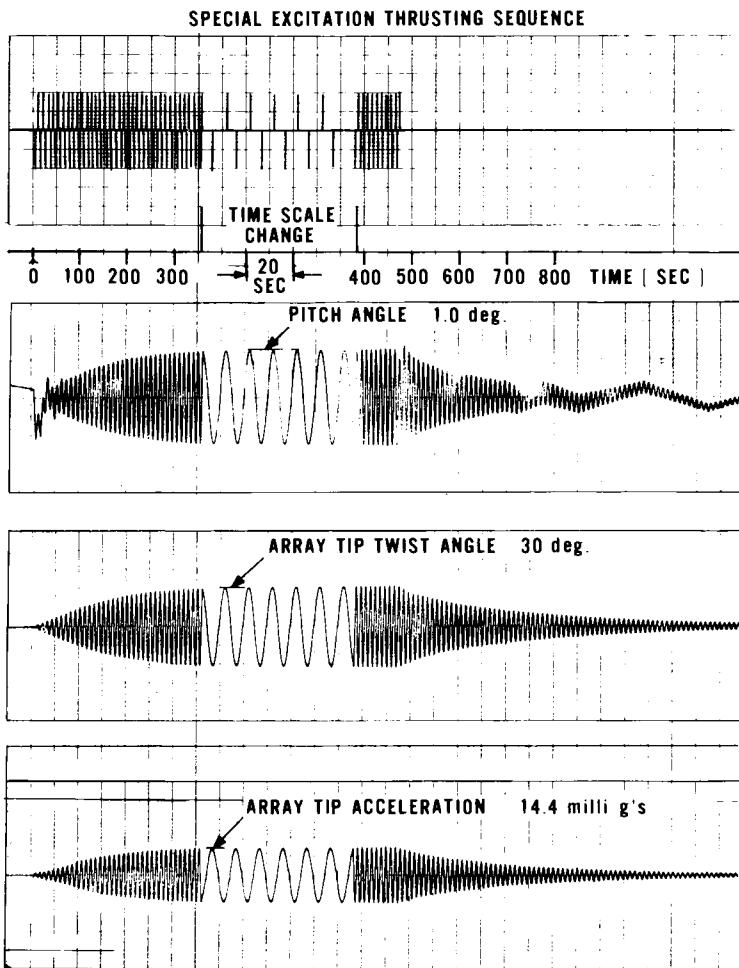


Fig. 10 Typical Special Excitation and Corresponding Response

PARTICIPATING AGENCIES

Dr. P. C. Hughes (University of Toronto Institute for Aerospace Studies (UTIAS), Dr. R. Ravindran (SPAR Aerospace Products Ltd.), Mr. B. Burlton (Telesat Canada), are active participants in the design and implementation of the mission events of the preceeding section. The six accelerometers are being supplied to the CTS Program by UTIAS and SPAR under the sponsorship of the National Research Council of Canada.

NASA representation on solar array technology aspects of the CTS program is provided by Dr. F. Shaker (NASA Lewis Research Center).

ESA participation in the solar array blanket technology is provided via Dr. K. Bogus (ESTEC/ESA). Development of instruments on the blanket (thermo-couples, solar cell test patch, and associated electronics) are supplied to the CTS program by ESTEC/ESA.

REFERENCES

- (1) G. Wolff, and A. Wittman, "The Flight of the FRUSA", AIAA Paper No. 72-510, AIAA 9th. Electric Propulsion Conference, April 17-19, 1972.
- (2) C.A. Franklin and E.H. Davidson, "A High-Power Communications Technology Satellite for the 12 and 14 GHz Bands", AIAA Progress in Astronautics and Aeronautics, Vol.32, MIT Press, 1973.
- (3) R. Buhs, "Layout and Technology of the CTS Solar Array Blanket", presented at the Photovoltaic Specialists Conference, IEEE? 1974.
- (4) E. Quittner, S. Sachdev, and D. Gossain, "Mechanical Design of a Deployable Flexible Solar Array System for a Communications Satellite", CASI/AIAA Joint Meeting, Toronto, Ontario, 30-31 October 1974.
- (5) T.D. Harrison, "Communications Technology Satellite Deployed Solar Array Dynamics Tests", CRC Report No. 1264, Department of Communications, Ottawa, Canada, January 1975.
- (6) T.D. Harrison, R. Buckingham, and F.R. Vigneron, "Functional and Dynamics Testing of the Flexible Solar Array for the Communications Technology Satellite", 8th. Conference Space Simulation, NASA SP-379, Paper No. 24, November 1975.
- (7) F.R. Vigneron, "A Structural Dynamics Model for Flexible Solar Arrays of the Communications Technology Satellite", CRC Report No. 1268, Department of Communications, Ottawa, Canada, March 1975.
- (8) P.C. Hughes, "A Model for the Attitude Dynamics of CTS With Reference to Attitude Control System Design", Aerospace Engineering and Research Consultants Ltd., AERCOL Report No. 75-14-4, March 1975.
- (9) P.C. Hughes and G. Sincarson, "Numerical Results of the Updated Dynamics Model for the Flexible CTS Spacecraft", Aerospace Engineering and Research Consultants Ltd., AERCOL Report No. 75-14-5, March 1975.
- (10) V.A. Wehrle, F.C. Yip, S.J. Zurawski, "On the Thermal Behaviour of the CTS Solar Array for Small Quasi-Static Displacements", Private Communication, to be published as a CRC Report.
- (11) J.S. Smith, S.S. Sachdev, and J.A. Hunter, "Testing of the Communications Technology Satellite Deployable Solar Array Subsystem", Paper No. 2.1, 11th. IEEE Photovoltaic Specialists Conference, May 6-8, 1975.

- (12) F.R. Vigneron, D.V. McMillan, and J.R. Kettlewell, "The Attitude Stabilization and Control System for the Communications Technology Satellite", NTC'72 Record, the 1972 National Telecommunications Conference, December 4-6, 1972, IEEE Publication 72CHO 601-5-NTC.
- (13) V.J. Sansevero, C.D. Arvidson, W. D. Boyce, and S.F. Archer, "Monopropellant Hydrazine Reaction Control System for the Communications Technology Satellite", AIAA Paper No. 73-1268, AIAA/SAE 9th. Propulsion Conference, November 5-7, 1973.
- (14) D.A. Bassett and K. Krukewich, "CTS Spacecraft Ground Station Batch Mode Simulation: Description", Vol.I and II, Document No. 5144-SW-102, SED Systems Ltd., Saskatoon, Canada, July 17, 1974.
- (15) G.B. Lang, C.P. Trudel, D.A. Staley, "CTS Attitude Control Subsystem Analysis Final Report", SPAR-R.626, SPAR Aerospace Products Ltd., Toronto, Canada, September 1974.
- (16) R.A. Millar and F.R. Vigneron, "The Effect of Structural Flexibility on the Stability of Pitch for the Communications Technology Satellite", Proceedings of the 1975 Canadian Conference on Automatic Control, University of British Columbia, June 23-24, 1975.
- (17) P.W. Likins, and H.K. Bouvier, "Attitude Control of Non-Rigid Spacecraft", Astronautics and Aeronautics, May, 1971, pp. 64-71.
- (18) "Attitude Control Systems Experiment Operations Plan", Private Correspondence, prepared by A.S. Brown, R.A. Millar and F.R. Vigneron, Communications Research Centre, Ottawa, Canada, April 1974.
- (19) R.A. Millar, "Deformation and Accelerations of the CTS Spacecraft", Private Correspondence to J. Phillips and F.R. Vigneron, CRC 6655-3 (SCS) Program, Communications Research Centre, Ottawa, Canada, 14 July 1972.

DEVELOPPEMENT ET REALISATION DES STRUCTURES DES RESERVOIRS D'ERGOLS 1er et 2 ème ETAGE ARIANE

Henri Abiven & Jean-Claude Groult

Aerospatiale Route de Verneuil — 78130 Les Mureaux, France

INTRODUCTION

Dans le cadre de l'organisation industrielle du programme ARIANE, l'AEROSPATIALE, au titre du contrat "Etagiste" :

- assure l'intégration des trois étages,
- réalise les structures des deux premiers étages, la coiffe et les interétages, le développement de la plupart de ces structures étant sous-traité à des coopérants étrangers.

Comme on peut le voir sur la figure 1, où est indiqué pour chaque structure principale le nom de la Société responsable du développement, le caractère européen du programme ARIANE est tout à fait respecté :

- la coiffe est développée par la Société CONTRAVES (SUISSE),
- la case à équipements par la Société MATRA (FRANCE), la fabrication de la structure étant assurée par la Société CASA (ESPAGNE),
- le réservoir du troisième étage par la Société L'AIR LIQUIDE (FRANCE),
- le bâti-moteur 3ème étage et les deux interétages par la Société FOKKER (HOLLANDE),
- le réservoir 2ème étage par la Société DORNIER (ALLEMAGNE),
- les jupes Avant et Arrière et le bâti-moteur 2ème étage par la Société ERNO (ALLEMAGNE),
- les impulseurs de séparation par la Société SNIA-VISCOSA (ITALIE),
- les réservoirs 1er étage par l'AEROSPATIALE (FRANCE),

Ariane. Structures principales
répartition industrielle

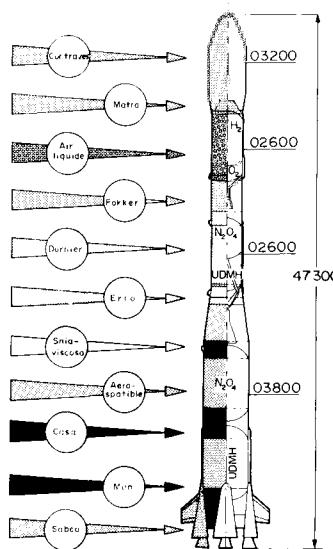


Fig.1 - Structures principales
Répartition industrielle

- les jupes Avant et inter-réservoir 1er étage par la Société CASA (ESPAGNE),
- le bâti-moteur 1er étage par la Société MAN (ALLEMAGNE),
- et les empennages et carénages des moteurs par la Société SABCA (BELGIQUE).

L'objet de cette communication est de présenter quelques aspects du développement des réservoirs des deux premiers étages (Figure 2) structures réalisées d'une part, par l'AEROSPATIALE (réservoirs 1er étage), et d'autre part, par la Société DORNIER (réservoir 2ème étage) en sous-traitance de l'AEROSPATIALE.

DEVELOPPEMENT DES RESERVOIRS 1er ET 2ème ETAGES

Fonction

La fonction de ces réservoirs est :

- de contenir les ergols principaux,
- d'assurer la continuité mécanique entre les structures adjacentes,
- de supporter divers équipements internes ou externes (anti-vortex, filtres, tuyauteries, gouttière électrique, anti-ballottants pour le réservoir deuxième étage, etc.).

Exigences principales

Les exigences principales spécifiées pour le développement sont indiquées sur la figure 3 :

- les ergols sont les mêmes pour les deux étages, N₂O₄ (peroxyde d'azote) et U.D.M.H. (diméthylhydrazine asymétrique),
- la pressurisation des réservoirs 1er étage est assurée par des gaz chauds (400°C) provenant d'un générateur utilisant les ergols principaux et dont la fonction principale est de fournir des gaz d'entraînement des turbines. Pour le réservoir deuxième étage, la pressurisation est assurée par de l'hélium ;

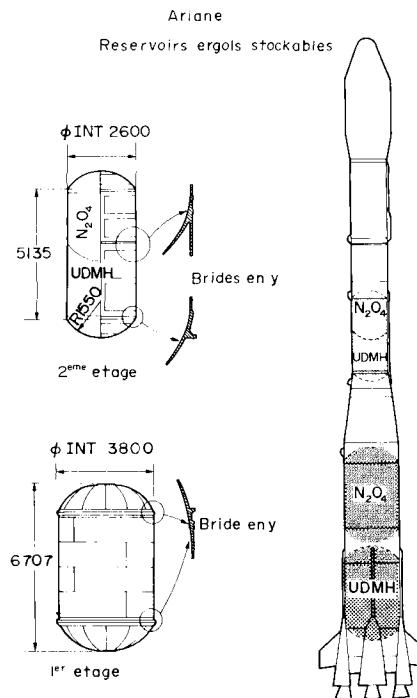


Fig. 2 - Réservoirs ergols stockables

- le réservoir deuxième étage comporte un fond commun séparant les compartiments N₂O₄ et UDMH ; par contre, pour le premier étage, les deux compartiments sont séparés : cette solution a été adoptée en raison surtout des dimensions du 1er étage ; la perte de performances correspondante est compensée par le fait que, compte tenu des masses d'ergols à emporter et de leur densité respective, les deux réservoirs peuvent être identiques, ce qui se traduit par un coût de développement moins important.

Exigences principales

	Réservoir 1 ^{er} étage	Réservoir 2 ^{ème} étage
Ergols	N ₂ O ₄ ~ UDMH	N ₂ O ₄ ~ UDMH
Architecture	2 réservoirs séparés identiques	2 compartiments avec fond commun
Diamètre interne	3,800 m	2,600 m
Volume	≈ 66 m ³ (Nou U)	≈ 16 m ³ (N ou U)
Pressurization	gaz chauds (N+U) θΔ 400 °C	Helium, θ ambiante
Pression	4 bars relatifs	3,5 bars absolus
Masse	< 2050 kg	< 805 kg

Concept et Principes de Fabrication

Les structures des réservoirs ont été conçues de manière à répondre à l'objectif du CNES, organisme chargé par l'Agence Spatiale Européenne de l'exécution du programme ARIANE, objectif qui est de développer un lanceur simple et rustique.

Cet objectif est atteint :

- en utilisant au mieux l'expérience acquise au cours de programmes antérieurs,
- en homogénéisant au maximum les solutions,
- en choisissant des solutions simples et éprouvées, les innovations technologiques étant limitées au strict minimum.

Du fait des exigences, le concept des deux réservoirs 1er et 2^{ème} étages présente quelques différences, la principale étant celle relative à la nature du matériau. Par contre, pour les liaisons fond-virole de réservoir et les liaisons réservoir-structure adjacente, des solutions identiques sont appliquées.

Matériau

Compte tenu de la température des gaz de pressurisation, l'acier 15CDV6 a été retenu pour le réservoir 1er étage ; ce matériau est largement utilisé dans les constructions aérospatiales, en particulier pour la fabrication des réservoirs du premier étage DIAMANT B, et il présente l'avantage de ne pas nécessiter de traitement thermique après soudage. Les caractéristiques principales du matériau sont indiquées sur la figure 4 : résistance à

Etage L 140

E 15CDV6 C3

Acier moyennement allié

Soudable sur état traité

Composition		Caractéristiques mécaniques			Applications technologiques	
		R	Forgeé	Laminé		
C	0,10 - 0,16					
Mo	0,8 - 1,0		1180 MPa	1180		
V	0,2 - 0,3					
Mn	0,8 - 1,1					
Si	0,20 max		Ro2	930 MPa	780	
Ni	0,50 max					
Cr	1,25 - 1,50					
S	< 0,006	A %	12	12		
P	< 0,015					
Fe	le reste		sens long			

Fig. 4 - Caractéristiques 15 CDV6

rupture 118 hb et résistance à 0,2 % limite élastique 93 hb pour les pièces forgées et 78 hb pour les produits laminés.

Pour ARIANE, le 15 CDV6 utilisé est du type E, les différences d'élaboration par rapport à l'acier coulé à l'air étant schématisées sur la figure 5 : le lingot obtenu à partir de la coulée à l'air est ensuite refondu, soit sous vide, soit au travers d'un laitier électroconducteur.

L'aptitude au soudage des tôles est améliorée et l'homogénéité des pièces forgées est meilleure. Du point de vue composition chimique, ce sont les teneurs en soufre et phosphore qui diffèrent :

$$S = 0,030 \% \text{ (coulée à l'air)} \quad 0,006 \% \text{ (coulée affinée)}$$

$$P = 0,030 \% \text{ (")} \quad 0,015 \% \text{ (")}$$

Les ergols U.D.M.H. et surtout N₂O₄ étant corrosifs vis à vis de l'acier, la surface intérieure du réservoir est protégée par un shoopage d'aluminium.

Pour le deuxième étage, l'alliage choisi est l'AZ5G (7020) dont les caractéristiques principales sont indiquées dans la Figure 6 (R = 35 hb, Ro2 = 28 hb).

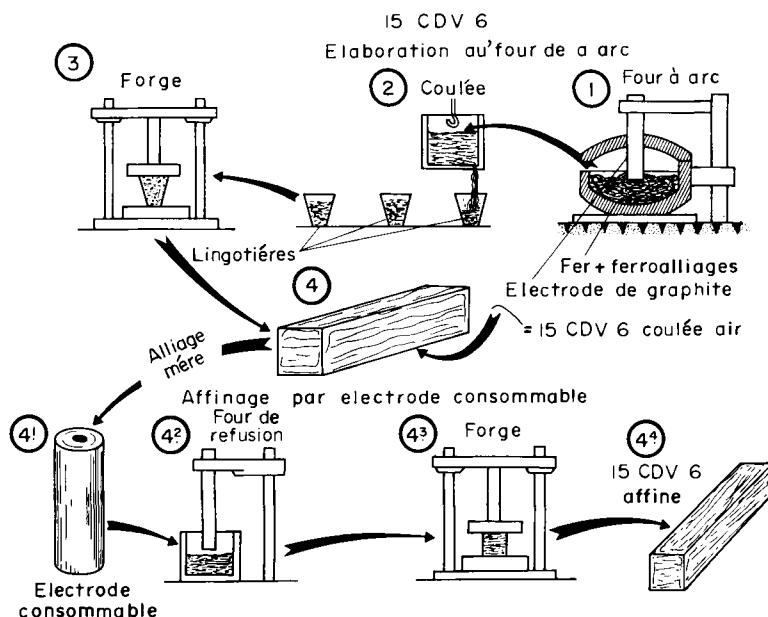


Fig. 5 - Elaboration du 15 CDV6 type E

Etage L33
AZ5G (7020)
Alliage d'aluminium de corroyage
A traitement thermique
Soudable sur état traité

Composition		Caractéristiques mécaniques		Applications technologiques
Zn	3,7 - 5	R	350 MPa	Emboutissage
Mg	0,9 - 1,5	RO,2	280 MPa	Soudage
Al	le reste	A %	10	Usinage chimique
Si	0,35			Repoussage
Fe	0,40			
Cu	0,20			
Mn	0,05 - 0,50			
Cr	0,35			
Ti+Zr	0,08 - 0,25			
Mn+Cr	≥ 0,15			
Zr	0,08 - 0,20			
		mini sens long		

Fig. 6 - Caractéristiques AZ5G

Principe de liaison fond-virole et réservoir-structure adjacente

Pour les réservoirs des deux étages, une bride en forme d'Y assure la liaison, d'une part entre virole et fond, et d'autre part entre réservoir et structure adjacente :

- une branche de la bride est soudée au fond,
- une deuxième branche est soudée à la virole,
- la troisième branche comporte des alésages pour le boulonnage avec la structure adjacente.

Ce type de liaison par boulonnage est appliqué pour toutes les structures principales du lanceur, l'interchangeabilité étant assurée par l'utilisation de gabarit étalon.

Soudage

Pour le soudage, le procédé T.I.G. (Tungstène Inert Gas) bord à bord est utilisé ; toute soudure à clin ou par points dans la structure résistante est proscrite pour des raisons de corrosion en présence des ergols.

Fabrication de l'ébauche des brides en Y

Le principe d'élaboration des ébauches des brides en Y (acier et AZ5G) est schématisé sur la figure 7 :

- poinçonnage du lingot,
- laminage circulaire,
- traitement d'homogénéisation,
- traitement de durcissement,
- expansion.

Pour la bride en AZ5G, la masse du lingot est de 1 tonne et celle du produit fini après usinage 55 kg ; pour la bride en acier, les masses sont respectivement 3,8 T et 270 kg.

Les vues suivantes montrent quelques phases de réalisation de ces ébauches :

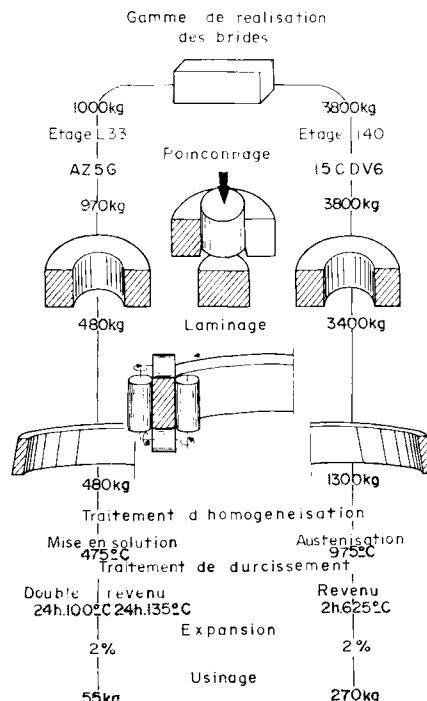


Fig. 7 - Brides acier et AZ5G

- lingot au manipulateur (figure 8)

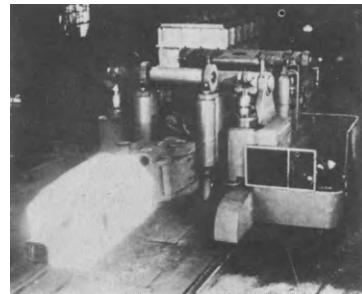


Fig. 8 - Laminage lingot au manipulateur

- laminage circulaire des
brides (figure 9)



Fig. 9 - Brides au laminage

- phases successives de laminage (figure 10)



Fig. 10 - Différentes phases de laminage

- préparation de l'expansion d'une bride en AZ5G (figure 11)

Fig. 11



- bride après expansion

Fig. 12



Fabrication des réservoirs

Réservoir 1er étage

Ariane.

Reservoir 1^e etage principe d'assemblage

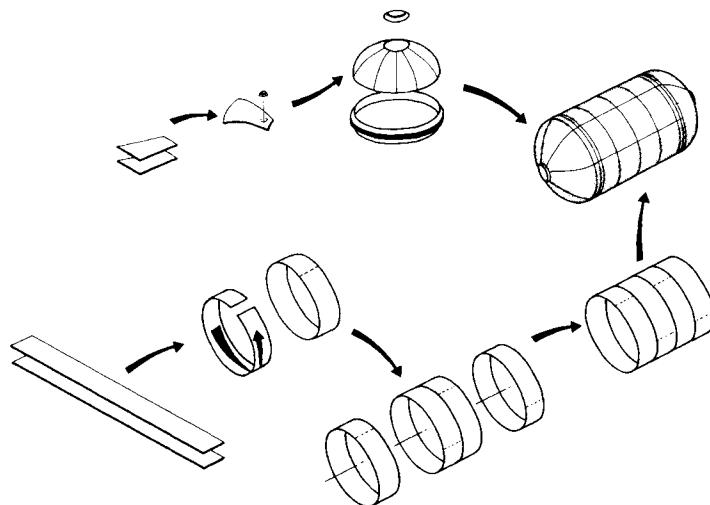


Fig. 13 - Principe d'assemblage du réservoir 1er étage

Les opérations successives d'assemblage du réservoir 1er étage sont schématisées sur la figure 13 :

- emboutissage de douze secteurs,
- soudage sur secteur des embases latérales,
- soudage des douze secteurs constituant la calotte,
- soudage sur la calotte de l'embase centrale et de la bride en Y,
- roulage puis soudage d'une virole,
- soudage successif de viroles aboutissant à un ensemble 4 viroles,
- soudage de l'ensemble 4 viroles avec un fond,
- fermeture du réservoir.

Les photographies suivantes montrent quelques procédés de fabrication utilisés :

- tournage d'une bride en Y Ø 3,800 m ; les résultats obtenus sont très satisfaisants. A titre d'exemple, l'ovalisation maximale obtenue est 0,7 mm, l'ovalisation moyenne étant inférieure à 0,5 mm (fig. 14).

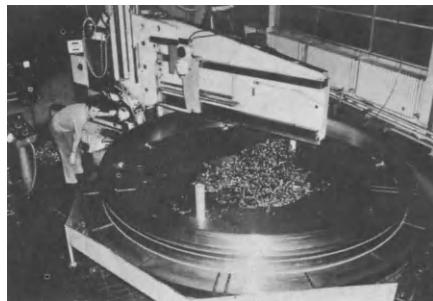
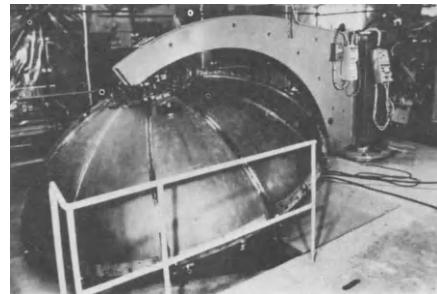


Fig. 14 - Tournage bride en Y 1er étage

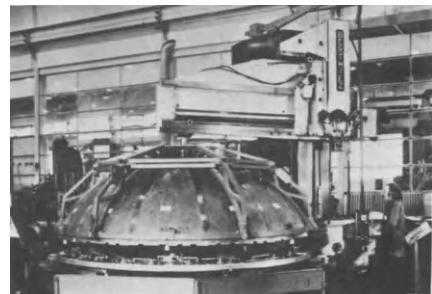
- soudage des secteurs de fond ; la tête de soudage est mobile et les éléments à assemblage fixes ; l'installation comporte également une tête de fraisage destinée à dresser les faces avant soudage (figure 15)

Fig. 15 Soudage secteurs de fond



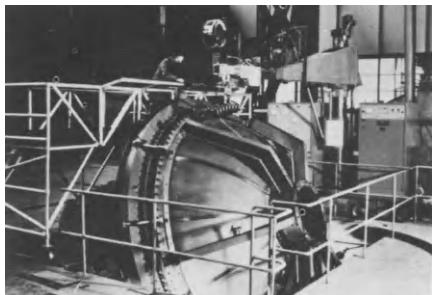
- tournage de la calotte fond (figure 16)

Fig. 16 Tournage d'une calotte de fond



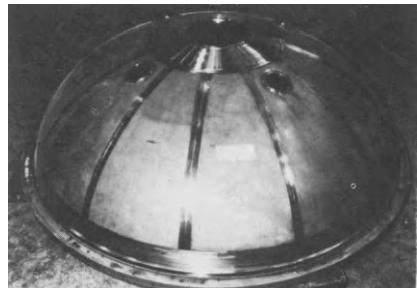
- soudage calotte-bride en Y ; la tête de soudage est fixe et les éléments à assembler tournent. Cette installation sert également au soudage de l'embase centrale sur le fond (Fig. 17).

Fig. 17 Soudage
bride en
Y calotte



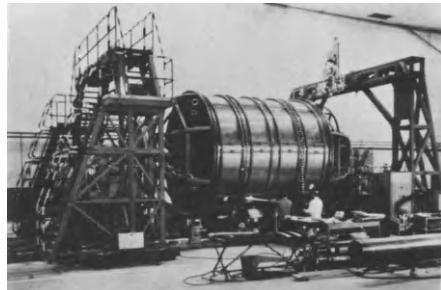
- exemplaire d'un fond terminé comportant l'embase centrale et les embases latérales (figure 18).

Fig. 18 Ensemble
fond
soudé



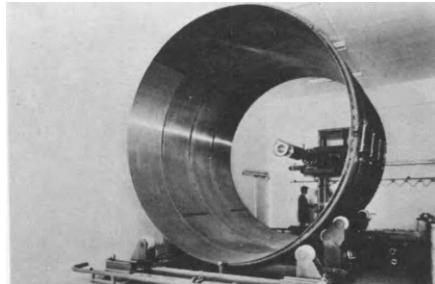
- soudage ensemble 2 viroles ; la tête est fixe et l'ensemble de la structure tourne. Pour la réalisation du cordon de fermeture du réservoir, une cerce de soudage démontable est utilisée (figure 19).

Fig. 19 Soudage ensemble 2 viroles



- contrôle aux rayons X après soudage (un contrôle radio complet des cordons de soudure est également effectué après embûrage des réservoirs (figure 20)).

Fig. 20 Contrôle radio ensemble 4 viroles



Réservoir 2ème étage

Les opérations successives d'assemblage du réservoir 2ème étage sont schématisées sur la figure 21.

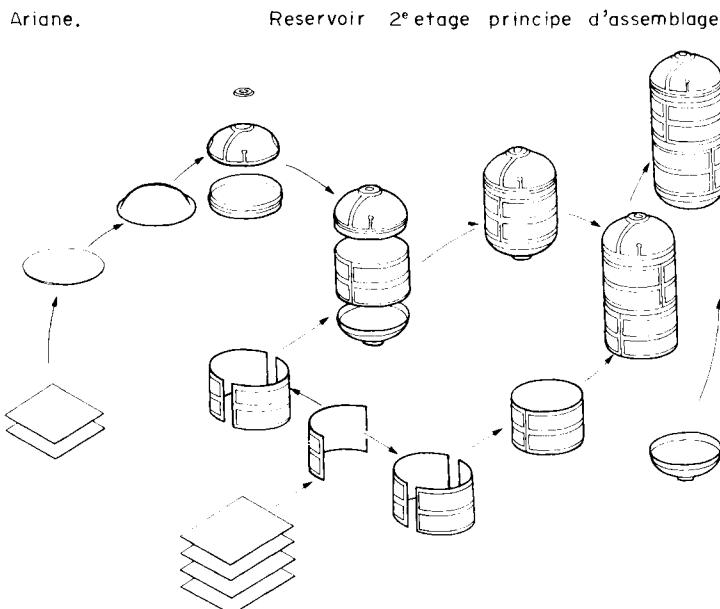


Fig. 21 - Principe d'assemblage du réservoir 2ème étage

- formage des calottes des fonds avant, arrière et intermédiaire par repoussage, à partir d'une tôle unique pour chaque fond,
- fraisage chimique des calottes,
- soudage des brides en Y et embases sur calottes,
- usinage chimique, roulage et soudage des viroles N et U,
- soudage fond avant sur virole N,
- soudage fond intermédiaire sur ensemble virole N - fond AV,
- soudage virole U sur ensemble ci-dessus,
- soudage fond arrière (fermeture du réservoir).

Parmi les procédés de fabrication utilisés, le formage du fond est le plus typique.

Le principe de repoussage est schématisé sur la figure 22. L'installation comprend :

Ariane. Principe de repoussage

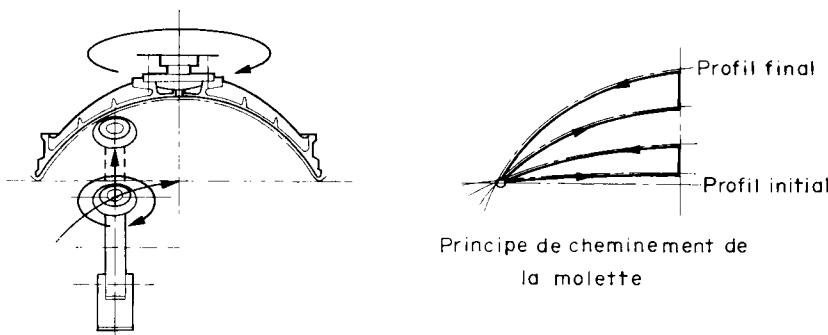


Fig. 22 - Principe de repoussage

- un tour à repousser équipé d'une molette et d'un dispositif de copiage,
- une matrice concave de profil identique à celui du fond fini.

Le mode opératoire est le suivant :

- la tôle plane est sertie sur le pourtour de la matrice,
- la matrice est mise en rotation,
- les déplacements de la molette au cours des différentes passes successives sont obtenus à l'aide d'un jeu de reproducteurs de courbure différente.

Les résultats obtenus par ce procédé sont très satisfaisants. Comme indiqué sur la figure 23, les caractéristiques du matériau ne sont pas altérées et sont comparables à celles des viroles et brides. D'autre part, l'évolution des épaisseurs le long d'une méridienne et la reproductibilité d'un fond à un autre sont correctes comme le montre la figure 24.

- la tolérance sur l'épaisseur d'un fond est inférieure à $\pm 0,3$ mm,
- les mesures relevées sur trois fonds différents montrent une

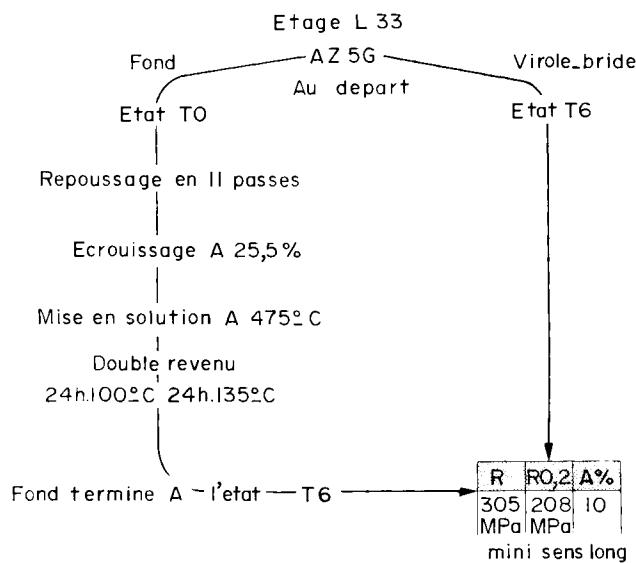


Fig. 23 - Caractéristiques du matériau après repoussage

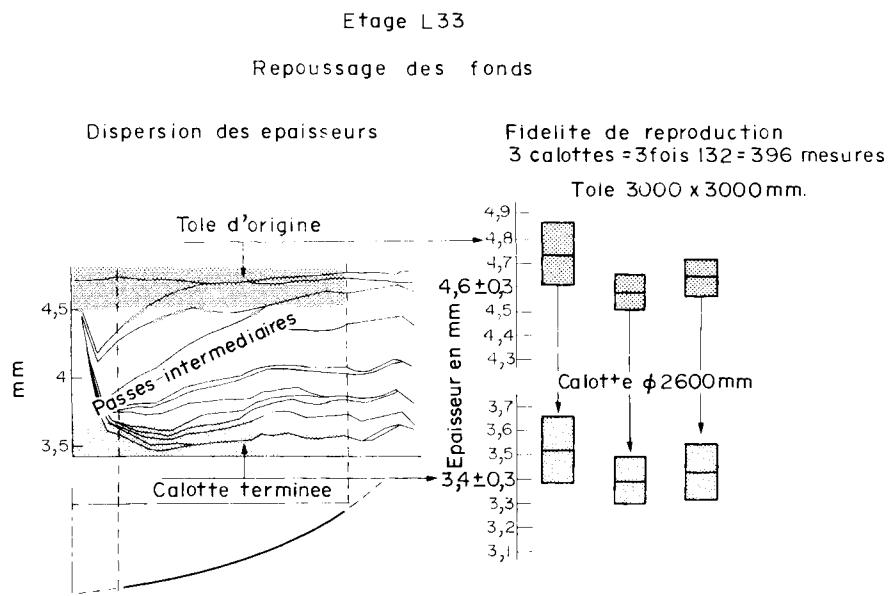


Fig. 24 - Evolution des épaisseurs après repoussage et reproductibilité

bonne reproductibilité du système de repoussage (les épaisseurs obtenues sont, bien entendu, dépendantes des épaisseurs initiales des tôles).

Les photographies suivantes montrent quelques autres procédés de fabrication utilisés :

- tournage d'une bride en Y Ø 2,600 m à l'AEROSPATIALE LES MUREAUX, sous-traitant de DORNIER pour cette réalisation. Les résultats obtenus sont satisfaisants (ovalisation inférieure à 0,4 mm) (figure 25).

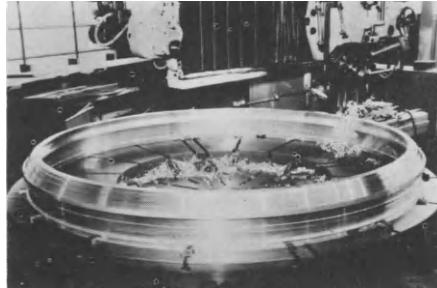


Fig. 25 - Tournage d'une bride en Y Ø 2,600 m

- soudage de la bride en Y et de l'embase centrale sur calotte. La tête de soudage est fixe et les éléments à assembler mobiles. Cette installation est également utilisée pour le soudage des embases latérales : pour cette opération, c'est la tête qui est mobile (figure 26).

Comme pour l'ensemble du réservoir, le fond est usiné chimiquement et toutes les soudures sont effectuées bord à bord (Fig. 27).

- soudage longitudinal de deux demi-viroles : la tête est mobile et les éléments à assembler fixes (figure 28).
- Soudage d'un fond sur virole : la tête est fixe et les éléments à assembler tournent (figure 29).

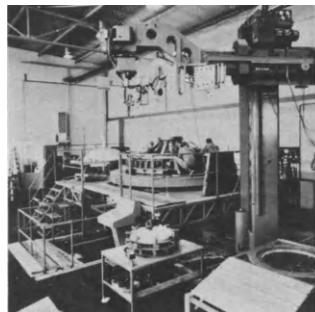


Fig. 26 - Soudage de la bride en Y sur calotte



Fig. 27 - Ensemble fond

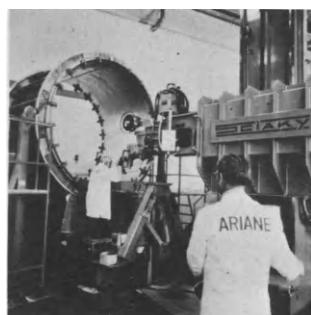


Fig. 28 - Soudage de deux demi-viroles

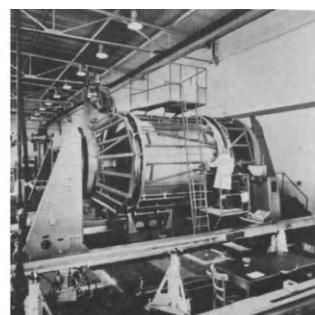


Fig. 29 - Soudage d'un fond sur virole

CONCLUSION

Comme souligné au début de cet exposé, le développement des structures présentées doit être considéré comme une application pratique des connaissances acquises jusqu'à présent par les industriels concernés. Pour ces derniers toutefois, ce programme n'est pas sans intérêt technique.

C'est la première fois en EUROPE que des structures de lanceur de près de 4 m de diamètre sont réalisées. Pour la fabrication des réservoirs du 1er étage, les principales difficultés proviennent de ces grandes dimensions, les technologies utilisées étant les mêmes que celles développées pour des structures ne dépassant guère 2 m de diamètre ; en particulier, les tolérances exigées pour les pièces de chaudronnerie sont pratiquement équivalentes à celles généralement spécifiées pour des pièces usinées. Le faible rapport épaisseur/rayon (1,9 mm/1900 mm) des éléments à assembler par soudage bord à bord a nécessité la réalisation d'outillages de soudage précis afin de limiter les dénivélés après soudage.

Pour le réservoir du deuxième étage, le principe de formage d'un fond par repoussage d'une tôle unique peut être considéré comme une innovation ; plusieurs fonds ont été réalisés avec succès suivant ce procédé. D'autre part, ce programme aura permis de parfaire la maîtrise de la technologie relative à l'AZ5G tant du point de vue formage que soudage. Ce matériau nous semble appelé à être largement utilisé dans les constructions aérospatiales, non seulement pour des structures pressurisées, mais également pour des structures dimensionnées au flambage ou à la rigidité, telles que des jupes ou des bâts-moteurs pour lesquels le paramètre le plus important est le module d'young, et non les résistances à rupture ou à limite élastique ; des études sont actuellement en cours pour l'introduction de ce matériau dans les normes AIR.

BIOLOGICAL SATELLITES AND THEIR CONTRIBUTION TO SPACE BIOLOGY AND MEDICINE

E. A. Ilyin

A biological satellite or a biosatellite as it is often abbreviated is one of the types of an automatic Earth satellite specially designed to carry out spacecraft-borne experiments with various representatives of plant and animal kingdoms as well as with isolated cells and tissues of plants and animals.

The first biological satellites were the Soviet second Earth artificial satellite that flew the dog Laika (1957) and Soviet recoverable spaceships-satellites (1960-1961) which flew dogs and many other living organisms. It is worth mentioning that these space vehicles were not termed biosatellites at the time of their flights.

The scientific literature began to use the term biosatellite during the period when the USSR and the USA developed research programs to be conducted onboard the space vehicles specifically designed for space biology experiments in near-Earth orbits (1962-1964). In the Soviet Union the term became firmly established after the flight of the Cosmos-110 biosatellite (1966).

One of the characteristic features of a biosatellite is the fact that the bioengineering equipment it carries as well as the flight itself, including the launch time and flight duration, are selected in such a way as to solve biochemical problems. Therefore, the term biosatellite cannot be applied to any automatic, i.e. unmanned, space vehicle which fly biological specimens.

At present it is not easy to analyze on a comparative basis the scientific and economic efficiency of biological experiments carried out in flights of manned spacecraft and biosatellites. However, the time of cosmonauts is obviously the most expensive thing in any manned space mission. This time should be planned and spent in a most precise manner; to date it must be primarily allowed for space navigation and economy-oriented investigations.

In other words, flights of manned spacecraft should be used, first of all, for the experiments which require the participation of cosmonauts. Only at that stage of the development of cosmonautics when physicians-researchers are crewmembers or when certain biochemical experiments are necessary from the practical point of view, plant and animal experiments in space flights may become as important as engineering and economic experiments and investigations.

Today, although biological experiments are very important for the further development of cosmonautics, we have to reconcile ourselves to the fact that plant and animal experiments will play a secondary role in space programs of manned flights during the immediate five years.

All this convinces us - the biomedical community - that in the near future it is an absolute necessity to fly biological experiments onboard biological satellites in order to study and to understand mechanisms of adaptation of living systems to prolonged weightlessness.

By now researchers of many countries, primarily of the Soviet Union and United States, have gained much experience in preparing and performing experiments onboard biological satellites.

Some information about the biological satellites is given in Table 1.

Results of space biology experiments, including biological satellites, were reviewed by O. G. Gazenko (1967), Jenkins (1968), Gazenko et al. (1974), Parfyonov and Lukin (1973), Parfyonov (1975). In addition, the investigations carried out on board the biological satellites of 1973-1974 were reported and discussed this May at the VIII Space Biology and Medicine Symposium in Varna, Bulgaria (Gazenko et al., 1975; Ilyin and Serova, 1975) and at the XVIII COSPAR Plenary Meeting in Varna, Bulgaria (Grigoriev et al., 1975; Ilyin et al., 1975).

Due to this, I should like to confine my presentation to some basic problems, for instance, the contribution which biological satellites have made and are making to the development of space biology and medicine.

In retrospect, looking back at the near past when mankind welcomed the sunrise of the space era, we have to emphasize the unique role which flight experiments with mammals and other biological specimens have played in the decision-making concerning the possibility and safety of manned space missions.

For instance, the results of the orbital flight of the dog Laika onboard the second artificial Earth satellite in 1957 have confirmed the opinions of Soviet scientists that highly organized living beings can tolerate a long exposure to weightlessness (Gazenko et al., 1962).

The purpose of the subsequent space flights of mammals onboard the recoverable spaceships-satellites of 1960-1961 was to verify and improve the equipment for operational medical monitoring of the physiological functions of cosmonauts and to test life support systems designed for man at every stage of the dog flight.

The results of these investigations and the evidence that cosmic radiation has no adverse biological effects have allowed the conclusion that manned flight into outer space is safe. From this point of view the period that preceded the first man-in-space flight can be considered as a stage of biological indication of the trajectories of orbital manned space missions which followed them.

This stage came to a close and space biology faced new tasks: study of the weightlessness effects on vital processes, investigations of mechanisms of adaptation of physiological systems of the animal and human body to prolonged weightlessness and of subsequent readaptation to Earth gravity, study of a combined effect of weightlessness and cosmic radiation, etc. In connection with lunar missions an additional problem arose - study of biological effects of protons and heavy ions of galactic cosmic radiation.

All these problems were subject to investigations in plant and animal experiments onboard biological satellites and in part onboard manned spacecraft.

Examinations of two dogs that made a 22-day flight onboard the *Cosmos-110* biological satellite (1966) revealed significant changes in their metabolism.

Table 1:
BIOLOGICAL SATELLITES

Biological Satellite	Year of Launch	Flight Time	Object of Study
Second Artificial Earth satellite	1957	USSR	The dog Laika
Second spaceship-satellite	1960	24 hr	Mammals (mongrel dogs, rats, mice, guinea-pigs), invertebrates, higher and lower plants, seeds, bacteria, viruses, biological preparations (cell cultures, enzymes, DNA), frog eggs
Third spaceship-satellite	1960	26.5 hr	
Fourth spaceship-satellite	1961	1.5 hr	
Fifth spaceship-satellite	1961	1.5 hr	
Cosmos-110	1966	22 days	Mammals (mongrel dogs), plants, seeds, bacteria
Cosmos-368	1970	6 days	Invertebrates, plants, seeds, bacteria, mammalian cell cultures
Cosmos-605	1973	21.5 days	Mammals (rats), reptiles (turtles), invertebrates, bacterial cells, lower plants
Cosmos-690	1974	20.5 days	Mammals (rats), reptiles (turtles), invertebrates, yeast, seeds, lower plants, bacterial spores
		USA	
Bios-1	1966	Abortive flight	Invertebrates, plants, seeds, bacteria, moulds, amoeba, frog eggs
Bios-2	1967	24 hr	Invertebrates, plants, seeds, bacteria, moulds, yeast, amoeba, frog eggs
Bios-3	1969	8.5 days	Primates (Macaca nemestrina)
OFO-1	1970	5 days	Amphibians (frogs)

They included appreciable weight losses, dehydration, muscle mass losses and decrease of bone density. Symptoms of cardiovascular deconditioning were also noted. These data showed the necessity of developing and using countermeasures against the unfavourable effects of prolonged weightlessness.

Real flights of Soviet cosmonauts and American astronauts gave evidence that these countermeasures were actually needed and that their use in long duration space flights was effective.

By physiological, biochemical and morphological methods specific and nonspecific changes in the functional systems of 45 rats flown onboard the Cosmos-605 (1973) were investigated. Special attention was given to the structural changes in the systems whose performance is to a greater or lesser extent related to Earth gravity.

These studies have shown that during 22-day weightlessness structural changes developed in skeletal muscles of limbs and in long tubular bones, i.e. in the systems that were incompletely loaded due to the weightless state. From the qualitative point of view the changes were similar to those found in hypokinetic animals. The difference was of a quantitative pattern.

Structural changes were also revealed in the hypothalamus-hypophyseal-adrenal system whose increased functional activity was responsible for the adjustment of the body to an unusual environment. Another proof of an activation of this system inflight was hypoplasia of the thymus, spleen and lymph nodes and reduction of the lymphocyte count in the peripheral blood and bone marrow observed postflight.

The rat experiment carried out onboard the Cosmos-605 biological satellite allows the basic conclusions that prolonged weightlessness gives rise to no disorders which should be considered pathological. The structural and functional changes detected in some systems were reversible and disappeared, as a rule, 25 days postflight.

The practical importance of these data for space medicine is obvious.

I should like to discuss radiological experiments carried out onboard the Bios-2 and Cosmos-690 biosatellites. In these experiments a great number of various organisms were irradiated while weightless. This was achieved with the aid of gamma-sources strontium-85 and cesium-137 flown onboard the Bios-2 and Cosmos-690, respectively. Obviously, such experiments cannot be performed onboard manned spacecraft because of the high risk of irradiation. Therefore, experiments flown onboard the biological satellites were the only possibility to study a combined effect of weightlessness and ionizing radiation - the problem which is very important from both the theoretical and practical point of view.

Brief results of this experiment flown onboard the Cosmos-690 biological satellite can be summarized as follows. Weightlessness slightly increased animal radiosensitivity, thus aggravating the course of radiation damage inflight and extending the recovery period postflight. It should be noted that the difference between the flight results and ground controls in the biological satellite mock-up was less significant than it was expected. This means that, taking into account inaccuracies of the existing methods of study and measurement as well as individual variations in the radiosensitivity of higher animals,

including human beings, the detected difference can be ignored from the practical point of view. With respect to the calculation and determination of permissible and critical doses of irradiation of crewmembers in space flights and computation of the thickness of spacecraft radiation shielding, it is necessary to proceed from the data of ground-based investigations and observations. These data need no correction as applied to space flight conditions.

The results of medical examinations of Soviet cosmonauts and American astronauts who have performed prolonged flights onboard the Salyut and Skylab orbital stations suggest an optimistic view on a further extension of manned space missions. However, in order to provide an adequate medical support of these flights it is necessary to continue plant and animal experiments in real flights. The main purpose of these experiments should be further study of mechanisms of adaptation of living beings to the effects of long-term space flight factors and testing of different countermeasures against their adverse influences.

From this point of view biological satellites can make important contributions to the development of cosmonautics. At present we are engaged in preparations for the experiments to be flown onboard a biological satellite of the Cosmos series.

The characteristic feature of this biological satellite will be a centrifuge to investigate biological effects of artificial gravity as a potential countermeasure against unfavourable effects of weightlessness. At this stage the main purpose of the study will be biological equivalence of Earth gravity of 1 g generated in space flight.

Figure I gives a drawing of the centrifuge with biocontainers and shows that the centrifuge consists of two parts - stationary and rotatory sections.

Biocontainers are mounted on the rotatory platform in the area where 1 g (12 containers) and 0.6 g (4 containers) will be generated throughout the entire flight. Identical biological specimens will be stationary mounted to be subjected to weightlessness per se and will be kept on the ground as controls.

Table 2 lists the experiments to be carried out on the rotatory and stationary platforms of the centrifuge. It can be seen that the biology program is characterized by its international realization.

This biological satellite will also fly 25 white rats. The purpose of the rat experiment is to study the resistance and structural-functional reserves of the mammalian body in the weightless state of 22 days and during readaptation to Earth gravity.

The rat experiment will be carried out using a life support of the type that was flown onboard the Cosmos-605 and Cosmos-690 biological satellites. Figure 2 gives a general view of a life support system unit.

During a month before flight rats will be kept under clinical and physiological surveillance and trained to live in isolated small cages. In addition, five rats will have implanted body temperature sensors and will be injected with 2-C¹⁴-glycine (to label red blood cells), declomycin (to label bone tissue) and Listeria monocytogenes antigen (to examine cell immunity).

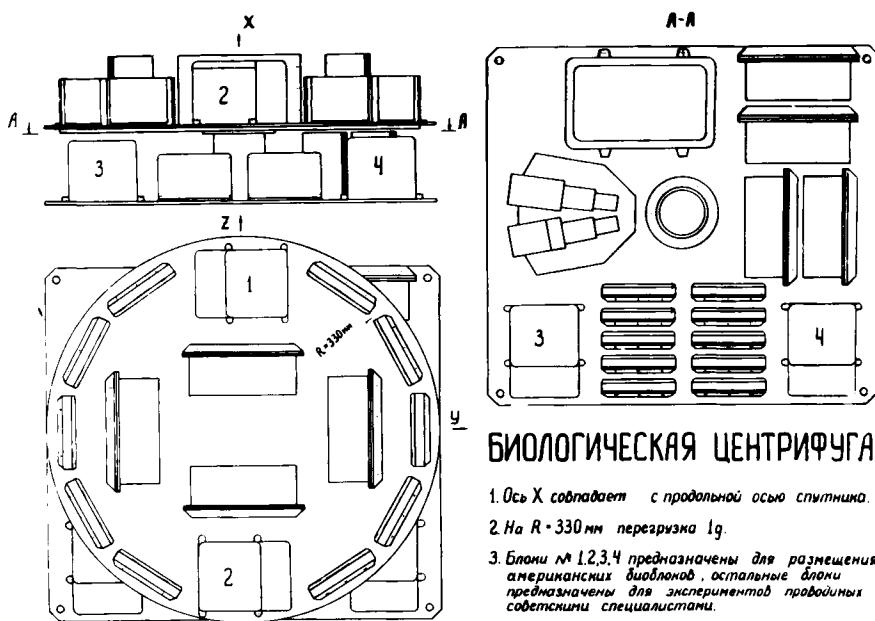


Fig. 1. The biological centrifuge

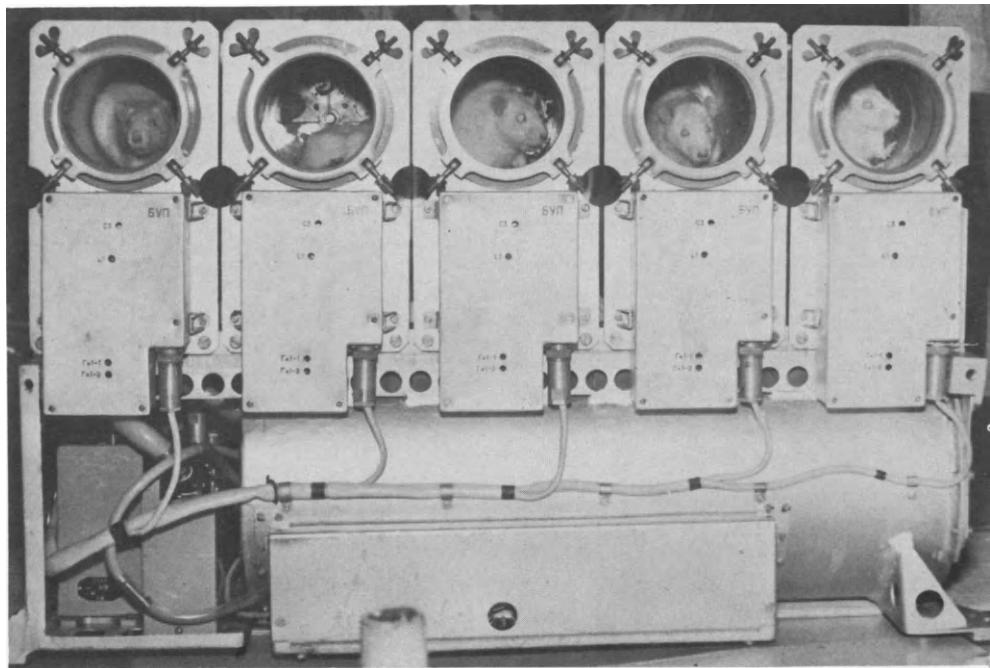


Fig. 2. A life support system unit for five rats.

Table 2:

BIOLOGICAL EXPERIMENTS TO BE FLOWN ONBOARD THE SOVIET BIOLOGICAL SATELLITE IN 1975

Biological Experiment	Purpose of Study	Investigators
1. Bacterial cells	Proliferation, mutagenesis and interaction of bacterial chromosomes and episomes	USSR
2. Tradescantia	Growth processes, chromosomes and genome mutations	USSR
3. Fungi	Morphogenesis	USSR
4. Seeds of different higher plants	Genetic growth processes and effect of sensitizing and protective substances on the processes Growth processes of higher plants in weightlessness as related to the initial orientation of seeds toward the nutrient substrate	USSR
5. Guppies	Embryonic development	USSR
6. Amphibians	Development of otoliths	USSR
7. Turtles	Effect of altered gravity (weightlessness and 0.6 g) on the structure and function of the vital systems	USSR
8. Mammalian tissue culture	Cell morphology and physiology	USSR
9. Drosophila melanogaster	Frequency of mutations and morphoses, effectiveness of mutagens and antimutagens, life cycle and quantitative indices of ageing	USSR and USA
10. Bioblock	Radiation effect of heavy ions on biological specimens (yeast, Arabidopsis seeds, <i>Crepis capillaris</i> seeds, lettuce seeds, etc.)	USSR and France
11. Cell tissue culture	Embryogenesis of higher plants during cultivation of somatic cells	USA and USSR
12. Plant tissue tumor	Biochemical and physiological study of bacterial tumors of higher plants	USA and USSR
13. Fundulus heteroclitus	General development and development of the vestibular organ	USA and USSR

During flight motor activity, food procuring reflexes and body temperature of the animals will be studied.

The main scope of investigations with the use of physiological, morphological, histochemical and biochemical methods will be done on the Earth postflight. The flight animals will be, therefore, subdivided into two groups.

The first group of the animals (12 rats) will be sacrificed on the R + 0 day, during the first 6 hours postflight, if possible. These rats will be used to study structural and functional changes that could develop under the influence of prolonged weightlessness.

The second group of the animals (13 rats) will be kept to investigate their readaptation to Earth gravity with the aid of different provocative tests. These rats will be sacrificed on the R + 25 day.

Scientists from Czechoslovakia, Poland, Hungary and the USA will cooperate in the postflight examinations of rats.

Thus, biological satellites are very important for biology and medicine from another point of view. They make it possible to integrate efforts of researchers from different countries with a view of resolving major scientific problems. Similar to the joint Apollo-Soyuz test mission, international cooperation of scientists in the field of space biology will promote peaceful coexistence and lasting peace on the Earth and will contribute to the advancement of our understanding of ecophysiological foundations of the life of various living beings, man including.

References

1. Gazeiko O. G., Space Biology. In: Development of Biology in the USSR. Nauka, 1967, 613-631.
2. Gazeiko O. G., Ilyin E. A., Parfyonov G. P. Biological investigations in space (results and perspectives), Izvestia AN SSSR, ser. biol., 1974, 4, 461-475.
3. Gazeiko O. G., Adamovich B. A., Grigoriev Yu.G., Druzhinin Yu.P., Ilyin E. A., Verigo V. V., Popov V. I. Results of radiobiological experiments on rats exposed onboard the Cosmos-690 biosatellite. VIII Space Biology and Medicine Symposium, Varna, 19-25 May, 1975, 27-29.
4. Grigoriev Yu. G., Ilyin E. A., Druzhinin Yu. P., Serova L. V., Popov V. I., Noskin A. D., Kuzin R. A., Kondratyev Yu. I. Study of radiosensitivity of mammals during prolonged weightlessness. COSPAR Eighteenth Plenary Meeting, 29 May-7 June, 1975. Varna, Bulgaria, Program Abstracts, 323-324.
5. Ilyin, E. A., Serova L. V., Study of physiological effects of prolonged weightlessness on the animal body. VIII Space Biology and Medicine Symposium, Varna, 19-25 May, 1975, 48-49.
6. Ilyin E. A., Serova L. V., Kondratyev Yu. I., Portugalov V. V., Tigranyan R. A., et al. The influence of 22-day weightlessness on functional and structural parameters of rats. COSPAR Eighteenth Plenary Meeting. 29 May - 7 June, 1975. Varna, Bulgaria. Program Abstracts, p. 340.
7. Jenkins D. W. USSR and US Bioscience. Bioscience, 1968, 18, 543-549.

8. Parfenov G. P., Lukin, A. A. Results and prospects of microbiological studies in outer space. Space Life Sciences, 1973, 4, 160-179.
9. Parfenov G. P. Biological indication of new space trajectories. Foundations of Space Biology and Medicine, 1975. Moscow, Nauka, Vol. II, Book 2, Pt. V, Ch. 19, 306-337.

ATMOSPHERE REVITALIZATION FOR SPACELAB

Herbert Eckert*, Alois Christian Gall* and Hamid Hassan**

**Dornier System GmbH, Friedrichshafen, Germany*

***European Space Research and Technology Centre (ESTEC) Noordwijk, Netherlands*

INTRODUCTION

Spacelab, shown installed in the Orbiter cargo bay in Figure 1, represents one of the major Space Shuttle payloads. Its flexible modular concept permits several different flight configurations, indicated in Figure 2, to satisfy the broad spectrum of planned experiments.

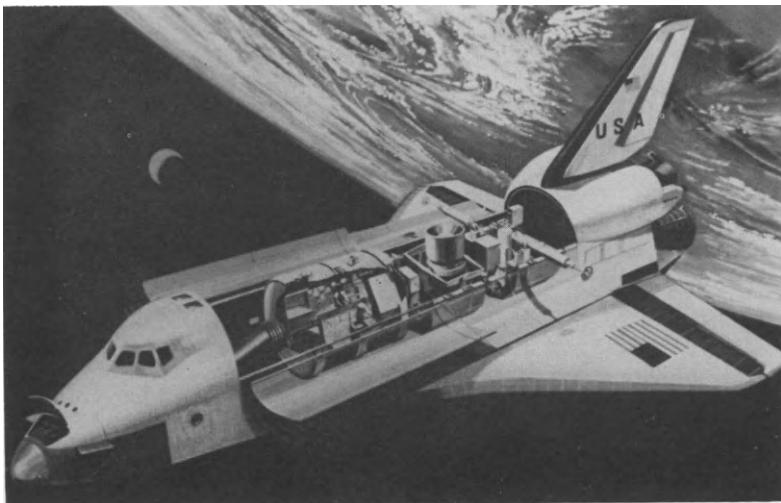


Fig. 1 ORBITER/SPACELAB configuration

Basically the Spacelab consists of two major elements, the Module and the Pallet. The Pallet element is used for experiments conducted directly in space; the Module provides the capability for scientists to perform experiments in a zero g condition under earth-like atmospheric conditions. The Shuttle Orbiter and the Spacelab Module are connected by a tunnel for transfer of scientists between the vehicles. One or two Module segments can be combined to form either a Short Module or a Long Module configuration. The dimensions of these two versions of Spacelab are shown on Figure 3.

The Environmental Control Subsystem maintains the environment in the Spacelab Module where the scientists will perform their experiments. This subsystem is controlled automatically and requires minimum training of the scientists for operation.

It consists of the following functional sections:

- Environmental Control and Life Support (ECS/ECLS)
 - o Atmosphere Storage and Control
 - o Atmosphere Revitalization
- Thermal Control (ECS/TC)
 - o Active Thermal Control
 - o Passive Thermal Control

This paper concentrates on the Atmosphere Revitalization Section (ARS) and its interfaces with the vehicle and other sections of the ECS.

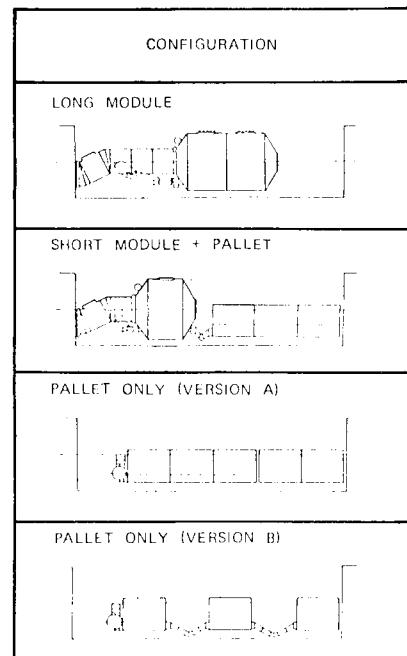


Fig. 2 SPACELAB base-line design configurations

FUNCTION

The ARS maintains a shirtsleeve conditioned environment for the crew members and a temperature controlled environment for avionics and experiments. In detail, the following functions are provided:

- Control of cabin atmosphere temperature
- Limitation of cabin atmosphere humidity

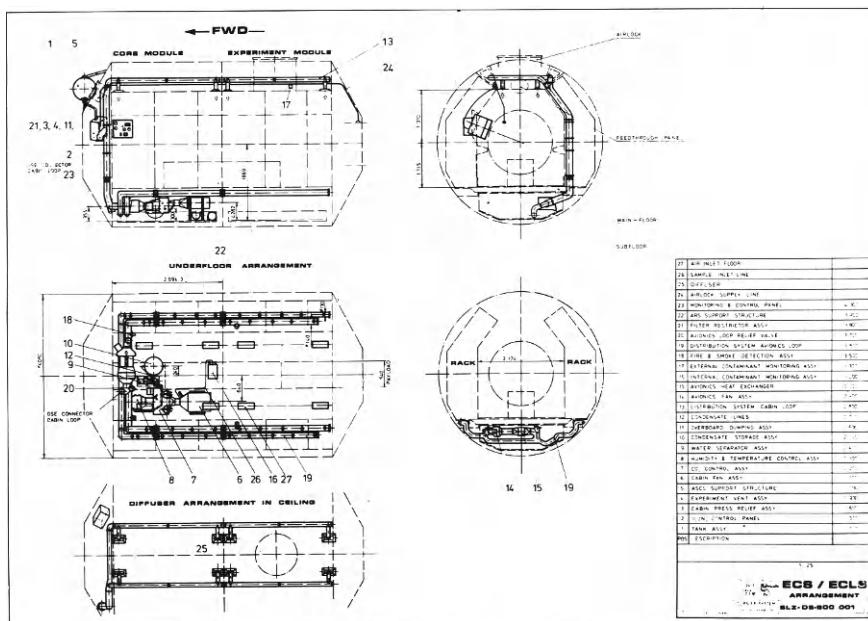


Fig. 3 ECS/ECLS arrangement

- Filtration of air supplied to cabin and electronic equipment
- Control of cabin atmosphere carbon dioxide partial pressure
- Removal of odors and trace contaminants
- Storage of condensate water, produced by condensing and separating latent heat loads
- Overboard dumping of condensate in contingency cases
- Cabin air distribution to maintain specified air velocities in the crew occupied areas
- Air cooling of electronic equipment
- Monitoring of contaminants internal and external the module
- Detection of fire and smoke
- Prevention of excessive pressure differentials between cabin and electronic equipment compartments.

The ARS provides these functions independent of the Shuttle Orbiter with the following exceptions:

- A water loop, which is part of the ECS/TC, is the heat sink of the ARS. It interfaces with the Orbiter payload heat exchanger where the heat loads produced in Spacelab

are transferred to the Orbiter freon loop and from there, via the radiator, rejected into space.

Electrical power, required for system operation, is supplied by the Orbiter via the Spacelab Electrical Power Distribution Subsystem.

DESIGN REQUIREMENTS

The major design requirements of the ARS are shown in Figure 4. In addition, the ARS is designed to meet the Spacelab life time of ten (10) years or 50 seven day missions. During a 7 day mission, no scheduled maintenance is required. The Spacelab ARS is also capable of supporting missions up to 30 days duration without principle changes; mainly, by the addition of expendables.

Ambient total pressure	1.013 bar	
Temperatures		
Atmosphere (range)	18 - 27 °C	
Coolant supply (max.)	7.3 °C	
Relative humidity	between 6 °C DP (min.) and 70% RH (max.)	
Heat loads (nominal)		
Metabolic	total	164 W/man
	sensible	103 W/man
	latent	61 W/man
Carbon dioxide removal	total	35 W/man
	sensible	23 W/man
	latent	12 W/man
Electrical power (nominal)	7000 W	
Electrical power (peak)	12 000 W	
Payload specialist station	300 W	
ORBITER heat rejection capability for SPACELAB	8500 W	
Workstation air velocity	5 - 12 m/min.	
Air filtration		
Cabin loop	5 micron nominal	
Avionics loop	280 micron nominal	
Carbon dioxide		
Generation (nominal)	990 g/man day	
Partial pressure (nominal) (max.)	5 mm Hg 7.6 mm Hg	

Fig. 4 ARS design conditions and requirements

SYSTEM DESCRIPTION

The baseline ARS configuration is shown on Figure 5. Basically, it is divided into two air loops, the cabin air loop and the avionics air loop. The separation of the two loops was selected from the standpoint of thermodynamics and contamination control.

Two additional major elements of the ARS are the external contamination assembly, which monitors contaminant deposition on externally mounted experiments, and the fire and smoke detection assembly consisting of sensors located in the cabin and the avionics air loops.

The major equipment of the ARS is located under the main cabin floor of the Spacelab, where the assemblies are mounted to Spacelab structural panels. The ARS arrangement is shown in Figure 3.

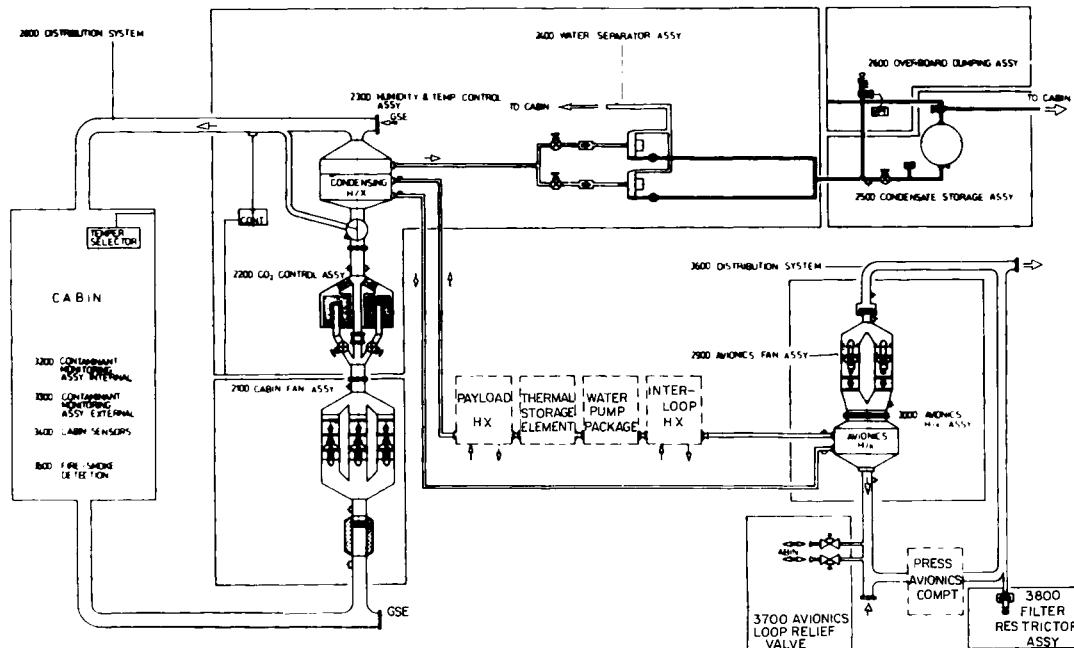


Fig. 5 ARS configuration

Cabin Air Loop

The cabin air loop consists of the following major assemblies:

- Cabin Fan Assembly
- CO₂ Control Assembly
- Humidity and Temperature Control Assembly
- Water Separator Assembly
- Condensate Storage Assembly
- Overboard Dumping Assembly

Cabin air distribution and ventilation is provided by ducting and air outlet diffusers located in each module segment in overhead structures. Figure 3 shows the ducting and diffuser arrangement. Air velocities within the habitable area of the cabin are kept within the specified range of 5 to 12 m/min to satisfy crew comfort requirements. Air delivered from the cabin air loop ventilation is supplemented by air induced in the diffusers and by secondary air patterns created within the cabin itself. The diffusers are adjustable to alter the air flow pattern.

Cabin air, which is maintained at a total pressure of 1.013 bar by the Atmosphere Storage and Control Section (ASCS) returns to the cabin air loop system from the cabin underfloor through a 5 micron wire mesh filter. The filter element, mounted on the inlet of the cabin fan assembly, is constructed of stainless steel to avoid corrosion and flammability problems. The filter is easily accessible and removable for changing or cleaning.

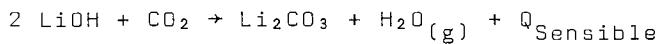
Downstream of the filter, redundant axial fans, operating with 115/200 VAC, 3 phase power are provided. Single fan operation is required for delivering the required flow rate. The remaining fans are for backup in the event of a failure of the primary fan. To avoid electromagnetic interference problems, the fans use electromagnetic induction motors rather than slip rings. The speed of the fans is approximately 10 000 RPM. Noise generated by high speed equipment is a major design consideration in manned space vehicles. The present Spacelab cabin fan has a noise criteria level of approximately NC 69, which represents a low value compared to formerly used fans. However, the total Spacelab system noise level design goal is NC 50. Therefore, during the progress of the Spacelab development, noise reduction techniques may be introduced.

Check valves are provided in the cabin fan assembly downstream the fans to avoid short circuiting of air by recirculation through inactive fans. The check valves have a low pressure drop in the open position to minimize the fan power consumption. However efficient sealing is required in the closed position to prohibit air recirculation within the assembly which would influence total system performance.

A differential pressure sensor is provided in the cabin fan assembly to monitor fan performance. In the event of a fan failure, a warning indication is provided on the Spacelab ECS/ECLS moni-

toring and control panel. In such a case, a redundant fan will be manually switched on. As a result, no repair action for a failed fan is needed in space.

Downstream of the cabin fan assembly, the CO₂ control assembly is provided. A bellow connects the two assemblies to minimize transport of fan generated noise via ducting structure and to avoid the rigid structural interface which could lead to deformations due to thermal or mechanical loads. The air delivered to the CO₂ control assembly is divided into 3 branches within the assembly. A fixed restrictor is provided in the air main stream to adjust the by-pass flow so that 50 kg/h goes through each of the two LiOH-cartridges simultaneously. The cartridges, which contain LiOH, KMnO₄ and activated charcoal, are mainly used to remove metabolic generated CO₂ by a chemical reaction in accordance with the chemical equation shown below. The activated charcoal and KMnO₄ are contained in the cartridges to provide odor and trace contaminants removal.



where

H₂O_(g) represents the produced latent heat.

The Spacelab cabin CO₂ partial pressure will normally be maintained below 5 mm Hg. The maximum allowable CO₂ partial pressure is 7.6 mm Hg. To obtain a high utilization of the cartridges, an alternate changeout procedure is foreseen; that is a fresh cartridge on-line to maintain CO₂ partial pressure control, while the depleted cartridge remains on-line in parallel until approximately 95 to 97 percent of its total capacity is utilized. The cartridges used in Spacelab are identical and replaceable with those used in the Shuttle Orbiter. Their normal capacity is 2 man days, so that, for a 7 day mission with a 52 man hours per day requirement, a total of 8 cartridges is needed.

The CO₂ control assembly also contains an air temperature control valve which controls cabin temperature by bypassing air around the condensing heat exchanger assembly which is packaged separately directly downstream. Air flow through the heat exchanger is modulated according to signals from a controller which responds to three sensor inputs: actual cabin temperature, selected cabin temperature and air temperature (anticipator) downstream of the condensing heat exchanger.

The temperature control valve has two actuators, each with synchronous gear motor drives. In case of a failure of any component involved in the control circuit, manual override capability is also provided.

One of the major items within the ARS is the condensing heat exchanger. It is a four pass cross counterflow design with plate/

fin construction. It is fabricated with stainless steel to withstand corrosion and provide long life capability.

The performance of the condensing heat exchanger is based on the design conditions as shown in the thermal balance in Figure 6.

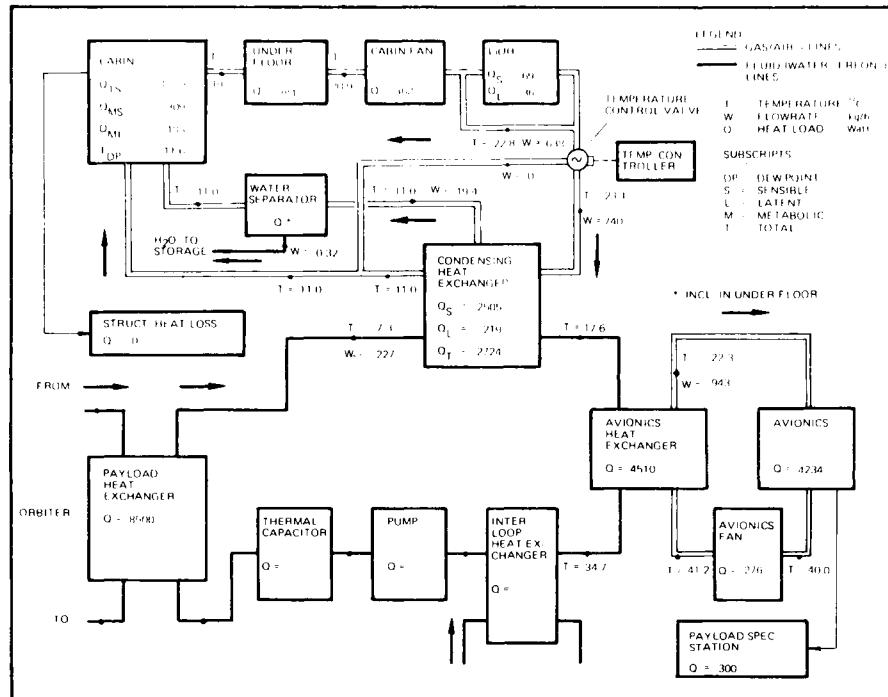


Fig. 6 ARS thermal balance

For the available water coolant flow rate of 227 kg/h and a coolant supply temperature of 7.3°C , the required air-flow for the given heat load distribution is 740 kg/h. As shown on Figure 7, the relative humidity is maintained within the specified range for all predicted operating conditions including cabin temperature and heat load variations. Sensible heat transfer mainly takes place in the forward section of the heat exchanger while most condensation occurs in the rear section.

Condensed moisture in the heat exchanger adheres to its internal surfaces by surface tension and is removed by means of an integral "slurper" device, located at the outlet end of the heat exchanger. The slurper consists of a series of holes provided at the end of each heat exchanger air passage. A suction flow of approximately 16 kg/h air draws the mixture of water and air from the condensing heat exchanger to two parallel mounted motor driven rotary water separators. Reliability estimates led to the selection of two separators. Only one separator operates at a time; the remaining one is for backup.

The water is separated from the air by centrifugal effect and delivered into a condensate storage tank. Fluid check valves are installed downstream of the water separator to provide sufficient back pressure to avoid air inclusion and to prevent reverse flow from the liquid storage tank. Speed sensors are provided on each separator for failure indication and isolation. Each separator also has an air check valve mounted at the inlet to prevent recirculation through the non operating separator so that all air leaving the separator returns to the cabin.

Separated condensate is stored in a bladder type tank with the cabin atmosphere total pressure as reference pressure. Manual valves are provided to isolate the tank in case of a bladder failure which can be indicated by the quantity measuring device. The capacity of the tank is sufficient for containing all generated condensate during a baseline 7 day mission.

A contingency capability to dump water overboard is provided for long duration missions or in case of excessive generation of condensate for the baseline mission. The operation of the overboard dump assembly basically involves manual switch actuators. This is acceptable because the crew is available for system operation for contingency situations.

Operation is initiated by switching on the heater on the dump nozzle. After it reaches a specified temperature, the dump solenoid valve is manually switched on and water is discharged to space using the cabin atmosphere pressure on the condensate tank bladder as the driving force. A fluid line pressure sensor and the tank quantity sensor indicate when the tank is empty and the operation will then be terminated by switching off the power supply.

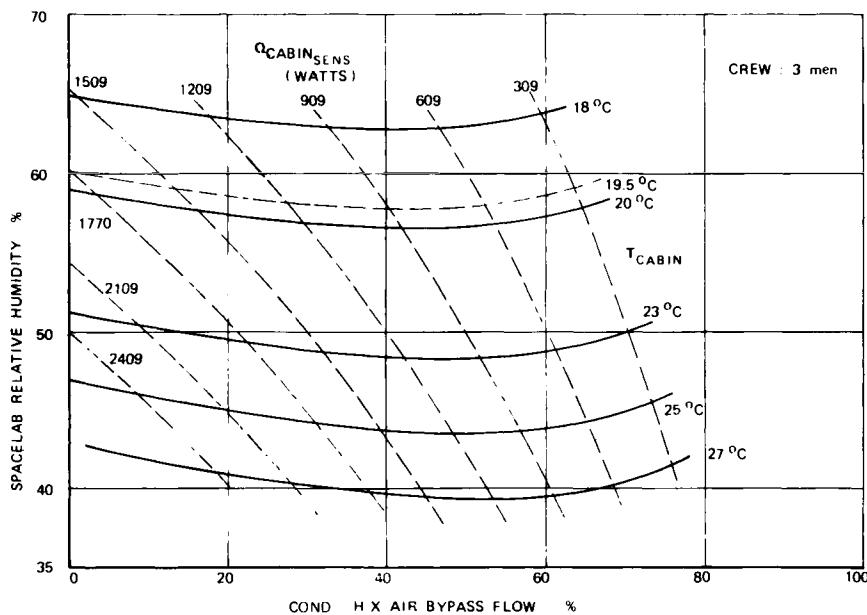


Fig. 7 Cabin temperature control performance

Avionics Air Loop

Avionics and experiments producing heat and requiring cooling are located in compartmentized racks which are isolated from the cabin air loop. Cooling is provided by forced air distributed to the racks. The avionics air loop consists of the following assemblies:

- Avionics Fan Assembly
- Avionics Heat Exchanger Assembly
- Distribution System (Avionics Loop)
- Filter Restrictor Assembly
- Avionics Loop Relief Valve Assembly

The avionics fan assembly is installed below the main subfloor and provides 943 kg/hr airflow to satisfy cooling requirements. Two redundant fans are arranged in parallel in a similar manner as the cabin fans. A 280 micron filter is provided at the fan inlet to protect the fan and the electronic equipment. Similar to the cabin fan filter, the filter is easily accessible for cleaning and changing. This is not normally required during a 7 day mission. A fan delta pressure indicator is used for failure indication and fault isolation. In addition, the avionics air outlet temperature is measured. Check valves are provided downstream the fans to avoid recirculation through the inactive fan which could produce performance reduction.

The avionics heat exchanger assembly has a direct interface to the fan assembly. The heat exchanger, using the interfacing ECS/TC water coolant loop as heat sink, removes sensible heat from the avionics loop air. The avionics heat exchanger mechanical design is similar to the condensing heat exchanger design except that no slurper is included as there is only sensible heat transfer. Temperature sensors are also provided for monitoring equipment performance.

Downstream of the avionics heat exchanger the cooled air is distributed to the equipment racks through a distribution system. Air entering the racks from the supply duct cools the equipment and is then drawn through the distribution duct to the air return duct. The distribution system ensures that sufficient cooled air circulation is provided near the front surface of the racks so that the transfer of heat, generated by the electronic equipment, to the cabin is minimized and surface touch temperature requirements are satisfied.

A slight negative pressure is maintained in the avionics loop to ensure that contaminants, which may be produced by the electronic equipment, are prevented from entering the cabin. It is accomplished by a relief valve between the avionics loop and the cabin and a fixed restrictor filter assembly between the avionics loop and space. The relief valve is full open at a pressure differential in either direction, ensuring flow from the cabin to the avionics loop under normal conditions and flow from the avionics

loop to the cabin for contingencies where the cabin may be depressurized. The restrictor filter assembly is sized to provide a vent flow of approximately 1.35 kg/day.

Cabin Sensors

Two cabin temperature sensors are located within the module cabin to give representative information about the cabin atmosphere temperature independently from the sensors required for system operation. Atmosphere total pressure and oxygen partial pressure sensors are also provided as part of the ASCS.

The ARS also contains a mass spectrometer which is used in the Spacelab cabin as an internal contamination monitor. Specifically, it has the capability to measure hydrogen, water vapor, nitrogen, oxygen, methane, carbon dioxide and any other gas with a mass to charge (M/C) ratio between 50 and 120.

PROPULSION SYSTEM CONCEPTS FOR SINGLE-STAGE SHUTTLES

Rudi Beichel

Aerojet Liquid Rocket Company, Sacramento, California

INTRODUCTION

Groups of scientists throughout the world have been earnestly studying how space can be used to solve the serious problems on earth concerning dwindling resources. One such enthusiastic and capable group met at Princeton University in May 1975 to reveal their progress and to determine a means to accomplish what they feel must be done to ensure the survival of man (1). Their concepts appeared to be technically sound, but the large space traffic requirements mean that more emphasis should be placed on space transportation economics. The only foreseeable propulsion system to meet this mission is exclusively chemical, is single-stage-to-orbit (SSTO), and utilizes the mixed-mode propulsion concept (Fig. 1). This propulsion system must be non-polluting and the propellant resources must be available in abundance. The propellants oxygen, hydrogen and hydrocarbon meet this specification.

MIXED-MODE

The propulsion concept involves combining in the same vehicle stage two propulsion modes operated sequentially (high-density propellants for liftoff and early ascent, and low-density propellants for final ascent to orbit) to produce greater stage performance than possible using either mode separately. The greater performance is achieved when the specific impulse for Mode 2 is greater than that for Mode 1, and the propellant density and density-impulse for Mode 1 exceed that for Mode 2.

MODE 1

Lift-off mode with engine burning LO₂/RJ-5 propellants

MODE 2

Ascent-into-orbit mode with engines burning LO₂/LH₂ propellants.

DUAL-FUEL

Application of the principle of density optimization to engine design leads to the concept of a high-pressure staged-combustion-cycle LO₂/RJ-5 rocket engine with a hydrogen-pumping-system attachment giving the engine the ability of also operating with LH₂ fuel.

PARALLEL BURN

The simplest trajectory for mixed-mode application involves a sequential burn—that is, startup and shutdown of Mode 1 engines followed by operation of the engines in Mode 2. It is also possible to have some overlap of Mode 1 and Mode 2 engine operations. Overlap from lift-off to orbit we call "parallel burn."

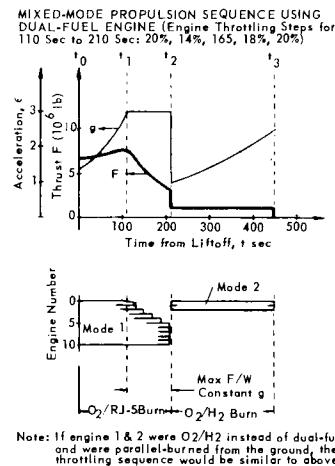


Fig. 1. Mixed-mode definition.

VEHICLE DESCRIPTION

There are two basic approaches to the SSTO space transportation system: a winged vertical takeoff and horizontal landing (VTOHL) vehicle and a vertical takeoff and vertical landing (VTOVL) vehicle.

The VTOHL vehicle shown in Fig. 2 is of conservative design and utilizes well established concepts. It has one of the largest base areas of all winged vehicles that have been considered. It consists of two main aluminum load-bearing LO_2/LH_2 tanks mounted atop a wing-and-carry-through structure containing RJ-5 fuel (2). The internal 15- by 90-ft cargo bay fits above the main tanks. The large base area of this configuration permits end loading and unloading of the cargo bay, circumventing heavy cargo bay doors (although doors could be included if needed to open up as much as 65 ft of the bay).

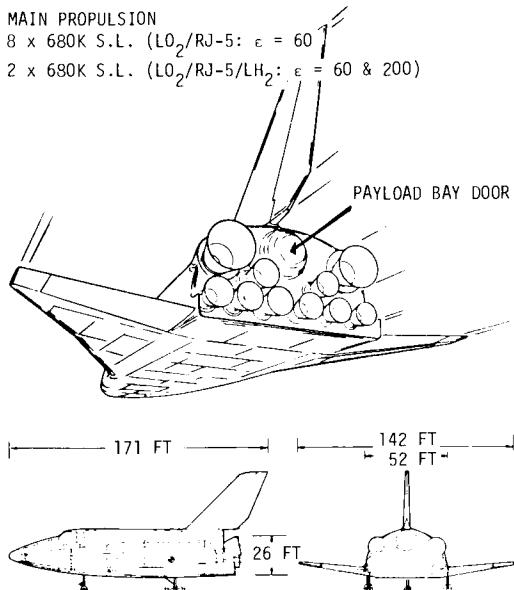


Fig. 2. Mixed-mode single-stage shuttle configuration after Salkeld (2).

Liftoff and first-mode propulsion come from eight high pressure ($P_c = 4000$ psia) $\text{LO}_2/\text{RJ-5$ engines and two dual-fuel engines burning $\text{LO}_2/\text{RJ-5}$, each with nozzle expansion area ratios of 60:1. Second-mode propulsion is from the two dual-fuel engines burning LO_2/LH_2 (at $P_c = 3000$ psia) and exhausting through an extendible nozzle of $\epsilon = 200:1$. The six engines aligned slightly above the wing planar are each separated by a distance of approximately two feet to allow for gimbaling. Propulsion area utilization is about 63% of the base area, excluding that allocated to the cargo bay.

It is seen that the bell cluster configuration of Fig. 2 limits the nozzle expansion ratio of the engines to about 60:1 (except in the case of the dual-fuel engines). The payload of the VTOHL vehicle into 100 NM orbit is given in Fig. 3 (3). Should larger payloads be required and the SSTO vehicle size be increased substantially over that shown, the area available for propulsion becomes less favorable because the vehicle volume (weight) increases as a cubed dimension, whereas the base area only is a squared dimension. Higher pressure engines would be able to achieve a higher ϵ (and thrust level) in the same base area, but lower pressure engines would not be able to meet the

performance and the vehicle thrust requirement in this base area. Better base area utilization might be possible, however, with other nozzle types.

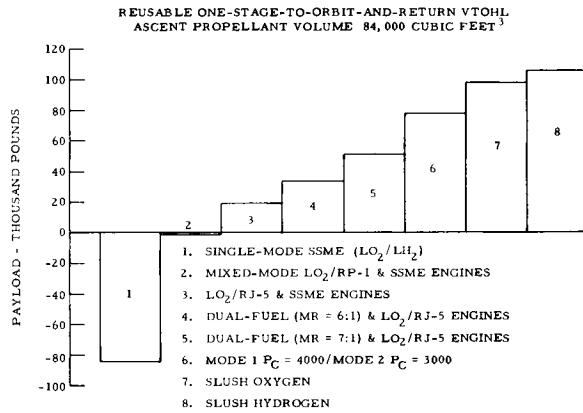


Fig. 3. Payload improvements through propulsion system optimization.

A problem common to flyback shuttle-type vehicles is the rearward shift of the center-of-gravity during ascent. Empty of propellant, the SSTO vehicle of Fig. 2 has its c.g. in the middle of the aft tank just above the landing gear. This makes the engine weight, which is sensitive to high area ratio nozzle design, very important. An excessive engine weight and resultant rearward c.g. shift, for example, can result in poor aerodynamic characteristics during flyback. Nozzle designs, therefore, should be optimized to achieve better integration, lighter weight, better lift and control capability, and improved altitude compensation.

For very large payloads into orbit, a single-stage VTOVL vehicle may provide a more economical approach. A concept proposed by Salkeld (4) uses mixed-mode propulsion to reduce the size and dry weight of the vehicle. The vertical landing feature was well established with the Apollo moon landing and with VTOVL aircraft. Compared to LO_2/LH_2 single-mode concepts, the mixed-mode approach makes possible more effective volume utilization, and cheaper tankage (since only 23% of the total propellant volume is LH_2 , compared to 73% in the single-mode system). The vehicle shown in Fig. 4 delivers 400,000 lb into 200 NM/28.5° orbit, utilizing 18 720K (sea level) $\text{LO}_2/\text{RJ-5}$ engines and 4 620K (vacuum) $\text{LO}_2/\text{RJ-5}/\text{LH}_2$ dual-fuel engines. Both bell and super-elliptical nozzles are shown in the figure clustered around a truncated plug nozzle giving an effective expansion area ratio of about 200:1. Since the engines are gimbaled, optimum sea level performance is achieved, with altitude compensation being provided by the plug nozzle when the engines are gimbaled toward the plug. Since the plug nozzle also serves as the re-entry heat shield, the super-elliptical nozzles are more shielded inside the shock than the bell configuration.

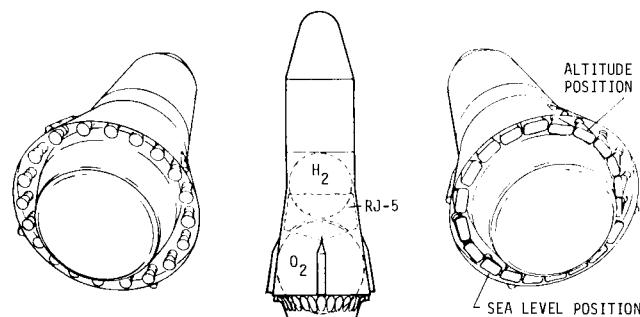


Fig. 4. Salkeld SSTO VTOVL vehicle for large payloads.

PHILOSOPHY OF INCREMENTAL IMPROVEMENTS

Through continued search and application of accumulative incremental improvements, an economical space transportation system will be achieved. The application (3) of this philosophy is dramatically illustrated in Fig. 3, where each somewhat marginal improvement in system density results in a significant improvement in the payload of a single-stage-to-orbit vehicle, and the sum total of these improvements in the accomplishment of a mission thought impractical only a short time ago.

As the figure shows, a reusable one-stage-to-orbit-and-return VTOHL vehicle with an ascent propellant volume of 84,000 cubic feet, corresponding to a propellant volume of 90,000 cubic feet for the 2-1/2 stage NASA baseline space shuttle, will not be able to achieve a 100 nautical mile orbit. On paper, the measure of this non-achievement is a negative payload of 84,000 lb. Application of the mixed-mode principle (substitution of a portion of the tankage and engines to handle RP-1 fuel instead of LH₂, and sequential burn of the propellants in high pressure LO₂/RP-1 and in SSME LO₂/LH₂ engines) is seen to provide a payload gain of 83,000 lb, but still a negative payload of 1000 lb. Replacing the RP-1 fuel with the heavier hydrocarbon RJ-5 and utilization of the mixed-mode principle leads to a delivered payload to orbit of 19,000 lb (a gain of 20,000 lb over RP-1, and a gain of 103,000 lb over use of SSME propulsion only). Further optimization of the propulsion density of the system is seen to result in a payload in excess of 106,000 lb, demonstrating the marked effect of density on the performance of a single-stage vehicle.

EFFECT OF PRESSURE ON PERFORMANCE AND DESIGN

Some of the confusion in the literature is the result of the conclusions derived from the analysis of low pressure engines being extrapolated to the design of high pressure engines. Therefore, prior to discussing concepts for the SSTO vehicle, I will show how the pressure influences design criteria and make clear the importance of high chamber pressure for this mission.

Performance

To illustrate the effect of chamber pressure on both sea level and vacuum performance, the range of pressures are shown from 500 psia up to 20,000 psia

in Fig. 5. Note that a 1000 psia engine operating at a nozzle expansion area ratio of 11:1 delivers 296.5 seconds (theoretical maximum) at sea level and 326.0 seconds in vacuum, the sea level value being 91% of the vacuum performance. A 4000 psia engine, on the other hand, produces 333.2 seconds ($\epsilon = 29:1$ optimum) at sea level and 352.8 seconds in vacuum, the sea level being 94% that of the vacuum value. The high pressure engine delivers closer to vacuum performance at low altitude and its sea level performance is 7.2 seconds higher than the vacuum performance of the low pressure engine. A low pressure engine with altitude compensating nozzle, therefore, can never compete (as implied in the literature) with a high pressure engine with any kind of reasonable nozzle design.

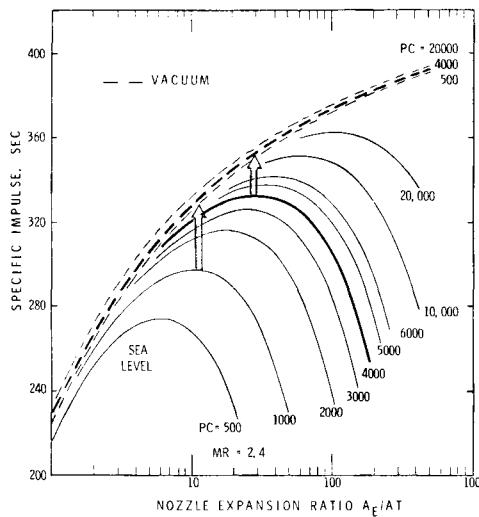


Fig. 5. $\text{LO}_2/\text{RJ-5}$ propellant performance.

The trend in performance continues as chamber pressure is increased, with a 10,000 psia engine generating 96% of its theoretical vacuum performance at sea level. Whether or not a 10,000 psia engine can be considered as practical at this time is not important. I am only trying to indicate the benefits that could accrue if higher pressures can be realized. But there are indications that the pressure limit within today's state-of-the-art is well in excess of 4000 psia.

Pressure has a major influence on nozzle design as shown in Fig. 6, where a comparison is made between engines operating at 1000, 4000 and 10,000 psia, using the highest performance that a given nozzle can deliver as represented by one-dimensional isentropic flow with chemical equilibrium. To avoid any confusion that could result from the use of parametric data (C_T vs. P_c/P_a), I have used specific impulse and altitude (U.S. Standard Atmosphere, 1966). Note that the performance for the higher area ratio ($\epsilon = 200:1$) nozzle is less than that for the lower area ratio ($\epsilon = 40:1$) nozzle below altitudes of about 35,000 ft and 66,000 ft, respectively, for chamber pressures of 4000 and 1000 psia. Also note the points where nozzle separation occurs; that is, separation of the gas flow along the nozzle wall when the exit pressure P_e is less than 0.3 (or 0.4) of the ambient (altitude) pressure P_a .

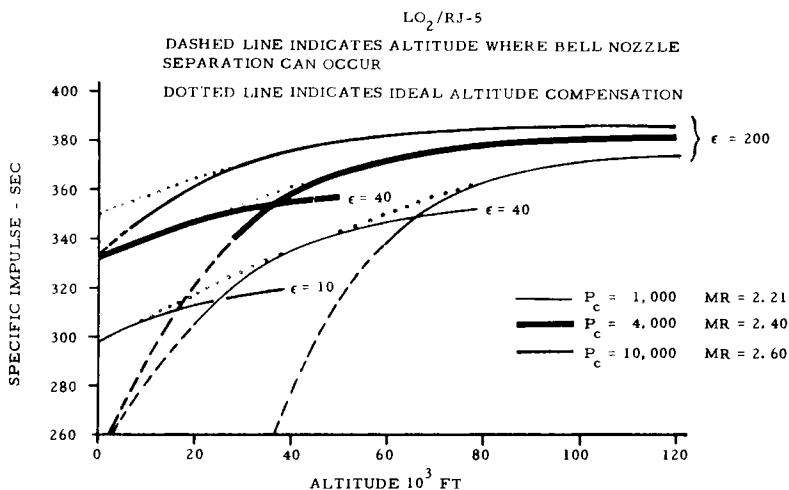


Fig. 6. Effect of chamber pressure on theoretical nozzle performance.

The performance of an ideal altitude compensating nozzle of design area ratio 200:1 is also indicated by the dotted lines in the figure. The ideal altitude compensating nozzle obeys as though it were a 'rubber' nozzle able to adjust its expansion area to the optimum condition for each altitude. For a bell nozzle to approximate this performance, it must have multi-position capability. It would have a minimum of three-position, two-position and one-position for a $P_c = 1000$, a $P_c = 4000$ and a $P_c = 10,000$ engine, respectively, as indicated in Fig. 6 and noted in Fig. 7. The multi-position capability can be accomplished in a number of ways, for example, as has been demonstrated with extendible nozzle types (5,6), and with gimbaled bell nozzles on a plug (7).

Engine Size

A typical sea level nozzle length is indicated in Fig. 7 by dashed lines and the lower value of ϵ . Each engine delivers approximately 680,000 lb thrust at sea level using the sea level nozzle length. The important thing to note is the large difference in nozzle diameter and length, and to consider this fact in conjunction with the packaging of the engines in the SSTO vehicle shown in Fig. 2.

Payload

The real benefit of any propulsion system is based on its effect on payload. The variation in SSTO payload for the vehicle in Fig. 2 versus chamber pressure, therefore, is of major importance. Two parametric curves are given in Fig. 8, one showing the variation in payload versus mode 1 chamber pressure, with mode 2 chamber pressure being held constant at 3000 psia, and the other showing this same variation with the mode 2 chamber pressure being 1000 psia less than that for mode 1 (8). The payloads are those achievable in polar orbit using LO₂/RJ-5 and LO₂/LH₂ propellants for mode 1 and mode 2, respectively. If

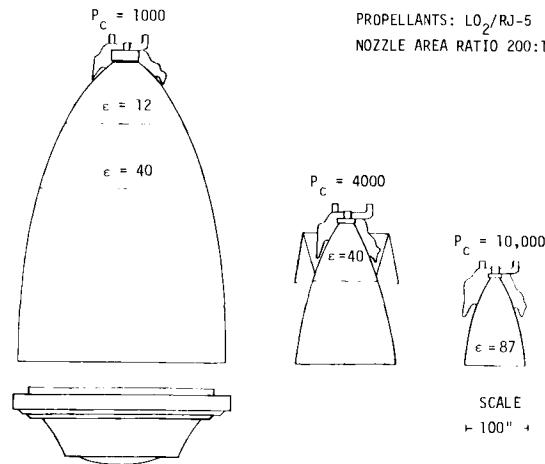


Fig. 7. Pressure effect on bell and annular spike engine size.

normal melting point (slush) oxygen were used instead of normal boiling point (liquid) oxygen, the payload would be 75,000 lb instead of 40,000 lb for the $P_c = 4000$ case. If a due east launch were utilized, the payload would be 80,000 lb using LO_2 .

Both curves must be considered somewhat academic at this point because simplifying assumptions (same propellant and vehicle weight for each case) were made, and no consideration was given to the feasibility of operating the rocket engines at the extreme pressures. The curves show important trends, however, that should be valid. These are that low pressure engines are not capable of delivering payloads for an SSTO mission, and that from 30 to 48% payload gain can be achieved by increasing mode 1 pressure from 4000 (the nominal engine design point for previous studies) to 6000 psia. It has already been shown that sufficient base area is not available in the vehicle of Fig. 2 to achieve the payloads listed for the low pressure engine.

Staged Combustion Cycle

The basic rocket propulsion considerations for a flyback vehicle involve the vehicle base area limitations and the engine thrust level. Because the vehicle acts as a vertical liftoff rocket during ascent to orbit and as an airplane during descent, the vehicle poses conflicting aerodynamic requirements on propulsion system packaging. A smaller base area is required for flyback compared with that allowed for liftoff. Since the thrust level of a rocket-propelled vehicle is a function of the gross liftoff weight as well as the vehicle base area, the smaller base area requirement for flyback must be accommodated by the propulsion system without a reduction in thrust level. In order to achieve the same rocket engine performance in the required vehicle base area, performance density optimization is required. This means the utilization of the highest chamber pressure possible, giving the maximum sea level specific impulse and thrust (maximum expansion ratio) in the allowable area, and in addition, the selection of a rocket engine cycle with the highest degree of

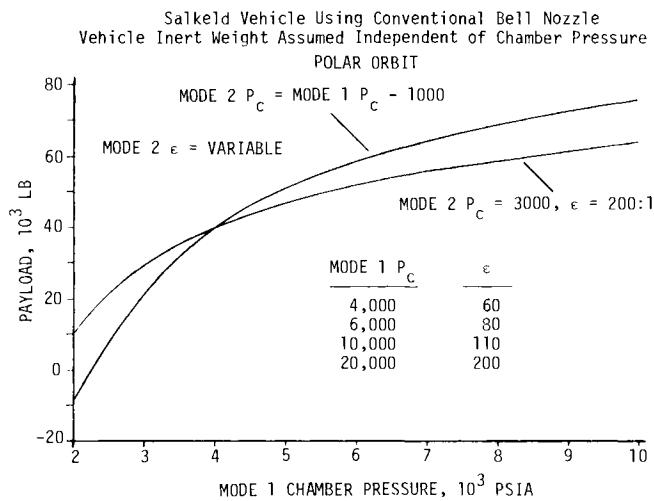


Fig. 8. Effect of chamber pressure on single-stage-orbit payload (8).

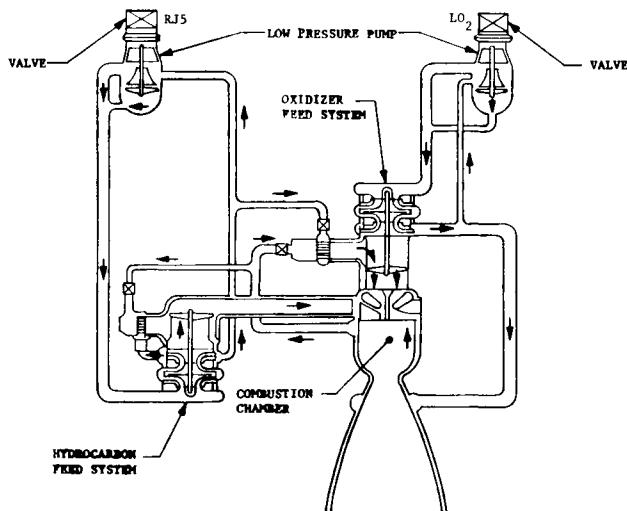
commonality (of components) and the maximum use of dense propellants. Further requirements of the propulsion system, relating to low costs, are a long operational life with minimum maintenance, with these requirements being achievable through maximum reuse of components and by maintaining low structure temperatures to reduce metal fatigue.

The production of high chamber pressure requires high horsepower input to the turbine. This requirement led to the development of the staged combustion cycle, which makes efficient use of the residual energy in the partially burned turbine drive gases. Instead of dumping these gases overboard, as is done in the low pressure gas generator cycle, they are combined at optimum mixture ratio with the remaining propellant in a second combustion chamber. The large pressure ratios that are achieved permit the use of nozzles with very large expansion ratios which, in turn, provide increased thrust. The only engine of this category presently under development in the U.S.A. is the Space Shuttle Main Engine (SSME).

Two other rocket engines have been under investigation to meet the requirements for single-stage vehicles. These are a high pressure $\text{LO}_2/\text{RJ-5}$ engine (Fig. 9 and Table 1) and a $\text{LO}_2/\text{RJ-5}$ engine with capability of also operating with LH_2 fuel (Fig. 10 and Table 2) (3,9). The high pressure $\text{LO}_2/\text{RJ-5}$ engine appears to be an absolute requirement for an effective mixed-mode shuttle. The dual-fuel conversion would give the highest possible vehicle performance, but an interim alternative clearly exists to use already-developed SSME's for mode 2.

$\text{LO}_2/\text{RJ-5}$ Engine

The basic features of the $\text{LO}_2/\text{RJ-5}$ engine are shown in the schematic. This version of the staged combustion cycle achieves efficient high pressure operation utilizing both an oxidizer-rich and a fuel-rich preburner and turbine, thus eliminating the need for an interpropellant seal common in most of the

Fig. 9. High pressure LO₂/RJ-5 engine schematic.TABLE 1 High Pressure LO₂/RJ-5 Engine

<u>Parameters</u>	
Thrust Sea Level, lbf	680,000
Thrust Vacuum, lbf	735,300
Number of Engines	1
Engine Cycle	Staged Combustion
Propellants, O/F	LO ₂ /RJ-5
Mixture Ratio, O/F	2.40
Chamber Pressure, psia	4,000
Area Throat, in ²	97.64
Area Ratio, A _E /A _T	40
I _{sp} Sea Level, lbf·sec lbm	319.6
I _{sp} Vacuum, lbf·sec lbm	345.6
Propellant Flow Rate, lbf/sec	2,127.7
Qualified Duration, hours	7.5 (Design)
Qualified Life, hours	7.5 (Design)
NPSH Oxidizer (ft)	171
NPSH Fuel (ft)	42.5
Engine Dry Weight, lb	5,335

engines in use today. Another benefit of having this combination of preburners is the low preburner and turbine gas temperatures required to achieve an engine power balance. The low temperatures satisfy the requirement of long life operation and result in lighter weight, uncooled turbopump assemblies and injector manifolding. A further benefit derived from this cycle arrangement is a gas-gas injector with both gases being very near the spontaneous ignition temperature for the mixture. It is well known that efficient combustion occurs under such reaction conditions, and that shorter combustion chambers are required. Because both propellants enter at approximately the same temperature, the structural stress on the injector are reduced, enhancing the operational life of this component.

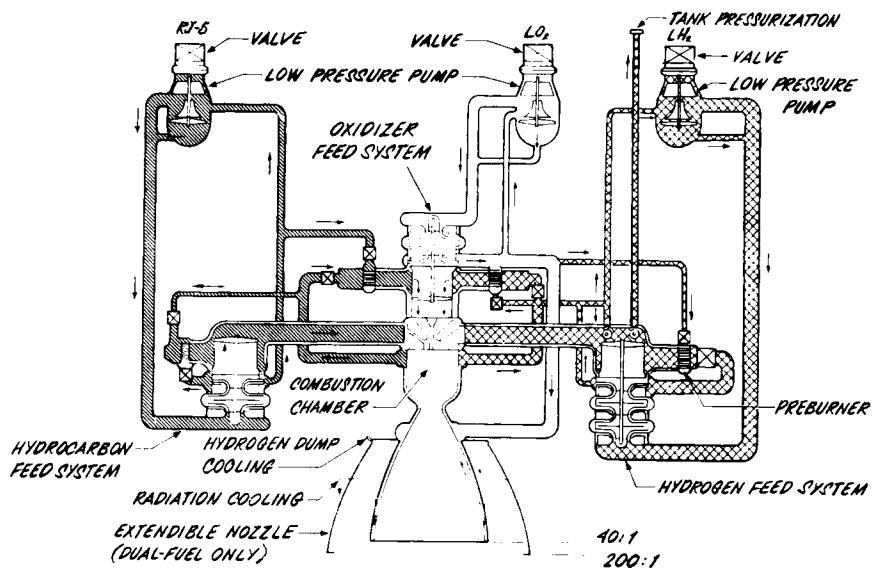


Fig. 10. Dual-fuel engine schematic.

TABLE 2 Dual-Fuel Engine

<u>Parameters</u>	
Thrust Sea Level, lbf	680,000 / -
Thrust Vacuum, lbf	735,300 / 588,390
Number of Engines	1
Engine Cycle	Staged Combustion
Propellants, O/F	LO ₂ /RJ-5 or LH ₂
Mixture Ratio, O/F	2.40 / 7.0
Chamber Pressure, psia	4,000 / 3,000
Area Throat, in ²	97.64
Area Ratio, A _E /A _T	40/200
I _{sp} Sea Level, lbf-sec/lbm	319.6 /
I _{sp} Vacuum, lbf-sec/lbm	345.6 / 461.5
Propellant Flow Rate, lbi/sec	2,127.7(SL) / 1,275(Vac)
Qualified Duration, hours	7.5 (Design)
Qualified Life, hours	7.5 (Design)
NPSH Oxidizer (ft)	171
NPSH Fuel (ft)	42.5 / 162
Engine Dry Weight, lb	8,920

Dual-Fuel Engine

The second engine under investigation is the dual-fuel conversion, which can be considered to be a LO₂/RJ-5 engine with a hydrogen feed system attachment kit and extendible nozzle. The main components of the engine--the combustion chamber, main injector, gimbal mount, thrust structure, controls, and the oxidizer feed system--are common to both modes of operation of the engine. Both fuel feed systems are self-contained, and have independent oxidizer-rich preburners to operate the oxidizer turbopump system. The controls for the engine are, therefore, relatively simple. As was the case with the LO₂/RJ-5 design, each preburner operates at approximately the same low temperature, thus eliminating the requirement for turbopump housing and injector manifold cooling.

Operational features of this engine worth noting are the change in chamber pressure from 4000 psia during mode 1 to 3000 psia during mode 2, the change in expansion ratio (through use of the extendible/retractable nozzle) from 40 to 200 to obtain near optimum performance at both sea level and altitude (mode 1 and mode 2), and the use of a LO₂/LH₂ mixture ratio of 7 for mode 2 operation. The change in chamber pressure aids the heat transfer situation for mode 2 operation, but more importantly, results in a mode 2 thrust level that is close to optimum for the single-stage-to-orbit mission.

The increase in the mixture ratio from 6:1, nominally utilized for LH₂-cooled LO₂/LH₂ engine operation, to 7:1 improves the heat transfer capability of liquid oxygen by making more of it available for cooling the engine. Since the ratio of the heat capacity of oxygen to hydrogen is about 1:9, the 7:1 quantity ratio of oxygen to hydrogen nearly compensates for its poorer cooling capability. The mixture ratio shift also reduces the hydrogen flow rate by 12%, improving the power balance of the mode 2 cycle, and reducing the hydrogen feed system weight slightly. An additional advantage is that it increases the bulk density of the mode 2 propellant combination sufficiently to improve the performance density of the system, despite the slight drop in propellant performance (specific impulse), and means a gain in payload of about 18,000 lb.

Regenerative Cooling

A plot of a typical heat transfer computer design analysis of the LO₂/RJ-5 thrust chamber is given in Fig. 11. It is seen that the peak flux in the nozzle throat region is 70.6 btu/in.²-sec and that the maximum wall temperature is between 930 and 1075°F. The band of wall temperatures is given because there presently is some question as to the best heat transfer correlation to be utilized in the calculations. Since the bulk temperature of the liquid oxygen passes through the critical temperature in the coolant jacket, heat transfer correlations that do not compensate for the large variation in fluid properties through the critical region give a high wall temperature, and conclusions are made that the chamber cannot be cooled. Limited experimental data on fluids such as oxygen and nitrogen at pressures of about 1000 psia in the critical temperature region, however, give correlations that lead to the lower band temperatures shown in the figure. Also, the utilization of a heat transfer calculation method that accounts for three-dimensional effects and the use of slush oxygen at the normal melting temperature both give even lower temperatures

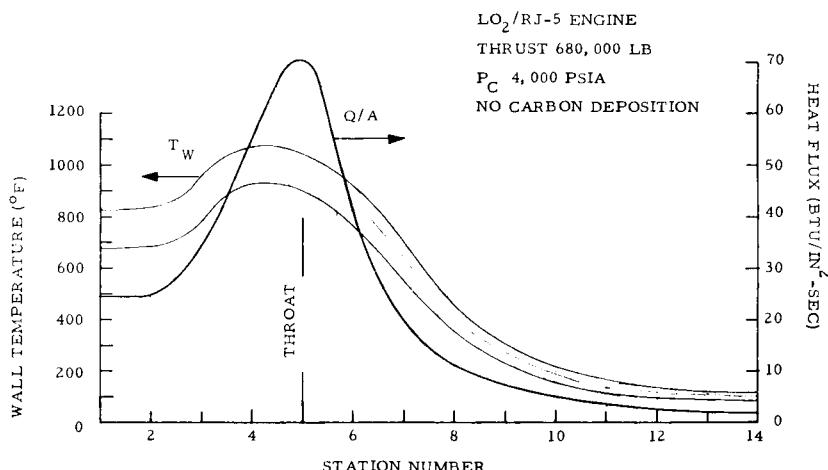


Fig. 11. Liquid oxygen regeneratively cooled engine.

than those indicated. Further, it is well known that the slightly fuel-rich mixture ratios, utilized to given optimum performance for LO₂/hydrocarbon burning engines, cause carbon to deposit on the walls of the combustion chamber and nozzle and form an effective insulating layer (10). Heat transfer calculations which utilize the experimentally determined carbon deposit gas-side thermal resistance result in throat wall temperatures of the order of 500°F. The data presented in the figure, therefore, are very conservative, and the conclusion can be made that the specified LO₂/RJ-5 engine can readily be regeneratively cooled with liquid oxygen.

The heat transfer conditions for mode 2 operation are depicted in Fig. 12 as a function of mixture ratio. The band of throat wall temperature is seen to range from as high as 1180° at a mixture ratio of 6:1 to a low of 850°F at a mixture ratio of 8:1 for liquid oxygen regenerative cooling. The upper bound of the curve refers to the heat transfer correlations obtained for fluids above the critical points, and the lower bound of the curve refers to more realistic correlations coupled with three-dimensional cooling effects and the use of slush oxygen. While the question concerning the capability of liquid oxygen to regeneratively cool the dual-fuel engine during mode 2 operation is not as certain as for mode 1 operation, it is believed that there will be no difficulty when the chamber contour and coolant channel designs are optimized. A factor that has not been considered in the heat transfer analysis is the good probability that the carbon film from mode 1 will remain on the chamber walls during mode 2 operation. Should this occur, the wall temperature in the throat region will be significantly reduced and long thrust chamber life will be assured.

Nozzle Concepts and Vehicle Integration (11)

Potential arrangements of several nozzle types are shown in the following figures utilizing SSTO vehicle of Fig. 2 as a base for comparison. The sea level thrust for all configurations is 6.8 million pounds, but the vacuum thrust varies with configuration due to the variation in nozzle area ratio.

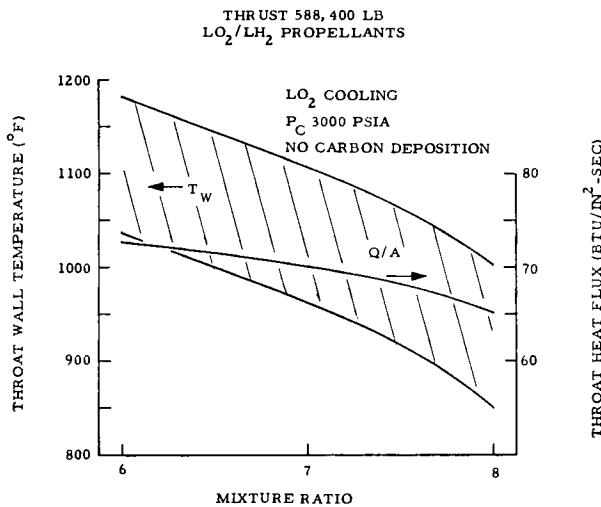


Fig. 12. Throat heat transfer conditions for mode 2 operation.

The chamber pressure for all bell nozzle engines is 4000 psia for mode 1 and 3000 for mode 2, and that for annular or linear throat engines is 2500 psia for mode 1 and 2000 for mode 2 (equivalent throat heat transfer). The internal expansion sections of the bell nozzles (except those of Fig. 2) have an area ratio of 40:1.

The nominal payload for the Fig. 2 vehicle is taken to be 40,000 pounds polar orbit as described previously (see Fig. 8). Since the SSTO vehicle payload depends so strongly on a complete vehicle design analysis, payload calculations for the vehicles with the new nozzle concepts were based upon simplifying assumptions--that the vehicle weight was assumed the same for all nozzle concepts and the nozzle performance was assumed the same for all configurations. This means that only potential gains or losses due to the nozzle contours are indicated.

The configuration with the best area utilization combines 20 bell nozzles in a linear forced-deflection nozzle, as shown in Fig. 13. The insert schematic of the figure shows the engines gimbaled away from the nozzle skirt to obtain optimum sea level performance. At the proper altitude (about 30,000 ft) the nozzles can be gimbaled such that the skirt contour provides an overall expansion ratio of 200:1 to obtain the maximum vacuum performance as indicated in the figure insert. This configuration thus improves the area ratio utilization from that in Fig. 2 by providing all engines with two-position high area ratio nozzles, and a resultant payload of 55,000 pounds (an increase of 38%). The four centrally located engines are dual-fuel engines operating on LO₂/LH₂ propellants during mode 2 operation. The space around the nozzles can be effectively used for propellant storage, as shown in the figure insert.

Also shown as an insert schematic in Fig. 13 is the potential use of this nozzle to achieve air augmentation during sea level and low altitude flight. To obtain the most benefits from air augmentation, a combined cycle engine should be utilized. A single-stage fan, powered by the rocket turbopump

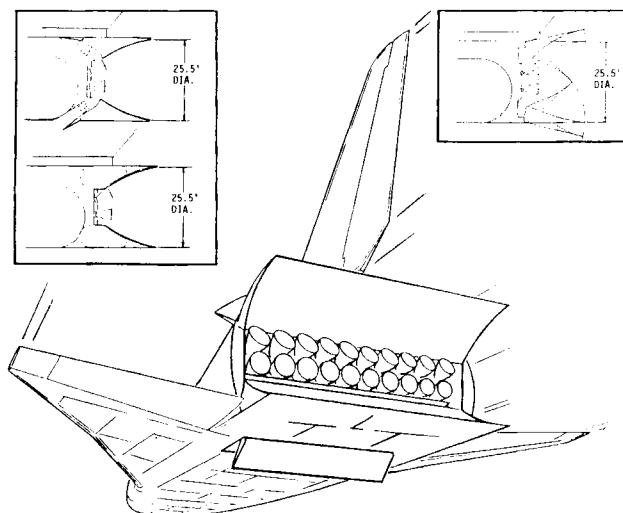


Fig. 13. SSTO vehicle with linear forced-deflection nozzle - gimbaled engines.

assembly, compresses air to provide a thrust increase. Some mixing of the rocket exhaust with the incoming air also provides a thrust increase. Since the rocket is a staged combustion cycle engine, all turbopump turbine-drive gases exhaust from the rocket nozzle. A performance loss would occur if the fan were driven by a separate gas generator. At approximately Mach 3 the air inlet duct can be closed and the rocket nozzles gimbaled onto the forced-deflection nozzle for normal operation.

Past attempts at achieving air augmentation were not successful because of the added air duct weight (12). In this instance, however, multiple use of the forced-deflection nozzle may make the system feasible. In addition, the linear skirt extension improves the lift capability during the return flight to, in effect, move the c.g. of the vehicle forward. It also can be rotated in to reduce base drag when necessary, as indicated in a schematic insert of Fig. 13, or used as an aileron. During powered flight the rotated out position offers a significant increase in area ratio (engine performance) not possible with other concepts.

A similar nozzle configuration is shown in Fig. 14, except that now the gimbaled engines are placed on a linear plug nozzle. Optimum sea level and altitude performance are achieved by gimbaling the nozzles off and on the plug at the proper altitudes. As with the configuration in Fig. 13, efficient TVC is obtained. The four centrally located engines are dual-fuel engines as before. Due to packaging on the plug, only a nozzle expansion area ratio of 180:1 is achieved with this configuration, which results in a payload of 51,000 lb.

The expansion ratio of the linear plug nozzle configuration can be increased to 185:1 by reducing the engine size (increasing the number of nozzles to 40). The resultant payload for this system is 44,000 lb. This arrangement, however, utilizes fixed rather than gimbaled engines, and suffers a sea level performance loss typical of altitude compensating nozzles of high area ratio. The TVC is also not as efficient in that it is provided by thrust modulation. Increasing

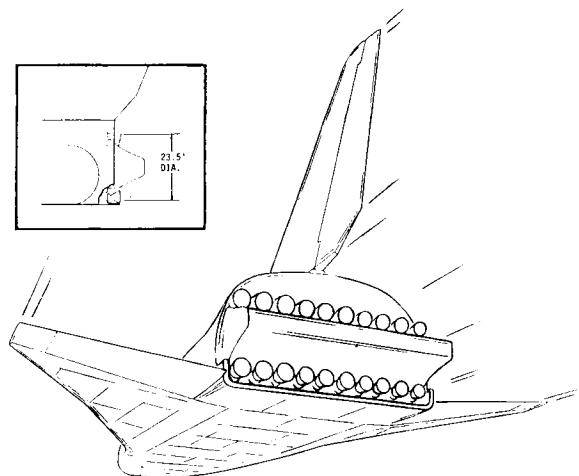


Fig. 14. SSTO vehicle with linear plug nozzle - gimbaled engines.

the number of nozzles, however, leads to a better aerodynamic profile on the plug than that achieved with larger discrete nozzles, while still allowing high pressure operation.

The best plug nozzle aerodynamics is obtained with the linear spike nozzle shown in Fig. 15. However, due to the lower chamber pressure operation dictated by heat transfer considerations, this engine achieves only a nozzle area ratio of 115:1 and a payload of 9,000 lb. The configuration also suffers a sea level performance loss and is less efficient in use for thrust vector control, similar to that for the 40-nozzle bell configuration.

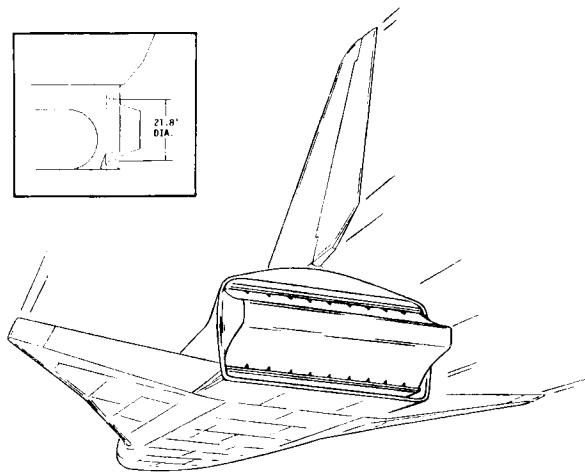


Fig. 15. SSTO vehicle with linear spike nozzle.

Another approach to improving the aerodynamics on a plug or in a forced-deflection nozzle is to provide the bell nozzle with a transition section from an axisymmetric throat to an elliptical or near rectangular cross-section (13). The transition section can be physically attached to a conventional bell nozzle at an area ratio of approximately 6:1 (Fig. 16) to obtain a super-elliptical nozzle with area ratio of about 40:1, or where the transition section is part of a telescoping nozzle. The computed performance for such a nozzle (circular to square-like) is 99.8% that for a conventional nozzle contour.

Because of the multiplicity of factors involved, no selection of nozzle configuration can be made without a full vehicle design study where weight and performance are carefully considered. Further, the type of vehicle (e.g., VTOHL, VTOVL or HTOHL) may dictate its own special requirements that involve a different nozzle configuration. The improvements that can be gained through the use of unique nozzle arrangements do not appear to be large in magnitude, especially if real performance and nozzle weights are taken into account.

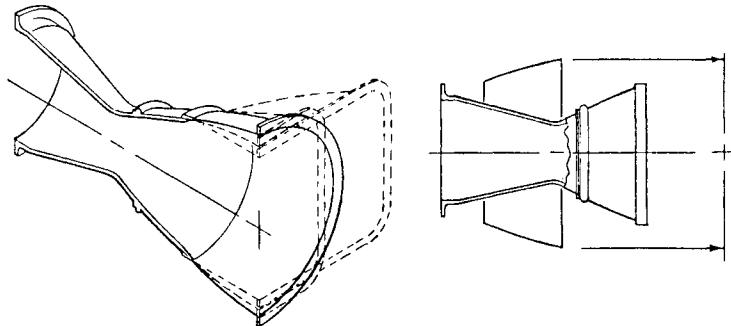


Fig. 16. Super-elliptical nozzles.

WHERE DO WE GO FROM HERE?

By defining the problem concerning the SSTO mission, we have been able to conceive propulsion system concepts that appear to ensure a successful accomplishment. Just as the staged combustion cycle was found to be the solution to achieving chamber pressures of the order of 4000 psia and the dual-fuel conversion the solution to obtaining greater density optimization, further cycle modifications may allow the attainment of higher pressures and thus greater payloads. The engineer must always remember to keep an open mind in looking for new incremental improvements. For example, his experience in the low pressure regime may indicate that a specific cycle may not be applicable, but at higher pressures, it might be the exact solution to the problem.

REFERENCES

1. G. K. O'Neill, J. Grey and E. Hannah, Organizing Committee, Space Manufacturing Facilities, Princeton University Conference No. 127 (May).
2. R. Salkeld, Single-stage shuttles for ground launch and air launch, J. Astronautics & Aeronautics, 12, No. 3, 52-64 (1974).
3. R. Beichel, Propulsion systems for single-stage shuttles, J. Astronautics & Aeronautics, 12, No. 11, 32-39 (1974).
4. R. Salkeld in D. Dipprey, D. Runkle and H. Davis, A Forecast of Space Technology, IV. Management of Energy, Jet Propulsion Laboratory (1975).
5. R. R. Atherton, Air Force Reusable Rocket Engine Program XLR129-P-1, Pratt & Whitney Aircraft Final Report AFRPL-TR-71-1, 1, 2, 3, Contract F04611-68-C-0002 (1971).
6. G. F. Kobalter, Development of expandable rocket nozzles, Aerojet Solid Propulsion Co. Final Report AFRPL-TR-74-9, Contract F04611-71-C-0070 (1974).
7. R. Beichel, Liquid rockets, J. Space/Aeronautics, 42, No. 4, 39-43 (1964).
8. R. Salkeld, private communication (1975).
9. R. Beichel and C. J. O'Brien, Dual-fuel engine technology, Aerojet Liquid Rocket Company Report 8774-06-F, (1974).
10. D. K. Huzel and D. H. Huang, Design of liquid propellant rocket engines, National Aeronautics and Space Administration Report NASA SP-125, 2nd Edition, 94 (1971).
11. R. Beichel, Nozzle concepts for single-stage shuttles, J. Astronautics & Aeronautics, 13, No. 6, 16-27 (1975).
12. C. J. O'Brien, R. W. Froelich and K. Morghen, Air augmented rocket concepts, Applied Physics Laboratory, Bumblebee Propulsion Panel Report TG 63053, The John Hopkins University, Maryland, 33-64 (1964).
13. V. H. Ransom and H. D. Thompson, Three dimensional supersonic nozzle flow field calculations, AIAA Paper No. 69-463, presented at the AIAA 5th Propulsion Joint Specialist Conference (1969).

RELIABILITY CONSIDERATIONS IN LONG-LIFE OUTER PLANET SPACECRAFT SYSTEM DESIGN*

E. Kane Casani

Jet Propulsion Laboratory

ABSTRACT

During the evolution of the Mariner Spacecrafts, a set of reliability design criteria has been developed. The Mariner Jupiter/Saturn Spacecraft, while designed with a four-year lifetime requirement, has a ten-year lifetime expectancy. The design has far more autonomy and fault detection and correction capability than previous Mariners.

The system design approach to achieve reliability has been to strive to eliminate single point failures by the use of redundancy, both block and functional, within the constraints of weight, cost, and schedule. This redundancy application is employed such that no single point failure can have a mission catastrophic effect. In cases where redundancy is not possible, reliability is improved by increased design margins, testing, and analysis.

Significant on-board correction capability is provided in the design since continuous tracking is not available and since the round trip light time (i.e., time to receive data from the spacecraft and transmit a command to the spacecraft) can be hours.

INTRODUCTION

Over the past decade, the Mariner spacecraft have made extensive scientific investigations of the inner planets. They have performed three flyby reconnaissance missions to Mars, Venus, and Mercury, and one Mars orbital mission. These Mariner designs evolved from the first three-axis-stabilized unmanned planetary spacecraft, designed and built by the Jet Propulsion Laboratory in the early 1960s. Having successfully accomplished the exploration of the inner planets, the design has matured into a third generation, which can be applied to the exploration of the outer planets. This new generation is first to be applied in the accomplishment of the Mariner Jupiter/Saturn 1977 mission.

Theoretical reliability analyses have been performed for the designs considered herein; however, that aspect of the subject is intentionally avoided. The historical development of the Mariner line is discussed to provide a background for the reliability design viewpoint. This discussion covers the actual vs design lifetime, the cause of mission end, the in-flight failures

*This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract NAS7-100, sponsored by the National Aeronautics and Space Administration.

and their consequences for the mission, and the use of redundancy in flight to avoid these failures.

The long-life outer planet system designs evolved from the previously developed medium-life inner planet system designs. This design process optimizes the use of proven subsystem and system designs and then makes the necessary improvements to increase the lifetime as required. In addition to longer lifetime, these missions also require additional on-board fault detection and correction capability to provide for the long periods of unattended flight necessitated by the longer round-trip communications time and the economic desire to reduce the requirement for continual ground coverage.

CONCLUSIONS

In the design of a long-life outer planet spacecraft system, the two most important considerations are lifetime and round-trip communications time. The lifetime design problem is handled in a way similar to its handling in previous designs. Additional improved reliability is attained by employing proven and well understood designs and increasing the fault tolerance of the design, particularly in the susceptible areas. Although parts quality, coupled with thorough design, analysis, and test programs, is basic to a reliable spacecraft design, these topics are not discussed here.

The increased round-trip communications time is so long that certain types of on-board failures could be mission catastrophic if corrective action is not taken on board. To prevent such failures, it is necessary to incorporate sophisticated on-board fault detection and correction capability.

MARINER MISSION SUMMARY

From time to time throughout this paper it will be necessary to refer to different Mariner missions, spacecraft, and the design generation. Table 1 shows the interrelationship of these parameters. This missions are designated by the launch year and target planet(s), and the spacecraft are designated sequentially with arabic numbers; those spacecraft that were not successfully injected due to a launch vehicle failure receive a designator, but are absent in most of the subsequent tables. The design generation shows that the first two Mariners were evolved from earlier designs; they are therefore not treated in this paper. The remaining ones are divided into three generations.

The first generation is characterized by the fact that it was the first design in which long lifetime was a serious design consideration. A sizeable effort was devoted to the question of long life during system design. Although the lifetime requirement was only 9 months, it was long compared to the previous experience with this class of spacecraft design, which was only a matter of days. It is true that the Mariner 1962 mission lasted 4 months, but since this was not a highly redundant design and the long-life feature was a "worry" rather than a design consideration, the mission is not treated herein.

The second generation began with Mariner 1969, and many (if not all) of the lessons learned from the previous generation were incorporated into its design. This design included such features as a programmable sequencer (the forerunner of the on-board computers now being flown), a significant improvement in telecommunications performance, which included many variable data

TABLE 1 Mariner Mission Summary

Mission	Launch Date	Spacecraft Designation	Cause of End of Mission	Design Generation
Mariner 1962 Venus	August 1962	Mariner 1 Mariner 2	Not injected Unknown	Earlier design
Mariner 1964 Mars	November 1964	Mariner 3 Mariner 4	Not injected Gas depletion	Mariner first generation
Mariner 1967 Venus	June 1967	Mariner 5	Unknown	Mariner first generation
Mariner 1969 Mars	February 1969	Mariner 6 Mariner 7	Gas depletion Gas depletion	
Mariner 1971 Mars Orbiter	May 1971	Mariner 8 Mariner 9	Not injected Gas depletion	
Mariner 1973 Venus	November 1973	Mariner 10	Gas depletion	Mariner second generation
Mercury				
Viking 1975 Mars Orbiter	August 1975	Not designated	Not launched	
Mariner 1977 Jupiter Saturn	September 1977	Not designated	Not launched	
Mariner 1979 Jupiter Uranus	November 1979	Not designated	Not approved	Mariner third generation

rates to take advantage of better than nominal performance, and a significant increase in redundancy (see Table 2).

While the Viking 1975 design is considered second generation, it contributed significantly to the transition to the third generation and perhaps should be considered second and a half generation. The Viking design is virtually fully redundant and represents the first significant effort devoted to the elimination of "single point failures." This important step in additional redundancy can also be seen in Table 2. Viking is still considered a second-generation design, since the basic system architecture and subsystem designs are not significantly different from the other second-generation designs. The single most significant change was made in the on-board computer. Mariner 1971 and 1973 contained a 512-word programmable computer with a maneuver sequencer that could be used for redundancy during the maneuvers; Viking 1975 contains a 4000-word, 64-operational-instruction, general-purpose, totally block-redundant computer. This computer is used to perform many standard and nonstandard (failure work-around) tasks on board the spacecraft.

The third generation begins with the Mariner 1977 design, where the lifetime requirement has taken another step from the class of 1 year to the realm of 4 years. This generation is categorized by significant system and subsystem redesigns to handle the longer lifetime and increased round-trip communications time. While the redundancy is not significantly greater than that of Viking, the management of that redundancy and the interactions between the ground and the spacecraft are different, chiefly in the additional on-board intelligence. In addition to the Computer Command Subsystem, there are two other programmable subsystems, the Attitude Control and Flight Data Subsystems. These combine to give a total on-board memory capacity of 64,000 words.

HISTORICAL DEVELOPMENT OF MARINER RELIABILITY

Before delving into the reliability considerations for the design of long-life outer planet spacecraft system design, it is informative to look into the past developments that have brought the design to the current point of departure. The Mariner spacecraft line has evolved over the past decade into a highly developed, sophisticated, unmanned planetary spacecraft. The first "true" Mariner was flown to Mars in 1964. The Mariner 1964 design was the first design with long life as a prime consideration; it therefore will be considered as the first of the Mariner line to be discussed in this paper.

The basic design approach was to translate the mission requirements into spacecraft design requirements and then to synthesize designs to satisfy those requirements. The critical failure modes and their effects on the achievement of the mission objectives were then identified. These critical failure modes were then removed by the application of block or functional redundancy to the extent permitted by the resources, particularly weight and power.

In block redundancy, two identical elements are provided to perform the same function; it can be implemented in an active mode or in a standby mode. In the active mode, both elements are powered and on-line simultaneously; in the standby mode, the redundant element is switched on only in the event of a failure of the primary element. The active mode avoids the switching logic and its associated complexity at the expense of additional power. Functional redundancy, sometimes called alternate mode redundancy, requires the

TABLE 2 Mariner Redundancy

	1964	1967	1969	1971	1973	1975	1977
<u>02</u>							
S-band exciter	2	2	2	2	2	2	2
S-band amplifier	2	2	2	2	2	2	2
X-band exciter	-	-	-	-	1	1	2
X-band amplifier	-	-	-	-	-	-	2
Receivers	1	1	1	1	1	2	2
<u>03</u>							
CDU decoding	1	1	1	1	1	-	-
CDU detector	1	1	1	1	1	2	2
CDU power	1	1	1	1	1	2	2
TMU interface	-	-	-	-	2	2	2
TMU modulator	-	-	-	-	2	2	2
TMU (block)	-	-	-	-	2	2	2
TMU oscillator	-	-	-	-	2	2	2
TMU mode cont.	-	-	-	-	2	2	2
TMU power	-	-	-	-	2	2	2
<u>04</u>							
400 1d	1	-	1	1	-	-	-
Booster regulator	2	2	2	2	2	2	-
2.4 inverter	1	2	2	2	2	2	2
400 3d	1	1	1	1	1	2	-
30 Vdc	-	-	-	1	1	2	-
Shunt reg/rad	-	-	-	-	-	-	1
Battery charger	1	1	1	1	1	2	-
Battery	1	1	1	1	1	2	-
<u>05</u>							
Sequencing	1	1	1	1	1	2	2
Decoding	-	-	-	-	-	2	2

TABLE 2 (contd)

	1964	1967	1969	1971	1973	1975	1977
<u>06</u>							
Timing	2	2	2	2	2	2	2
A/DC	2	2	2	2	2	2	2
Modulation	1	1	2	2	-	-	-
Mixer and output	1	1	2	2	-	-	-
Eng. mix	1	1	2	2	2	2	2
Memory	-	-	-	-	2	2	2
Memory cont.	-	-	-	-	1	2	2
Power converter	1	1	1	1	2	2	2
<u>07</u>							
ACE	1	1	1	1	1	2	2
IRU	1	1	1	1	1	2	2
RCA	2	2	2	2	2	2	2
Canopus tracker	1	1	1	1	1	1	2
Sun sensor	1	1	1	1	1	1	2
<u>08</u>							
Pyro	2	2	2	2	2	2	2
<u>16</u>							
DSS	1	1	1	1	1	2	1
Number of redundant elements	14	16	22	22	34	60	56
Number of elements	33	32	37	38	50	63	58
Redundant elements, %	42	50	59	58	68	95	97
Redundant electronics, %	22.6	24.9	20.5	18.6	19.1	41.6	38.5
Number of redundant functions	7	8	11	11	17	30	28
Number of functions	26	24	26	27	33	33	30
Redundant functions, %	27	33	42	40	51	91	93

capability of performing the same, or nearly the same, function with two different implementations. An example of functional redundancy is maintaining the spacecraft inertially referenced by using either celestial sensors which can recognize the sun, the earth, or the star Canopus and orient the spacecraft with respect to these celestial bodies, or a gyro package inside the spacecraft that has some a priori inertial reference system about which the spacecraft can be oriented.

Functional redundancy has the advantage over block redundancy of avoiding systematic design deficiencies, but on the other hand, the functional alternate usually does not have the same performance as the primary.

This design process gave birth to the Mariner 1964 spacecraft with the following redundant features:

- Two RF power amplifiers
- Two RF excitors
- Two power regulators
- Two PN generator-A/D converter combinations
- Two pyrotechnic subsystems
- Dual gas systems, including storage tanks, regulators, valves, and jets

Using Mariner 1964 as a baseline, the evolution of major subsystems as well as the system will be explained through the Mariner Jupiter/Saturn 1977 mission, including:

- Mariner Mars 1964 (MM'64)
- Mariner Venus 1967 (MV'67)
- Mariner Mars 1969 (MM'69)
- Mariner Mars 1971 (MM'71)
- Mariner Venus/Mercury 1973 (MVM'73)
- Viking Mars Orbiter 1975 (VO'75)
- Mariner Jupiter/Saturn 1977 (MJS'77)

Although no Mariner missions are approved after Mariner Jupiter/Saturn 1977, there are missions under consideration that use the MJS'77 generation. These include a Mariner Jupiter/Uranus 1979 (MJU'79) mission and a Mariner Jupiter Orbiter 1981 mission. Activities relative to increasing the 4-year lifetime of the MJS'77 design for a 6-year MJU'79 design are under way at the time of this writing.

In studying this evolution of subsystems, it is informative to consider three major subsystems: the Computer Command Subsystem, the Radio Frequency Subsystem, and the Power Subsystem. These subsystems have evolved from little or no redundancy to total block redundancy with some functional redundancy, particularly in the Computer Command and Radio Frequency Subsystems.

Computer Command Subsystem

Although the Mariner Jupiter/Saturn 1977 Computer Command Subsystem (CCS) is only one of the three computers on the spacecraft, it performs all the prime on-board computations and control functions. It is a 4000-word,

64-operational-instruction, general-purpose, totally block-redundant computer. It can be completely reprogrammed from the ground. This machine has evolved from the simple Central Computer and Sequencer (CC&S) of Mariner 1964, which had two hardwired sequencers, used for the launch and cruise sequences, and a maneuver sequencer with three variable registers into which could be stored the maneuver parameters (Fig. 1). For the Mariner Mars 1969 mission, this CC&S was replaced with a new design (Fig. 2). The new CC&S retained the maneuver sequencer and replaced the two hardwired sequencers with a 128-word, in-flight reprogrammable, 8-operation, sequence generator. In addition to this mission sequencing flexibility, the maneuver reliability was significantly enhanced, since it was possible to perform the maneuver under the maneuver sequencer control and computer control. The in-flight reprogrammability feature of this machine was not really exploited until the Mariner Mars Orbiter of 1971. For the 1971 mission, the memory size was increased from 128 to 512 words. This was possibly the single most contributing factor to the success of that mission. When the spacecraft arrived at Mars, the planet was totally engulfed in a dust storm, requiring a complete redesign of the entire preplanned mission sequence. This flexibility was accommodated by the CC&S.

The basic design was successfully flown on three missions (1969, 1971, 1973), and the potential of in-flight reprogramming was recognized. Therefore, the fixed maneuver sequencer was removed, and the computer capability of the machine was expanded to the fully redundant, 4000-word, 64-operational-instruction, general-purpose equipment of the Mariner Jupiter Saturn 1977 spacecraft (Fig. 3).

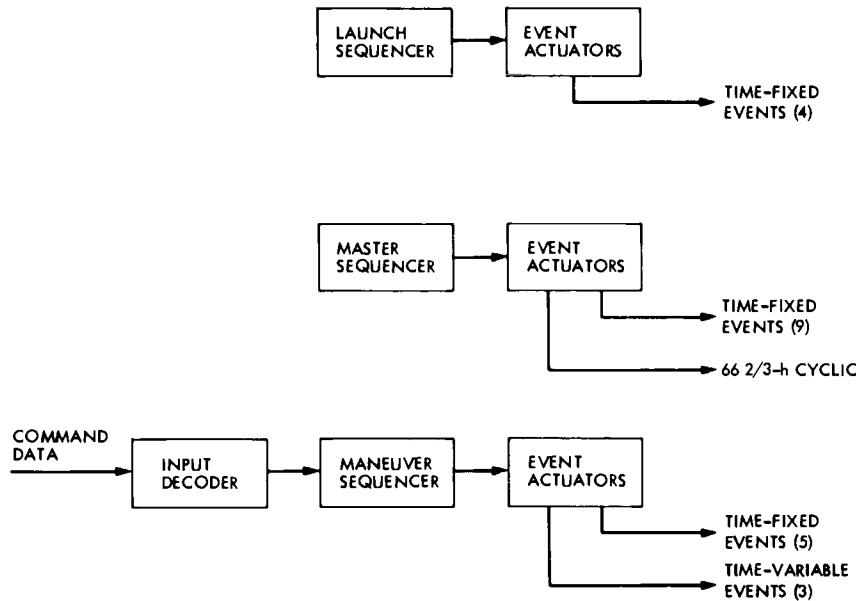


Fig. 1. Mariner 1964/1967 CC&S.

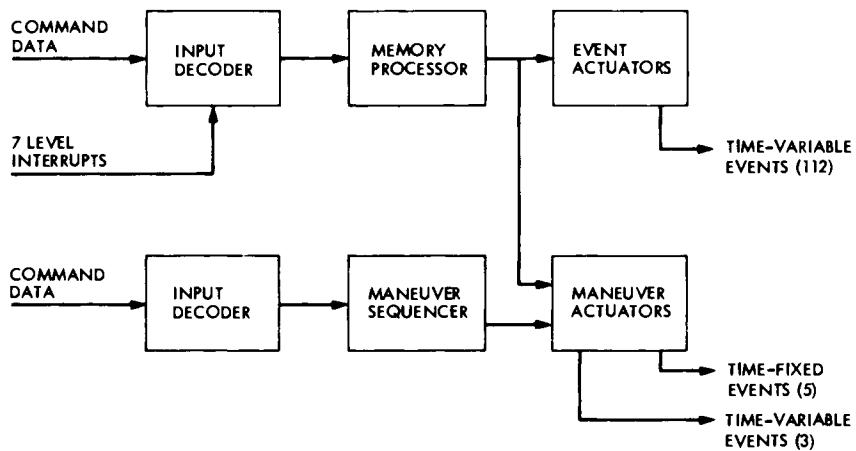


Fig. 2. Mariner 1969/1971 CC&S.

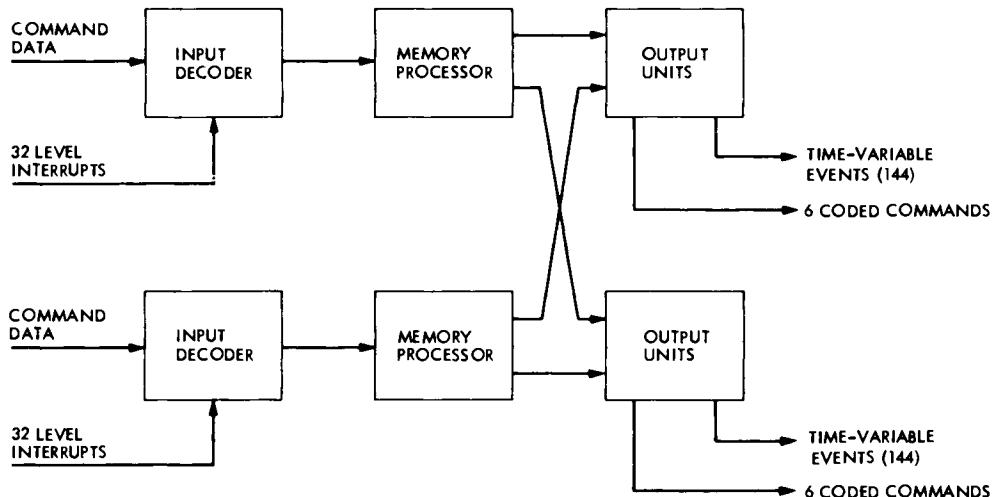


Fig. 3. Viking 1975/Mariner 1977 CC&S.

Radio

The Mariner Jupiter Saturn 1977 radio is fully redundant, with two S-band excitors, two S-band power amplifiers, two X-band excitors, two X-band power amplifiers, and two receivers (Fig. 4). In both the S-band and X-band systems, the excitors are fully cross-strapped into the amplifiers. The Mariner Mars 1969 design, from which this radio evolved, contained a fully redundant

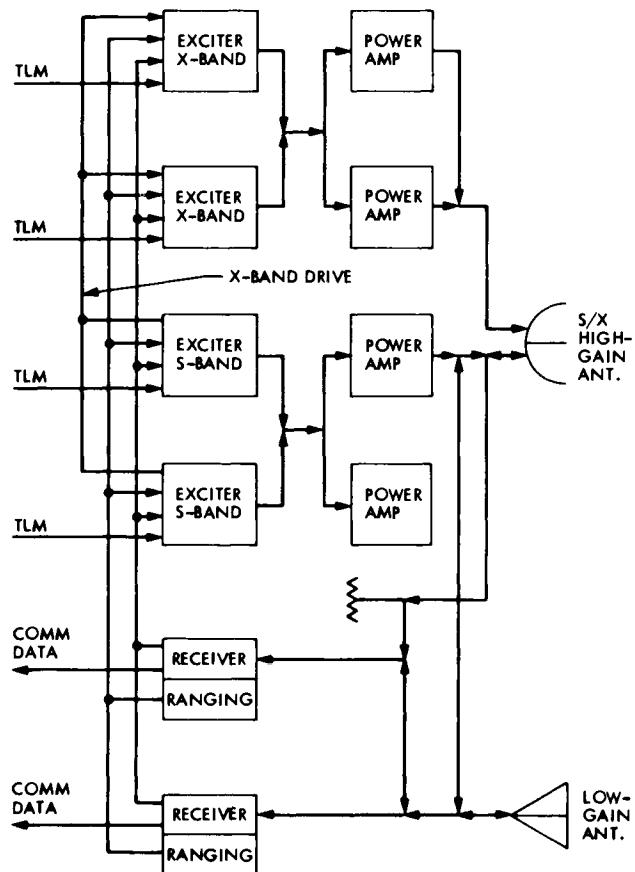


Fig. 4. Mariner 1977 radio.

S-band system, with two S-band exciters fully cross-strapped into two power amplifiers and a single receiver (Fig. 5). While the single receiver (which receives only ground commands) was a reliability concern, the mission sequence was contained in the CC&S as a functionally redundant backup. The prime function of the ground command link was to load the maneuver parameters needed to perform the trajectory correction maneuver into the maneuver sequencer, after which the ground commands were sent to back up the on-board sequenced events. This basic radio configuration (Fig. 4) was used on MM'64, MV'67, MM'69, MM'71, and MVM'73, and it was not until the Viking Mars Orbiter 1975 that the design was changed by the addition of a redundant receiver (Fig. 6). Subsequent to that, the redundant X-band system was added for the MJS'77 spacecraft design.

Power

The Mariner Jupiter Saturn 1977 Spacecraft Power Subsystem uses radioisotope thermoelectric generators (RTGs) as a power source. All previous designs

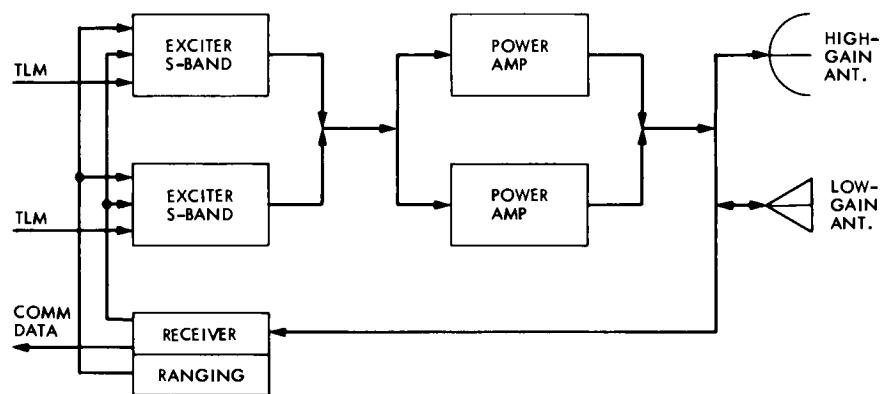


Fig. 5. Mariner 1964/1967/1969/1971/1973 radio.

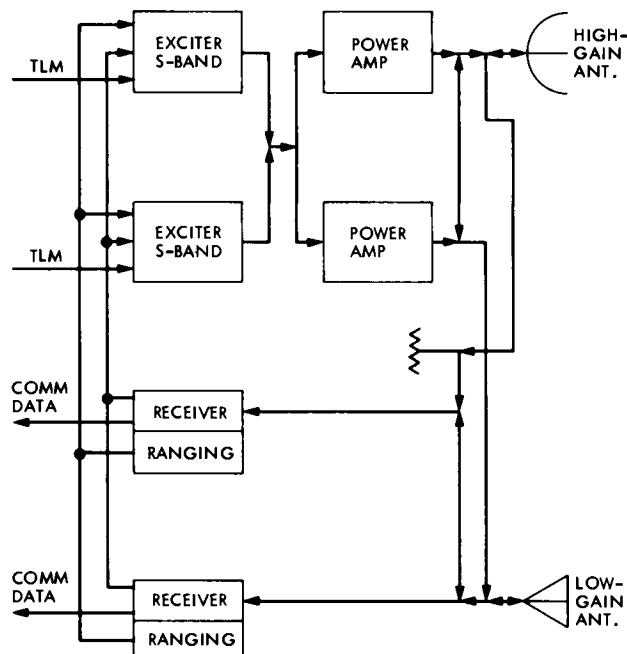


Fig. 6. Viking 1975 radio.

were solar cell powered. Unlike solar cells, whose voltage varies with solar radiation intensity, the RTGs have essentially a constant output voltage. The power from the RTGs is used to drive the 2.4-kHz inverter, which supplies all the ac loads to the spacecraft. There are two block redundant inverters. Should the prime inverter fail, an on-board failure detector automatically switches in the standby inverter. This switch will be effected if the output voltage is low for a prescribed length of time. The dc loads are supplied directly from the RTG bus. A shunt regulator is incorporated to keep the total load on the RTG essentially constant; that is, when the user demand is small, the power is shunted from the RTGs through the shunt regulator to a radiator plate. Two redundant batteries are provided to perform specific functions during the injection phase. These batteries are no longer used after the injection phase (Fig. 7).

The previous Mariner power configurations were slightly different, the power source being solar panels instead of RTGs. They included a booster regulator, which supplied a fixed regulated voltage to the inverters, since the solar panel voltage changes as the spacecraft moves inward on a Venus or Mercury mission or outward on a Mars mission. A battery, which is continuously on-line, is provided to supply power when the spacecraft is not sun-oriented. In the MM'69, MM'71, MVM'73, and VO'75 configurations (Fig. 8), the main booster and main 2.4-kHz inverters were tied directly together, forming a main power chain; similarly, the standby units formed a standby chain. This equipment was an updated version of the Mariner 1964/1967 designs (Fig. 9). In these designs, the main inverter handled all the cruise loads and the maneuver inverter all the maneuver loads. There was no redundancy in these inverters, although there was some load sharing that provided operational work-around capability should the main inverter fail. Two boosters were provided so that either booster could supply power to the main inverter. Thus,

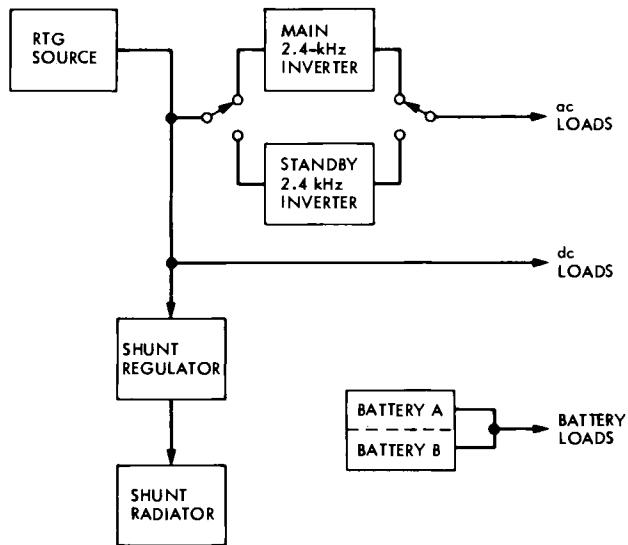


Fig. 7. Mariner 1977 power.

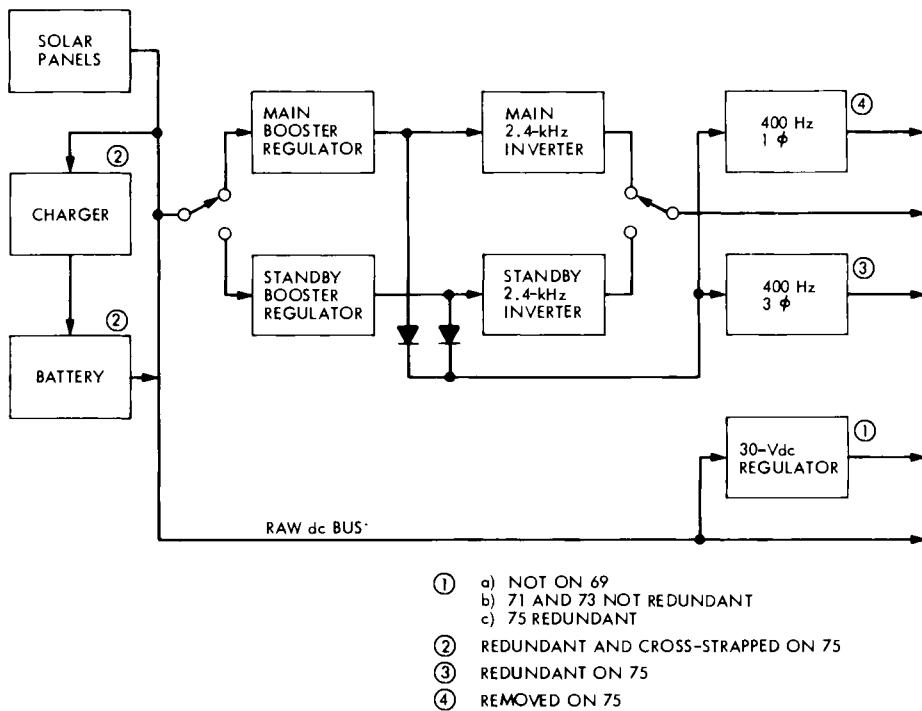


Fig. 8. Mariner 1967/1971/Viking 1975 power.

the power subsystem of Mariner 1964 with limited redundancy evolved into the Viking 1975 design of full redundancy, which was then modified to meet the requirements of the RTG power source. Both the 1975 and 1977 configurations have no single point failure which can cause loss of power.

Weight of Redundancy

One of the drawbacks of the application of redundancy is the additional system weight incurred. Table 3 shows the weight of the spacecraft under discussion in both their fueled and unfueled configurations, the weight of their redundancy, and the percentages of the total weight it represents. All of these missions have had high performance demands, and this additional weight was difficult to accommodate.

Redundancy Evolution

To further clarify the design evolution, Table 4 displays the way in which the functional element block redundancy has increased. Here it can be observed that the basic number of functional elements is about 20 for all designs, and the redundancy percentage has increased by a factor of two to the point where the system is almost 100% redundant. It should be further noted that, where block redundancy does not exist in the MJS design, there is functional redundancy.

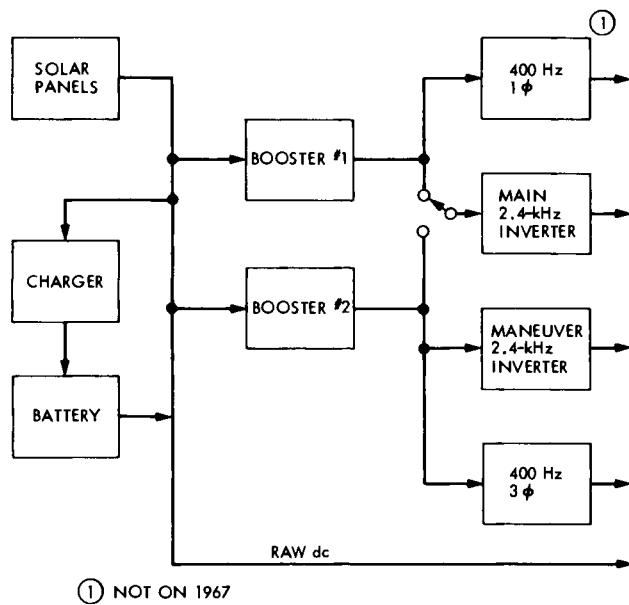


Fig. 9. Mariner 1964/1967 power.

TABLE 3 Weight and Percentages of Redundancy

	MM'64	MV'67	MM'69	MM'71	MVM'73	VO'75	MJS'77
Spacecraft total weight, kg	285.6	268.3	420.6	1029.9	513.3	2369.3	772.8
Spacecraft dry weight (less propulsion, adapter, kick stage), kg	239.3	221.9	370.9	422.6	448.7	713.2	681.4
Spacecraft electronics weight, kg	95.7	88.7	147.4	169.0	179.5	285.3	152.0
Block redundancy weight, kg	21.6	22.1	30.2	31.5	34.4	118.7	58.6
Spacecraft total redundancy, %	7.6	8.2	7.2	3.1	6.7	5.0	7.5
Spacecraft dry redundancy, %	9.0	10.0	8.2	7.4	7.7	16.6	8.6
Spacecraft electronics redundancy, %	22.6	24.9	20.5	18.6	19.1	41.6	38.5

TABLE 4 Number of Block Redundant Functional Elements

Hardware	MM'69	MM'71	MVM'73	VO'75	MJS'77
Command receivers	1	1	1	2	2
S-band power amplifiers	2	2	2	2	2
S-band excitors	2	2	2	2	2
X-band power amplifiers	0	0	0	0	2
X-band excitors	0	0	1	1	2
Command detectors	1	1	1	2	2
Command decoders	1	1	1	0	0
Telemetry modulators	2	2	2	2	2
2.4-kHz inverters	2	2	2	2	2
400-Hz inverters	1	1	1	2	0
CCS redundancy	2 ^a	2 ^a	2 ^a	2	2
Telemetry processor	1.5	1.5	1	2	2
A/C torquer sets	2	2	2	2	3
Gyro/IRU sets	1	1	1	2	2
Sun sensor sets	1	1	1	1	2
Canopus trackers	1	1	1	1	2
ACE	1	1	1	2	2
Pyro systems	2	2	2	2	2
Data storage	1	1	1	2	1
Propulsion valves	2	1	1	1	2
Timing oscillators	2	2	2	2	2
Articulation	1	1	1.5	1	1
Redundant elements	10	9	9	15	18
Basic elements	20	20	21	20	20
% redundant elements	50	45	43	75	90

^aRedundancy in maneuver control only.

Table 2 further expands the redundancy picture by identifying the redundancy at the lowest hardware element. That is, the functional elements of Table 4 may have several internal independent redundant elements. We also see the increasing number of individual elements as the design sophistication increases. In addition to showing the percentage of redundant elements, the table also presents this percentage of redundant electronics by weight, eliminating the weight of such major nonelectronic elements as the structure, mechanical devices, power sources, propulsion, cabling, and adapter. In both tables, the evolution to higher redundancy is clearly evident.

Spacecraft Configuration

The Mariner Jupiter/Saturn 1977 spacecraft consists of a ten-sided (ten-bay) basic bus with an articulated science scan platform for pointing the instruments at the planet, an RTG power source, and an earth-pointing high-gain antenna. These same features, except for the RTG power source, have been included on all Mariner spacecraft. (The others were solar cell powered.) The number of electronic bays has increased from the earlier designs that had eight bays. The size and the packaging within the bays has remained essentially constant throughout the evolution. The details within the bays have changed as the component technology evolved from discrete components and printed circuit boards to integrated circuits and multilayer boards.

The high-gain antenna has been used for high-data-rate transmission from the planet to earth, mostly of the imaging data. The antenna has been body fixed or articulated in 1 or 2 degrees, depending on the planetary encounter geometry. Similarly, the articulation scheme on the science platform has varied from mission to mission.

HISTORY OF MARINER FLIGHT PERFORMANCE

In examining the applicability of the current Mariner designs to long-life outer planet missions, it is interesting to study the performance of the past Mariners. The Mariner flight performance can be characterized by the following five significant observations:

- (1) All primary mission objectives were achieved; minor objectives not accomplished were due to an instrument failure, not an engineering (or support) functional failure.
- (2) The spacecraft sustained a large number of failures/anomalies (10 to 20).
- (3) A small percentage of the available redundancy, both block and functional, was needed in flight to accomplish the mission.
- (4) There is no indication that the failure rate is constant with time and that the classic random failure prediction technique prevails (i.e., λT is constant).
- (5) Actual lifetime data are difficult to infer from flight performance since the majority of the missions were terminated by the depletion of the attitude control gas necessary to maintain the spacecraft attitude.

The achievement of all primary mission objectives can be attributed to the fundamental design philosophy used in the system and subsystem designs of the Mariner spacecraft.

This philosophy is characterized by taking advantage of the lessons learned from previous designs and missions and striking a proper balance between maximizing the possibility of significant scientific return and minimizing the possibility of catastrophic failure. Specific reliability criteria developed are:

- (1) The spacecraft design must be governed by the policy that the requirement for reliable operation takes precedence over the requirements for additional capability/flexibility beyond that required to achieve the basic mission.
- (2) The design should take advantage of the hereditary equipment and designs and of the experience gained in the previous projects to the maximum extent feasible.
- (3) Where there are new designs or modifications to the previous designs, care must be taken in the choice and control of new fabrication techniques and operational procedures to enhance the probability of success.
- (4) Within cost, mass, and schedule constraints, block redundancy or functional redundancy should be employed such that no single failure mode of any component (electronic, mechanical, pyrotechnic, electromechanical, or structural) can cause the loss of all data return from more than one science instrument or the loss of more than 50% of the engineering data. Acceptable single point failure risks must be identified and approved by the Project Manager.
- (5) Functional interdependency between subsystems can be used only where such design contributes to the flexibility and/or the economy of the system without reducing the system reliability. It must be a goal that a failure-independent element or its interface may cause no more than a graceful degradation of the system performance.
- (6) All scientific instruments must be designed to be functionally independent of one another. The spacecraft must be designed in such a manner that a failure in one instrument or in a nonscience subsystem common to the support of several instruments will have a minimum effect on the total data transmitted and other spacecraft functions.
- (7) Particular emphasis must be placed upon proven designs plus new designs that contribute to the system's flexibility and/or economy without reducing the system's reliability, along with a complete program of component, subsystem, and system testing.

These philosophies have evolved over the past decade of Mariner development. To some extent, the celestial mechanics and fiscal constraints have contributed in that the launches have been on 2-year centers. This is enough time to benefit from mistakes of the previous missions and correct them in the next, and yet not enough time to introduce radical new designs or deviations from flight-proven successful designs, which is so often the designer's desire.

It is understandable that failures to achieve mission objectives have been the result of a scientific instrument failure. Most of the instruments flown are new designs and do not have the benefit of the kind of evolution typical of the engineering subsystems. It is also true that, in a given instrument, one can consider taking a higher risk if the scientific results justify it. On the other hand, the Mariner imaging subsystem has been handled more like the engineering subsystems. It has evolved along with the rest of the design and has exhibited the same high quality of performance.

Table 5 is a summary of the Mariner in-flight performance relative to anomalies and failures. It presents both anomalies and failures as well as their causes according to the following definitions and descriptions.

An anomaly is an occurrence in a given subsystem or system state (configuration) that is not expected or predicted. It is not considered an anomaly if the component, subsystem, or system reacted as it was designed to react, but an imperfect understanding of the design led to an erroneous prediction or interpretation of that reaction. Thus, the component, subsystem, or system may experience an anomalous environment or event and react unexpectedly but not anomalously, and therefore the anomalous environment or event is the anomaly. If the component, subsystem, or system subsequently fails, that failure is considered an induced or secondary failure.

The failures are classified into mission failure, spacecraft system failure, spacecraft subsystem failure, component failure, or piece part failure.

Mission failure is any anomaly that results in

- (1) Permanent loss of control of the spacecraft prior to the last planned computer update.
- (2) Termination of the spacecraft life prior to the design life.
- (3) Permanent loss of data from more than one scientific instrument.

Spacecraft system failure is any anomaly that results in

- (1) Temporary loss of control of the spacecraft.
- (2) Temporary loss of communications from the spacecraft.
- (3) Loss of data from any one scientific instrument.
- (4) Loss of more than 50% of the engineering data.

Spacecraft subsystem failure is any anomaly that results in

- (1) Required use of block or functional redundant elements.
- (2) Required use of unplanned sequences.
- (3) Loss of any scientific or engineering data.

Component failure is any anomaly that results in failure of the component to function as designed.

Piece part failure is any anomaly that results in specific electronic piece part failure, excluding TWTs, TWTAs, image dissector tubes, photomultipliers, vidicons, and batteries.

TABLE 5 Mariner Performance Summary

Spacecraft	Number of Prime Anomalies	Number of Failures ^a					Cause of Anomalies and/or Failures ^b				
		M	S	U	C	P	D	M	S	R	W
Mariner 4	12	0	3	5	2	8	1	6	2	1	3
5	5	0	0	0	5	0	2	0	0	0	3
6	16	0	0	12	2	3	8	2	11	3	1
7	13	0	1	10	4	4	4	1	10	6	0
9	27	0	0	13	5	6	3	1	19	6	3
10	26	0	2	15	12	4	7	4	14	5	3

^aM: Mission failure

S: System failure

U: Subsystem failure

C: Component failure

P: Piece part failure

^bD: Design

M: Manufacture

S: Secondary

W: Searout

R: Random

U: Unknown

The causes of the failures are classified into design, manufacture, secondary, random, wearout, or unknown.

Design failures are characterized by reproducibility, imperfect understanding of the environment, improper use of components, or inadequate understanding of the interface.

Manufacturing failures are characterized by an error in the operations or assembly of a component, unit, subsystem, or system.

Secondary failures are a direct result of another anomaly or failure, a human error, or an act of God.

Random failures are ones that cannot be attributed to a design, manufacturing, secondary, or wearout problem.

Wearout failures are a result of a known, accepted, end-of-service-life process.

Unknown failures are those in which the cause of failure cannot reasonably be attributed to a design, manufacturing, secondary, wearout, or random failure.

While Table 5 shows that there have been a rather significant number (approximately 10) of failures/anomalies on each flight, it also demonstrates that there has actually been a relatively small number of piece part failures. Also it should be noted that, although these piece part failures are considered random, they are not consistent with any predictable failure rates.

It is interesting to observe the small number of block and/or functional redundancies that have had to be used in flight because of an anomaly. It was pointed out earlier that the fault tolerance of the design can be closely correlated with the available redundancy. In Table 6, it can be seen that this fault tolerance has not been severely taxed. Two out of 22 redundant elements were used on the Mariner 6 flight and seven on Mariner 7; Mariner 9 used 4 out of 22 and Mariner 10 used 9 out of 34.

Table 6 also shows the actual and the design lifetimes, illustrating that all flights met or exceeded their design lifetime, and that the dominant cause of the end of the mission was the depletion of the attitude control gas. Without this gas, the spacecraft can no longer maintain its attitude and keep the solar panels pointed at the sun. After loss of attitude, the battery is depleted in several hours and the mission terminated.

Functional redundancies are identified, but only for those cases in which no block redundancy is available. In actuality, the functional redundancies for the second- and third-generation Mariner designs are virtually impossible to compute and surely beyond the scope of this paper. Again referring to Table 6, it can be seen that little functional redundancy has been employed.

The lack of need to use the redundancy available is a strong indication of the highly fault-tolerant design of the Mariner line, and this is indeed one of the strong bases for the belief in inherent long-life capability.

Figure 10 shows the Mariner failure history in terms of the number of failures vs flight time in days. In studying this figure, one will make two

TABLE 6 In-Flight Redundancy and Mission Duration

	Redundancy		Lifetime, days		Cause of End of Mission
	Block	Functional	Actual	Design	
Mariner 4	0	1	1117	250	Gas depletion
Mariner 5	0	0	173	180	Unknown
Mariner 6	1	1	770	250	Gas depletion
Mariner 7	2	5	643	250	Gas depletion
Mariner 9	2	2	515	300	Gas depletion
Mariner 10	4	5	506	186	Gas depletion

interesting observations. First, there is no correlation between the number of failures and the flight time. In other words, there is no indication of constant failure rate. Second, the failures are in a band from approximately 10 to 18. The following explanation of these two observations is offered. Referring again to Table 5, it can be seen that 37% of the total failures are component, and a smaller number of known failures are random, say perhaps 10-15%. Thus, the random component is small, making the total slightly dependent on time. A large percentage of the failures are design, manufacture, and secondary. While it is true that these should be eliminated by the design, test, and quality assurance programs, there is a practical limit to the effectiveness of such effects. In other words, what is shown here is a measure of the effectiveness of that screening process. Figure 11 shows an idealized graph of percent perfection vs resource, where resource can be money, time, people, etc. It can be seen that to achieve 100% perfection requires significantly more resources than to settle for 80 to 90%. These additional resources can be better applied to increase the fault tolerances of the system, which is exactly what has been done. The design of the spacecraft recognizes about 10 to 20 expected failures/anomalies. These will occur unpredictably in time, location, and consequences. The only thing predictable is that some number of failures, larger than 10 and hopefully less than 20, will occur, and the system must have the fault tolerance to accept these failures.

The observation that the missions have been terminated by the depletion of the attitude control gas is disturbing because it has precluded the gathering of flight data that would be most informative in supporting or refuting speculations about the true inherent lifetime capabilities of the third-generation long-life outer planet Mariner spacecraft.

MARINER DESIGN CHARACTERISTICS

Early in the design stage of the Mariner Jupiter/Saturn spacecraft, a specific definition of the design characteristics was established. This process has been used on all previous Mariners to focus and guide the development of the design.

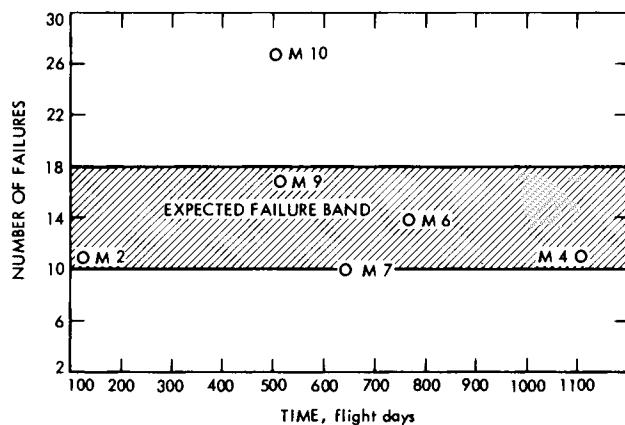


Fig. 10. Mariner failure history.

The Mariner Jupiter/Saturn 1977 spacecraft will be attitude stabilized in three axes, using the sun and Canopus as primary reference objects while maintaining communications with the earth. The spacecraft will have two-way communication equipment that permits the transmission of science data to the earth, receipt of command transmission at the spacecraft, two-way doppler tracking, and range measurements. The spacecraft will be capable of executing on-board stored sequences for trajectory correction and science instrument pointing maneuvers, antenna pointing, science instrument pointing,

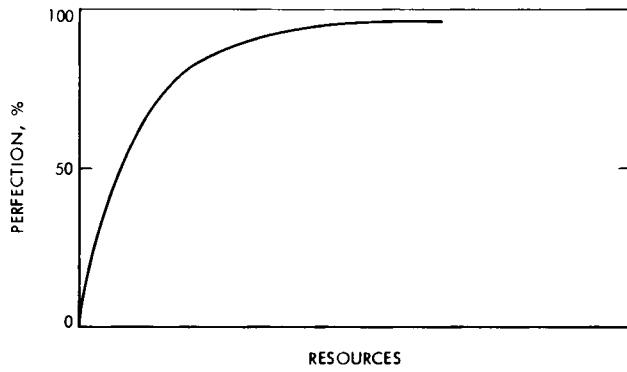


Fig. 11. Percent perfection vs. resources.

science and engineering data acquisition, and data formatting. Electrical power will be derived from solar-independent sources (i.e., RTGs).

Experiments will be performed during the cruise phase between planets and in the vicinity of Jupiter and Saturn, which includes selected satellites and Saturn's rings. In addition, it is expected that the science instruments can be used for an extended mission to investigate the region beyond Saturn until the mission is terminated by a depletion of attitude control fuel, inadequate performance capability, or spacecraft failure. The radio science experiment is performed using the spacecraft telecommunications equipment. The science subsystems are listed below:

Cosmic Ray Subsystem	(CRS)
Planetary Radio Astronomy Subsystem	(PRA)
Ultraviolet Photometer Subsystem	(UVP)
Low Energy Charged Particle Subsystem	(LECP)
Photopolarimeter Subsystem	(PPS)
Plasma Subsystem	(PLS)
Ultraviolet Spectrometer Subsystem	(UVS)
Magnetometer Subsystem	(MAG)
Imaging Science Subsystem	(ISS)
Infrared Interferometer Spectrometer and Radiometer Subsystem	(IRIS)

OUTER PLANET MISSION RELIABILITY CONSIDERATIONS

To date, all the Mariner missions flown have been to the inner (or terrestrial) planets and had mission lifetimes of less than 2 years and round-trip communications times of less than an hour. These two considerations most significantly influence the design of the outer planet spacecraft design.

The increased lifetime of approximately 4 years is dealt with by continual application and extension of the techniques used previously, that is, the application of redundancy, elimination of single point failures, use of reliable parts, thorough design analysis, application of previous designs or modifications thereof, and thorough test programs.

The long round-trip communications time has made it necessary to incorporate more on-board intelligence on the spacecraft. The Mariner Jupiter/Saturn 1977 spacecraft has significantly more on-board failure detection and correction capability than previous spacecraft.

The fault detection and correction algorithms used to provide the required spacecraft autonomy reside in the CCS, except for the power and attitude control subsystems, which are shared with the CCS. Figure 12 is a simplified system diagram showing the information sent to the CCS and its command capability over the remaining subsystems. The CCS can execute over 180 commands in controlling the engineering subsystems. The science instruments are controlled by the Flight Data Subsystem; this command capability is on the order of 160 commands. As can be seen in Fig. 12, the only subsystem that is controlled by two sources, the CCS and the FDS, is the Data Storage Subsystem.

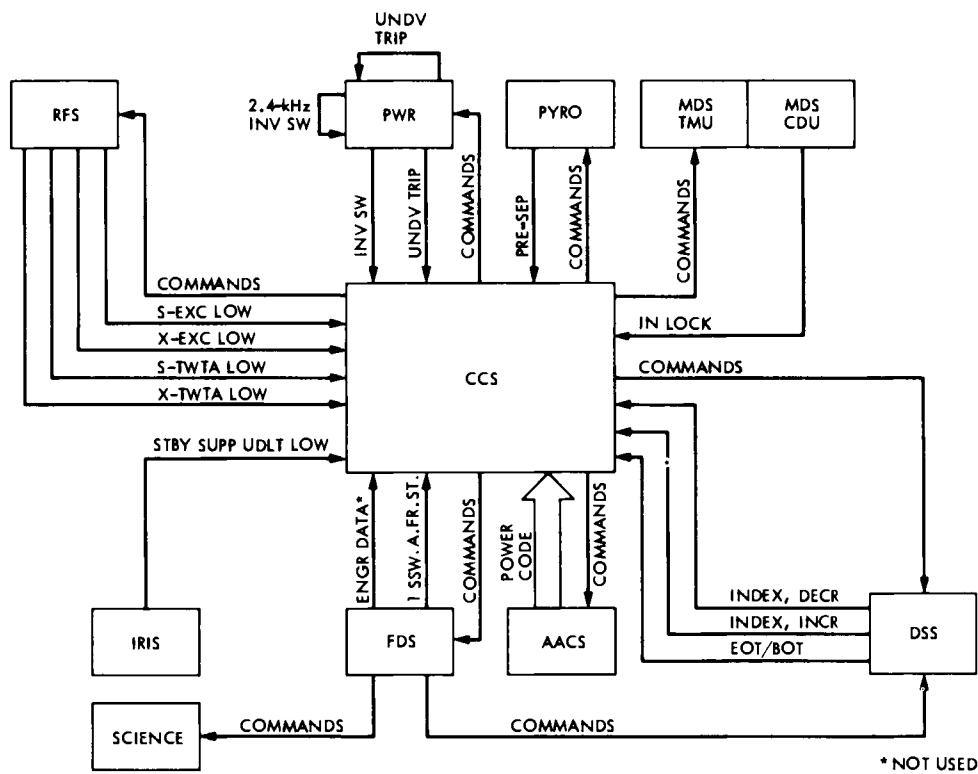


Fig. 12. Mission module mode control diagram.

After a failure is detected, the fault detection and correction algorithms will perform prescribed diagnostic tests and put the spacecraft into a corresponding safe state, allowing a thorough analysis to be made on the ground, after which the spacecraft can be commanded into the desired operational state.

These algorithms are:

- Power recovery
- RF loss
- Command loss
- Attitude control recovery
- IRIS automatic switchover

The power recovery algorithm can detect a system undervoltage condition caused by either an internal power fault or a load-induced fault. Having detected such a condition, the spacecraft will determine whether the fault is in the power subsystem or the loads. If the fault is in the power subsystem, the spacecraft will switch to the backup inverter; if the fault is in a load,

the nonessential loads will be switched off, thus either removing the fault and/or allowing the maximum power margin to blow fuses should the fault persist.

The RF loss algorithm monitors the output of the switchable RF elements. If the output power drops below a prescribed level, the affected element is switched off, and the redundant element is switched on.

The command loss algorithm protects against the loss of commandability of the spacecraft. Six serial elements are required to receive and process a command on board the spacecraft; all of these elements are block redundant and switchable from either ground or on-board commands. The command loss algorithm expects a command to be processed on board the spacecraft at least once during a prescribed interval (e.g., every 32 hours). If a command is not processed during this interval, the command loss algorithm is invoked, and the redundant elements are systematically switched, according to the sequence shown below, until a command is successfully processed.

- (1) Command Detector Unit B Select
- (2) Receiver 2 Select
- (3) Command Detector Unit A Select
- (4) Low-Gain Antenna Select
- (5) Command Detector Unit B Select
- (6) Receiver 1 Select
- (7) Command Detector Unit A Select
- (8) Attitude Control 2 Select
- (9) Command Detector Unit B Select
- (10) Receiver 2 Select
- (11) Command Detector Unit A Select
- (12) High-Gain Antenna Select
- (13) Command Detector Unit B Select
- (14) Receiver 1 Select
- (15) Command Detector Unit A Select
- (16) Attitude Control 1 Select
- (17) S-Band Low Power
- (18) X-Band Low Power

In addition to providing protection against a command processing element failure, the command loss algorithm also protects against a radio self-lock situation, in which the radio locks up on its own downlink signal and ignores the ground-generated uplink signal. To circumvent such a possibility, the command loss algorithm sequences the radio through nine different configurations. These radio configuration changes are contained in an outer loop of the algorithm (as shown below), while the other switched uplink elements are contained in an inner loop of the algorithm.

- (1) S-Band Exciter 1 Select and TMU End Circuit Command
- (2) S-Band TWTA 2 Select

- (3) S-Band Exciter 2 Select and TMU End Circuit Command
- (4) S-Band TWTA 1 Select
- (5) X-Band Power Off
- (6) S-Band TWTA Off and S-Band Exciter Off
- (7) S-Band Exciter On and S-Band TWTA On
- (8) X-Band Power On and USO Off
- (9) USO On

The attitude control algorithm will maintain the spacecraft three-axis stabilized, with the high-gain antenna pointed toward the earth in the event of a failure within the attitude control subsystem. The redundant elements under control of this algorithm are:

HYBIC power supply
HYBIC
Sun sensor
Canopus star tracker
Processor power supply
Flight control processor
Memory
Gyros

The IRIS automatic switchover algorithm is used to switch in a backup heater controller power supply. The heater is used to maintain the required thermal environment of the Infrared Interferometer Spectrometer (IRIS) telescope. Should this thermal control be lost in flight, the telescope would cool to 195 K and thermally distort itself beyond recovery before ground action could be taken. Therefore, this algorithm is designed to maintain the output of the heater controller power supply and, in the event of a failure of that supply, take the necessary on-board action to switch on the backup power supply. This is a classic example of the types of problems which the long round-trip communication time causes. In an inner planet flight, this failure could be detected on the ground and the standby unit switched in ample time since the time constant is approximately 2 hours.

During the design process of the Mariner Jupiter/Saturn spacecraft, there is a continually ongoing effort to identify all single point failures in the system. A single point failure is any failure of any component. The single point failures are identified and then eliminated by the application of either block redundancy or functional redundancy. The redundancy can be at the component level (redundant fuses, quad relays), at the functional level (redundant oscillators, redundant gyros), or at the subsystem level (redundant radios, redundant computers). In some cases, the single point failure is acceptable; these cases are identified and consciously approved by the Project manager. Such single point failures in the Mariner Jupiter Saturn design may occur in

Structure
Microwave components
Injection propulsion unit engines and drivers

Scan actuators, drivers, and heaters
Selected cabling
Injection stage solid motor
Propellant tank and lines
Thermal blankets and shield
Micrometeoroid shields
Radioisotope heaters
Deployment mechanisms
Data storage subsystem
S/X-band antenna

From an inspection of this list, it can be seen that the exemptions are in the area of mechanical and electromechanical devices, where redundancy in general is extremely difficult to implement. It should also be noted that previous flight experience with these devices has been excellent.

The reliability diagram of the Mariner Jupiter/Saturn spacecraft is shown in Fig. 13. This system configuration has been used in the calculation of the overall system reliability. Unfortunately, it is overly conservative in such an application because it does not properly represent the degraded modes of operation available or the ways in which the functional redundancy can be used. These modes are so varied and numerous that they cannot be represented diagrammatically. On the other hand, this diagram is most informative in identifying the block redundancy and its switching implementation. The diagram also illustrates the location and type of single point failures and their effect on the overall system.

The diagram has been divided into major subsystems for convenience; the overall system is a serial linking of these figures.

Four basic types of interface redundancy are employed on the Mariner spacecraft; these are shown diagrammatically in Fig. 14. Each interface is reviewed relative to suitability for a particular type of interface redundancy.

The simple channel interface is used where there are two circuits, each uniquely identified in the system functional block diagram. This type of redundancy is used on all the redundant science interfaces with the FDS; it is also used for the 2.4-kHz power into CCS A and CCS B, for example. There are 202 such interfaces on the spacecraft.

The half cross-strapped interface is used where there is one circuit uniquely identified in the system functional block diagram. There are 72 such interfaces on the spacecraft.

The full cross-strapped interface is used where there is one circuit uniquely identified in the system functional block diagram. Here the actual cross-strapping is implemented in the cabling and can be accomplished in one of three ways, depending on the particulars of the interfaces and their application. There are 43 such interfaces on the spacecraft.

The 2×4 channel interface is used where there are four circuits, each uniquely identified in the system functional block diagram. Here there is no cross-strapping in the cabling; it is all accomplished in the end circuits themselves. There are 16 such interfaces in the spacecraft. Table 7 summarizes the total number of interface circuits and the type of interface redundancy used.

TABLE 7 Interface Redundancy Summary

Total interface circuits	1012
Science interface circuits	358
Simple channel	49
Engineering interface circuits	654
Simple channel	153
Half cross-strapped	72
Full cross-strapped	43
2 4 Channel	16

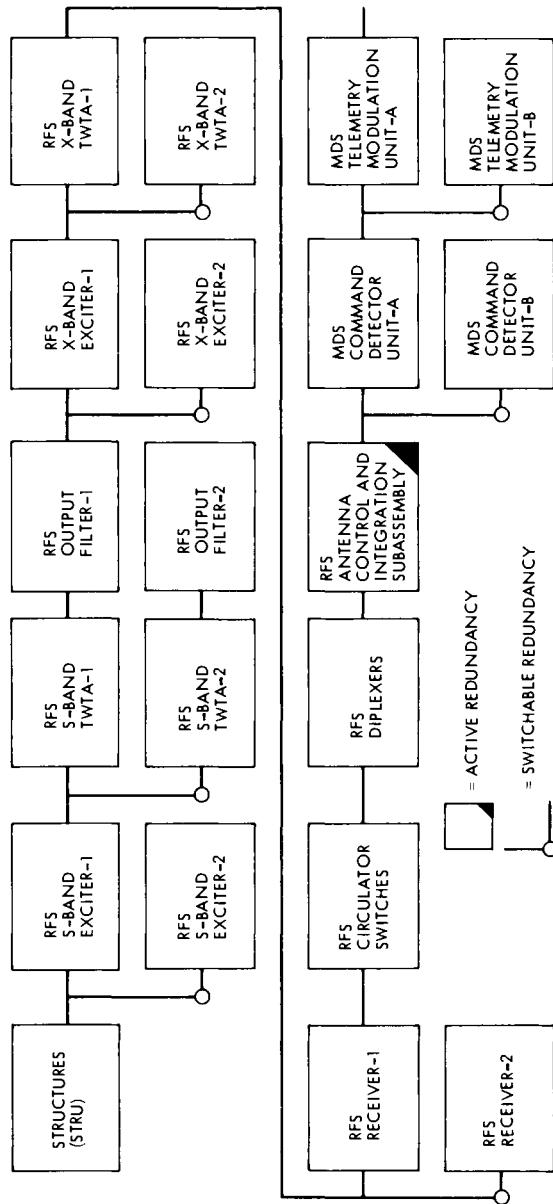


Fig. 13. Mariner Jupiter/Saturn 1977 reliability block diagram.

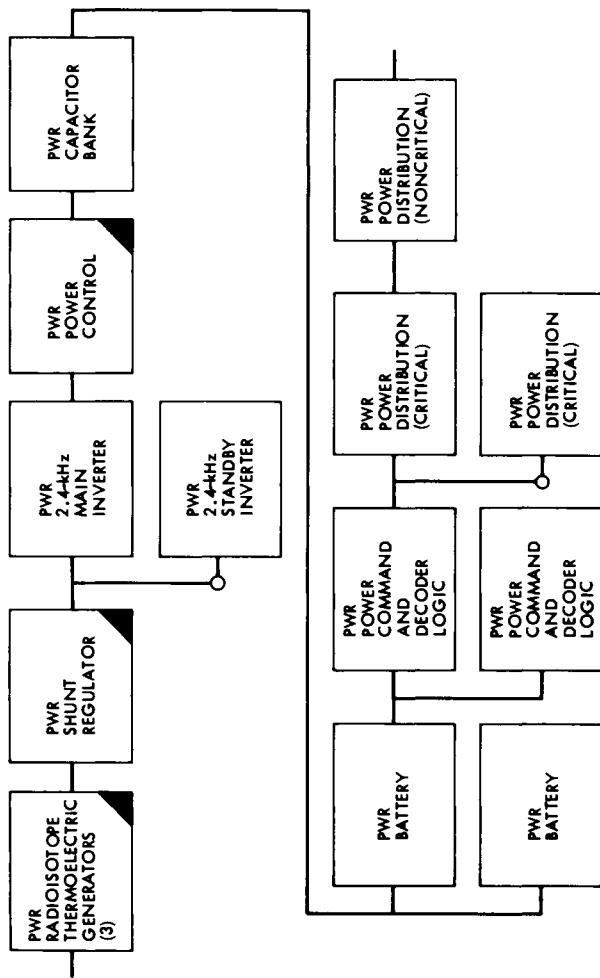


Fig. 13 (contd)

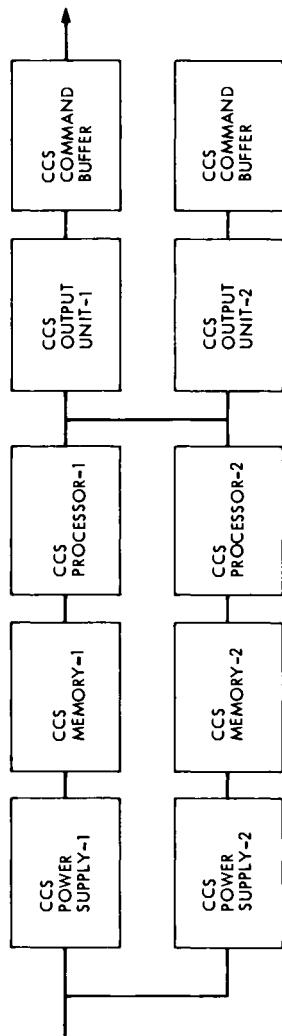


Fig. 13 (contd)

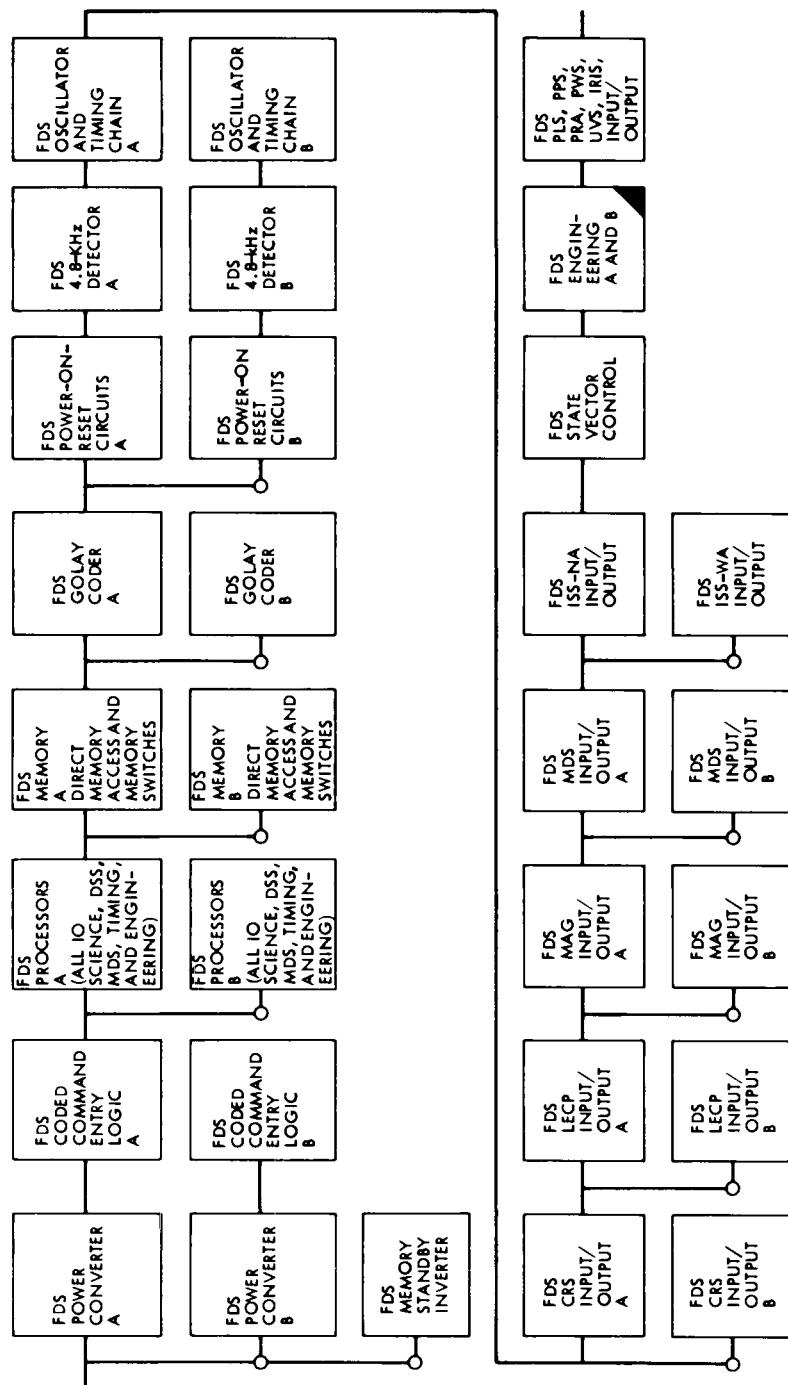


Fig. 13 (contd)

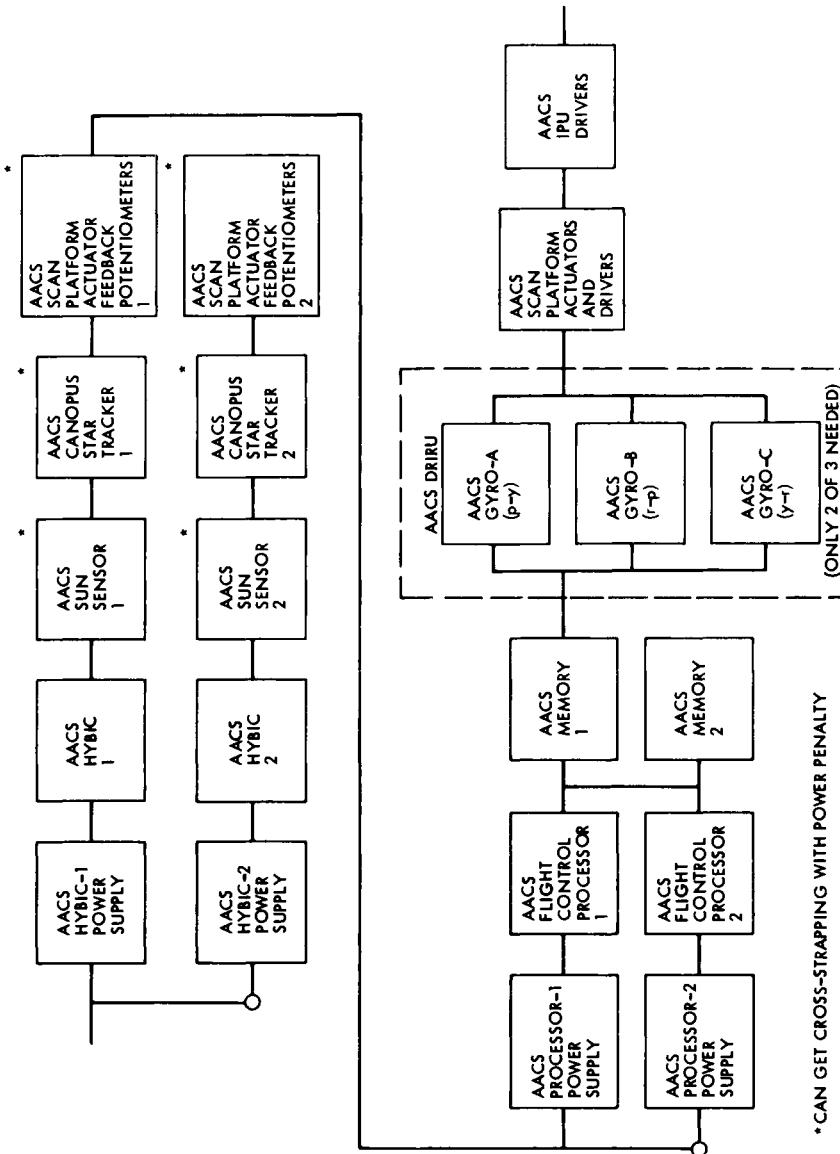


Fig. 13 (cont'd)

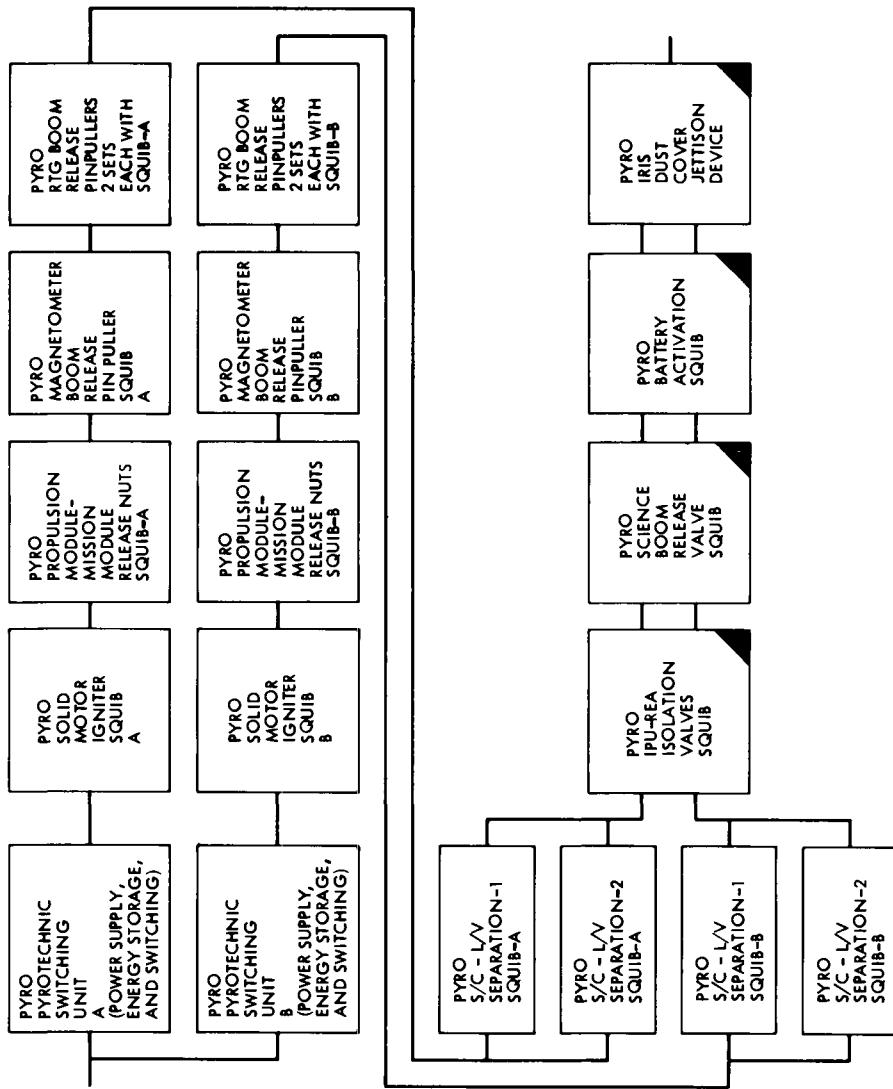


Fig. 13 (contd)

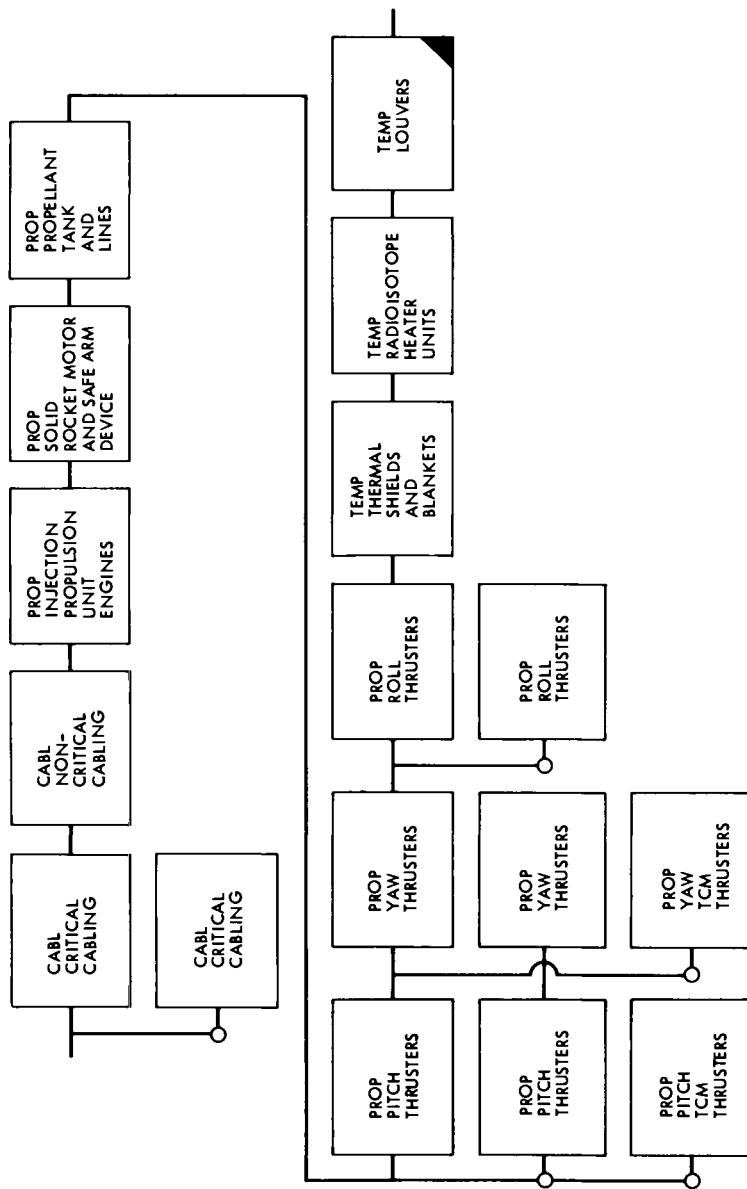
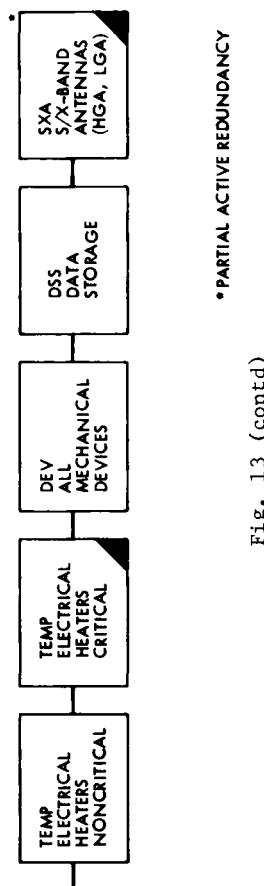


Fig. 13 (contd)



* PARTIAL ACTIVE REDUNDANCY

Fig. 13 (contd)

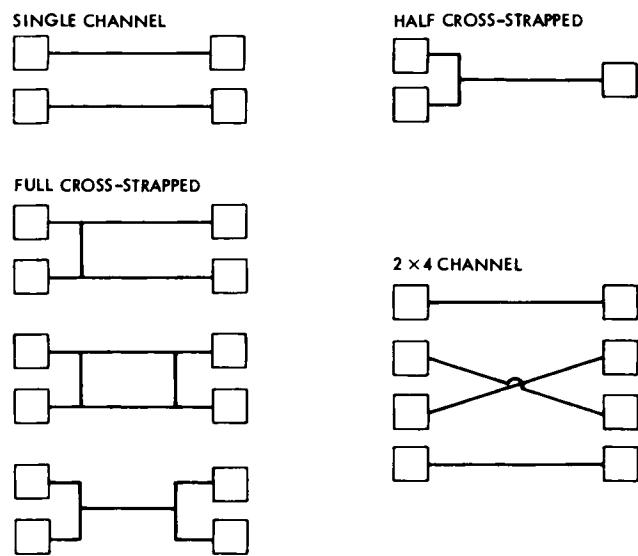


Fig. 14. Types of interface redundancy.

ABBREVIATIONS AND ACRONYMS

AACS	Attitude and articulation control subsystem
ANT	Antenna
BAT	Battery
BOL	Beginning of life
BOM	Beginning of mission
BOT	Beginning of tape
B/U	Backup
CABL	Cabling subsystem
CC	Coded commands
CCS	Computer command subsystem
CDU	Command detector unit
CHRGR	Charger
CMD	Command
CRS	Cosmic ray subsystem
CS/S	Cruise sun sensor
CST	Canopus star tracker
DC	Discrete command
DEV	Mechanical devices subsystem
DRIRU	Dry inertial reference unit
DSS	Data storage subsystem
DTR	Digital tape recorder
ENG	Engine
EOM	End-of-mission
HGA	High-gain antenna
HYBIC	Hybrid buffer interface circuit
HYPACE	Hybrid programmable attitude control electronics
INV	Inverter
I/O	Input/output
IPU	Injection propulsion unit
IRIS	Infrared interferometer spectrometer and radiometer subsystem
ISS	Imaging science subsystem
LECP	Low-energy charged particles subsystem
LGA	Low-gain antenna
MAG	Magnetometer subsystem
MDS	Modulation demodulation subsystem

MM	Mission module
NA	Narrow angle
PCE	Power conditioning equipment
PLS	Plasma subsystem
PM	Propulsion module
POR	Power on reset
PPS	Photopolarimeter subsystem
PRA	Planetary radio astronomy subsystem
PROP	Propulsion subsystem
PSU	Pyrotechnic switching unit
PWR	Power subsystem
PWS	Plasma wave subsystem
PYRO	Pyrotechnic subsystem
RFS	Radio frequency subsystem
RSS	Radio science subsystem
RTG	Radioisotope thermoelectric generator
S/C	Spacecraft
SCAN	Science platform
SCI	Science
S/S	Subsystem
STRU	Structure subsystem
SYS	System
TC	Temperature control
TCM	Trajectory correction maneuver
TEMP	Temperature control subsystem
TLM	Telemetry
TMU	Telemetry-modulation unit
TWT	Traveling wave tube
TWTA	Traveling wave tube amplifier
UNVD	Undervoltage voltage detector
UVP	Ultraviolet photometer subsystem
UVS	Ultraviolet spectrometer subsystem
WA	Wide angle

DEVELOPMENT OF THE EXPERIMENTAL HYBRID ROCKET BARBARELLA—AN EXAMPLE OF CLOSE COOPERATION BETWEEN A STUDENT ASSOCIATION AND A UNIVERSITY INSTITUTE

Robert H. Schmucker* and Winfried M. Schauer**

**Dr.-Ing., Privatdozent, Head Research Section, German Aerospace Establishment
—Institute of Chemical Rocket Propulsion (DFVLR — ICR), 7101 Hardthausen
—Lampoldshausen, Germany*

***Dipl.-Ing., Institute of Astronautics, Technical University, Munich, 8000 Munich, Germany)*

ABSTRACT

On March 12, 1974, the experimental hybrid rocket BARBARELLA was launched from a platform in the Baltic Sea. This launch represents the culmination of a research program at the Technical University of Munich (TUM), which was started in 1964. In that year, a student association with the name "Wissenschaftliche Arbeitsgemeinschaft für Raketentechnik und Raumfahrt (Scientific Study Group for Rocketry and Astronautics)-WARR" was founded at the Technical University in Munich, Germany. In 1966, the Institute of Astronautics at the TUM was founded and a close co-operation developed between the student association WARR and the University Institute. After 1966, the student activities were carried out under the supervision of the Institute, so that the group members were able to combine academic training and independent research, thus making their activities more effective. In 1968 the BARBARELLA program - the development of an experimental hybrid rocket - was started, and student activities were concentrated into three fields: basic research of various phenomena in hybrid rockets, especially those related to the vehicle; development of the various components of the rocket; and systems studies for the vehicle design. The development of the vehicle took 5 years. As a result of the program, many study papers, a dozen masters theses and about the same number of Institute reports, a Ph.D. Thesis, and a book were published. The BARBARELLA program was supported by various other organizations, such as the Institute of Propulsion and Institute of Fluid Mechanics of the Technical University Munich, the German Aerospace Establishment (DFVLR), the German Society for Aeronautics and Astronautics (DGLR) and the Department of Defense, Procurement Office (BWB). The successful launch of the vehicle, in which all goals were achieved, demonstrates the effectiveness of a program based on close cooperation between a student association and a University Institute.

* Dr.-Ing., Privatdozent, Head Research Section, German Aerospace Establishment – Institute of Chemical Rocket Propulsion (DFVLR – ICR), 7101 Hardthausen-Lampoldshausen, Germany

** Dipl.-Ing., Institute of Astronautics, Technical University Munich, 8000 Munich, Germany)

INTRODUCTION

The following report is a documentation of years of research and development work carried out by the "Wissenschaftliche Arbeitsgemeinschaft für Raketentechnik und Raumfahrt (Scientific Study Group for Rocketry and Astronautics) - WARR" in the field of hybrid rocket propulsion. The activities carried out in cooperation with the Institute of Astronautics at the Technical University of Munich range from basic research to the development of an experimental rocket which was launched successfully on March 12, 1974. WARR, the Scientific Study Group for Rocketry and Astronautics, was founded in 1964 by students of the Technical University of Munich as a student group of the Deutsche Gesellschaft für Raketentechnik und Raumfahrt (German Society for Rocketry and Astronautics) DGRR e.V.

Objectives. The aim of this association was to carry out both theoretical and experimental activities in the field of astronautics, particularly as, at that time, there were hardly any specialized Institutes for Astronautics at the Technical University of Munich. For this purpose seminars were held in cooperation with industry and research establishments, theoretical studies were made in the various sectors of astronautics and small-scale experiments were carried out with solid propellant, liquid propellant and hybrid rocket engines. The following educational aspects were the main consideration during these studies:

Academic education by means of theoretical studies in the following areas: Basic science, Engineering, Rocketry and Astronautics and Computer science.

Gathering of practical experience in: System analysis, Planning of missiles and spacecrafts, Design, construction and production, Test and qualification, and Project management.

In order to realize these objectives, cooperation was sought with various University Institutes, to give the group members the opportunity of preparing and publishing independent scientific reports, particularly with theoretical content, under the supervision of these Institutes. At the same time, the support of institutions such as the Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt e.V. (German Aerospace Research Establishment), the Deutsche Gesellschaft für Luft- und Raumfahrt (German Aerospace Society), the German Forces with air bases and test fields, and many German industrial companies made it possible to begin an experimental program on a larger scale. The main objective of WARR was to promote the work of the members in the

framework of a total concept and to give support by means of its contacts to relevant organizations and institutions. In 1966 the group became associated with the Institute of Astronautics, which was established in that year under the direction of Prof. Dr.-Ing. Harry O. Ruppe. The, at first, loose contacts, restricted to the supervision of study papers and various discussions or engaging members of the group as auxiliary assistances, soon developed into intensive cooperation which must be regarded as the key to the success of the BARBARELLA Project. Here was not to be found the normal type of cooperation between students and lecturers, whereby the students are set a task by the lecturer and, under his direction, work on a clearly defined subject. WARR members' activity was truly independent. The Chairman/academic advisor, in his capacity as member of the Institute, as far as supervising study papers and masters theses were concerned, however, he was not bound in his decisions and activities within the group, and was only responsible to the group members. Thus, group activities were carried out solely by group members, at their own responsibility. In order to accelerate the work in keeping with the program, it was necessary to persuade a small permanent body of members into agreeing to remain active within the group for several years, so that new members and people interested in membership could easily be made familiar with past results, and also in order to preserve the basic characteristics of the original course.

Publications concerning the BARBARELLA Project. The results gathered during the course of the project and which are directly concerned with the activities have, to a certain extent, been published by the Institute of Astronautics in various reports, as study papers and masters theses, and have appeared in the relevant journals. A total of 46 reports were prepared, characterizing the amount of work which was involved.

The studies covered the following areas: missile program BARBARELLA with vehicle, components and system studies, basic research in the field of chemical rocket engines under special consideration of hybrid propulsion and, finally, further studies for larger vehicles. The main reports are the following:

Project BARBARELLA: Vehicle

- Schauer, W., : Experimental Hybrid Rocket BARBARELLA
20. IAF-Congress, Brussels, 1971
- Schauer, W. : The history of hybrid rocket
Schmucker, R. Herion Nachrichten, 13, (1974)
- Schauer, W., : Experimental hybrid rocket BARBARELLA
Schmucker, R. LRT-TUM-21 (1974)
- Schauer, W. : The development of an experimental hybrid rocket
Schmucker, R. Raumfahrtforschung, 1975, Nr. 1
- Schröder, B. : Aerodynamics of BARBARELLA
LRT-TUM-Masters Thesis 1973

Components

- Besser, H.-L. : Feed system of BARBARELLA
LRT-TUM-Study, Report SA 74
- Gesch, H.,
Langkau, R. : Parachute recovery system
LRT-TUM-Study, Report SA 73
- Lechner, A. : Design of launch facilities
LRT-TUM-Study, Report SA 71
- Schauer, W. : Design of integrated tank bottom
LRT-TUM-Study, Report SA 69
- Schmucker, R. : Development of a small hybrid rocket engine
20. IAF-Congress, Brussels, 1971

System studies

- Schauer, W. : Effect of propellant loading of flight
trajectory. LRT-TUM-KB-7 (1973).
- Schmucker, R. : Preliminary design of a hybrid sounding rocket
LRT-TUM-Master Thesis 1967
- Latsch, K. : Outer and internal ballistics of a sounding
rocket. LRT-TUM-Master Thesis 1970

Basic research

- Katterloher, R. : Thrust and pressure measurement
LRT-TUM-Study, Report-SA 69
- Schmucker, R. : Hybrid Rocket Propulsion
Goldmann Verlag, München 1972
- Schmucker, R.,
Schauer, W. : Start transient of hybrid rockets
LRT-TUM-Master Thesis 1970
- Tuzinsky, W. : Optical measurement of regression rate
LRT-TUM-Master Thesis 1970
- Rath, K. : Cooling of a small engine with nitric acid
LRT-TUM-Study, Report 1969
- Schmucker, R. : Cooling principles of small engines
LRT-TUM-TB: 70-2 (1970)

Further studies

- Borufka, H. : Design of a 1000 N test stand
LRT-TUM-Master Thesis 1972
- Schauer, W.,
et.al. : Design of a 2-stage rocket
LRT-TUM-Joint-Study, Report 1971

CHRONOLOGICAL DESCRIPTION OF THE PROJECT

The BARBARELLA project began in 1968 with the development of a rocket engine. This work was based on fundamental theoretical and experimental studies which had been started in 1964 with the formation of the group. Fig. 1 shows the chronological order of the main activities during the preliminary phases and during the project itself.

As rocket engines play a dominant role in astronautics, activities were concentrated, from the very beginning on engine development. Studies of the various types of engines proved that hybrid systems have more advantages to show than the classical solid and liquid propellant engines. Moreover, at that time, the lack of relevant scientific data available left scope for own ideas and made work in that field particularly attractive. All efforts were, therefore, concentrated on this sector.

During the first years, theoretical studies and combustion tests with experimental engines were carried out to clarify fundamental phenomena. The experimental engines were of varying construction and thrust level, used primarily for measuring the regressionrate and for observation of combustion.

Fig. 2 shows these experimental engines.

These studies yielded a fund of information, concerning the working performance of this kind of propulsion and led eventually to the system analysis of an experimental rocket. In order to keep the technical effort and the costs of the development within reasonable bounds, a small rocket was planned. Emphasis was laid on the realisation of a complete system rather than on performance.

The first step to realising this project was the development of a rocket engine for a vehicle. The main problems here were injection and cooling. The first prototype was completed in 1970. In the same year, work began on the other rocket components. The essential elements here were the feed-, regulation- and fuel-system, and the tank structure.

Parallel to this, work began on the construction of a test stand for static firings with an integrated missile. Following this work, which took about a year to complete, the first successful test was carried out. The tremendous amount of experience gathered in the following static tests was applied in the further development of the device.

The rocket used for the test flight represents the third stage of development in the BARBARELLA project. By the end of 1973 two devices of this type were ready for use.

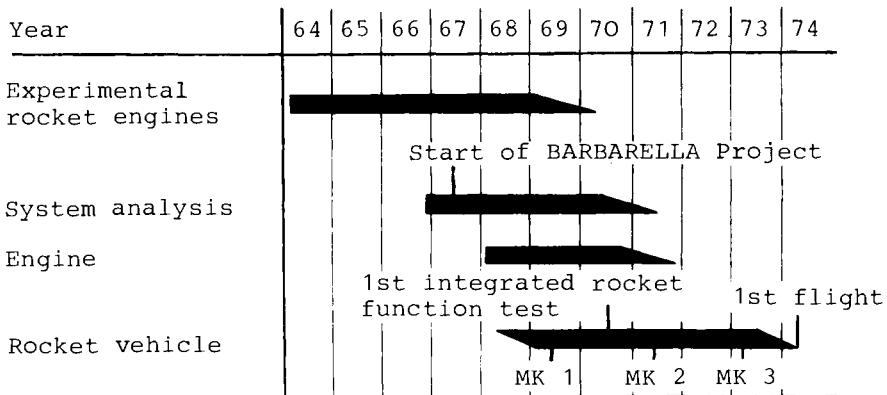


Fig. 1 Main activities during the development program

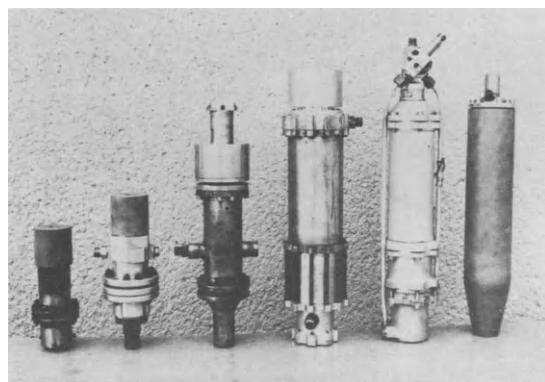


Fig. 2 Experimental engines from the period 1964 - 1968 and two prototypes of the rocket engine

DEVELOPMENT STAGES OF BARBARELLA

The outer diameter of BARBARELLA was fixed at 10 cm, so as to keep the flight performance loss by drag as small as possible. This size was considered as the smallest diameter possible for accommodating valves, regulator feed systems and cooling. Take-off weight should be 8 kg with a thrust of 400 N. A pressurized-gas system was planned to feed the oxidizer.

After a short constructive phase, production began on the MK 1 version which was essentially completed in the middle of 1969. MK 1 was a rocket with separate components, three fins and a structure of corrugated aluminum. However, it became apparent that the planned take-off weight could not be kept to. Problems with the design of the individual parts led to a radical change of the construction principle for the individual components. The separate construction methods of the feed, fuel and injection systems of version MK 1 were abandoned in favour of an integration of these components. The valve configuration was also simplified. Despite the resulting difficulties - apart from the combustion chamber and nozzle, all components had to be newly designed - a new prototype, MK 2, was completed in the summer of 1971. In the same year, a first test stand run was carried out, whereby all components proved their functionability. The information obtained from several test series was used for improvement, and thus a shorter construction length and a further reduction of the weight were made possible. The ramp designed in conjunction with German Forces test station 91 required the fitting of 4 to the vehicle which also resulted in a complete alteration of the fin-fixtures and the shell structure. The corrugated material was abandoned in favor of a smooth coat of which showed a relatively high drag coefficient. The cause, the stepped tail, was removed by adding a suitable shell structure. These modifications led to version MK 3 whose static tests were concluded in the year before launch. In Table 1 are listed the data and characteristics of all 3 versions

Table 1 Data and characteristics of the development versions

Development version	MK 1	MK 2	MK 3
Length (m)	1.31	1.08	1.04
Weight (kg)	12.0	10.5	10.1
Diameter (m)	0.10	0.099	0.099
Construction method	separate		integrated
Number of fins	3	3	4

HYBRID ROCKET BARBARELLA

The main characteristic of BARBARELLA is a hybrid rocket motor. A pressurized gas system is used to feed the oxidizer. As the main objective was the testing of the systems in flight, a payload was dispensed with. Figure 3 shows the rocket.

Concept. The main technical aspects of the concept were as follows: Take-off weight under 10 kg, outer diameter under 10 cm, minimal amount of moving components and simple construction production methods.



Fig. 3 Hybrid rocket BARBARELLA

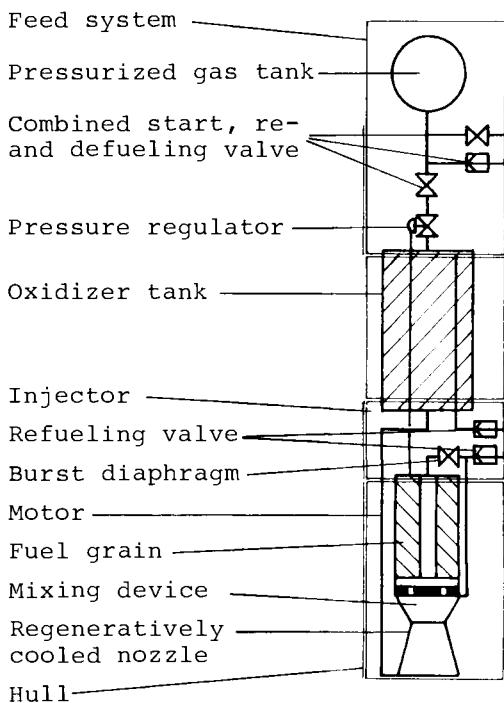


Fig. 4 Diagram of the vehicle

This concept led to an integral construction method with a minimum of valves and remote regulators. For the protection of the test team controlled refuelling systems were planned. Figure 4 shows a simplified diagram of the vehicle with its individual components. The system contains 2 main valves, one to release the pressurized gas, the other in the form of a burst diaphragm for the oxidizer. All tubes are laid inside the structure, are needed for all integral parts and the motor.

INDIVIDUAL COMPONENTS

Figure 5 shows a model of the interior of BARBARELLA with its most important elements.

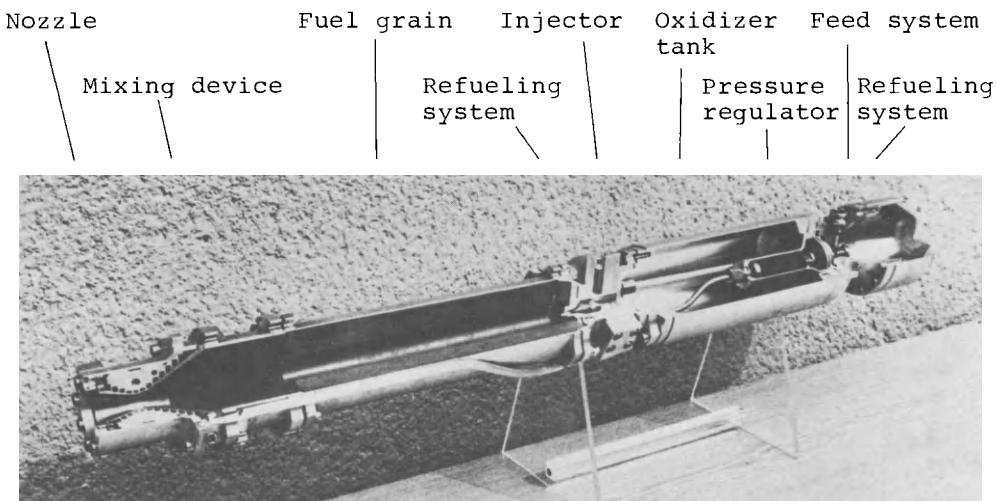


Fig. 5 Cross-section of rocket

Motor. Development work on motor Hy 7 began in 1968. In the years which followed, different versions were tested. These versions differ from each other in injector design, the length of the grain and in the construction of the nozzle. The main materials used were aluminum alloys. The injector is built as integral component. It forms the end-closure to the combustion chamber and is also the bulkhead of the oxidizer tank. Through a port pressure of the combustion chamber regulates the feed-system. The combustion chamber is screwed onto the injector with a flange ring which has been milled off to reduce weight. The grain, which is shown as a hollow cylinder, consists of 4 parts. These are the end grain, the propellant grain, the shell and the ignition mixture.

The propellants are medium energy combinations based on nitric acid. 100 % pure nitric acid is used as an oxidizer. The solid propellant is a mixture of 30 % p-toluidin and 70 % p-aminophenole. In order to increase combustion efficiency, a reaction chamber is added to the combustion chamber, causing the inflowing to react. In order to increase the effect, a mixture plate was installed between combustion chamber and reaction chamber. Due to the high thermal load, these plates are not easily cooled regeneratively. For this reason, a reaction chamber was developed with a material which would absorb the heat. Graphite was ideal because of its high melting point and its low heat conductivity.

One of the main problems during the development of the motor Hy 7 arose due to the cooling of the nozzle. As well as regenerative cooling other types of cooling were examined. The cooling problem is characterized by the following three points: small amount of oxidizer for cooling, high loss of pressure in the

cooling passages and production of the nozzle with helical cooling passage.

Feed system. In the first design phase the feed system was planned as a component with a series of valves and individual parts. This solution was abandoned in favour of an integral construction method. By this means, the problems of the exactness of the longitudinal axis could be solved, more reliability and a considerable weight saving and reducement of length were attained. The same materials were used as for the motor.

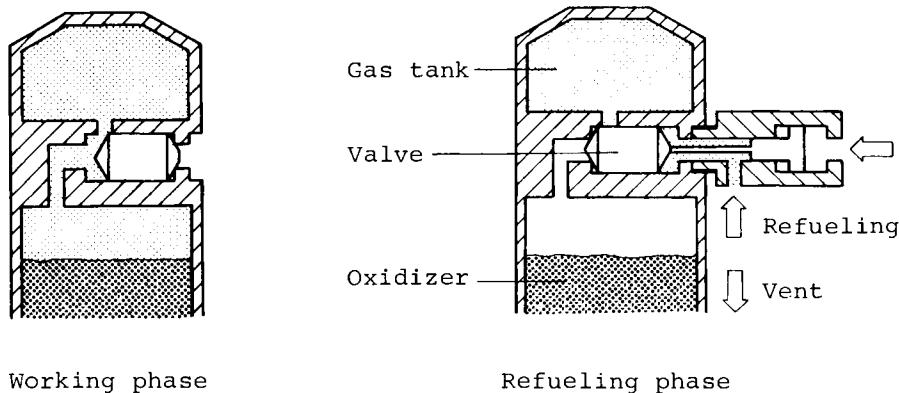


Fig. 6 Diagram of feed gas valve positioning

The pressure in the gas tank is reduced by a regulator mechanism to the pressure in the oxidizer tank. By using the combustion chamber pressure as control, an optimal output pressure can be obtained. That guarantees a low oxidizer mass flow during ignition phase and stabilized working conditions.

The pressurized-gas valve has the function of a pressurization-, (depressurization) relief- and tank valve.

Figure 6 shows diagrams of feed gas valve positioning for the different working phases.

The valve functions are taken over by a balance piston with dynamic seats at the ends, the position of which is controlled automatically from outside and also by gas pressure.

Oxidizer tank. The oxidizer tank consists of a thin-walled cylindrical tank section into which are fitted the combustion chamber pressure line, the relief cock and the end section of the regulator. The tank bulk head is an integral part which contains the injector and the oxidizer refueling unit. Figure 7 shows this in diagram form.

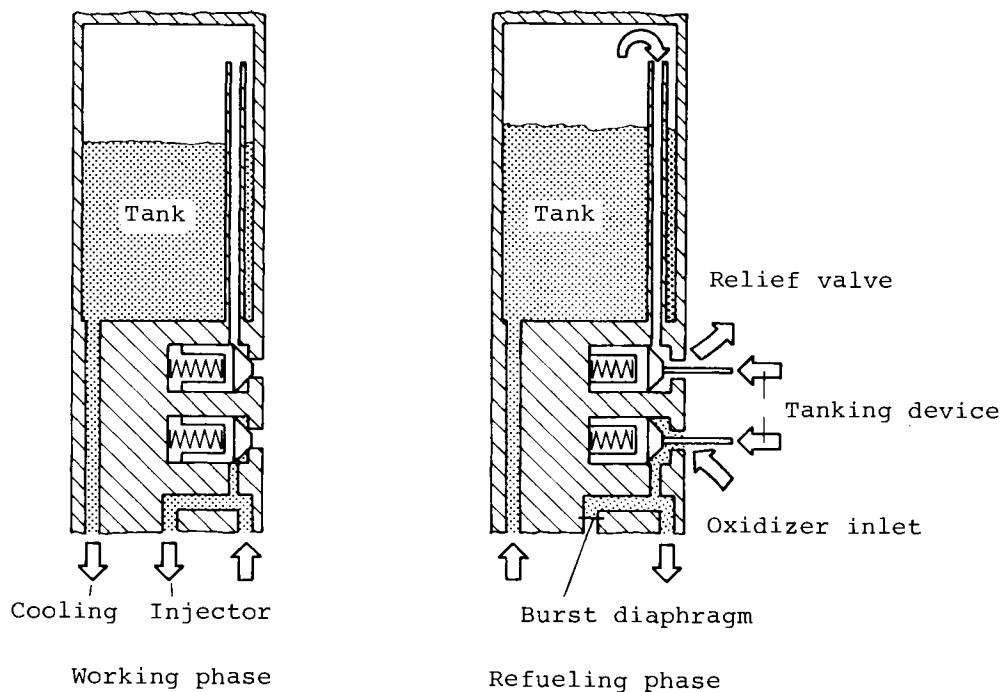


Fig. 7 Diagram of refueling system

Two check valves controlled from outside act as oxidizer and relief connections. The oxidizer refueling valve is constructed in such a way that oxidizer tank, cooling and all main fuel ducts are filled up to the burst diaphragm.

BARBARELLA DATA

The most important BARBARELLA Data are listed in the following table.

Table 2 Rocket Data

Measurements	Weight		
Length (m)	1.04	Total weight (kg)	10.1
Diameter (m)	0.009	Net weight (kg)	6.5
Length of fuel grain (m)	0.275	Propellant weight (kg)	3.6

Table 2 Rocket Data (continued)

Performance data	Pressures		
Thrust (maximum) (N)	360	Chamber pressure (bar)	15
Specific impulse (s)	198	Tank pressure (bar)	42
Total impulse (Ns)	6900	Load pressure of the feed-gas tank (bar)	120
Burning time (s)	22	Burst pressure of the diaphragm (bar)	7.5
		Tanking pressure (bar)	4
		Initial oxidizer tank pressure (bar)	18

Calculations with wind influence, thrust vector- and aerodynamic fin faults show that these dimensions lead to considerable scattering in relation to the nominal course. In order to reduce these deviations to a minimum, the rocket is mounted on a special adjusting device.

GROUND AND TAKE-OFF DEVICES

For the hybrid rocket BARBARELLA a refueling unit was developed that could be used both for the static tests with the complete rocket and for the test flight.

Principle Construction of the Refueling Unit. For re- and defueling of the rocket and starting take-off procedures a control unit was used that could be remotely controlled from a control desk. It contains an oxidizer tank with the corresponding acid-proof valves, a high-pressure nitrogen reservoir with pressure reducers and a series of magnetic valves for controlling the individual functions.

Automatic Refuel Connections. The oxidizer and pressure refueling occurs with the help of two automatic connections.

1. Automatic oxidizer refuel connection. The automatic oxidizer refuel connection which is fitted onto a movable arm contains an oxidizer refueling inlet and a relief outlet. This is shown in Fig. 8 opened, thus forming the connection to the tank.

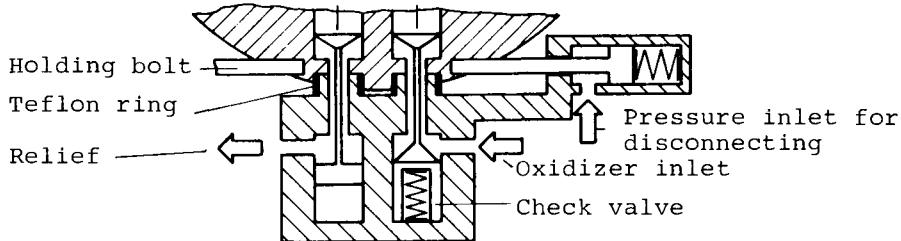


Fig. 8 Diagram of automatic oxidizer refueling system

The sealing of these refuel connections is done by teflon rings. There are also two spring-loaded bolts fitted onto the side which slide into the corresponding part of the rocket and thus hold the refueling unit. After refueling the bolts return pneumatically to their original position. The unit, which has now been released from the rocket vehicle, is pushed away from the rocket by two more pneumatic bolts and swings away from it, due to the effects of gravity. The disconnecting and the pushing away occur simultaneously.

2. Automatic refueling system for pressurized-gas. The automatic refueling system is attached to the rocket filler socket by 2 pivoted clamps. Figure 9 shows a diagram of this device. A pressure-locked piston brings the refueling valve into a position in which the oxidizer tank inlet is closed and the pressurized gas tank inlet opened. After fueling, the clamps are released pneumatically and the pressure in the filler socket pushes the unit away from the rocket. Thus, the inlet is closed to the outside and the oxidizer tank inlet opened. Relief can occur in the same way as fueling. Due to the pressurization of the oxidizer tank, the diaphragm bursts and the ignition sequence starts.

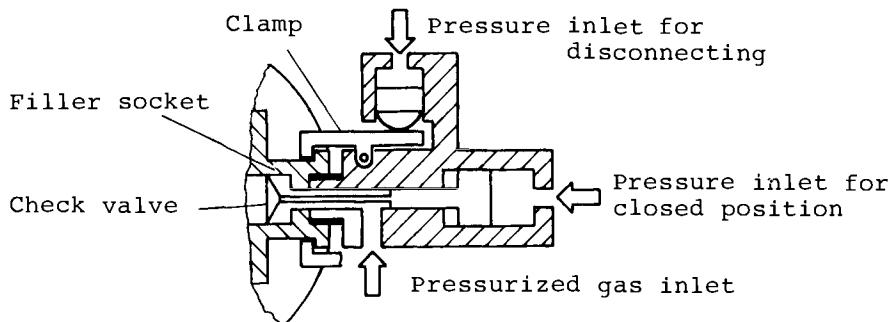


Fig. 9 Automatic refueling system for pressurized gas

ROCKET LAUNCHER

For the test flight, a rocket launcher - adjustable in height - from the German Forces Test Station 91 was used. On this a ramp was constructed, with two slide rails to steer the rocket during lift-off. The distance between the two opposite slide rails and, therefore, the free space between the rocket and the rails is adjustable. The length of the launching ramp - 5 m - was determined after examining flight behaviour during the lift-off phase. Thrust generation of the BARBARELLA engine occurs within 0.2 s. Therefore, the idea of holding the rocket during the ignition phase was abandoned and only the position of the automatic re-fueling unit was fixed.

TEST FLIGHT

The launching of the hybrid rocket BARBARELLA was carried out on the swimming platform BARBARA which was stationed in the Baltic Sea, off the ECKERNFÖRDE Bay. Based on trajectory calculations which effectively took into account not only the normal parameters but also the wind conditions in the Baltic Sea, a test-flight security area of 20 km in diameter was agreed upon. The platform stood 3 km south of the center. The trajectory led north. Launching was performed at an inclination angle of 55°. Figure 10 shows BARBARELLA in the ramp directly before lift-off.

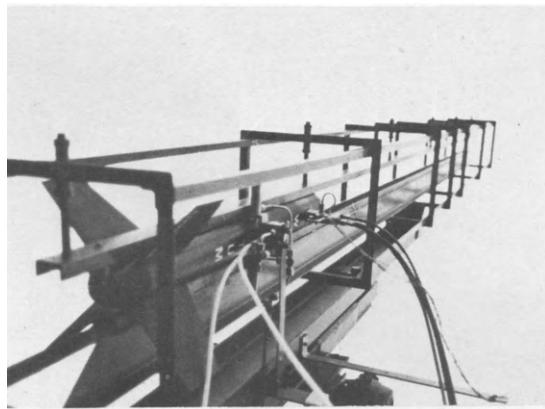


Fig. 10 BARBARELLA in the ramp

At the time of the test flight, there was a lateral ground wind of approx. 12 m/s from an easterly direction. After the disconnection of the automatic pressure refueling unit and perfect engine ignition, the rocket left the launcher without any complications. In the first phase of the flight, which was stable throughout, a slight deviation from the flight course of 20° was noticed. This effect was caused by the relatively strong side-wind. Figure 11 shows a cutting from high speed film. As predicted by the theoretical studies, a rolling moment was observed, which was caused by the momentum of the nozzle cooling and injection systems. There were no disturbances during the ignition phase.

According to optical observations, a burn-out altitude of approx. 700 m, a burn-out range of 3 km and a total range of 7 km were determined.

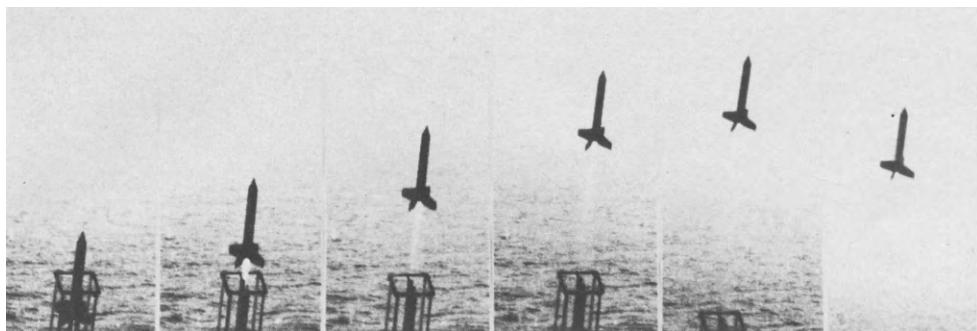


Fig. 11 BARBARELLA leaving the launching ramp
(cutting from a high-speed movie,
time difference 0.04 s)

These flight performance data confirm the theoretic studies, as far as aerodynamic data and flight behaviour are concerned. The various components, which had until then only proved their functionality in static tests, also worked perfectly in flight. The flight test in which all test objectives were fulfilled, confirms the developed concept and shows the unlimited efficiency of the used components.

The successful conclusion of the BARBARELLA program demonstrates the effectiveness of close cooperation between a University Institute and a student group.

THE LARGE SPACE TELESCOPE — A NEW ADVENTURE

Frank M. Friedlaender

Lockheed Missiles & Space Co., Inc., Sunnyvale, California

INTRODUCTION

The Large Space Telescope, illustrated in Fig. 1, will be one of the first scientific payloads to be launched and supported by the Space Shuttle. The telescope will have a near diffraction limited aperture of 2.4 meters and an operational life expected to be longer than 10 years. When this telescope is launched in 1982, it will become a facility with broad scientific capabilities and international significance. It can accommodate four scientific instruments at any given time, permitting spectrographic and photometric observations from 1200 Å in the ultraviolet to 1 mm in the infrared.

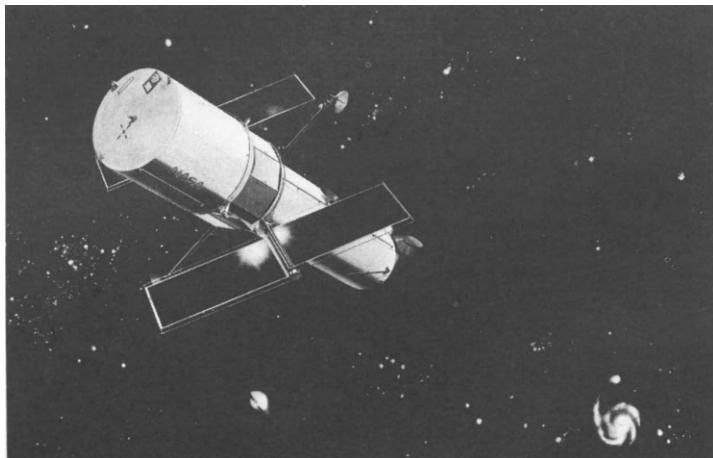


Fig. 1 Large Space Telescope

The capability of this facility exceeds anything that has been used on the ground or in space. Figure 2 shows the broad spectral range of the facility and its limiting magnitude capability in the ultraviolet, visible, and infrared regions. Because the telescope will operate outside of the perceptible atmosphere, the resolution will be significantly

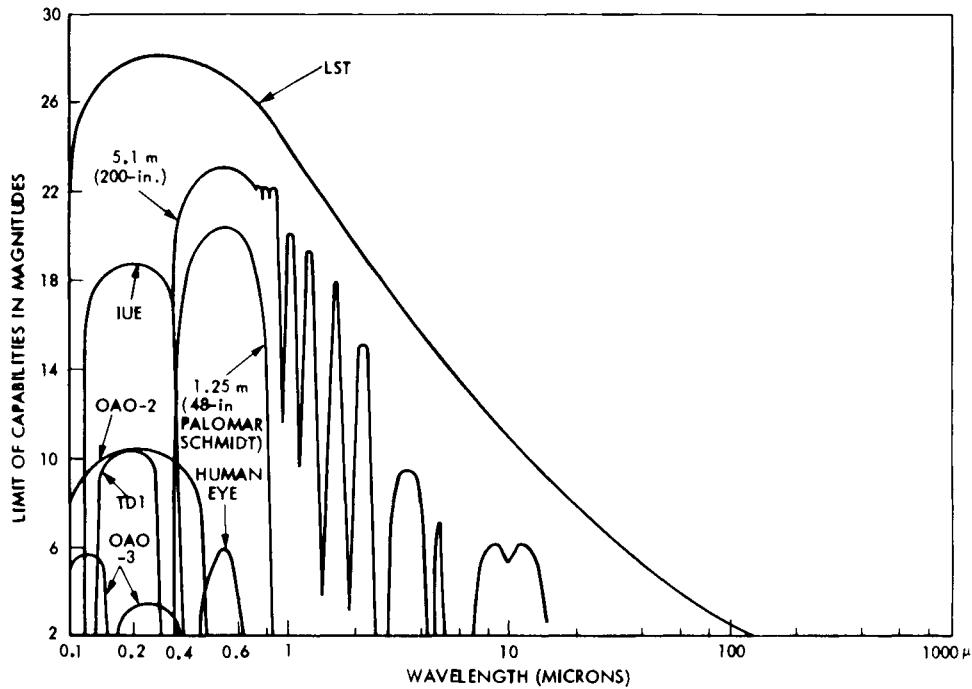


Fig. 2 LST imaging capability

better than that obtainable with a similar ground-based telescope. The physical and environmental conditions are expected to permit observation of stars fainter than the 27th magnitude. The observation of such faint targets must be carefully scheduled to consider the position of the sun and moon, the earth's shadow, natural environmental constraints, effects imposed by zodiacal light, integrated starlight radiance, and effects of traversing through the South Atlantic Anomaly at an altitude of 500–600 km.

The Large Space Telescope (LST) is an assemblage of three basic elements, as shown in Fig. 3. These three elements are the Optical Telescope Assembly, the Scientific Instruments, and the Support Systems Module. The Optical Telescope Assembly consists of the primary and secondary mirrors, their support structure and the internal light shields and baffles. The Scientific Instruments element consists of the focal plane assembly and the primary viewing instruments. Four instruments can be accommodated at any given time. The instruments currently being considered for inclusion are an f24 wide-field camera, an f48/f96 planetary camera, a high-resolution spectrograph, a faint-object spectrograph, a high-speed point/area photometer, an infrared photometer, and an astrometer. The Support Systems Module is the spacecraft bus, which provides power, data management, communications, thermal control, meteoroid protection, stabilization, and pointing.

DEVELOPMENTAL CONSIDERATIONS

The Space Shuttle with its capability to deliver and retrieve payloads from orbit as well as service them in orbit provides a new dimension in operational flexibility for any spacecraft. This capability has a strong impact on the design and the development

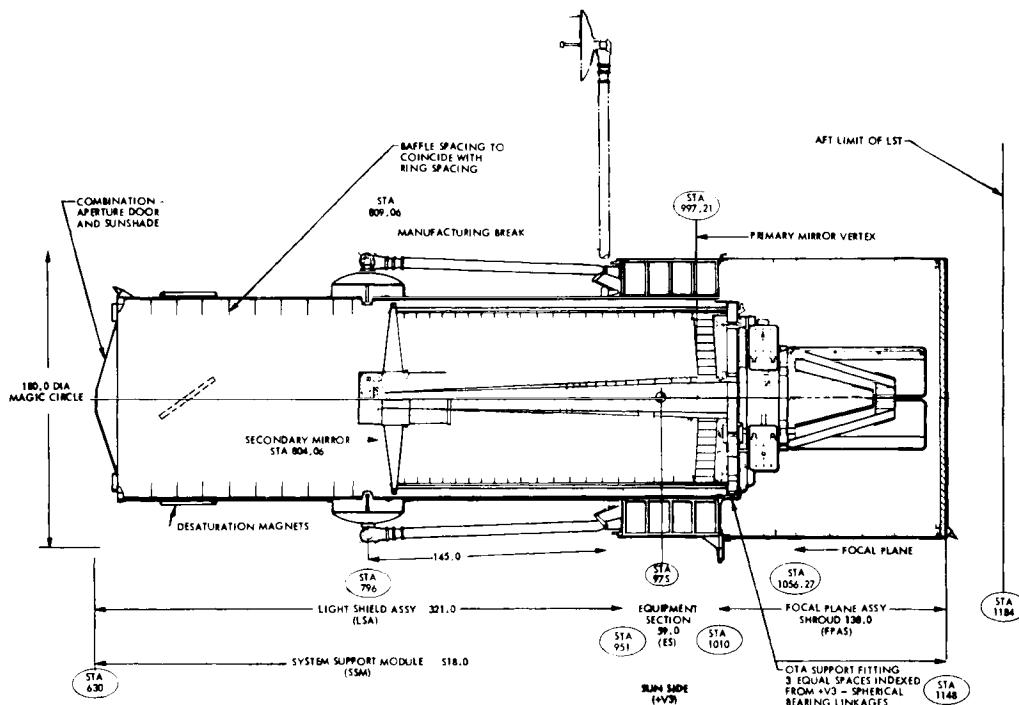


Fig. 3 LST inboard profile

of the LST. One of the first questions to be asked is: Should all of the LST testing be accomplished on the ground as is currently being done, or should some testing be done in orbit where no artificial environment need be provided? To make this determination, another question must be asked: If the testing is done in orbit, can the components that are being tested be replaced in orbit? If they can, the determination should be made on a cost basis. Since there are many levels of testing, the effect of each test must be evaluated separately to establish the total impact.

Closely tied in with testing is the maintenance of the LST during its operational life. The two primary questions are: Where will the maintenance be done? and How? Since the Shuttle is available, the LST can either be maintained in orbit or returned to earth for maintenance. The How of maintenance manifests itself really into the level at which replacements are made. This can be at the component, assembly module, subsystem, or element level. The two parameters Where and How are deeply interactive. For on-orbit replacement, modules are most appropriate. They should be neither too small nor too large. If they are too small, too many operations are required; if too large, they are cumbersome to handle by the astronauts. Ground replacement, however, permits both of these small and large extremes to be handled.

On the LST, a mix of approaches is utilized. Regular on-orbit replacement of modular units is being planned, with less-frequent ground return for major refurbishment, such as recoating of the mirrors. Another question that comes to mind is: Should one buy spare parts for 5, 10, 15, or more years initially. Space parts for a 10-year period could represent 60 to 70 percent of the initial hardware cost. Three significant factors must be considered. The first is the technological life of any given component,

the second, is the shelf life of any given component, and the third is the cost premium that must be paid for short-term procurements. The first two parameters establish the range of consideration, and the third becomes a cost tradeoff to be made for each component, assembly, and subsystem. Instruments, however, fall into a slightly different category, since they can be both interchanged and improved.

An additional factor that affects both the maintenance and logistics questions is that of long-term degradation of electronic components from radiation. Some typical lifetimes for electronic components are shown in Fig. 4. Although the minimum lifetimes of the most sensitive components are on the order of 15 years, the normal distribution of failures and degradation dictates that the replacement of these components be factored into the maintenance policy. A 25 percent planned replacement every 3 to 4 years of all electronic components is probably the prudent approach to this problem. One final question in this sequence is: How should the data flow be handled and when and where should the data be processed? The options are wide and they range from direct data transmission from the instrument to the ground, where processing is performed, to reading-out instruments to an on-board storage buffer, performing some data processing on board, and conducting the final data processing on the ground. It is planned that the data be read out of instruments into an on-board buffer, where it is formatted, and that all of the data reduction be conducted on the ground.

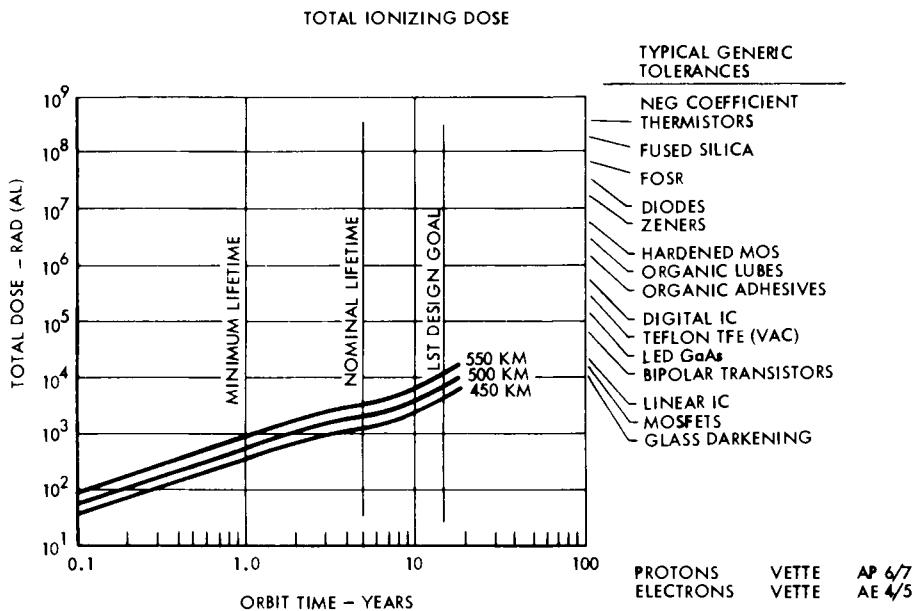


Fig. 4 Radiation effects

OPERATIONAL CONSIDERATIONS

How does one plan for this mission, maximizing the time available for scientific observation, yet minimizing cost? If one dismisses the matter of observing strategy for maximizing viewing efficiency for the available time, one is left merely with

maximizing the rate of available observing time to total clock time. The primary elements to consider, in decreasing order of significance, are maintenance and refurbishment, calibration, checkout, and housekeeping. By far the most important element from a planning point of view is maintenance and refurbishment. Typical downtimes for on-orbit maintenance are on the order of weeks, while for ground refurbishment they are on the order of months. This can be seen in Fig. 5.

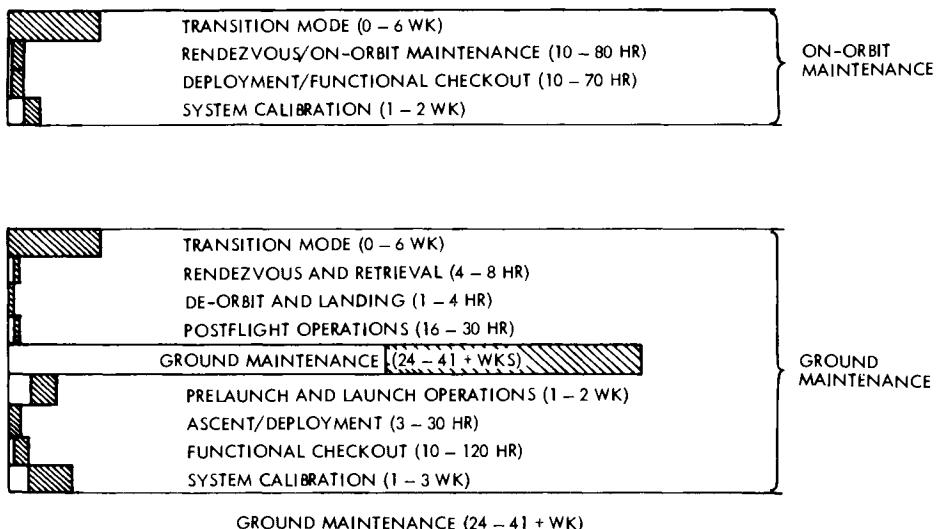


Fig. 5 Ground vs on-orbit maintenance

Once the LST is returned to the ground, the procedural activities add a significant time block to the total time span, even if the same maintenance activities were to take place as during orbit. In practice, a ground return implies additional maintenance/refurbishment activities. Consequently, from a time point of view, one would minimize ground return. From a cost point of view, it is somewhat more complex. It is certainly more costly to design the LST for on-orbit maintenance, crew training, and Shuttle support. On the other hand, with ground refurbishment, there is ground transportation and need for additional facilities. How can one actually determine the costs? Are shuttle costs assigned to the project? Can the Shuttle costs be shared, thereby reducing LST allocated costs? Finally, what is the value of each lost day of scientific observations? If one considers only the direct operating costs of supporting LST scientific operations, then one can determine the daily costs for this and have a measure of effectiveness against which to trade each day lost to observations.

Calibrations relate primarily to the instruments and can therefore be allocated to science time; nevertheless, they should be minimized. They can easily use up one-half of the data capability and one-fourth of the observing time. Two precautions should be taken: instruments should have internal calibration sources covering the spectral requirements of the instruments as much as possible, and external calibrations should be minimized after completion of the initial operation checkout sequence. Internal calibrations can generally be performed without any time penalty — only power and data penalties. Internal calibrations can often be performed simultaneously with an observation if another instrument is used and if power and data capability is

available, or during earth occultation for bright targets and during the sunlit portion of the orbit for faint targets that are observed only in the shadow portion of the orbit. Checkout of all systems is a function of each maintenance action; therefore, total maintenance action, independent of location, should be minimized. Finally, house-keeping actions can be performed during the off-time of the scientific operations; they do not require a separate optimization.

MISSION SCENARIO

A mission scenario can now be synthesized. The major phases of the flight are shown in Fig. 6. After the integration of the instruments into the telescope assembly, the combined payload is mated to the spacecraft bus. The LST is then tested and shipped to the Kennedy Space Center by airplane, where it is received and inspected. Final preparations are then made, it is moved to the launch pad and installed in the Space Shuttle Orbiter Payload Bay, a final verification is made, and the launch countdown is initiated. The LST is essentially passive during ascent, and when the 500-km, 28.8° - inclination orbit is obtained, the deployment sequence is begun. This sequence, shown in Fig. 7, consists of erecting the LST into a vertical position on the Orbiter sill with the Orbiter Remote Manipulator System, acquiring the Tracking Data Relay Satellite with the omni-antenna, commanding the solar arrays into an extended position, extending the high-gain antennas and acquiring the Tracking Data Relay Satellite, and conducting a final verification test. The LST will then be lifted off the sill with the Orbiter Remote Manipulator System, oscillations will be allowed to damp out, and the LST will be released. The LST is now an autonomous satellite under control of the LST Mission Operations Center. The Orbiter will be backed off several kilometers and flown in an escort mode while functional checkout of the LST occurs. After the checkout sequences are completed, approximately one week or less, the Orbiter is released to return to earth and/or perform its next assigned mission tasks.

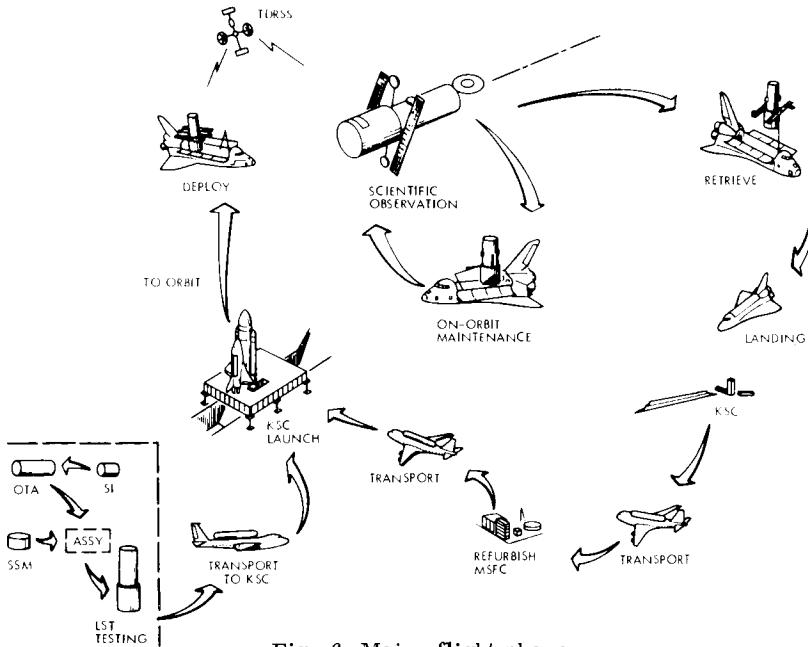


Fig. 6 Major flight phases

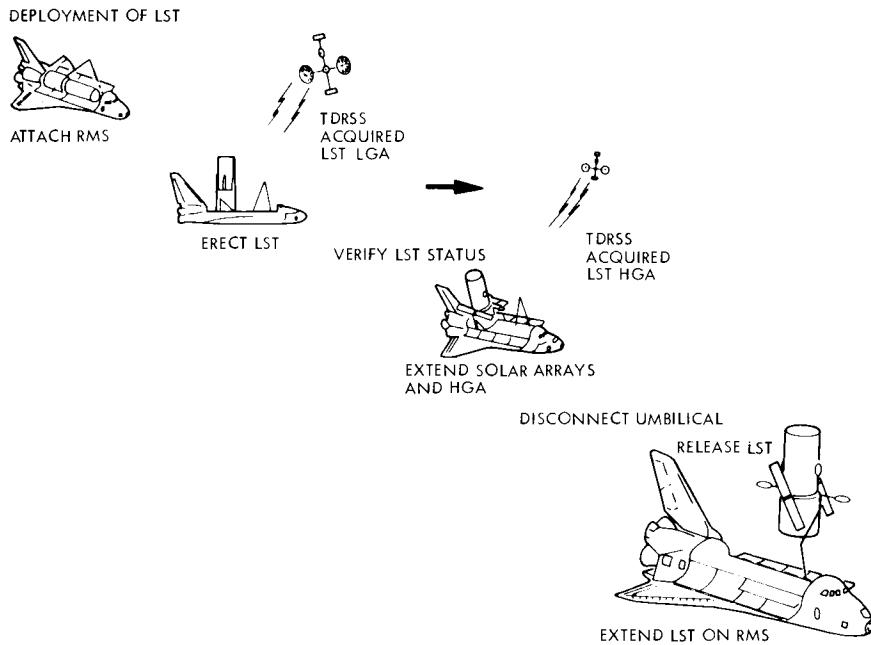


Fig. 7 Deployment sequence

Meanwhile, LST system calibrations will be in process for several weeks, preparing the spacecraft bus and the instruments and optical system for the scientific observations, which will be performed by the facility staff observers, guest observers, or the principal investigators responsible for the instrument development. The observers will be located at a Science Institute, where the experiments are planned and the data processed. The actual control of the LST will be from the Mission Operations Center by controllers and an operations staff (Fig. 8).

After 1 to 2 years of operation, maintenance will be expected. Specific criteria for maintenance action will be determined, a Shuttle flight scheduled, and all equipment prepared and tested. The actual maintenance will be conducted on-orbit, with the LST berthed on the Orbiter Payload Bay sill in a manner similar to the deployment sequence. Subsequent to the replacement of components and possibly the exchange of some instruments by suited astronauts, the LST will be deployed, a functional checkout conducted, and systems calibrations performed as in the original sequence. The LST will then be ready to resume its scientific observations. After the next observation interval, another maintenance action may be required, or a major refurbishment action may be required. Some typical times for this activity are shown in Fig. 9. If a major refurbishment is required, the LST will be returned to earth with the Shuttle. In this case, the LST will be commanded to a dormant mode, the Shuttle sent to the LST orbit, the LST and Shuttle oriented for rendezvous, the LST appendages retracted and, finally, the LST captured by the Orbiter Remote Manipulator System and stowed within the Orbiter Payload Bay for the descent to the Kennedy Space Center.

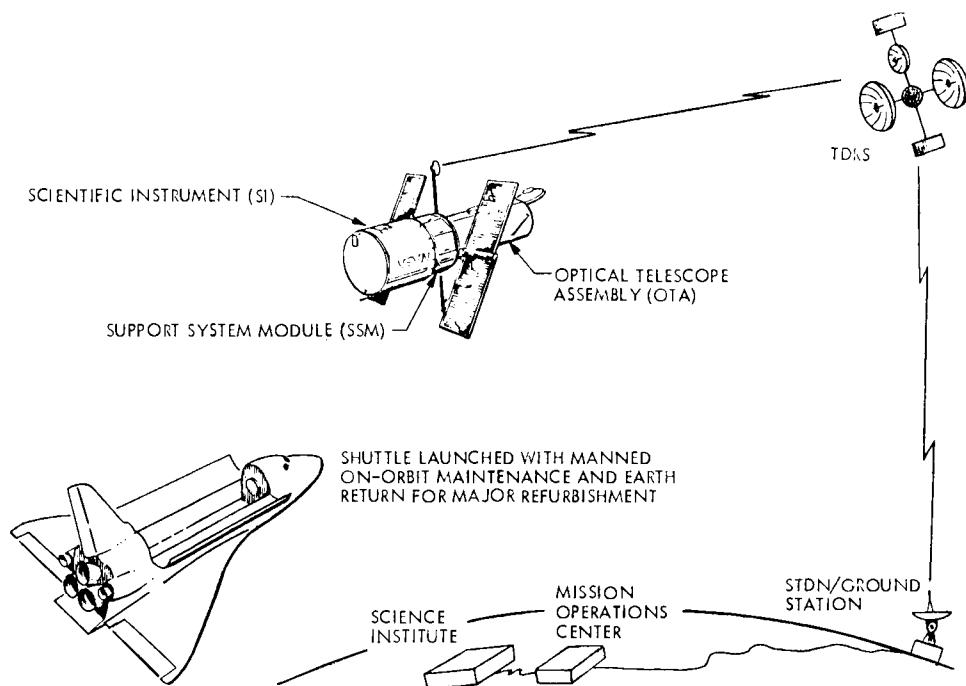


Fig. 8 System elements

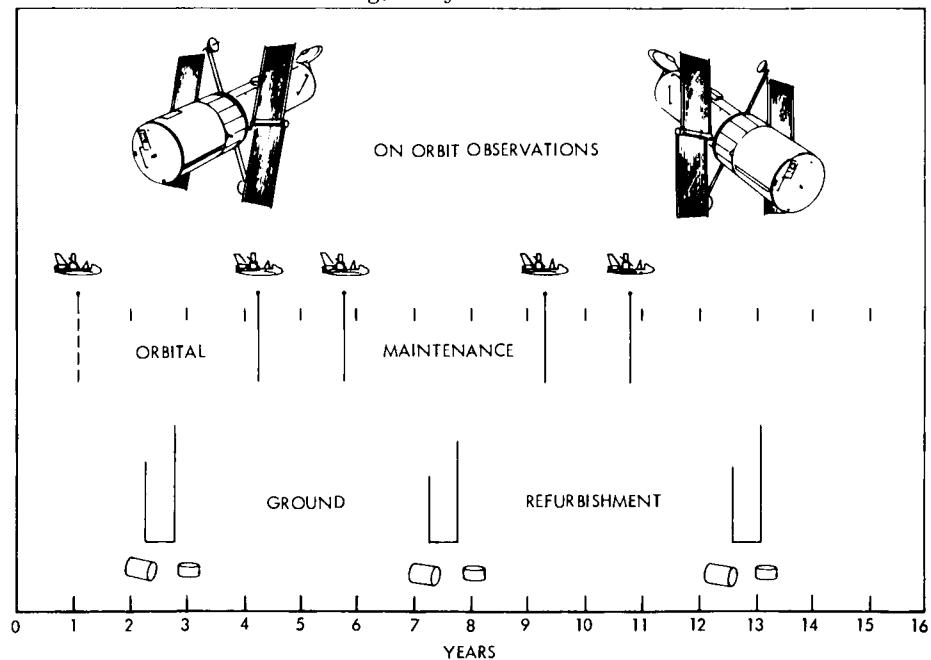


Fig. 9 Life cycle times

After the normal postflight operations sequences for the Shuttle, the LST will be removed from the Orbiter bay, inspected, and sent to the Marshall Space Flight Center for the refurbishment. Transport will again be by air, probably by a Boeing 747 Shuttle Transporter equipped with external pod for spacecraft transport. The LST will then be disassembled into its major elements, which will either be refurbished at Marshall Space Flight Center or, in the case of such equipment as instruments and optics, be sent to the original fabricators. Only major activities, such as re-coating the mirror, changes of interfaces, and other complex repairs will be performed in this mode. A basis will also be provided for observing the environmental effects on long-term spacecraft. When all of the work is completed, the elements will be reassembled, tested, and sent back to Kennedy Space Center. The same sequence as on the initial launch will be followed, and scientific observations resumed. This sequence of events can be sustained almost indefinitely.

CONCLUSION

The LST is planned as a long-term observational facility. What additional opportunities does that provide? One that readily comes to mind is growth - growth in accuracy, data capability, and instrument performance. But why should one design the ultimate facility immediately? Certain aspects of such a facility development are not cost-effective at this time. Both the absolute pointing accuracy and the stability can be improved with time. Another area for consideration is data storage and processing; there is considerable historical evidence that in another decade significant improvements will occur in this area. Data storage and processing instruments will not only be improved but also changed. Finally, a larger-diameter mirror may someday be incorporated. The exact mechanization of the initial LST is only one aspect of making this a useful, growing, and long-lived facility. The system must be designed to allow for changes in both mechanization and operation. The LST promises to be a new adventure, not only in its ability to explore the unknown aspects of the universe, but also to initiate a new era in long-term spacecraft that can be maintained by the Space Shuttle.

MISSION DESIGN FOR A 1980 PIONEER JUPITER ORBITER

Wolfgang Kokott and Walter Müller

Messerschmitt — Bölkow — Blohm GmbH, 8012 Ottobrunn, F.R. Germany

INTRODUCTION AND SUMMARY

As a potential ESA/NASA cooperative project, under ESRO contract an orbiter mission has been studied by MBB as the logical next step toward a further, detailed exploration of the Jovian system. A mission of this type could be launched in 1980 and would involve, to a large extent, hardware already flown successfully. In particular, a spacecraft of modified Pioneer 10/11 design can be equipped with orbiter-specific subsystems of largely European origin, notably a Symphonie-type propulsion unit for insertion into an orbit around Jupiter, and a Helios-type data storage unit. Other necessary modifications include advanced command distribution and power (RTG) subsystems.

Major objectives of this mission are magnetospheric investigations, studies of the inner five satellites of Jupiter and their interactions with the Jovian magnetosphere, and improvements of our knowledge of the gravitational/dynamical properties of the Jovian system.

A feasible mission was found to involve, after entry probe separation, a two-impulse insertion utilizing a 1.8/150 Jupiter radii intermediate orbit (which is needed for fuel economy and some of the scientific objectives, as well as for entry probe data relay). This orbit is, by the second powered maneuver, transformed into a 14/150 Jupiter radii orbit which avoids radiation hazards and provides the opportunity of subsequent encounters with the satellites Ganymede and Callisto, which may be used for modifying the orbit.

PROJECT RATIONALE

After three flyby missions to Jupiter (Pioneer 10, Pioneer 11, MJS — 77), at the end of this decade the phase of tentative first exploration of the Jovian System which is done, literally, "en passant" (Fig. 1) by spacecraft following interplanetary trajectories, will be finished. The logical next step to follow is the orbiter mission to Jupiter, in order to facilitate sufficiently detailed investigations of the planet, its satellites, and its magnetosphere over an extended period of time.

Any planetary mission should, in order to be cost-effective, be based upon

- results of its precursor mission
- existing spacecraft types to be improved and adapted step by step, thus building

- up reliability
- new experiments, subsystems and components which have already been successfully flown on satellite missions.

According to these guidelines, a Jupiter Orbiter mission may

- be launched after MJS — 77 encounter
- use a Pioneer 10/11 type spacecraft
- employ, as far as new subsystems are concerned, flight-proven hardware, e.g. the Symphonie propulsion unit, the HELIOS data storage unit, and VIKING-Prospector type line-scanning imagery.

Optimum performance is, under such guidelines, only feasible if the mission is adapted to the existing facilities rather than vice versa. Only by a repetitive iteration involving parameters of existing equipment, mission geometry, and scientific objectives, a mission can be defined which yields maximum return at minimum cost.

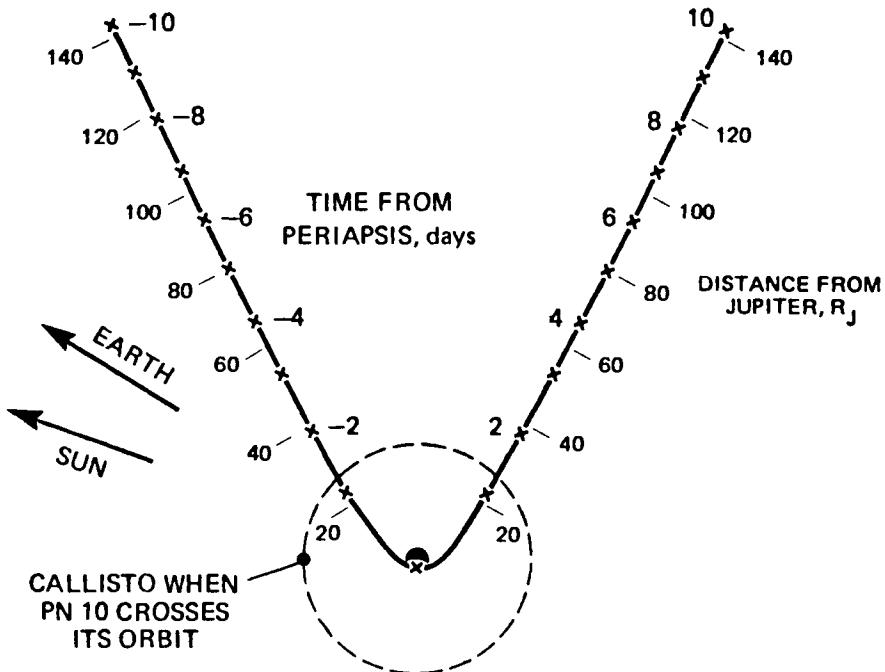


Fig. 1 Pioneer 10 Jupiter swingby

MISSION WINDOW ENERGY

The sequence of launch windows for any Jovian mission is determined by Jupiter's synodic period (399 days). As usual, eccentricities and mutual inclination of the orbits of Jupiter and Earth lead to small but significant changes in energy and geometry of the mission windows from one year to the other. Beside being the earliest launch opportunity which makes sense from a planning point of view, the 1980 window is characterized by somewhat average conditions within Jupiter's 12 year orbital cycle.

Both Earth departure and Jupiter arrival conditions may be considered to be representative for other launch years.

Thus, feasibility conclusions established here may be cautiously generalized as long as no very special, extreme parameter values are involved.

Fig. 2 shows, for an energy-optimum launch window of 15 days duration, departure and arrival hyperbolic excess velocities. As is obvious from the right-hand diagram, arrival energy is for a given arrival date over a quite wide range nearly independent of the launch date. This special feature (which is by no means common to other missions) enables us to assure the impulse needed for orbit insertion at Jupiter to be, for given arrival date and target orbit, nearly constant throughout the launch window. From the left-hand part of the diagram, we find that departure energy is but slightly depending on the arrival date within the four months plotted, and on face value the late (March/April 1983) arrivals are preferred because of their significantly lower arrival excess velocities. However, operational considerations to be discussed below lead to earlier arrival dates for the mission under discussion.

Throughout the window, departure asymptote directions are compatible with Cape Canaveral (ETR) safety restrictions. Arrival asymptotes at Jupiter are, within $\pm 2^\circ$, equatorial.

For the Titan III E/Centaur/TE-364 launch vehicle (Fig. 3), we find for a departure hyperbolic excess velocity of 9.6 km/s a nominal overall payload of 1100 kg. This in turn leads, for a 275 kg Pioneer spacecraft and a 40 kg adapter, to an excess mass of 795 kg which is used

- for the retro propulsion unit, needed for orbit insertion at Jupiter with propellants
- optionally, for an atmospheric entry probe to be included in the mission.

The mass budget is summarized in Table 1.

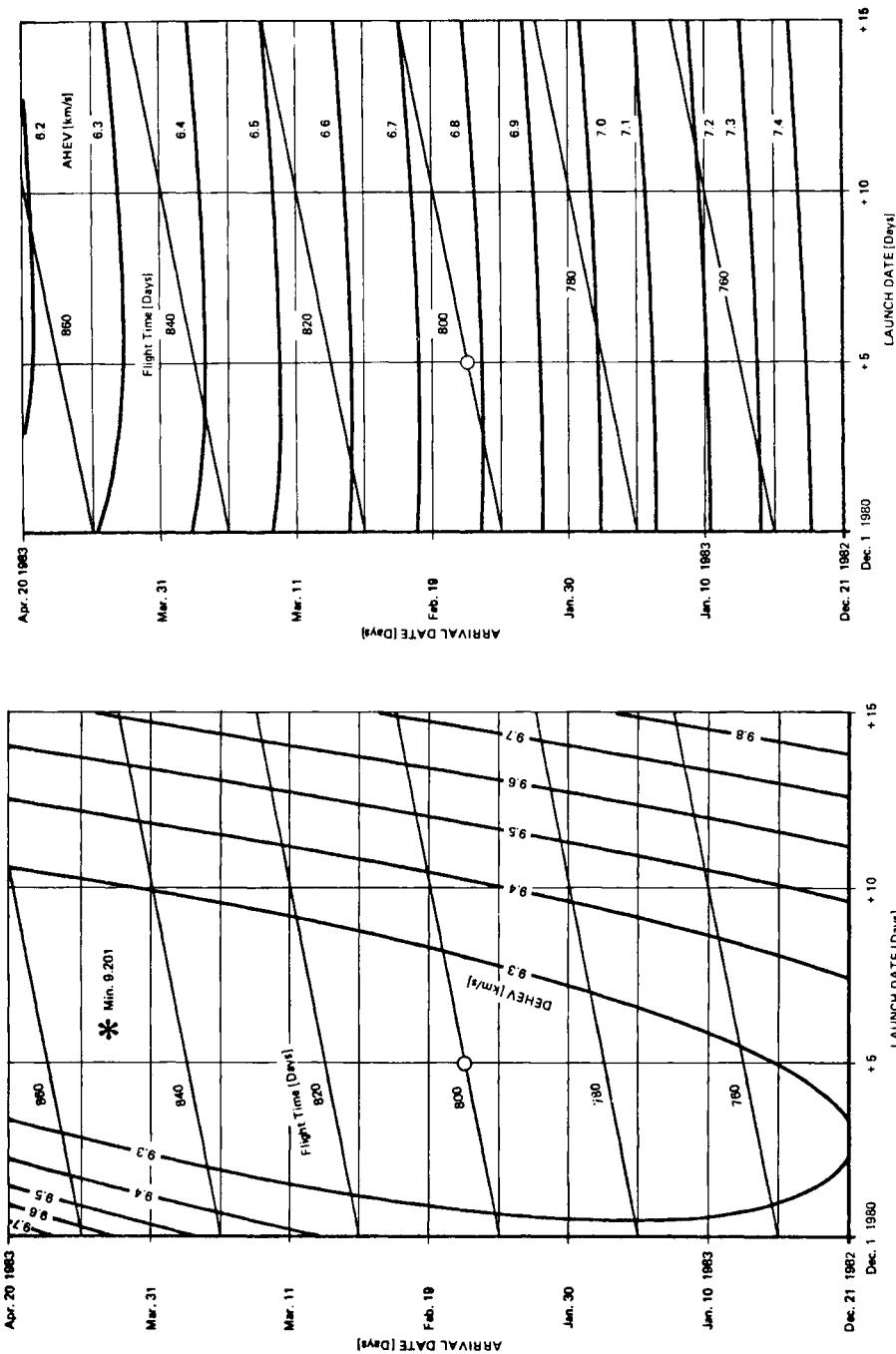


Fig. 2. Departure and arrival energy

TABLE 1 PJO - 1980 mass budget

PIONEER Mass (without prop. system)	275 kg
Probe Mass	147 kg
RPU Net Mass (incl. prop. & Gas residuals)	85 kg
Max. RPU Useful Propellant Load	553 kg
<hr/>	
Separated Spacecraft Mass	1 060 kg
Attach Fitting (incl. telemetry)	40 kg
Launch Payload	1 100 kg

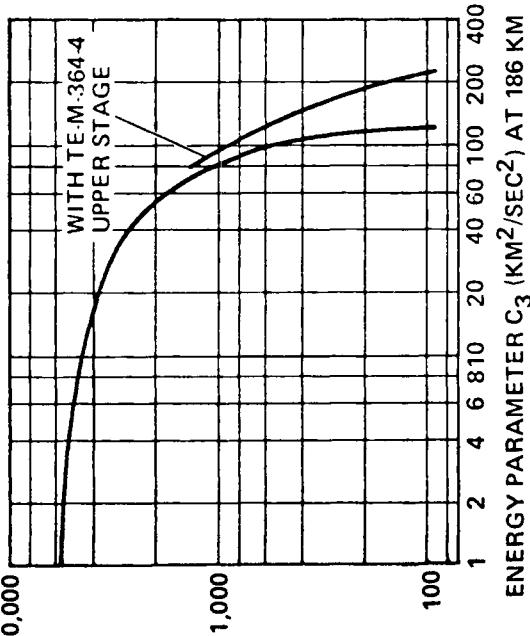


Fig. 3. Titan - Centaur performance

For selected arrival dates, classes of elliptical orbits at Jupiter may be defined in a parametric way (Fig. 4).

For the masses distributed according to Table 1, the selected Retro Propulsion Unit (Symphonie type, cf. Table 2) is capable of delivering a total impulse of approx. 2.3 km/s.

TABLE 2 Retro propulsion unit performance

- Spinning Operation Capability (120 rpm)
- Longterm Steady State Burn Capability (longest continuous burn 3 hours)
- Demonstrated Reliability (0.98 at 90 % L.C.L.)

Manufacturer	MBB
Status	Qualified for SYMPHONIE
Vacuum Thrust (Nominal)	392 N
Vacuum Spec. Impulse	303 ± 2 * sec
Mass	1 950 gr ± 50 gr
Propellant Combination	N ₂ O ₄ / AZ 50
Mixture Ratio	1.6
Expansion Ratio	1 000
Area Ratio	76.9 ± 1 %
Flow Rate (Nominal)	0.132 kg/sec
Inlet Pressure N ₂ O ₄ (Nominal)	10 bar ± 0.5 bar
Inlet Pressure AZ 50 (Nominal)	12 bar ± 0.5 bar
Chamber Pressure (Nominal)	7 bar ± 0.3 bar
Engine Length	400 mm
Nozzle End Diameter	179 mm
Operating Temperature Range (Ignition)	- 5 °C to + 70 °C
Thermal Capacity	950 J/°C
Absorbtivity of Nozzle	$\alpha = 0.82$
Emissivity of Nozzle	$\epsilon = 0.22$
Material	Nimonic / V2A
Ignition Time (Command to 90 % Thrust)	< 0.25 sec
Ignition Current (Valve actuation)	< 1.5 A
Voltage	27 VDC
Duration (power ON)	100 m sec
Reliability (90 % Confid.)	0.986

* Measurement tolerance of test equipment

For any given perizoenon distance, the insertion impulse decreases with increasing apozenon distance (or orbital period, or excentricity). In other words, for any given apozenon, or orbital period, the most economic mode of orbit insertion is to go as close to Jupiter as possible. However, radiation hazards must be taken into account; they limit the permanent perizoenon locations to a safe minimum distance of 10, better 14 Jupiter radii. It is not only obvious from Fig. 4 that the energy penalty from such a high orbit is significant; it also can be shown that an orbit of 14 r_J perizoenon and 100 — 150 r_J apozenon is more economically established by two-impulse insertion, i.e.

- targeting of the arrival hyperbola to 2 r_J flyby distance, establishing an elliptical orbit with the desired apozenon distance by a retro impulse at perizoenon;

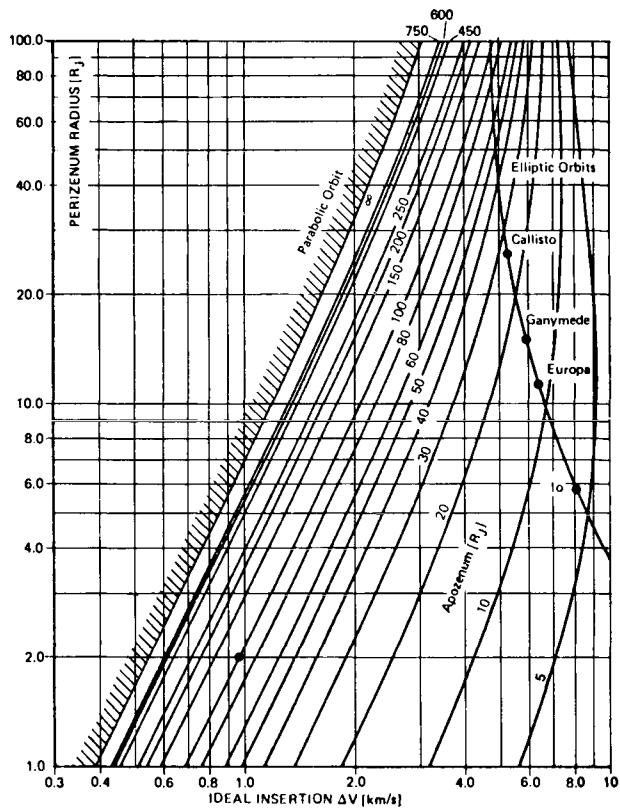


Fig. 4. Insertion impulse for elliptical orbits at Jupiter

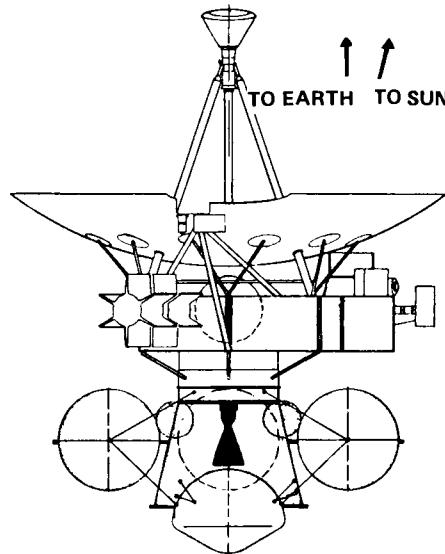


Fig. 5. PJO spacecraft configuration

- at apozenon, increasing orbital energy (and thus perizennon distance) by another powered maneuver.

The impulse budget for this sequence is shown in Table 3, together with several other maneuvers occurring through the mission.

TABLE 3 Retro propulsion unit impulse budget

	Velocity Increment	Orientation
Interplanetary Transfer (midcourse and targeting maneuvers)	104 m/s	any
Deflection on Arrival Asymptote	45 m/s	Earth
Orbit Insertion (perizennon)	780 m/s	Earth
Orbit Change (apozenon)	891 m/s	Earth
In - Orbit - Corrections	500 m/s	any
	2 320 m/s	

MISSION REQUIREMENTS

Taking into account all the above considerations, the guiding parameters for a 1980 Pioneer Jupiter Orbiter (with Entry Probe) mission were defined first, keeping in mind, as basic objectives,

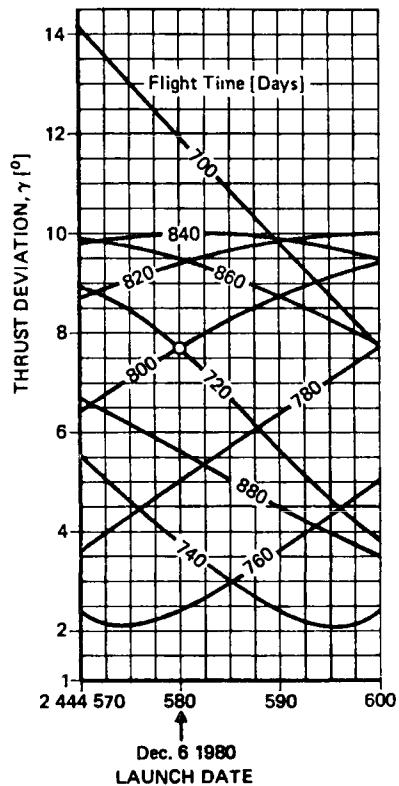
- “close-up and continuous studies of cloud dynamics and atmospheric structure, evaluation of the higher harmonics of the gravitational potential, investigation of the Galilean satellites and a thorough study of the magnetosphere”

Accordingly, the model payload (Table 4) is based upon the Pioneer 10/11 scientific instrumentation, with modifications resulting from Pioneer 10/11 mission experience as well as from the somewhat different criteria of an orbiter mission. Operational criteria, too, were based upon the Pioneer 10/11 concept, emphasizing, beside less critical requirements,

- Earth-line orientation of the spacecraft (Fig. 5) as far as possible, especially during critical maneuvers.
- Sun — spacecraft — Earth angle of at least 5° for all operations which call for full performance of the attitude measurement system.

TABLE 4 NASA - ESRO model payload (1974)

800 Days TRANSFER	
IDEAL INSERTION MODE	EARTH LINE MODE
$\gamma = 0^\circ$	$\gamma = 7.7^\circ$
$i = 0.1^\circ$	$i = 0.02^\circ$
$\omega = 60.44^\circ$	$\omega = 71.7^\circ$
$\Omega = 317.7^\circ$	$\Omega = 302.9^\circ$
$r_p = 5.8 R_{J}$	$r_p = 5.7999 R_{J}$
$r_A = 150 R_{J}$	$r_A = 150 R_{J}$
$\Delta V = 1.373 \text{ km/s}$	$\Delta V = 1.385 \text{ km/s}$

 ΔV LOSS 12 m/s

Magnetometer	MAG	3.0 kg
Plasma Magnetosphere	PMS	5.5 kg
Plasma Magnetosheath	PSH	2.5 kg
Electron Gun	PEG	1.0 kg
Energetic Particles	ENP	10.0 kg
Plasma Waves	PLW	6.0 kg
Micrometeoroids	MET	2.0 kg
UV Photometer	UV	0.7 kg
Imaging Photopolarimeter	IPP	6.3 kg
Infrared Radiometer	IRR	2.0 kg
Radiophysics (X - Band)	RPH	1.6 kg
Margin and Radiation Protection		6.9 kg
TOTAL		47.5 kg

TABLE 5 Probe — spacecraft telecommunications geometry

Telecommunications Range	76 000 km
Telecommunications Period	30 min
Data Rate	44 bps
Data to be stored	26 kbit
Antenna Half Beamwidth	66°
Probe Perizone	$0.985 r_J$
Pressure Lifetime Maximum	30 bar
Entry Angle	7.5°

FIG. 6. Earth-line powered insertion maneuver

While it turns out that the Earth-line mode may be kept for the major maneuvers with only a slight propellant penalty for using an impulse vector not exactly cotangential to the orbit (Fig. 6), the 5° elongation requirement (if observed throughout the window) turns out to be decisive for the selection of the proper arrival date.

An additional requirement is introduced by the Atmospheric Entry Probe, which has no propulsion system and is designed to use orbiter telecommunications as a data relay. That means that

- the orbiter must first be targeted into the proper entry trajectory for the probe ($< 1 r_J$ perizelon)
- after probe release the orbiter must be re-targeted (deflected) into its own arrival trajectory which, moreover, must provide for unobstructed probe — to — orbiter communications contact.

It turns out that, for fuel economy reasons, the separation maneuver must occur 1 — 2 months before perizelon, and that the necessity to finish data relay operations before the powered insertion into the intermediate orbiter orbit further limits the choice of the intermediate perizelon. The finally adopted value of 1.8 Jupiter radii for the latter was arrived at after several iterations of the parameters involved (Fig. 7, Table 5).

The location of the orbit with respect to Jupiter's magnetosphere is determined by arrival geometry and energy; the typical case is shown in Fig. 8. Subsequent changes are very small, apsidal rotation is negligible. For improving magnetospheric coverage, Friedman has proposed a multiple-swingby technique using the Galilean satellites for rotating (and, if desired, tilting) the spacecraft orbit throughout the mission (Fig. 9).

In order to initialize this type of mission, which is analyzed in detail in other publications, the arrival sequence must end with a close encounter of one of these large satellites, i.e., for $14 r_J$ minimum perizelon, Ganymede or Callisto.

Since, because of this minimum perizelon resulting from radiation safety considerations, the inner parts of the Jovian system (radiation belts, Io, Europa, and Amalthea) are only encountered during the arrival sequence discussed here, we concentrate on optimizing the arrival orbits with respect to these encounters.

ARRIVAL SEQUENCING

From the above discussion follows the fact that proper arrival sequencing is crucially depending not only on arrival geometry and energy but also on the positions of the five inner satellites of Jupiter. If radiation hazards to orbiter and probe are to be minimized, the rotation period of Jupiter's magnetosphere is to be considered as well.

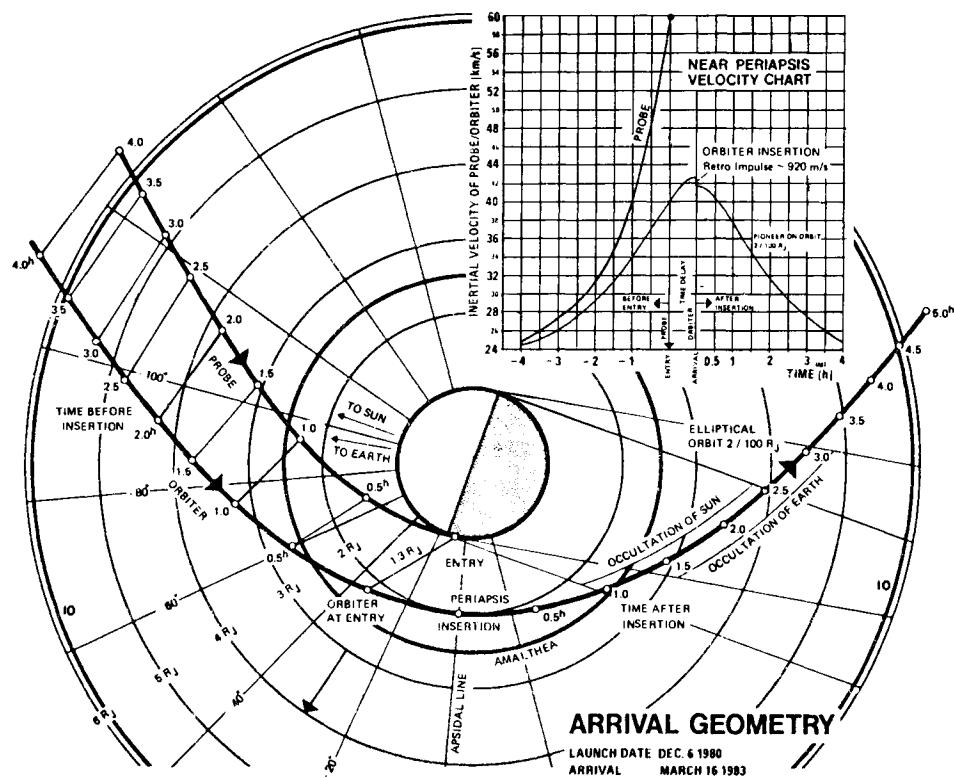


Fig. 7. Probe entry geometry

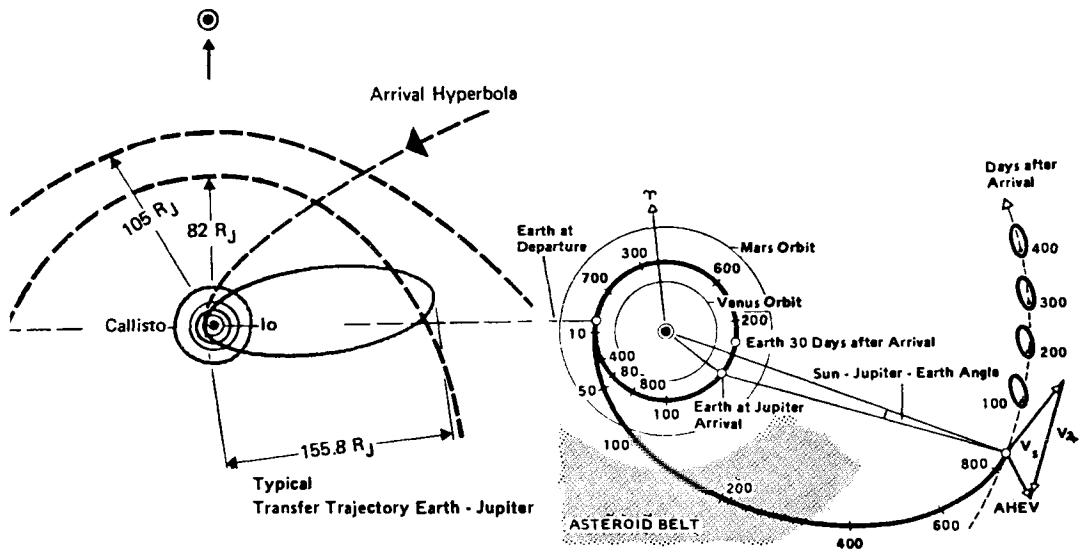


Fig. 8. Baseline mission geometry

Very conveniently, the periods involved are quite short (Table 6), and it is thus possible to define quite easily a sequence which

- conforms with the arrival window defined by the 5° Sun/Earth elongation requirement
- allows for moderately close encounters with all five satellites
- brings probe entry/retro maneuver/occultations into a temporal relation which is operationally feasible

This sequence (Fig. 10, Table 7) leaves plenty of options for improvement, especially as far as

- encounter distances
- terminator locations
- passage in the magnetospheric wakes of Io and possibly Amalthea
- magnetospheric coverage

are concerned. Also, probe mission designers may wish to avoid satellite encounters before entry for reasons of operational safety. Due to the short periods involved, there exist numerous possibilities, depending on the exact scale of priorities, for achieving an ultimately optimum mission. The required adjustments of perizeneron time will be of the order of hours; the Ganymede encounter at the end of the sequence may be kept unchanged by a slight adjustment of apozenon distance.

Further alternatives are

- shifting encounter time by full orbital periods of Ganymede
- using the outgoing rather than the incoming branch of the final ellipse for Ganymede encounter
- using Callisto instead of Ganymede for a "final" target.

Another alternative would imply additional phasing orbits — this is not recommended because of the limited duration of favourable Sun/Earth elongation conditions.

"FLOWER" SEQUENCE OPTIONS

Subsequent to the encounter sequence, numerous multiple-swingby options at Ganymede and Callisto exist for transforming the spacecraft's orbit as mentioned above and discussed by Friedman (6)

A list of priorities may conceivably correspond more or less to the tentative listing shown in Table 8; there are numerous other options, e.g.

- encountering the "median" satellites Himalia and Elara (major changes in "flower" geometry but not in the basic orbit)
- risking lower perizeneron altitudes after evaluation of initial encounter results may

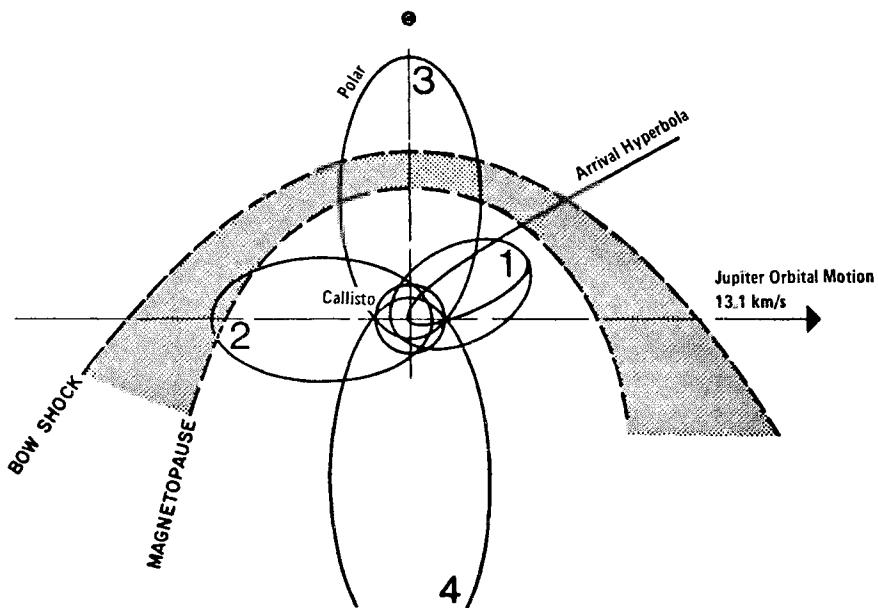


Fig. 9. Sample "Flower" orbit (Jupiter satellite gravity assist, after (6); phasing orbits deleted)

TABLE 6 Critical periods for sequencing

Jupiter System III (solid/magnetospheric)	0.41353 d
V Amalthea	0.49818 d
I Io	1.76914 d
II Europa	3.55118 d
III Ganymede	7.15455 d
IV Callisto	16.68902 d

TABLE 7 Baseline arrival events

1982 Dec. 18	Probe Separation
	Orbiter Deflection (degraded TM/TC)
1983 Jan. 7	Orbiter Deflection/Correc-
	tion
Jan 20/30	Bow Shock/Magnetopause
Feb. 6	Europa Encounter
	Io Encounter
	Probe Entry
	Terminator Crossing
	Perizenon Maneuver
	Amalthea Encounter
	Sun Occultation
	Earth Occultation
Mar. 19	Apozenon Maneuver
May 2	Callisto Encounter
May 4	Ganymede Encounter

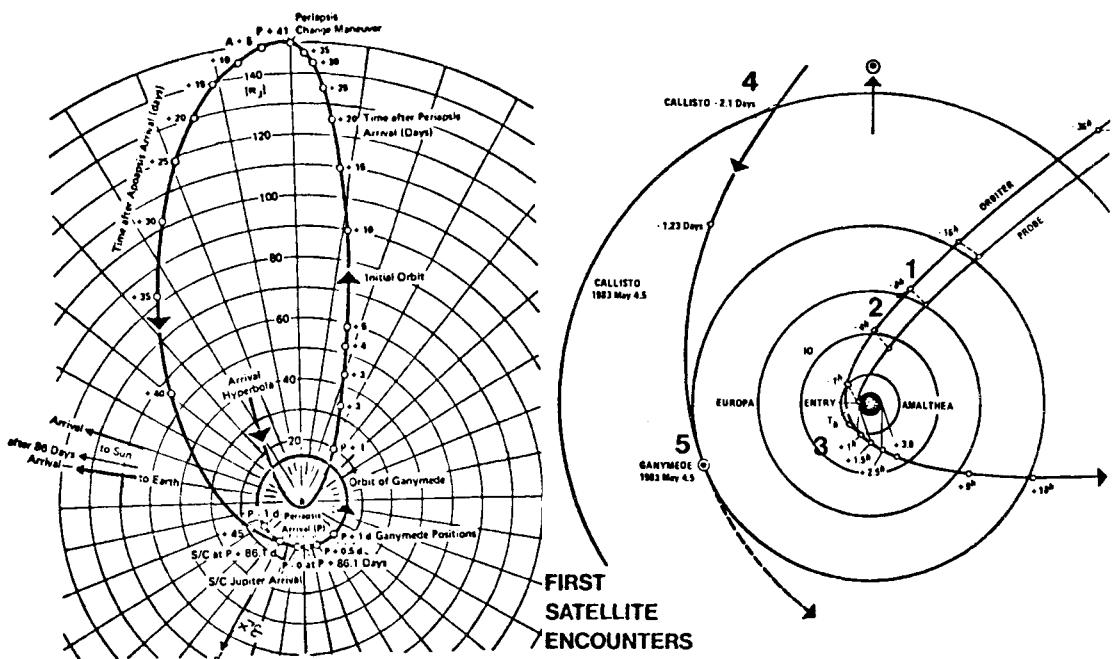


Fig. 10. Sample encounter sequence

TABLE 8 ESRO/NASA mission priorities

	PHASING REQUIREMENTS	FEASIBILITY CONDITIONS
PRIMARY OBJECTIVES		
1. Probe Data Relay	Entry to end of transmission	Perizoenon $1.8 r_2$, Arrival time optimization within 5 hours interval
2. Magnetosphere	Arrival to 30 months after	3 - petal flower (dawn - tail - dusk) Apozenon $150 r_2$
3. Satellites: — Ganymede/Callisto — Amalthea	Throughout 30 months mission In arrival sequence after probe operations	— Arrival time optimization within 12 hours interval
SECONDARY OBJECTIVES		
4. Io/Europa early	In arrival sequence	As timeline permits
5. Io/Europa late	After dusk petal	Pump down
6. Inclined orbits	After dusk petal	Crank up
7. Himalia and small satellites	Any time	As timeline permits

have shown evidence for safe possibilities,

and so on.

All these options must, of course, await a more rigorous analysis.

REFERENCES

- (1) D.J. Shramo, B.R. Foushee, and P.J. O'Leary, Centaur — a major element of the current space transportation system,
IAF Preprint 74 — 075 (XXVth Congress 1974)
- (2) K. Bohnhoff, D.E. Koelle, and W. Kokott, Mission criteria and propulsion system concept for a Pioneer Jupiter Orbiter,
AIAA Preprint 75 — 1154 (1975)
- (3) NASA — ARC, Pioneer F/G spacecraft operational characteristics,
Doc. PC — 202 (1971)
- (4) ESRO, Pioneer Jupiter Orbiter and Probe,
Doc. MS (74) 33 (1974)
- (5) MBB, Study on critical aspects of a joint ESA/NASA Jupiter Orbiter,
Report UR — V — 77 (1974)
- (6) L.D. Friedman, and R.R. Nunamaker, Mission design of a Pioneer Jupiter Orbiter,
AIAA Preprint 75 — 1135 (1975)

PROGRESS AND TRENDS IN COMMERCIAL SATELLITE COMMUNICATIONS—A SURVEY

P. L. Bargellini and B. I. Edelson

COMSAT Laboratories, Communications Satellite Corporation U.S.A.

INTRODUCTION

Satellites have changed the patterns and pace of world communications. In their effects upon the daily lives of people and commerce among nations, communications satellites constitute an outstanding benefit derived from the international investment in space technology.

Communications by satellite is now a practical business, having progressed from science fiction through stages of scientific analysis and engineering development to trials and operational use. Satellite and earth station operations and their interfacing with terrestrial systems are all well understood. The second decade of operational service has begun.

Today, the system operated by INTELSAT (the International Telecommunications Satellite Organization) provides routine reliable communications to 112 earth station antennas in 60 countries. Other satellite communications systems in Canada, the Soviet Union, and the United States provide domestic communications carrying telephone, television, facsimile, and data traffic. Communications services via satellite have been extended to small villages, individual buildings, ships at sea, and even to aircraft in flight.

This paper reviews the origins of satellite communications, surveys the technological and operational progress in communications satellites and systems employing them, and examines trends for the future.

HISTORICAL PERSPECTIVE

In 1945, Arthur C. Clarke conceived the development of satellites in geosynchronous orbit providing global telecommunications (Clarke, 1945). Clarke showed that three

geosynchronous satellites, powered by solar energy converted into electricity by silicon cells, could provide worldwide communications "for all possible types of service with unrestricted use of a frequency band at least 1000 MHz wide (providing), with the use of beams, an almost unlimited number of channels."

The realization of Clarke's concept required further development of adequate rockets for launching satellites and advanced electronic devices for radio transmission by satellites. However, even by that date, both rocketry and radio transmission had a considerable history. The pioneering work of Constantin Tsiolkowsky, Robert Goddard, Herman Oberth, Theodore Von Karman, and Wernher Von Braun led (directly or indirectly) to the V-1's and V-2's of World War II, and eventually to the first Sputnik, launched by the USSR in October 1957. The development of electrical communications had been going on for almost a century; after telegraphy, which came into use during the 19th century with wires spanning first countries, then continents, and eventually oceans, Maxwell's formulation of electromagnetic wave theory led to the discovery of radio by Marconi in 1895.

Electronic devices operating at ever increasing frequencies on the one hand, and a better understanding of radio wave propagation phenomena on the other, were major causative factors in electrical communications. However, curiously enough, at least until the late 1940's all electrical communications systems from telegraph systems to radio were developed without a precise quantitative understanding of the commodity (information) which was being handled. Eventually, in 1948, long after practical communications systems had been implemented, a body of knowledge known as statistical communication theory was developed

which focused on the fundamental relationships among information transmission rate, bandwidth, signal and noise power, and transmission impairments (distortion and attenuation) (Shannon, 1948).

Communications satellite systems benefited from this knowledge. As soon as rocket engineering provided the vehicles for safely injecting reliable electronic packages into orbit, success was ensured, and a substantial increase in communications capacity with respect to previously available wire and wireless systems was achieved, especially in the case of transcontinental communications. Long distance telegraph cables before the advent of inductance loading techniques and electronic repeaters were characterized by extremely low values of communications capacity (about 400-600 characters/minute, i.e., about 25 to 50 bits/second). Although inductance-loaded cables, as well as cables with repeaters, allowed a fivefold increase in capacity, it was not until 1950 that repeatered coaxial telephone cables with polyethylene dielectric permitted the spanning of considerable distances (Key West-Havana cable: 214 km, one cable for each direction). The first transatlantic telephone cable went into service in 1956 (Newfoundland-Scotland: 3,612 km, 52 repeaters with 68-km spacing, 36 voice channels/cable).

Radio communications advanced from the early very long wave (VLF) high-power transmitters capable of providing intercontinental telegraph service at speeds up to 30-40 words/minute (about 15 bits/second) to short wave (HF) transmission techniques developed in the middle 1920's. With power one or two orders of magnitude lower than that used at VLF, speeds about 10 times higher were achieved in telegraphy. In regard to voice communications, although the first transatlantic commercial telephone circuit (New York-London, 1927) operated at LF (60 kHz), much progress was made in the 1930's and 1940's with an extensive network of intercontinental HF telephone links. In spite of the advances made possible by the use of single sideband modulation, directive antennas, and diversity reception, the major drawback of HF radio was its low reliability (of the order of 70 percent), resulting from the vagaries of the medium (the ionosphere). Tropospheric communications developed after World War II, although more reliable

than HF, were inherently limited in range (about 1,000 km).

The stage was thus set for the communications satellite. When the first commercial satellite was orbited (Early Bird, 1965), the status of transoceanic communications was drastically altered. Two hundred forty voice circuits were available, and for the first time high-quality transatlantic TV was possible.

INSTITUTIONAL ARRANGEMENTS AND EARLY DEVELOPMENTS

The combined efforts and contributions of governments and industry were required to promote space exploration leading to satellite communications. In the United States, the Communications Satellite Act, enacted by the 87th Congress in August 1962, was signed by President John F. Kennedy on 30 August 1962. The Act declared it to be the policy of the United States "to establish, in conjunction and in cooperation with other countries, as expeditiously as practicable a commercial communications satellite system, as part of an improved global communications network, which will be responsive to public needs and national objectives, which will serve the communications needs of the United States and other countries, and which will contribute to world peace and understanding" (U.S. Congress, 1962).

The Communications Satellite Corporation (COMSAT), incorporated in February 1963, pursued a development program for the global satellite system and its financing. On 20 August 1964, 11 participant nations entered into a unique international partnership to be known as the International Telecommunications Satellite Consortium (INTELSAT), and designated COMSAT as its Manager.

Since its inception, INTELSAT has operated under two sets of agreements, referred to as interim and definitive. On 12 February 1973, INTELSAT, whose membership had grown to 87, became the International Telecommunication Satellite Organization, with the definitive agreements becoming effective.

It must be understood that, while the space segment, consisting of the satellites and the facilities required to support their operation, is owned and operated by INTELSAT, the earth segment, which comprises the earth stations in various countries, is owned by designated telecommunications entities in each country.

COMSAT, in its threefold role of U.S. representative in INTELSAT, owner and operator of U.S. earth stations, and Management Services Contractor for the INTELSAT system, has played a significant part in the development of commercial satellite communications.

Arthur Clarke's prescient concept had gone practically unnoticed by communications engineers for several years. But then, in 1954, J. R. Pierce of Bell Telephone Laboratories independently studied the fundamentals of radio relaying via artificial satellites, and about two years before the first Sputnik prepared the first concrete technical proposal for satellite communications (Pierce, 1955). Passive satellites had the important advantages of simplicity, reliability, and potentially unlimited multiple access, but studies of both active and passive satellites conducted from 1954 to 1962 indicated that the inverse distance square law applicable to active satellites would give them a great advantage over passive satellites, to which the inverse distance fourth power law applies. Although experiments have been performed with reflecting balloons, all operating systems have used active satellites.

Height and inclination of the orbit are fundamental parameters of satellite systems. The orbital period, derived from Kepler's Laws, is

$$P = \frac{2\pi a^{3/2}}{\mu^{1/2}}$$

where a is the semimajor axis of the ellipse, and μ is the gravitational constant x earth mass = $3.99 \times 10^{14} \text{ m}^3/\text{s}^2$. Since only modest payloads ($\sim 100 \text{ kg}$) could be placed in low or medium orbits ($< 10,000 \text{ km}$) by early rockets, these orbits were used for experimental satellites. However, with short periods, the satellites passed rapidly overhead and required tracking. Continuous communications between points on the earth's surface required traffic handover from a satellite setting beyond the horizon to another satellite rising to take its place. Many satellites would be required to ensure continuous coverage. In spite of these difficulties, systems were proposed with as many as 50 spacecraft in medium-altitude orbits to serve the Atlantic Ocean region. Early experiments with single spacecraft (TELSTAR, 1962-1963, and RELAY, 1962-1964) proved the feasibility of active satellites.

Increasing the orbital altitude to 35,863 km, i.e., about six times the earth's radius, increases the period to a sidereal day (23 hr, 56 min). In an equatorial orbit, the satellite appears to hover above a fixed spot, subtending an angle of about 18° , thus effectively providing coverage to about four-tenths of the earth's surface. Additional advantages of this so-called "geostationary" orbit are no Doppler effect on the signal, infrequent thermal stress cycles and low energy requirements for eclipse operation, mild radiation environment, and low perturbations by the earth's magnetic field.

Despite these advantages, the difficulties of achieving operation in geostationary orbit were considerable. The available version of the Thor-Delta vehicle, which had been used for launching the TELSTAR (AT&T, 1962-63) and RELAY (RCA-NASA, 1962-64) satellites, was inadequate to inject the corresponding payload, around 80 kg, directly into geostationary orbit from Cape Kennedy (28° north latitude). A solution, consisting of first injecting the payload in a highly elliptical orbit with apogee at synchronous altitude, was adopted (Rosen, 1973). An added rocket, whose weight was equivalent to about one-half of the payload in transfer orbit, fired at apogee, allowed circularization of the orbit while auxiliary thrusters were used to change the orbital plane, thus achieving zero inclination. Such a mission was regarded as highly complicated, given the state of the art of space technology at the time it was first proposed. Although the first experimental geosynchronous satellite, SYNCOM I (February 1963), failed to attain its orbit, the second and third launches of SYNCOM II (July 1963) and SYNCOM III (August 1964) were successful. These satellites were developed for NASA by the Hughes Aircraft Co.

A geostationary satellite must be kept in position since solar and lunar gravitational fields, the oblateness of the earth, and a slight ellipticity of the earth's equator are major causes of perturbations. Solar radiation pressure, on the other hand, is the major nongravitational perturbation. Stationkeeping is necessary to counteract these forces, corrections being obtained by activating onboard thrusters upon command from the earth. Ideally, longitudinal drift and orbit inclination can be controlled independently, but in practice some interaction is encountered. Usually drift is allowed to build up to some set limit

and then corrective maneuvers are effected.

In addition to orbital positioning, attitude control is necessary to point the satellite's directional antennas toward the earth. A common method of attitude control consists of spinning the satellite body for gyroscopic stiffness. Despinning the communications antenna permits beam pointing accuracy of the order of about one-tenth of a degree. This simple, reliable method eases the problem of thermal control of the structure, but also results in only partial utilization of the total area of the solar cell panels, limiting the amount of electrical power available.

Early Bird (INTELSAT I) was to some extent an "experimental satellite" with an operational capability (Metzger, 1972). It provided answers to questions related to the applicability of geostationary satellites within the existing global terrestrial telephone network. Early Bird was successful in proving, for commercial applications, the capability of achieving geosynchronous orbit, the practicality of stationkeeping and attitude control, the acceptability of the transmission delay inherent to the geosynchronous orbit, and the successful interconnection of terrestrial networks. Although designed for a lifetime in orbit of only 18 months, it remained in operational commercial service for more than four years.

SYSTEMS DEVELOPMENT

Newton's orbital mechanics, Maxwell's electromagnetic theory, and Shannon's communications theory provide the elements needed to define the theoretical bounds of communications satellite systems' performance.

Satellite communications channels involve at least two cascaded links, the up-link from an earth station to the active repeater in orbit, and the down-link from the satellite to another earth station; the two channels in opposite directions constitute a full circuit. A satellite in geosynchronous orbit sees about four-tenths of the earth's surface, and it can link stations separated by great circle distances up to 17,000 km. It can, in principle, link any pair of such stations. It is this capability of multiple access which makes the satellite a truly unique tool for communications.

For m stations visible by a sat-

ellite, the number of potentially available communications circuits is

$$n = \frac{m(m - 1)}{2} \quad (2)$$

Thus, the satellite system flexibility as an n -port network is conspicuous in comparison to the inflexibility of 2-port communications networks such as cable and land circuits.

In a space link, the rate of information transmission relates to certain fundamental parameters through the equation

$$R = \frac{BP_t G_t A_r}{4\pi r^2 \beta N} \quad (3)$$

where

R = information rate (bps)

r = link distance (m)

B = bandwidth (Hz)

P_t = transmitter power (W)

G_t = transmitting antenna gain

A_r = receiving antenna area (m^2)

β = communications efficiency

= $\frac{\text{energy per bit}}{\text{noise power density}}$

N = noise power (W)

The upper bound for R is the Shannon channel capacity:

$$R \leq C = B \log_2 \left(1 + \frac{S}{N} \right) \quad (4)$$

When B is allowed to go to infinity, the signal-to-noise ratio goes to zero and the communications efficiency β assumes the asymptotic minimum value $\log_2 2 = 0.693$. The actual value of β , greater than the above-mentioned minimum, depends upon the modulation-demodulation scheme (Sanders, 1960). Equation (3) has practical importance since it indicates that communications capacity is proportional to bandwidth and power when all other quantities are fixed. Operating frequencies and bandwidth, theoretically defined by spectrum selection in terms of minimum noise power density and favorable propagation, depend also on availability in terms of international and regional agreements. Power, easily available on the ground, is limited by

spacecraft mass and hence launch vehicle capability.

Commercial satellite repeaters have, until now, used a frequency translation process to separate the up- and down-links with identical modulation in both links. Modulation schemes trading signal-to-noise ratio for bandwidth expansion have been used to decrease the power requirements, and frequency modulation has prevailed because of its effectiveness, simplicity, and adaptability to the interfacing of satellite and ground communications systems. Frequency modulation is combined with frequency-division multiplexing of voice channels at baseband and frequency-domain multiple access for several RF transmissions through the satellite. This total transmission scheme is known as FDM/FM/FDMA.

Unfortunately, the situation is complicated by the presence of nonlinearities which cause modulation conversions (AM/AM and AM/PM) in the satellite repeaters and also at the transmitting earth stations. Ultimately, the maximum information transmission rate—in practice, the number of telephone circuits a satellite can handle between specified pairs of earth terminals—is dictated by the overall signal-to-noise ratio, S/N_{tot} , whose inverse equals the sum of the inverse signal-to-noise ratios arising in the up-link, the down-link, and the intermodulation processes:

$$\frac{1}{S/N_{tot}} = \frac{1}{S/N_{up-link}} + \frac{1}{S/N_{down-link}} + \frac{1}{S/N_{im}} \quad (5)$$

Other transmission impairments arise from earth RF out-of-band emission, cochannel interference due to imperfect beam isolation, transponder group-delay and dual path distortion, and adjacent transponder interference. Finally, in the case of frequency reuse through orthogonal polarizations, another kind of interference appears in the form of cross coupling due to rain depolarization effects (amplitude and phase). Clearly, transmission system planning requires careful efforts involving several complex interactive effects (Bargellini, 1972, 1973).

With FDM/FM/FDMA, the intermodulation noise contribution is reduced

by backing off the traveling wave tube amplifiers (TWTAs) until a maximum value for S/N_{tot} is achieved in

Eq. (5). Otherwise, the choice of a particular method of multiple access, or a combination of the methods of multiple access with specific forms of modulation and multiplexing, yields solutions which can be evaluated in terms of the number of channels which can be provided.

For a spacecraft of given configuration, mass, and available prime power, communications capacity can be increased through reuse of the available spectrum. The fundamental concept here involves sharing (or reuse) of the spectrum several times by separate narrow antenna beams (Dicks, 1975) and/or doubling of the available spectrum by overlapping antenna beams handling separate signals carried by orthogonally polarized waves (Kreutel, 1973). These methods, which are applicable to current FDM/FM/FDMA systems, can be further enhanced, as shown by numerous system studies, by using them in combination with digital modulation (PCM/PSK), time-division multiplexing (TDM), and time-domain multiple-access (TDMA) techniques. In order to take full advantage of the SDMA concept, onboard switching is necessary; hence, systems of this kind are designated as satellite-switched/space-domain multiple-access, or SS/SDMA, systems.

Figure 1 shows a possible system configuration combining TDMA with SS/SDMA. A spaceborne distribution center, consisting of a switching matrix and a control unit, provides the interconnection of different transponders. Information about traffic flow is stored in onboard memory circuits that control the switching matrix. Command signals from the earth can rearrange the connections in the matrix whenever necessary. Three important advantages result from SS/TDMA. With a single carrier present at any given time, the TWTs operate at saturation with maximum conversion efficiency, no intermodulation noise, and without the weight penalty of the multiplexing filters required in FDMA. An experimental TDMA system, TDMA-1, operating at 50 Mbps, was tested successfully over the Pacific Ocean via an INTELSAT III satellite in 1970 (Maillet, 1971). Satellite switching has not been tested in orbit.

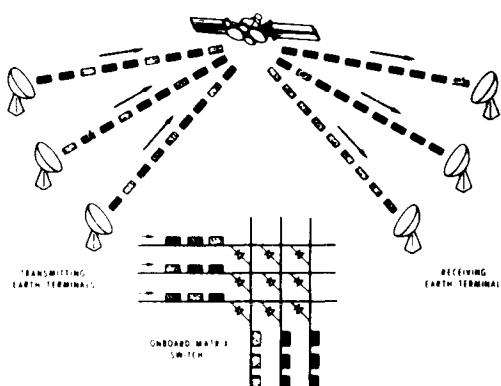


Figure 1. Satellite-Switched TDMA

As previously mentioned, a multi-beam frequency reuse satellite may also employ FDMA. In this case, filters are cross connected, with the arrangements of the connections commanded by ground signals. The weight of the filters and the backing off of the TWTs necessary in any FDMA system are disadvantages of this approach, although its operational implementation may be somewhat easier than that of an SS/TDMA system.

As communications capacity is proportional to transmitter power, power is in turn related to spacecraft mass and size. Since the communications subsystem of a satellite is the useful payload which provides the above-mentioned capacity, it is important to keep the mass of all other subsystems as low as possible. In the apportionment of mass among the different subsystems (communications, structure, power, positioning and orientation, and TT&C), communications satellites of different types have been investigated (Kiesling, 1972). These investigations have made it possible to identify trends which indicate that the introduction of certain advanced technologies would allow sizeable increases in communications capacity for a given total satellite mass. Although these concepts have been verified in some experimental programs, the first decade of satellite communications has chiefly evolved around the gradual growth from the technology introduced in the SYNCOM and Early Bird satellites. In this sense, communications capacity has increased in terms of larger and more powerful satellites. The launch vehicle combinations and weight capability are shown in Figure 2.

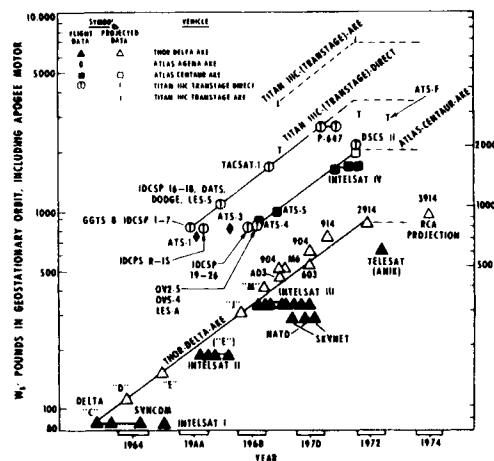


Figure 2. Launch Vehicle Evolution

In spite of the variety of potential launch vehicles, from small rocket probes to SATURN V, the choice of a practical vehicle for placing communications satellites in geosynchronous orbit is limited (in the U.S.A.) to the Thor, Atlas, and Titan rockets in combination with specific upper stages.

Additional important factors are the availability of launch facilities (sites and pads), the program life of a given vehicle, the reliability record of each program, and naturally, the cost of the vehicle. An analysis applicable to the INTELSAT IV system has appeared elsewhere (Edelson, 1974).

There have been altogether 138 known attempts to launch communications satellites of one type or another, experimental and operational, civil and military. By country (of ownership), these break down as follows:

Canada	3
France/Germany	2
INTELSAT	21
U.S.S.R.	47
U.K.	4
U.S.	61

Of these satellites, 38 were for experimental purposes and 100 were intended to provide operational service. Four of the experimental satellites were passive (Echo and Westford) but none of them have been launched since 1964. Of the total launch attempts, it might be of interest to note that 128 were orbited successfully. The orbits deployed were as follows:

geostationary	41
12 hour	46
low and medium altitude	51

The INTELSAT global system was the first commercial system in operation (1965), and is today by far the largest and most extensive. However, there are, of course, other systems in operation and under development, including regional and domestic systems, and special purpose systems for military, maritime, data collection, or aeronautical service. The INTELSAT system and several domestic systems are treated in the following sections.

THE INTELSAT SYSTEM

All INTELSAT satellites have operated in synchronous equatorial orbits—so-called "geostationary" orbits—35,700 km above the equator. This position allows for continuous coverage, with tracking by earth stations necessary only to maintain their radio beams pointed at the satellite through very small angles of drift. The satellites receive on frequencies in the 6-GHz band and transmit on frequencies in the 4-GHz band. All of the satellites so far launched are spin stabilized; that is, they maintain their orientation in space through rotation of the body of the satellite along an axis parallel to the earth's axis. All satellites are powered by arrays of silicon solar cells mounted on the spinning body.

The transponders employ TWTAs generating a few watts of RF output power.

The effectiveness of a communications satellite is dependent upon its capacity, which in turn depends upon radiated power and bandwidth, and also upon its lifetime in orbit. As the technology developed during the past decade, INTELSAT satellites improved in all three areas:

- (a) Power. More powerful launch vehicles placed greater weight in orbit, thus providing a greater area for solar cells and more electrical power generated and available for radio transmission.
- (b) Bandwidth. With increased power and more sophisticated transponder design, more of the allocated bandwidth was used and eventually reused.

(c) Lifetime. Improved components, devices, and design, and more effective quality assurance techniques, all improved satellite reliability and operating lifetime.

Early Bird, built by the Hughes Aircraft Company, was launched in April 1965. It weighed only 38 kg in orbit and had a total effective radiated power of 10 W in each of two transponders, using only 50 MHz of bandwidth. Its limited power output required that the spinning antenna be "squinted" to cover the heavy traffic route between North America and Europe. Its potential capacity of 240 two-way telephone circuits could allow linking of only two earth stations at a time; i.e., there was no multiple access available, nor was it possible to carry television and telephone simultaneously.

A second series of commercial satellites was developed to support the manned spaceflight operations of NASA. Reliable communications were urgently needed to connect a worldwide network of tracking stations for Project Apollo, some on islands and others on ships at sea. The Hughes-built INTELSAT II satellites were larger than Early Bird, having more power and bandwidth, and thus were able to provide coverage of a wider area of the earth. An important innovation in INTELSAT II was the introduction of multiple-access capability: many pairs of earth stations could be connected through the satellites, each transponder carrying several radio frequency carriers simultaneously. The first INTELSAT II satellite entered service in January 1967.

The larger INTELSAT III series, built by TRW, was introduced in late 1968. By 1969, three of these satellites made possible the realization of a true global communications system. Each INTELSAT III had a nominal capacity of 1,200 telephone circuits. The increase was achieved by using a mechanically despun antenna always pointed toward the earth, providing a so-called "global beam" which covered all of the earth visible from a given position of the synchronous orbit.

The first INTELSAT IV satellite was launched in January 1971. Also built by Hughes, and weighing 720 kg in orbit, these satellites are currently the mainstay of the global network. The major advance of INTELSAT IV over its predecessors is the use of

"spot-beam" antennas covering only a small portion of the visible earth, in this case a beam angle of about 4.5°. The resulting concentration of radiated energy provides another increase in capacity. The INTELSAT IV satellites are therefore rated at about 4,000 circuits or greater, depending on the number of transponders connected to spot beams and the multiple-access system in use. The INTELSAT IV electrical power subsystem provides about 470 W generated by some 45,000 solar cells and includes Ni-Cd batteries used during solar eclipse.

INTELSAT IV is the first communications satellite to be bandwidth limited. The communications subsystem is channelized into 12 transponders of 36-MHz bandwidth. All transponders receive from a global-beam antenna. Both global- and spot-beam transmit antennas, providing about 180 and 2,500 W of equivalent radiated power (22.5- and 34-dBW e.i.r.p.), respectively, are available.

Extended capacity satellites, improved versions of the present INTELSAT IV models (to be known as INTELSAT IV-A), have been under construction by Hughes Aircraft Company for three years. These satellites are expected to see service in late 1975 in the Atlantic region to meet that area's higher level of traffic requirements. As already mentioned, these will be the first satellites to

incorporate a frequency reuse technique, expanding the number of simultaneous transmissions per frequency bandwidth through spatial separation. The spacecraft body will be similar to that of INTELSAT IV, but will include a modified communications subsystem and advanced antenna array. The antenna patterns will allow simultaneous use of the same portion of the spectrum in two separated areas; e.g., in the Atlantic region one shaped beam will cover Europe and Africa, and the other will cover North and South America. Global beams will also be included. The satellite will have 20 transponders connected in various combinations to global, eastern, or western receive or transmit beams. With approximately the same weight and power as INTELSAT IV and using the same 6/4-GHz frequency band, the "A" version satellite will have 50 percent greater capacity (about 7,000 circuits) through frequency reuse.

Table I shows how the first five generations of INTELSAT satellites have increased in size and weight, communications capacity, and design lifetime with an attendant increase in satellite construction and launching costs. However, the overall cost effectiveness of each successive generation has improved, and during the 10-year period, the investment per circuit year in orbit has decreased from \$30,000 to about \$1,000 from INTELSAT I to IV (Edelson, 1975).

Table I. The Five Generations of Satellites in the INTELSAT System

	INTELSAT I	INTELSAT II	INTELSAT III	INTELSAT IV	INTELSAT IV-A
YR OF 1st LAUNCH	1965	1967	1968	1971	1975
DIMENSIONS (cm) {DIA / HEIGHT	72.3 59.6	142 67.3	142 104	238 282 (DRUM) 528 (OVERALL)	238 282 (DRUM) 590 (OVERALL)
MASS (kg) {AT LAUNCH / IN ORBIT	68 38	162 86	293 152	1385 700	1469 790
LAUNCH VEHICLE	THOR-DELTA	IMPROVED THOR-DELTA	LONG-TANK THOR-DELTA	ATLAS/CENTAUR	ATLAS/CENTAUR
PRIMARY POWER (W)	40	75	120	400	500
TRANSPOUNDERS	2	1	2	12	20
BW/TRANSPONDER (MHz)	25	130	225	36	32-36
ANTENNA	OMNI-SQUINTED	OMNI	DESPUN	DESPUN	DESPUN
COVERAGE	N HEMISPHERE	GLOBAL	GLOBAL	GLOBAL & SPOT BEAMS	GLOBAL & HEMI BEAMS
e.i.r.p./BEAM(dBW)	11.5	15.5	23	22.5 (GLOBAL) 33.7 (SPOT BEAM)	22 (GLOBAL) 29 (HEMI BEAM)
NO. OF TEL. CIRCUITS	240 (NO MULT ACCESS)	240	1200	4000	6000
DESIGN LIFETIME (YRS)	1.5	3	5	7	7
COST/CIRCUIT YEAR (\$/000)	30	10	2	1	1

Along with the spacecraft, the earth segment has grown enormously. The global satellite communications system has grown from 5 earth stations in 1965 to 112 earth station antennas in 60 countries.

Figures 3 and 4 show the INTELSAT system today, and its traffic growth from 1967 to 1975. Figure 5 depicts the distribution of the communications paths in the three ocean regions: Atlantic, Pacific, and Indian.

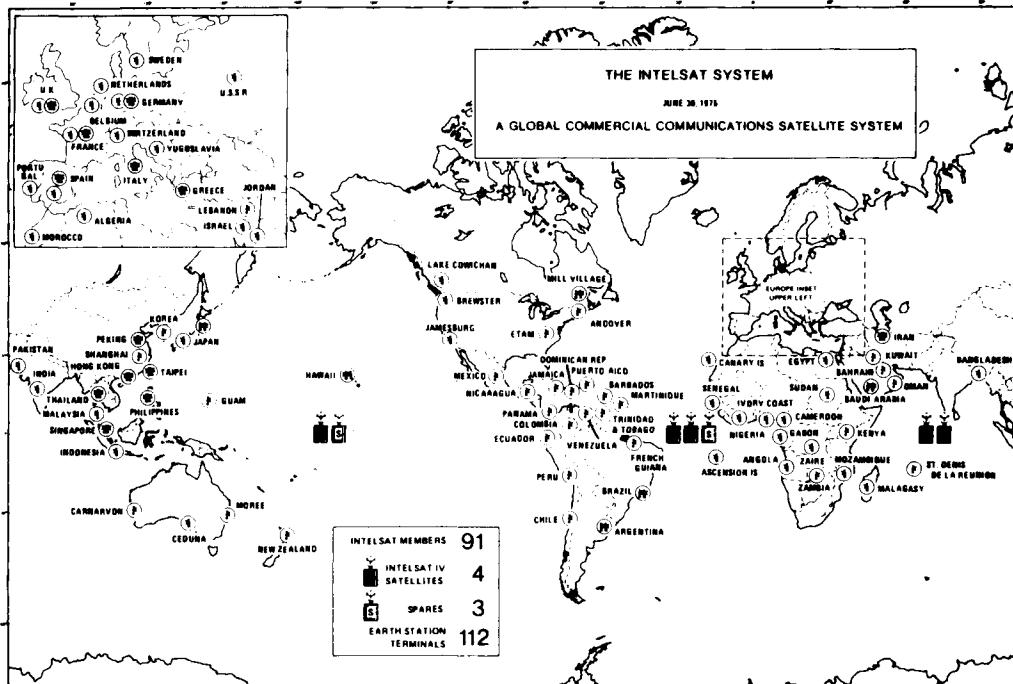


Figure 3. The INTELSAT System Today

The INTELSAT system provides continuous global telephone service in addition to television, telegraph, and data transmission. Temporary service is also available on a short-term basis during special world events and times of special communications requirements. All such services meet or exceed international standards of quality (Browne, 1974).

The primary operational mode of the system is multiple access, which is preassigned in the frequency domain. A demand-assignment multiple-access system (known as SPADE) was introduced in 1973 and has been implemented in the Atlantic Ocean region

with 21 terminals in operation. Direct digital transmission has been established between a number of points, and more are planned.

The system has provided restoration service for interrupted submarine cables on very short notice, and permitted the rerouting via satellite of a number of such cable circuits which would have remained suspended indefinitely while the repairs were made at sea. During the first half of 1974, temporary service for submarine cable restoration via satellite comprised 24,926 half-circuit days. At one period in 1972, 34,100 half-circuits were used in this way.

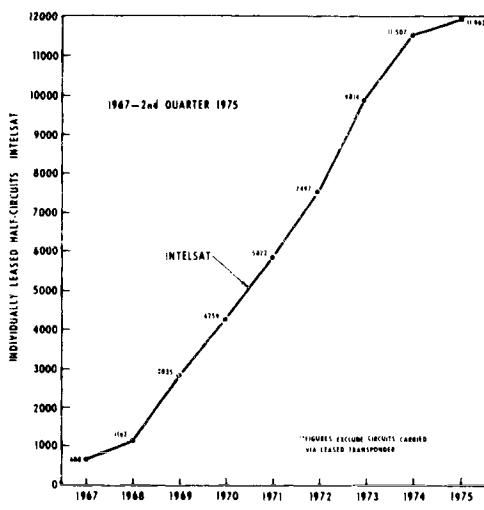


Figure 4. Traffic Growth in the INTELSAT System

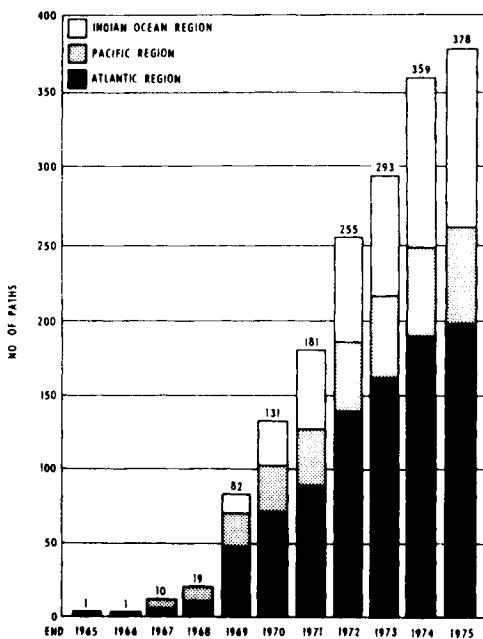


Figure 5. Earth-Station-to-Earth-Station Paths in the INTELSAT System

All real-time transoceanic television is transmitted via satellite. TV service is characterized by large fluctuations, caused by singular world events such as international political happenings and sports events.

THE CANADIAN DOMESTIC SYSTEM

Introduction

Canada has been the first country in the Western hemisphere to establish a multipurpose communications satellite system for domestic service (Chinnick, 1973; Ahmed, 1973). The space segment comprises three satellites in orbit, the first launched in November 1972 and becoming operational in January 1973 at 114° west longitude; the second launched in April 1973 at 109° west longitude; and the third launched in May 1975 at 104° west longitude.

The satellites were manufactured by Hughes Aircraft Co. with Canadian firms as subcontractors for the on-board communications equipment (Northern Electric) and the main structure (Spar Aerospace). The launch vehicles (Thor-Delta) and facilities were provided through NASA.

Telephone, television, data, and facsimile transmission services are provided throughout Canada, increasing the capacity of the existing terrestrial systems and also permitting their interconnection. The overall transmission performance is comparable to or better than that of terrestrial systems.

The earth segment consists of over 50 stations of different categories: 2 for heavy route traffic, 6 for network television, 18 for thin route traffic, 2 for northern telecommunications, 26 for remote television service, and finally, 1 TT&C station. The operating frequency bands are from 3.202 to 4.178 GHz in the down-link (horizontal polarization), and 5.927 to 6.403 GHz in the up-link (vertical polarization).

The ANIK Spacecraft

The satellites are cylindrical with a diameter of 190.5 cm and height of 161 cm. The mass in orbit is 600 kg. The platform carrying 12 transponders spins at 100 rpm, with the outer drum, which supports 20,048 silicon cells, producing 300 W of electrical power at beginning of life. The despun upper structure carries a

152.4-cm-diameter lightweight parabolic reflector which is illuminated by a multiple form offset feed. The radiation pattern (about $3^\circ \times 8.5^\circ$) is elliptically shaped to cover the territory of Canada. The e.i.r.p. per transponder is 33 dBW (minimum at the beam contour) obtained by a 5-W TWT in each transponder, and the transmit antenna gain is 27 dB. A TT&C antenna which is positioned on top of the parabolic reflector brings the overall satellite height to 358.1 cm.

Figure 6 is a photograph of the spacecraft. The repeater is a fixed gain, single-conversion, 12-transponder design with 36-MHz bandwidth allocated to each channel and a 4-MHz guard band between adjacent channels. A redundant (switchable) wideband receiver using tunnel diode amplifiers at 6 GHz is common to the 12 transponders. The receive G/T ratio is -7 dB/K. Each transponder can operate with the TWT at saturation in the single-access mode, providing a capacity of 960 voice circuits or one color TV and two 5-kHz audio circuits (earth station G/T = 37 dB/K). It can also operate at backoff in the multiple-access FDM/FM/FDMA mode with varying communications capacity, depending on the circumstances. During eclipses, only 10 of the 12 transponders can be fully operational.

The Earth Segment

The locations of the Canadian earth stations are shown in Figure 7, and as mentioned previously, their variety is considerable. The TT&C facilities, located at the two heavy route stations on the east (Allan Park, Ontario) and west (Lake Cowichan, Vancouver Island) coasts, operate in conjunction with the Ottawa control center. During the transfer orbit phase of a mission, a third TT&C station operates on the island of Guam in the Pacific.

The two heavy route stations satisfy high-density traffic requirements as well as television services. The antennas are parabolas of 29.87-m diameter, exhibiting a gain of 63 dB in transmission and 59 dB in reception. The combination of the antenna and a parametric front end yields a G/T of 37 dB/K. The e.i.r.p. per carrier is 84 dBW with 960 voice circuits, or color TV plus two audio channels. Normally, there are five transmit channels and seven receive channels with hot standbys. The stations operate continuously and have tracking

capabilities. Power is provided via a rotary converter, batteries with 15-min capacity, and a diesel-driven generator with automatic starting in case of primary power failure.

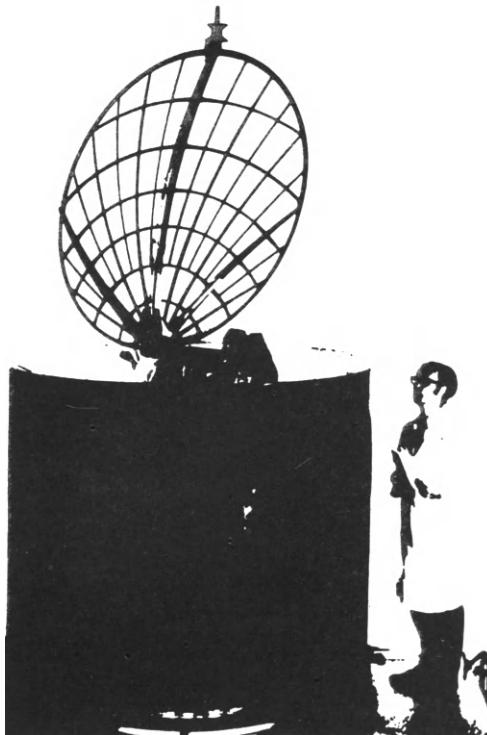
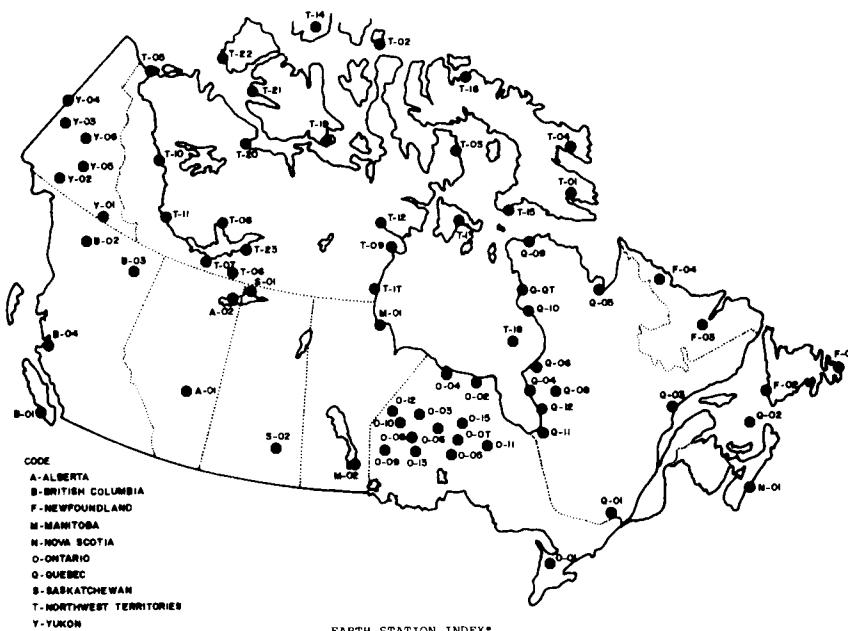


Figure 6. The ANIK Spacecraft
(courtesy of TELESAT Canada)

The six network television stations located near major cities provide TV signal transmission and reception for use by the CBC network. The antennas are parabolas of 10.5-m diameter with gains of 52.5 and 50.5 dB, respectively, in transmission and reception. The antennas are fixed and, although no tracking capability exists, manual steering over a limited range is available. Thus, in general, precise satellite stationkeeping is required in order to remain within the antenna beamwidth. The G/T is 28 dB/K and the e.i.r.p. per carrier is 83 dBW. The communications capacity per carrier is one TV signal plus two audio channels. Each station can simultaneously receive on three channels and transmit on one of three preassigned channels. Some stations are permanently manned, while others require only part-time staffing. Primary power is supplied by commercial power



EARTH STATION INDEX	
ALBERTA	NORTHWEST TERRITORIES (Continued)
A-01 Huggett	T-06 Ft. Smith
A-02 Ft. Chipewyan	T-07 Pine Point
BRITISH COLUMBIA	T-08 Yellowknife
B-01 Lake Cowichan	T-09 Rankin Inlet
B-02 Cassiar	T-10 Norman Wells
B-03 Ft. Nelson	T-11 Ft. Simpson
B-04 Bella Bella	T-12 Baker Lake
MANITOBA	T-13 Coral Harbour
M-01 Churchill	T-14 Red Point
M-02 Belair	T-15 Cape Dorset
NEWFOUNDLAND	T-16 Pond Inlet
F-01 Bay Bulls	T-17 Eskimo Point
F-02 Port au Port	T-18 Sanikiluaq
F-03 Goose Bay	T-19 Cambridge Bay
F-04 Nain	T-20 Coppermine
NOVA SCOTIA	T-21 Holman
N-01 Harrietsfield	T-22 Sachs Harbour
NORTHWEST TERRITORIES	T-23 Snowdrift
T-01 Frobisher	ONTARIO (Continued)
T-02 Resolute	O-12 Sachigo Lake
T-03 Igloolik	O-13 Slave Falls
T-04 Pangnirtung	O-14 Webique
T-05 Inuvik	QUEBEC
	O-01 Rivière Rouge
	O-02 Magdalen Is.
	O-03 Sept Isles.
	O-04 Ft. George
	O-05 Ft. Chimo
	O-06 Poste de la Baleine
	O-07 Povungnituk
	O-08 Radisson
	O-09 Saglouc
	O-10 Inoucadjuac
	O-11 Ft. Rupert
	O-12 Nouveau Comptoir
	ONTARIO
	SASKATCHEWAN
	S-01 Uranium City
	S-02 Qu'Appelle
	YUKON
	Y-01 Watson Lake
	Y-02 Whitehorse
	Y-03 Dawson
	Y-04 Clinton Creek
	Y-05 Faro
	Y-06 Elsa

*Stations are numbered by province in order of start of service.

Figure 7. TELESAT Earth Stations (courtesy of TELESAT Canada)

systems with local backup of diesel generators, batteries, and static inverters.

The remote TV stations, which are unattended, provide TV reception in areas which would not otherwise have TV coverage. An attached TV transmitter rebroadcasts the program locally.

The antennas at these stations are 7.92-m parabolas with receive gain of 48.5 dB and G/T of 26 dB/K. The antennas are fixed and no tracking is available. The power and the communications equipment are housed in a small modular shelter (2.44 x 5.80 m) which can be quickly assembled and easily removed.

The northern telecommunications stations provide telephone service up to 132 voice channels at two far north locations. The stations are unattended but controlled from a nearby supervision and maintenance center. The antennas used at these stations are 10.15-m parabolas (52.5-dB transmit and 50.5-dB receive gain), with a G/T of 28 dB/K and an e.i.r.p. of 73 dBW per carrier. Klystrons rated at 1.5 kW are used in the final power amplifier stage.

The thin route stations provide small communities with a limited number of telephone circuits (usually two to eight). Single-channel-per-carrier techniques with delta modulation (32 kbps) and FDMA preassigned access are used, and in addition to telephone services, combinations of voice, teletype, facsimile, and data services are available. Increased capacity will be achieved by voice activation and demand-assigned multiple-access techniques. Major characteristics of these stations are a G/T of 20 dB/K and an e.i.r.p. per voice carrier of 55-58 dBW.

The TELESAT system was the first to introduce TDMA, starting with early tests in 1971 and with commercial operation beginning in September 1975. Four hundred telephone circuits will be handled at an overall transmission rate of 61.248 Mbps between a station at Harrietsfields, Nova Scotia (G/T = 31 dB/K) and another station at Allan Park, Ontario (G/T = 37.5 dB/K). This TDMA system will replace a previous 240-circuit FDMA system (Kwan, 1975).

The outstanding characteristic of the TELESAT system is its ability to provide a variety of services. In general, as system optimization for a given type of service does not necessarily coincide with optimization for other services, this conflict could be resolved by optimization of all services, matching different earth stations with satellite transponders of different bandwidth and power. However, this solution is not desirable because flexibility would be lost, and in addition, spacecraft mass and service costs would increase. Standardized transponder design, with compromises in communications capacity and quality of service, is thus preferable. For given values of earth station G/T, satellite e.i.r.p., and bandwidth, boundaries can be defined in terms of physical parameters, available tech-

nologies, operational margins, and costs.

The careful planning of the TELESAT system seems to have paid off well in terms of the operational success and experience since 1973. The advanced design of the spacecraft itself has resulted in its adoption by other users (e.g., the U.S. domestic system of Western Union and the Indonesian domestic system).

U.S. DOMESTIC SATELLITE SYSTEMS

In December 1972 the U.S. Federal Communications Commission, after almost seven years, concluded deliberations on the matter of domestic communications satellites and announced a so-called "open-sky policy," whereby the arena was open to competing private enterprise.

About a dozen companies filed applications with the FCC for domestic systems. Three companies (American Satellite, RCA, and Western Union Telegraph) established ground stations and initiated services in 1973-74. Initial service was established through leasing available transponder capacity from the Canadian TELESAT satellites.

Western Union launched the first U.S. domestic satellite, WESTAR I, in April 1974, followed by WESTAR II in June of the same year. These satellites, essentially identical to the ANIK satellites of TELESAT, were built by Hughes Aircraft Co. Commercial service was established in August 1974 with five earth terminals in New York, Los Angeles, Chicago, Dallas, and Atlanta. These stations have 15-m parabolic antennas.

RCA is soon expected to establish a U.S. domestic system of its own involving body-stabilized spacecraft operating at 4 and 6 GHz, providing about 15,000 voice circuits by means of 20 transponders, each having a 36-MHz bandwidth using dual polarization (horizontal and vertical) for frequency reuse of the 500-MHz spectrum (Hume, 1975).

In the meantime, COMSAT was authorized to furnish satellite capacity to AT&T through a subsidiary corporation, COMSAT General, established in 1973 for this purpose. In January 1974, the FCC formally approved the plans for a system which will consist

of three satellites in orbit and one spare on the ground. The major characteristics of these satellites, called "COMSTAR," currently under construction at Hughes Aircraft Co. (Rusch, 1974), are shown in Table II. The communications subsystem is shown in

Figure 8. Beacons at 19 and 28.6 GHz will be flown on COMSTAR satellites to measure attenuation, phase coherence, and depolarization (Cox, 1974; Briskman, 1974).

Table II. Major Characteristics of COMSTAR Satellites

Configuration	Dual-spin spacecraft
Satellite Mass	1,470 kg
Launch Vehicle	Atlas-Centaur
Primary Power	550 W
Frequencies	4 and 6 GHz (up- and down-links with frequency reuse by means of orthogonal linear polarizations)
Polarization Isolation	33 dB (31 dB for CONUS/Alaska)
Antenna Coverage	CONUS, Alaska, Hawaii, and Puerto Rico
Number of Transponders	24, each with 34-MHz usable bandwidth
e.i.r.p. (at beam edge)	33.0 dBW (31 dBW for CONUS/Alaska combined)
Lifetime	7 years

Launches are planned in early 1976. AT&T will own and operate five earth stations, one in Pennsylvania, one in Illinois, one in Georgia, and two in California. In addition, stations operated by GT&E in Hawaii, Alaska, and Puerto Rico will be part of the system. COMSAT General Corp. will provide two TT&C earth stations (Connecticut and California) and a system control center in Washington, D.C. With two satellites in orbit fully utilized and the third serving as a backup for peak traffic situations, the system will be capable of handling about 30,000 voice circuits. (It should be pointed out for compari-

son that AT&T is operating some 400,000 domestic telephone circuits on its present land facilities.) The earth stations will be fully steerable parabolas of 30-m diameter with G/T of 41.4 dB/K and e.i.r.p. of 92.3 dBW.

In summary, it seems proper to state that domestic U.S. satellite service now just begun will be greatly expanded in the next few years. Satellites will be used in conjunction with the extensive microwave and coaxial communications facilities available in the U.S. and should provide a wide range of services, including public switched telephone service, private line telephone service, television distribution, facsimile, data, computer intercommunication, and perhaps eventually electronic mail.

OTHER SATELLITE SYSTEMS

In addition to the INTELSAT international system, five separate satellite communications systems should be operational in 1976. They are, in order of establishment: ORBITA (U.S.S.R.); TELESAT (Canada); and WESTAR, COMSTAR, and RCA (U.S.A.). The American Satellite Corporation system using the WESTAR satellite will also be operational. Indonesia has contracted for a separate domestic system using an ANIK-type satellite. Other countries and groups have indicated that they are studying or planning independent domestic or regional systems. These include a 9-nation European system, an Arab system, and systems for Brazil, Iran, and Japan.

A maritime satellite communications system (MARISAT) is now being developed by COMSAT General Corp. (Lipke, 1974). Also, an aeronautical system (AEROSAT) is being planned by Canada, ESA, and the U.S.A.

In the meantime, the practice of leasing available transponder capacity for special system use has continued. In the INTELSAT system, participants in this kind of activity have been, or will be, Algeria, Brazil, Chile, Malaysia, Mexico, Nigeria, Norway, Spain, U.S.A., and Zaire, for domestic communications or dedicated TV transmissions.

FUTURE TRENDS

Satellite communications systems will continue to expand because traffic growth forecasts indicate the need for greater communications capacity. For example, the yearly average growth

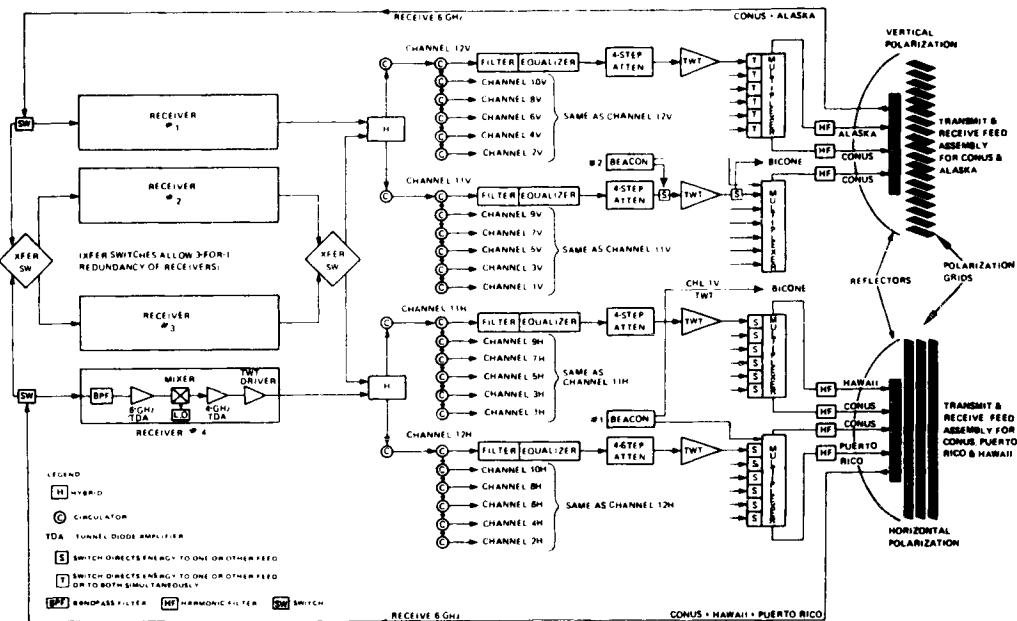


Figure 8. COMSTAR I Communications Subsystem

rate in the INTELSAT system is expected to be at least 15 percent in the next decade. Satellites will become more competitive vis-a-vis other systems—in terms of magnitude of communications capacity, high reliability, flexibility, and progressively increasing cost effectiveness.

The two natural resources, the electromagnetic spectrum and the geosynchronous orbit, do, of course, constitute the bounds of satellite communications growth, and will be briefly discussed in the following sections.

Spectrum Utilization

As latecomers among the users of the electromagnetic spectrum, satellites have had difficulty in being assigned "optimum" spectral regions. The assignment to fixed satellite services of 500 MHz at around 4 and 6 GHz for the down- and up-links, respectively, is close, but not optimum since the overall noise power density exhibits its broad minimum between 1 and 2.5 GHz. In addition, power flux density limitations were imposed on satellite systems because the above-mentioned 4- and 6-GHz frequency bands had to be shared with existing terres-

trial microwave systems. In spite of these constraints, satellite systems have been successful, the most serious limitation probably being that which derives from the constraints on the site selection for the earth terminals.

All commercial satellite systems at this time operate at 4 and 6 GHz. New frequency bands were assigned to satellite communications at the World Administrative Radio Conference of 1971. There are two new pairs of bands: one at 11 and 14 GHz, and another at 19 and 29 GHz. At 11 and 14 GHz, 500 MHz is available for the down- and up-links, respectively. These bands will have to be shared with terrestrial services and are therefore subject to power flux density limitations. There is 2,000 MHz of shared spectrum available between 17.7 and 19.7 for down-links and another 2,000 MHz between 27.5 and 29.5 GHz for up-links. Finally, bandwidths of 1,500 MHz will be exclusively available for down- and up-links, respectively, between 19.7 and 21.2 GHz and 29.5 and 31.0 GHz.

Aside from frequency reuse techniques by separate antenna beams applicable in principle to any frequency

band, the WARC 1971 frequency assignment results in bandwidth availability eight times greater than that at 4 and 6 GHz. The decrease in antenna size and the requirements of sharing, which are reduced or eliminated, make the new frequency bands extremely attractive. Future plans for INTELSAT and domestic or regional systems include the use of the 11/14-GHz bands as the RF hardware at these frequencies is in an advanced state of development and ready to be space-qualified. The 19/29-GHz bands will be used later as the art progresses and as the traffic needs require. Attenuation resulting from rain will be counteracted by diversity techniques and by adequate power margins when possible.

The use, and reuse, of all of the above-mentioned frequency bands will provide striking improvements. It has been pointed out (Pritchard, 1972; Knopow, 1974) that communications capacities of the order of 100,000 telephone circuits per satellite or greater can be attained without exceeding the mass and weight limits already encountered in satellites of the current generation (e.g., the INTELSAT IV or IV-A). The increase in capacity, which amounts to at least one order of magnitude, will be made possible by the expanded adoption of frequency reuse techniques via multi-beam antennas and orthogonal polarization, and also by the introduction of the following technologies:

- (a) 3-axis body stabilization,
- (b) higher efficiency solar cells,
- (c) higher efficiency energy storage devices,
- (d) electrical propulsion for stationkeeping and positioning,
- (e) onboard switching combined with time- and space-domain multiple-access techniques,
- (f) linearized transponders,
- (g) hybrid modulation techniques,
- (h) source encoding, and
- (i) intersatellite links.

The consequence of the above, plus the overall increased reliability

and longer lifetime in orbit of the spacecraft, will be a progressive reduction of the space segment cost of satellite systems. Further overall cost reductions will derive from the application of new concepts in earth terminal design (Pollack, 1974). Finally, it is expected that the development of new space transportation systems (space shuttle and tug) will also be beneficial.

Orbit Utilization

On the basis of the following assumptions, it can be stated that the noise in each circuit is produced by radiation spillover:

- (a) ideal modulation-demodulation processes;
- (b) neglect of thermal noise, i.e., no power limitations;
- (c) finite available communications bandwidth;
- (d) sharing of this bandwidth by satellites uniformly spaced in equatorial synchronous orbit; and
- (e) earth station antennas as uniformly illuminated apertures.

Thus (Bradley, 1968), the noise power is constrained by an upper bound that is a function of a single geometric variable, that is, the satellite spacing. Neglecting differences in slant range and choosing an optimum spacing $\Delta\theta = \lambda/D$ results in a communications capacity per unit angle of

$$C = 2 \frac{D}{\lambda} \text{ bps/Hz/rad} \quad (6)$$

Consequently, the information rate (bps) which can be handled by a segment of synchronous orbit spanning θ rad with an available (common) bandwidth B is

$$R = 2 \frac{D}{\lambda} B\theta \text{ bps} \quad (7)$$

A bandwidth of 500 MHz and antennas of 30-m diameter at 6 GHz would yield a theoretical global capacity of 3.77×10^{12} bps, or roughly 10^8 telephone channels.

In practice, these theoretical results must be corrected to take into account mechanical problems of tight

orbital spacing, real modulation-demodulation processes, effects of thermal noise, and actual antenna configuration. While the first three items would lead to lower communications capacities, the last item leads to an increase in communications capacity.

Other possibilities of augmenting the communications capacity are inter-satellite relaying, increase in the allowable interference ratio, channel interleaving, reversed use of frequencies, and pseudostationary satellites and 2-dimensional orbit space (Bargellini, 1969).

Since the two problems of spectrum and orbit utilization are inseparable, studies of all these possibilities are in progress.

CONCLUSIONS

Striking progress has been made in technology, operational use, institutional arrangements, and cost effectiveness of communications satellite systems in their first decade of deployment. This progress is continuing with expanded use. Existing international, domestic/regional, and mobile systems will grow and new ones will come into being.

Future systems will have greater capacity as well as flexibility to serve a variety of users, increased reliability, and longer lifetime. The problems of interconnecting satellite systems with ground networks and also among themselves will pose unique challenges.

The higher frequency bands will be progressively occupied in addition to the continued use of the lower bands already occupied. Advanced spacecraft, signal processing, and microwave technologies, combined with earth terminal development slanted toward automated operations and greater reliability, will contribute to the progress of satellite communications systems well beyond the achievements recorded in the first decade of commercial operation.

REFERENCES

Ahmed, N., and T. Simms, "The Promise of Canadian Satellite Communications," Telecommunication Journal, Vol. 40-IV/1973, pp. 198-206.

Bargellini, P. L., "Extension of the

Concept of Satellite Communication System Capacity to a Two-Dimensional Model," Conference Record, 1969 IEEE International Conference on Communications, June 9-11, 1969, Boulder, Colorado, pp. 37-25-37-29.

Bargellini, P. L., "Commercial Communications Satellites," Proceedings. Microwave '73 Conference, 19-21 June 1973, Brighton, England, pp. 483-493.

Bradley, W. E., "Communications Strategy of Geostationary Orbit," Astronautics and Aeronautics, Vol. 6, No. 4, April 1968, p. 34.

Briskman, R. D., et al., "19/28 GHz Radio Propagation Research," IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-10, No. 4, July 1974, pp. 535-536.

Browne, S., "The INTELSAT Global Satellite Communications System," COMSAT Technical Review, Vol. 4, No. 2, Fall 1974, pp. 477-488.

Chinnick, R. F., "The Canadian Telecommunications Satellite System," Journal of the British Interplanetary Society, Vol. 26, 1973, pp. 193-202.

Clarke, A. C., "Extraterrestrial Relays," Wireless World, October 1945, pp. 305-308.

Cox, D. C., "Design of the Bell Laboratories 18 and 28 GHz Satellite Beacon Propagation Experiment," Conference Record, 1974 IEEE International Conference on Communications, June 17-19, 1974, Minneapolis, Minnesota, pp. 27E-1-27E-5.

Dicks, J., and M. Brown, Jr., "INTELSAT IV-A Transmission System Design," COMSAT Technical Review, Vol. 5, No. 1, Spring 1975, pp. 73-104.

Edelson, B. I., et al., "INTELSAT System Reliability," International Astronautical Federation XXVth Congress, Amsterdam, The Netherlands, 30 September-5 October 1974.

Edelson, B. I., et al., "Cost Effectiveness in Global Satellite Communications," to be presented at the XXVIth International Astronautical Federation Congress, Lisbon, Portugal, September 21-27, 1975.

Hume, R., "A Body-Stabilized 24 Transponder Domestic Satellite," AIAA

Annual Meeting, 25 February 1975,
Washington, D.C.

Kiesling, J., et al., "A Technique for Modeling Communications Satellites," COMSAT Technical Review, Vol. 2, No. 1, Spring 1972, pp. 73-104.

Knopow, J. J., "Next Generation Communications Satellites," Communications Satellite Systems (AIAA Progress in Astronautics and Aeronautics, Vol. 32), P. L. Bargellini, ed., Cambridge, Massachusetts: M.I.T. Press, 1974, pp. 63-86.

Kreutel, R. W., "The Orthogonalization of Polarized Fields in Dual-Polarized Radio Transmission Systems," COMSAT Technical Review, Vol. 3, No. 2, Fall 1973, pp. 375-386.

Kwan, R. K., "The Telesat TDMA System," Conference Record, 1975 IEEE International Conference on Communications, June 16-18, 1975, San Francisco, California, pp. 44/1-44/3.

Lipke, D., and D. Swearingen, "Communications System Planning for MARISAT," Conference Record, 1974 IEEE International Conference on Communications, June 17-19, 1974, Minneapolis, Minnesota, pp. 29B-1-29B-5.

Maillet, W., "INTELSAT's 50 Mbit/s TDMA-2 System," Second International Symposium on Digital Communications, Paris, France, November 1971.

Metzger, S., "Foreword," COMSAT Technical Review, Vol. 2, No. 2, Fall 1972, pp. vii-xi.

Pierce, J. R., "Orbital Radio Relays,"

Jet Propulsion, Vol. 25, April 1955, pp. 153-157.

Pollack, L., and W. Sones, "An Unattended Earth Terminal for Satellite Communications," COMSAT Technical Review, Vol. 4, No. 2, Fall 1974, pp. 205-230.

Pritchard, W. L., and P. L. Bargellini, "Trends in Technology for Communications Satellites," Astronautics and Aeronautics, Vol. 10, No. 4, April 1972, p. 26.

Rosen, H., "Synchronous Communications Satellites," Space Communications, A. V. Balakrishnan, ed., New York: McGraw-Hill, 1963, Chapter 17.

Rusch, R. J., "COMSAT General Domestic Communications Satellite," Proceedings of the National Telecommunications Conference 1974, December 2-4, 1974, San Diego, California, pp. 122-127.

Sanders, R. W., "Communications Efficiency of Several Communications Systems," Proceedings IRE, No. 48, April 1960, p. 575.

Shannon, C. E., "The Mathematical Theory of Communication," Bell System Technical Journal, Vol. 27, No. 3, July 1948, pp. 379-428; Vol. 27, No. 4, October 1948, pp. 623-656.

"The INTELSAT IV Communications System," P. L. Bargellini, ed., COMSAT Technical Review, Vol. 2, No. 2, Fall 1972, pp. 437-572.

U.S. Congress, Communications Satellite Act of 1962, Sec. 102(a), Public Law 87-624, 87th Congress, H.R. 11040, August 31, 1962.

CURRENT AND FUTURE COMMUNICATION SATELLITE TECHNOLOGY*

Walter E. Morrow, Jr.

Lincoln Laboratory, Massachusetts, Institute of Technology, Lexington, Massachusetts 02173, U.S.A.

BACKGROUND

It is difficult for most of us to realize that less than 20 years ago we were struggling to launch a grapefruit-size satellite into low-earth orbit. Since that time, we have become accustomed to many space spectaculars. The later manned launches and recoveries were even pre-empted from full television coverage by ordinary television programming. Despite the rapid acceptance of the conquest and use of space, very few truly recognize the revolutionary effects which space capability has had on our lives and the way we do business. This is particularly true in the field of communications. And these changes will be followed by even more dramatic uses of space communications.

I would like to discuss briefly some of the technical background of this rapidly developing field.

I have not made an extensive literature search for the first serious suggestion for the use of satellites as communication transponders. But a remarkably accurate forecast of the potential was given in 1945 by Arthur C. Clarke in an article, "Extra-Terrestrial Relays," in Wireless World (see Fig. 1).

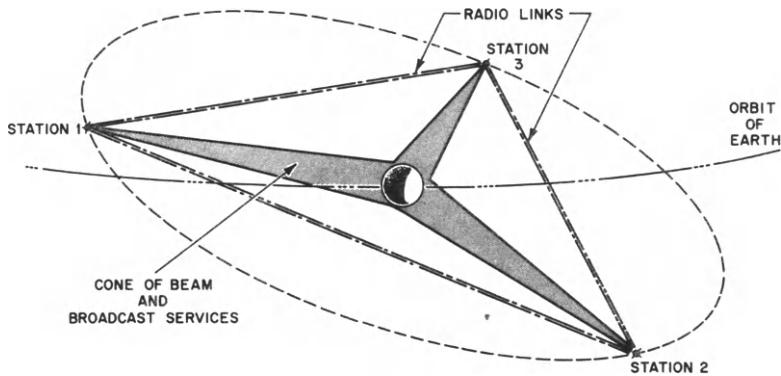


Fig. 1. Three satellite stations would ensure complete coverage of the globe.

Clarke relates the orbiting of a spacecraft to the development of the long-range rocket of which the V-2 was the prototype. He then goes on to explain

*This work was sponsored by the Department of the Air Force.

"astronautics" or the simple fundamentals of orbital mechanics, and explains the stationary equatorial orbit — or synchronous orbit — and discusses use of solar power to energize the transmitters.

In 1955, and again in 1959, Dr. John R. Pierce of the Bell Telephone Laboratories proposed the design of possible transoceanic satellite communication systems employing large numbers of metallized balloons in 1000-mile altitude orbits.

The art developed rapidly from 1960 onwards. Some of the key events and dates are listed in Table I. The prime emphasis was on telephonic communications between large, fixed antenna installations using microwave frequencies. The rapid increase in capacity is notable.

TABLE I Events in History of Civilian Communication Satellites

Synchronous Satellites Proposed by Clarke	1945
Medium Altitude Satellites Proposed by Pierce	1959
ECHO Launched — 1 Voice Channel	1960
TELSTAR Launched — TV	1962
COMSAT Act Passed	1962
SYNCOM II Launched	1963
International Consortium Formed	1964
Early Bird Launched — 240 Voice Channels	1965
MOLNIYA — 60 Voice Channels	1965
INTELSAT IV — 5000 Voice Channels	1971

Only a single voice channel could be transmitted by the passive ECHO balloon. TELSTAR (Fig. 2), developed by BTL, could transmit TV using an active microwave radio repeater in the satellite. Both of these satellites operated at altitudes of 1000 to 4000 miles. An operational system required dozens of such satellites for continuous service.

By 1963, satellites operating in synchronous orbit, such as SYNCOM (Fig. 3), developed by Hughes, and LES-4 (Fig. 4), developed by M.I.T. Lincoln Laboratory, had been launched. These satellites achieved roughly the same capacities as TELSTAR, but required only one satellite to achieve continuous communications. In 1965, commercial service was established with the launch of INTELSAT I into synchronous orbit.

Also, in 1965, the USSR launched the MOLNIYA satellite (Fig. 5) into a highly inclined and elliptical orbit which favors communication in northern areas. Only two satellites are required in this orbit to achieve continuous communications across the USSR.

The relationship between the three types of orbits can be seen in this sketch (Fig. 6). About 28,000 feet per second mission velocity is required for the low orbits. About 40,000 feet per second is required for synchronous orbit. Thus, the price for reducing the number of required satellites from dozens to one or two is about 12,000 feet per second or about a factor of three or four in payload mass — a well worthwhile trade-off.



Fig. 2. TELSTAR Satellite

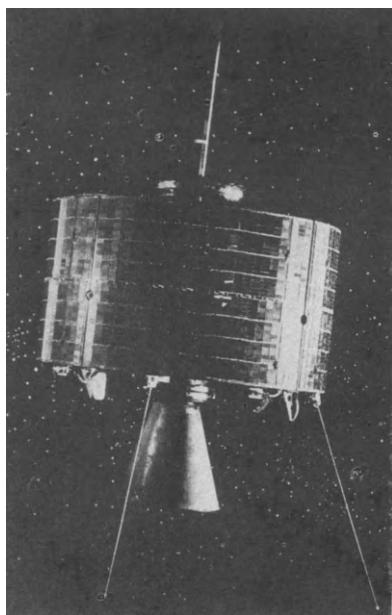


Fig. 3. SYNCOM Satellite

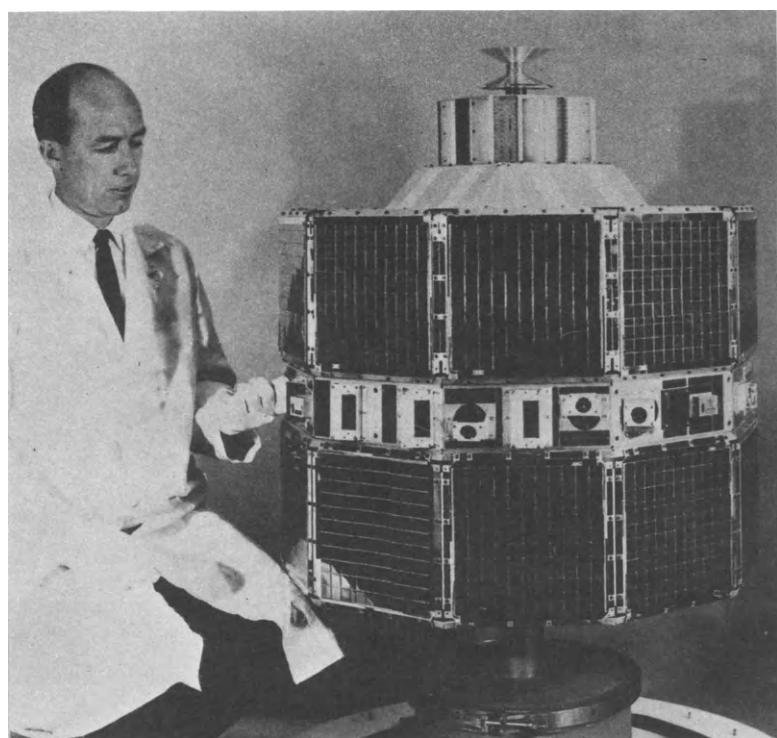


Fig. 4. Lincoln Experimental Satellite No. 4 (LES-4)



Fig. 5. MOLNIYA I

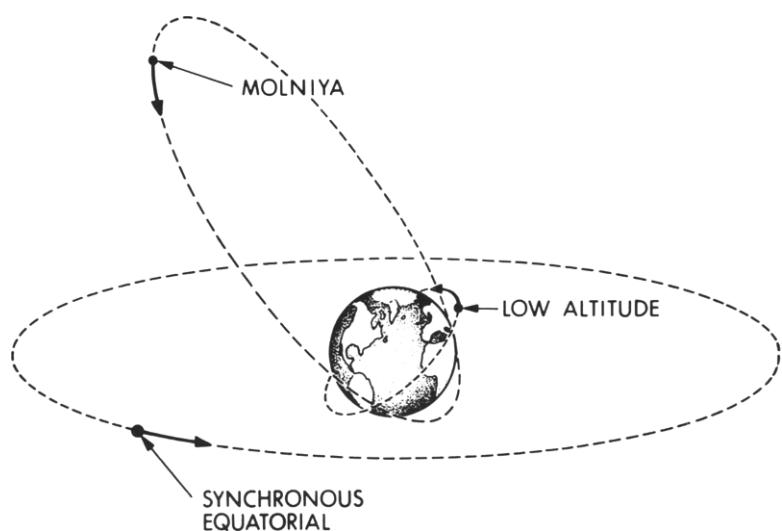


Fig. 6. Types of Communication-Satellite Orbits

Up to that time, all of the satellites used antennas with beamwidths which covered the whole earth. The next innovation involved the use of antennas with less than earth coverage beamwidth. The INTELSAT IV, launched in 1975 (Fig. 7), had antennas with beamwidths of only a few degrees and, by this means, together with more radio transmitter power, achieved 5,000 voice channels capacity between large, fixed terminals.



Fig. 7. INTELSAT IV

The rapid advances in communication satellite technology have led to a marked decrease in the cost per satellite telephone circuit between large, fixed ground terminals. The effect of these advances is shown in Table 2, which uses data taken from the "1973 Annual Report of the COMSAT Corporation".

TABLE 2 INTELSAT Series

	I	II	III	IV	IVA
Year Launched	1965	1967	1968	1971	1976
Antenna Gain (dB)	6	10	17	31	31
RF Power (W)	8	8	8	50	80
Weight (lb)	85	190	332	1585	~1600
Capacity (in voice circuits)	240	240	1500	4000	7500
Design Lifetime (years)	1.5	3	5	7	7
Investment per Operational Satellite (\$M)	12.8	10.8	15.7	40.4	60.0
Cost per Circuit Year (\$K)	35.6	15.0	2.1	1.4	1.1

You will note that the cost per circuit year has decreased by 20 or 30 fold over a 10-year period.

The principal technical advance which has permitted this improvement is the increase in antenna gain from 6 dB (4 times isotropic) in SYNCOM to 31 dB (1200 times isotropic) in INTELSAT IV and IVA. The increase in antenna gain just about matches the cost reduction.

Of course, other technical advances have also been significant. For instance, improved reliability of satellite propulsion systems and electronics has improved satellite lifetimes by a factor of 5.

FUTURE SERVICES

To date, most communication satellite service has been between large, fixed terminals. In the future, a growing desire for new forms of service can be anticipated. A list of some of these possible future services is shown in Table 3.

TABLE 3 Possible Future Communication Satellite Services

- A. Single-Channel Circuits to Ships
- B. Single-Channel Circuits to Aircraft
- C. TV Broadcast Service to Small Installations
- D. Single-Channel Voice Circuits to Land-Mobile-Installations
- E. Single-Channel Voice Circuits to Land-Fixed-Installations

Some experience with these classes of service has been obtained as a result of experimental satellites launched by NASA and by the Department of Defense. Generally, these experimental systems have used frequencies between 130 Mc and 1000 Mc because the mobile terminals can use inexpensive omni-directional antennas. Some of these early systems are shown in the following figures.

Lincoln Laboratory Experimental Satellite No. 6 (Fig. 8) has provided service

to mobile Department of Defense platforms over the past 6 years at frequencies of about 300 Mc.

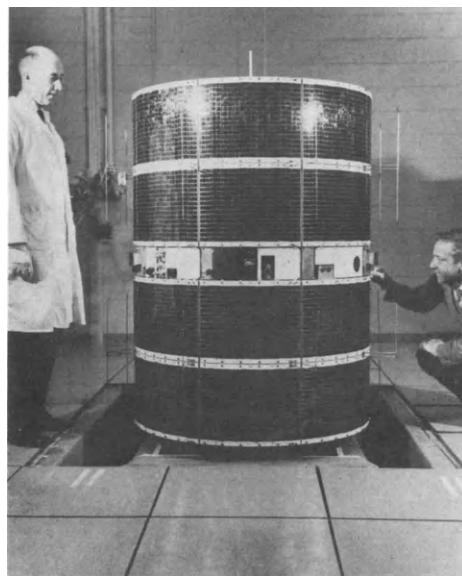


Fig. 8. Lincoln Experimental Satellite No. 6 (LES-6)

TACSAT (Fig. 9), another Department of Defense satellite, launched after LES-6, provided a somewhat greater capacity to mobile terminals by means of a more directive 300 Mc antenna.

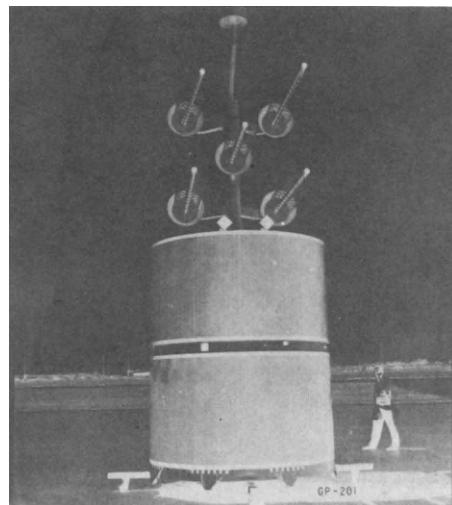


Fig. 9. TACSAT

MARISAT (Fig. 10), which is planned for launch in the Summer of 1975, will provide services to civil and Navy ships.

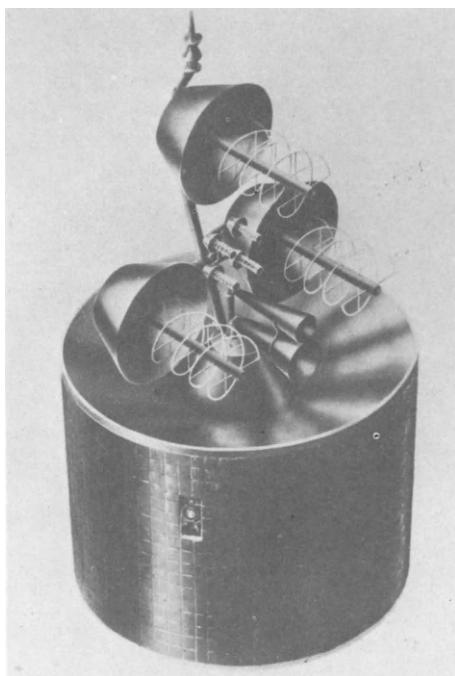


Fig. 10. MARISAT

The general characteristics of these satellites are more varied than the INTELSAT series. Table 4 indicates some of the general features. It should be noted that the capacities to the mobile terminals are quite limited compared with that between large, fixed INTELSAT terminals. Of course, considerable sharing between mobile users is possible because of the intermittent nature of the traffic. Even so, the effect of the small antennas used by the mobile terminals is clear.

TABLE 4 Communication Satellites Providing Mobile Service

	<u>LES-6</u>	<u>TACSAT</u>	<u>MARISAT</u>
Launch Date	1968	1969	1975
Weight (lb)	350	1500	720
Frequency (MHz)	300	300	300 1600
Antenna Gain (dB)	10	14	10 13
Transmitter Power (W)	100	240	100 50
Approximate Capacity to Small Terminals in Numbers of Simultaneous Half-Duplex FM Voice Channels	1	5	1 5

If satellites of the TACSAT class have lifetimes and costs similar to INTELSAT IV, the yearly costs per half-duplex voice circuit amount to about \$1 million or nearly 1000 times the cost of satellite circuits between large, fixed terminals. The difference is due, of course, to the very small size of the surface terminal antennas. Some communication users, such as the Department of Defense and owners of large ships and transoceanic aircraft, can profitably employ such costly circuits, particularly if they share the circuit among a number of terminals. More general application of satellite communications to inexpensive mobile and fixed terminals awaits a substantial reduction in cost through technical innovation.

IMPACT OF FUTURE TECHNOLOGY DEVELOPMENTS

In this section, the possibilities of future technology developments will be considered. It should be pointed out that current communication satellites already represent a very high state of technology. Further reductions in cost of satellite communication for mobile and small, fixed terminals will require significant technical advances.

In most satellite communication systems, the system capacity is determined by the downlink from the satellite to the ground. This follows, since it is invariably easier to generate radio frequency power on the ground than in a satellite. Hence, the uplinks have larger margins than the downlinks, assuming equal antenna gains for both directions.

The key parameters which determine downlink capacity are listed in Table 5. The satellite transmitting antenna gain determines the ability of the satellite to focus power on the surface receiver; the role of the satellite transmitter power is obvious. The terminal receiving antenna area determines the amount of signal that can be captured. The receiver noise level then sets the received signal-to-noise ratio and, finally, the modulation system determines how well the signal power can be used.

TABLE 5 Key Technical Characteristics which determine Capacity of Downlinks

- A. Satellite Transmitting Antenna Gain
- B. Satellite Transmitting Power
- C. Terminal Receiving Antenna Area
- D. Terminal Receiver Noise
- E. Modulation System

Table 6 indicates typical downlink parameters for contemporary fixed multi-channel systems and for small single-channel terminal systems. The capacity advantage of the large, fixed multi-channel systems is mostly due to the large size of the receiving aperture. This, of course, is the one parameter not available to a small, inexpensive terminal. How might this disadvantage be made up? The principal opportunity lies in the use of large, high-gain satellite antennas. Only modest gains can be anticipated in modulation system efficiencies, receiver noise level, and satellite transmitter power.

TABLE 6 Comparison of Contemporary Satellite Communication Systems Servicing Large Multi-Channel Terminals and Small Single-Channel Systems

	<u>Large Multi-Channel</u>	<u>Small Single-Channel</u>
Satellite Antenna Gain (dB)	15 to 30	15
Satellite Transmitter Power (W)	100	200
Terminal Antenna Size (ft in diameter)	85	1 to 2
Terminal Receiver Noise Level ($^{\circ}$ K)	100	600
Modulation System for Voice	FM	FM
Typical Capacity	5000	5

Large Spacecraft Antennas

The possibility of large spacecraft antennas seems antithetical to the notion of spacecraft weighing, at most a few thousand pounds. After all, the standard 85-foot ground antennas weigh hundreds of thousands of pounds. There is one large difference, however, between the surface of the earth and space; namely, gravity and wind forces are absent from space. It turns out that rather large space antennas can be built which weigh only a few hundred pounds.

The NASA ATS-6 spacecraft incorporates a 30-foot parabolic antenna shown in Fig. 11. It weighs between 100 and 200 pounds and is operable to 10 GHz.

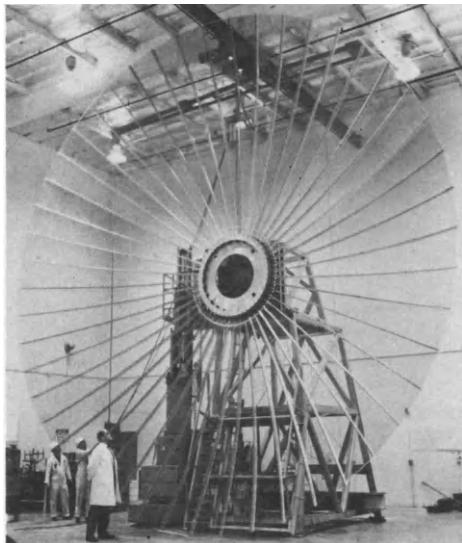


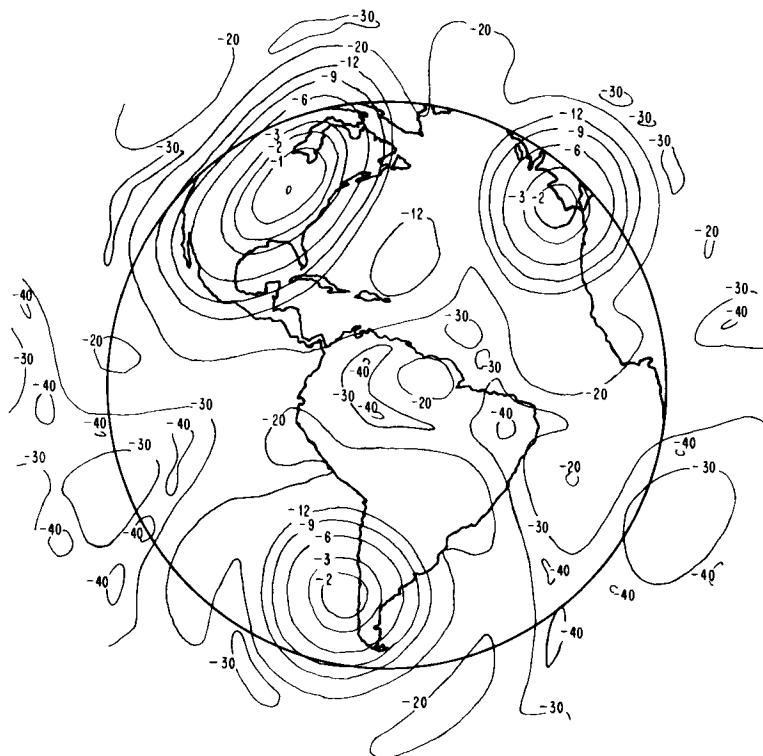
Fig. 11. ATS-F 30-foot diameter antenna

This antenna consists of a series of sheet aluminum ribs on which is stretched a metallized net. During launch, the antenna is packed into a small container by wrapping the ribs and mesh around a central hub. Upon reaching orbit, the ribs are released, whereupon they unwind into their deployed position.

Multiple Beams

One difficulty with these large antennas is that they produce very narrow beams and, therefore, will have limited coverage on the earth's surface. For instance, the ATS-6 30-foot antenna has a beamwidth of about 2° at about 1000 Mc.

Figure 12 indicates the beam size when it is aimed at various points on the earth's surface. If such an antenna is to be usefully employed over the earth's surface, it will be necessary to generate a number of antenna beams as are shown in Fig. 13. A total of 37 beams in this case is necessary to cover the earth. Thirty-seven beams do not mean that we need to launch thirty-seven 30-foot diameter antennas into space. It is entirely possible to share one antenna aperture among many beams.



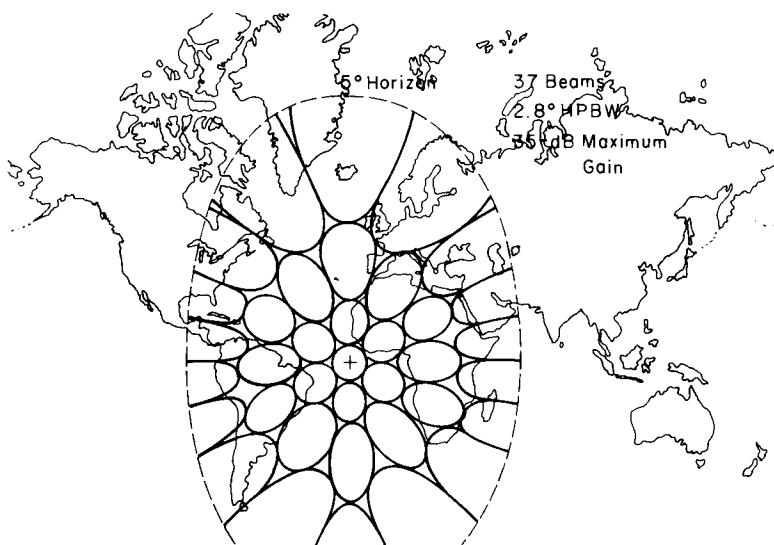


Fig. 13. Coverage of Multibeam Sync Orbit Satellite

As an example, Fig. 14 shows a Lincoln Laboratory developed X-band lens antenna of about 30-inches diameter, which is illuminated by 19 feed horns and which can simultaneously produce 19 beams, which will just cover the earth from synchronous orbit. The equipment in back of the feed horns permits the satellite transmitter to be connected to any combination of from 1 to 19 of the feed horns. The whole antenna system weighs less than 40 pounds.

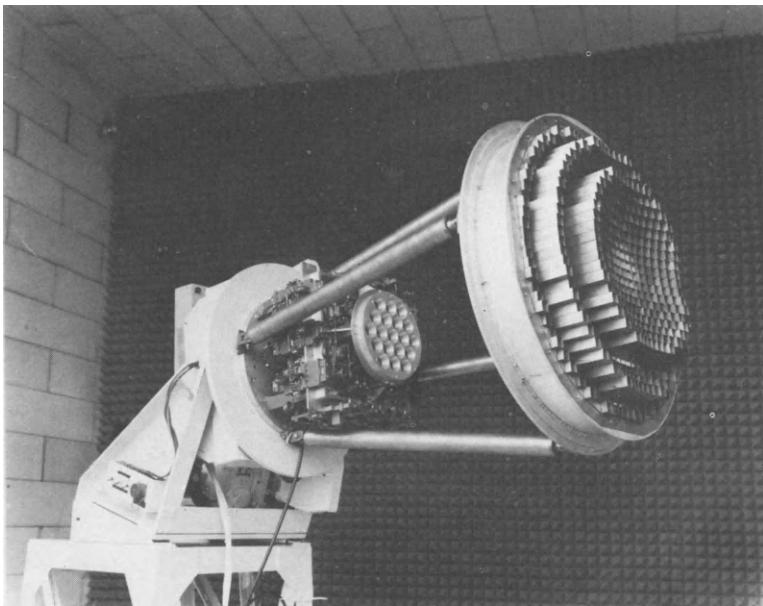


Fig. 14. 19-Beam X-Band Lens Antenna

Similar arrangements can be made for large parabolic reflector antennas. In that case, a cluster of antenna feeds is located at the focal plane of the parabola as shown in Fig. 15.

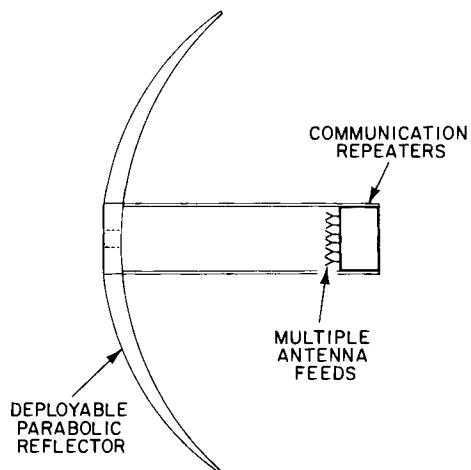


Fig. 15. Multibeam Communication Satellite with Large Deployable Antenna

Precision Antenna Aiming

With currently available spacecraft technology, precision aiming to 0.1° is relatively easy. With large antenna structures, it is advantageous to attach rigidly the rest of the spacecraft to the antenna and aim the structure as a whole.

Figure 16 indicates the geometry involved. In order to maintain beam-pointing, the satellite location in space must be known, the directional vector to the earth determined, and then pitch, roll, and yaw maneuvers performed to aim the collection of antenna beams towards the required locations on earth or to another spacecraft.

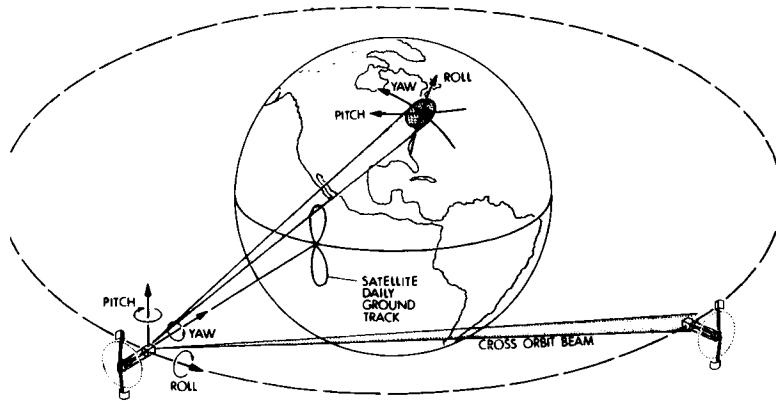


Fig. 16. Motions of Spot Beam on Earth Surface

The spacecraft location can be determined by means of a series of ground-based observations of satellite range and range rate, or by means of an on-board system shown in Fig. 17.

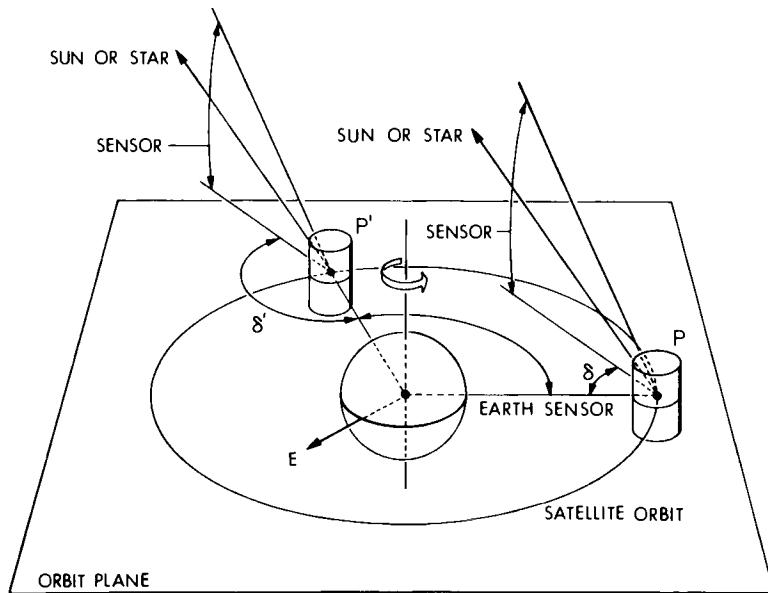


Fig. 17. Sensing Geometry for an Automatic Station-Keeping Satellite System

The on-board system, which was incorporated in Lincoln Experimental Satellite No. 6, involves the use of a satellite precision chronometer and visual and/or infrared sensors angle sightings of the sun and earth's edge. The satellite location in orbit is determined by noting the time at which the observed angle between the sun and earth reaches a given value.

These same sensors can also be used to obtain pointing directions towards the earth.

A spacecraft, with a large antenna, can be turned in space by means of an on-board momentum wheel or wheels as shown in Fig. 18. By speeding up and slowing down the wheel speed, pitch maneuvers can be made. Pivoting of the wheel axis can produce roll and yaw motions.

The spacecraft must also be kept in proper orbital position. This is customarily accomplished by hydrazine-fueled thrusters; a cross-section of a typical design is shown in Fig. 19. Ammonia thruster systems, such as the system shown in Fig. 20, can also be used. Electrically-powered thrusters have also been used. Figure 21 shows a pulse plasma thruster which was used on LES-6.

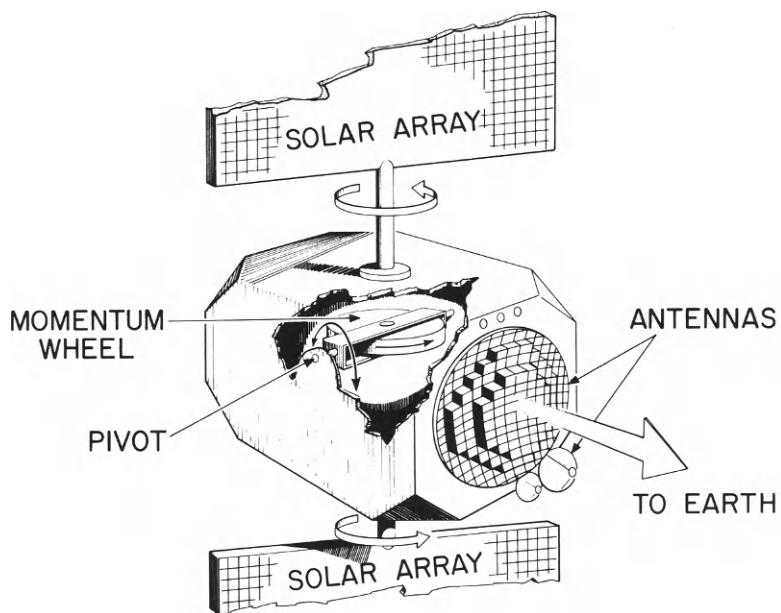


Fig. 18. LES-7 Stabilization

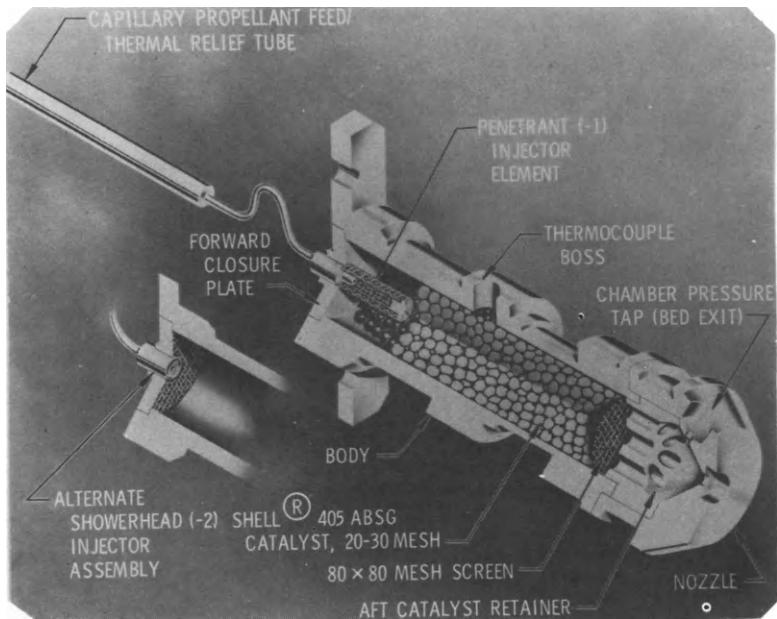


Fig. 19. Typical Hydrazine Thruster



Fig. 20. Ammonia Cold Gas Thruster System

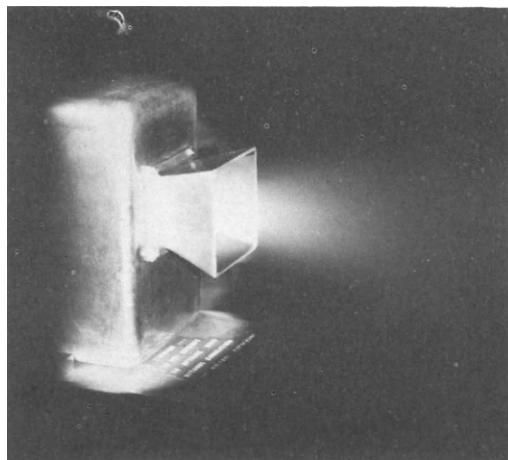


Fig. 21. Arc Thruster

On-Board Message Switching

The use of multiple high-gain satellite antennas will clearly provide substantial increases in capacity to small terminals. On the other hand, it raises the problem of how to interconnect users who are located in different beams.

One solution to this problem might be to transmit the signals from the various beams to the ground on a very wide-band downlink to a large ground terminal as shown in Fig. 22. The interconnection could then be made by conventional switching equipment and multibeam transmit signals returned to the spacecraft on a wide-band link for distribution to the multiple downlink beams.

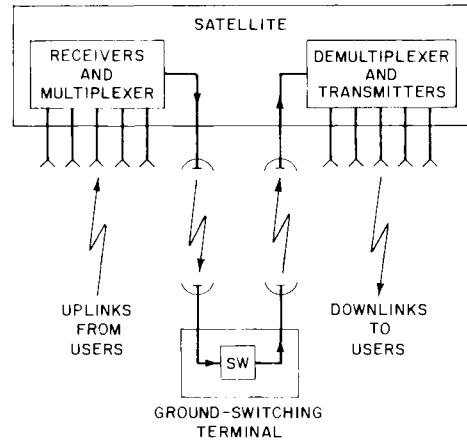


Fig. 22. Ground-Switched Multiple Beam Satellite System

Such a solution, while permitting the complex switching equipment to be located on the ground, would require very wide-band channels to the control station. In addition, the problems of extra satellite power and an additional time delay of 0.25 second would add to the already existing 0.25 second time delay.

A better solution seems to be to perform the switching in the satellite, as shown in Fig. 23. The figure shows the switching of demodulated signals.

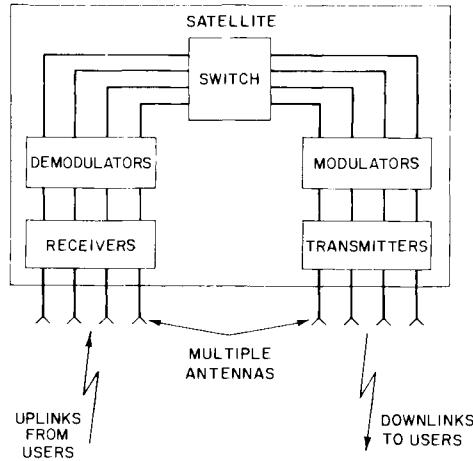


Fig. 23. Satellite-Switched Multiple Beam System

An alternative is to switch a radio frequency which avoids the complexity of demodulation, but raises problems of time conflicts in the use of the downlink transmitters.

The solution of these on-board switching problems is a significant technical challenge; however, recent advances in high-speed digital signal processors point the way to a successful solution.

For instance, Fig. 24 shows a digital voice signal processor recently developed at the Laboratory. This equipment can perform more than 20 million computer observations per second. It is sufficiently small to be carried in a satellite. A variation of this design should be able to carry out the demodulation and switching of a large number of channels.

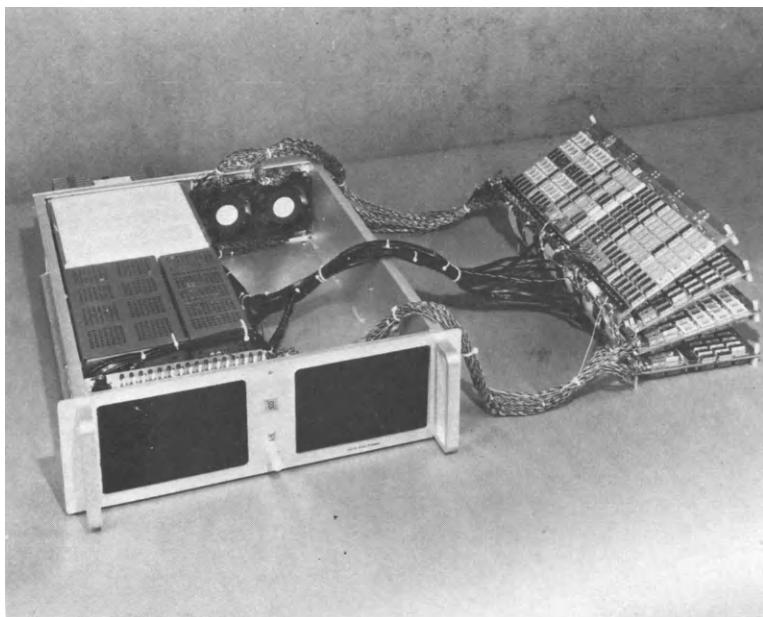


Fig. 24. Digital Voice Processor

Higher Satellite Power

Another obvious way to increase satellite capacity is to increase the satellite transmitter power. The transmitter power output is the product of the available prime power and the efficiency of the transmitters.

There is relatively little opportunity to increase the efficiency of current satellite solid-state transmitters operating at frequencies up to 2000 Mc. Figure 25 shows a 60% efficient 300 Mc solid-state transmitter developed for the Lincoln Experimental Satellite No. 6.

At frequencies above 2000 Mc, traveling wave tubes are commonly used. Efficiencies of up to 40% are common.

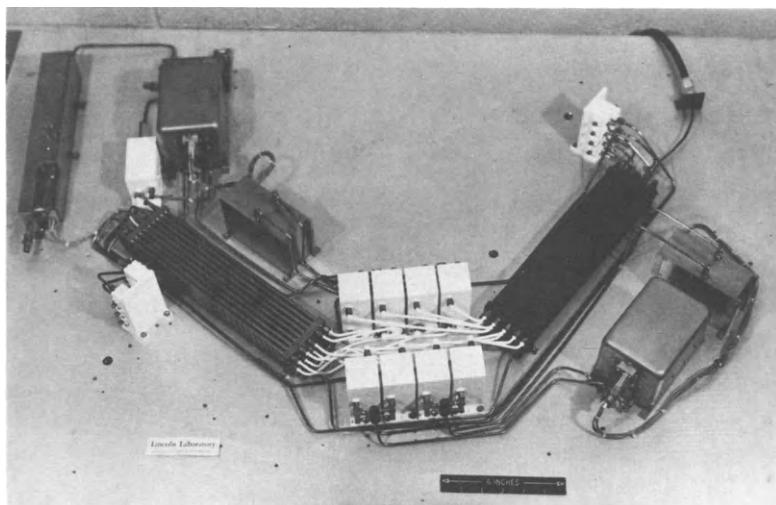


Fig. 25. 80-Watt Solid State UHF Transmitter

It is possible that significant advances might be made in the area of prime power systems. Most current satellites employ silicon solar cell power systems, having efficiencies of the order of 10%. The lightest arrangement involves solar-oriented planar arrays, as shown in Fig. 26. In such arrays, 10 watts of power per pound of array is not uncommon. New designs, involving more efficient cells on lightweight flexible substrates, should be able to produce 50 watts per pound, a five-fold increase.

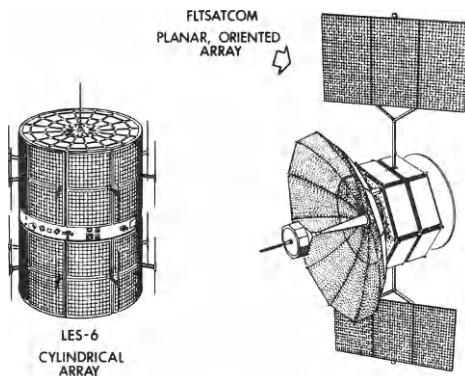


Fig. 26. Solar Arrays

It may also be possible to develop even larger power per pound of weight by means of deployed parabolic solar concentrators which could be used with either solar cells or perhaps Brayton closed-cycle turbo-alternators.

Terminal Receiver Noise

Contemporary equipments are already obtaining noise levels which are approaching or below that of the ambient background, especially for mobile terminals operating at the lower frequencies. There is little opportunity for further improvement in this area.

Modulation Systems

Most contemporary systems employ analogue frequency modulation voice and TV transmission. For FM voice systems, a ratio of signal power-to-noise power in a one-cycle band of 50 dB is required as shown in Table 7. Digital speech transmission systems operating at 2400 bits per second with very efficient modulation systems have been demonstrated to operate at signal power-to-noise ratios of about 40 dB. While currently these systems are far too expensive to be employed by inexpensive mobile terminals, the recent advances in the reduction of the cost of digital equipment indicates the possibility of low-cost voice systems operating at significantly lower signal-to-noise ratios.

TABLE 7 Voice Modulation Systems

<u>System</u>	P_R/N_o (dB)
FM Voice	50
2400 Bits per Second Speech	40

Future System Possibilities

The potential of future satellite systems can be roughly determined by sketching out a hypothetical system using parameters derived from the projections given above.

Table 8 indicates the possible parameters and compares them with current small terminal systems. The Table indicates a 4000 to 1 increase in capacity. The advanced satellite will undoubtedly cost several times more than the contemporary satellite, still the cost per channel in the advanced system should be reduced by the order of 1000 times. Such a reduction should bring the cost of such channels down to a level comparable with that of channels between large, fixed multi-channel terminals. It is conceivable that costs as low as \$1000 per channel-year might be achieved. Since these channels could be shared by a number of users, using a demand access system, one can anticipate rather general availability of satellite communications to small mobile and fixed terminals.

TABLE 8 Advanced Satellite Communication System for Small Single-Channel Users

	<u>Current System</u>	<u>Advanced System</u>
Frequency (MHz)	300	1000
Satellite Antenna Gain (dB)(ft in diameter)	15 (15)	47 (90)
Satellite Transmitter Power (W)	200	1000
Terminal Antenna Size (ft)	1	1
Receiver Noise Level (o K)	600	600
Modulation (dB P_R/N_o)	FM (50)	Digital Speech (40)
Typical Capacity (voice channels)	~5	~20,000

THE SIRIO PROGRAMME

F. Carassa* M. Macchia** and L. A. Ciavoli Cortelli

*Politecnico di Milano and Centro Telecomunicazioni Spaziali of Consiglio Nazionale delle Ricerche

**Project Manager Sirio, Servizio Attività Spaziali del Consiglio Nazionale delle Ricerche

***TELESPAZIO S.p.A. per le Comunicazioni Spaziali

Introduction

This paper summarizes the features and the objectives of the Italian Sirio satellite programme and the characteristics of its major systems. Among these objectives the principal one is to investigate about the applicability of frequencies above 10 GHz to satellite communication systems, through an SHF propagation and communication experiment.

Organizational aspects

The Sirio programme is developed in the frame of the Italian space activity by the Consiglio Nazionale delle Ricerche (CNR); the projects management is at the Servizio Attività Spaziali (SAS) of CNR, whilst the experimental activity is under the scientific direction of the Centro Telecomunicazioni Spaziali of CNR at the Politecnico di Milano.

Prime contractorship for the Sirio satellite has been entrusted to Compagnia Industriale Aerospaziale (CIA) whilst the contractorship for the Sirio Operation Control Center and for the Sirio SHF ground stations has been entrusted to Telespazio. Telespazio is also giving technical and administrative assistance to CNR for the development of the programme. NASA will take care of the satellite launch.

The SIRIO spacecraft

SIRIO is a spin stabilized geostationary satellite with a mechanically despun antenna for the propagation and communication experiment. The nominal spin rate is 90 rpm.

The satellite will be launched from Cape Kennedy by means of an improved Thor Delta vehicle and put in geostationary orbit at a longitude of 15° West by an Italian non magnetic apogee boost motor.

- - Politecnico di Milano and Centro Telecomunicazioni Spaziali of Consiglio Nazionale delle Ricerche.
- - Project Manager Sirio, Servizio Attività Spaziali del Consiglio Nazionale delle Ricerche.
- - TELESPAZIO S.p.A. per le Comunicazioni Spaziali.

Its configuration, of cylindrical shape, is typical of the spin stabilized geostationary satellites. The cylindrical body has a diameter of 1433 mm and a height of 945 mm. The overall height, inclusive of apogee motor nozzle and SHF mechanically despun antenna, is 1999 mm. The antenna has paraboloidic shape with 320x300 mm main dimensions.

The total weight at launch, inclusive of apogee motor weight, is 398 kg; at the beginning of life in operative geostationary conditions it is 220 kg, decreasing at the end of life to 188 kg. The apogee motor weights 198 kg.

The electrical power, supplied by means of a solar array, is about 145 W at beginning of life in equinox conditions, decreasing to about 102 W at the end of life in summer solstice sunlight conditions.

The spacecraft is designed for an orbital lifetime of two years.

Technological, operational and managerial objectives of the programme.

The Sirio programme gives a very noticeable contribution to the qualification of the Italian industries in the field of space technology and related techniques.

From the operational point of view the programme will allow the acquisition of experience in carrying communications of various types, using the new frequency bands and advanced methods of modulation and transmission.

Also advantages in developing modern managerial techniques will be gained by Italy through the Sirio programme. These techniques will be useful also for the development of other programmes in the field of applied research. A tight cooperation between scientific and industrial groups is another major aspect of the Sirio programme.

Scientific objectives.

The SHF experiment aims to investigate about the propagation of frequencies above 10 GHz for application to satellite communication systems. As well known, the propagation of these frequencies through the atmosphere is affected by water vapour and oxygen absorption and by precipitation effects. The latter effects are the most important at the frequencies of interest and are mainly produced by rain, whose drops absorb and scatter energy. The principal effect is attenuation, but also crosspo-

Iarization coupling and possibly bandwidth limitations have to be considered.

The frequencies chosen for investigation in the Sirio-SHF experiment are 12 and 18 GHz, which will be used on the satellite-to-earth and on the earth-to-satellite paths, respectively. The on-board equipment is a transponder which has been designed to allow propagation measurements as well as communication experiments; it has therefore three modes of operation, namely: - a propagation mode; - a narrow-band communication mode (1,5 MHz at -1 db); - a broadband communication mode (32 MHz at -1 db). These modes of operation have been described in 1 and 2 (F. Carassa, 1971) and will be summarized very briefly in the present paper.

In the propagation mode, up to four frequencies have access to the satellite as shown in fig. 1a.

In this figure the input spectrum to the satellite is shown consisting of two pairs of lines separated by a frequency interval of 770 MHz. These four lines are converted in a narrow-band (20 KHz wide) and added with an amplitude calibrated line (REF) generated on-board to be used as reference for attenuation measurements on the earth-to-satellite path (Fig. 1b). After further conversion to a final intermediate frequency of 265 MHz, the spectrum composed by the given lines is finally transmitted to earth by amplitude modulation of a carrier at 11.6 GHz (Fig. 1c). In addition to the sidebands the carrier is also transmitted at a calibrated level, and a telemetry carrier is added. The measurements reported in table 1 may be performed.

Special care has been taken in fixing the frequency plan, due to the need of avoiding possible problems arising from intermodulation products generated within the feed of the ground Antenna by flange connections.

Concerning communication experiments, the input spectrum at the satellite is shown in fig. 2a: the signal may be allocated 292 MHz either above or below the suppressed carrier at 17.4 GHz. On the satellite, the signal band is converted to intermediate frequency, and retransmitted 70 MHz below the carrier at 11.6 GHz (Fig. 2b). The carrier is also transmitted to allow attenuation measurements on the down link to be performed continuously.

In the narrow-band configuration the present experimental programme foresees to transmit up to 12 carriers digitally modulated by voice or data but a final decision on this experiment has not been taken yet.

In the wide-band configuration, experiments of television transmission will be performed. Frequency modulation will be used for

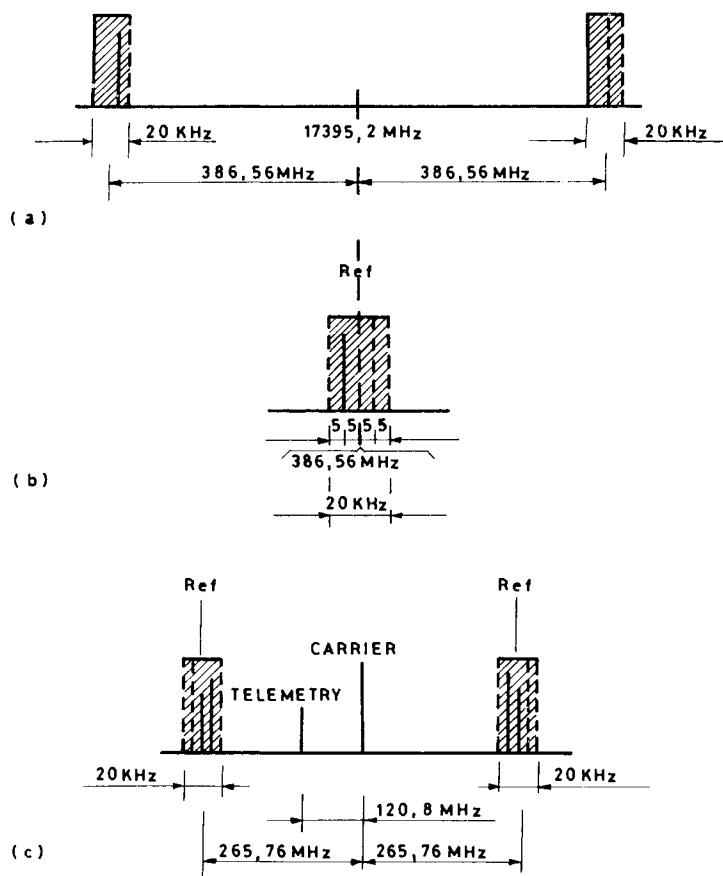


Fig. 1

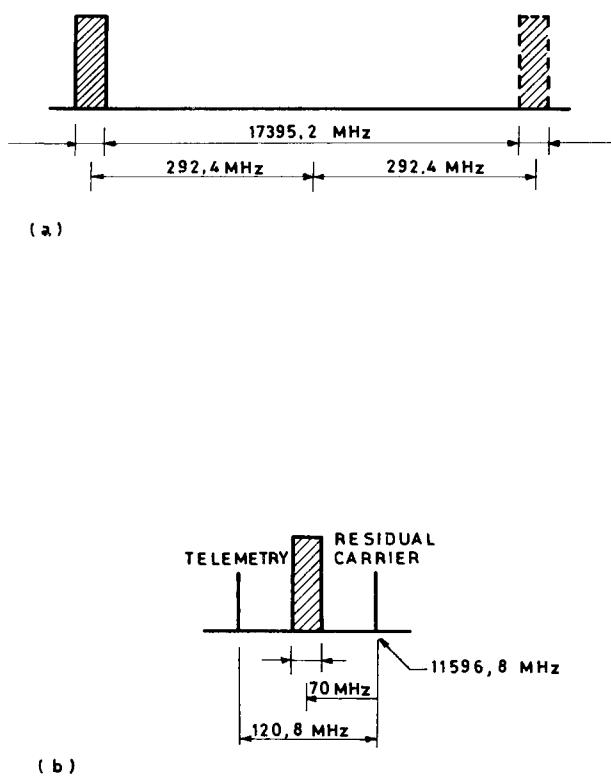


Fig. 2

Table 1 - Propagation measurements

Measurement	Method of measurement
- Attenuation at 11.6 GHz	Measure the level of the carrier received from the satellite
- Differential attenuation at 12 GHz on a frequency interval of 590 MHz	Compare the level of the two side bands received from the satellite
- Phase distortion at 12 GHz.	Compare the phases of the sidebands with the phase of the carrier.
- Attenuation in the 18 GHz band	Compare in a sideband of the signal received from the satellite the level of the line transmitted from earth with the reference line
- Differential attenuation at 18 GHz in a frequency interval of 770 MHz.	Compare, in a sideband of the signal received from the satellite, the levels of two lines transmitted from earth by a single station with the frequency interval of 770 MHz.
- Space diversity efficiency at 12 GHz.	Measure in two different sites at the same location the level of the carrier received from the satellite
- Space diversity efficiency at 18 GHz.	Transmit two different lines by two transmitters located in two different sites at the same location.

black-and-white or for colour signal with the sound channel transmitted on the synchronizing pulses. The threshold performance of the system will be improved by using a phase-lock demodulator followed by a spike suppressor developed at the Centro Telecomunicazioni Spaziali.

Experiments on digital television transmission at 17 Mb/s are presently under consideration, in association with studies on redundancy reduction presently under development at the Centro Telecomunicazioni spaziali, in cooperation with the Centro Televisione at the Istituto Nazionale Galileo Ferraris in Torino.

Digital transmission of a random test signal at 34 Mb/s is also being considered on request of the Research Institute of the German Post Office, which is one of the Institutions participating to the experiments.

Coverage areas of the SHF experiment

The Sirio satellite will be positioned in geostationary orbit at a longitude of 15° W. The mechanically despun antenna has two pointing positions, one over Europe and one over the Atlantic, so that, taking into account the characteristics of the antenna and the pointing errors, it is convenient to define (fig. 3) three conventional coverage areas (two when pointing over Europe and one when pointing over the Atlantic) to be used for the definition of the performance of the system. Accordingly, the minimum guaranteed values at the edges of the various coverages are the ones given in table 2.

Italian earth stations

Two main stations will be located in Italy in the same sites of the Telespazio stations, i.e. at Fucino, near Rome, and at Lario, near the Como lake, in the northern part of Italy. These stations use Cassegrain parabolic antennas having diameter of 17 m. and have a G/T of 39.5 and 36.5 dB/°K in good and bad weather conditions respectively and an EIRP at each transmitted frequency of 94 dBW. These stations are able to measure with a large margin, supplementary attenuations up to 20 dB at 12 GHz and up to 35 dB at 18 GHz. They are also able to perform the communication experiments described above.

One or two small stations for the propagation experiment only (and for the down link attenuation only) are also under final consideration for installation at the experimental field of the Centro Telecomunicazioni Spaziali at Spino d'Adda near Milano.

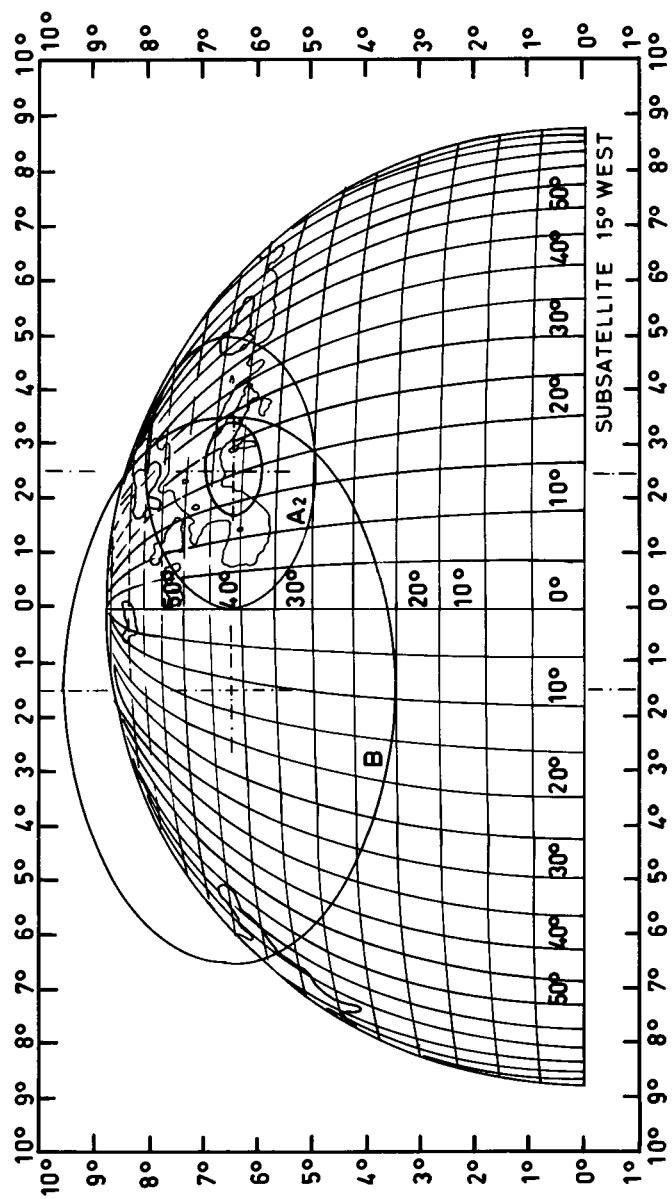


Fig. 3

Table 2 - Guaranteed values for the various coverages

Coverage	A ₁	A ₂	B
propagation	G/T	-15 dB/°K	-22 dB/°K
	Carrier EIRP	+44.5 dBm	+41 dBm
	Sidebands EIRP	+44 "	+40 "
	Carrier EIRP	+44.5 "	+41.5 "
	Signal EIRP	+54.5 "	+51.5 "
	Carrier EIRP	+42 "	+39 "
wide-B comm.	Signal EIRP	+57 "	+54 "

In this experimental field a meteorological radar will be operational in a few months. Also installations are foreseen in the field for participation to the propagation experiments at 13 and 18 GHz with the NASA ATS-6 satellite. The small station or stations for Lario, are foreseen to be subsequently used with the ESRO OTS satellite.

The Italian Navy will participate to the experiments with two small stations installed on two ships.

At present the possibility of performing a diversity experiment at 18 GHz is under study also. To avoid the frequency diversity existence together with the space diversity, it will be necessary to transmit from the two diversity sites (most likely Fucino and Celano) a pair of frequencies selected in the same up-link sideband.

The diversity terminal in Celano would be equipped with a small dish (1.5 m. diameter), a Klystron power amplifier and a $\times 100$ frequency multiplier.

The frequency transmitted from the diversity terminal would be

generated in the main site, divided by 100 and transmitted to Celano at VHF. Information concerning the power level to be radiated from the diversity site would be transmitted on a separate carrier.

The measurement dynamic range for the small terminal would be 20 dB, with the satellite antenna pointing to Europe.

Data acquisition, recording and analysis

The two stations of Fucino and Lario will be equipped with a data acquisition and recording system composed as follows:

- strip-chart recorders for quick-look
- digital system, including:
 - analog-to-digital converter
 - minicomputer with 16 Kwords memory
 - magnetic tape unit
 - disc unit
 - line printer, etc.

The digital system will acquire propagation data, rain data (clicks transmitted from tipping-bucket pluviometers), telemetry data transmitted from the satellite, plus some calibration data generated in the earth station.

These raw data will be recorded on magnetic tape. They will also be translated in engineering units and corrected from some major error sources in real time. The corrected data will be used to produce on line some significant statistical information, such as cumulative distributions, regression coefficients, correlation coefficients, variances, etc. It will also be possible to display on a video-terminal the mean values on a short-term basis.

A more refined data analysis will be performed off-line by a large computer. The production of tapes including the most significant events is also foreseen; these tapes will be produced according to a common format agreed upon with other experimenters, to ease the exchange of data of particular interest.

International participation.

Several important European Institutions have indicated, since the beginning, their interest in the experiment, and some of them are ready with the equipment. We quote among them: the Netherland Post Office (and the University of Eindhoven), the Research Institute of the German Post Office, the Appleton Labo-

ratory (U.K.), the CNET (French Post Office), the Helsinki University of Technology (Finland), the Institute for Broadcasting Techniques in Hamburg (Germany), the Technical University of Graz (Austria).

Present status of the programme.

The Development Prototype of the Satellite has been completed and laboratory tests on the overall system, including the ground transmitting and receive chains, have been carried out. Tests on a complete inclined link along the side of a mountain in the Fucino area, will be carried out at the beginning of the next year.

References

1. CARASSA F.: SIRIO SHF Experiment, Symposium on Long Term Prospects for Satellite Communications, 2-4 June 1971, Istituto Internazionale delle Comunicazioni.
2. CARASSA F.: Research Programme at Frequencies above 10 GHz and the Italian Satellite SIRIO-Proceedings of the fifth colloquium on microwave communication - Budapest, 24-30 June, 1974.

THE JAPANESE DOMESTIC COMMUNICATIONS SATELLITE SYSTEM

Kinnosuke Kawakami

Radio Research Laboratories, Ministry of Posts and Telecommunications, Koganei-Shi, Tokyo, Japan

OUTLINE

It can be said that the space development of Japan got into full scale activities in 1967 when the National Space Activity Council started its activities in a practical manner.

Later the space development was divided into two categories, namely, the scientific satellite program and practical satellite program. And with respect to each program, research in rockets, satellites, missions and ground facilities have been carried on.

As for the communication satellite including broadcasting, it was planned at the beginning to conduct experiments of both radio wave propagation and satellite communication covering a wide range of frequency band including millimeter wave by launching a geostationary satellite weighing 130 Kg by means of N-rocket which had been developed in Japan. This plan is being called "ECS-plan". The main object of this satellite is to survey how to use the radio wave having the frequency which is higher than the frequency band now in use for satellite communication, and to acquire techniques necessary for putting a satellite into geostationary orbit.

According to this plan, it was expected to launch the ECS in 1977, but for various reasons this plan has been postponed until 1979 and preparations are being pushed forward with this end in view.

On the other hand, apart from the survey of wide frequency band, the necessity of conducting experiments by satellite concerning the communication and broadcasting is stressed, and such experiments should be directly connected with the practical use and consequently the CSE and BSE projects are being carried forward.

It may seem strange to attempt to conduct the communication and broadcasting by satellite in such a country as in Japan which is limited in area and is situated at the middle latitude, and besides which is considered to be fairly well developed both in communication and broadcasting.

Nevertheless, Japan decided to venture this attempt mainly for the following reasons:

1. In order to secure necessary information transmission and communication in case of natural disasters.
2. In order that the communication may be available in the surrounding islands in the same manner as in the main island.

3. In order to acquire through experiments, operational techniques which are compatible with the information transmission other than the communication by satellite, providing for the future needs of communication.
4. In order to evaluate temporary circuits by satellite relay to cope with the sudden increase of communication demands occurring regionally.
5. In order to perform the TV broadcasting towards remote places and isolated islands, and to realize an equal distribution of its quality and quantity to any place in Japan.

With a view to coping with the above objects, the program of communication and broadcasting by satellites will be carried forward in such a way that researches and investigations of the first stage will be conducted by means of the experimental communication satellite (CSE) having medium capacity and the experimental TV broadcasting satellite (BSE) making the community reception an object.

The processes to be taken at the second and third stages mentioned below are considered to be reasonable by the persons concerned. At the second stage, by making the best use of experiences and results thus obtained, the communication satellite and the broadcasting satellite having similar capacity will be put to practical use. At the third stage, the program of communication satellite having large capacity and direct broadcasting satellite will start.

In connection with the communication, researches are under way concerning the maritime communications system to be used for small ships in Japanese waters. In the same way, researches will be carried on concerning the high quality TV and other new broadcasting systems.

It has been decided that the CSE and BSE now under development will be launched by the US 2914 rocket, and the control operations necessary for the stationariness of both satellites will be performed by the National Space Development Agency of Japan and afterward experiments will be conducted under the presidency of the Ministry of Posts and Telecommunications and with collaboration of Nippon Telegraph and Telephone Public Corporation (NTT) and the Japan Broadcasting Corporation (NHK).

With this object in view, preparations for the establishment of a master station, a control station, transportable stations and experimental TV receiving stations are being made in parallel with detailed examination of experimental contents.

Further the outline of both satellites will be given in the following table.

TABLE 1 Outline of CSE and BSE

	CSE	BSE
Launch Vehicle	Delta 2914 (USA)	Delta 2914 (USA)
Position of Space Station	135°E	110°E
Attitude Control System	Spin	Three Axis(Zero Momentum)
Orbital Mass	337 Kg	350 Kg
Usable D.C. Power	500 Watts	1000 Watts
Main Communication System	PCM/PSK (100 MB/s)	FM (25 MHz TV)
Transponder	C-Band 2-pairs, 6W K-Band 6-pairs, 4W	14/12 GHz Direct Conversion, 2-pair, 100W each
Communication Antenna	Shaped Beam Horn-Reflector C-Band 25 dB, K-Band 33 dB	Triple Front Fed Parabola Main Islands 37 dB Whole Territory 28 dB

Further, I would like to add some informations or explanations on this table:

- 1st: It is natural that CSE should be located at the south equator of the center of Japan islands. In contrast with this, the position of BSE is 25 degrees west compared with that of CSE. This position of BSE is decided both by the calculation of the time and duration of eclipse as a function of satellite position and by the reason that the broadcasting is not absolutely necessary in the midnight time.
- 2nd: The attitude control system adopted to each satellites are not decided only by their logical study. Examples of practical use of spin stabilized satellites and the conception that we must have experiences on the three-axis-stabilized satellite techniques are affected to our decision.
- 3rd: The main reason to apply PCM/FSK system as main communication system for CSE is its high capability of random access. As for the transponders of CSE and BSE, they have no individuality except their frequency compared with the conventional on-bord transponders. The details of the communication antennas performances are described in Material B-2.

Thirteen materials concerning our domestic satellite plan are introduced.

For our convenience, I divided them into three series A, B, C,: Series-A is concerned with the background of our satellite plan, and B is concerned directly with our satellite already planned including main items of experiments using CSE and BSE and their expected results. Further, series-C is concerned with some general activities on the developments of satellite communications and broadcasting in Japan.

Material: A-1, JAPAN, ITS STATUS RELEVANT TO TELECOMMUNICATION COMMUNITY

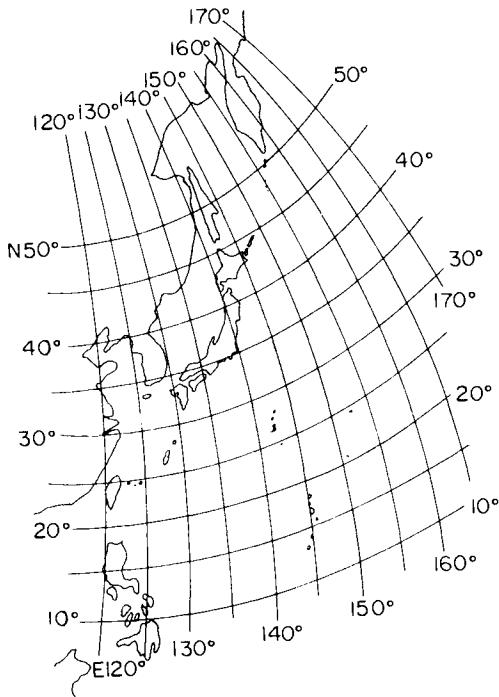
As shown in the Fig. 1, Japan is situated at the North-West corner of the Pacific Ocean, and consist of four larger islands closing to one another together with two trains of small islands.

The latitude of the center of Japan is about 35° North and North-South and East-West length are a little longer than 2000 Km. And the total land area is about 3.7 million square Km with population of one hundred and twenty million.

About 30% of people live in the North-East sea-side cities including Tokyo and Osaka, but other are scattered almost uniformly in our main islands.

In addition, the facts that our main islands have neither high mountain nor large open fields and that Japan is damaged so often by typhoon make Japan's communication plan so difficult.

Fig. 1. Geostationary satellite (110°E) eye view of Japan and its surroundings.



Material: A-2, PRESENT SITUATION OF RADIO WAVE BROADCASTING IN JAPAN

The under-mentioned tables show the number of broadcasting stations in Japan.

The number of TV broadcasting stations using VHF and UHF bands has increased at a high rate, 6000 stations being in operation at the end of March 1975.

However, 72% of these stations were opened under the definition of the "extremely small power station", and they are stations of transponder only which re-broadcast TV signals by receiving broadcasting waves of the regular "plan station". In most cases 3 to 4 times relay signals and even 5 times relay signals are used as their broadcasting source. Therefore, the quality of service of these stations are not always satisfactory.

In spite of the fact that all possible efforts have been hitherto made, 3% of the populations of Japan are not in a position to enjoy the benefits of TV service in their home life.

Besides, the degradation of TV signal quality has been caused by the disturbances of massive buildings increasing in city area.

The above two conditions are giving rise to public discussion.

TABLE 2 Total number of TV stations as of August 1975

	Plan station	Small power station	Total
UHF	851	3636	4487
VHF	838	676	1514
Total	1689	4312	6001

TABLE 3 Increase in number of TV stations in Japan

Finan. Year	1971	1973	1975
No. of Stations	3554	4759	6001
Increase	0	34%	26%

Material: A-3, OUTLINE OF RESEARCHES IN INFORMATION TRANSMISSION SYSTEM OF LARGE CAPACITY COMMUNICATION BY OTHER SYSTEMS THAN SATELLITE RELAY SYSTEM

In Japan, researches and developments are being conducted by NTT with the intention of coping with the increase in demands of public communications. Needless to say, certain manufacturers are coping with NTT for such researches and developments. They are also contributing to meet the communication needs other than the following categories, namely, the development of new communication systems planned by private enterprises:

1. Large capacity FDM system using coaxial cable:

Efforts are being made to put the 36 Ch. telephone T.V. of 1 MHz bandwidth to practical use by means of a circuit of 10,000 telephone Ch., and besides, researches are under way concerning the coaxial FDM system having tens of thousands of Channels.

2. Ultra high speed PCM system using coaxial cable:

The practical use of the PCM-24 and PCM-100M system has been realized, and developments are being made to put the PCM-400M (telephone 5700 Ch.) to practical use.

3. Micro-wave network:

The combination of the radio relay network below 10 GHz and the coaxial system is the main circuit of domestic communications. The communication system used for the said combination is being extended to higher frequencies.

4. Quasi m/m PCM system:

A plan has been framed to the effect that in accordance with the concept of domestic communication network using the PCM system of 20 GHz band, the telephone 5,700 Ch. per carrier system (20G-400M) has been completed, and the practical use experiment of the said system is to start, and finally the domestic circuit network employing this system is to be perfected in the very near future.

5. m/m waveguide system:

For the purpose of realizing the telephone 300,000 Ch. or 208 bi-directional color-TV by the use of 43-87 GHz and PCM, the W-40G experimental system of 50 Km has succeeded and it is now under evaluation test.

6. Combined system of marine cable and micro-wave:

Large capacity cables are laid in the deep seas of the surrounding waters of Japan, and a system connecting large-sized buoys streamed at different points with the land by micro-wave circuits aims at tens of thousands of telephone channels. As the first step of this project, experiments already started concerning the 36 MHz coaxial cable CS-36M (telephone 2,700 Ch.), the micro-wave circuit (4/6 GHz, 25,000 Ch.) as well as the buoys in which repeaters are to be installed are now under experiment.

Material: B-1, FREQUENCY ARRANGEMENT OF JAPAN'S EXPERIMENTAL
COMMUNICATION SATELLITES

We have four experimental satellite plans concerning the communication including broadcasting.

Fig. 2 shows the frequency arrangement to be used for each satellite, and Fig. 3 shows the channel assignment of CSE and BSE.

Fig. 2. Frequency arrangement of experimental communication satellites

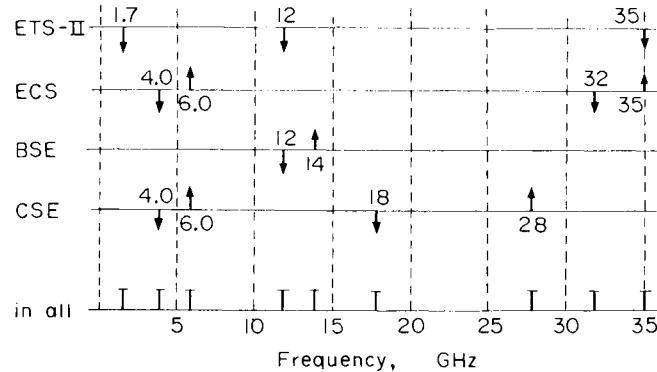
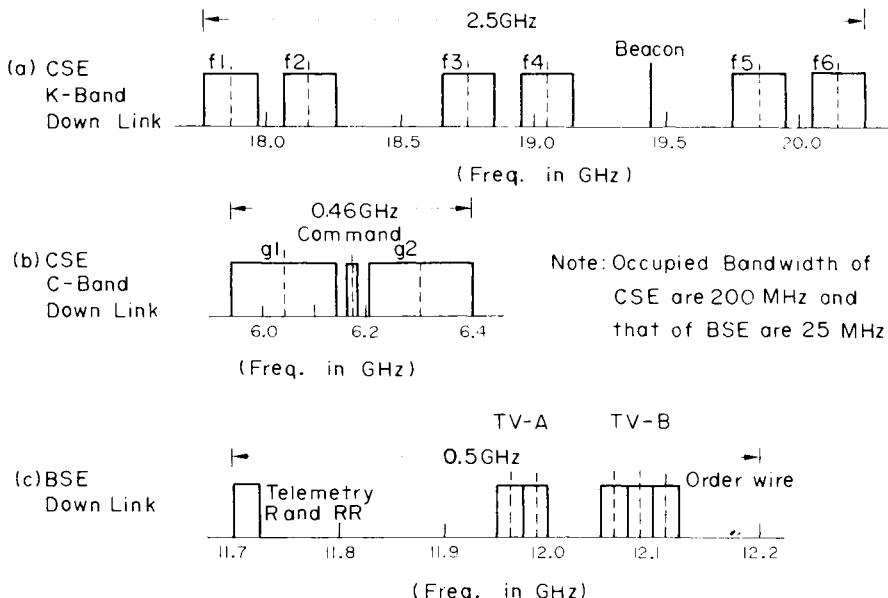


Fig. 3. Channel assignment of CSE and BSE



Material: B-2, SPACE BORNE ANTENNA OF CSE AND BSE

In general, the footprint of the space borne antenna is to approximate to their service area. By this condition, not only the power efficiency of the transmitter is improved but also the interferences caused by spillover will decrease.

In case of our BSE, three-axis-stabilization is used. Antenna system having proper designed parabolic reflector of elliptical shaped and the combination of the three kinds of primary horn antenna offer us a fairly favourable footprints as shown in Fig. 4. On the earlier phase of the antenna design, properly deformed parabolic mirror type antennas are also discussed, but the estimated difficulties on production made us abandon this attempt.

CSE is a spin stabilized satellite, and 28/18 GHz chain is planned to use for main islands and 6/4 GHz chain for isolated islands. Specially designed horn and reflector type antenna is adopted for these purposes to produce proper patterns. Considerations both for the antenna system and feed systems are also paid concerning the common use for C-band and K-band. Figs. 5 and 6 are the converted footprints from the data of the experimental antenna.

Further studies on the environment tests and the effects of pointing errors of these antennas are continued.

Fig. 4. Footpring of BSE antenna.

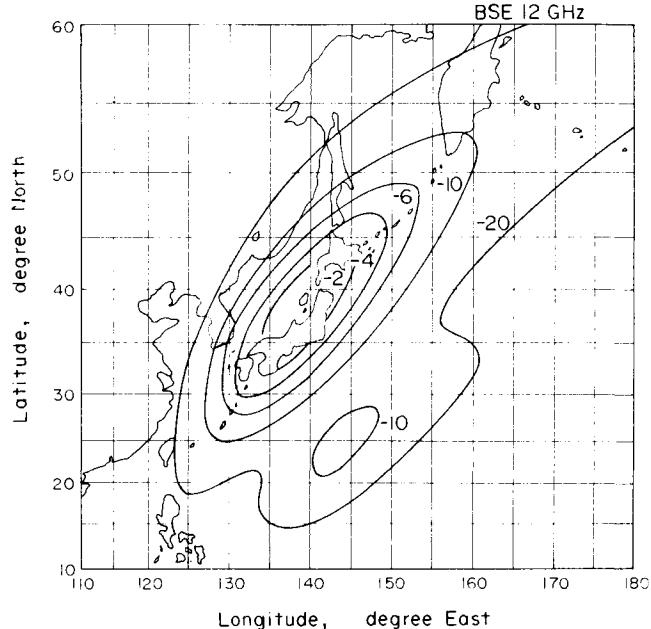


Fig. 5. 18 GHz band of footprint of the communication antenna of CSE.

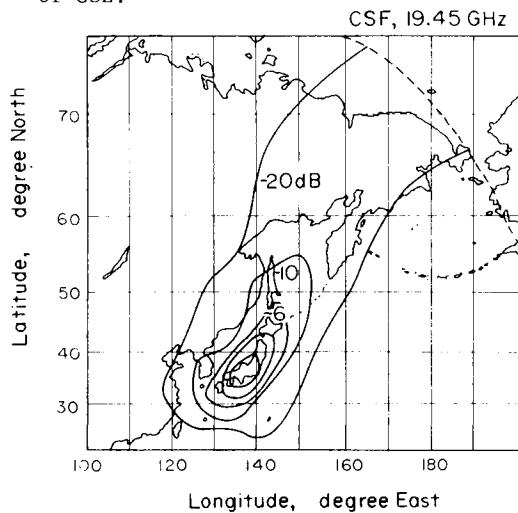
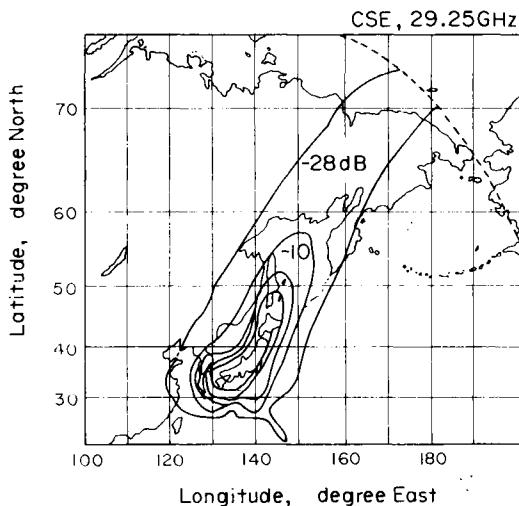


Fig. 6. 28 GHz band footprint of the communication antenna of CSE.



Material: B-3, ITEMS OF EXPERIMENTS USING BSE AND THEIR EXPECTED RESULTS

(A) Main items of experiments to be conducted:

- 1) Pointing power measurements are to be made in Japan proper and its surroundings for a long period of time.
- 2) Observations of polarization change caused by rain drops.
- 3) Measurements of performance characteristics of transponders and their support systems on board are to be made for a long period of time.
- 4) Evaluation tests of TV signal quality are to be made by changing parameters of each part of the total systems under various conditions of propagation path.
- 5) Experiments are to be conducted concerning the compatibility of satellite broadcasting with terrestrial TV broadcasting both being in the same frequency band.
- 6) Experiments are to be made on the switching techniques of plural transmitters at distant places.
- 7) Execution of the attitude and position control of three-axis-stabilized satellite.

(B) Expected results:

- 1) System optimization based on the existing techniques. Condition is to be given as to how to decide the output power of transmitter, the pattern and gain of the space-born antenna, the modulation specifications, etc., according to the number of channels in the assigned band.
- 2) Decision of the policy for further developments of receiving systems developed in Japan.
- 3) Establishment of switching techniques of plural ground transmitters of distant places.
- 4) Clarification of technical prospects in case 12 GHz band is used in

common both for the satellite and terrestrial broadcastings.

- 5) Point out problems to be studied concerning the frequency reuse by cross polarization techniques for 12 GHz band broadcasting.
- 6) Establishment of the operation and maintenance techniques of the total system to carry on the desired service without hitch.

Material: B-4, ITEMS OF EXPERIMENTS USING CSE

- 1) Measurement of 18 and 28 GHz radio wave propagation characteristics:
Transmission loss and polarization change caused by raincore are to be observed by properly distributed earth stations.

Simultaneous observations should be conducted among stations of 10 to 25 Km apart.

Further, unwanted radio propagation in these frequency band should be observed together with the radar observation of the raincore.

- 2) Measurement of performance characteristics of the transponder and their support systems on board for a long period of time:

Data obtained on 2 transponders of C-band and 7 transponders of K-band must be effective to make our space born electronics design more rational.

- 3) Failure mode analyses of the space born equipments.

- 4) Measurement of space born antenna performance:

Effect of solar radiation on the radiation characteristics of our shaped beam antenna.

Transient responses of our despin antenna system which may be caused by position change and attitude control are also studied.

- 5) Measurement of information transmission characteristics:

Evaluation of signals quality relayed by satellite as a function of types of modulation, modulation specification, access system and number of channels.

- 6) Experiments on some location diversity systems.

- 7) Experiment on connecting satellite circuit to terrestrial relay networks.

- 8) Experiment to establish tentative circuit using satellite relay system.

- 9) Experiment on some counter plan against delay and singing.

- 10) Compatibility test between the satellite communication circuit and terrestrial relay network.

- 11) Experiment and experience concerning the operation and maintenance techniques of the total system.

Material: B-5, RESULTS OF CALCULATION OF QUALITY OF RECEPTION TO BE EXPECTED BY RECEIVING BSE SIGNALS

The output of the two transmitters are fed to their antenna systems to form different radiation characteristics, therefore, two kinds of radiation beam are to be used for our experiments. One beam is designed for main islands having half power beam width of 1.4 degree and the other covers the whole islands having beam width of 4.0 degree.

As shown in Table 4, the quality of reception within these service areas are expected to satisfy TASO grade 1 by 99% of time, and S/N of 39 dB is to be guaranteed by 99.99% of time by using the 1.4 meter antenna in Japan proper and 4.5 meter antenna at isolated islands, respectively.

TABLE 4 Expected quality of reception on BSE experiments

Category of receiving system		Radiation beam B, for main islands, 1.4° beam width	Radiation beam A, for whole islands, 4.0° beam width
Antenna diameter	Receiver location		
4.5 m	Inside of 1.4° beam	CCIR standard (99.9%) S/N = 54 dB (99%) = 52 dB (99.9%) = 48 dB (99.99%)*	Over TASO grade 1 (99.9%) S/N = 45 dB (99%) = 43 dB (99.9%) = 39 dB (99.99%)*
4.5 m	1.4° beam edge	Over TASO grade 1 (99.9%) S/N = 45 dB (99%) = 43 dB (99.9%) = 39 dB (99.99%)*	Over TASO grade 1 (99.9%) S/N = 45 dB (99%) = 43 dB (99.9%) = 39 dB (99.99%)*
4.5 m	4.0° beam edge		Over TASO grade 1 (99.9%) S/N = 45 dB (99%) = 43 dB (99.9%) = 39 dB (99.99%)*
1.6 m	Inside of 1.4° beam	Over TASO grade 1 (99.9%) S/N = 45 dB (99%) = 43 dB (99.9%) = 39 dB (99.99%)*	

Note: assumed noise temperature of receiver is 500°K, and * denotes threshold condition.

Material: B-6, RESULTS OF CALCULATION TO BE EXPECTED ON CSE EXPERIMENTS

The preliminary design of the total system decides that 6/4 GHz link is to be used for isolated islands and 28/18 GHz link is for main islands.

Table 5 shows the performance specification of transponders of each frequency, and the design values of faculty of equipments to be installed in each earth stations of three categories are shown in table 6.

Furthermore, some examples of the expected performances of relay networks are shown in Table 7 together with their chain parameters.

TABLE 5 Performance specification of CSE transponders

	C-band	K-band
Receiving frequency band in GHz	6	28
Receiving antenna gain in dB	34.0	37.0
Noise figure of receiver in dB	9.0	13.0
Transmitting frequency band in GHz	4	18
Output power of transmitter in dBW	4.5	4.0
Transmitter antenna gain in dB	31.0	39.0

TABLE 6 Design value of faculties of earth station

Class of station	C-band			K-band		
	FIC	TRC	STC	FIC	TRC	STC
Transmitting frequency band in GHz	6.	6.	6.	28.	28.	28.
Transmitter output power in dBW	23.0	23.0	23.0	27.0	27.0	27.0
Transmission antenna gain in dB	56.0	54.0	43.0	69.0	67.0	56.0
Receiving frequency band in GHz	4.	4.	4.	18.	18.	18.
Noise temperature of receiver in dBK	20.0	20.0	20.0	26.0	26.0	26.0
Receiving antenna gain in dB	52.0	50.0	39.0	66.0	64.0	53.0
Diameter of parabolic reflector in meter	10	10	2.8	13	10	2.8

Note 1: FIC, TRC, STC are the abbreviation of Fixed, Transportable and Small Transportable station, respectively.

Note 2: Further small scale stations are also under planning.

TABLE 7 Example of expected performance of relay circuit

- a) From the fixed station to a transportable station situated on the beam edge (25 dB or contour for C-band and 33 dB contour for K-band)

over-all down link up-link	C-band	K-band
	Frequency in GHz	6.31
	EIRP in dBW	77.0
	Free space loss in dB	199.8
	C/T in dB	-126.4
	Frequency in GHz	3.82
	EIRP in dBW	29.5
	Free space loss in dB	196.0
	C/T in dB	-137.5
	System C/T in dB	-137.8
over-all down link	C/N _o in dB/Hz	90.8
	Necessary C/N _o for 50 MBPS in dB	89.0
	Margin in dB	2.0
		1.2

- b) Between small transportable stations situated on beam edge (25 dB contour for C-band and 33 dB contour for K-band)

over-all down link up-link	C-band	K-band
	Frequency in GHz	6.31
	EIRP in dBW	64.0
	Free space loss in dB	199.8
	C/T in dB	-144.4
	Frequency in GHz	3.82
	EIRP in dBW	29.5
	Free space loss in dB	196.0
	C/T in dB	-148.5
	System C/T in dB	-149.9
over-all down link	C/N _o in dB/Hz	78.7
	Necessary C/N _o for 6 MBPS in dB	80.0
	Margin in dB	-1.3
		-1.0

Material: C-1, A NEW DEVELOPED 12 GHZ TV RECEIVING SYSTEM

1. Outline:

Fig. 7 shows the block diagrams that convert 12 GHz FM input to VHF-AM signal. In this figure, (a) denotes the conventional system already developed, and (b) denotes the so-called "Direct Conversion System" developed by NHK. It is noted that (b) is far simpler in constitution than (a).

The direct conversion system of this type consists of a specially designed SHF-UHF frequency converter suitable for mass production which is followed by a simplified FM/AM converter.

NHK has already developed this converter and passed through some experience in mass production.

According to these facts, this receiving system is expected to be used for BSE community reception test except for some special cases.

2. First converter:

The pattern and its equivalent circuit are shown in Fig. 8. It is a simplified low-noise 12 GHz band converter with a planer circuit mounted in waveguide, and for these reasons, this converter is suitable for mass production.

Production tests show that the conversion loss and noise temperature are 3.2 dB and 500°K, respectively, by the use of a diode whose cut-off frequency is about 180 GHz. Further, the theoretical reduction indicates that the conversion loss and noise temperature can be improved up to 2.0 dB and 300°K, respectively, by using 600 GHz diode.

3. FM/AM converter of direct conversion type:

Fig. 9 is a block diagram showing a circuit that converts directly UHF-FM signal to VHF-AM output with neither demodulation nor modulation.

By employing this circuits, the number of parts used is so small as compared with the conventional converter. Consequently, the size and cost of the converter becomes reasonable.

It is to be added that the principle of this circuit can be explained as follows:

The impedance change caused by the frequency change of the incoming signal produces amplitude variation of the reflected wave of the VHF oscillator installed.

The bridge circuit in Fig. 9 is used instead of a circulator in view of cost reduction.

4. Example of complete system:

Fig. 10 shows a block diagram of the total receiving system designed for the reception not only of FM and AM signals of 12 GHz but also of UHF and VHF-TV broadcasting signals.

Fig. 7. Block diagram of the SHF-FM to VHF-AM converter

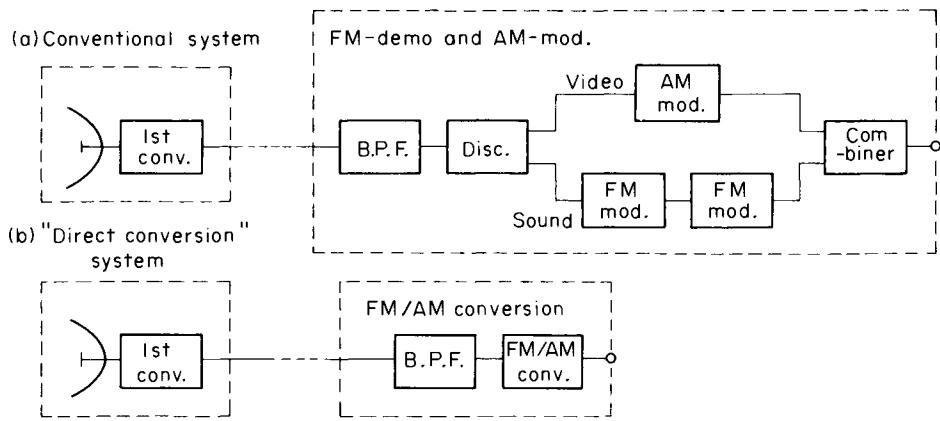


Fig. 8. Pattern and its equivalent circuit of the first converter

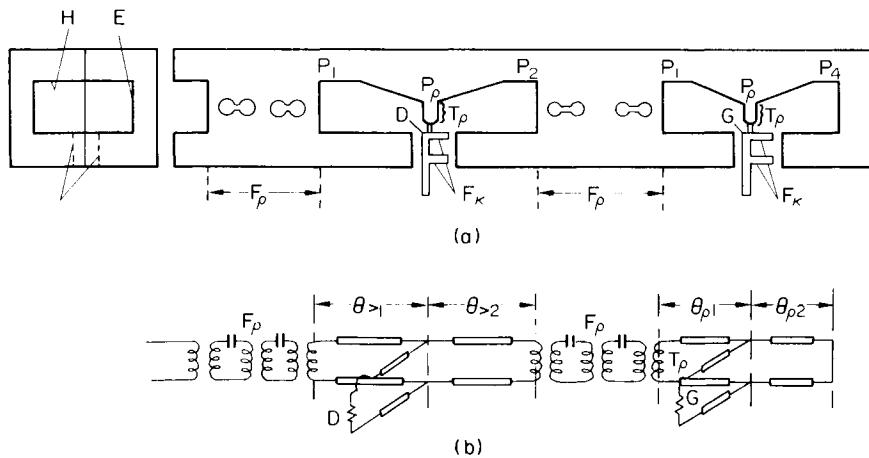


Fig. 9. Block diagram of the FM/AM converter of a direct conversion type

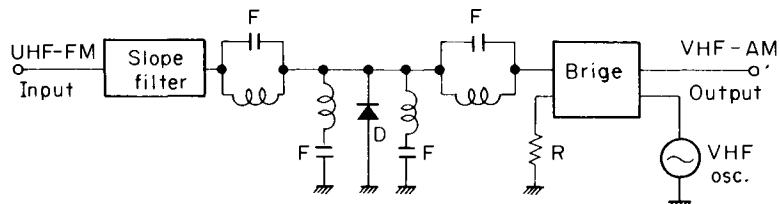
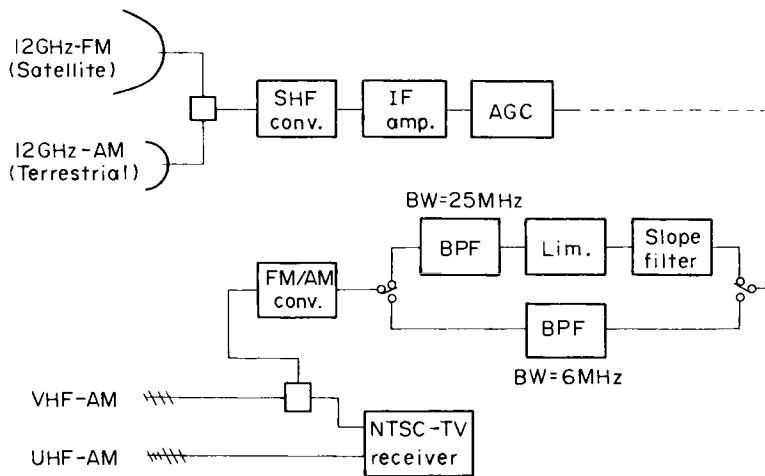


Fig. 10. Block diagram showing an example of multifrequency-band receiver



Material: C-2, RESEARCHES IN RADIO WAVE PROPAGATION FOR SATELLITE COMMUNICATIONS IN JAPAN

At the early stages of researches and developments in space communication in Japan, researches on radio wave propagation started. The main object of such researches is to study the effect of precipitations on radio wave propagation whose frequencies are higher than 10 GHz. Such researches are being based on the prospect that in view of the present situation of Japan's domestic radio communication, radio waves of fairly high frequencies will be used for the satellite communications.

The above researches are summarized as follows:

- 1) Studies of radio wave attenuation by observing solar radio noise:
Observations of the received noise intensity up to 48 GHz concerning radio communication are being conducted by five organizations.
Joint studies and analyses of 12 GHz radio wave attenuation caused mainly by rainfall have been completed by Malaysia and Japan.
- 2) Presumption of radio wave attenuation by measuring "sky temperature.":
This method forms a contrast to the abovementioned method as to the advantages and disadvantages. Since the equipment of observation is simple, many observations can be performed. However, there is some limit in calculating the attenuation by this method. Therefore, researches are under way concerning a more effective method of conversion in the light of the observed values obtained by the abovementioned method.
- 3) Study plan of radio wave propagation by launching artificial satellites:
Since we have no satellite which transmits radio waves with frequencies higher than 10 GHz, experiments on radio wave propagation will be conducted on the occasion of experiments on communication and broadcasting by the use of satellites of several kinds, and preparations are being made for this purpose. In addition, this study plan covers observations both on the attenuation and the depolarization caused by rain drops.

4) Relation between meteorological data and radio wave attenuation:
If the relation above captioned is realized, the design of the total systems of satellite communication will become more rational using the observed data obtained by a small number of earth stations.

For this reason, observations on rainfall are being made by one to ten rain gauges at the option of each station. Besides, specially designed radars suitable for this object are installed in some stations.

The concept of effective path-length of rain core has been introduced and the evaluation of this concept is now in progress. The object of this concept is to generalize the observed values of radio wave attenuation and to establish the relations between the radio wave attenuation and the meteorological observed data accumulated.

5) Researches in radio wave propagation concerning "location diversity effects":

In case the frequencies susceptible to rain attenuation are put to practical use, location diversity effects should be evaluated. For this reason, the relationship between the observed values obtained by plural stations arranged at reasonable intervals to be analysed.

6) Theoretical studies of the attenuation of radio waves caused by rain drops:

This study has been put forward by Dr. T. Oguchi together with that of the scattering characteristics. He adopted model rain drops which have the similar shape, dimensions and density to real raindrops and transformed his calculation into the form applicable to practical use.

His calculation of attenuation and scattering characteristics will play an important part in the design of satellite communication.

7) Others:

As an example, investigations, both theoretical and experimental, being carried on concerning the relations between the above sea-level of earth station and the radio wave attenuation caused by rainfall.

Material: C-3, THE RESEARCH ON NEW ORBITS FOR OPERATIONAL SATELLITES

The recurrent perigee, near-recurrent perigee and sub-synchronous orbits are under investigation to apply to operational satellites.

Considering the concepts of the effective use of the geostationary orbit, studies have been conducted by our group concerning the evaluation of other orbits for communication and broadcasting.

The term recurrent perigee or near-recurrent perigee orbit is used to refer to the recurrent or near-recurrent orbit whose longitudes of the perigee and apogee remain almost the same for a certain period of time. The recurrent orbit is the orbit whose ground track is almost the same every day or in every orbital period, and the near-recurrent orbit is the orbit whose ground track becomes almost the same in every recurrent period which is larger than one day and one revolutional period.

The term sub-synchronous orbit is used to refer to the orbit for which the mean solar day is an integral multiple of the satellite's revolutional period relative to the longitudinal plane of the earth.

For example, highly elliptical recurrent and near-recurrent perigee orbits will be applicable to communication satellites in high or medium latitudes where stationary satellites are difficult to use for their lower elevation angles, or be applicable to meteorological satellites to observe high latitudes continuously every day over longer time than the observation time in case of sun-synchronous orbits.

Circular sub-synchronous orbits with zero inclination will be applicable to broadcasting satellites or meteorological satellites which broadcast or observe globally except high latitudes at certain time every day.

Table 8 shows advantages or disadvantages of recurrent perigee, near-recurrent perigee, sub-synchronous, constant argument of perigee, sun-synchronous and stationary orbits.

In the table, the recurrent perigee, near-recurrent perigee and constant argument of perigee orbits are the case of highly elliptical orbits and of being used in high or medium latitudes. The orbit of Molniya Satellite is an example of the highly elliptical constant argument of perigee orbit with about a half day period.

References

- (1) Takahashi, K. and Kawakami, K., "Determination of Elements of Recurrent and Near-Recurrent Orbits and Their Application to Operational Satellites," J. of Radio Research Laboratories, 21, No. 107, 1974.
- (2) Takahashi, K., "Determination of Elements of Sub-Synchronous or Super-Synchronous Orbits and Their Application to Operational Satellites," J. of Radio Research Laboratories, 22, No. 108, 1975.

TABLE 8 Comparison of recurrent perigee, near-recurrent perigee, sub-synchronous,
constant argument of perigee, sun-synchronous and stationary orbits

Orbits	Orbital Period	Orbital Inclination	Launch into Orbit	Maximum Elevation Angle	Visible Time	Recurrence	Periodicity	Effect of Eclipse
Recurrent, Near-Recurrent Perigee	Arbitrary	Restricted to $50^\circ - 70^\circ$	Easy	Large	Medium	Yes	No	Small
Sub-Synchronous	Discrete 24 h/N N: integer	Arbitrary	Easy	Indefinite	Medium	No	Yes	Large
Constant Argument of Perigee	Arbitrary	Nearly Constant About 63°	Easy	Large	Medium	No	No	Small
Sun-Synchronous	Arbitrary	Restricted to $> 90^\circ$	Easy	Large	Short	No	No	No
Stationary	Constant	Constant 0°	Difficult	Small	Long	Yes	Yes	Large

Material: C-4, SSRA COMMUNICATION SYSTEM

SSRA communication system has fascinating features for some communication networks. The advantages of this system are as follows:

No timing requirement, no strict frequency allocation, simultaneous range measurement, nearly perfect free random access, less interference and message security.

Radio Research Laboratories conducted theoretical and experimental study on this communication system by using ATS-1 through the courtesy and nice suggestion of NASA.

Principle of this communication system is shown below:

SSRA stands for Spread Spectrum Random Access. As this name implies, this system uses spectrum spreading technique. A carrier wave is frequency-modulated by a baseband signal (primary modulation) resulting in a signal of narrow-band spectrum and, then it is bi-phase modulated by a PN code with extremely high clock frequency (secondary modulation) resulting in a wide-band signal wider than the former by about 10^2 to 10^6 times. The PN code is a noise-like sequence having a long frame period and having a sharp correlation characteristic. These SSRA signals are transmitted to a satellite from different stations. Widely spread spectra overlap each other at the input of the satellite repeater and retransmitted to ground stations. At the ground receiver, all the SSRA signals are received, but the only desired signal is reproduced in the following manner.

In the synchronization circuit, the PN code for the desired SSRA signal is locally generated, whose phase is circulated and compared bit by bit with the incoming PN code of the received SSRA signal. When the perfect coincidence happens, synchronization completes. This synchronized PN code is then fed to the spread demodulator and inversely modulates the desired SSRA signal of the wide-band spectrum into the narrow-band FM signal. This signal is finally frequency-demodulated and information is reproduced. Undesired SSRA signals, however, remains as signals of wide-band spectrum and work as noise to the desired signal.

Note: Further informations printed are available by request.

CONCLUDING REMARKS

Let me conclude my speech with the remarks that my personal opinion on the developments in Japan of satellites for practical use is as follows:

- 1) I shall be very happy if you could understand, through my presentation, with what intention Japan is planning to develop the communication satellites and broadcasting satellites, and in what direction Japan is hoping to advance such satellites.
- 2) Generally speaking, in designing the satellite communication circuits, various researches and discussions have been conducted as to the selection of frequency band to be used. I am much interested in this problem.

3) In this connection I am of the following opinion:

4) When we plan the satellite communication circuits to be used in Japan, a comparatively low frequency is sometimes considered to be most suitable according to the purpose. In Japan, however, the density of terrestrial radio communication using a frequency band lower than 10 GHz is so high that there is only small room left for the use of these frequency band for satellite communication. Therefore, great efforts should be made toward technical development so that this frequency band may be used also for satellite communications.

5) Consequently, frequency band from 10 GHz to perhaps 60 GHz should be mainly used for satellite communication, but it is necessary to decide this frequency band after due consideration of its common use for satellite and terrestrial communications.

6) In case such a high frequency is used, the effect of shaped beam or spot beam can easily be expected, and thereby the reduction of spillover as well as the effective use of transmitting power can be also expected. In addition, the common use of frequency for satellite and terrestrial communications becomes easier, and as a result the effective use of radio frequency spectrum can be realized.

7) In order to use a high frequency, strenuous efforts should be made towards the improvements in electronic technology and control technology of satellite as well as in the study of radio wave propagation characteristics.

8) For some reasons mentioned above, I think that for the time being, efforts will be concentrated upon experiments in radio wave from 10 GHz to 60 GHz when researches and developments are made in Japan on satellite communication.

9) Experiments based on such intention are considered to be very useful for Japan's domestic communications, and on the other hand, I think that technical developments should be made aiming at being useful also for the plans formed by some countries or regions, the communication situation of which is different from that of Japan.

10) Needless to say, judging from the nature of satellite communication, various kinds of plans should be carried out with international cooperation. In this respect, Japan has already established a policy of positive participation in such international cooperation.

11) As you are probably aware, with regard to Japan's satellite plan, arrangements have been made between Japan and the U.S.A. for support, and between Japan and each of Canada and ESRO for mutual cooperation, and besides, useful technical information is being provided to Japan by other countries.

12) I myself, as a member of participants in Japan's satellite plan, I do hope that this plan is useful for our communication and culture, and that the same time, technical information obtained through the performance of the said plan can be profitable not only for the plans of different countries but also for international cooperation.

13) Finally, I would like to tell you that for any technical information of which you think Japan is in possession, please apply to us without hesitation.

THE MARISAT SYSTEM

Charles Dorian

COMSAT General Corp., Washington, D.C. 20024, U.S.A.

INTRODUCTION

Maritime mobile communications on the high seas in the HF radio frequency band, 3-30 MHz, suffers from the inherent limitations of signal propagation. Ionospheric and magnetic storms cause serious disruption of HF transmissions. Continuous, real-time, highly reliable communications to ships on the high seas is not available today. These will become a thing of the past with the introduction of a new maritime mobile-satellite service via MARISAT. MARISAT will be the world's first commercial maritime satellite system. It is scheduled to be placed in operation in early 1976. Two satellites will be launched, one to cover the Pacific Ocean and the other to cover the Atlantic Ocean.

BACKGROUND

In 1972 the U. S. Navy stated a need for satellites to provide communications in both the Atlantic and Pacific Oceans for a two to three year period. This need resulted from the expected failure of their two operating satellites and "gap" in time before replacements would be available. The "gap" was to be of limited duration. In order to find a reasonable and economical solution to this problem another user was needed. Examination revealed that the international maritime shipping industry had the greatest need and also potentially offered a good source of revenue.

COMSAT General conducted studies and decided to develop an integrated satellite system which could meet the U. S. Navy's needs on their assigned frequencies as well as those of commercial shipping on newly assigned frequencies in the UHF band (1535-1660 MHz). The system could meet the Navy's limited needs and permit a gradual transfer of satellite capacity to the commercial maritime service as the needs of the Navy decreased.

The U. S. Navy accepted COMSAT General's proposal. System designs were finalized, and contracts were let. Hughes Aircraft Company was selected to construct the satellite and Aeronutronics-Ford the earth station.

SATELLITE

Three satellites are being procured from Hughes Aircraft Company. Characteristics of the spin stabilized cylindrical consist of five repeaters: a 480 kHz wideband UHF repeater, two narrowband 24 kHz UHF repeaters, a L-band to C-band repeater and a C-band to L-band repeater. The UHF repeaters which operate in the 250 to 400 MHz band are used for Navy services. The other two repeaters are used for commerical services with transmissions between the shore station and the satellite in the 4 and 6 GHz bands and transmissions between the ships and the satellites in the 1.5 and 1.6 GHz bands. Repeater operating states and commercial communications capacity are given in Tables 2 and 3.

The two satellites will be placed in geostationary orbits at 150° west longitude and 176.5° east longitude where they will provide communications to about two-thirds of the world's ocean area from 70°N to 70°S latitude. (See Fig. 1)

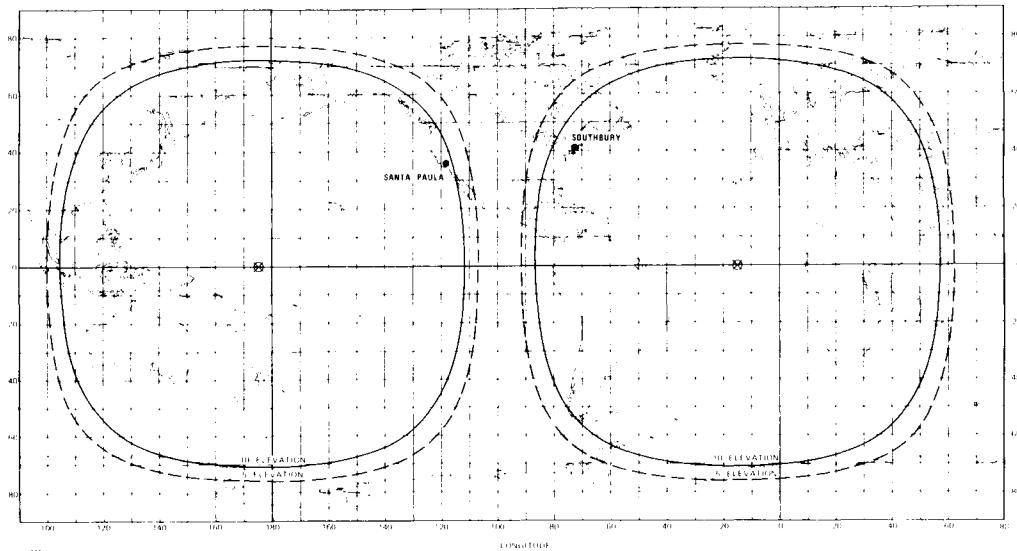


Fig. 1 Coverage area

TABLE 1 Marisat Spacecraft System Characteristics

Size	
Solar Panel Dia.	216 cm (85 in.)
Solar Drum Height	160 cm (63 in.)
Overall Height	382 cm (150 in.)
Weight	
Launch Weight	655 kg (1,445 lbs.)
Apogee Motor Expendables	264 kg (582 lbs.)
Hydrazine	74 kg (163 lbs.)
Dry Weight	317 kg (700 lbs.)
Power Subsystem	
Solar Array	> 7,000 N-ON-P cells
D.C. Power	> 300 watts
Eclipse Capability	Full
UHF Communications Subsystem	
Receive Band	300-312 MHz
Transmit Band	248-260 MHz
EIRP	
1 Wideband Channel	28 dBW
2 Narrowband Channels	23 dBW (each)
Antenna	Three element Helix Array
Polarization	RHC (transmit and receive)
Half Power Beamwidth	> 19° (earth coverage)
Gain	12.6 dB
L-to-C Band Channel	
Receive Band	1638.5 - 1642.5 MHz
Transmit Band	4195 - 4199 MHz
Receive G/T	-17.0 dB/K
C-Band EIRP	18.8 dBW (saturated output)
L-Band Antenna	Four element Helix Array
Receive Gain	14.4 dB
Polarization	RHC
C-Band Antenna	Circular horn
Transmit Gain	16 dB
Polarization	LHC
C-to-L Band Channel	
Receive Band	6420 - 6424 MHz
Transmit Band	1537 - 1541 MHz
Receive G/T	-25.4 dB/K
C-Band Antenna	Circular horn
Transmit Gain	16 dB
Polarization	RHC
L-Band Antenna	Same as L-to-C

TABLE 2 Repeater Operating States

	BEAMEDGE E.I.R.P. Satellite/Ship (L-Band)	FOR COMMERCIAL Satellite/Shore (C-Band)
All UHF Channels On	20 dBW	18.8 dBW
Wideband UHF Only	26 dBW	18.8 dBW
No UHF On	29.5 dBW	18.8 dBW

TABLE 3 Marisat Commercial Communication Capacity*Shore-to-Ship

	<u>VOICE</u>	<u>TELEPRINTER CH.</u>
Full Navy Service L-Band E.I.R.P. = 20 dBW	1	44
Wideband Navy Service L-Band E.I.R.P. = 26 dBW	5	66
No Navy Service L-Band E.I.R.P. = 29.5 dBW	9	110

Ship-to-Shore

Regardless of Navy Service	9	110
----------------------------	---	-----

* These are typical quantities. Voice and teleprinter capacity can be traded as needed within the constraints of available satellite power and frequency plans.

EARTH STATION

Initially, only the Southbury, Connecticut and Santa Paula, California earth stations will access the MARISAT satellite, although system design permits other earth stations to work through the satellites with direct access to ship terminals should other countries desire to participate.

The RF portion of the earth station equipment consisting of a 12.8 meter (42-foot) diameter full motion antenna, uncooled parametric amplifier receivers, and high power amplifiers has been built under contract by Aeronutronics-Ford. Fabrication and installation of this equipment are complete. The ground communication electronics which will be used for commercial maritime satellite services is being constructed by Scientific-Atlanta, Inc. and its subcontractors, the Digital Communications Company and the International Computing Company. Construction of this equipment was started in September of 1974 and is scheduled for completion in 1975 and 1976. Earth station characteristics are given in Table 4.

In addition to being equipped for 6 GHz transmission to the satellites and 4 GHz reception from the satellites, each earth station has a L-band transmit (1.6 GHz)/receive (1.5 GHz) capability. This is used for test and monitor purposes and also for the transmission and reception of pilot tones looped from the shore station through the satellite and back to the shore station. These tones are used to compensate for frequency shifts introduced by satellite oscillator instabilities and doppler shifts. Since system frequency shifts can be large in comparison to the bandwidth of the signals employed, it is necessary to correct for them.

TABLE 4 Major Earth Station Characteristics

Antenna Diameter	12.8 meters (42 feet)
C-Band Transmit (communications)	
Frequency Band	6174.5 - 6424.0 MHz
Antenna Gain	56.0 dB (on axis)
Nominal EIRP	90.5 dBW (max) (on axis)
Voice Channel	75 dBW
Data Channel	72 dBW (TDM)
AFC Pilot	59 dBW
Polarization	RHC
C-Band Receive (communications)	
Frequency Band	3945.4 - 4199.0 MHz
Receive G/T	31.4 dB/K
Polarization	LHC
L-Band Transmit (test and control)	
Frequency Band	1638.5 - 1641.5 MHz
Antenna Gain	G = 38.0 dB (on axis)
Nominal EIRP	69 dBW max., normally 38 dBW
Voice Channel	37 dBW
Data Channel	37 dBW
AFC Pilot	35 dBW
Polarization	RHC
L-Band Receive (test and control)	
Frequency Band	1537.0 - 1542.0 MHz
Receive G/T	12 dB/K
Polarization	RHC

SHIP TERMINAL

Two hundred ship terminals presently are being manufactured for COMSAT General under a contract with Scientific-Atlanta, Inc. These will be made available by COMSAT General to its maritime customers for use in providing voice, telex, facsimile and data services. It is expected that as usage grows, ship terminals will be manufactured by various companies and either sold or leased to maritime users directly by the manufacturers and the established maritime communication companies.

The COMSAT General terminals consist of two portions; an above deck equipment group and below deck equipment. The above deck portion consists of a 1.2 meters (4-foot) antenna, 40 watt solid state L-band power amplifier, L-band transistor receiver pre-amplifier, low loss radome, and a four axis antenna pointing system which consists of an azimuth-elevation pedestal which is stabilized on a pitch-roll platform. A signal from the ship's gyrocompass is used to correct the azimuth drive and a step-track system, which tracks an L-band signal from the satellite, is used for antenna pointing.

The electronics equipment located below deck consists of an antenna control unit, communications electronics used for transmission, reception, signalling, network control and access, and telephone and teleprinter equipment. COMSAT General ship terminals will be equipped with a telephone and a 50 baud international speed teleprinter, a five level machine with a private wire call control unit, automatic answer-back, paper tape punch and reader, and two color printing (red on receive, black on transmission).

The ship terminals are designed for unattended automatic operation insofar as is possible. The receiver is always on so that incoming calls are always received. Also, control of the terminals (frequency retuning, channel selection, message type) is performed automatically upon command from the shore station. The minimal operating controls which are located on the ship terminal include those for selection of message priority (emergency, normal, etc.), type of message (voice, TTY, data, etc.), shore station destination, and an ocean code which indicates the region in which the ship is operating. Ship terminal characteristics are given in Table 5.

TABLE 5 Ship Terminal Characteristics

Communications

Transmit Band	1636.5 - 1645 MHz
Receive Band	1535.0 - 1543.5 MHz
Transmit EIRP	37 \pm 1 dBW
Receive G/T	-4.0 dB/K

Antenna

Pointing Method	4 Axis Slaved
Diameter	1.2 meters (4 feet)
Gain	23.5 dBi
Beamwidth	10° @ 1.64 GHz
Polarization	RHC

Environmental

Ambient Temperature	-40°C to +65°C Above Deck -15°C to +55°C Below Deck
Relative Humidity	Up to 95%
Icing	Up to 2.54 cm (1 inch)
Precipitation	Up to 10 cm/hr.
Wind	Normal Operation in Winds up to 75 Knots

Power Supply	115V \pm 10%, 60 Hz \pm 4Hz Single Phase or 115V \pm 10%, 50 Hz \pm 3Hz Single Phase or 220V \pm 10%, 60 Hz \pm 4 Hz Single Phase or 220V \pm 10%, 50 Hz \pm 3 Hz Single Phase
--------------	---

Physical Characteristics

DIMENSIONS

Above deck antenna and radome	Base 79" in diameter Height 80"
Below deck console	24" x 24" x 60"

WEIGHT

Above deck antenna	500 lbs (227 kilos)
Below deck antenna	350 lbs (159 kilos)

The types of communications signals and their major parameter values are shown in Table 6.

TABLE 6 Marisat Communications Systems

Signal Types

Outbound -- Shore-to-Ship

- Voice Frequency modulated (FM) single channel per carrier (SCPC) voice grade channels employing compandors and having a subjective signal-to-noise ratio greater than 28 dB.
- Telex Time Division Multiplexed (TDM) carrier containing an assignment channel (used for access and control) and 22 Assign- fifty baud channels.
ment

Inbound -- Ship-to-Shore

- Voice FM/SCPC as above.
- Telex Time Division Multiple Access (TDMA) system for narrow-band data. Frequency for TDMA carrier paired with TDM carrier for duplex operation. Timing information derived from TDM signal which is received at all times.
- Request Burst mode random access transmission at same bit rate as TDMA.

<u>Signal Parameters</u>	<u>Shore-to-Ship</u>		<u>Ship-to-Shore</u>	
	Voice	TDM	Voice	TDMA/Request
Modulation	FM/SCPC	2 Ø CPSK	FM/SCPC	2 Ø CPSK
Transmission Rate (BPS)	--	1200	--	4800
Peak Deviation (kHz)	12	--	12	--
RF Bandwidth (kHz)	27	1.6	27	6.25
Subjective S/N (dB)	28	--	30	--
Bit Error Rate	--	10^{-5}	--	10^{-5}
C/N Unfaded at 5° el.	50.4	47.4	53.8	53.8

SERVICES

All of the services - telex, voice, facsimile, data, distress and broadcast - are available today in the existing MF and HF maritime mobile communication system. But they are not universally available throughout the world, nor are they available 24 hours per day. Additionally, there is not available a standard, installed and operating, worldwide selective calling system. Hence, for all practical purposes the present maritime mobile communication system has not improved very much in the past 50 years. If anything, the radio frequencies are becoming more crowded as new countries become users and increased interference occurs as more messages are transmitted. Some countries have predicted they will reach saturation of their allocated HF radio frequencies within the decade 1970-1980.

The solution to this problem is communications via satellite. Inherent in a satellite system is selective calling and the capability to offer all the communication services desired by maritime users. The following are more details on the individual services:

Telex: The worldwide agreed system for telex is the 50 baud CCITT No. 2 telegraphic code. This code and speed has had the greatest acceptance throughout the world. Standards and interchange facilities already are established and in use for those countries employing higher speeds and different codes. Telex service to ships equipped with appropriate terminals will be automatic, direct dialing, in a duplex mode.

Voice: Once contact has been established between the ship station and the shore station the telephone will operate like any other telephone operator-assisted call. On calls from the ship the user has only to advise the telephone number of the party he wants. The operator will then place the call into the terrestrial network. Interconnection to the ship is made and real-time, full duplex, voice communications can proceed. In the future, after the reliability of the automatic interconnection features have been proven, it will be possible to have a direct dial, real-time, automatic telephone system.

Facsimile: This service will be available over the voice channel. Signals may be either digital or analog depending upon the equipment used. Normally, the user will probably employ the same equipment that his company has in its various offices. The message format or use of special forms will be those now being sent or received in the company's shore organization.

Data: This term includes data at speeds from 300 bps to 240,000 bps. Here again the particular modem in use is the controlling factor. While the U. S. Navy is using the satellite the system is power limited in the shore-to-ship direction. It is not limited in the ship-to-shore direction. Thus, while there is only one voice and 44 telex channels available in the out-bound direction, there are 14 voice equivalent channels available in the in-bound direction. Because of present day limitations

with existing terrestrial data circuits, only 2400 bps in the U. S. A., and 1200 bps in Europe, can be sent. If conditioned direct circuits are installed between the shore station and the customers officer, 4800 bps data communications will be available. Additionally, in the in-bound direction it is possible to have data communications at speeds up to 240,000 bps provided the proper modem equipment is installed and the output power of the shipboard terminal is increased in relation to the data speed desired.

Distress: An emergency request from a ship terminal receives priority over all other traffic. A telex channel will be assigned immediately upon receipt of an emergency request. In the unlikely event that all telex channels are being used, the network control equipment will immediately clear a channel and assign it to a ship in distress. At the earth station an immediate connection is made to a direct circuit to the rescue authorities. On the East Coast of the U.S.A. this is the Commander Atlantic Area, U. S. Coast Guard, located in New York City. On the West Coast it is the Commander Pacific Area, U.S. Coast Guard in San Francisco, California. These commands have direct control of rescue ships and aircraft and access to many supporting organizations and communication facilities. If voice communication is desired, such a circuit would be made available. Considering the average of one distress case per day and the more limited number of ships expected to be outfitted with a MARISAT terminal, there is virtually no likelihood of delays in the handling of distress calls. The terminal itself has been reduced to utmost simplicity. It is only necessary to press the DISTRESS button. This button, which is protected by a short, round metal tube, is immediately accessible in front of the operator.

It should be recognized that the inclusion of this distress feature is not required by any existing Safety Convention nor governmental requirement since a maritime satellite terminal is entirely a voluntary installation. It has been included in MARISAT because in the future it is expected to be made mandatory requirement and the experience with MARISAT will permit the maritime safety authority (IMCO) to gain factual knowledge to assist in their decisions at future revisions of the Safety of Life at Sea Convention.

Broadcasts: It is possible to use the MARISAT System for transmission to ships of broadcasts, such as weather, hydrographic or navigational information, data on ship collisions or wrecks, or anything of interest to more than one ship. It can be used for a general message to ships of a specific shipping line, to ships of a common nationality, or those sailing in a certain geographical area. Such ship groups all can simultaneously receive a teletype message transmitted by the earth station provided the ship is not in the midst of handling normal communication via the satellite.

International Aspects

International efforts have been underway in the International Governmental Maritime Consultative Organization (IMCO) to reach long-range policy decisions regarding maritime satellite communication services. In March 1972 IMCO formed a Panel of Experts on Maritime Satellite (POE) to study the technical, operational and economic factors involved in a commercial maritime satellite system. It is the stated intent of the MARISAT System design not to prejudice such efforts.

Over the past decade a number of study efforts have investigated requirements for a global maritime satellite system. Much of this work was summarized in the report of the 1971 CCIR Special Joint Meeting (SJM). The IMCO-POE three year study, drawing heavily on the SJM results, has arrived at some measure of international consensus concerning initial operational features to be incorporated in such a system.

Generalized operational requirement guidelines derived from the IMCO-POE study indicate that a first-phase (an initial seven years of operation) international maritime satellite system should provide the following:

- a) Full-time coverage between $\pm 70^{\circ}$ latitude. Periodic daily coverage to $\pm 82^{\circ}$ latitude is desirable.
- b) Simplex and duplex channels for voice, data, facsimile and teleprinter.
- c) Special handling for vessel distress and safety messages.
- d) Collection and dissemination of environmental data on individual and fleet basis.
- e) Rapid and automatic operation with interconnection to the worldwide telecommunications networks.
- f) The option for radiodetermination and other possible future requirement.

THE FUTURE

The first session of the "International Conference on the Establishment of an International Maritime Satellite System" was held in London, England in April and May of last year. This Conference was the result of a number of years of effort in the Inter-Governmental Maritime Consultative Organization and the International Telecommunications Union. The Conference agreed that there was a need for a worldwide maritime satellite system in order to improve maritime communications. It also agreed that if such a system were established, there was a need for an International Inter-Governmental Organization to administer and manage it.

A second session of this Conference met in London in February 1976. Between these two Conferences there were meetings of working groups to resolve various issues and prepare for consideration by the Conference draft agreements relating to the establishment of an International Maritime Satellite Organization.

The results of these Conferences will be a vital factor in the development and implementation of a follow-on maritime satellite system to MARISAT. This interest on the part of many countries to ensure that ships on the high seas continue to have improved satellite communications leaves little doubt as to the successful future of maritime satellite systems.

THE NORWEGIAN DOMESTIC COMMUNICATION SATELLITE SYSTEM

John R. Veastad

Director, Radio Services Division, Norwegian Telecommunications Administration, Oslo

GENERAL COMMUNICATION REQUIREMENTS

Radiocommunication for off-shore oil rigs has for years been carried out by conventional MF (medium frequency) radiotelephony, usually on a one-channel simplex basis. This has been an established method in the Mexican Gulf and other well-known areas and it was also the most likely means of communication to and from drilling platforms, barges, supply vessels and so on in the North Sea.

However, with the establishment of production units, often comprising several platforms of different functions, it became apparent that conventional means would not be able to carry the number of telephony, teleprinter, facsimile and even high capacity data circuits necessary for monitoring and controlling the mechanics of the entire extraction process of such big plants that far off shore.

For exploration most oil rig operators have solved their communication requirements by one simplex MF radiotelephone channel between the shore base and several drilling platforms. In some cases teletype or facsimile circuits have been established. For operations on the Norwegian Sector of the shelf these radio installations have been licensed by NTA. Similar arrangements have been used during the construction phase where several users, the production unit owner, barges, supply vessels, and so on have operated on the same "construction frequency".

As the oil field develops to a production complex requirement for 2-10 duplex telephone circuits, several telex and some facsimile circuits as well as some data circuits (usually 2400 Baud) have been put forward. One important factor here is the need for a very high reliability (time availability factor of 99.9% or higher).

To take care of the needs for reliable radiocommunication systems of moderate capacity, conventional tropospheric scatter systems of 12-120 channel capacity seemed to be the solution. This because distances to be covered are 100-250 miles, well beyond the maximum distance for line of sight microwave links. However, in the future the distances to be covered might extend out to 600 km or more, especially north of the 62nd parallel and in the Barents Sea.

The troposcatter system in question as an alternative for the connection to Norway would need 1-5 kW transmitters, two 8 m diameter antennas on each terminal giving a possibility of quadruple diversity arrangements to provide 99.9% (of the worst month) or better time availability. This could give rise to a considerable installation and maintenance problem. Since Norway would

have had to establish at least three highly reliable links to Ekofisk (PPC), Frigg (ELF) and the Statfjord (Mobil) or more fields, this would require at least two troposcatter terminals on the mainland with 6 or more parabolic antennas, see Fig. 1. For every new prosperous field a need could arise for a new link, with its terminal equipment and two antennas at each site for quadruple diversity. Depending on location and distance to be covered it might be necessary with a new station building and associated constructions also on the mainland.

A far more elegant technical solution is the use of telecommunication satellites, since one single earth terminal in Norway via one satellite can cover a great number of terminals out in the North Sea, see Fig. 2. And each production unit or platform would need only one antenna regardless of number of links in different directions. Within a dedicated system it could be possible with one operating satellite to cover all communication requirements for a relatively great number of production units, compressor or pump stations, as well as drilling platforms. The INTELSAT system does not, at present, due to wording of the basic agreement, offer possibilities of multilateral arrangements covering such communication complexes.

THE AGREEMENT BETWEEN INTELSAT AND NTA

The Norwegian Telecommunications Administration (NTA) started discussions on the possible use of satellites with the North Sea operators early in 1972. However, at that time less suitable lease arrangements within INTELSAT for satellite capacity would, together with, for instance, the tariffs for leased lines from Tanum (Nordic earth station) to Stavanger, lead to excessive costs, even for telex circuits. Besides, such a technological solution seemed unproven to most oil companies. Relatively high reliability figures for the INTELSAT satellites and earth stations in this system did not count, since the price would be too high.

As INTELSAT in early 1974 indicated a possibility of leasing capacity in their spare satellites on a "preemptible basis" for an acceptable price, the situation became completely different. After negotiations within INTELSAT, NTA was offered the possibility of obtaining a lease arrangement for capacity equivalent to half a transponder, i.e. sufficient capacity, bandwidth and powerwise, for a total of more than 150 channels in a domestic single channel per carrier (pcm psk) system comprising some 10 small earth stations. The lease contract is for five years at a yearly rate of \$ 0,5 Mill.

It should be mentioned that the reliability of the INTELSAT system today is beyond questioning. Figure 3 shows the continuity of the global service for the entire INTELSAT system, i.e. 79 earth stations with 100 antennas. As will be seen the average global continuity exceeds 99.9%.

The average number of system outages per earth station per month is shown in Fig. 4 which gives reason to expect that, in a limited, domestic system, the continuity will be well above 99.9%. Problems in connection with a changeover of all antennas from one satellite to another will occur. In most cases however, such interruptions will be known on beforehand and the operation can be planned.

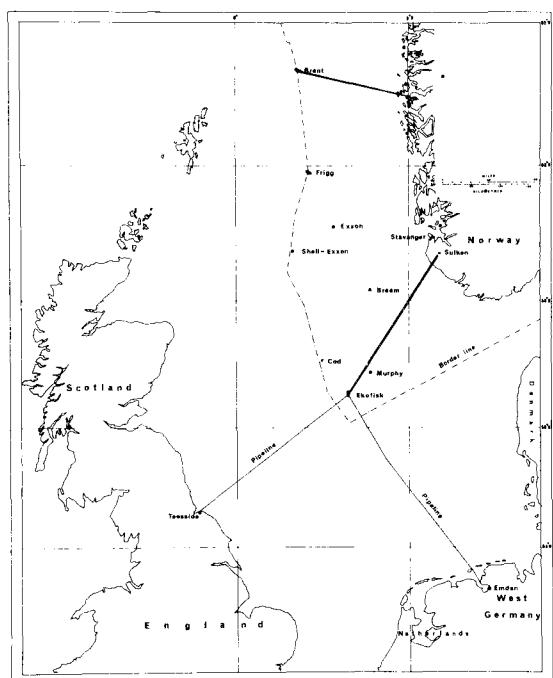


Fig. 1.

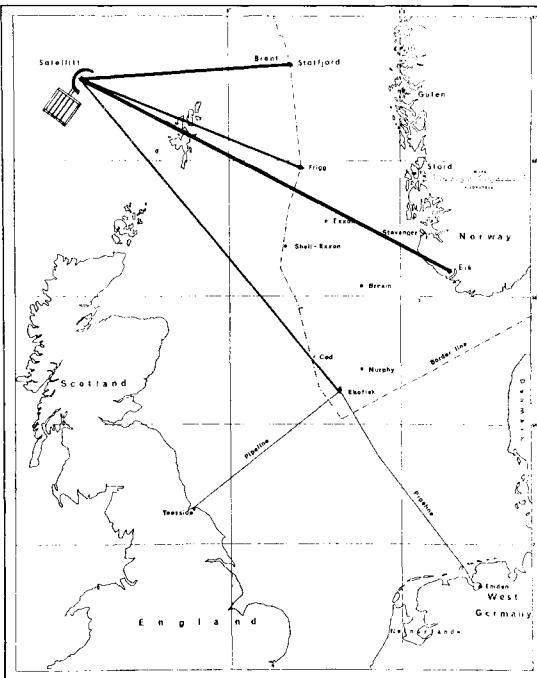


Fig. 2.

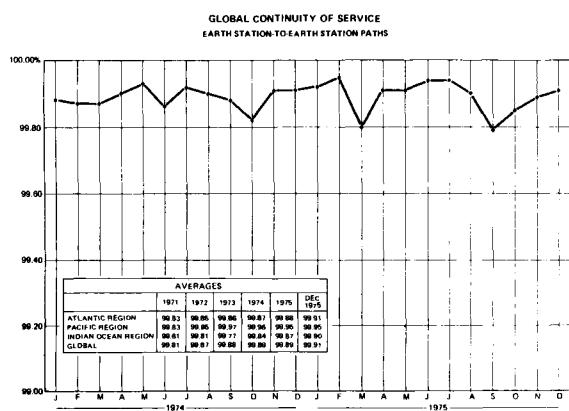


Fig. 3.

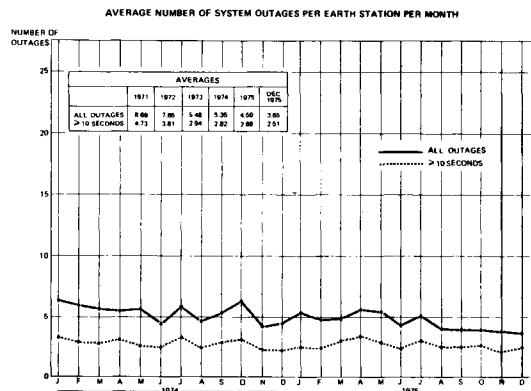


Fig. 4.

ARRANGEMENTS BETWEEN THE OIL FIELD OPERATORS AND NTA

According to Norwegian law there is a monopoly situation where Televerket (NTA) can give a licence according to certain rules for a limited, single private communication system, - but not a carrier where several users are involved. This means that the use of a high capacity troposcatter system from the mainland out to one platform with a distribution via line of sight (or tropo) links to several users would have to be established, owned and operated by NTA. A low capacity tropo or satellite link would be a different situation. In the case of the present domestic satellite system, NTA is the formal owner and operator of the system as such, whereas the oil field operators are operating their own satellite terminals on the platform according to a NTA licence. This also fulfills the INTELSAT requirement of NTA having full technical control over the system.

The agreement between the three oil field operators Phillips Petroleum Company Norway (for Ekofisk), ELF Norge A/S (for Frigg), Mobil Norge A/S (for Statfjord) and NTA states that all investments in the main station and the space-segment capacity (increase in the total equivalent number of units allotted to Norway) will be covered in equal parts by the users, whereas the operating costs (including lease of 1/2 transponder) will be divided according to their use.

An expansion of the system to carry 12-24 telephone channels to Svalbard is foreseen, whereby NTA will contribute 1/4 of the total investments. Present requirements for each oil company are some 12-24 channels including voice, teleprinter, facsimile and data. The system has the advantage and flexibility of allowing future expansion of capacity for each user as well as an inclusion of new partner stations on the Norwegian shelf or in Norwegian territories. Each additional operator (oil company) joining the system in the future will have to contribute his part of the depreciated investment.

According to the requirements which were stated by the oil companies, the system was planned to be operational by 1 November 1975, one year after the signing of the equipment contract. This very tight schedule will be met for the main station, and it is based on the spirit of interest and excellent cooperation shown by the manufacturers as well as the oil companies.

DESCRIPTION OF THE SYSTEM

The site which has been chosen for the main station is in Rogaland, about 60 km south-east of Stavanger. There are hills to the east and the west, but the horizon in the direction of the satellites is suitably low, see Fig. 5. The station will mainly use the satellites above the Atlantic, the ones above the Indian Ocean will serve as alternates. There is space for additional antennas, and the building is planned to accommodate more equipment than will be needed at the first stage.

The Cassegrain type 13 m antenna will be erected near the station building, see Fig. 6. It has been designed to permit quick repositioning between the satellites above the Atlantic and the Indian Oceans.

The transmit gain is 56 dB at 6 GHz while the receiving gain is 53 dB at 4 GHz. The noise temperature at an elevation of 10° is 45 K.

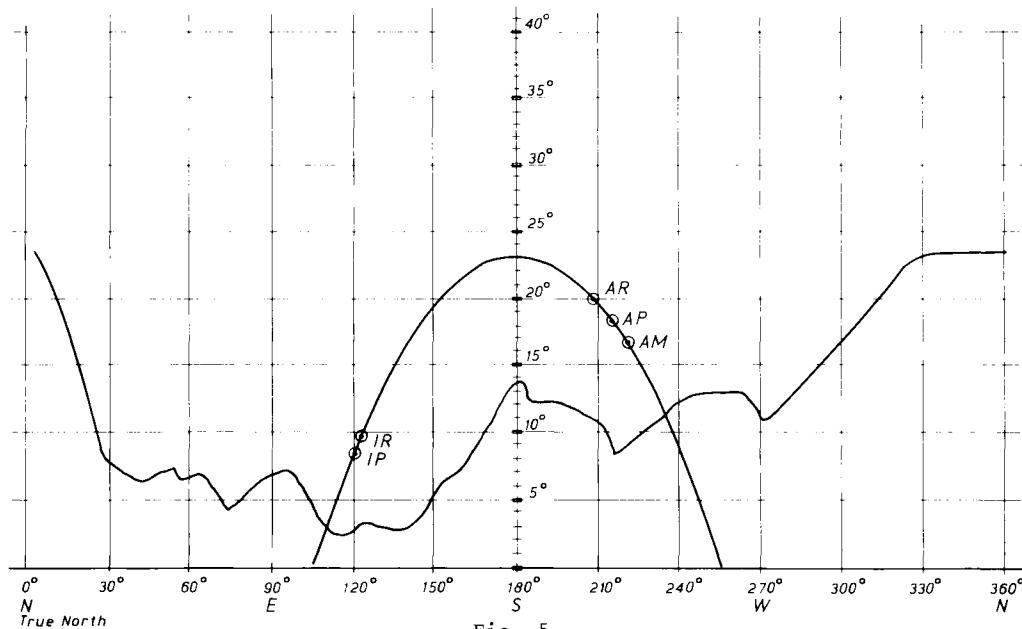


Fig. 5.

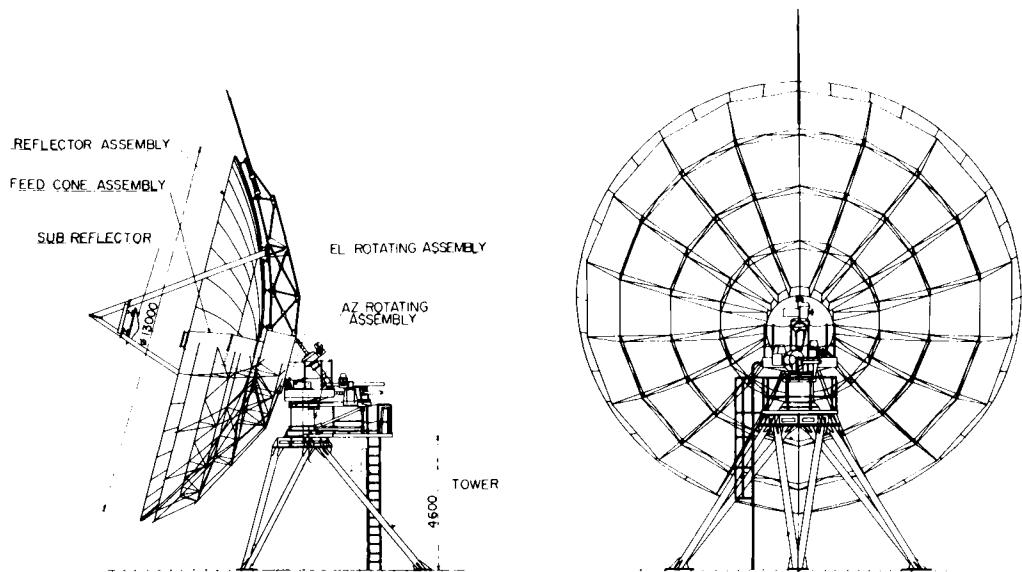


Fig. 6.

The antenna and feed can be used over the full 500 MHz bandwidths of the 3.7 to 4.2 and 5.925 to 6.425 GHz bands used by INTELSAT.

The station is equipped with redundant transmitter using 1.5 kW klystron highpower amplifiers.

The receiver front end consists of a redundant 500 MHz bandwidth low-noise receiver amplifier operating at a noise temperature of 55 K, in the 4 GHz band.

The channel equipment is of the single channel per carrier type for transmission of Delta-M encoded voice signals. There are also channel units for transmission of digital data signals and, finally, signalling control and supervision signals.

A modern computer will supervise the main station and also remotely control some functions at the outstations.

A 425 kVA Diesel generator and no-break supply with supporting battery will make the station completely self-sufficient during periods of possible power failures. Normally the no-break rotary converter runs from the mains supply.

The first three outstations will be established on the Ekofisk, Frigg and Statfjord production platforms in the North Sea, see Fig. 7. Apart from the smaller antenna with a diameter of 8 m, the outstation equipment will be almost identical to that of the main station.

STATION AND OPERATION COSTS

Many attempts have been made, within and outside Norway, at comparing costs between troposcatter systems and satellite systems in the North Sea. A salesman, the oil company or NTA as provider of the service, may all come to different conclusions. It is obvious that, whereas a troposcatter system would score most points by the provision of a limited number of circuits (i.e. 12-24) between two single points, a satellite solution will be more rational in a complex system with a greater number of circuits going to many destinations.

Within a relatively short period three operators in the North Sea expressed their future requirements, 12-24 telephone channels plus highly reliable data circuits. Since NTA had, and still has, a strong interest in establishing 12-24 telephone circuits to Svalbard, our cost comparison had to be based on three or more links with a capacity of at least 48-120 circuits. (With some modifications of the equipment and an increase in capacity to a full transponder, program channels for broadcasting and a TV programme channel can be established.)

In addition, drilling is still going on in the North Sea and the establishment of additional production facilities can be expected.

The two competing systems would have to consist of:

1a) troposcatter with three (or more) links with at least one terminal station at each platform, two stations on the main land, each of 60, 72 and 120 channel capacity to allow distribution to neighbouring production units via

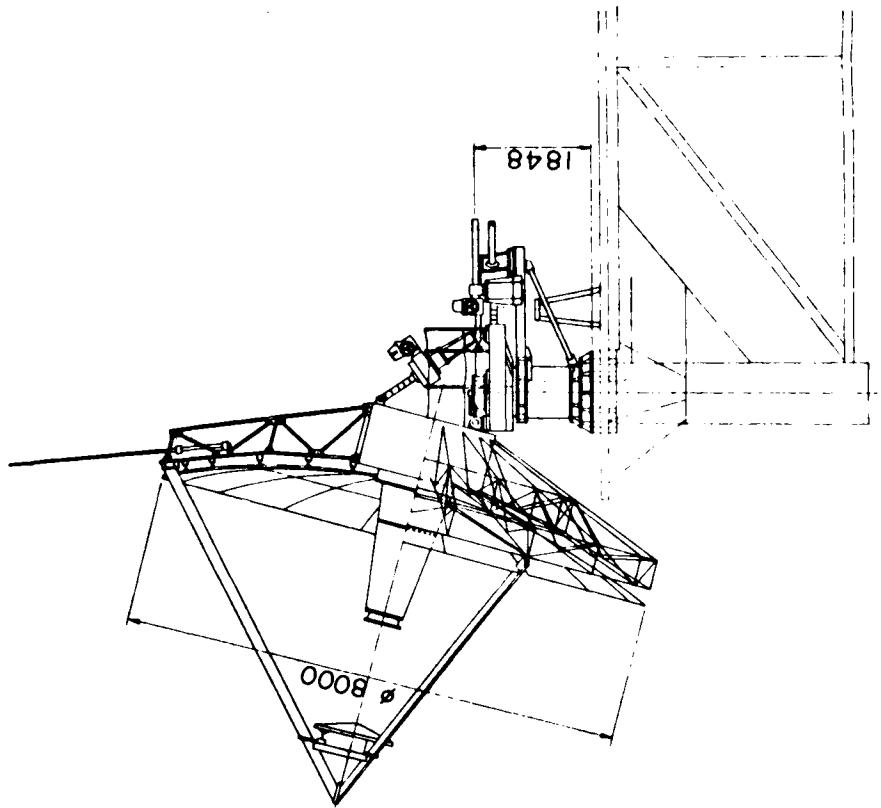
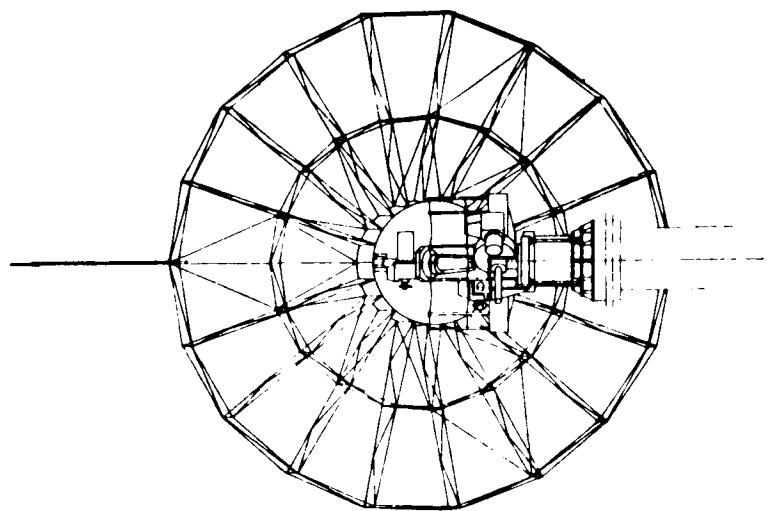


Fig. 7.

line of sight links (LOS),

1b) as 1a) plus a terminal station with a 12 channel capacity link to Svalbard, including a relay station at Bjørnøya (a greater capacity than 12 channel is not practical at these hop lengths),

2a) satellite service to three (or more) platforms, which is relatively flexible as regards transmission capacity (at least 150 channels via 1/2 transponder), comprising one terminal in Norway and three (or more) in the North Sea,

2b) as 2a) plus a 24 channel link to Svalbard.

According to our calculations the investments for these alternatives would be: (See Fig. 8.)

1a)	2a)	1b)	2b)
\$ 3,75 Mill.	\$ 6,6 Mill.	\$ 6,7 Mill.	\$ 8,55 Mill.

Additional links to the North Sea:

tropo system: \$ 1,33 Mill. (average)
sat. " : \$ 0,93 Mill.

These curves show that it would be highly advantageous to use a satellite system if a link to Svalbard is established. However, for links to the North Sea only, the difference is less stringent. The crossing point seems to be more close to 5 links than 3 or 4, as earlier anticipated.

As regards operating cost, see Fig. 9, estimates for the future development of the system seem positive. It must be admitted that with only 3 links the total operating costs may be somewhat higher in the satellite system. However, since so many unknown factors are involved, no definite statement in that respect should be made.

We feel confident that the flexibility of the satellite system, which is of great importance for the future development in the North Sea, is a factor of great value to Norway and all other oil consuming countries as well.

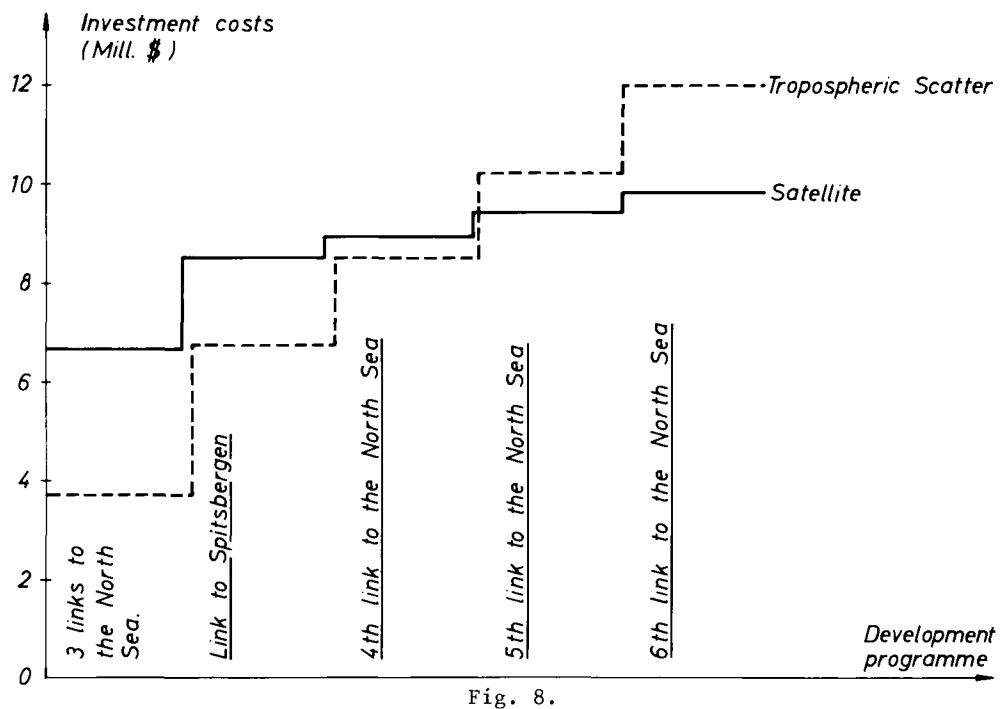


Fig. 8.

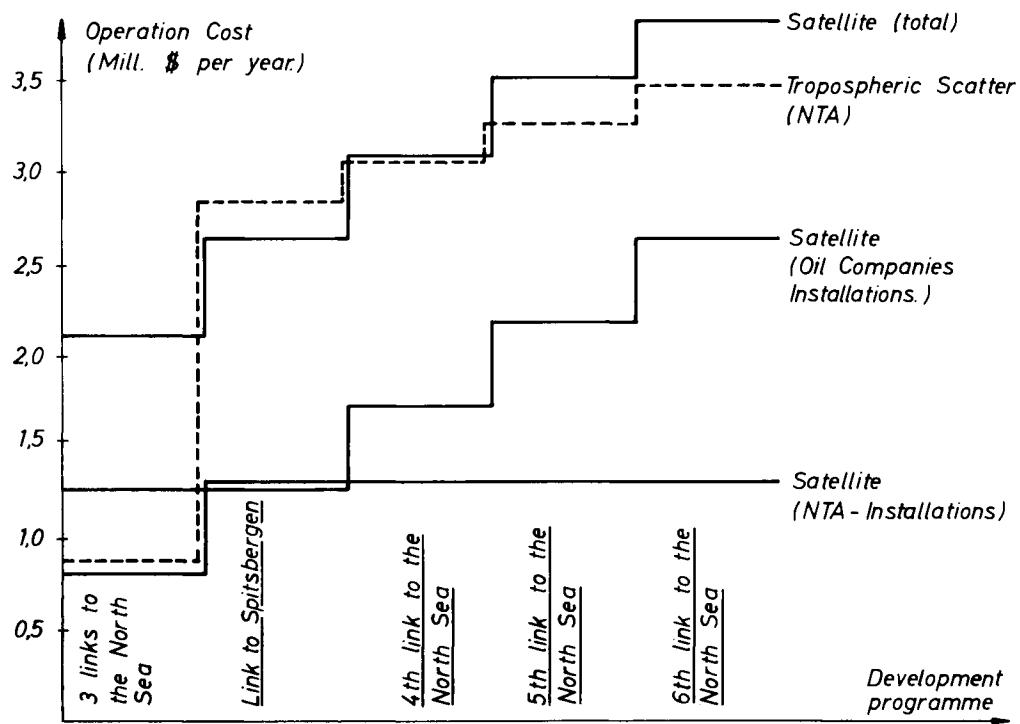


Fig. 9.

ACQUISITION AND PROCESSING OF METEOROLOGICAL DATA FROM THE SMS/GOES SATELLITE

Henry J. Goett and Fred W. Stang

Aeronutronic Ford Corporation, WDL Division, Palo Alto, California 94303

SUMMARY

The first Synchronous Meteorological Satellite (SMS) was launched May 17, 1974; the second, February 6, 1975. Thus, at this time, some 23 months of satellite operational experience have been acquired on the SMS system. The current phase of the SMS program consists in the development of techniques which adapt the satellite data to the needs of the eventual user -- the operational meteorologist.

This paper describes the data obtained from the two SMS/GOES satellites now in orbit and certain uses of these data by the National Oceanic and Atmospheric Administration (NOAA) in conjunction with its operational National Weather Service. In addition, this paper discusses a data acquisition station and minimal data processing equipment appropriate for regional users without access to NOAA's weather network. This equipment will enable regional users to incorporate SMS data into regional operational systems. Potential uses of these data are also discussed.

INTRODUCTION

The first SMS was launched May 17, 1974; the second, February 6, 1975. Thus, at this time, some 23 months of satellite operational experience have been acquired on the SMS system. After initial checkout by the National Aeronautic and Space Administration (NASA), these satellites were turned over to NOAA for operational use. Future satellites of identical design, designated Geostationary Operational Environmental Satellites (GOES), will be launched for NOAA's operational purposes during 1975-80.

The current phase of the SMS program consists in the development of techniques which adapt the satellite data to the needs of the eventual user -- the operational meteorologist. This adaptation is a phase in the development of all application satellite systems that does not attract the attention given the satellite development itself. However, it is a highly innovative, experimental process that is vital to the eventual utility of an application satellite.

Initial use of SMS/GOES data has been in conjunction with the NOAA National Meteorological Center (NMC) and the National Weather Service (NWS) in Suitland, Maryland. Primary interest has therefore been the incorporation of cloud data into weather analyses heretofore based primarily on more conventional weather data accumulated by ground-based facilities. New techniques have involved the use of fairly extensive photographic facilities and large computers. This paper will give a progress report on the use of SMS data by these new means.

Use of SMS satellite data by regional facilities without access to the internal United States weather network, over which NWS predictions are transmitted, is in the more formative stage. This use requires a small data acquisition station, for reception of

data directly from the satellite, and minimal data processing equipment. This paper describes such a station and associated processing equipment. In addition, potential uses of SMS data of particular interest to regional users are discussed.

I. DESCRIPTION OF THE SMS SYSTEM

The Synchronous Meteorological Satellite system and the objectives of this NASA program are described in some detail in Ref. 1. More complete details on the satellite and the system data flow are provided in Refs. 2 and 3.

An SMS system overview (Fig. 1) indicates the various functions this system performs. The SMS primary objective is to obtain cloud images from synchronous orbit. These digitized images are transmitted to the National Environmental Satellite

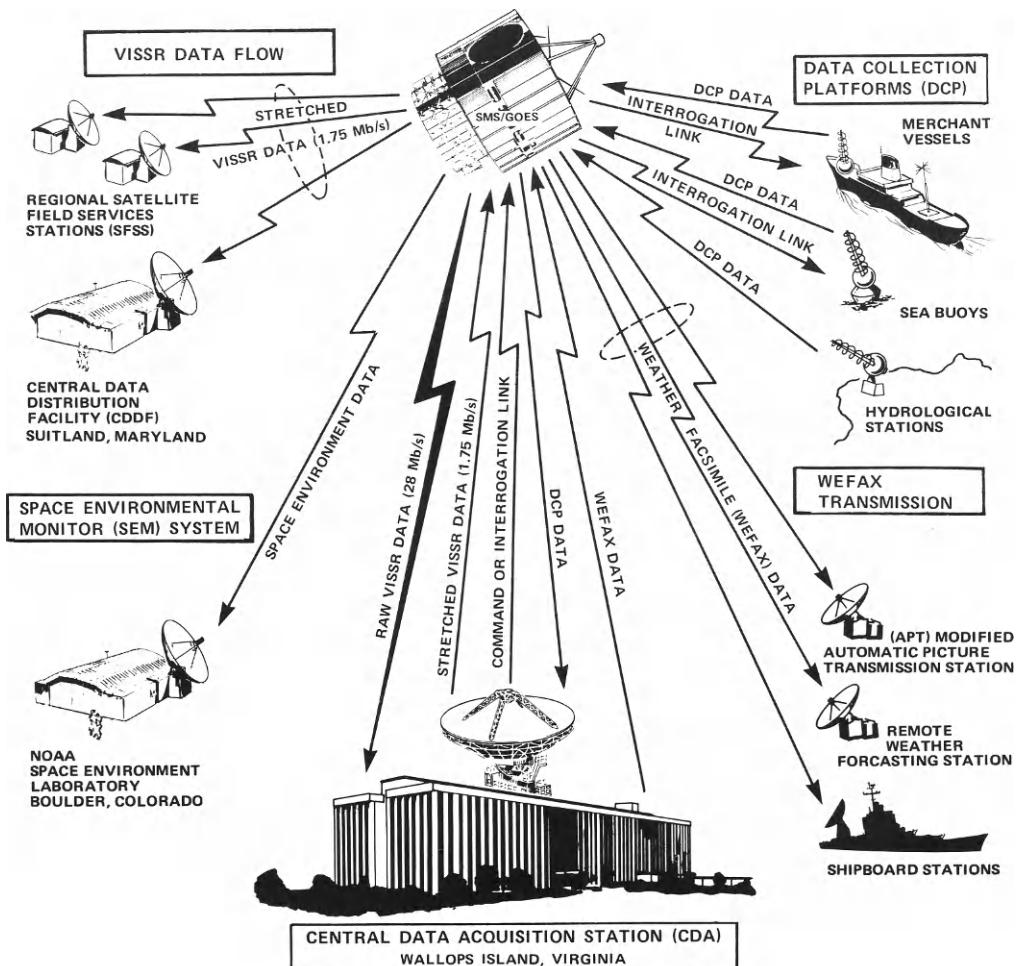


Fig. 1. SMS system and data links.

Service (NESS) Command and Data Acquisition Station (CDA) at Wallops Island, Virginia, over a wideband (30 MHz) channel at 28 megabits per second (Mb/s). Images are normally taken at 30 minute intervals, although this can be reduced to 5 minute intervals if the North/South field-of-view is reduced. It is this frequency of images from geosynchronous orbit that distinguishes SMS from lower-orbiting satellites of the TIROS/ESSA/ITOS/NOAA family and opens up expanded cloud image usage for weather forecasting.

The SMS performs additional functions, making it not only a meteorological satellite but also a communication satellite, collecting and disseminating meteorological data. As Fig. 1 indicates, these auxiliary functions consist of the following:

a. Retransmission of images at narrow bandwidth (4.0 MHz) with a reduced data rate (1.75 Mb/s). This enables both the NESS Data Processing Facility in Suitland, Maryland, and any regional user within view of the satellite to receive the images in real time but with an antenna smaller than the 60-foot antenna at Wallops Island, which receives the primary wideband image.

b. Collection of data from small data collection platforms (DCP's) on ocean buoys, unattended hydrological stations, ships, and earthquake monitoring stations. Through these links, data can be obtained from unattended stations relative to river heights, rain and tide measurements, wind speed and direction, air and water temperature, radiosonde data, and earthquake monitoring. The interrogation link can command a readout from selected stations and send a teletype message to any ship equipped with a relatively simple antenna.

c. Data collection from a space environmental monitor (SEM) on board the satellite. The SEM measures the direction and magnitude of the earth's magnetic field, solar X-ray flux, and energetic particle distribution at the satellite location. These data are used by NOAA's Space Environment Laboratory at Boulder, Colorado to monitor solar disturbances.

d. Transmission of narrowband weather facsimile (WEFAX) data to automatic picture transmission (APT) stations for use in conjunction with lower-orbiting NOAA satellites which have been modified as described in Ref. 4.

Reference 5 gives a concise description of the spacecraft; its communication, power, attitude control, and propulsion subsystems; and its instrumentation payload. Figures 2 and 3 and Appendix I are abstracted from Ref. 5.

Two SMS spacecraft are currently in orbit. They are in a 35,000 km geosynchronous orbit at 75° west longitude (SMS-1) and 115° west longitude (SMS-2). SMS-1 has approximately a 1.9° inclination; SMS-2 has an inclination of 0.7°. From these locations the satellites see all of North and South America and adjacent ocean areas. Figure 4 shows the useful SMS camera coverage.

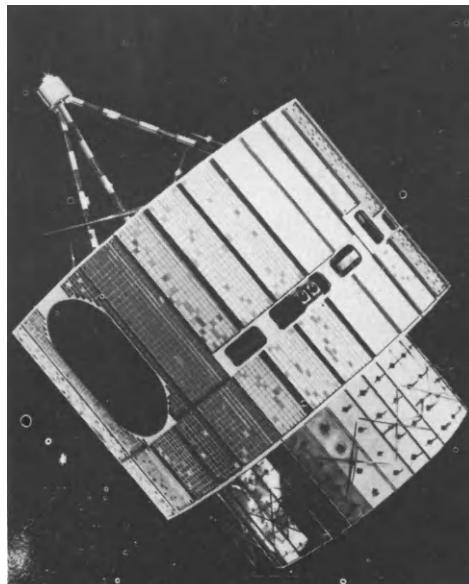


Fig. 2. SMS/GOES spacecraft.

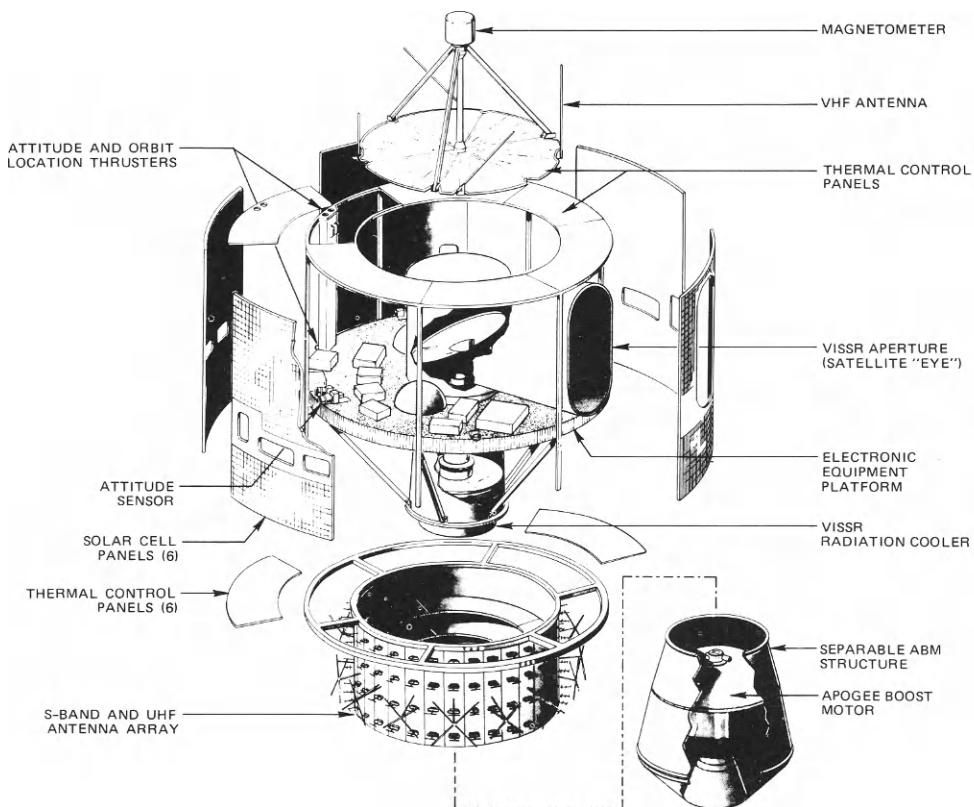


Fig. 3. SMS/GOES spacecraft, exploded view.

The satellites rotate at 100 revolutions per minute, and the visible and infrared spin scan radiometer (VISSR) scanning mirror (Fig. 5) scans the earth from west to east for the 17° when the earth is in view. As shown schematically in Fig. 6, there are eight identical visible channels viewing the earth on each scan (in a north/south alignment), and one (redundant) infrared channel. During spacecraft rotation when the mirror is not viewing the earth, it steps so that on the next earth pass it views the next southward line.

With such a scan pattern, the VISSR accomplishes the 1821 scan steps required to provide a high resolution image of nearly one-quarter of the earth's surface within 18.2 minutes. (North/south coverage can be decreased on command, and the time per reduced-coverage image decreased to as low as 5 minutes.) Resulting visible images (0.55 to $0.70 \mu\text{m}$ band) contain 14,568 lines and have a resolution of 0.9 km; infrared (IR) images (10.5 to $12.6 \mu\text{m}$ band), with 1821 lines, have a resolution of 8.9 km. Figure 7 is a typical SMS visible image, and Fig. 8, an IR image.

It is to be noted that the images presented have lost resolution in the photographic reproduction process; a photo 22 by 22 inches would be needed to preserve the resolution inherent in the basic data. Such photos are made with a laser scanner, but this equipment is not generally available. To avoid this loss in resolution, a normal

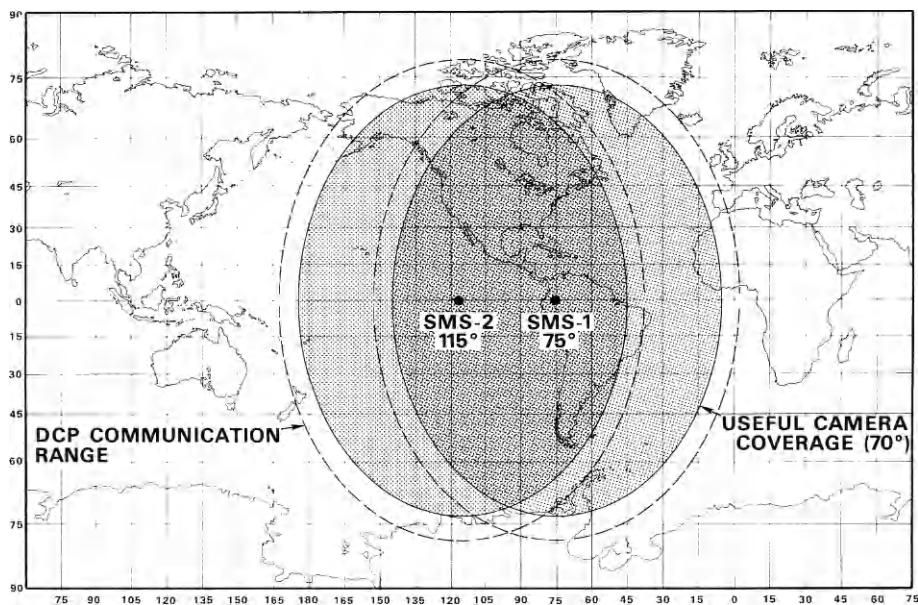


Fig. 4. SMS/GOES camera coverage.

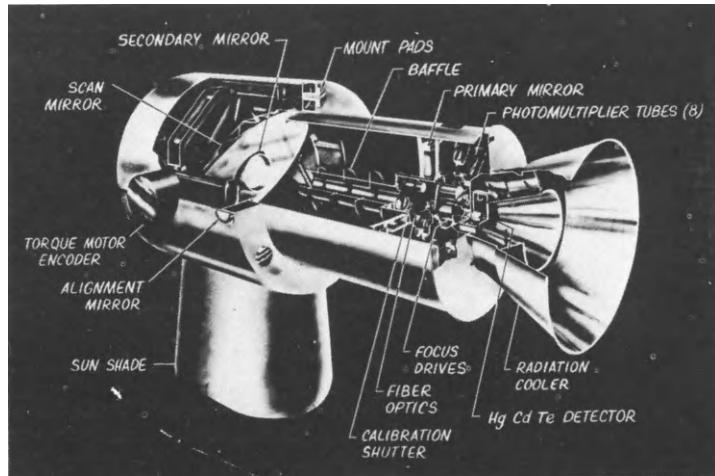


Fig. 5. Visible and infrared spin scan radiometer.

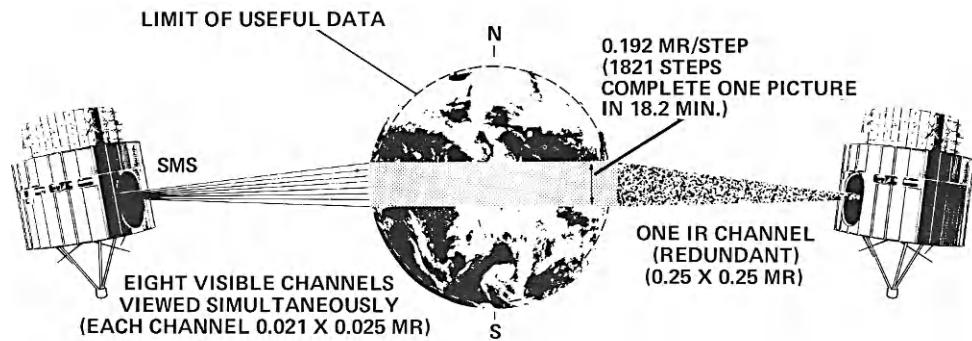


Fig. 6. VISSR visible and infrared scan elements.

operational procedure is to "sectorize" the full-earth image and show only a particular region of interest. Figure 9 is a sectorized SMS image. (The cyclonic disturbance shown in the middle bottom half of this image is 1975 Hurricane Connie; the disturbance to the east of it, just off the coast of Mexico, is 1975 Hurricane Debbie prior to its full development.)

II. OPERATIONAL USE OF GEOSTATIONARY SATELLITE IMAGES

The availability at half-hour intervals of visible and IR cloud images, such as those shown on Figs. 7, 8, and 9, affords potential use for a number of meteorological, oceanographic, and hydrologic purposes. However, before this can be done, processing and analysis techniques must be developed for the extraction of quantitative observations on a routine operational basis. NOAA Technical Memo NESS 64 (Ref. 6) provides an excellent summary of some techniques in use or under development by NESS. Additional NOAA memoranda, in a series from NESS 21 through NESS 63, describe applications of SMS-type data which may eventually find their way into operational use. The development of these uses and techniques was given a headstart by the prior availability of visible spin scan camera images from Applications Technology Satellites (ATS-1 and -3). However, with continuous day and night imagery now available from SMS, plus the additional information contained in IR images, an impetus has been given to the development of techniques for operational use of this type of data.

The following sections describe processing techniques and applications already in daily operational use by the NOAA National Meteorological Center and the National Environmental Satellite Service. More complete details are provided in Refs. 6, 7, and 8.

Image Alignment

A comparison of cloud positions on successive SMS images is used to estimate wind velocity. Before the images can be so used, "apparent motion," due to slight pitch and yaw of the camera with respect to the earth, must be removed. It is probably feasible to maintain a stable spin axis perpendicular to the orbital plane and thus eliminate resultant "look-angle" variations from this source. However, it is most difficult to keep the subsatellite point fixed. A geosynchronous orbit inclined only 1° will allow subsatellite motion up to 120 km/h. This shows up in successive images as



Fig. 7. Visible full-earth SMS image.

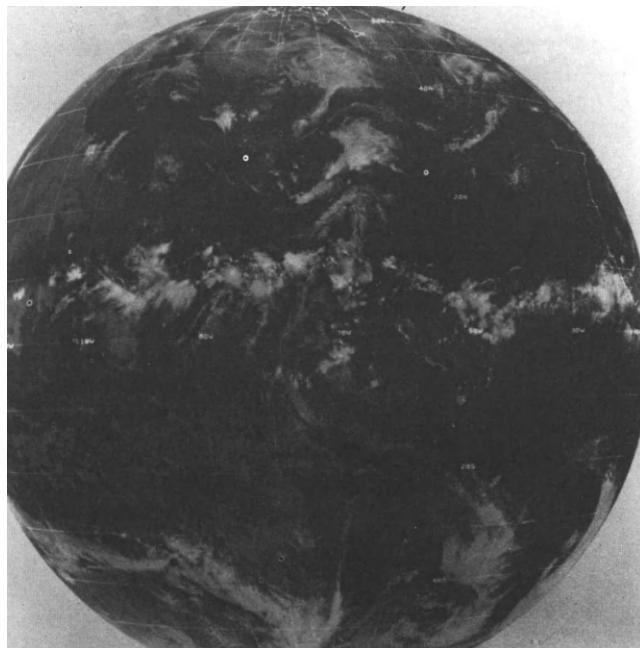


Fig. 8. Infrared full-earth SMS image.

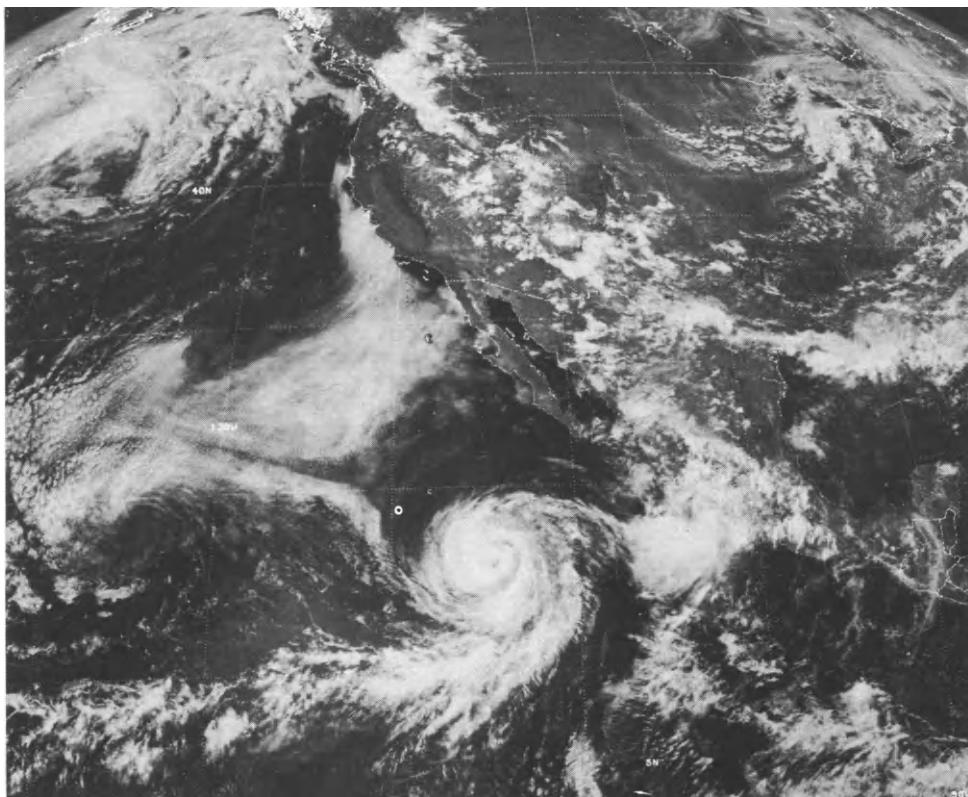


Fig. 9. Visible sectorized SMS image.

equivalent motion of earth and clouds. An analysis of the error introduced into the calculated wind vector at various points in the field of view is presented in Ref. 9.

To compensate for this apparent motion, successive images must be aligned before cloud position can be used to determine wind vectors. This is done by landmark matching, a procedure described in Section 12 of Ref. 6. Figures 10a, 10b, and 10c illustrate successive steps in this process.

Successive images are registered by matching sufficient locations of cloud-free landmarks, or control points. Figure 10a shows a map of the Salar de Uyuni Salt Flats of Bolivia, which are generally visible. These salt flats stand out as a bright spot on visible SMS images (Fig. 10b). Figure 10c shows the gray scale depiction of this area as printed out by the computer. The distance between the cross (+) and the circled cross, (Θ) represents satellite motion error which is removed by registering two successive images on this point.

Wind Estimates

The first step in wind estimation is the assembly of a film loop, a time-lapse motion picture of four successive IR cloud images that have been aligned as described in the

previous section. The film sequence covers a real elapsed time of 2 hours but is played in approximately 6 seconds. The result is repetitive cloud pattern movement which emphasizes the individual cloud movement limits during the 2-hour real-time interval. Projection of such a film loop permits selection of very well defined clouds on any given image. In the infrared mode, the analyst has little difficulty distinguishing between low, middle, and high altitude clouds suitable as motion tracers to define the wind in which they are immersed. SMS IR imaging is presented photographically in approximately 18 shades of gray, ranging from warm temperatures (300 K), which are black on the image, to cold temperatures (220 K), which are white. Thus, the sea surface (dark), low clouds (medium gray), middle clouds (light gray), and high clouds (white) can be readily identified, and any given cloudy area can be traced from one image to the next.

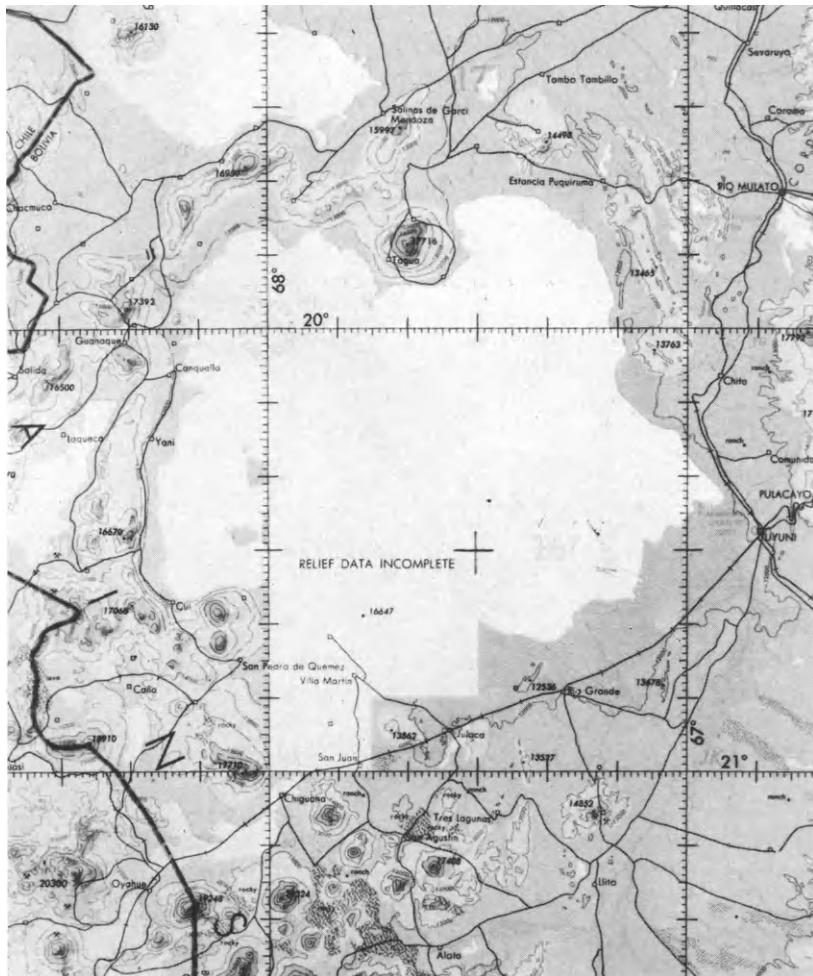


Figure 10a. Detailed map of Salar de Uyuni Salt Flats in Bolivia.

† 15:00 300:74 11-A-2 0025 1911 WDS 2MI 7A2 CH1



Fig. 10b. SMS-1, 2x2-mile resolution visible image, 1500Z, October 2, 1974.

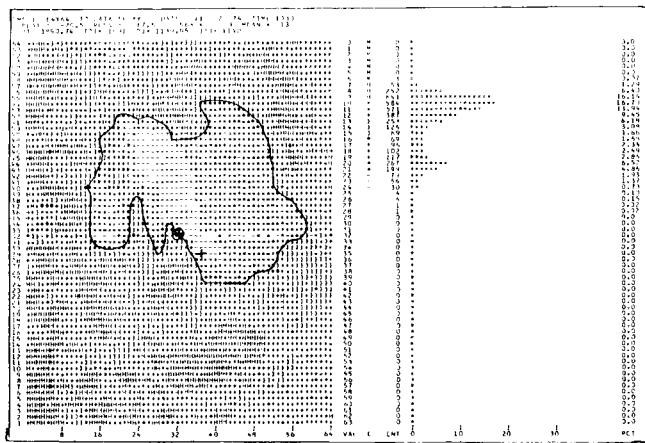


Fig. 10c. Gray scale depiction, using alphanumeric characters, of the Salar de Uyuni as observed by SMS-1 at 1030Z on November 2, 1974.

Figure 11 shows the relatively simple apparatus used to extract cloud position data from such film loops. The overhead closed-loop projector displays the film loop on a paper worksheet placed on the plotting board. A pencil mark is made at the initial position of the cloud element on the worksheet, a second mark at the final position, and an arrow to indicate direction. Translation of the above information into longitude and latitude and an identification of time interval provides sufficient data for the computation of cloud motion, which is assumed to be wind velocity. Computations are performed most efficiently on a computer. For this purpose, the data is digitized, ie, the points marked on the worksheet are translated into X, Y components. Coupled with a conventional card punch, the digitizer supplies properly formatted numerical data for computer computations. (The foregoing is a summary description of the wind estimation technique in operational use at NESS in June 1975. For more details, see Ref 6, Section 13, The GOES Wind Operation, by M. T. Young.)

An alternative approach is described in Ref 6, Section 12, The Automatic Extraction of Wind Estimates from VISSR Data, by R. Green, G. Hughes, C. Novak, and R. Schreitz. This method is more automatic than the one described above but is not yet in routine operational use. By computerized crosscorrelation, it compares a digital 32 by 32 inch sample image array, with an overlying 64 by 64-inch array from a later image. Cloud displacement during the interval is identified by crosscorrelation of the two images. By this method there is less human involvement in the process, but there is a need for a larger computer than with the previously described technique.

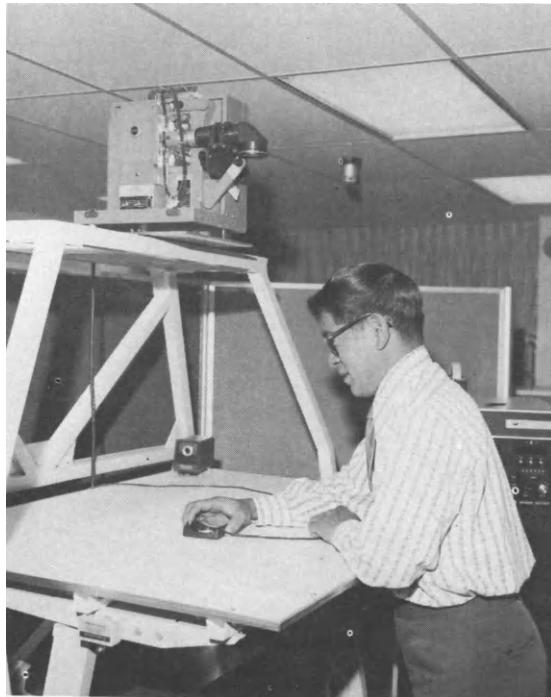


Fig. 11. Wind vector end point extraction station, with loop movie projector.

Cloud Height Estimation

VISSR IR data measures temperature in 256 distinct gray levels in the digitized image data. In cloud height determination, these gray levels are first translated into radiative temperature. Contours are added at 10 K intervals. A typical machine printout of a 17 by 17-inch array in kelvin is shown in Fig. 12. A single cirrus cloud top is shown at the center of the printout; higher temperatures at the printout edges are sea surface temperatures.

The analyst's judgement is required to translate these temperatures into cloud heights. The blackbody temperature of a cloud equals the actual cloud temperature only if the cloud is sufficiently dense to shield the satellite sensor from radiation below the cloud. The analyst only selects clouds that meet this test. The effective cloud temperature is then compared to the estimated lapse rate, and thus the altitude of the cloud is estimated. Figure 13 shows a typical end result with cloud-top heights noted (in thousands of feet) for a number of high-altitude clouds.

Synoptic Judgements and Atmospheric Model Inputs

Wind vector and cloud height determinations represent relatively quantitative data that can be extracted from the SMS cloud images. These images are also used in a more subjective fashion to extract synoptic data and atmospheric model computer inputs for longer range weather predictions now in use at NOAA.

Information of this type is derived from a 12-hour real-time film loop of IR or visible cloud images. Such a loop (in contrast to the 2-hour loop used for wind estimates, which shows clouds hopping back and forth) reveals the general movement of weather over a 12-hour period. Front movements can be followed, jet streams identified, and hour-by-hour location of major weather disturbances located. Figure 14 shows a single IR image, enhanced to emphasize lower temperatures and to accentuate major storm areas. These areas are defined by the white (cold), high-altitude cloud tops, generally ringed by black (warm) lower-altitude regions. On a 12-hour endless loop with this type of enhancement, movement of major storm areas (about 12 to 18 in satellite view on a typical day) can readily be traced. These observations are now fed on a daily basis into weather predictions based on more conventional data. These data are particularly useful to the National Weather Service's location of major storm areas to be avoided on transoceanic flights.

291	289	283	277	283	290	292	293	293	293	293	293	293	293	293	293
290	283	270	274	288	292	292	293	293	293	292	293	293	293	293	293
282	281	280	287	291	292	292	291	290	291	292	292	293	292	293	293
279	288	291	291	292	289	283	283	284	284	286	291	292	292	293	293
274	287	292	292	289	278	266	262	270	275	276	280	285	289	292	293
274	283	288	287	279	265	261	253	249	249	249	257	270	282	288	290
289	287	284	277	269	259	252	247	245	245	244	242	253	270	281	287
293	289	286	280	270	263	256	249	246	245	244	243	249	261	274	284
289	285	279	274	266	262	255	249	244	239	243	246	255	266	278	282
285	284	281	271	262	254	249	246	241	236	242	246	254	269	284	287
274	270	287	285	271	261	251	245	240	235	243	247	255	268	279	285
265	282	288	287	278	264	255	250	248	244	242	244	258	268	272	288
271	285	288	285	278	274	271	259	250	245	244	248	262	271	280	288
279	285	284	282	276	278	271	257	251	253	254	267	269	265	270	283
275	282	285	284	283	279	268	258	261	270	275	279	276	265	260	271
267	281	289	290	291	287	270	274	280	282	279	280	279	266	256	266
274	284	289	291	292	292	287	282	285	285	276	275	278	264	254	265

Fig. 12. Sample 17 by 17-inch temperature array, depicting a single cirrus cloud tracer overlying sea surface.



Fig. 13. IR image with derived cloud heights.

In addition to the foregoing, regions of dry air (black on the image) and moist air (white) are easily located on the film loop. These moisture data, along with the temperature data, are used to supplement and adjust atmospheric model estimates.

Data Dissemination

Reference 7 describes the dissemination of satellite weather data within the United States Meteorological Service. This is done over the internal NOAA data network interconnected with the NESS Central Data Distribution Facility (CDDF), Fig. 15, in the World Weather Building (WWB) in Camp Springs, Maryland (Ref 8).

As indicated in the data network, the SMS satellite transmits all data to the NESS CDA station at Wallops Island, where a 60-foot antenna is dedicated to the SMS. The visible VISSR data are "stretched" at this station and retransmitted in real time via narrow-band links to a 24-foot antenna on the NESS Building in Suitland, Maryland. These



Fig. 14. Enhanced IR image emphasizing cloud tops.

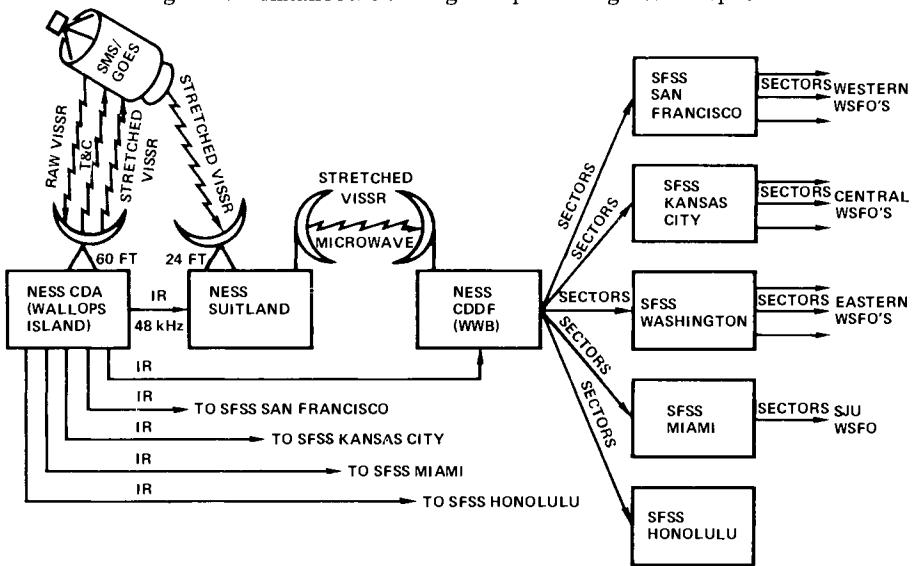


Fig. 15. NESS central data distribution system.

data in turn are transmitted via microwave link to the NESS CDDF, about 2 miles from the Suitland receiver site. Lower resolution IR data are transmitted to NESS-Suitland and other users via 48 kHz telephone lines.

The primary use of these data is by the NESS Operational Processing and Analysis Division (in the WWB), which is responsible for day-to-day centralized processing of data from various spacecraft. This group is responsible for the operational production of satellite-derived winds and synoptic analysis. This requires a photographic "production line" for producing negatives from the electronically digitized data received from the satellite, quality control of these negatives, 20 to 30 prints of each negative at a speed of 30 feet per minute, enlargements, and the 16 millimeter film loops. In this form, the images go to the NESS Operational Processing and Analysis Division for analysis in the manner described previously in this section.

Secondary users of SMS data are the Satellite Field Service Stations (SFSS), shown in Fig. 15, located in Washington, D.C.; Miami, Florida (collocated with the National Hurricane Center); Kansas City, Missouri (collocated with the National Severe Storms Forecast Center); San Francisco, California, and Honolulu, Hawaii (both collocated with the National Weather Service Forecast Office [WSFO]). These groups focus on local weather conditions. They act as the first-line interface with the WSFO's for real-time injection of satellite data into the regional forecasting process. For this purpose they receive, via telephone line, sectorized visible images extracted from full-earth images. Nineteen sectorizers extract the various sectors required by users scattered throughout the United States. Regional SFSS stations retransmit selected sectors to local users in their regions.

Through this network the NESS Operational and Analysis Group in Washington and the Satellite Field Service Stations bring the data to the eventual users, the National Weather Service Forecast Offices. These offices merge the satellite data with conventional weather information received from other sources. NESS analysts at each SFSS location are in ready communication with local forecasters and offer a formal briefing before daily weather forecasts. This interaction is a most effective approach to SMS data use.

As previously stated, regional stations tend to be more interested in local weather conditions than is the National Weather Service in Washington, D.C. Much regional interest therefore focuses on sector images covering 800-mile-square regions. As in the World Weather Building in Washington, D.C., regional NESS analysts who analyze the satellite data are collocated and interface with their counterparts in the regional Satellite Field Service Stations. WSFO use of the data is the same as at the NESS Central Data Distribution Facility at Suitland, but with a local emphasis.

DCP and WEFAX Links

As of June 1975 the data collection platform links had not gone into truly operational use. Only 37 DCP's had been deployed, in comparison with the 10,000 that can be serviced every 6 hours by a single SMS satellite. It will be some time before DCP's are deployed in quantities large enough to realize this potential. The capacity they will eventually possess to collect data from remote places is sizable.

The DCP has a second possible use. It not only transmits data to the central station but can receive teletype interrogation from it. The required DCP antenna is only a 3-foot conical spiral. Use of this link to transmit information to small fishing boats, for instance, offers a broad range of possibilities that have not yet been explored.

The weather facsimile link was put in operational use in June 1975. Through a WEFAX link, stations within view of the satellite will be able to receive cloud image facsimile data, now available only to stations of the internal United States National Weather Service network. A primary WEFAX virtue is that data can be received by small ground stations located throughout the world for the reception of APT pictures from low-orbiting ESSA satellites. A relatively simple conversion on these ground stations, described in Ref. 4, will allow WEFAX reception.

III. SMS DATA ACQUISITION AND PROCESSING SYSTEM FOR REGIONAL USE

The techniques for use of SMS-type data described in the previous sections are those currently in use by NOAA to support the primary function of the National Meteorological Center, the projection of weather trends on a global and national basis. Initial use of the data, therefore, has been to support this function rather than local or regional applications. Furthermore, the techniques used were adapted to the availability of relatively large, existing facilities, likely to be justified only in a central receiving facility. These facilities include photographic processing equipment which, though highly automated, requires a significant number of personnel for operation and maintenance. In addition, some of the data processing requires fairly large computers (IBM 360/195's), which are available and necessary for other purposes, such as large-scale atmospheric modeling.

The above situation has restricted the use of SMS data to primarily large-scale weather phenomena to date. A number of more local SMS applications have not yet found full operational use. These include local and finer-grained weather forecasting, severe storm tracking at shorter intervals than that normal for national forecasting, identification of sea temperatures influencing fishing conditions, local rainfall or snow precipitation, and identification of local clouds most likely to produce precipitation if seeded. For these potential uses to be fully exploited, regional groups with specific interests in these phenomena should have both an onsite data acquisition station for real-time access to SMS data, and simple data processing capability.

Such a station is feasible because of the stretched VISSR data link on the SMS. As mentioned previously, this link transmits in real time -- with no processing delays -- the same image data that now goes to the central Wallops Island facility. However, this stretched VISSR link is a narrower bandwidth, enabling reception by smaller antennas. A 16- to 21-foot antenna is adequate, in contrast to the 60-foot antenna at Wallops Island needed for reception of wideband data. A small antenna also gives real-time access to DCP data (which might be used to monitor local flood conditions) and WEFAX data from the central facility.

Note: Variation in antenna size quoted above is a function of data quality desired (in terms of bit error rate) and the conservatism of design margins. Actual experience with the SMS shows an EIRP of 54 dB at earth edge. This translates into a 15 foot/35.5 dB ground antenna which, with zero margin, will give a 1 in 10^6 bit error rate.

Error rate need varies with intended data use. A 1 in 10^6 rate is needed if the data are to be computer processed by autocorrelation techniques and the like for image matching. However, if it is to be viewed by a human operator on a photograph or CRT display, bit error rates as high as 1 in 10^3 are tolerable. The latter rate can be achieved with a 10-foot/31.5 dB ground antenna.

For regional data processing, self-sufficient equipment, requiring only the analyst to operate, is desirable. The fact that the VISSR image is received from the satellite in digital form makes it readily adaptable to magnetic tape recording and storage. This in turn enables ready input to a small computer for input to a CRT display in either numeric or pictorial form. Processing in this manner eliminates the need for photographic reproduction. Self-sufficient equipment can perform many of the data processing functions now done in the NOAA central facility by the large computers and photoprocessing line. In addition, this simpler equipment offers the advantages of man-machine interaction in analysis and processing, which is not feasible with photographic reproduction techniques.

Data Acquisition Station

The hardware for a small regional data acquisition station exists. Figure 16 shows a typical assembly of the necessary subsystems. Such a data acquisition and handling station, described in Ref. 10, includes:

- a. A 16- to 21-foot antenna with microwave elements. (A typical 21-foot antenna with drive is shown in Fig. 17. A block diagram of the antenna-mounted microwave elements is in Fig. 18.)
- b. Cabinet-mounted receiver/demodulator equipment. (A functional block diagram of such equipment is in Fig. 19. This equipment detects the biphasic shift-keyed VISSR data. The advanced, general purpose telemetry receiver features solid-state design through the use of integrated circuits and subminiature components.)
- c. A data-handling subsystem for bit and frame synchronization; sectorizing, formatting, and recording the data; and monitoring and controlling operations via a CRT/keyboard terminal. (Diagnostic functions are also performed by this equipment. Figure 20 is a functional block diagram of the data-handling subsystem.)

Data Processing Equipment

A data-handling subsystem of the type shown in Fig. 20 records and stores SMS data in such a manner that it can be accessed by minicomputer, processed as required, and displayed to the analyst on a CRT. A number of such image processing facilities exist. Section 11 of Ref. 6 describes a Man-Machine Interactive Processing System in use by NOAA. A Man-Computer Interactive Data Access System (McIDAS), described in Ref. 11, has been assembled by Dr. Suomi's group at the University of Wisconsin Space Science and Engineering Center. A third system, Image Processing and Display (IPAD) equipment (Ref. 12), is in use at the Western Development Laboratories (WDL) of Aeronutronic Ford Corporation. The IPAD system is described below.

All these systems are capable of approximately the same data processing and image enhancement functions and differ only in the equipment and computer used. WDL's IPAD consoles and associated equipment are shown in Fig. 21. Major IPAD subsystems, shown in block diagram form in Fig. 22, perform the following functions.

- a. The computer subsystem controls image handling and input/output, has two seven-track magnetic tape drives, employs a 50 k-word analog-to-digital converter, responds to touch point entry device (TPED) commands, and processes the imagery accordingly. The Honeywell DPP-16 computer used has 12 k-words of 16-bit core memory.

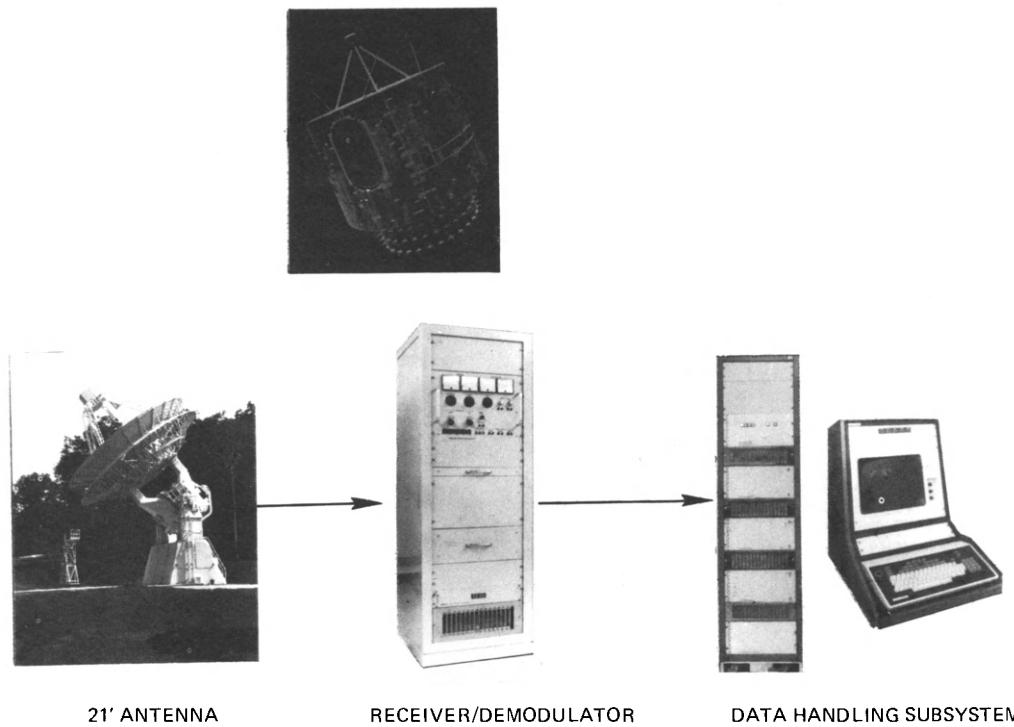


Fig. 16. Data acquisition station.

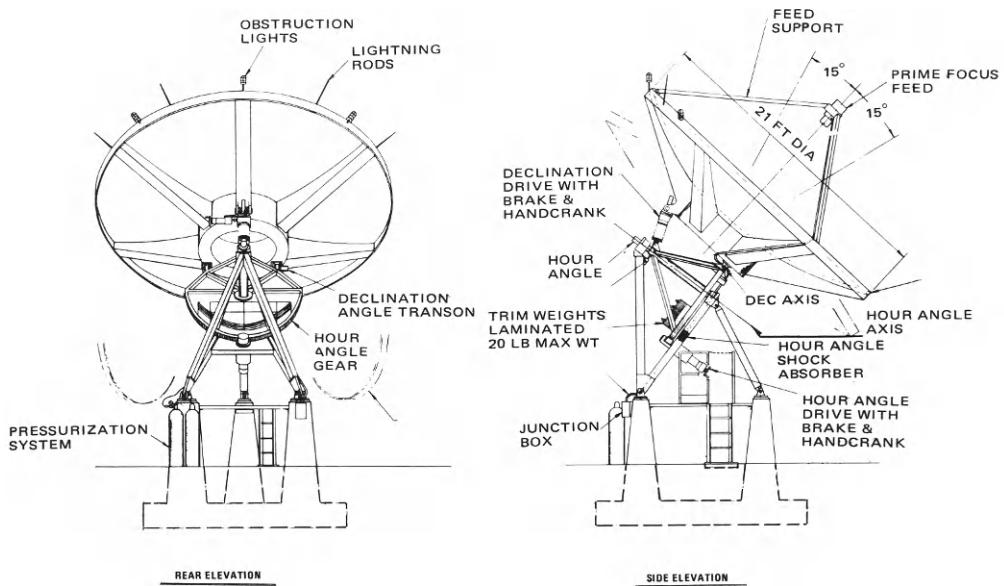


Fig. 17. Typical 21-foot antenna.

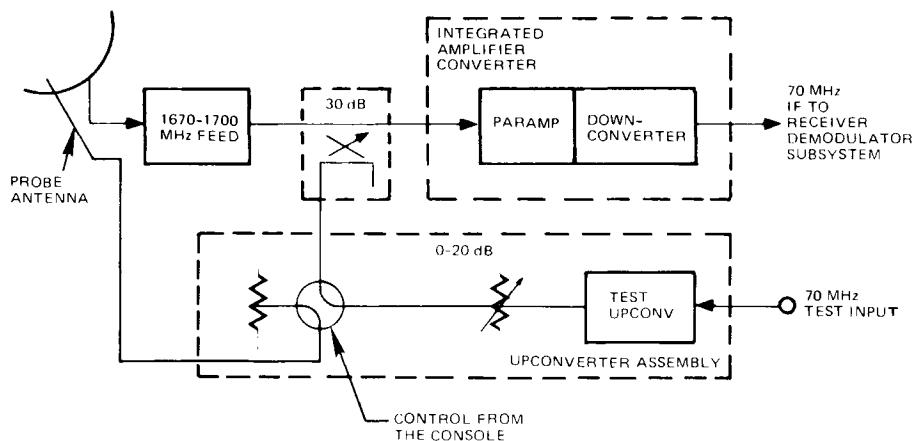


Fig. 18. Antenna-mounted microwave electronics.

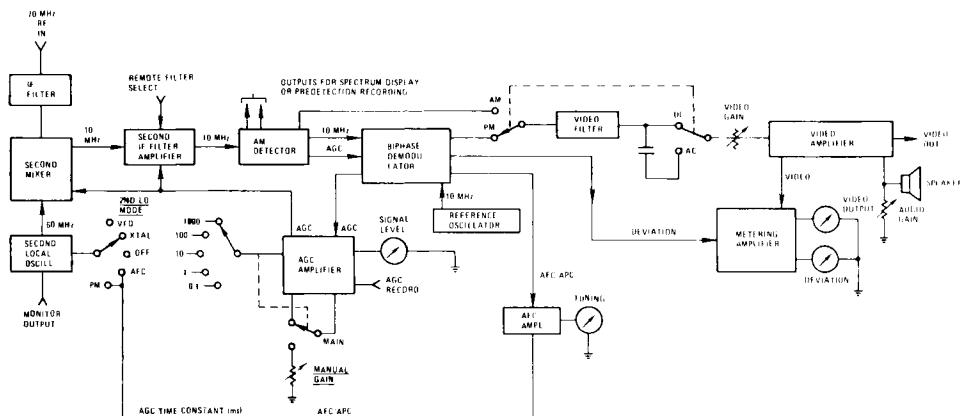


Fig. 19. Receiver/demodulator subsystem functional block diagram.

b. The memory subsystem implements all interface and mode controls from the Central Processing Unit (CPU): life-addressable, pixels, color programs, TPED addresses, and status tests. It synchronizes the CPU to the refreshed images at the rate of 60 frames/second, giving a flicker-free image on the CRT display. This subsystem consists entirely of special devices with high-volume, high-resolution, fast memory capabilities.

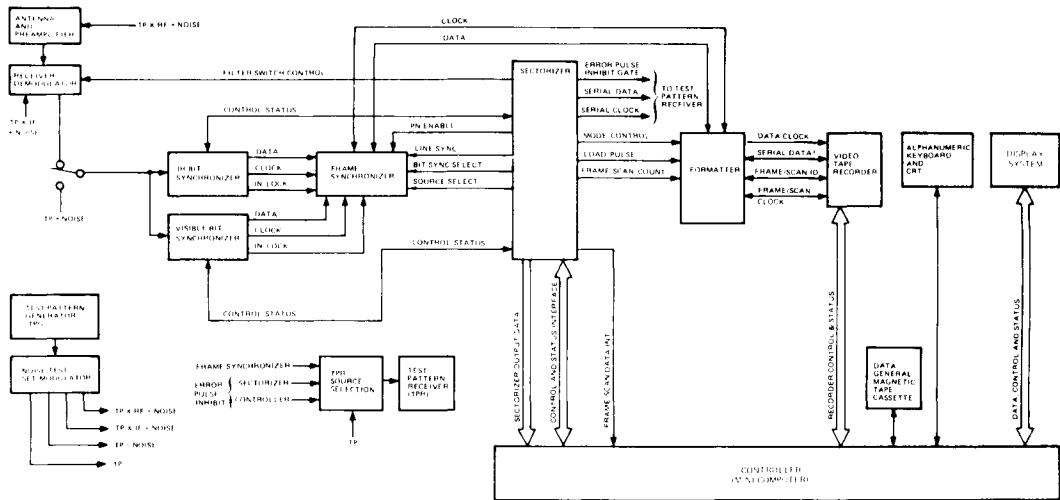


Fig. 20. Data-handling subsystem functional block diagram.

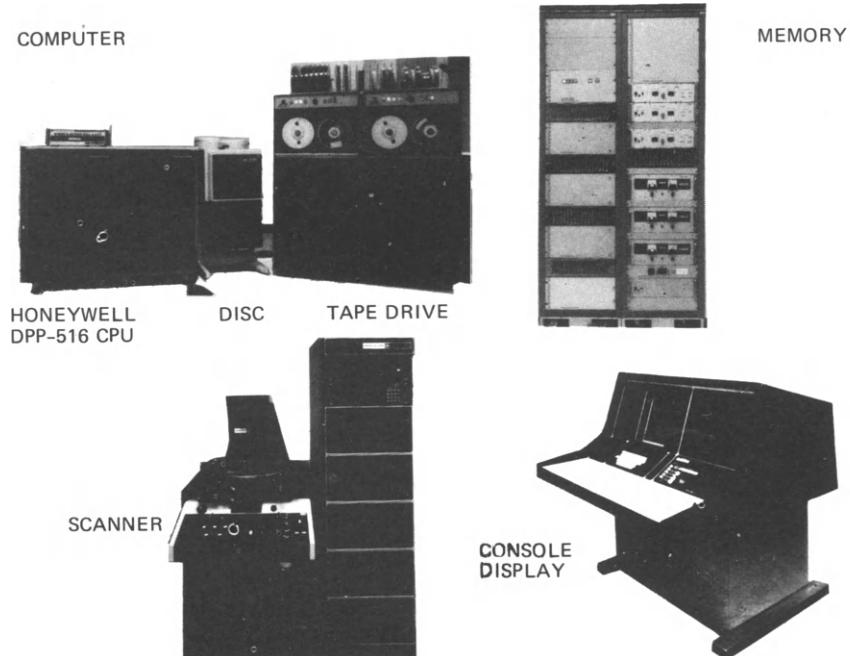


Fig. 21. Image processing and display (IPAD) console and associated equipment.

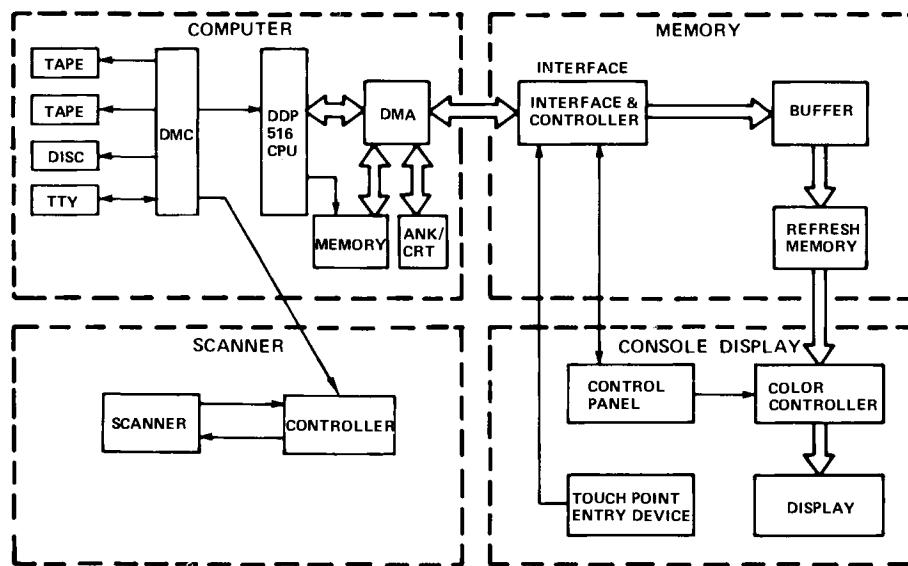


Fig. 22. IPAD Laboratory equipment functional block diagram.

c. The console display subsystem performs parallel-to-serial video conversion; displays imagery on a 1024 line, 12- by 12-inch CRT display; performs gray and color control changes; and commands the interactive functions called for by the operator. A touch point entry device is the main vehicle by which the operator designates or points out any one of the million pixel brightness values.

d. The scanner subsystem digitizes film transparencies at 2 micron resolution.

With IPAD equipment, it is feasible to perform all functions described in Section II which were based on photographic techniques and projections. This includes sequential display of cloud images which gives the same result as film loop projections for estimating cloud movement. It also is feasible to perform multicolor enhancement of such images. Cloud position measurement to determine wind velocity is possible by a process similar, in principle, to the photographic projection of the NOAA unit previously described.

SMS data processing equipment, in addition to its relative simplicity and minimum manpower requirements, has one additional advantage and one disadvantage relative to the photographic/large-computer technique. The disadvantage is the reduction of picture resolution that is inherent in current, commercially available CRT's. The image transmitted by the SMS has 1821 lines in infrared, giving an 8.9 km resolution, and 8 by 1812 lines in the visible, giving an 0.9 km resolution. Thus, any CRT projection with less than 1821 lines degrades the IR image; less than 14,496 lines degrades the visible. Commercially available CRT's in the United States have 512 lines. The WDL IPAD equipment described has 1024 lines. Ongoing developments should make displays available in a year or two with sufficient lines to maintain at least the IR image resolution. (It is noteworthy that the VISSR gives an image resolution not yet fully exploited by either photographic or CRT means. However, full use is made of the resolution and gray scale capabilities in digital computer processing for such uses as cloud temperature determination.)

Data Processing Software

The hub of a self-sufficient data-processing system is the real-time software operating system and attendant software program packages that control the devices, respond to operator commands, and process the data. Dr. Suomi and his associates have pioneered in the development of software for use in conjunction with the McIDAS hardware assembled by the Center. (Development of this type of software requires a very close working relationship between the meteorologist, the programmer, and the hardware designer.) McIDAS also can perform functions similar to those done by the photographic processing techniques in use at the Central Data Distribution Facility of NOAA. These functions include archiving the image in digital tape form for ready access, registering the images on a CRT display, enhancing the images, and extracting wind vector and cloud height information. Reference 13 describes the McIDAS system, its software, and the meteorological basis on which this software is built. The following paragraphs give a brief description of some software equipment discussed in Ref. 13 and demonstrated by Mr. J. T. Young of the University of Wisconsin Space Science and Engineering Center.

Archiving and data access. The digital image data is accepted from the computer and stored on an Ampex analog video disk with capacity for storing 250 TV frames on each of two channels. The operator can select any of the TV frames for display. Further commands enable the operator to focus on any desired sector and to alter enhancement to aid in cloud identification.

Navigation program. The first step in calculating wind vectors (as in the photographic process) is the alignment of a series of satellite images and referencing these images to ground coordinates. The navigation program transforms a satellite coordinate system to an earth reference system, taking into account movement in the earth image due to satellite inclination and a spin axis that may not be perpendicular to the orbit plane. In addition, landmarks are identified and a least-squares procedure used to center the earth in each frame. With this adjustment the location of any given cloud trace in earth coordinates can be read into the computer.

Wind vector and cloud height measurement. To measure wind vectors, the operator selects a cloud target area from two successive pictures. Cloud images in these target areas are compared, and a best match is derived by the computer. The displacement of the two images relative to earth coordinates, by reference to the navigation program, enables computation of the wind vector.

To determine cloud height, a visible SMS image is first used to determine the optical thickness of the cloud, assuming that reflected intensity is a function of cloud thickness. Functional relationships between optical thickness and cloud emissivity in the 11.5 μm band are generated, and the measured emissivity used, along with a measure of fractional cloud cover, to compute blackbody temperature of the radiating cloud surface. Standard atmospheric profiles or available soundings are then incorporated to yield an atmospheric level.

McIDAS software is such that wind vector measurement and cloud height computation are activated by a single command, after designating the cloud target position with a "joystick" on the CRT. The subprocessor stores this position and target size, and the operator repeats his target designation on the next image in the sequence. Wind vector and cloud height are then computed.

Image enhancement. A capability of image-processing facilities, such as McIDAS, that deserves special emphasis is image enhancement to bring out a desired feature. Enhancement can be accomplished by photographic means, but the photoprocessing

steps are not at the fingertip control of the analyst. Thus, the man-machine interaction capability of the McIDAS-type equipment is not available in the photographic processing approach.

Enhancement consists in slicing the image to emphasize particular gray scale levels or gradients. If the feature to be viewed is cloud, as distinguished from water or land masses, slicing can be spread over a 100 K temperature range. In contrast, if water temperatures only are desired, a 10 K range in the higher water temperatures can be selected. This emphasizes water temperatures but loses other features. The Gulf Stream images discussed in Section IV were enhanced in this manner.

Another example of image enhancement is contouring, illustrated in Fig. 23. Figure 23a is an image of a cyclonic disturbance without enhancement; Fig. 23b shows the same image contoured. Contouring is accomplished by enhancing in the manner shown in Fig. 23c; a sawtooth cycle is set into the computer so that the full black-to-white (or color) range is repeated three times to cover three major temperature ranges. By this means, the central cumulus is enhanced, while retaining full cloud pattern definition. A white (cold) cloud top is surrounded by a lower darker (warmer) island (Fig. 23d). Lower clouds go through another black-to-white range and are also emphasized, while water or land masses fall primarily in the third range. Contouring can be done by photographic means, but it would still need the analyst's prior selection of slicing levels. On the computer-CRT display system, this is done by the turn of a knob; the enhancement can thus be more readily optimized.

The above-described data storage, access, display, and processing techniques enable SMS data to be processed in near real time by equipment practical for use in regional weather stations. Many types of analysis of interest for local users, other than those described, are feasible. Reference 14 describes some additional uses developed by the University of Wisconsin group. For example, the simpler equipment has been used to enhance Earth Resources Test Satellite (ERTS) images and to derive land-use classifications from them. It can enhance an SMS IR image to show warmer and colder ocean waters, locating large areas such as the Gulf Stream (of interest for ocean navigation) and smaller areas of interest to fishermen. Rainfall estimates have also been made.

IV. FUTURE DEVELOPMENTS

A current progress report on SMS system development points up that the full potential of this system is still to be realized. System developments that can be anticipated are discussed in the following paragraphs.

It will be noted that minimum mention has been made of DCP links. At this time, only 37 DCP's have been deployed in comparison with the 10,000 that can be serviced every 6 hours through the SMS communication link. Experience will build up as more platforms go into use. It is noteworthy that the DCP link not only transmits data to a central facility but also (by way of its interrogation link) can communicate with remote stations. Teletype communication (75 b/s) can be achieved with a 3-foot conical spiral antenna (10 dB gain) and a simple receiver at the DCP end. A small fishing boat, which otherwise might be out of communication with land stations, could receive storm warnings, local weather, or sea temperature information on the likely location of fish by use of such simple equipment.

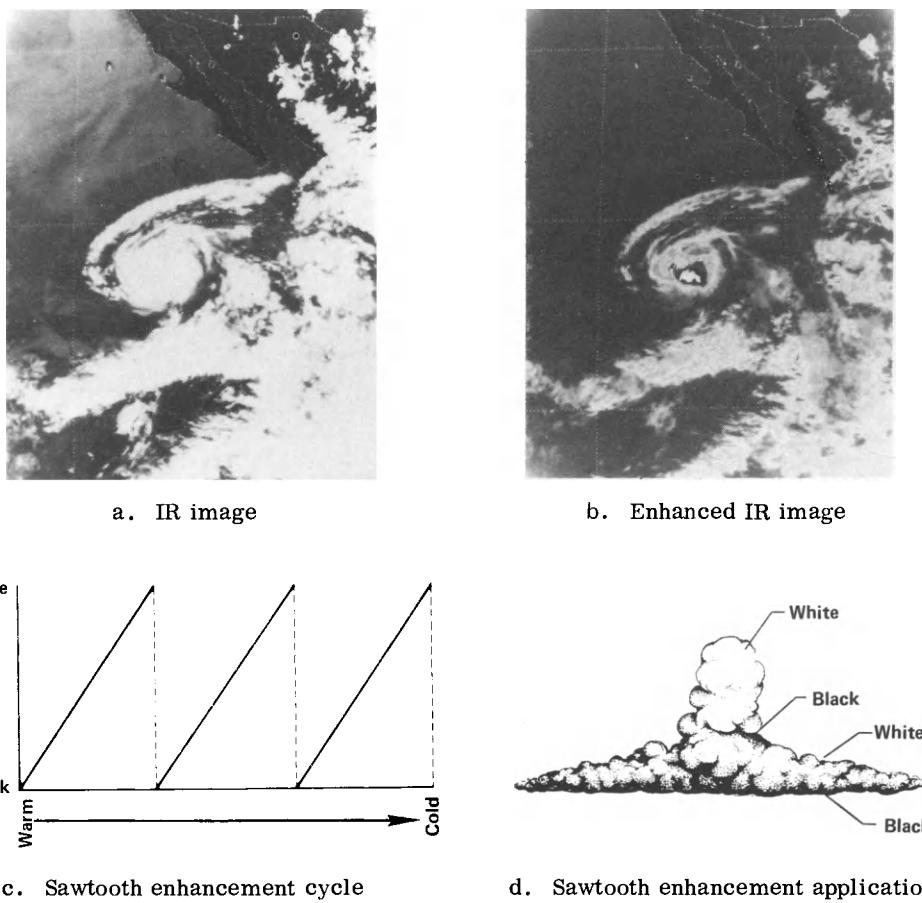


Fig. 23. Contouring technique.

The WEFAZ transmission link was about to be put in operational use at the time this paper was prepared. The fact that existing APT stations used in conjunction with lower orbiting satellites can be readily adapted for the WEFAZ link (Ref 4) should encourage its operational use.

Plans are already underway for giving the SMS satellite a vertical sounding capability. This will be done on a modified VISSR, to be called the vertical atmospheric sounder (VAS). The optical chain will remain essentially the same as on the VISSR. However, in place of the single IR sensor used for the 10.5 to 12.6 μm band on the VISSR, there will be six infrared detectors with filters giving the following spectral bands on the VAS: two window bands (the VISSR 11.2 μm band and an added 3.7 μm band); two water vapor bands (6.7 and 7.2 μm) and eight carbon dioxide bands (seven from 12.7 through 14.7 μm and one at 4.3 μm). The window and carbon dioxide bands penetrate at varying depths in the atmosphere and can therefore be used to derive a vertical temperature profile. The water vapor bands yield information on the water vapor content of the sample atmospheric column.

The most significant SMS system development anticipated, without any change in the satellite, is the general use of small data acquisition and processing stations. A number of these are planned in the United States for special purpose users. Perhaps of more importance is the potential use by areas within view of the satellite, but not within the United States meteorological ground network, the NWS. With stretched visual and IR images available on a real-time basis, and with the use of the man-machine processing equipment described herein, such regional areas should be able to benefit from the SMS system at minimum cost. The necessary hardware and software are already available. One desirable technological development would be the larger, increased resolution displays discussed above. These are needed if full SMS system capability is to be realized.

Available research literature gives an indication of expanded applications that can be feasibly employed by regional users. Their more local interests will give an impetus to such added uses. Following are some examples of likely applications.

Hurricane, thunderstorm, and tornado tracking. Reference 15 discusses use of SMS-type cloud images for identification of intensities and wind velocities in tropical cyclones and hurricanes. Successive images of a developing hurricane, such as that in Fig. 9, can be analyzed by the wind vector measuring technique previously discussed to measure the wind vectors around the core of the hurricane. These vectors can be used to calculate the vorticity and rate of change of vorticity as the hurricane develops. These two characteristics are indicators of hurricane intensity and thus are of interest in tracking a hurricane.

The half-hour frequency of SMS images is useful in tracking more rapidly developing, shorter-life weather disturbances. Exploratory use is being made of this "near-real-time" observation of thunderstorms in the Chesapeake Bay area in support of "NOW-CASTING" by the Washington Field Services group of NESS. Hourly or half-hourly tracking of these thunderstorms is used for a short-time forecast of the location of these storms.

The Severe Storm Warning Center in Kansas City, Missouri also has an interest in exploring the use of SMS images for tornado tracking. In view of a tornado's rapid development and short lifetime, it is desirable to have successive cloud images at intervals shorter than the half-hour normally programmed for SMS. This is possible when the north/south camera scan is reduced by ground command. SMS has been commanded into this shorter mode of operation for a period long enough to permit forecasters to determine the utility of tornado-tracking data. Although this use is still in the research phase, an insight is being obtained into the mechanism of tornado development from a line squall. If this application reaches operational use, it will pose a ground problem of assimilating quantities of data of an order of magnitude more than is now processed operationally by the SMS system. This special purpose might require an added satellite in orbit, since it conflicts with normal SMS use. (The Severe Storm Satellite is being studied by NASA for this dedicated purpose.) Further development of quick-reaction data processing will be necessary to fully realize this potential.

Snowstorm forecasting. Reference 16 reports on studies made by airplane, surface, and more recently by low-orbiting satellite, observations of developing snowstorms along the east coast of the United States in an attempt to forecast location and time of snow precipitation. Figure 24 (from Ref. 16) shows the typical cloud pattern from which there is precipitation (hatched area). Forecasting requires a prediction of the time when the dry slot reaches the center of the cloud vortex. The frequent cloud images now available from SMS and the ability to observe nighttime clouds will greatly

improve the capability of studying developing cloud patterns during the critical hours prior to storm maturity. This is another potential "local" use for regional stations with real-time data acquisition and processing capabilities.

Sea surface temperature. Radiated energy in the 10.5 to 12.5 μm waveband indicates sea surface temperature with an accuracy of 1.5° to 2°C when the surface is viewed through a cloud-free atmosphere at an incidence angle less than 60° (Ref. 17).

Figure 25 shows an enhanced image of the Gulf Stream taken under these conditions.

The slicing range of this image is narrowed to emphasize the 10°C difference between the Gulf Stream and the surrounding ocean. The Gulf Stream, as defined by water temperature, is readily discernible. Since strong currents are associated with the boundaries of the stream, this information is useful in the navigation of oceangoing vessels. Exxon Corporation utilizes a daily SMS Gulf Stream analysis to route both its north- and south-bound oil tankers, realizing a significant fuel savings.

Sea surface temperature is also an indicator of the likely location of fish. The San Francisco Satellite Field Services Station has taken the lead in testing the use of satellite data for this purpose and to determine what is necessary to apply the information to operational conditions. This experiment, conducted during the 1975 fishing season is described in Ref. 18, from which the following is summarized. Existing studies indicate that albacore (a species of tuna) and salmon are "temperature sensitive". Albacore prefer temperatures in the range of 14.4° to 18.3°C; salmon prefer an 11.1° to 13.3° range. Both species tend to concentrate along "thermal fronts", i.e., the line of demarcation between warmer and colder waters where the temperature gradient is 0.5°C per nautical mile or greater. Figure 26a (Ref. 18) is an enhanced IR image showing water temperature variation in the geographical test area extending from San Francisco Bay to the mouth of the Columbia River, and from the California coastline to 130° W longitude. This image was obtained under cloudfree conditions. The black-to-white shading represents a water temperature variation of somewhat less than 10°C.

Figure 26b shows the thermal front lines derived from Fig. 26a, along which the gradient is 0.5°C per mile or greater. The location of these fronts is influenced by a number of factors. They involve fixed conditions such as ocean bottom contours, seasonal effects such as coastal upwelling during the summer due to prevailing northerly winds, and the movement of successive weather fronts through the area. At present, this type of plot is made available on a weekly basis to fishing fleets out of Eureka, California. The use of this front information in an operational system would require transmission over radio networks, or via the satellite, on a daily basis.

This use of SMS data is now in the experimental stage. Feedback from the fishing fleets to evaluate the method is being collected by the Humboldt (California) State Marine Advisory Extension Service in cooperation with the NESS San Francisco Satellite Field Service Station.

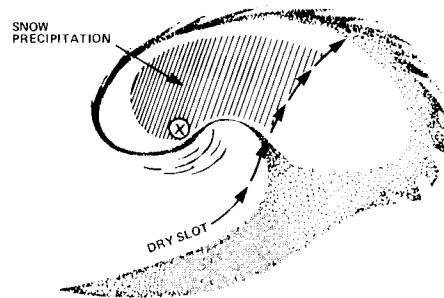


Fig. 24. Location of precipitation in a cyclone.

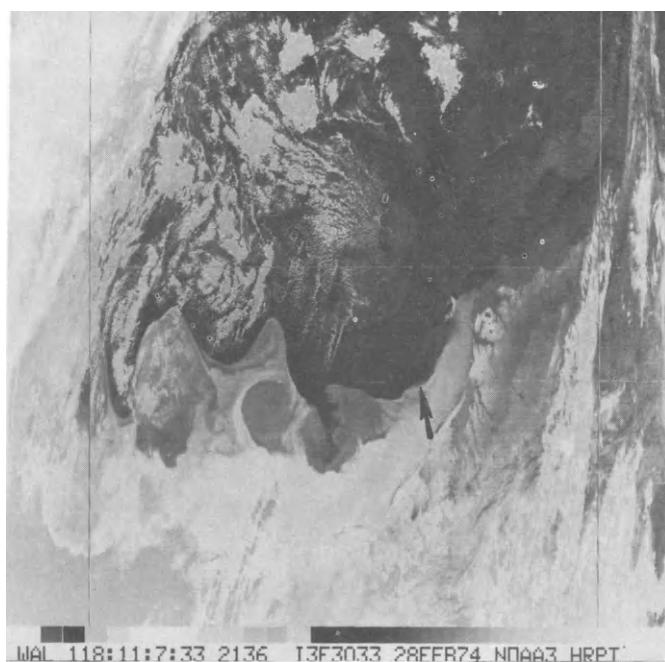
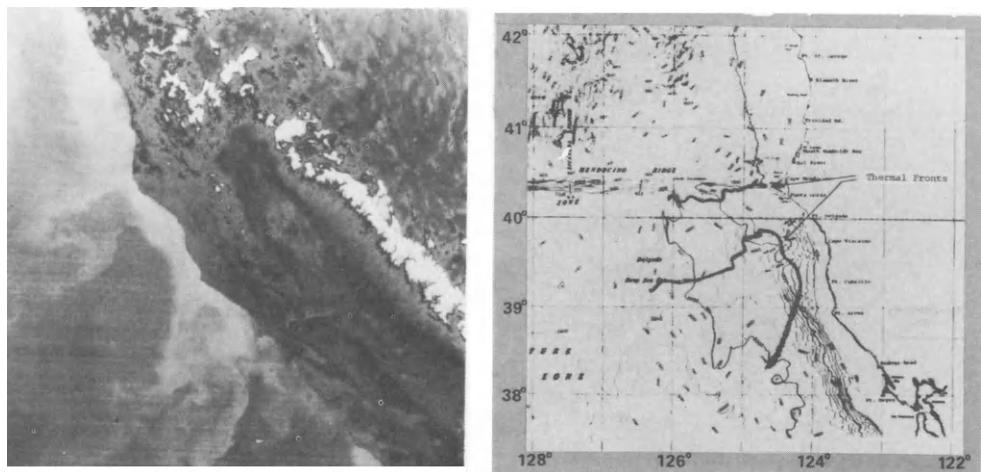


Fig. 25. Enhanced IR image of Gulf Stream.



a. IR image of test area.

b. Thermal fronts derived from IR image.

Fig. 26. Use of sea-surface thermal pattern to indicate favorable fishing areas.

Rainfall estimates. Experiments have been conducted (Refs 19 and 20) on the feasibility of estimating daily rainfall from satellite cloud images. In Ref. 19, daily rainfall estimates were made by estimating the percentage of cumulonimbus cloud cover in a region. Studies were made from visible image cloud observations made at 3 pm local time and from infrared images obtained at 9 pm local time. Seven-day totals thus estimated for Alabama, Georgia, and South Carolina, during July, August, and September 1973, were compared with corresponding ground measurements. The agreement was of sufficient accuracy to compute crop yield. Earlier tests are reported in Ref. 20 for regions in Zambia, Thailand, California, Florida, and south central United States. Results in these areas are also encouraging. It is anticipated that empirical modifications will be required for differing regions, such as mountainous dry regions or valleys. A regional station could "calibrate" its local area of interest. Such estimates should be of value in areas of low-population density where runoff estimates are important.

Earth resources information. The Earth Resources Satellite (Landsat) and the lower-altitude, polar-orbiting weather satellites have the advantage of higher resolution, more spectral bands, and/or polar coverage needed for many earth resource observations. However, there are uses (snow melt, depth and extent; sea temperature; sea-ice location; definition of water/land boundaries) where the lower resolution of a geosynchronous satellite is acceptable, and the higher observation frequency may be preferable. SMS frequency of observation may be as desirable as the higher resolution of lower orbiting satellites. (Would you rather miss catching a fish by 8 km or 8 hours?) Potential earth resource uses will increase when the additional spectral bands on the VAS are available. For example, the 0.4 to 4.7 μm band will measure ocean surface color and indicate plankton concentrations that attract larger fish.

Although no useful near-polar images can be obtained by SMS, data collection platforms can be serviced by SMS as far poleward as 70° to 75° latitude and longitudinally to about 35° either side of the satellite. On this basis, the SMS system can be used as backup for more specialized satellite systems, with no added investment in ground acquisition or data processing equipment. Data of this type will represent a compromise from that obtainable from a lower orbiting satellite with specialized sensors. However, these data would be a zero-cost fringe benefit from the SMS system.

The thought expressed in the Introduction deserves repetition here. The adaptation of a new application satellite system to the needs and procedures of the eventual user is a highly innovative, experimental process. It requires the merging of diverse technology and scientific disciplines. The SMS story typifies this, and judging from future potential, this story is just beginning.

ACKNOWLEDGMENTS

A survey paper of this type is based on the work of many people. As emphasized in the paper, the success of the SMS system required a broad spectrum of technical skills supplied by a number of groups. Typical of the weather satellite program in the United States, this has been a team effort led by the NASA Goddard Space Flight Center and the National Environmental Satellite Service.

Credit is given to associates at the WDL Division of Aeronutronic Ford for their contribution to a major portion of the SMS system. They were involved in the design and manufacture of the satellite, the design of the data acquisition and handling station, and the image processing and display equipment described herein. The length of the list of contributors to this phase of the work and to the paper precludes

individual mention. Their credit line will emanate from the continued successful operation of satellites in orbit.

Acknowledgment is also extended to Messrs. W. J. Hussey, M. J. Young, and F. E. Bittner of NESS. The description of the use of SMS data is based on their work and that of their co-workers. Discussions with them and a tour of the Central Data Distribution Facility was of major benefit. The help of J. Bottoms of the NESS San Francisco Satellite Field Service Station is also acknowledged. His assistance generated an understanding of regional station operation and a number of potential data applications discussed in the paper. In addition, Figs. 10a, 10b, 10c, 11, and 12, from Ref. 8, were furnished by the U.S. Department of Commerce/National Oceanic and Atmospheric Administration.

The section of the paper dealing with the McIDAS could be written only with the help of Dr. Suomi and his associates at the Space Science and Engineering Center of the University of Wisconsin. The data processing technique that this group pioneered is expected to be a major factor in exploiting SMS data and expanding its availability to additional users.

APPENDIX I. SYNCHRONOUS METEOROLOGICAL SATELLITE SUBSYSTEM CHARACTERISTICS

Communications

The Communication Subsystem of the SMS performs the following functions:

- Image formatting
- Image transmission (including VISSR and WEFA)
- DCP interrogation and reporting

<u>Characteristics</u>	<u>S-band</u>	<u>UHF</u>
Receive frequency	2030 MHz	402 MHz
Transmit frequency	1680 MHz	469 MHz
Effective radiated power (on earth center)	56.0/35.6 dBm	43.5/49.5 dBm
Receive figure of merit G/T (on earth center)	-24.4 dB/ $^{\circ}$ K	-23.8 dB/ $^{\circ}$ K
Bit rates: Raw VISSR	28.0 Mbps (max)	
Stretched VISSR	1.75 Mbps	
Bit error rate (max)	10^{-6} at 14.5 dB E_b/N_o	

Telemetry, Tracking and Command (TT&C)

The TT&C function is performed at both S-band and VHF frequencies. It is fully redundant with turn-around ranging capability compatible with Goddard Range and Range Rate System. The VHF antenna provides complete spherical coverage.

Characteristics

VHF receive frequency	148 MHz
VHF transmit frequency	136 MHz
Transmit power	2 watt/8 watt
Receiver noise figure	2.5 dB
Receive antenna gain	-6.5 dB min
Transmit antenna gain	-5.0 dB min

Attitude Determination and Control (ADAC)

This subsystem processes data for determination of spin axis orientation, spin rate, and satellite timing signals. Active nutation control is used during transfer orbit operation only for unstable inertia ratios. This gives a maximum nutation of 0.35 degree in automatic mode. A real-time sun-synchronous timing pulse provides accurate line-to-line registration of VISSR images.

Characteristics

Attitude determination	0.1° capability with the on-board sensors; landmarks on the visible images are used for more precise determination
Attitude stability	0.5 arc sec max
Antenna pointing	Anywhere within $\pm 11^{\circ}$ in E/W relative to sub-satellite point. Resolution $\pm 0.176^{\circ}$
Passive damping	Time constant 3 minute max Damping residual 0.5 arc sec max

Visible and Infrared Spin Scan Radiometer (VISSR)

The VISSR provides the satellite imaging capability in both the visible and infrared spectra. The spinning motion of the satellite provides line scan in the west-to-east direction. North-south tracing is accomplished within the VISSR.

<u>Characteristics</u>	<u>Visible</u>	<u>Infrared</u>
Spectrum	0.55 to 0.75 μm	10.5 to 12.6 μm
Number of detector channels	8	2
IGFOV each channel	0.021 x 0.025 mrad	0.25 x 0.25 mrad
Ground resolution (nadir)	0.9 km (0.5 nmi)	8.9 km (4.85 nmi)
Data bandwidth per channel	210 kHz	26 kHz
Number of scan lines	1821	
Frame time	18.2 min, 100 rpm	
Retrace time	1.71 min	
Frame size	20° x 20°	
Passive cooled detectors (Hg-Cd-Te) noise equivalent radiance	0.9 x 10 ⁻⁵ watts cm ⁻² sterad ⁻¹	
Effective aperture	40.64 cm (16 in)	

Space Environmental Monitor (SEM)

The SEM detects unusual solar activity such as sun flares, which have a high level of radiation; intensity of solar winds; and strength and direction of the earth's magnetic field, which affects global weather patterns. It consists of a magnetometer assembly, X-ray instrumentation, and an energetic particle sensor.

- X-ray Sensor - Measures the intensity of solar X-rays in two wavelength bands: 0.5 - 3 Å and 1 - 8 Å
- Energetic Particle Sensor - Determines the intensity of charged particle flux in 7 proton energy ranges, 6 alpha particle ranges, and 1 electron energy range
- Magnetometer - Monitors magnitude and direction of ambient magnetic field, parallel field ($\pm 1200\gamma$), and transverse field in 4 selectable ranges ($\pm 50\gamma$, $\pm 100\gamma$, $\pm 200\gamma$, or $\pm 400\gamma$)

REFERENCE

1. D. V. Fordyce, R. J. Wirth, and W. E. Shenk, "The Synchronous Meteorological Satellite (SMS) System," Sixth Conference on Aerospace and Aeronautical Meteorology, El Paso, Texas (1974).
2. P. T. Burr and F. B. Pipkin, "The Synchronous Meteorological Satellite (SMS) System," NASA/Goddard Space Flight Center, Greenbelt, Maryland.
3. S. A. Ames, R. J. Martel, and B. R. Perkins, "Data flow in the Synchronous Meteorological Satellite System," NASA/Goddard Space Flight Center, Greenbelt, Maryland.
4. J. J. Nagle, "A method of converting the SMS GOES WEFAX frequency (1691 MHz) to the existing APT/WEFAX frequency (137 MHz)", COM-74-11294/AS.
5. Philco-Ford (now Aeronutronic Ford), "SMS/GOES--Introducing a New Era in Global Weather Observation" (1974).
6. C. L. Bristor, Editor, Central processing and analysis of geostationary satellite data, NOAA Tech. Memo. NESS 64 (1975).
7. W. J. Hussey, "The Geostationary Environmental Satellite System," National Environmental Satellite Service, Suitland, Maryland, EASCON (1974).
8. E. P. Weigel, The World Weather Building, NOAA, V, 1 (1975).
9. L. F. Whitney, Jr., "Effect of orbital inclination and spin axis attitude on Wind Estimates from Photographs by Geosynchronous Satellites," in NOAA Tech. Memo. NESS 41 (1972).
10. Aeronutronic Ford Corporation, "GOES Acquisition/Data Handling System," WDL-TP6050, Palo Alto, California (1975).
11. V. E. Suomi, "Man-Computer Interactive Data Access System (McIDAS): Continued Development of McIDAS and Operation in the GARP Atlantic Tropical Experiment," University of Wisconsin-Madison (1975).
12. Aeronutronic Ford Corporation, "IPAD (Image Processing and Display) Laboratory Operating Description," WDL-B807, Palo Alto, California, (1974).
13. E. A. Smith, "The McIdas System," IEEE Trans. on Geoscience Electronics, GE-13, 3 (1975).
14. G. C. Chatters and V. E. Suomi, "The Application of McIdas", IEEE Trans. on Geoscience Electronics, GE-13, 3, New York (1975).
15. C. O. Erickson, "Use of Geostationary Satellite Cloud Vectors to Estimate Tropical Cyclone Intensity," NOAA Tech. Memo. NESS 59, (1974).
16. F. C. Parameter, "Use of Satellite Data in East Coast Snowstorm Forecasting," NOAA Tech. Memo. Ness 33 (1972).
17. C. Braun, "Limits on the Accuracy of Infrared Radiation Measurements of Sea-Surface Temperature from a Satellite," NOAA Tech. Memo. NESS 30 (1971).

18. L. Breaker and F. Jurick, "Providing Near-Real-Time Sea Surface Temperatures to the California North Coast Fishing Fleet: An Approach," National Environmental Satellite Service, NOAA, California (1975).
19. W. A. Breaker and V. J. Oliver, "A Comparison of Infrared Imagery and Video Pictures in the Estimation of Daily Rainfall from Satellite Data," NOAA Tech. Memo. NESS 62 (1975).
20. W. A. Follansbee, "Estimation of Average Daily Rainfall from Satellite Cloud Photographs," NOAA Tech. Memo. Ness 44 (1973).

SARLAB PROJECT
PRELIMINARY IDEAS ON SIDE LOOKING RADAR SYSTEMS
FOR SPACELAB EARTH OBSERVATION PAYLOADS

J. P. Guignard*, Thomson and H. Mooney**

*CSF, France **BAC, England

ABSTRACT

A preliminary study on side looking radar systems for Spacelab (ref. 4) Earth observation payloads (SARLAB study) was carried out under ESRO contract. This short paper is aimed at reporting the purpose and the main results.

The aim of this study was to examine the use of Spacelab to act as a bridge between the present and planned experimental Earth Resources aircraft programmes and an eventual operational automatic satellite (such as the proposed SARSAT (x)).

After presenting the scope of the Sarlab study, emphasis is made on the new concepts introduced with respect to the earlier Sarsat study.

They are mainly dealing with :

- the beam orientation
- the on-board preprocessing
- the antenna technology
- the monopulse systems

Apart from the above points, the technical developments are briefly summarized. The main characteristics and the basis for the selection of the preferred options are presented, leading to a final Sarlab preferred option.

The conclusions are presented in terms of :

- main SARLAB concept based on Spacelab capabilities
- applicability of the SARLAB concepts to the SARSAT design
- validity of using Spacelab in view of the SARSAT development.

SARLAB STUDY PRESENTATION

After demonstrating the feasibility of an operational side looking radar satellite programme (SARSAT project, ref. (1)) it is considered desirable to have an experimental phase. This experimental activity has two main and interrelated objectives.

These are:- a technology objective to test, evaluate and qualify the radar sensors in orbit prior to their inclusion in operational automatic satellites or as standard facility sensors in manned space missions.

- a geoscientific objective aimed at evaluating the application of

(x) Synthetic Aperture Radar Satellite - See ref. 1

radar sensors for earth observation purposes and developing measurement methods and techniques for use in later operational systems.

Bearing in mind these objectives, it turns out worthwhile to examine the use of Spacelab to act as a bridge between the present and planned experimental aircraft programmes and the eventual operational automatic satellites, as the proposed SARSAT.

Therefore a preliminary study on side looking radar systems for Spacelab Earth observation payloads (SARLAB study) was carried out from February 1974 to May 1974, under ESRO contract SC/24/73 HQ by THOMSON-CSF (Prime Contractor) and the British Aircraft Corporation.

GUIDELINES AND DESIGN PHILOSOPHY

The starting point of the SARLAB study was the SARSAT baseline as presented at the Ninth International Symposium on Remote Sensing (April 1974 - see ref. 1) and characterized by the following features

spacecraft altitude	565 km
radar resolution cell	25 x 25 m ²
swath width	77,6 km
RF power	151 W
scattering coefficient	- 20 dB (design value)
antenna length	8,09 m
antenna width	0,85 m
polarisation	single
beam steering	none
SNR after azimuth processing	6 dB

With respect to the user's requirements, the main limitations deal with :

- only X - Band (instead of X and L - Bands)
- only one polarisation
- no beam steering

The main difficulties and critical areas are related to :

- the antenna (technology and manufacturing)
- the antenna deployment mechanism
- the data processing and data storage
- the attitude control system

Based on this SARSAT project, the main purpose of the SARLAB study was to examine the technical implications of using Spacelab in the area of imaging systems, and the impact these implications may have on the performance and possible design of the Spacelab and Shuttle systems.

Initially the radar sensor will be a relatively simple equipment ("basic S.L.R. System"), evolving in time to a more complex system ("advanced S.L.R. system") as our experience and knowledge of the capabilities of these sensors grow.

Basic SLR System

Side looking radar systems are expected to form a part of the early Spacelab

missions, either as a sensor in an Earth resources Survey Facility Payload, or as the main sensor in a smaller earth resources "add-on" payload forming a part of a multidisciplinary Spacelab payload.

During these early missions, the radar sensor is expected to be a relatively simple, single frequency synthetic aperture system similar to the SARSAT baseline option as far as the performance is concerned.

Advanced SLR System

Experiments using multipolarization, multifrequency radars are expected to follow from the initial experimental phase using the basic system.

In order to cope with the above and overcome the SARSAT critical areas (by allowing for Spacelab capabilities) particular attention was paid to the following concepts :

- beam orientation from the attitude control system
- on-board data preprocessing
- light structure type antenna and beam steering
- monopulse systems (azimuth and slant range directions)

These concepts are expanded in the following sections.

SARLAB MAIN CONCEPTS (REF. 2)

Beam orientation from attitude control system information

The Doppler effect (due to the earth's rotation) is corrected (SARSAT baseline) by using two antenna channels (rough correction) and three phase lock loops (fine correction).

However because the target rotates during the measurement, the measurement duration has to be short enough to prevent any mixing of the slant ranges. Then only a very limited onboard preprocessing is possible with the following drawbacks :- high data bit rate (about 50 Mb/sec.)
- limited storage capability, etc...

The idea, as far as the SARLAB study is concerned, is to take advantage of a gimballed platform, used to support a light structure antenna, to achieve a suitable beam orientation.

The main condition to be satisfied during the radar measurement is the orthogonality between the beam orientation vector and the velocity vector (relative velocity Spacelab/target). A detailed study demonstrated the possibility of implementing an attitude control processor to compute the beam vector components (from attitude information, position, time, earth rotation, etc..) and to carry out the beam orientation by using a three-axis gimballed platform.

Onboard preprocessing

Using a L - meter long antenna, the theoretical achievable resolution is $L/2$ meters when using a focused synthetic aperture antenna.

Based on the previous considerations (beam orientation), an opportunity exists of taking advantage of the difference between the required resolution (e.g. 25 or 50 m) and the achievable resolution (e.g. 4 m) to reduce the data

bit rate and the storage requirement.

The preprocessing is shown to be equivalent to a low pass filtering. Such a preprocessing is carried out for each of the slant range channels. To reduce the sidelobe level a two-step filtering is used on two channels (sine and cosine channels) for further recovery of the radar signal phase.

Light structure and beam steering

Allowing for the Shuttle bay dimensions, a light structure type antenna is adopted consisting of :

- a parabolic cylinder reflector
- a line of waveguides as feed line (see fig. 1)

Two feed waveguides are used to get the suitable beam shaping (as required to reject the radar ambiguities). Each of the waveguides consists of various sections to achieve the required bandwidth (25 MHz), compared to the waveguide antenna structure (e.g. SARSAT antenna), the electrical efficiency is slightly reduced (0,5 dB in gain).

A dual polarisation antenna concept is shown in figure 2, using a second feed set (cross polarisation) and a second reflector (grid) to maintain the focal length.

To overcome the relationship between the swath width and the depression angle an additional feed (2 waveguides) could be implemented to allow a variation of the beamwidth (2 possible values).

For the sake of simplicity it turns out desirable to consider either dual polarisation capability or beam shaping, the feasibility of an antenna offering both possibilities being not considered as demonstrated.

Monopulse systems (azimuth and slant range)

An azimuth monopulse system is used to center the Doppler frequency spectrum of the radar return as required to allow further processing (correlation).

On the other hand, any altitude variation is shown to have a drastic impact on the radar design : the target to be imaged has to be kept within the range between two ambiguity zones.

In case of altitude variation, the target can be continuously centered by modifying the pulse repetition frequency.

The preferred system to modify the PRF is shown to be a slant range monopulse system.

One antenna difference channel is added leading to an error signal proportional to the angle between the beam axis and the direction dealing with the centre of the recurrence period.

SARLAB OPTIONS

Definition of the various proposed configurations allows for the statement of the typical support services to be provided to payloads by the Spacelab.

As far as the orbits are concerned, it is presumed that, as with SARSAT, mission users will require two types of orbit i.e.

- orbits having a zero drift swathing pattern

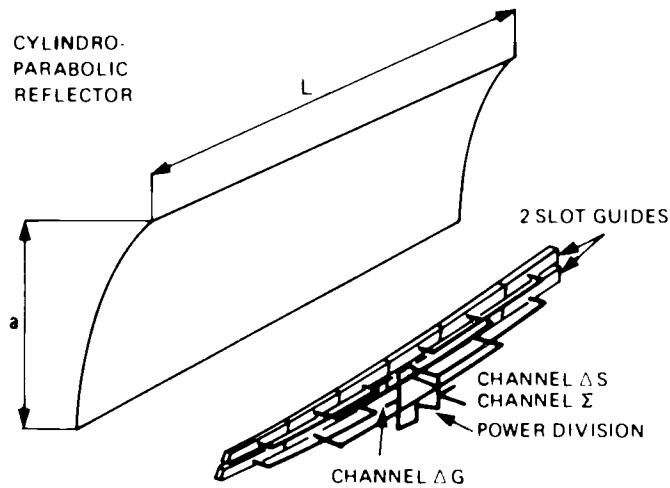


FIG. 1 - SINGLE POLARISATION ANTENNA

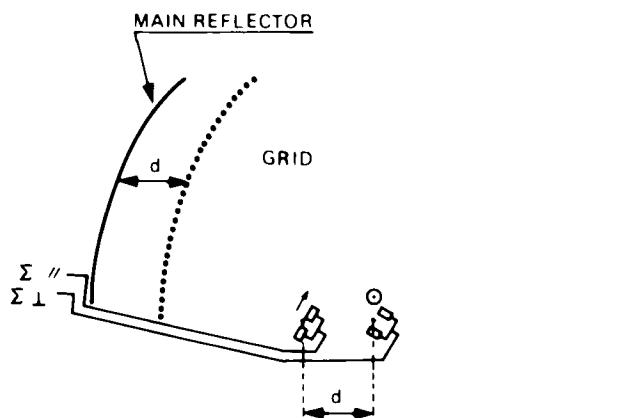


FIG. 2 DUAL POLARISATION ANTENNA CONCEPT

- orbits giving a minimum drift swathing pattern

In order that the radar's ground swath can intercept, in a desired way a set of ground truth sites, it may be necessary to alter the phase of the swathing pattern.

Basic SLR payloads definition (e.g. option 1 - fig. 3)

Earth resources payloads and "add-on" payloads and associated pressurized boxes resulted in the following :

- limiting the antenna length according to pallet module length
- preventing any development effort as far as antenna and data processing are concerned
- using existing airborne optical imaging chain

As a matter of fact, all basic options are using existing optical recording and storage unit (airborne THOMSON-CSF unit). After a detailed study of all the possible optical imaging chains, the following processing is selected :

on board : flying spot + optics
 film storage

on ground : optical correlation

Laser recording, electron beam recording etc... have been discarded mainly because of the required development programme.

Desirable improvements

Desirable improvements of the basic payloads mainly deal with :

- improving performance
- improving the system flexibility (e.g. beam steering with constant swath width, dual polarisation antenna, focused and unfocused modes, X and L bands radars).

In general the possibility exists to use the preprocessing unit to achieve an unfocused mode processing (real time) and to take advantage of such a device :

- when flying over the oceans, with a poor resolution (e.g. 500 m) but allowing measurement with a scattering coefficient of - 30 dB (instead of - 20)
- when flying over the ground, for onboard monitoring (real time display with 250 or 500 m resolution)

After defining a focused aperture system, it turns out :

- possible to use such a radar in unfocused mode with same swath width
- not possible to increase the swath width when using the unfocused mode (ambiguity problems)
- not possible to work in real aperture mode (typically 4 km resolution but with wider swath width with the same radar equipments.

Advanced SLR performances (e.g. options 2 to 4, fig. 3)

Based on the above, advanced SLR systems are defined for both earth resources and add-on payloads for later Spacelab missions.

All advanced options use digital imaging chain (for the sake of flexibility) with :

OPTION	1	2	3	4	REMARKS
ALTITUDE (km)	185	370	370	370	
RESOLUTION CELL (m ²)	50 X 50	50 X 50	50 X 50	50 X 50	
FREQUENCY BAND	X	X	X	X,L	
SWATH WIDTH (km)	28,5	28,5	75	98	
DEPRESSION ANGLE (DEG)	55	55	40-70	40-70	
GIMBALED PLATFORM	*	*	**	**	* beam orientation only ** beam orientation + beam steering
IMAGING CAPABILITY	12300 km PER FILM REEL	10000 PER TAPE	10000 PER TAPE	10000 PER TAPE	
SNR AFTER PROCESSING (dB)	6	6	6	6	
POLARISATION	1	2	2	2	X - BAND L - BAND

RF POWER (W)	56	112	112	127 127	X - BAND L - BAND
ANTENNA LENGTH (m) WIDTH (m)	3,05 0,78	3,10 1,53 } X2	9,06 0,51	12,35 0,37 12,35 3,11 }	X - BAND L - BAND
SLR MASS (kg)	132	173	176	199	
SLR CONSUMPTION (W-DC)	1395	2400	2485	2995	400 Hz 3 φ
OPT. REC. UNIT	1	-	-	-	
V.T.R. UNIT	-	2	2	4	20 MHz 8 ANDW

FIG. 3 SARLAT OPTIONS ; MAIN FEATURES

- on board preprocessing (presumming) and digital recording (video tape recorder)
- ground based digital correlator

Advanced "add-on" payloads consist of pressurized boxes.

Four preferred options are selected :

- option 1 : the simplest one ("add-on" option)
- option 2 : the main SARLAB option
- option 3 : an intermediate option
- option 4 : a very sophisticated one

These options are characterized in figure 2.

As far as the radar system is concerned, the 370 km altitude has been considered as a worst case condition.

A typical antenna layout is given in figure 4 (for option 2)

It must be noticed that :

- the antenna should be mounted along the shuttle's fore and aft line
- the antenna should preferably be mounted on a pedestal. This pedestal can be moved on the pallet to positions which suit the accomodation requirements of adjacent pallets.
- the various resources offered by Spacelab (i.e. data, power, etc) are expected to be entirely adequate for SARLAB.

For the sake of argument, we can remark that :

- option 1 is based on the use of optical means as far as on-board recording and ground processing are concerned (however, presumming is carried out in digital form).
- option 2 to 4, deal with a completely digital imaging chain, featuring a first step towards the SARSAT imaging chain.
- option 2 provides significant geoscientific potential (dual polarization system) based on the use of a simple technology (2 antennas of 3 m length each)
- option 3 requires a more extended antenna technology study (9 m long antenna with dual polarization capability).
- option 4 is mainly consisting of twice the option 3 radars, respectively at X-band and L-band.

Finally, as far as the overall programme is concerned, option 2 is recommended as the most promising initial phase of such a programme.

The overall programme for option 2 is depicted in figure 5.

CONCLUSIONS

These conclusions are based on the following considerations :

- main SARLAB concepts with respect to the Spacelab capabilities
- applicability of the SARLAB concepts to the SARSAT design
- the validity of using Spacelab in view of the SARSAT development.

Main Sarlab concepts based on Spacelab capabilities.

In order to understand how far the SARLAB design is depending on the Spacelab capabilities, the following reasoning is required :

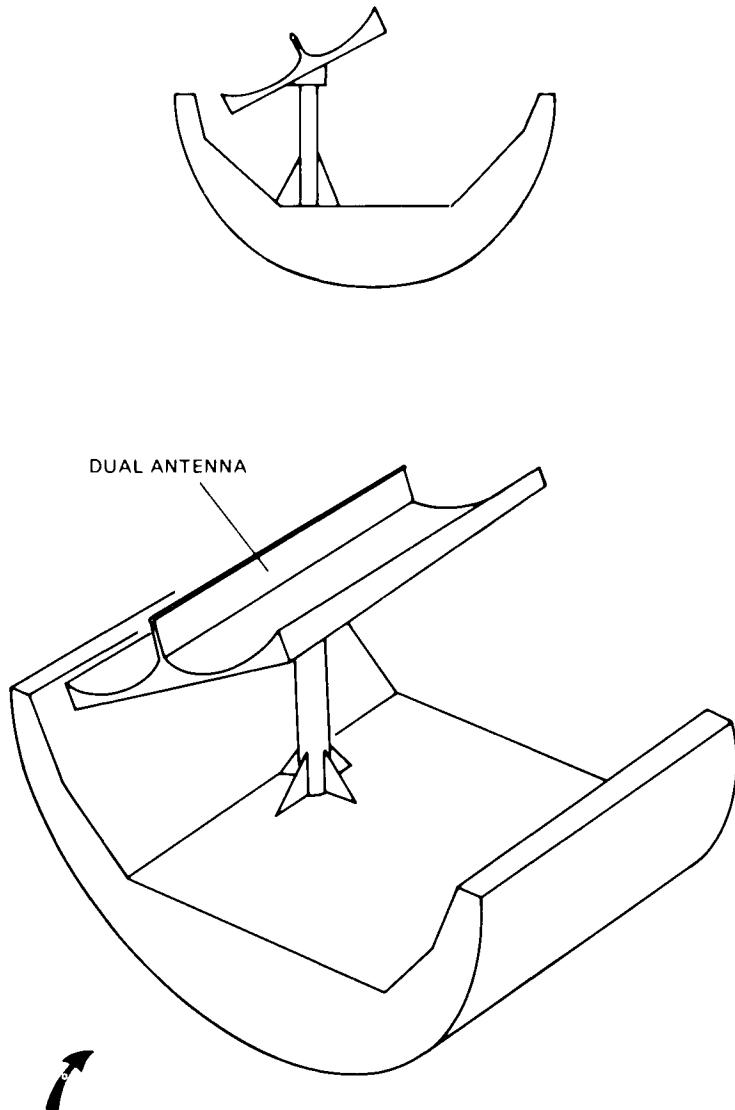


FIG. 4 RELATIVES SIZES OF ANTENNA AND 3-M LONG PALLET MODULE

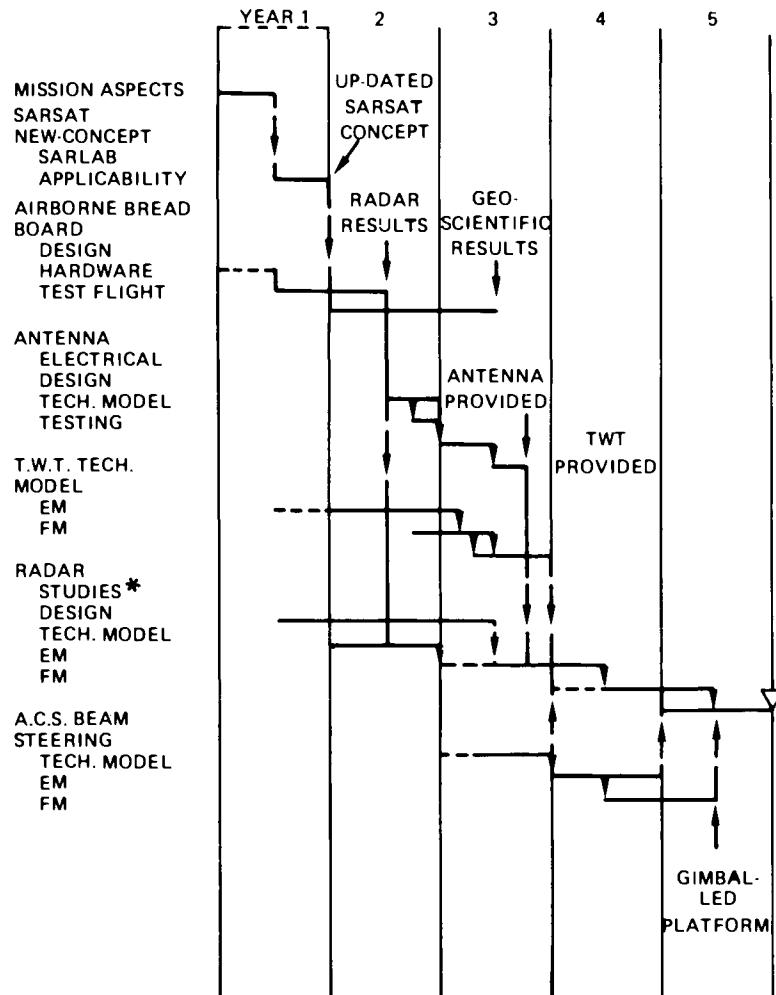


FIG. 5 OVERALL PROGRAMME FOR OPTION 2

a) Due to the lack of significant and frozen Spacelab specifications, the SARLAB design is aimed at being compatible with the Spacelab capabilities irrespective of possible later Spacelab constraints.

For example : - as far as the power supply is concerned, the radar design isn't affected in case of using either fuel cells (SARLAB) or solar panels (SARSAT).

- as far as the digital storage capability is concerned, the SARLAB radar is provided with its own video tape recorders. Then if the opportunity arises (depending on the overall Earth Resources Payload Definition) to use the Spacelab recorders, the Sarlab radar design is compatible to this adoption.

b) As far as the radar itself is concerned, the defined subsystems are valid for both SARSAT and SARLAB concepts (e.g. low noise parametric amplifier, T.W.T., etc...) The main difference is relevant to the qualification level.

c) The main modification with respect to the initial SARSAT design (see ref. 3) is the beam orientation capability, and the consequent advantage of decreased data bit rate by using an onboard preprocessing technique. The adoption of such concept is justified below :

- the radar design is not dependent on the antenna technology. Whatever be the project (SARSAT or SARLAB) a study is required on antenna technology (e.g. dual polarization antenna).
- the monopulse systems are aimed at coping up with both beam steering implications and attitude variation. Such an improvement is also desirable in the case of SARSAT.
- consequent possible adoption of existing tape recorder (20 MHz bandwidth) for SARLAB is not contradictory to the use of improved tape recorders (100 MHz bandwidth) in case of non-implementation of a SARSAT presumming unit.
- achieving of beam orientation by using a gimballed platform is then the only critical area as far as the SARLAB-SARSAT concept is concerned. This point is discussed in the next paragraph.

By summing up these points, it can be stated that :

- most of the SARLAB radar design concepts are relevant to the SARSAT radar design
- the main difference is in the SARLAB beam orientation concept.
- the applicability as such capability for SARSAT has to be demonstrated (see next paragraph).

Applicability of the SARLAB concepts to the SARSAT design

As previously explained the only critical area when comparing the SARLAB design is consisting of :

- the mechanical beam steering
- the beam orientation from onboard computation
- the antenna technology

All the other aspects are either common (e.g. RX, TX, etc...) or relevant to the above (e.g. preprocessing unit).

It yields a first important conclusion :

- apart from the above points, all the SARLAB concepts are common enough to the SARSAT concept to be directly applicable.

During the SARSAT study, two kinds of systems have been defined (among six options) : - radar systems relevant to a Thor Delta launch
- radar systems relevant to the Shuttle (used as a launcher).

Therefore, the second aspect was not developed to the level of the first one. Use of Thor Delta for launching, being the main constraint in the Sarsat study.

Before taking any final decision, the Shuttle launch concept is to be extended in the light of the SARLAB study results (e.g. focused/unfocused and real aperture modes).

Such a recommended study would likely lead to a more refined SARSAT concept dealing with :

- either an experimental phase using a performance limited Spacecraft launched by a Thor Delta.
- or an experimental phase using Spacelab and compatible with a SARSAT operational system launched by the Shuttle.

It must be borne in mind that an operational system as the proposed SARSAT would not require the same level of flexibility (e.g. beam steering) as an experimental system.

Further the development of an experimental system with high flexibility does not preclude the development of a further operational system launched by an expendable vehicle.

Validity of using Spacelab in view of the SARSAT development

The previous chapter was aimed at discussing the compatibility of the SARSAT/SARLAB concepts as far as the performances are concerned (e.g. beam steering)

Development Programme

As established during the study, most of the required SARLAB developments are understood as preliminary steps of the SARSAT developments. This is valid, for instance as far as following items are concerned :

- radar T.W.T.
- radar electronics circuits

The implications of using Spacelab have more impact on the spacecraft aspect, e.g. : - non deployable antenna
- limited structure study
- less severe environmental constraints, etc...
leading to significant cost savings.

As a final conclusion, we would like to stress on the following argument :

- the results of the SARSAT study (preliminary feasibility study) cannot be considered as final and frozen results.
- most of the SARSAT limitations could be solved through more detailed studies. The SARLAB study has demonstrated the ability of achieving such an aim, mainly due to the intrinsic side looking radar concept flexibility. (Ref. 3)

BIBLIOGRAPHY

(1) A preliminary feasibility study on a synthetic Aperture Radar Satellite (SARSAT) for earth resources surveys. Proceedings of the ninth international symposium on remote sensing of environment

15-19 April 1974, Volume III - Ann Arbor Michigan

(2) Preliminary feasibility study of a side looking radar satellite for earth resources surveys.

SARSAT Project under ESRO contract SC/24/73/HQ

Final Report Volume II October 1973

(3) Preliminary study on side looking radar systems for Spacelab earth observation Payloads

SARLAB Project under ESRO Contract 1834/72 PP

Final Report May 1974

(4) SPACELAB

An orbital laboratory for Science, Applications and Technology

European contributions to :

Applications Summer study, National Academy of Engineering, Snowmass, Colorado, 30th June, 13th July 1974

COLLECTION AND ANALYSIS OF SPECTRAL REFLECTANCE DATA AND ITS USE IN THE DESIGN OF A MULTIBAND PHOTOGRAPHIC SYSTEM

R. J. Drewett

Plessey Radar Research Centre, Eastleigh Road, Havant, Hampshire PO9 2RE

ABSTRACT

The basic concepts of a complete multiband photographic system are discussed, stressing the importance of spectral reflectance data in the choosing of camera filters, and the need for high quality output material. A system to provide these needs has been developed at the Plessey Radar Research Centre, and the elements of the system are described. The paper concludes with a section on current and potential applications in the field of remote sensing.

INTRODUCTION

Over the past five years the Plessey Radar Research Centre has been involved in research which has required investigation of various remote sensing systems. In particular, the use of a multiband photographic system has been studied resulting in the evolution of an operational capability. The main purpose of this paper is to describe the design and development of the system, with particular reference to the theory and basic assumptions which it is felt are vital to the production of useful imagery.

Without a basic knowledge of target signatures it is most unlikely that a useful set of filters will be chosen for the system; further, unless careful consideration is given to the optical equipment the quality of the final hard copy output will be poor and thus cancel any extra information which the system could potentially provide. In many cases neither of the above considerations are taken seriously, with the result that potential users of the system become rapidly disenchanted.

AIMS OF THE SYSTEM

Before describing the design approach, it is useful to define the main aims of a multiband system as follows:

- A. To improve upon the information content of available imagery, either colour or panchromatic.
- B. To maintain image quality to a level which does not reduce the information content of the image, taking into account each particular application.

Obviously the level of improvement which justifies the use of a more expensive system will be a matter of individual choice. A criticism often levelled at the multiband system is that it very rarely makes visible that which was invisible on conventional images; no-one would argue with such a statement - nor would it necessarily be accepted as a valid criticism. The

fundamental theory of photographic sensing shows clearly that only in very special circumstances could the invisible be made visible. Consider the following example - suppose that two targets requiring to be separated on the image had the spectral signatures of Fig. 1, and that a colour film used to record the targets had three emulsion sensitivities $f(B)$, $f(G)$ and $f(R)$ as shown. The final density difference between the targets TA and TB on each emulsion layer of the colour film would then be proportional to the mean reflectance difference within the emulsion sensitivity band. More precisely, this difference would be expressed as an integral; for example, in the blue sensitive emulsion the density difference between the targets would be proportional to:

$$\frac{1}{\lambda_2 - \lambda_1} \int_{\lambda_1}^{\lambda_2} f(B) \cdot \{f_R(TA) - f_R(TB)\} d\lambda (= I_B)$$

where λ_1 and λ_2 would be 0.4μ and 0.5μ , and $f_R(TA)$ and $f_R(TB)$ represent the reflectance functions of each target varying with wavelength. It can be shown mathematically, and it is also clear intuitively, that this mean or integral value will be in general less than the maximum reflectance difference in the blue spectral region (at μ_a in Fig. 1), but greater than zero except in the special case where $I_B = 0$. This latter situation is approximated when the mean difference in reflectance within the blue spectral region is zero.

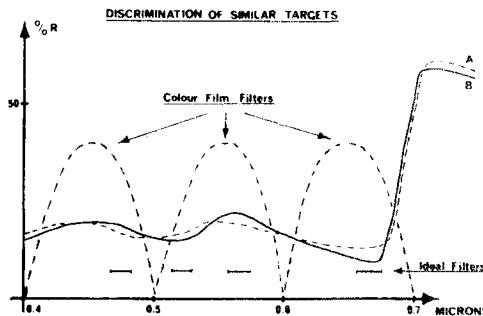


Fig. 1. Discrimination of similar targets

To summarise, the above analysis shows that where two targets with similar but not identical reflectance spectra are imaged onto colour film, the contrast obtained will be less than that which could be achieved if optimum sensing bands were chosen, and in some cases the contrast may be very small or even zero. In the example of Fig. 1, using film sensitive at the wavelengths μ_a , μ_b and μ_c , and with a sensitive bandwidth less than, say 0.03μ (30 nanometres), would ensure maximum target contrast.

Thus, to restate the first aim of the multiband system, it is to improve upon the performance of conventional film by choosing sensitive bands which maximise the final image contrast. This in turn depends on a study of the

reflectance spectra of the targets under a representative range of environmental conditions.

The second aim, although obvious, is often not achieved in practice with the result that much or all of the increased information is negated by poor image quality. The multiband system depends upon producing a colour image from up to four separate panchromatic images, and this requires a high standard of equipment throughout with very well matched optics. Both the cameras and the additive viewer used to recombine the images must be closely matched between optical channels. The viewer in particular must be capable of accurate register and easy operation, and the production of high resolution hard copy (prints or transparencies). This latter requirement is very rarely considered, except in the most expensive commercial instruments.

SPECTRAL REFLECTANCE MEASUREMENT

The first stage in the design of a multiband photographic system is to gain a good understanding of the reflectance spectra of all the targets of interest, under the conditions in which photographs will be taken. The problem of measuring spectra, and then relating these to the photographic appearance, is a very complex one. Variables such as angle of view, illumination type and direction, and atmospheric effects can drastically alter the apparent reflectance of a target, especially with vegetation and other coarse textured materials. This makes the relationship between spectra and photograph very uncertain, and spectra must therefore be treated with some caution. When comparing spectra from different locations, two main factors are important. Firstly, a stable and repeatable reference must be used; secondly, careful record of all relevant conditions is essential so that variations can be allowed for in the analysis. The extent to which spectral changes due to genuine reflectance differences can be separated from those due to changing conditions will depend on the severity of the latter, and on the amount of data available. Since this implies a large set of spectra, the use of computer storage, display and processing techniques becomes a vital part of the system.

Instrument Design

A remote reflectometer operating in the wavelength range 0.2μ to 2.4μ was built during 1972, and is shown in block diagram form in Fig. 2. In standard form the instrument scans six ground targets, one of which is a reference panel, and records the radiance at each of 128 or 256 wavelengths within the range. Standard scans are $0.2\mu - 1.2\mu$ and $0.4\mu - 2.4\mu$ using two gratings in the monochromator. The wavelength scan is driven by a stepping motor which increments the wavelength between target scans. Resolution at each measurement may be varied between 0.5 nanometres and 6 nanometres (1 micron = 1000 nanometres = $10,000 \text{ \AA}$). The instrument is normally operated on a trailer roof (Fig. 3), or from a tower, and data is sampled and recorded digitally for later computer input. Changes in target size, spacing and number to suit a given application require only simple modifications to the scanner.

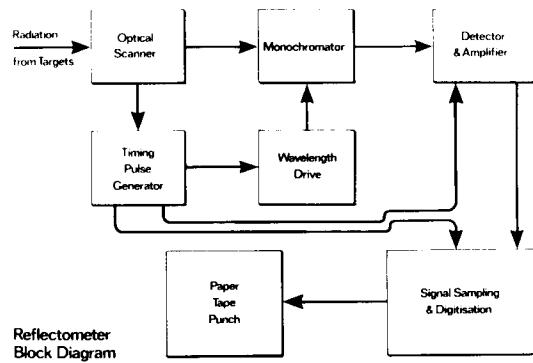


Fig. 2. Reflectometer block diagram

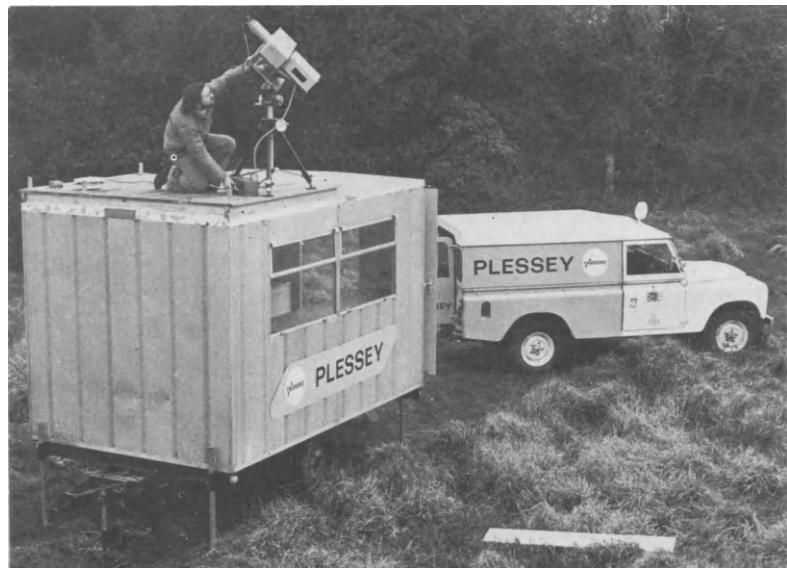


Fig. 3. The reflectometer in use

Each target measurement is accompanied by a reference measurement made within a few seconds, and the complete scan may take between two and ten minutes depending on the detail required. Any variations in illumination intensity during the run are thus easily normalised at the display stage. The reference panel is coated with a stable white paint (barium sulphate) whose reflectance has been accurately measured at the National Physical Laboratory.

Data Gathering and Storage

Spectral data is recorded on paper or magnetic tape, and stored on disc in the computer. During the run, measurements of sun angles, intensity and quality of illumination, cloud cover and other relevant variables are made and entered into the computer in coded form as a 'label' with the spectral data. In addition each spectral set is numbered for identification. Display programmes include percent reflectance (the reflectance reference panel, of known and stable reflectance, is compared to the target at each wavelength), reflectance difference, and mean value curves of a selected set. Sets of spectra are selected by number, or by specifying a limited set of conditions - the programme then selects all runs conforming to those conditions. Examples of percent reflectance, reflectance difference and mean value curves are shown in Fig. 4, (taken from the storage tube display).

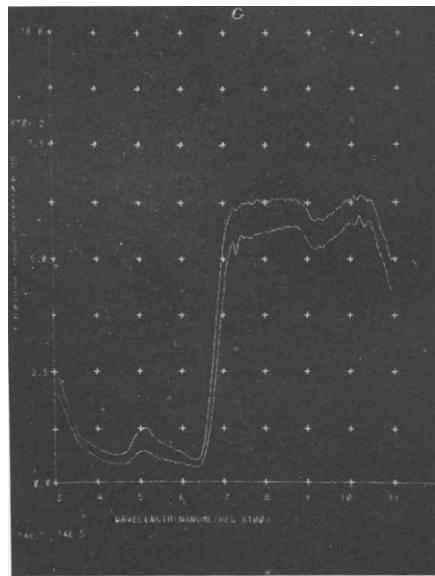


Fig. 4a. Percent reflectance plots

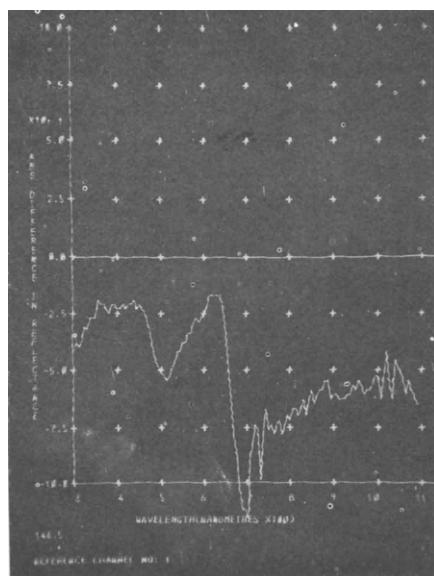


Fig. 4b. Reflectance difference plot

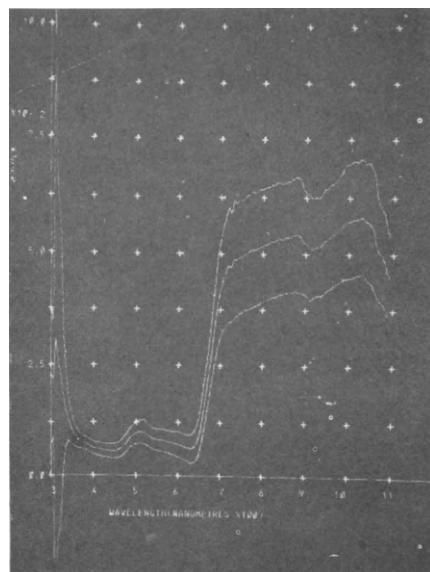


Fig. 4c. Mean and deviation curves for a set of spectra

Data Analysis

Using the display programmes the operator can rapidly view the spectra from target classes of interest, and in many cases establish where significant differences exist between them. Usually maximum separation between spectra will be found at several different wavelengths and over varying bandwidths about those wavelengths, and this will give a good guide to the filters required for the multiband camera system. In most cases, however, the variability of the spectra within a class will be significant and a more reliable prediction of performance can be obtained by discriminant analysis.

Selected wavebands, or features, are evaluated by plotting the spectra according to their feature values in an n-dimensional space; n is the number of features chosen, each representing an axis or dimension in the space. If the features chosen are significant, the spectral points will group or cluster according to class; the separability of the clusters then determines how well the spectra can be classified using the chosen features. This analysis therefore allows an optimum feature set to be developed - the aim being to achieve maximum separability with the smallest and simplest feature set. In many cases a suitable set of three or four features can be found, and these are then used to design the filters in the photographic system.

In all cases the data analysis will assist greatly in the design of an operational imaging system by predicting performance and by indicating the number and complexity of the sensing channels required. It will also indicate whether it is possible to achieve the performance required of the imaging system without the need for full-scale trials and inspired guesswork in the choice of spectral bands.

THE PHOTOGRAPHIC SYSTEM

Assuming that a suitable set of three or four filters has been indicated by the spectral analysis, a multiband photographic system is then required, having the following elements - camera unit, aircraft (or tower for some studies), processing facilities and an additive viewer. A description of the system recently developed at the Plessey Radar Research Centre is given below.

Cameras and Platform

A set of four discrete cameras was chosen in preference to a multilens camera, in order to maintain flexibility in the choice of films; for example the use of one colour or false colour film and three filtered panchromatic films gives the opportunity for comparison of the multiband image with a conventional film. The camera initially chosen was the Hasselblad 500 EL/M, mounted in a four-in-line array. One problem in the use of some 70mm format cameras stems from the nature of the filters used. These are often narrow-band interference filters which are required to optimise detection of pigment absorption effects (most pigments such as chlorophyll, carotene, have absorption bandwidths as low as 20nm). These filters are photographically very 'slow', and require very fast film and relatively long shutter openings; many survey and reconnaissance cameras have a lowest speed of 1/1000 sec., which can be difficult.

More recently a set of four Vinten 360 reconnaissance cameras has been acquired. These are fitted with automatic exposure control, and specially modified shutters to achieve extended exposure times; the Vinten camera is specifically designed for air survey, and is able to take long lengths of film more easily than the Hasselblad. In general, the Hasselblad array would be used for small experimental exercises, and the Vinten array for larger areas.

To date, the cameras have been mounted on a Wallis Autogyro for experimental work (Fig. 5), and this has proved very successful over small and medium areas - say up to 30 square miles. The aircraft has considerable flexibility, and also very low vibration. It operates close to the site, requiring only a short takeoff and landing area; several flights over the area using different filter sets can be made very simply, and image quality is excellent even at shutter speeds of 1/250 sec. Typical image ground resolution is 2" per 1000 ft. altitude, which is very acceptable considering the very fast film used (e.g. Tri-X uprated to 2000 ASA, or 2485 at 5000 ASA). Accurate line and stereo imagery can be flown to a standard which is quite satisfactory for most work, except that for mapping. For larger operations a survey aircraft has been used with the cameras replacing the normal 9" x 9" camera.



Fig. 5. Wallis autogyro fitted with four Hasselblad cameras

Film and Print Processing

In order to ensure that negatives are suitable for additive viewing and printing it is necessary to maintain close control over the film development. In particular, the overall density of the image and its gamma should be optimised and made as consistent as possible. The problem is complicated by the use of narrowband filters and the consequent difficulty in accurately estimating exposure; this has been mainly overcome with the Hasselblad cameras by taking exposure meter readings through each filter from a standard reflecting card and relating the camera settings to the readings by careful tests. The Vinten cameras, with automatic exposure control, cause no problems in this respect.

For good results in viewing and printing each negative should have a low mean density, but without loss of detail in the darkets areas of the scene, and the gamma should be between 2.0 and 3.0 depending on the exact situation. The main factors which determine the optimum gamma are scene brightness range, viewer illumination (e.g. diffuse, condenser, etc.) and the colour print paper used. It is important that each film should have the same gamma so that the final composite colour scale is not distorted. This in turn requires careful tests, since for most films the gamma will vary with wavelength; the complete system may be tested by using a grey scale target - if this can be reproduced in the composite print without any colour tint then the negatives are well matched. Three negatives may be best tested by projecting through yellow, magenta, and cyan filters respectively onto a colour sensitive paper.

Various types of film and paper processing systems are available, and the choice really depends on available funds. A continuous processor is the ideal method for films, but has a very high cost. At present, spiral reel tanks are used for panchromatic and colour films, and with careful operation these have produced excellent results. For papers, batch processing in tanks is used, which is satisfactory provided that replenishment is frequent and accurate. For both film and paper processing, temperature control is very important and is maintained by placing processing tanks in thermostatically controlled water baths.

Image Viewing and Recording

Provided that the cameras, aircraft and processing facilities are carefully used, it should be possible to produce panchromatic narrowband negatives whose resolution is better than that of a conventional aerial colour film. The task of the viewing and recording equipment is to combine the negatives into a colour composite image without significantly reducing the resolution or adding unwanted colour or tone variations. Further, the additive viewer should have a good range of filtering on each channel which includes control of hue, brightness and saturation.

In many commercial instruments the range of hue available is very limited - often only three primary filters. Brightness is often controlled by lamp voltage, which also interferes with colour balance, and desaturation is very rarely provided. Another major deficiency in all but the most expensive instruments is the impossibility of obtaining high resolution copies of the composites, either in print or transparency form, and in some cases it is even difficult to obtain a well registered composite at all. In view

of these severe limitations it was decided to design and build an instrument; this has now been in use for a year, and has produced some excellent results. The main features of the viewer are described below.

The three projector units are modified colour printing enlargers (Durst M800), which are arranged in skew format to project their images in register onto a common baseboard. Each enlarger is fitted with a colour filter unit which allows secondary filters to be progressively added to the light, giving from zero to full saturation. These filters used in combination give a very wide range of hue and saturation to each light source. Brightness control is by infinitely variable lens diaphragm. Each projector has scale, focus and registration controls and these enable rapid and very accurate registration. The image is viewed by reflection from a white diffuse screen on the common baseboard, the operator being seated. Controls are arranged conveniently to hand, and include motorised registration, focus, scale and colour controls. modified negative carriers accept up to 6 metres of 70mm film and enable rapid viewing of a series of frames with the minimum of re-registration.

The instrument has been very satisfactory in use, and produces colour prints whose quality is never lower than that of a conventional colour print from 70mm transparency film. The reflective viewing screen is generally more suitable than the usual rear projection screen, which always gives granularity and a bright spot. If rear projection were necessary for map overlaying, a simple mirror and screen unit could be constructed; for registration and general viewing however, the direct reflected image is superior.

In order to make a print of the image, the screen is simply replaced by colour sensitive paper and an exposure made. To assist in correct exposure and colour balance, a four-channel light meter is used, with digital readout. A transparency is simply made either by using film instead of paper, or by photographing the screen with a suitable camera (this must have a displaced lens to produce an undistorted image).

APPLICATIONS AND RESULTS

Police Searches

The existing multiband system was developed to aid in the considerable problem of large area searches, particularly for buried or concealed bodies. Extensive study has shown that the anomaly produced by a concealed body can be more reliably detected by multiband techniques than by conventional film. The anomaly often consists of a subtle change in vegetation growth, particularly in the level of pigments in the leaf, and this can be enhanced by choosing narrowband filters which match the pigment absorption characteristics. Results of trials have shown that consistent improvement in target visibility can be obtained compared to colour films. The imagery taken on these trials over a period of twelve months has also shown that other features can be considerably enhanced. These include: vegetation changes due to subsurface water anomalies, delineation of surface water, tree species separation, archaeological marks.

Forestry

Potential applications of the multiband system to forestry survey appear very

promising from the initial studies so far carried out. Spectral measurements of young spruce trees having controlled levels of phosphate deficiency have shown that carefully chosen filter sets should give considerable enhancement of the small reflectance changes due to deficiency; this could provide a much more convenient method of planning phosphate application than the present sampling and chemical analysis of leaves. Other information which a well-planned multiband aerial survey could provide may well enable early detection of insect and fungal diseases, and also allow more accurate and rapid inventories - species differentiation has been shown to be much more marked on composites than on colour or false colour film.

Agriculture

Potential applications of the system to agriculture are well known, and include crop disease detection, insect infestation, soil moisture studies and crop identification. Again, the use of specific narrowband filter sets would provide optimum enhancement, and this would require considerable study. A great deal of work is in progress in this field, particularly in the USA, although a more co-ordinated approach would be desirable in the UK than is evident at present. A recent study carried out with the Plessey multiband equipment in Nigeria showed promising results for crop discrimination, but a great deal more work is required before a useful system can be evolved.

Water Resources

This field has a number of applications, such as river pollution and the location of water in arid areas. Those particularly relevant to the UK are the estimation of snow yield for hydroelectric catchment areas, and the study of water movement in river and reservoir catchments. Both these applications are the subject of a current study at the Plessey Radar Research Centre, and both seem likely to rely partly on the provision of aerial imagery with maximum information on water-related phenomena. Earlier work has shown that, for example, variations in water table level can be detected by the subtle effect on vegetation growth which is not visible on conventional colour imagery. It is also possible that the distinction between old and new snow could be enhanced, which is valuable information in assessing snow water equivalent.

Archaeology

The use of narrowband imagery for site surveys has recently shown that enhancement of archaeological features, particularly cropmarks, can be achieved. These features are notoriously unreliable, and success will always depend on fortuitous timing, but there is still a need for optimum detection of growth anomalies. The use of stereo multiband photography seems to offer the maximum probability of detection, since there is often a height or texture anomaly in archaeological marks as well as a small reflectance anomaly.

CONCLUSIONS

Development of the overall multiband system comprising spectrometer, data processing, aerial photography and image analysis is now at an advanced stage. Results have been very encouraging, and a number of applications have been identified in addition to that of police searches. Some of these

applications are being actively pursued, although due to lack of resources it is often difficult to obtain funding for potentially very useful studies.

The need for careful choice of filter sets is of primary importance in obtaining a final image with maximum enhancement of the required features, and this part of the system is often neglected. Results of spectral studies have shown that, in general, both the position and bandwidth of the camera filters is important; further, the optimum bandwidths are often narrower than those normally used in available multiband camera units. Without a study of detailed reflectance measurements as outlined in the section headed Spectral Reflectance, an optimum filter set is very difficult to design.

In all of the applications so far identified, the need for high quality composite imagery is very apparent, and many potential users have so far been unable to realise the maximum advantages of multiband photography because of poorly designed equipment - particularly the viewing and printing instrument. Because of the unavailability of a satisfactory additive viewer at reasonable cost, an instrument was constructed from readily available components and this has proved capable of meeting the requirements outlined earlier.

ACKNOWLEDGEMENTS

The major part of the work described in this paper was funded by the Home Office, Police Scientific Development Branch, who have kindly given permission to publish.

EARTH AND OCEAN DYNAMICS SATELLITES AND SYSTEMS—AN OVERVIEW

F. O. Vonbun

International Astronautical Federation (I.A.F.) XXVIth Congress

ABSTRACT

The purpose of this introductory overview is to provide a possible guide for future IAF session activities in the general area of Earth and Ocean Dynamics, application satellites, ground systems, their necessary analytical tools and software. Some of these topics have been discussed in two previous IAF papers, but this is the first time that a serious and concentrated effort in this rather new area of space applications is to be made.

The major potentials of these earth oriented Earth and Ocean Satellite and Systems objectives is in the area of "utilization" of space science and technology. This is to help and further geophysical exploration, earthquake hazard assessment, synoptic monitoring of the dynamics of the earth and oceans, the exploration of the polar ice regions and the refinements of the earth's gravitational and magnetic fields. These objectives have practical usage in such applications disciplines as protection of life and property, protection and prevention of disasters, exploration of mineral and energy resources, the shipping and fishing industry, coastal zone construction and off-shore drillings. It is evident that rather large economical benefits will result by achieving the above objectives. It is, however, very difficult at this time to determine a real monetary value for the benefits expected. Fortunately, some work has already been performed in the past years. Methods have developed, ground and flight systems have been constructed, operated and launched to satisfy at least some of the elements mentioned before.

Geomagnetic spacecraft launched by the U. S. (OGOs) and the USSR (COSMOSS) have contributed to the construction of world magnetic field maps important for geologic studies. Other spacecraft, together with precision ground tracking systems (range and range rate, Doppler, LASER, Very Long Baseline Interferometer), have been used to develop rather sophisticated earth gravity models and geoids important to such applications as tectonic motion and earthquake studies, future global geophysical explorations and sea current studies. The Skylab and Nimbus missions contributed to the development of ocean surface wave and wind field models, sea surface temperature fields, ice fields and details of the ocean topography, to quote a few specific examples. The real intent of this paper is, as mentioned, to show one hand of some typical examples of how these applications objectives can be achieved. It is hoped that it will provide and act as a catalyst and driver for future IAF sessions to come in this area of space applications.

INTRODUCTION

The purpose of this paper is to present to you an overview of the present state of satellite and ground systems making observations of the Dynamics of the solid earth and the oceans. It is necessarily confined to spacecraft and ground systems with which I am somewhat familiar, that is, NASA spacecraft and systems. Many results will not be presented because I may not be sufficiently informed about them and partly because there are too many to include in this paper.

The central theme of this application program is to provide a platform for a rather broad effort for the development of practical tools – predictive models and observational spaceborne and ground systems to further the social and economic benefit of mankind. Its primary goals are to develop, identify, demonstrate and utilize relevant space science-applications, techniques and methods for practical usage. In essence, "APPLY AND USE" to the benefit of all we have learned in the past two decades of space exploration in these particular areas. The objectives of this program fall almost naturally into two major categories – Earth Dynamics Application and Ocean Dynamics Application.

The discipline of Earth Dynamics includes the study and observation of such phenomena as tectonic plate motion, fault motion, earth rotation and polar motion, solid earth tides, as well as the motion of the earth in space. It further includes the exact determination of location of tracking stations on the earth surface, the determination and improvements of the earth gravitational and magnetic fields and in particular the study of their anomalies. The latter ones are important for practical applications in the area of mineral and energy resources, to quote an example.

The discipline of Ocean Dynamics encompasses studies and observations such as currents, circulation, waves, surface wind temperature fields, sea ice, polar ice and its structure, age and dynamics. It is understandable that all these phenomena are closely related and are of great interest to the shipping and fishing industries, the coastal zones as well as for the general area of weather forecasting. The vast ocean areas covering about 70% of the globe's surface play a large role in interacting with the atmosphere, thus influencing the world's weather and climate. It should be emphasized here that this paper is more or less result-oriented and does not particularly stress any specific planning activities. Few future systems and spacecraft are only briefly mentioned.

Geomagnetic spacecraft launched by the U. S. (OGOS) and the USSR (COSMOSS) have contributed to the construction of a world geomagnetic scalar field map important for future geologic studies. Other spacecraft, together with precision ground tracking systems (range and range rate, Doppler, LASER and Very Long Baseline Interferometer (VLBI)) have been used to develop rather sophisticated earth gravity models and geoids important to such applications as tectonic motion and earthquake studies, future global geophysical explorations and sea current studies. The Skylab and Nimbus missions contributed to the development of ocean surface wave and wind field models, sea surface temperature fields, ice fields and details of the ocean topography, to quote a few specific examples.

Tables 1 and 2 show a listing of the present state of the art and possible future requirements in the Earth and Ocean Dynamics disciplines. The numbers shown are approximate ones and in no way fixed, nor should they be. On the contrary, it is hoped that in future sessions these tables will be modified and extended so that they can be used as a guide for future national and international efforts in these disciplines.

OBJECTIVES AND REQUIREMENTS

As indicated, the major potential earth-oriented Earth and Ocean Dynamics Satellite Systems objectives are to "utilize" space science and technology to further and help develop and validate methods and models:

- (a) to support our continuous and ever growing needs and thus search for mineral and energy resources by studying the anomalies and possible correlation of the earth gravity and magnetic fields, the geologic composition and structure of the earth surface;
- (b) to study plate tectonic, fault and polar motions, solid earth tides and earth rotation leading together with ground observations to better predict probable time, location and intensity of major earthquakes;
- (c) to refine our knowledge of the earth's gravitational and magnetic fields, particularly its globally distributed anomalies, to support studies mentioned under (a) above and to refine the earth geoid (sea surface topography) to support and further our knowledge on ocean currents, circulation, storm surges, and other surface phenomena;
- (d) to synoptic monitor the world's oceans, transient phenomena including the magnitudes and geographical distribution of sea state, surface salinity, eddies, tides, surface winds, storm surges, swells with emphasis of identifying potential hazards for the shipping and fishing industries as well as the coastal zone areas to provide needed ocean surface (air/sea interaction) information for weather and climate forecasting; and
- (e) to assess the general ocean circulations, currents and their transport of mass, heat, nutrients, polar ice, ice structure, drift and age, and open water areas important for the fishing industry, weather and climate forecasting.

In order to accomplish these main objectives listed above, it is obvious that proper methods, instruments, ground and flight systems have to be developed, built, launched and operated. Fortunately, an appreciable amount of work has been done over the past years in this respect.

SATELLITE MISSIONS AND RESULTS

Ever since the first artificial satellite was launched almost two decades ago, real progress has been made through earth viewing and sensing spacecraft in the

TABLE 1 Practical Uses of Outer Space Ocean Dynamics Requirements

PHENOMENA	STATE OF THE ART	POSSIBLE FUTURE NEEDS
WAVE HEIGHT	$\pm 30\%$ ACCURACY, LIMITED OCEAN COVER-AGE	0.5 m OR $\pm 10\%$ GLOBAL
WAVE DIRECTIONAL SPECTRUM	LIMITED IN VERY LOCALIZED AREAS	DIRECTION $+5^\circ$, 10% MAGNITUDE $\lambda = 50$ TO 1000 METERS
SURFACE TEMPERA-TURE FIELD	± 1 TO 2°C , 5 km RESOLUTION IR RADIOMETER, 0 TO 35°C	$+0.3^\circ\text{C}$ GLOBALLY, RESOLUTION 10 km, 2 TO 40°C
SURFACE WINDS	SPEEDS DETERMINED 3 TO 15 m/s FROM MICROWAVE RADIANCE (NIMBUS)	DETERMINE SPEED TO $\pm 10\%$ AND DIRECTION TO $\pm 15^\circ$
CURRENTS, CIRCULATION	BOUNDARIES DETERMINED BY TEMPERA-TURE MAPPING (SKYLAB, NIMBUS)	VELOCITY DETERMINATION TO ± 1 TO 3 cm/s (GEOSTROPHY)

TABLE 1 (Cont'd) Practical Uses of Outer Space Ocean Dynamics Requirements

PHENOMENA	STATE OF THE ART	POSSIBLE FUTURE NEEDS
TIDES	TIDAL MODEL PREDICTIONS: 1.5 m AMPLITUDE IN OPEN OCEAN, 5° TO 15° IN PHASE (MANY METERS NEAR SHORE)	GLOBAL TO 2 TO 10 cm IN AMPLITUDE & 0.3° IN PHASE
ICE, POLYNYAS	ICE AREAS MAPPED, AGE DETERMINED, OPEN AREAS GROSSLY MAPPED	ICE STRUCTURE, & DRIFTS, POLAR ICE CAPS VARIATION, cm/yr., DEPTH, AGE, OPEN WATER WITH HIGH RESOLUTION
RESOLUTION ~ 30 km (NIMBUS)	10 to 20 m	
SURFACE SALINITY	SPARSE DATA, L-BAND, SKYLAB, 12%	WORLDWIDE, 1% ACCURACY OVER LARGER TEMPERATURE RANGE
ACCURACY AT 30°C ONLY		
OCEANIC EDDIES	ONLY SPARSE INFORMATION AVAILABLE FROM TEMPERATURE MAPS	SCALE 100 TO 300 km
		2 TO 10 cm TOPOGRAPHY,
		POSITION TO ~10 km

TABLE 2 Practical Uses of Outer Space Earth Dynamics Requirements

PHENOMENA	STATE OF THE ART	POSSIBLE FUTURE NEEDS
FAULT MOTION	KNOWN NEAR THE FAULT LINE (± 20 km)	0.5 TO ± 1 cm/yr ACCURACY NEAR & FAR FROM THE FAULT
TECTONIC MOTION	NO REAL GEODETIC MEASUREMENTS, AVERAGE OVER MILLIONS OF YRS TO ± 0.5 cm/yr	0.5 cm/yr OVER 300 TO 500 DAYS
POLAR MOTION	± 80 cm OVER 6 TO 12 HRS. n-s COMPONENT, LASER OR CORNER CUBE SPACERCRAFT, VLBI	± 2 TO ± 5 cm TOTAL
EARTH ROTATION	1 msec OR 50 cm	LESS THAN 0.05 TO 0.1 msec OR ~ 2 TO 5 cm
VERTICAL MOTION	INSUFFICIENT DATA OVER LARGE DISTANCES (500 km)	0.5 cm/yr OVER FEW HUNDRED KM REGIONS
SOLID EARTH TIDE	± 10 cm AMPLITUDE, PHASE ERROR LARGE (OCEAN LOADING)	± 1 cm AMPLITUDE & $\sim 10^\circ$ IN PHASE, GLOBALLY
GRAVITY FIELD	25 x 25 ORDER & DEGREE	1 TO 3 mgals, ($2^\circ \times 2^\circ$) GLOBALLY

TABLE 2 (Cont'd) Practical Uses of Outer Space Earth Dynamics Requirements

PHENOMENA	STATE OF THE ART	POSSIBLE FUTURE NEEDS
SURFACE GEOLOGY	DETECTED LINEARS, SKYLAB, LANDSAT	ACCURATE LINEAR STRUCTURE IDENTIFICATIONS
GEOID	ACCURACY 3 TO 5 m LOCAL 1500 km, 5 TO 25 m GLOBAL	ACCURACY 2 TO 10 cm, $\lambda =$ 200 km, GLOBAL MEAN OCEAN < 1 cm
SURVEYS (LAND) MAPPING	± 10 m USING TRANSIT AND OTHER POSITIONING SYSTEM	m RANGE
SURVEYS (OCEANS)	± 10 m USING TRANSIT	5 TO 10 m
MAGNETIC FIELD	$\pm 20\gamma$ SCALAR, GLOBAL	$\pm 3\gamma$ SCALAR $\pm 6\gamma$ IN VECTOR COMPONENTS GLOBAL AT 300km HEIGHT

understanding of the environment of our own planet. Here lies the great contribution of present and future applications satellites in general.

Earth Dynamics

In this discipline, such spacecraft as Vanguard 1, launched in 1958, together with the NASA Minitrack System (Mengel, J. T., et al., 1959), have contributed to the analyses of our earth's gravity field. The pear-shape of the earth was computed based upon orbital perturbations (O'Keefe, J. et al., 1959), to quote an example. In the following years, great strides have been made by many workers to steadily improve our knowledge in the earth gravity field (References 3 through 16). The latest model, the Goddard Earth Model (GEM 8) was developed using 400,000 optical and electronic space tracking data from 27 spacecraft (Lerch, F., 1974). These data were taken by space tracking systems such as Minitrack, Radars, Doppler, Range and Range Rate, Cameras, and over the last few years, high precision LASERS. In addition, about 1,600 ground gravity data ($5^\circ \times 5^\circ$ average) from the Air Force Chart and Information Service in St. Louis have been used in constructing this 25 degree and order model having an accuracy of about ± 8 mgals over $4^\circ \times 4^\circ$ segment of the earth surface.

Polar motions and earth tides have further been detected and analyzed using precision laser tracking of Beacon Explorer-C (Smith, D. E. et al., 1972, Kolenkiewicz, R. et al., 1973). Two laser stations, one located near Seneca, New York, and one located at Goddard, have been used for this experiment. The N-S component of the pole motion was determined to within an accuracy of ± 80 cm over a half a day's time interval. Figure 1 depicts the result of this first laser N-S polar motion experiment superimposed on the BIH values. As can be seen, a fairly good agreement was obtained for this first test. Interesting enough, it turned out that only one station was sufficient to perform this experiment. In addition, another result was obtained, namely, the distance between these two stations separated about 400 km could be established. The ability to repeat this distance determination from separate satellite laser rangings turned out to be in the dm-range as shown in Fig. 2. This result gave the actual impetus to a further new experiment, namely, the San Andreas Fault Experiment (SAFE) (Smith, D. E., Vonbun, F. O., 1973).

Improvements in the past have not only been made in the determination of the earth's gravity field as mentioned but also in the development of the earth magnetic field. Particularly instrumental were the Orbiting Geophysical Observatories launched in 1965, 1967, and 1969 (Cain, J. C. et al., 1966). From POGO data field in terms of spherical harmonics to the order and degree of 11 was developed (Cain, J. C. Langel, R. A., 1971). This field, computed for an average height of 550 km was constructed with 390,000 improved OGO observations and has an overall accuracy of about $\pm(6$ to 8) gammas.

More important than the magnetic survey maps as such are maps of magnetic anomalies. These are more closely related to the crust, that is, to mineral and earth resources. A recent global anomaly map has been published (Regan, et al., 1975). In this map, POGO satellite magnetometer data are used as a basis for the determination of crustal magnetic anomalies. Verification of very distant anomalies was obtained by examining individual satellite profiles as well as data collected during the

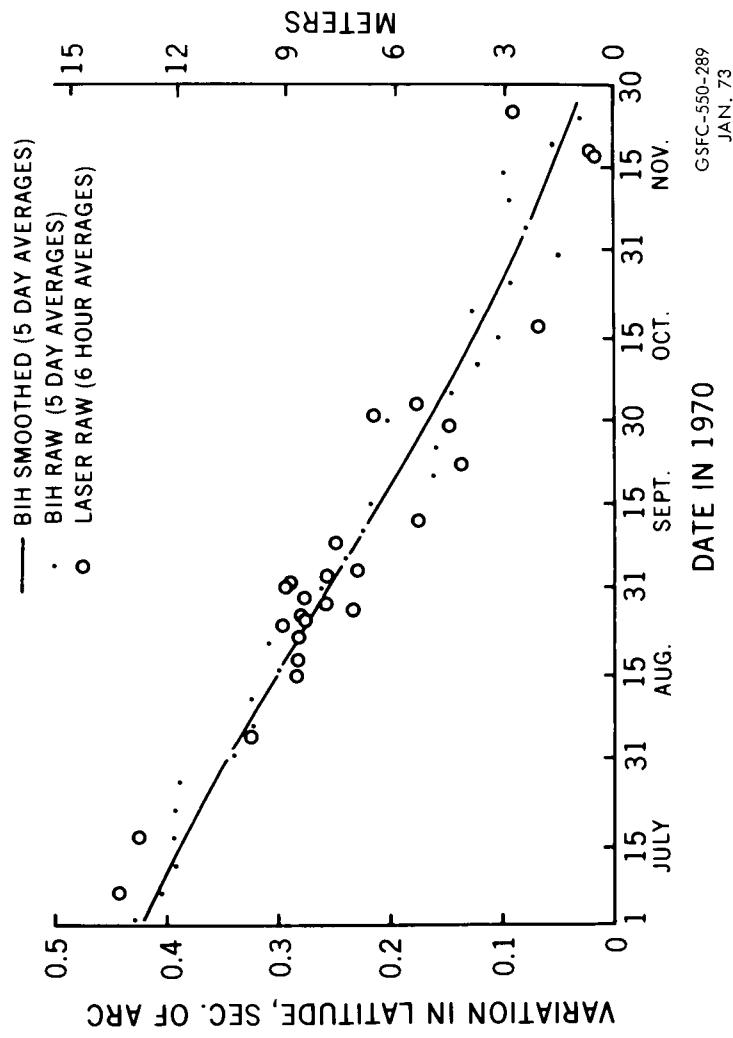
VARIATION IN LATITUDE OF GODDARD LASER

Figure 1

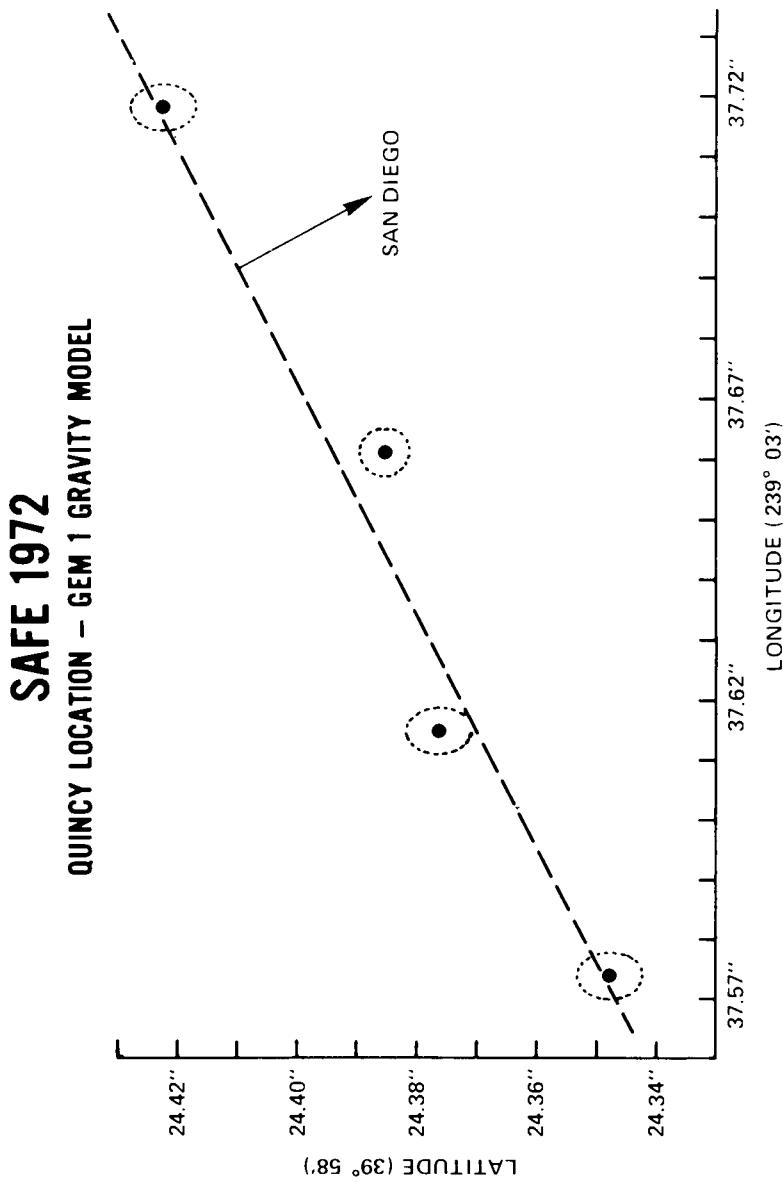


Figure 2

Project MAGNET (Stockard, H. P., 1971). The persistence of anomalies of all satellite data independent of local time indicate that these anomalies are real and not the result of any magnetospheric or ionospheric disturbance. These results demonstrate the real utility of satellite magnetometer data as a geological and geophysical survey tool.

It is suspected that the greatest range in magnetic anomalies may turn out to be areas that had the most intense geotectonic activity such as crustal movement, faulting and volcanism. The so-called Bangui anomaly depicted in Fig. 3 in Africa, for example, shows up clearly in the map and lies over the tectonic uplift zone between Chad to the north and the Congo Basin to the south. It should also be noted that in the same area rather large gravity anomalies have recently been detected using the latest Goddard gravity fields. Thus, it seems that there exists a correlation between the gravity and magnetic field anomalies in this region of geography. Work will continue to study these correlations in more detail.

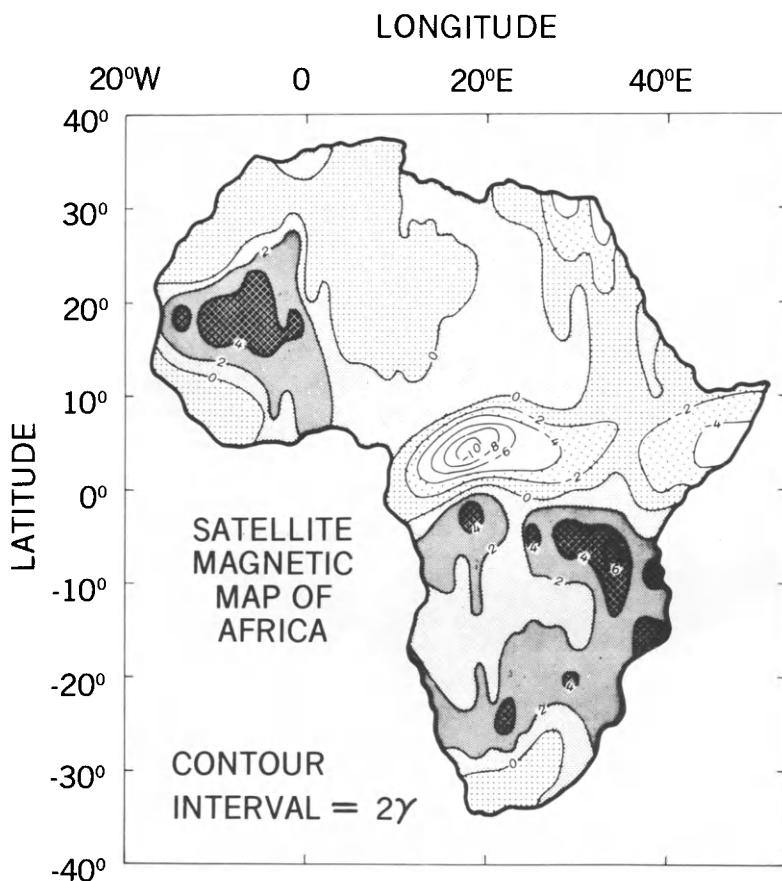
During the last few years, a new tracking/sensing method, namely, satellite-to-satellite tracking has been developed to specifically detect and determine global gravity anomalies of the earth's field (Vonbun, F. O., 1975).

Spacecraft such as GEOS-3, Nimbus-6 and recently the Apollo-Soyuz have been tracked via ATS-6. Gravity anomalies, which could not be sensed in the past by normal ground tracking systems have been "seen" by this method during the Apollo-Soyuz mission. Figure 4 shows a ground track together with an anomaly in the Himalayas. It is the anomalies as mentioned in both the earth's gravitational field, as well as magnetic field, that are of particular and practical interest for the applications disciplines. These are possible indicators of mineral and earth resources as well as crustal structures in general.

We have recently developed a new satellite geoid at Goddard as shown in Fig. 5, together with the Skylab 4 ground track (Marsh, J. and Vincent S., 1974) based upon the Lerch-field. During the last Skylab mission we had the first opportunity to make an actual partial accuracy test of this geoid (Vonbun, et al., 1975). The on-board radar altimeter "measured" the "computed" sea surface which is shown in Fig. 6. As can be seen, a very good agreement between these two surfaces to say within ± 4 to ± 8 m was obtained. It should, however, be pointed out here that we did "adjust" the Skylab orbital height which in this case is nothing else than a small height bias error as far as the sea surface variation is concerned and thus is not important in this special case.

This latter investigation brings us already into the area of Ocean Dynamics and demonstrates clearly that these earth observation disciplines are related to each other. The study of the deviation of the real ocean surface from the geoid does reveal such information as geostrophic currents, wind pile ups, eddies, storm surges, etc. providing, of course, that the sea surface is accurately known.

MAGNETIC ANOMALIES FROM POGO



FROM: Regan R. D. et al
J. Geoph. Res. Vol 8, No. 5, Feb. 10, 1975

Figure 3

ATS-6 APOLLO-SOYUZ
INDIAN AND HIMALAYAN GRAVITY ANOMALIES

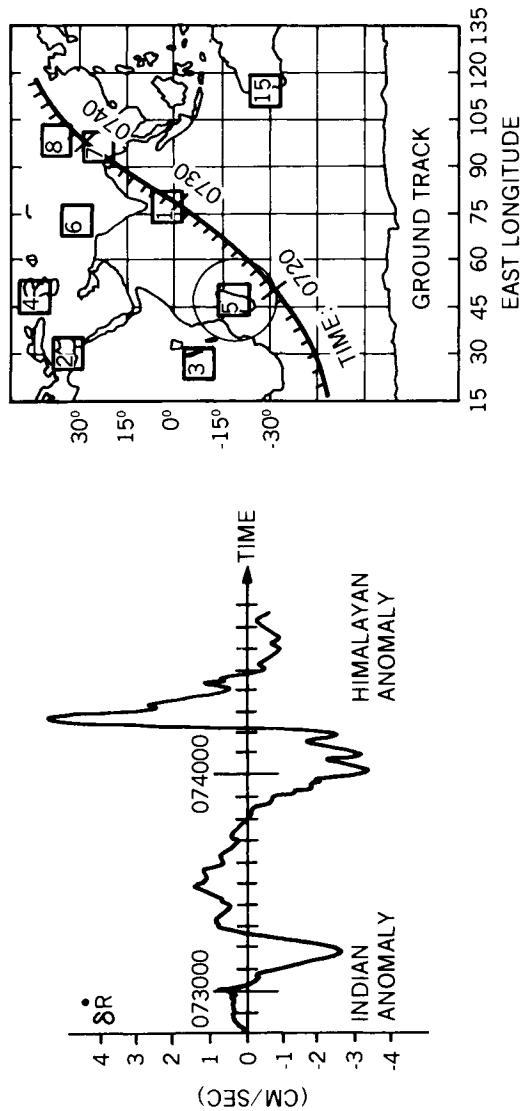


Figure 4

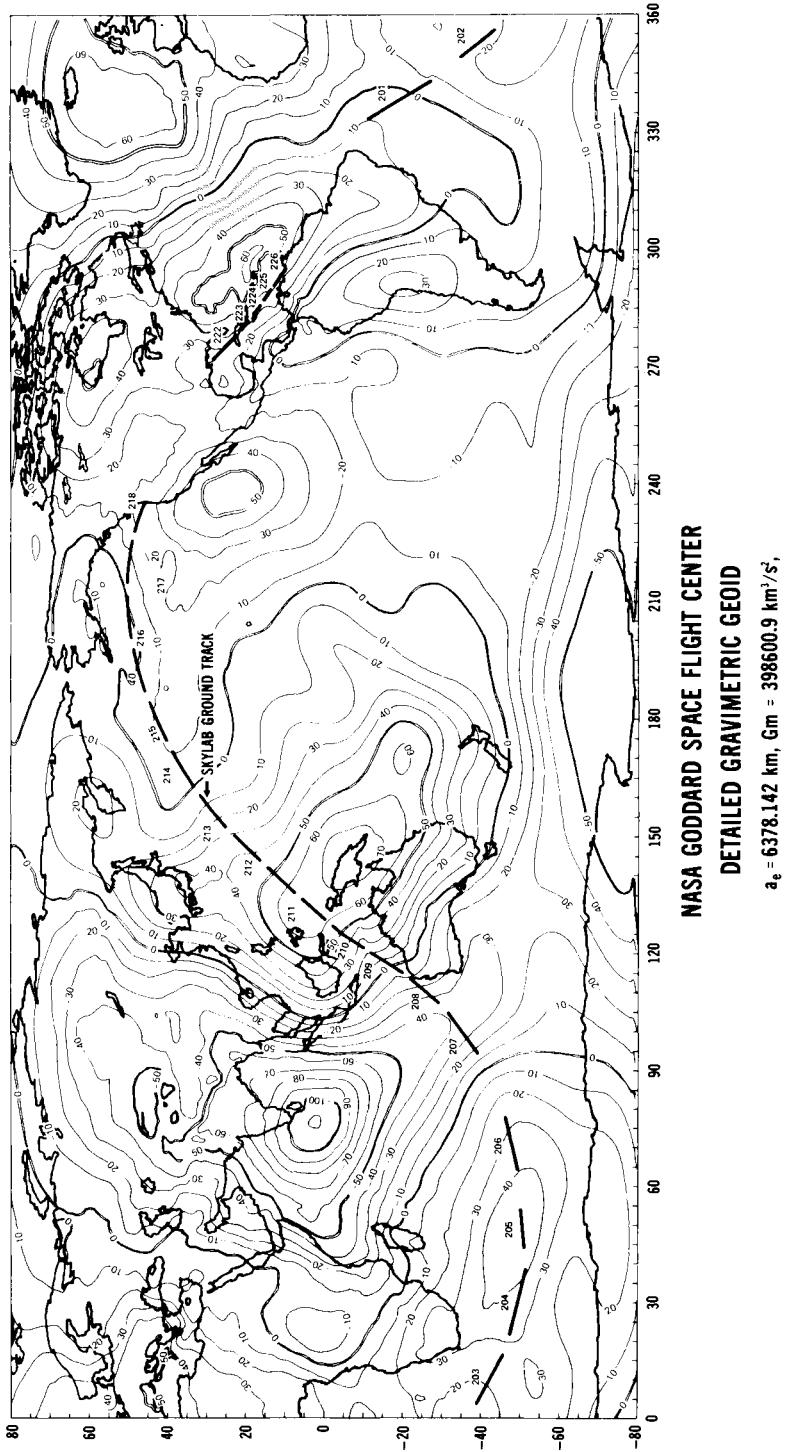


Figure 5

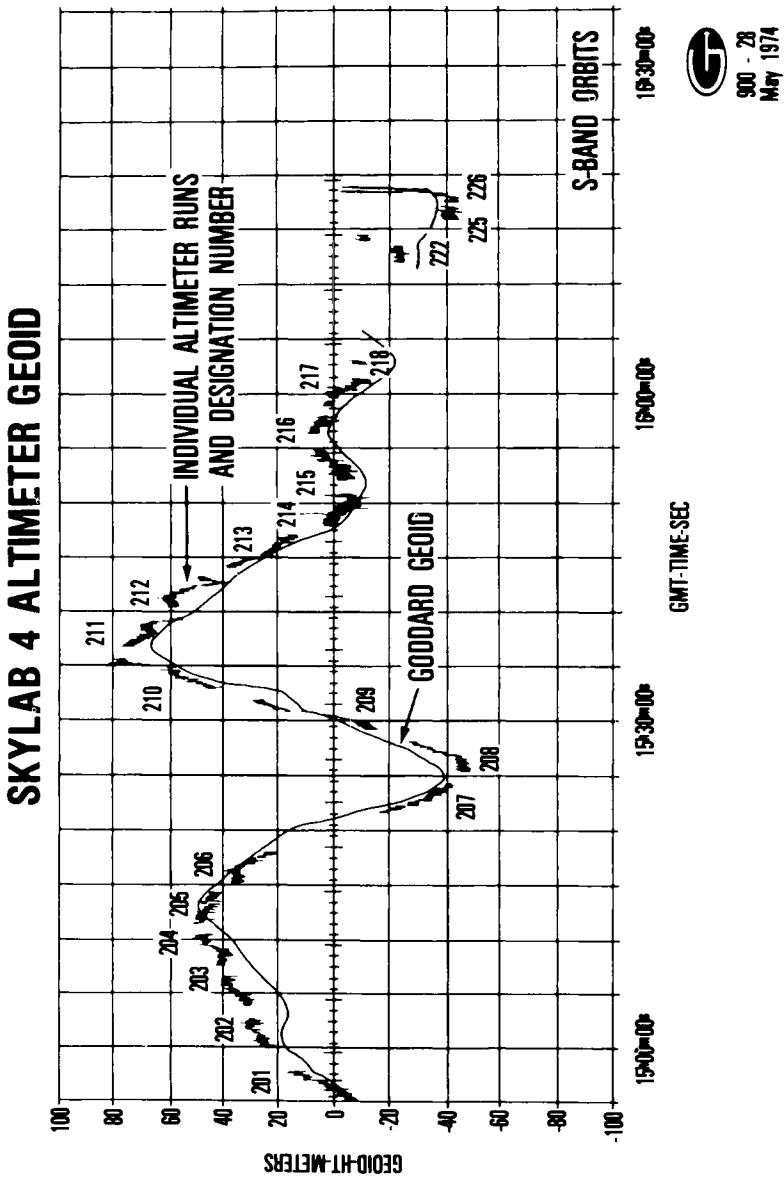


Figure 6

Ocean Dynamics

As mentioned, Skylab radar altimeter data are bridging the Earth and Ocean Dynamics disciplines. The altimeter data show such features as water "grabens" and "mountains". Figure 7 depicts the Puerto Rican Trench and Fig. 8 shows a rather distinct local elevation of the ocean surface near the Cape Verde Islands from Skylab III, Pass 13, on September 3, 1973 (McGoogan, 1975 and Vonbun, et al., 1975). This clearly indicates the extreme usefulness of radar altimetry for detailed topographic studies of the oceans.

Radar altimeter data from Skylab not only refined the sea surface topography as shown but gave new information on wave height determination, rain cell and surface wind detection. Ocean wave height data is obtained by studying the shape of the leading edge of the return pulse from the sea surface. The slope of the leading edge is an indicator of the sea surface roughness (McGoogan, 1974 and Miller, et al., 1972). Similar, rain cells have also been detected using the Skylab altimeter as shown in Fig. 9. W. Pierson further was successful in determining the surface wind from the S-193 Skylab scatterometer as shown in Fig. 10. (Pierson, W., 1975).

Figure 11(a) shows the variation of the emissivity ϵ with ocean windspeed for vertical polarization and 45° nadir angle, the variation of ϵ with frequency for nadir observations of smooth and rough seas and the variation of ϵ with surface temperature from nadir observations. The variation of ϵ with speeds above 7 m/s, the contributions of white water becomes dominant over the contribution of waves (Nordberg, Conaway, Ross and Wilheit, 1971). The variation of ϵ with frequency is also dominated by the white water for rough seas. The white water, streaks and foam combined, acts as a thin matching layer and raises the brightness temperature to nearly the thermodynamic temperature for wavelengths shorter than the electrical depth of the layer. As the wavelength becomes longer, the matching becomes less effective and the brightness temperature approaches the smooth water values (Webster, Gloersen and Wilheit, 1974). The variation of ϵ with surface temperature has a maximum near 5 GHz. At this frequency, a change in the thermodynamic temperature of 2° K results in about 1° K change in the brightness temperature.

Figure 11(b) shows the variation of ϵ with ice age for observations at vertical polarization and 45° zenith angle and the variation of ϵ with frequency and age for nadir observation. The appearance of scattering centers within the ice and the lowering of the loss tangent as the brine pockets in the ice drain is responsible for the lowering of ϵ at the high frequency and end of the spectrum.

This demonstrates that both active and passive microwave systems play, and will even more so in the future, an important role in the applications disciplines discussed.

It should, however, be mentioned that the only active systems flown in orbit (Skylab) was rather new and was operated on a limited schedule. GEOS-3 is the first unmanned spacecraft carrying an active radar which is intended to operate at least one full year in orbit but unfortunately has no scatterometer capability for wind determination. Thus, the output for the ocean dynamics area using this spacecraft is also somewhat restricted.

SEASURFACE SKYLAB ALTIMETER MEASUREMENTS

(PUERTO RICAN TRENCH, SL-2 MISSION)

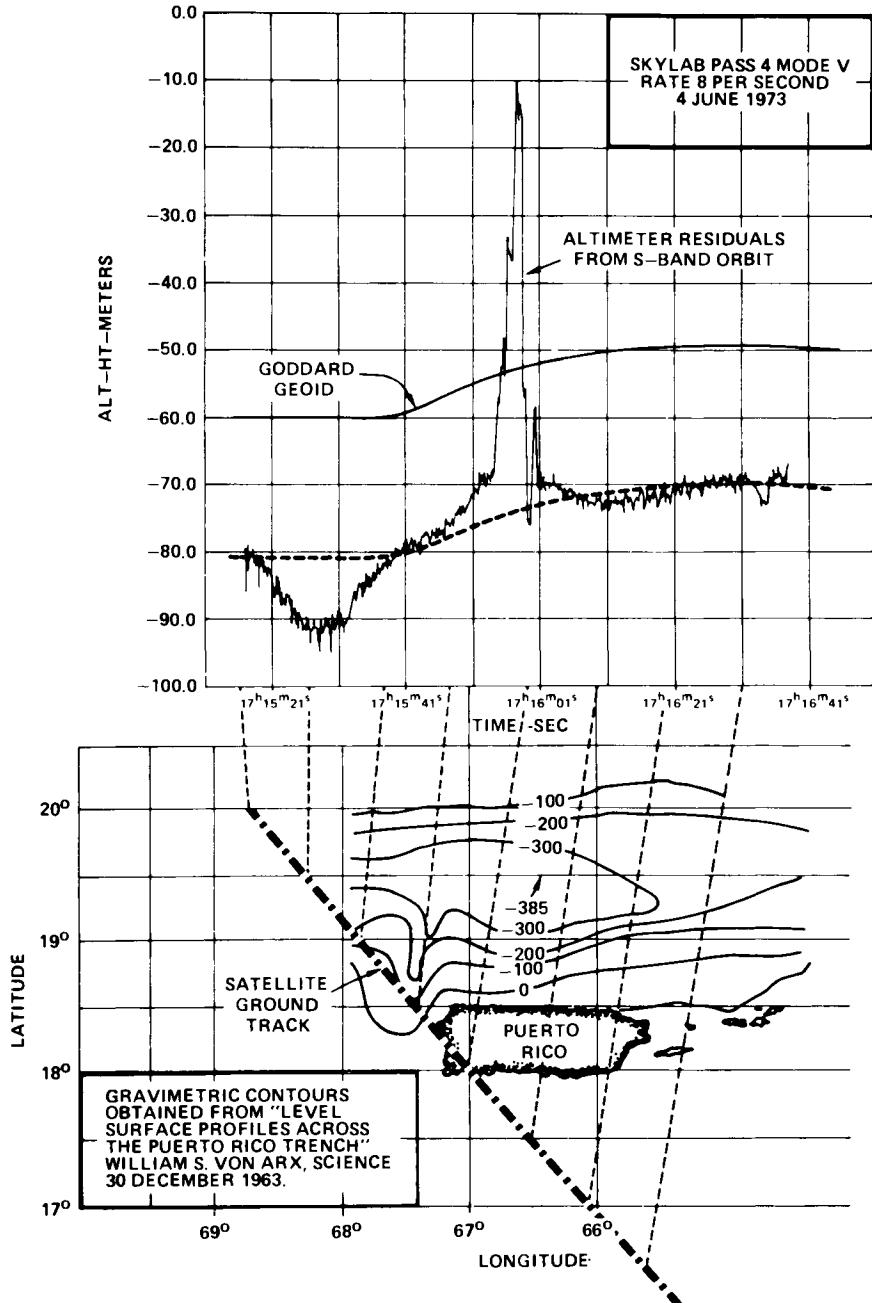


Figure 7

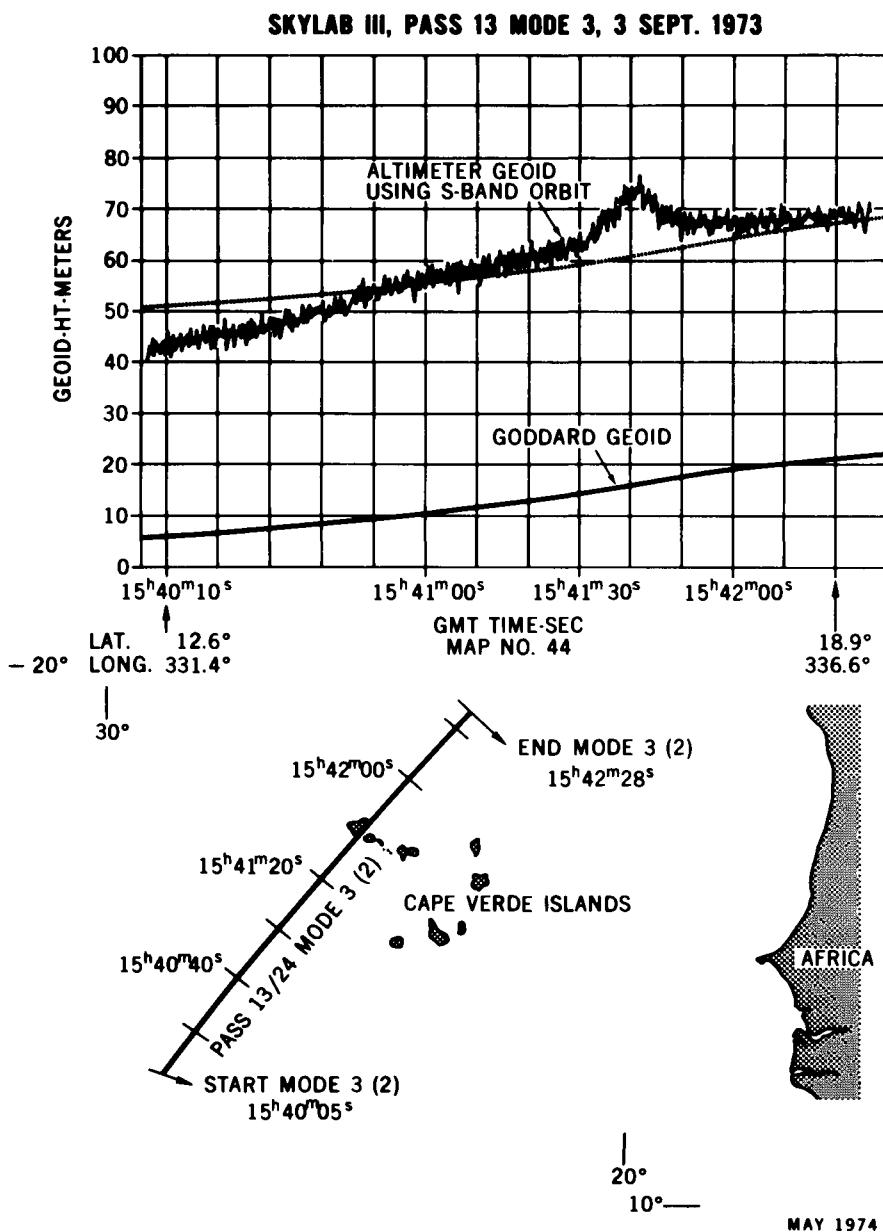


Figure 8

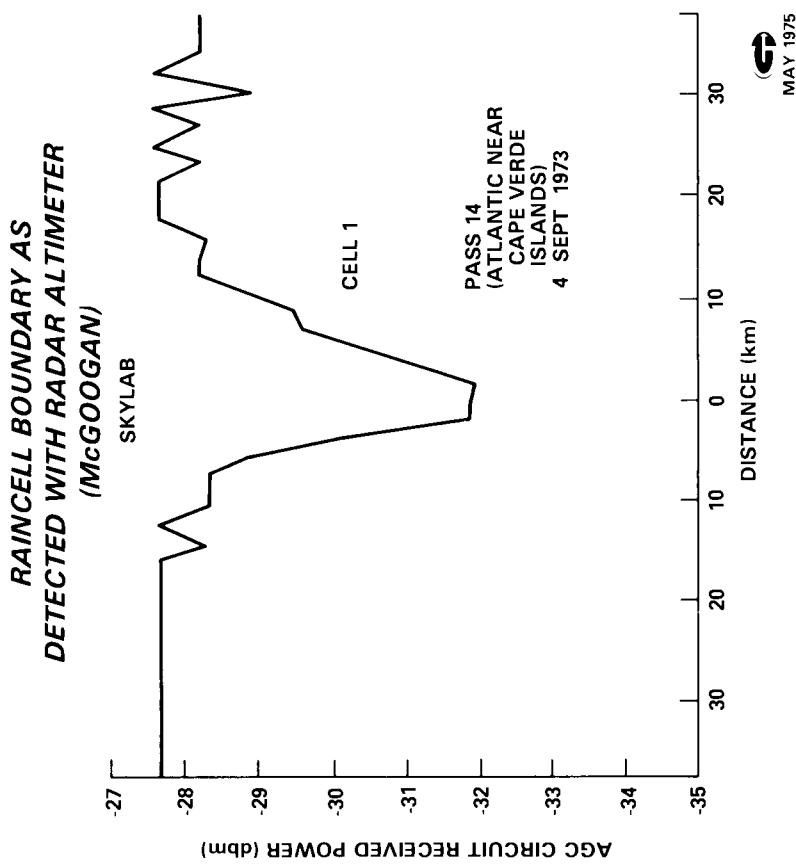


Figure 9

Meteorological (V_M) Versus Radar (V_R) Winds
Preliminary

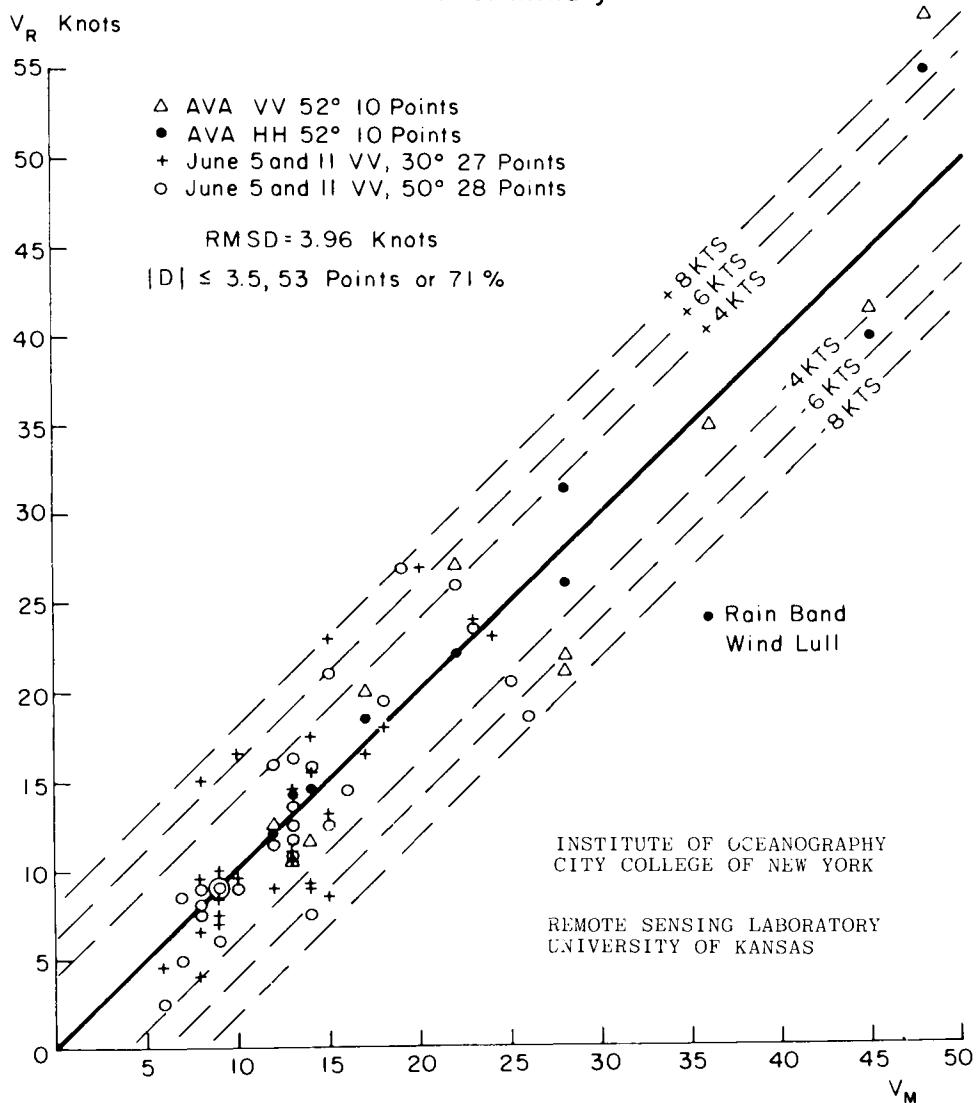
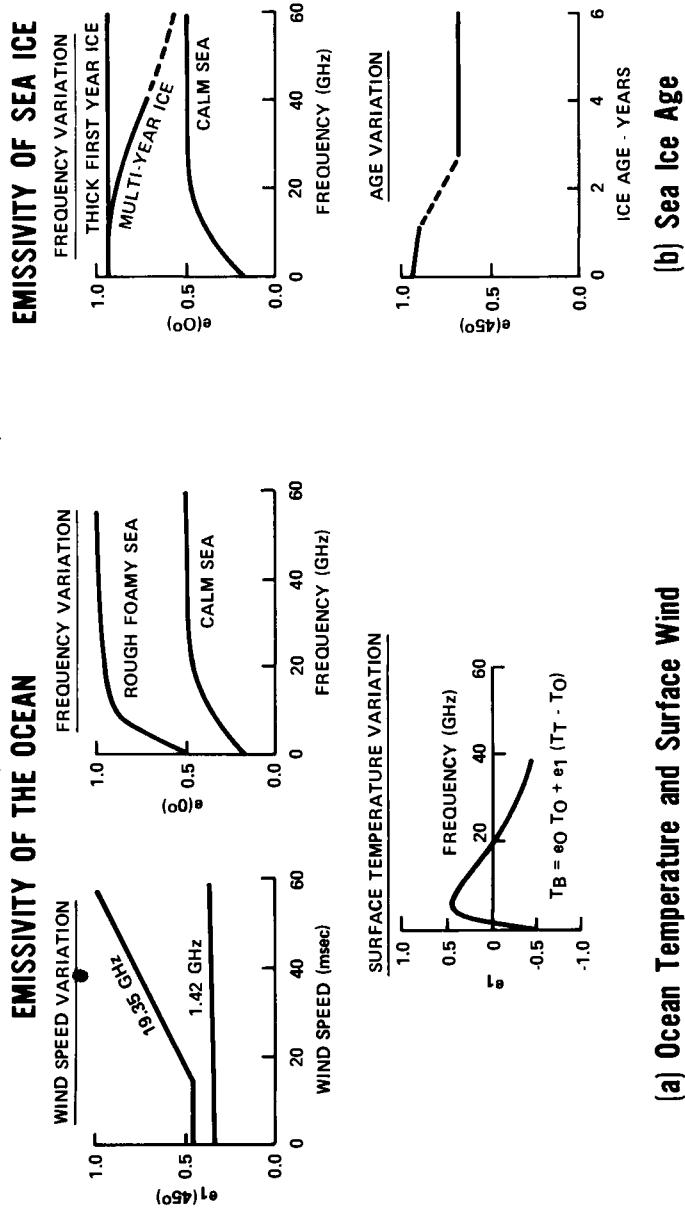


Figure 10

OCEAN OBSERVATIONS [Passive Microwave]



(a) Ocean Temperature and Surface Wind

Figure 11

(b) Sea Ice Age

SEASAT-A, to be launched in 1978, will however be the first dedicated spacecraft for both the Earth and Ocean Dynamics disciplines. The phenomena to be observed by this spacecraft are shown in Fig. 12.

GROUND SYSTEMS

High precision laser ranging systems have been developed at Goddard as mentioned over the past decade which enables us to determine satellite orbits accurately in the m-range and intersite distances in the cm-range, as shown in Fig. 13, important for geodynamics. Today's laser systems have bias and noise errors in the order of 5 to 8 cm as can be seen from Fig. 14 (noise in this case). On-going development of laser ranging system at the Goddard Space Flight Center and elsewhere will hopefully bring these errors down to the 2 cm level. It is further planned to expand the Goddard laser operations to a global basis for tectonic plate motion determination. Using properly distributed laser stations over the globe (Goddard, SAO and possible stations of other nations) will be used to measure intercontinental distances in the cm-range, together with the Lageos spacecraft as shown in Fig. 15 to be launched next year, that is, in 1976. This will enable us to actually determine the relative motions with great accuracies over these distances. Verifications of computed tectonic motions will then be possible for the first time. At present, all plate motions are computed average values over missions of years.

In addition to the LASER effort, real progress has been made over the last 8 years in the area of Very Long Baseline Interferometry (VLBI). Since 1963, a rather intensive cooperative effort between MIT and Goddard Space Flight Center has been under way to develop and improve Very Long Baseline Interferometer techniques and systems. These systems are used for geodynamic observations such as polar motion, earth rotation, and the determination of intercontinental distances. The NASA worldwide tracking network with its large dishes is being utilized for this purpose. Distances between Haystack, Massachusetts and Goldstone, California have been determined with a precision in the 10 to 20 cm range as shown in Fig. 16. Again, this shows a relative measure, that is, LASERS and VLBI, are being pursued at present and will be in the future for precision geodynamics work. This dual approach is essential since there is at this time no other way in checking either one of these extremely accurate systems alone.

CONCLUSIONS

In summary, it can be stated that very good progress has been made in the area of Earth and Ocean Dynamics since this program was initiated in the 1969 to 1971 time frame (Vonbun, F. O., 1972).

EODAP missions such as GEOS-3 have been launched; Lageos is almost finished and will be launched next year; SEASAT has just been started and is planned for a 1978 launch; and, finally, Magsat is our latest effort and is to be launched in 1980. Experiments such as SAFE, together with polar motion and earth rotation are under way.

USER REQUIREMENTS FOR SEA SURFACE PARAMETERS

PHYSICAL PARAMETER	INSTRUMENTS	RANGE	PRECISION	RESOLUTION OR FOV	TOTAL FOV	COMMENTS
WAVE HEIGHT, $H_1/3(x, y)$	PULSE ALTIMETER COHERENT ALT.	1.0 - 20m	± 0.5 m or $\pm 10\%$	2x7 km SPOT	2-km SWATH	ALONG SUBSATELLITE TRACK ONLY
DIRECTIONAL WAVE SPECTRUM $S(\lambda, \theta, x, y)$	IMAGING RADAR (2-D TRANSFORM) 2-D WAVE SPECTROMETER	S: UNKNOWN λ : 50-1000 m θ : 0-360°	S: ... λ : $\pm 10\%$ θ : $\pm 10^\circ$	50 m RESOLUTION	20x20 km SQUARES	GLOBAL SAMPLES AT 250-km INTERVALS
SURFACE WIND FIELD, $\bar{U}(x, y)$	S: UNKNOWN λ : 6-500 m θ : 90° SECTOR	S: ... λ : $\pm 10\%$ θ : $\pm 90^\circ$	8 x 25 km SPOT	300-km SWATH ABOUT NADIR	GLOBAL SAMPLES AT 150-km INTERVALS	
SURFACE TEMPERATURE FIELD, $T(x, y)$	SCATTEROMETER U: 0-360° θ: UNKNOWN	U: 3-25 m/s θ : 0-360°	± 2 m/s, $\pm 10\%$ $\pm 20^\circ$, $\pm 10\%$	≤ 50 km SPOT	200-450-km SWATHS	GLOBAL, 36 HRS. (LOW SPEEDS)
GEODIAL HEIGHTS, $h(x, y)$ (ABOVE REFERENCE ELLIPSOID)	IR RADIOMETER θ: 20° TO + 35°C	θ: UNKNOWN	± 2 m/s, $\pm 10\%$	≤ 100 km SPOT	900-km SWATH ABOUT NADIR	GLOBAL, 36 HRS. (HIGH SPEEDS)
SEA SURFACE TOPOGRAPHY, $j(x, y)$ (departures from geoid)	PULSE ALTIMETER COHERENT ALT.	7 cm - 200 m	± 7 cm	2x7 km SPOT	18-km SPACING ALONG EQUATOR	SAMPLED THROUGHOUT ONE YEAR
OCEANIC, COASTAL, & ATMOSPHERIC FEATURES (PATTERNS OF WAVES, TEMP. CURRENTS, ICE, OIL, LAND CLOUDS, ATMOSPHERIC WATER CONTENT)	IMAGING RADAR IR RADIOMETER μW RADIOMETER	HIGH RESOLUTION HIGH RESOLUTION LOW RESOLUTION	ALL WEATHER CLEAR AIR	25 OR 100m 1.7 km 15-100 km	100 OR 200 km 1500-km SWATH 900-km SWATH	SAMPLED DIRECT OR STORED IMAGES BROADLY SAMPLED IMAGES GLOBAL IMAGES



Figure 12

*BASE LINES DETERMINED
LASER AND SATELLITES*

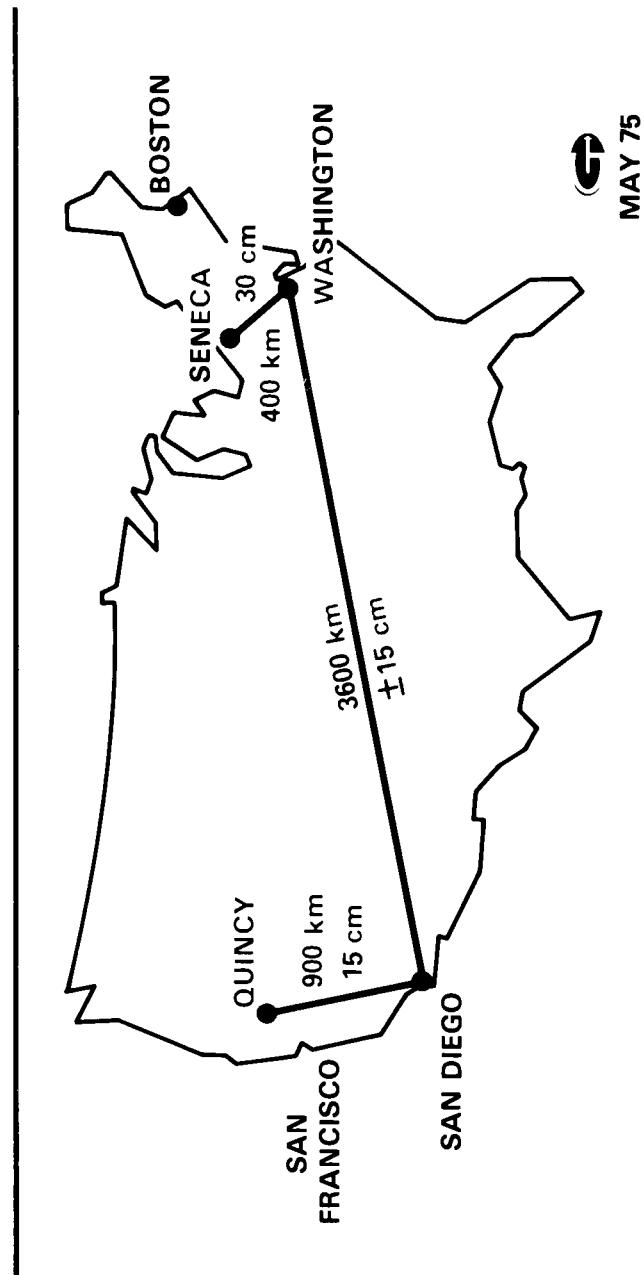
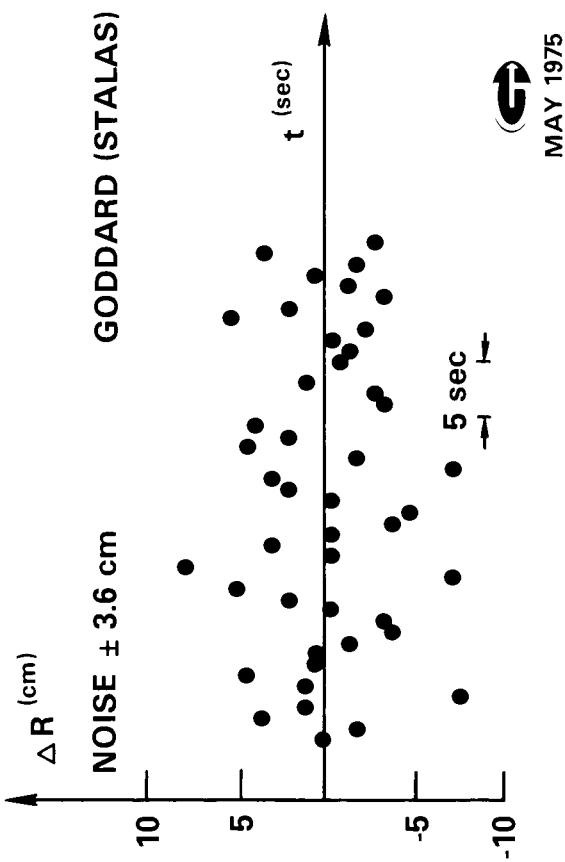


Figure 13

LASER RANGE RESIDUALS
(15TH DEGR. POL YN.)
GEOS-3
22 APR. 75



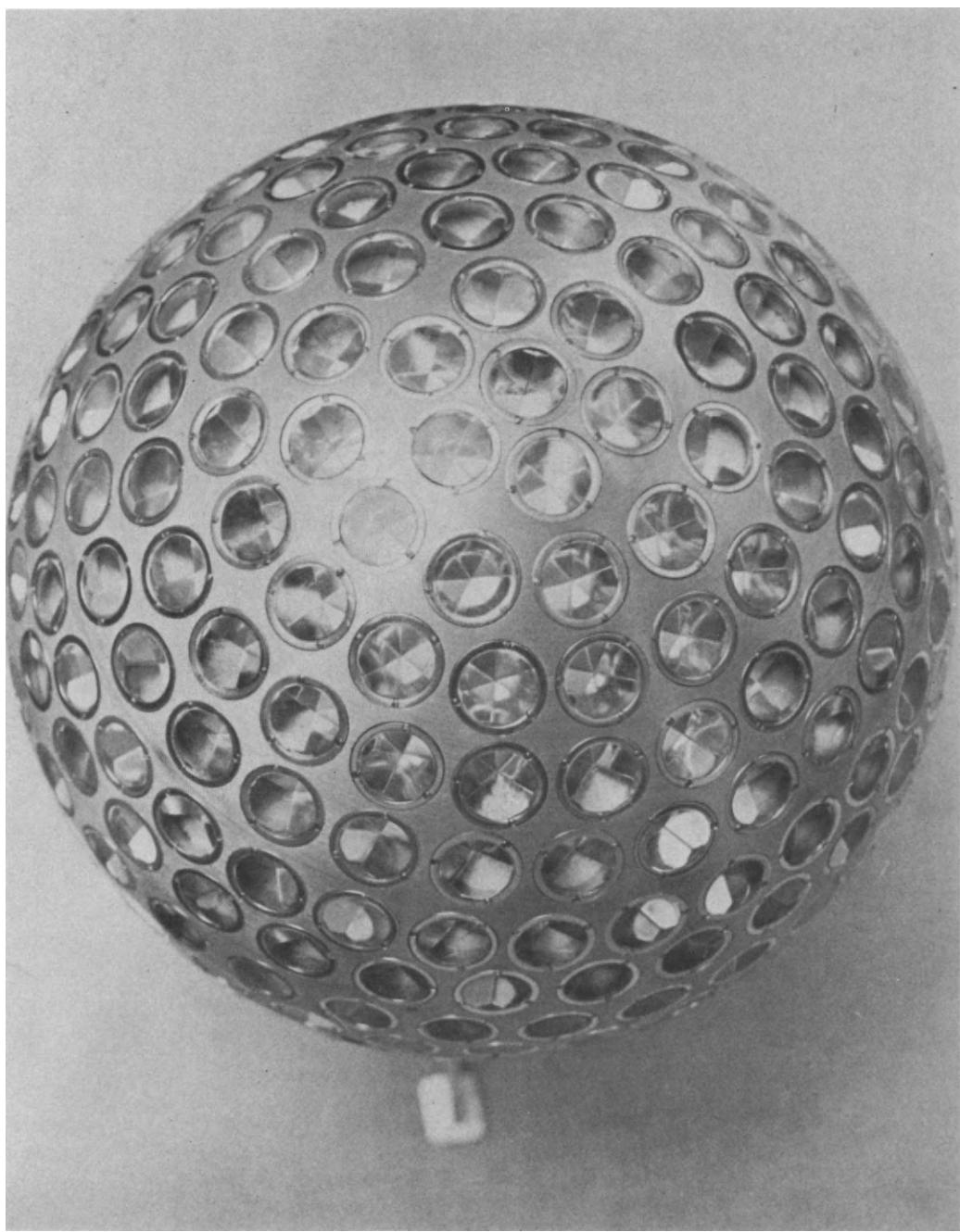


Figure 15

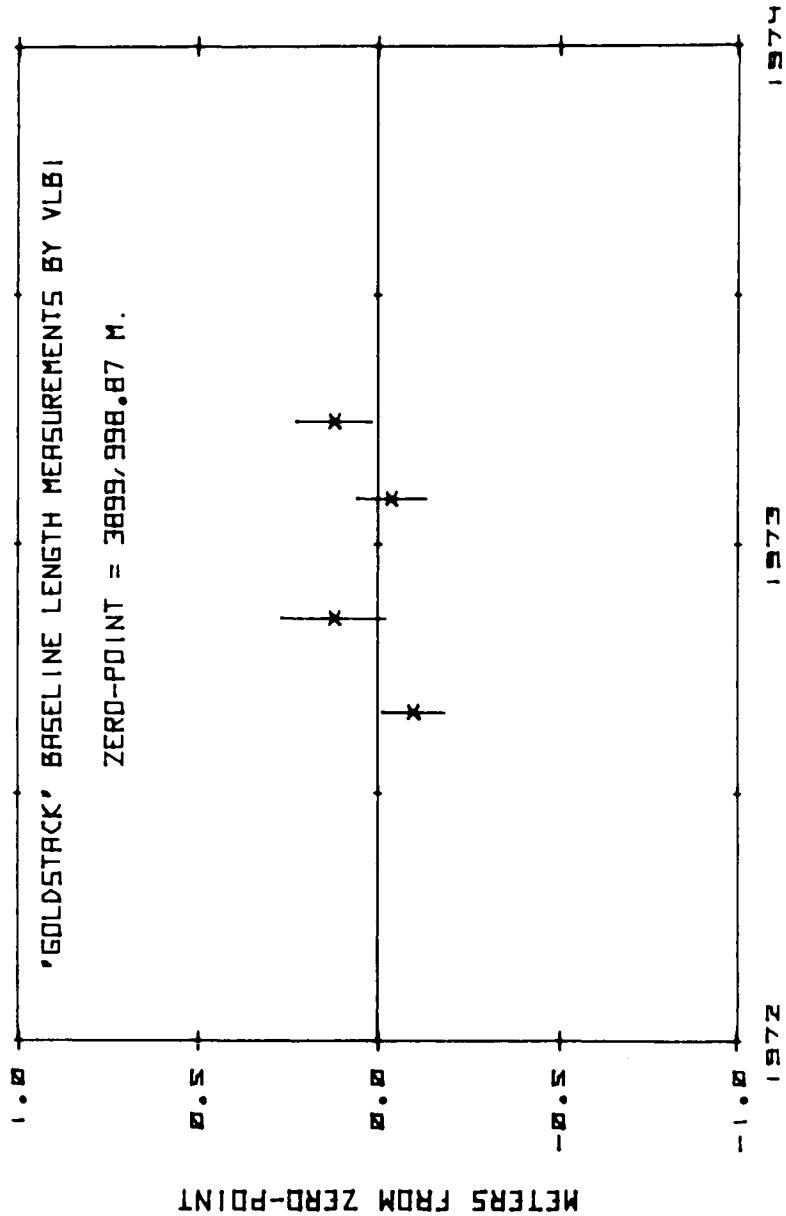


Figure 16

Construction of the needed mathematical models for data interpretation and analyses for earth dynamics phenomena as well as for ocean dynamics are also in progress. However, it should be realized that the ocean dynamics part of the program is behind its earth dynamics counterpart. This is mostly due to the fact that the proper systems needed for the latter one have not been flown in space yet. This is true particularly for active microwave systems.

However, it should be realized that we are still in a starting phase and a large effort, both national as well as international, is needed over the next decade in order to economically and socially benefit from the result of this new application program.

REFERENCES

1. Mengel, John T., "Tracking the Earth Satellite and Data Transmission, by Radio", Proc. IRE 44, pp. 755-760, 1956.
2. O'Keefe, J. A., A. Eckels, and R. K. Squires, "Vanguard Measurements Give Pear-Shaped Component of Earth's Gravure", Astron. J 64, No. 7 p. 245, 1959.
3. Veis, G., "Geodetic Uses of Satellites", Smithsonian Contrib. to Astrophys., 3, pp. 95-161, 1960.
4. Kaula, W. M., "Analysis of Gravitational and Geometric Aspects of Geodetic Utilization of Satellites", Geophys. J., 5, pp. 104-133, 1961.
5. Kaula, W. M., "Determination of the Earth's Gravitational Field", Revs. Geophys., 1, pp. 507-552, 1963.
6. Cook, A. H., "The Contribution of Observations of Satellites to the Determination of the Earth's Gravitational Potential", Space Sci. Rev., 2, 1963.
7. Anderle, R. J., "Geodetic Parameters SET NWL-5E-6 Based on Doppler Satellite Observations", G. Veis, ed. Proc. 2nd Int. Symp. Geod. on Use of Satellites, Athens, in press, 1966.
8. Guier, W. H., and R. R. Newton, "The Earth's Gravity Field Deduced from the Doppler Tracking of Five Satellites", J. Geophys. Res., 70, pp. 4613-4626.
9. Isask, I. G., "A New Determination of Non-Zonal Harmonics by Satellites", J. Kovalesky, ed. Trajectories of Artificial Celestial Bodies as Determined from Observations, Berlin: Springer Verlag, in press, 1966.
10. King-Hele, D. G., and G. E. Cook, "The Even Zonal Harmonics in the Earth's Gravitational Potential", Geophys. J., 10, pp. 17-30, 1965.
11. King-Hele, D. G., G. E. Cook, and D. W. Scott, "The Odd Zonal Harmonics in the Earth's Gravitational Potential", Planet. Space Sci. 13, pp. 1213-1232, 1965.
12. Kozai, Y., "New Determination of Zonal Harmonics Coefficients in the Earth's Gravitational Potential", Publ. Astron. Soc. Japan, 16, pp. 263-284.
13. Smith, D. E., "A Determination of the Odd Harmonics in the Geopotential Function", Planet. Space Sci., 11, pp. 789-795, 1963.
14. Smith, D. E., "A Determination of the Odd Harmonics in the Geopotential Function", Planet. Space Sci., 13, pp. 1151-1160, 1965.
15. Wagner, C. A., "A Determination of Earth's Equatorial Ellipticity from Seven Months of Syncor 2 Longitude Drift", J. Geophys. Res., 70, pp. 1566-1568, 1965.

16. Gaposchkin, E. M., and K. Lambeck, "1969 Smithsonian Standard Earth II", SAO-Special Report 315, 1970.
17. Lerch, F. J., C. A. Wagner, J. A. Richardson, and J. E. Brownd, "Goddard Earth Models (GEM 5 and 6)", GSFC Report X-921-74-145, December 1974.
18. Smith, D. E., R. Kolenkiewicz, P. J. Dunn, H. H. Plotkin, and T. S. Johnson, "Polar Motion from Laser Tracking of Artificial Satellites", Science Vol. 178, p. 405, October 27, 1972.
19. Kolenkiewicz, R., D. E. Smith, and P. J. Dunn, "Polar Motion and Earth Tides from Beacon Explorer-C", GSFC Report 592-73-236, May 1973.
20. Smith, D. E., and F. O. Vonbun, 24th IAF, Baku, USSR, October 7, 1973.
21. Cain, J. C., R. A. Langel, and S. J. Henricks, "First Magnetic Field Results from OGO-2 Satellite", Space Res., 7, pp. 1466-1476, 1966.
22. Cain, J. C., and R. A. Langel, "Geomagnetic Survey by the Polar Orbiting Geophysical Observatories", Magnetic World Survey, 1971, IAGA Bull., No. 28, pp. 65-74.
23. Regan, R. D., J. C. Cain, and W. M. Davis, "Global Magnetic Anomaly Map", J. of Geophys. Res., 80, No. 5, February 10, 1975.
24. Stockard, H. P., "Worldwide Surveys by Project Magnet", IAGA Bull. No. 28, pp. 60-64, 1971.
25. Vonbun, F. O., "The ATS-F/Nimbus-E Tracking Experiment", Symp. #48, IAU, Marioka, Japan, 1971, D. Reidal Publ. Comp., Dordrecht, Holland, 1972.
26. Vincent, S., and J. G. March, "Global Detailed Gravimetric Geoid", paper presented at the First International Symposium, The Use of Artificial Satellites for Geodesy and Geodynamics, Athens, Greece, GSFC Document X-592-73-266, 1973.
27. Vonbun, F. O., J. McGoogan, J. Marsh, and F. Lerch, "Sea Surface Determination from Space - The Goddard Geoid", GSFC Document X-900-75-216, August 1975.
28. McGoogan, J. T., "Precision Satellite Altimetry; 1974 IEEE Intercon Session 34", pp. 1-7, March 26-29, 1974, 34/3.
29. Miller, L. S., and D. L. Hammond, "Objectives and Capabilities of the Skylab S-193 Altimeter Experiment", IEEE Transact. on Geosc. Electr. GE-10, No. 1, pp. 73-79, January 1972.
30. Pierson (private communication).

31. Webster, W. J., T. C. Chang, P. Gloersen, T. J. Schmugge, and T. T. Wilheit, "Satellite Microwave Radiometry", 1974 IEEE Intercon Session 34, pp. 1-6, March 26-29, 1974, 34/3.
32. Gloersen, P., W. Nordberg, T. J. Schmugge, and T. T. Wilheit, J. G. R., 78, p. 3564, 1973.
33. Nordberg, W., J. Conaway, D. B. Ross, and T. T. Wilheit, Journ. Atmos. Sci., 28, p. 429, 1971.
34. Schmugge, T. J., A. Rango, L. Allison, and T. T. Wilheit, Bull. Amer. Meteor. Soc., 54, p. 1117, 1973.
35. Schmugge, T. J., P. Gloersen, T. T. Wilheit, and F. Geiger, to appear in J.G.R.
36. Vonbun, F. O., "Earth and Ocean Physics Applications Program, EOPAP", Astron. Res. 1972, pp. 239-245, D. Reidel Publ. Comp., Dordrecht, Holland, 1973.

AN INTEGRATED APPROACH TO THE REMOTE SENSING OF FLOATING ICE

W. J. Campbell*, R. O. Ramseier**, W. F. Weeks*** and P. Gloersen****

*U.S. Geological Survey, University of Puget Sound Tacoma, Washington 98416, U.S.A.

**Dept. of the Environment, 580 Booth Street, Ottawa, Ontario, K1A OH3, Canada

***U.S. Cold Regions Research and Engineering Lab, Hanover, New Hampshire 03755, U.S.A.

****NASA—Goddard Space Flight Center, Greenbelt, Maryland 20771, U.S.A.

ABSTRACT

The current increase of scientific interest in all forms of floating ice - sea ice, lake ice, river ice, ice shelves and icebergs - has occurred during a time of rapid evolution of both remote-sensing platforms and sensors. The application of these new research tools to ice studies has generally been both piecemeal and sporadic, partly because the community of ice scientists has not kept up with the rapid advances in remote sensing technology and partly because they have not made their needs known to the space community. This paper seeks to help remedy the latter shortcoming. The remote sensing requirements for floating ice studies are given, and the capabilities of various existing and future sensors and sensor combinations in meeting these requirements are discussed. The desirable future sensors are also discussed from both the research and operational points of view.

1. INTRODUCTION

Within the last few years there has been a sharp increase in the variety of satellite and aircraft borne remote-sensing techniques that show promise for the investigation of the floating ice covers of the world's oceans, lakes, and rivers, as well as shelf ice, and icebergs. Associated with this increase in technical capability there has also been an increase in actual remote sensing studies of ice-related problems (Campbell et al., 1975a; Page and Ramseier, 1975b). Many of these studies are, of course, focused on examining what a given sensor can and cannot detect. Two examples are, the recent studies of the variation of microwave signatures with changes in sea ice type (Wilheit et al., 1972; Gloersen et al., 1973; Ramseier et al., 1974) or the investigation of whether or not the intensity of a SLR (side looking radar) return can be used to indicate if a lake is frozen to its bottom (Sellman et al., 1975). Other studies focus on using a specific sensor to study a specific type of problem; for instance, the use of LANDSAT imagery to study ice drift and deformation (Hibler et al., 1975; Ramseier et al., 1975) or the use of IR imagery to estimate sea ice thickness (Poulin, 1973, 1975; Gloersen et al., 1974b). Only a few studies have attempted to explore the potential of applying a set of varied remote sensing observations to resolving a problem, even though this approach obviously has great promise (Campbell et al., 1974). The major current example of trying to utilize an integrated, multi-sensor remote sensing program as a fundamental input to a coordinated investigation of sea ice dynamics is presently underway in AIDJEX (Arctic Ice Dynamics Joint Experiment) (Weeks and Campbell, 1975). It is our experience with this project that has convinced us that in the future the great majority of floating ice problems will require data to be collected simultaneously by a variety of satellite-borne sensors. After all, it is only by

remote sensing technology that the vast areas of the earth's surface covered by floating ice can be sampled on the large space-scales and short time-scales that are required.

2. REMOTE SENSING REQUIREMENTS

First we will discuss exactly what sort of information investigators who study the different types of floating ice would like to acquire and why. Here no attention will be paid to whether or not a remote sensing technique exists, or is likely to exist, that will collect the requisite data. A summary of the required information is given in Table 1 as an estimate of the desired time and space scales involved. In the following it should always be kept in mind that in most floating ice problems extremely large expanses of ice are involved. Therefore, one will always be faced with the compromise between obtaining enough data to give adequate estimates of the parameter of interest and collecting so much data that processing becomes an intolerable burden. This is particularly true when one considers the rapidity with which drifting ice may undergo significant changes.

2.1 Sea Ice

Of first importance is the need to sense the presence (or absence) of ice. This is essential for defining the edges of the ice pack which may assume rather complicated wave and stream type forms. It is also essential for sensing the presence of open leads within the pack. Although a horizontal resolution of 100 m is adequate to resolve most features along the ice edge, a higher resolution (taken here as 5 m) would be useful in delineating the many narrow leads/cracks that occur in the pack. At the present there are no studies of the relative contribution of narrow leads that are below the resolution of current satellite systems such as LANDSAT or NOAA, to the actual observed strains (see Fig. 1). It is certainly conceivable (but doubtful) that many small unobserved leads may be as important as the few large leads that show clearly in the imagery.

Once the ice/open water distinction is made, it is necessary to give estimates of the areal amounts of ice of different thicknesses. These data coupled with the amount of open water specify the ice thickness distribution (G). At the present time G is believed to be the most important single parameter that controls the rheological response of pack ice to external forces (Thorndike et al., 1975). In this sense the critical end of the ice thickness distribution is the thin ice portion since it is known that thin ice is invariably deformed before thicker older ice. In fact thick (3-4 m) multiyear ice is rarely deformed during ice drift. Therefore in Table 1 the recommended vertical resolution of 10 cm is particularly applicable to the determination of ice thickness in thinner ice (< 1 m). In thicker ice a vertical resolution of 25 to 50 cm would be adequate. The spatial resolution of 5 m is believed to be necessary in order to resolve the rapid local changes in thickness that are commonly caused by ridging and by the refreezing of newly formed leads. These striking thickness variations are well shown in Fig. 2 which presents a portion to the sonar profile of the underside of the ice in the Arctic Ocean obtained by the British submarine **HMS Dreadnought** during its March 1971 cruise to the North Pole (Swithinbank, 1972). For many purposes monthly estimates of G should be adequate in areas with a heavy ice cover such as the Beaufort Sea, inasmuch as it is now possible to calculate the changes in G from existing ice drift and deformation models (Thorndike et al., 1975). Therefore, the observed G values would primarily be used to calibrate and check the model calculations. However, there are areas of thin first-year ice such as the Bering and Baltic Seas where more frequent (weekly) measurements of G would be desirable. This is particularly true during time periods when the frequent passage of cyclones causes pronounced ice deformation.

Table 1. Desired Remote Sensing Capability for Floating Ice.

Parameter to be Measured	Maximum Desired Resolution (H = horizontal, V = vertical)	Minimum Usable Resolution	Frequency of Observations
1. Sea Ice			
Presence or absence of ice	5 m (H)	100 m (H)	Daily
Ice thickness	10 cm (V), 5 m (H)	1 m (V) 25 m (H)	Weekly
Roughness of upper and lower surfaces	10 cm (V), 1 m (H)	10 cm (V), 5 m (H)	Monthly
Ice type characterization	5 m (H)	100 m	Monthly
Ice temperature	10 cm (V), 5 m (H)	10 cm (V), 100 m (H)	Weekly
Ice salinity	10 cm (V), 5 m (H)	10 cm (V), 100 m (H)	Weekly
Ice and lead patterns	5 m (H)	100 m (H)	Daily
Ice concentration	5 m (H)	0.1-25 km (H)	Daily
Ice drift	± 50 m	0.1-25 km (H)	Daily
Ice deformation	± 0.1%	± 0.1%	Daily
Surface albedo			
regional	10 km	100 km	Weekly
local	100 m	1 km	Weekly
Surface heat flux			
regional	10 km	100 km	Daily
local	100 m	1 km	Daily
2. Lake Ice - same as above, except for salinity which is not applicable.			
3. River Ice - again the factors are similar to those given for sea ice (deleting ice salinity). Because of the extremely dynamic nature of many river ice covers, ice roughness should probably be measured weekly.			
4. Shelf Ice			
Ice thickness	5 m (V)	Same	Yearly
Roughness of upper and lower surfaces	upper 10 cm (V), 1 m (H); lower 1 m (V), 10 m (H)	Same	Yearly
Ice type characterization	5 m (V), 100 m (H)	5 m (V), 0.1-25 km (H)	Yearly
Ice temperature	5 m (V), 100 m (H)	5 m (V), 0.1-25 km (H)	Monthly
Ice salinity	5 m (V), 100 m (H)	5 m (V), 0.1-25 km (H)	Yearly
Crevasse patterns	100 m (H)	Same	Yearly
Ice motion	± 10 m (H)	± 100 m (H)	Quarterly
Ice deformation	± 0.1%	Same	Quarterly
Surface albedo	10 km (H)	100 km (H)	Quarterly
Surface heat flux	10 km (H)	100 km (H)	Daily
5. Icebergs			
General dimensions	5 m	Same	Weekly
Location	1 km	Same	Daily
Any geometric or internal characteristics that would allow one to trace the motion of specific icebergs		Same	Daily

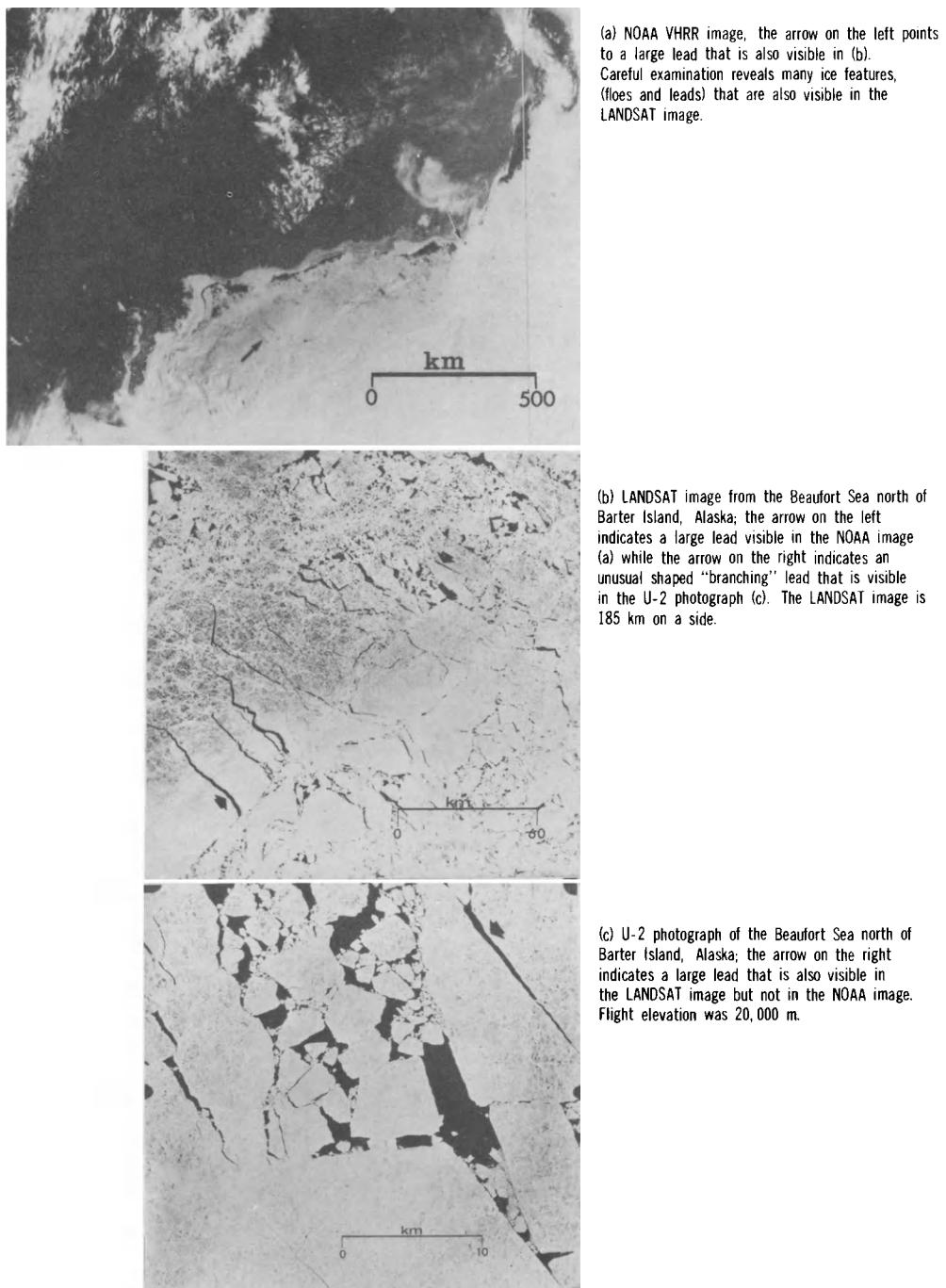


Fig. 1. NOAA, LANDSAT and U-2 imagery obtained over the same ice area in the Beaufort Sea at essentially the same time on June 21, 1974.

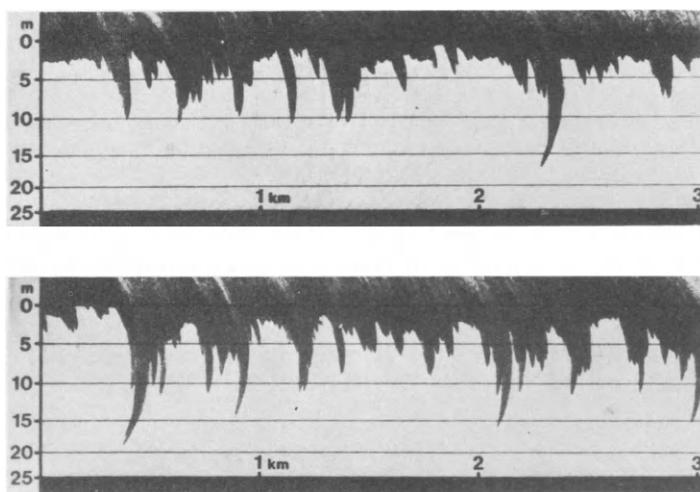


Fig. 2. Sonar profile of the underside of the Arctic pack ice obtained from **HMS Dreadnought** during the North Pole cruise in March 1971: both profiles show the ice canopy at 83° N, 06° E (Swithinbank, 1972).

If sea ice were in local isostatic equilibrium then accurate measurements of ice thickness coupled with a knowledge of the density of the ice would specify the roughness of both the upper and lower ice surfaces. However it is known that in the vicinity of pressure ridges pronounced local deviations from isostatic equilibrium are common (Weeks et al., 1971). Therefore, it is desirable to measure the roughnesses of both ice surfaces independently so that any changes in these values can be incorporated into changes for the estimated surface drag coefficients for wind and water. In ridged areas the changes in roughness are striking. Current information (Hibler, 1975) suggests that although a vertical resolution of 10 cm is desirable in measuring ridge sails, a resolution of 1 m is probably quite adequate for ridge keels. One problem associated with adequately measuring roughness, is that currently available techniques (laser, sonar) provide profile measurements, while, in fact, what is needed is the characterization of the roughness of broad areas. Fortunately during most periods of the year roughness is not believed to change rapidly. The principal exception to this is probably the early portion of the melt season when ablation rapidly rounds the jagged sails of first-year ridges.

It is also desirable to be able to categorize sea ice depending upon its physical characteristics. The most obvious sub-groups here would be first-year ice (thickness < 2 m, 5 to 15‰ salinity) and multiyear ice (thickness between 2 to 4 m, 0 to 4‰ salinity). This pronounced difference in salinity is caused by the extensive brine drainage and flushing of nearly fresh water through the ice during the summer melt period. Therefore multiyear ice and multiyear pressure ridges are much stronger than first-year ice because of the lower brine volume resulting from the lower salinity. The first-year/multiyear separation is also a useful one in that multiyear ice to a large degree forms the "undeformable" matrix within which first-year ice grows and is deformed.

Even better than the first-year/multiyear ice distinction would be to measure directly some pertinent physical characteristic of the ice, either as an average value, such as the average salinity, or even better as a profile. In thin ice a 10 cm vertical resolution of the property being profiled is desirable while in thicker ice a 50 cm resolution would be adequate. The most important single parameter that could be measured to specify the mechanical, electrical and chemical characteristics of sea ice is the brine volume profile which in turn is specified by the salinity and temperature profiles (Weeks and Assur, 1969). Measuring the temperature profile or at least the ice surface temperature is also very important in estimating the ice growth rate and the surface heat balance of the ice. The principal problem in relating variations in the snow surface temperature to variations in ice thickness is caused by the snow cover on the ice. Inasmuch as there is no simple relation between time and snow thickness, one cannot in general use surface temperatures as determined by

satellite-borne radiometers to make estimates of the temperature of the underlying ice without independent information on the characteristics of the snow cover (Poulin, 1975). The principal exception to this is found in thin, newly formed ice which because of the brief period between its formation and the time of observation usually has not accumulated an appreciable snow cover. Fortunately it is this thin ice and the areas of open water that are of prime importance in specifying the thermal exchange between the ocean and the atmosphere in the Arctic. Note that in Table 1 two different scales are designated as of interest in studying both the surface heat flux and the surface albedo: the 10 km or larger scale which would be of interest for regional studies, and the 100 m scale which is needed if one wishes to examine specific uniform ice types and/or conditions. The frequency of measurement required for either meso- or macroscale heat flux studies would be daily because of the possibility of the rapid opening or closing of systems of leads.

Information is also needed on the drift and the deformation of the pack as well as on the general geometry of the floes and leads. Typical drift velocities within the pack are roughly 5 km/day (Dunbar and Wittmann, 1963). The largest values (up to 20 km/day) are invariably found near the free edge of the pack where no lateral constraint exists on the ice movement. Strains within the ice as large as several percent are commonly noted, and are associated with the passage of a cyclone through or near the area of interest. Associated with these strains differential motions along highly oriented systems of leads are commonly observed. Because one system of leads may close (or become inactive) while another system rapidly forms, it is desirable to have information on these changes collected at least daily.

Finally a parameter of particular importance to shipping interests is the state of divergence in the ice pack - whether it is increasing or decreasing. Converging ice is difficult or impossible to negotiate even with an ice breaker, whereas diverging ice presents little impediment to a properly designed vessel.

2.2 Lake and River Ice

As noted in Table 1, the problems encountered in the study of lake and river ice are quite similar to those encountered in sea ice. This is particularly true in fresh water bodies in which a continuous fixed (fast) ice cover does not form and the ice continues to drift during the period of ice growth.

The principal differences, when contrasted with sea ice, are that the ice that is formed does not have an appreciable salinity. However, pronounced layers of entrapped air bubbles that can serve as scattering centres are quite common. It also should be remembered that inasmuch as the dielectric loss is lower in fresh water ice than it is in sea ice, electromagnetic radiation can penetrate through quite thick layers of fresh ice. Because of this the internal properties of the ice are important in determining the nature of the return signal when radar or microwave techniques are used. Details of which specific aspects of the ice are important are, however, many times poorly understood.

Two particular problems that are unique to rivers are the occurrence of frazil ice crystals in turbulent water and ice jams at constrictions in streams. It would, of course, be of interest to be able to sense both of these phenomena. It should be remembered that although the resolution requirements for most lake ice studies are similar to those for sea ice, many river ice problems require a significantly higher resolution since as it is the constrictions in a river, which may only be a few tens of meters wide, that are commonly of interest.

2.3 Shelf Ice and Icebergs

In general the information that is required in the study of ice shelves is surprisingly similar to that required for pack ice. The time scales are, however, quite different in that, with the exception of surface heat flux which would be required daily, most observations are required either quarterly or yearly. This reduces the amount of data processing. It should, however, be noted that although the items of interest are similar, the values that would be measured are quite different as are the processes involved (Swithinbank and Zumberge, 1965). For instance, ice shelf thicknesses vary from 100 to 500 m; the small scale roughness of the upper shelf surface is primarily the result of aeolian processes as contrasted with pressure ridges; the ice itself is primarily glacier ice or superimposed snow ice; and the salinity is due to lateral percolation of brine from the seaward edge of the shelf or from crevasses that permit the entry of sea water. Typical ice velocities on Antarctic ice shelves are 500 m/year. There are also pronounced variations in these velocities associated with the presence of faster flowing ice streams within the shelves.

Icebergs are primarily of interest as hazards to be avoided by shipping although it is conceivable that at some future time they may be utilized as a fresh water source (Weeks and Campbell, 1973). It is necessary to keep track of icebergs accurately within shipping lanes until melting has reduced them to quite small sizes. Because iceberg drift in the open ocean can be rapid it is not necessary to know their positions more accurately than 1 km at any given time of sensing. If there are features that allow the identification of specific icebergs, then their paths can be studied and analyzed in terms of the different forces that are involved.

3. SENSOR CAPABILITIES

Here we will briefly discuss the capabilities of a number of satellite systems that are either currently operational or presumably will be so in the near future. When possible we will also give references where more detailed information can be found. The characteristics of the sensors that are discussed are summarized in Table 2.

3.1 Visible and Near Infrared

3.1.1 LANDSAT. LANDSAT (formerly known as ERTS) imagery is of cartographic quality on a scale of 1:10⁶. Each image is 185 km on a side. The Multi-Spectral Scanner (MSS) sensor system collects information in 4 spectral bands ranging from visual (0.5 μm) to near IR (1.1 μm) with a stated surface resolution of 100 m. Experience has, however, shown that high contrast linear features such as leads with widths significantly smaller than this are commonly visible in the imagery. For instance, compare the LANDSAT imagery shown in Fig. 1 with the U-2 imagery shown in the same figure. The imagery reveals a wealth of sea ice information and shows quite clearly the distinction between water and ice, the different classes of thin ice (< 50 cm) as opposed to thicker ice, the presence of melt ponds, of brash ice, and in some cases the difference between old and first-year ice (Barnes and Bowley, 1974; Barnes et al., 1975). Because of its scale, resolution, and map-like accuracy LANDSAT imagery is ideally suited for studying lead patterns, ice drift and deformation. It is now being used as the basis for several such investigations (Hibler et al., 1974; Nye, 1975; Rothrock and Hall, 1975). There are, of course, a number of common ice features that cannot be distinguished on LANDSAT imagery (ridges, hummocks, rafted ice, individual melt features, and in most cases old ice).

The drawbacks of the LANDSAT imagery are that the MSS is limited by cloud cover. It is also limited by darkness, although present indications are that highly useful data can be obtained at much lower light levels (sun angles) than those at which the satellite is normally operated. Also the latitude and longitude of the centre of each image as calculated from the satellite orbit have been shown to contain significant errors (up to 8 km difference between the real and the stated position). This induces uncertainties in the utilizations of LANDSAT to study ice drift at sites far from land, although average daily drift rates commonly exceed these uncertainties. Problems can also be encountered in calculating strains and vorticities because of the errors in estimating the orientation of the longitude lines. These problems have been discussed by Hibler et al., (1975) and a procedure for minimizing them has been developed.

The major drawback inherent in LANDSAT imagery is its lack of continuity. At the equator LANDSAT images exactly the same area once every 18 days. In the polar regions the orbital overlap permits coverage of the same site during 3 to 5 consecutive days (depending on the latitude) followed by a 13 to 15 day "holiday" before the satellite returns to the study area. When the days lost because of cloud cover are added to these "holiday" periods, LANDSAT coverage has, in fact, proven to be quite spotty. To obtain the optimum use of LANDSAT identifiable ice features must be located on sequential images. The long holiday periods make this identification difficult except during periods of exceptional weather which, of course, are not periods of storm passage when most interesting changes occur in the ice pack. Nevertheless, even with these drawbacks, LANDSAT imagery has, because of its cartographic quality and good resolution, proven to be the satellite system that is presently most useful for quantitative studies of ice drift and deformation.

3.1.2 NOAA. The NOAA satellite series has also produced a large amount of imagery that is useful in the study of sea ice (McClain, 1974(a) and (b)). These satellites operate at appreciably higher altitudes than do satellites in the LANDSAT series (1465 km above the earth's surface as compared with 900 km). The prime sensor involved for sea ice studies is a dual channel very high resolution radiometer (VHRR) which

TABLE 2. CHARACTERISTICS OF PRESENT OR PLANNED SENSORS PARTICULARLY USEFUL IN SEA ICE STUDIES AND THEIR VEHICLES

Vehicle	Sensor	Wavelength	Surface Resolution	Repetitive Coverage	Maximum Latitude	Restricted by	Features of Note	Status as of Sept. 1975
LANDSAT	Multi-spectral Scanning System (MSS)	0.5-1.1 μ m	80 m	Daily for 2 to 5 days, followed by a 13 to 16 holiday	81° N	Clouds/Darkness	High resolution cartographic quality	Operational
NOAA	Scanning Radiometer	0.5-0.7 μ m 10.5-12.5 μ m	3.5 km 8 km	2 x daily 2 x daily	Near Polar Orbit	Clouds/Darkness Clouds	- Thermal IR	Operational
	Very High Resolution Radiometer	0.5-0.7 μ m 10.5-12.5 μ m	1 km 1 km	Daily Daily		Clouds/Darkness Clouds	- Thermal IR	
DMSP	Line-scan Radiometer	0.4-1.1 μ m 8-13 μ m	0.6 or 3.7 km 0.6 or 3.7 km	4 x daily 4 x daily	Near Polar Orbit	Clouds/Darkness Clouds	- Thermal IR	Operational
Nimbus 5	Electronically Scanned Microwave Radiometer (ESMR)	1.55 cm (19.35 GHz)	32 km	2 x daily	Near Polar Orbit		All Weather	Operational
	Temperature-Humidity Infrared Radiometer (THIR)	6.5-7 μ m 10.5-12.5 μ m	22 km 7.5 km	2 x daily 2 x daily		Clouds Clouds	Thermal IR Thermal IR	
Nimbus 6	ESMR	0.8 cm (37 GHz)	25 km	2 x daily	Near Polar Orbit		All Weather	Operational
Nimbus G	Scanning Multichannel Microwave Radiometer (SMMR)	0.8-6 cm (5 bands, 10 channels)	16 to 87 km as a function of frequency	2 x daily	Near Polar Orbit		All Weather	1978
	THIR	See Nimbus 5						
SEASAT-A	Synthetic Aperture Radar (SAR)	22 cm (L band)	25 m	36 hours	75° N		All Weather (100 km swath)	1978
	Radar Altimeter	2.2 cm (K band)	1.6-12 km				All Weather	
	Microwave Scatterometer	2.2 cm	50 km				All Weather (400 km swath)	
	SMMR	See Nimbus G						
	Visible/IR Scanner	0.49-0.94 μ m 10.5-12.5 μ m	3 km 5 km			Clouds/Darkness Clouds		
Earth Observatory Satellites (EOS)	MSS		10 m	6-9 days at equator		Clouds/Darkness		1980s
	SAR		15 m				All Weather	
Close Grid Geodynamic Measurement Satellite System (CLOGEOS)	Determines differential motion of passive targets.		< 10 cm	?	Near Polar Orbit	Clouds		1980s

obtains information in both the visible (0.5 - 0.7 μm) and the infrared (10 - 12.5 μm) regions. The images produced are 2000 km on a side and have a surface resolution of 1 km. In the polar regions a given area is usually imaged once daily. These satellites also carry a scanning radiometer that gives a twice daily image of any location on the earth's surface in both the visible and IR but at a reduced resolution (8 km). It is also possible to enhance the resulting imagery by distributing the 16 distinguishable grey tones over the 30 degree temperature range usually encountered in the Arctic as opposed to the 100 degree range commonly utilized for meteorological purposes.

The advantages of NOAA imagery are that during both light and dark periods, it provides regular broad scale coverage of the polar regions on a scale at which many useful large ice features can still be distinguished. The disadvantages are that images cannot be obtained through a cloud cover, the resolution and scale are sufficiently large so that many interesting and important ice features are lost (for instance in Fig. 1 compare the NOAA image with the LANDSAT image and the U-2 photograph), and finally the projection of the image is such that quantitative drift and deformation measurements are difficult.

At the present, most sea ice studies based on NOAA imagery have been semi-quantitative analyses of the regional behaviour of pack ice areas such as the Beaufort and Bering Seas and the Southern Ocean (Ackley and Hibler, 1975; Streten, 1974; DeRycke, 1973). If the data could be displayed in a format that would facilitate quantitative ice drift and deformation measurements, it would definitely enhance the usefulness of the imagery to the sea ice community.

3.1.3 DMSP (Defense Meteorological Satellite Program). This satellite series which was formerly called DAPP (Data Acquisition and Processing Program) is operated by the U.S. Air Force for the U.S. Department of Defense. The principal purpose of the satellites is the observation of cloud patterns. The imagery is collected by both high resolution and very high resolution 2 channel radiometers operating in both the visual (0.4 to 1.1 μm) and the IR (8 to 13 μm) ranges. The satellites orbit at a slightly lower altitude (833 km) than satellites in the LANDSAT series and significantly lower than those in the NOAA series. Usually two DMSP satellites are in orbit at the same time which provides coverage of any spot on the earth's surface four times daily. The VHRR system which is of primary interest here has a spatial resolution of 0.61 km at nadir. The imagery is currently archived at the University of Wisconsin Space Science and Engineering Center. A detailed description of the satellite system is given by Dickenson et al., (1974). DMSP imagery provides a view of sea ice on a scale that is intermediate between that of LANDSAT and of NOAA. Several images are obtained each day, darkness is not a limitation, and the map format, while not as convenient as LANDSAT, is easier to use than NOAA. In addition there are a variety of options available for specially processing both the visual and the IR data, so that the imagery has a higher apparent resolution and the temperature (or albedo) scale is expanded relative to some arbitrarily set base value such as the freezing temperature of sea water. Figure 3 shows the ice north of Alaska presented in such an expanded format. Note the large amount of detail that is revealed about the lead structure within the pack. This type of specially processed data should be available on a routine basis in the fall of 1975. The principal limitations of the sensing system are that it does not penetrate clouds and does not possess as high a spatial resolution as might be desired.

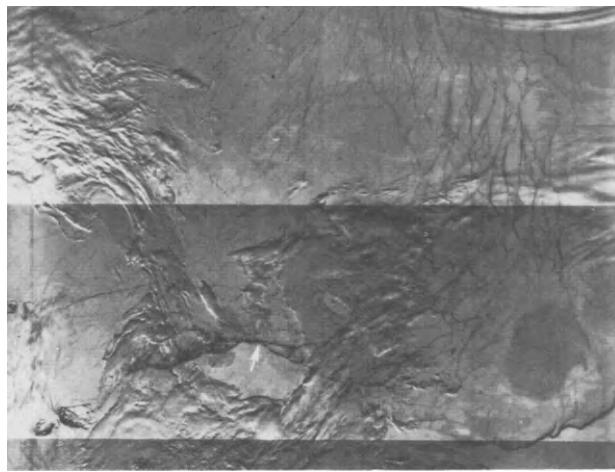


Fig. 3. DMSP image of Northern Alaska on 22 April 1975 (with Barrow indicated by the arrow) and the Arctic Ocean. The lead patterns which were only poorly visible on the normal VHR image were enhanced by spreading the 16 grey shades over a 25° K temperature interval bounded on the upper end by the freezing temperature of sea water.

3.2 Thermal Infrared

As discussed in the previous sections both the NOAA (and its predecessor ITOS) and the DMSP satellites have thermal IR systems operating respectively in the ranges of 10.5 - 12.5 μm and 8 - 13 μm and with resolutions of 1.0 and 0.6 km. Other satellites with IR systems of interest are in the NIMBUS series. The High Resolution Infrared Radiometer (HRIR) systems on NIMBUS 1, 2 and 3 operated in the 3.4 to 4.2 μm window which was only useful at night because of contamination by reflected solar radiation during daylight hours. The approximate spatial resolution at nadir was between 3.3 and 7.4 km. On NIMBUS 2 and 3 the HRIR system was joined by a five channel medium resolution system (MRIR) that operated in a scanning mode with a ground resolution of approximately 55 km at the subpoint. On NIMBUS 4 the HRIR and MRIR systems were replaced by a Temperature Humidity Infrared Radiometer (THIR) experiment that was designed to provide both day and night cloud and surface temperature information, as well as information on the moisture content of the upper troposphere and stratosphere. The THIR system had a ground resolution of 7 km for the temperature channel. NIMBUS 5 was equipped with a Scanning Microwave Radiometer with 3 detectors of which the thermal IR operated at a wavelength of 10.5 to 11.3 μm with a resolution of 0.8 km. NIMBUS G will also fly a scanning radiometer operating in a similar range (10.5 to 12.5 μm) with a resolution of 0.8 km.

While the list of satellites and radiometers is rather overwhelming, the trend is clear. The first IR sensors were useful only at night and had a low resolution, while the more recent versions operate successfully either during the day or night with increasing resolution. Because of their resolution limitations the IR systems on the early satellites were primarily useful for the mapping of gross ice boundaries such as the edge of the pack or the location of large polynyas (Barnes *et al.*, 1972). Even as the resolution of the IR systems improved to values of 1 km or less so that detailed observations of features within the pack became possible, little was done with the IR imagery except use it during dark periods as a replacement for visual imagery which was usually of an even better resolution. This was done by simply noting that open water or thin ice areas that appeared dark on visual imagery (low albedo) also appeared warm on the IR imagery. However, these studies indicated that the large reflectance changes that made the water-ice transition so clear in the visual range often corresponded to much less obvious thermal contrasts as recorded by the IR system. This was particularly true when the IR image was displayed by distributing the distinguishable grey shades over the 100 °C of interest for meteorological purposes. It was the problem of improving these sorts of contrasts that stimulated the

development of the special enhancement techniques that are now being applied to both the IR and visible imagery from the NOAA and DMSP satellites.

Admittedly Poulin's (1975) cautions concerning the difficulties introduced in the quantitative interpretation of snow (usually of unknown thickness and density) are quite correct and should be heeded by all considering studies in this field. Nevertheless it is our opinion that there will prove to be considerable profit in quantitative studies of regions of thin ice (with presumably thin snow covers) via IR techniques alone or, if possible, IR coupled with other techniques. Examples of such recent studies are Barnes et al., (1972) and Chase (1972). Figure 4 gives an example of the quality of IR images that are currently available from aircraft platforms (Dunbar and Weeks, 1975).

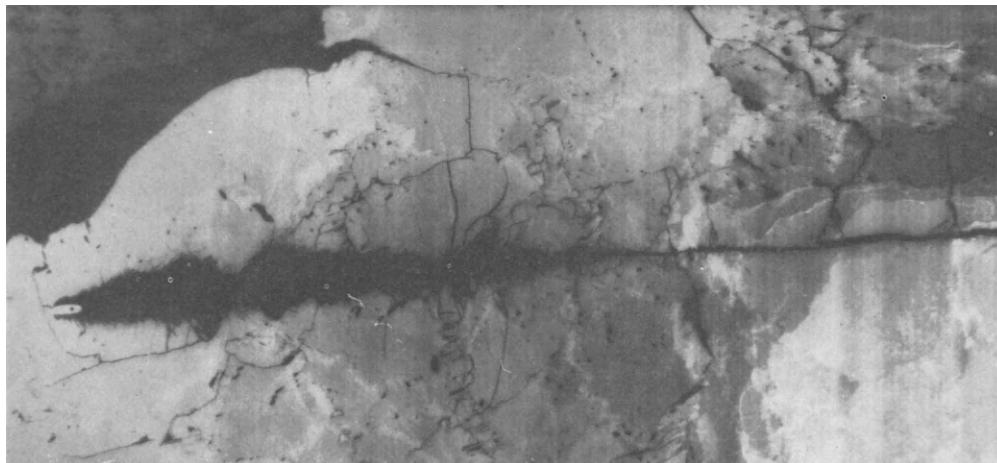


Fig. 4. IR image of **CSS Dawson** travelling south through the ice in the Gulf of St. Lawrence during January 1974. The warmest temperatures (black) are given by the open water and the ships smoke-stack. The thickest ice (white) was approximately 40 cm (Dunbar and Weeks, 1975).

3.3 Passive Microwave

Perhaps the most rapidly evolving area of remote sensing research in floating ice has been in passive microwave applications. Both aircraft and satellite-borne microwave radiometers have been used to observe a wide variety of ice types during the last five years. Although the time span is short, sufficient data now exist to cause us to conclude that such measurements will be a fundamental part of future ice observation programs.

3.3.1 Aircraft Microwave Observations. The most important series of aircraft flights which demonstrated the feasibility and usefulness of ice observations by means of passive microwave sensors were those that occurred during the NASA-AIDJEX (Arctic Ice Dynamics Joint Experiment) program and the U.S./U.S.S.R. BESEX (Bering Sea Experiment) joint program. A series of three AIDJEX pilot field experiments were performed during the springs of 1970, 1971, and 1972 in the southern Beaufort Sea. During each of these experiments, the NASA CV-990 "Galileo I" performed a variety of flights ranging in altitude from 150 m to 11 km. A wide variety of visual and infrared sensors were operated in addition to a 19.3 GHz imaging radiometer and 1.42 GHz, 4.99 GHz, 10.7 GHz and 37.0 GHz radiometers.

The 1970 data (Wilheit et al., 1972) showed that it was possible to distinguish sea ice from liquid water both through the clouds and in the dark. This finding was useful because it pointed the way to an "all-time" ability to observe leads and polynyas. These data also showed that strong microwave emissivity differences occur

on the ice surface itself. However, the lack of sufficient ground truth data prevented a determination of the reason for these differences.

The ground truth measurements and mesoscale microwave mosaic maps ($10,000 \text{ km}^2$) acquired during the 1971 AIDJEX-NASA experiments allowed Gloersen *et al.*, (1973) to show that the observed emissivity differences of sea ice at a frequency of 19.35 GHz are associated with the age of the ice, with multiyear ice having cold brightness temperatures ($\approx 210^\circ \text{ K}$) and first-year ice having warm ones ($\approx 235^\circ \text{ K}$). This was important because it suggested that passive microwave imagery could provide an all-time capability of distinguishing between old (thick) and new (thin) ice and of tracking ice motion as well as lead and polynya dynamics. Figure 5 shows a 19 GHz microwave image from that paper depicting the observed emissivities of a large ice area ($10,000 \text{ km}^2$) of the eastern Beaufort Sea. The multiyear floes, having an average thickness of 3-4 m, have cold (blue) brightness temperatures while the warm (yellow) temperatures are associated with first-year floes having average thicknesses of less than 2 m.

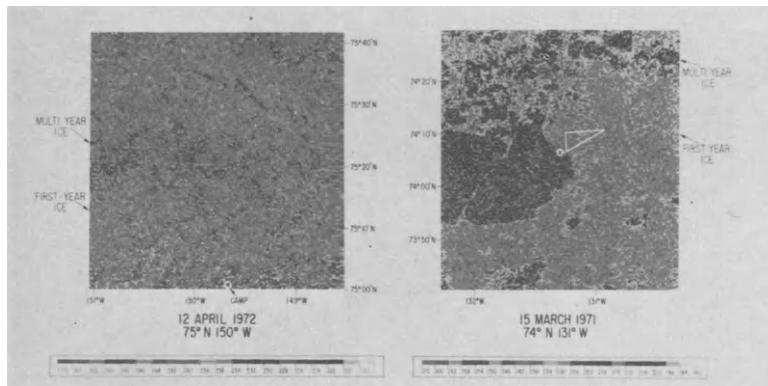


Fig. 5. False-colour ESMR image ($\lambda = 1.55 \text{ cm}$) collected on the NASA CV-990 aircraft over the AIDJEX areas, Beaufort Sea:

- (a) Mosaic obtained on 15 March 1971 from 11 km altitude with image centred at about $74^\circ 6' \text{N}$, $131^\circ 17' \text{W}$. The AIDJEX campsite is within the circle.
- (b) Mosaic obtained on 12 April 1972 from 11 km altitude with image centred at about 75°N , 150°W . The AIDJEX camp is within the circle at the bottom of the image.

During the 1972 AIDJEX-NASA field work two experimental innovations were made: a sled-mounted 13.4 GHz microwave radiometer was used to map the surface brightness temperatures of areas of various ice types where detailed crystallographic measurements had been made; and during each of the seven flights a mosaic microwave image was obtained of a segment of the southern margin of the ice cover adjacent to the north coast of Alaska (Campbell *et al.* in press). Analysis of the surface truth data by Meeks *et al.*, (1974) showed that surface salinity and pore size were primarily responsible for determining the microwave signatures of sea ice of various ages. It also showed that for mixtures of ice of different ages in which the floes are smaller than the radiometer footprint, the percentage of first-year versus multiyear ice can be estimated using microwave techniques. Figure 5 also shows a microwave image of the 1972 AIDJEX area in the western Beaufort Sea. A comparison of this image with one obtained a year earlier in the eastern Beaufort Sea reveals a markedly different morphological makeup of the ice in that no large multiyear floes were present. Campbell *et al.*, (1975a) noted that numerous small multiyear floes, smaller than the radiometer footprint of 500 m , were observed in the area, and by matching and averaging radiometer signatures of the various ice types they found that the 1972 AIDJEX area was made up of approximately equal amounts of multiyear and first-year ice.

The AIDJEX-NASA pilot experiments showed that microwave imagery can be used to study the morphology and dynamics of ice at the edge of the ice pack as well as ice within the pack. A time sequence of microwave mosaic images of the sea ice extending from Harrison Bay, on the north of Alaska, north into the Beaufort Sea for 350 km is shown in Fig. 6. The seven images cover the time from 4 to 23 April 1972. Three distinct zones of sea ice can be observed in each image: (1) a shore-fast zone made up of undisturbed first-year ice having a uniformly high brightness temperature; (2) a shear zone composed of three distinct ice types - multiyear floes with cold brightness temperatures, refrozen polynyas with brightness temperatures similar to those of shore-fast ice and a mixture of first-year and unresolvable chunks of multiyear ice having a brightness temperature between those of the two ice types; and (3) a mixed zone of first-year ice and unresolvable floes of multiyear ice with many refrozen polynyas. The extremely dynamic nature of the shear zone is clearly seen, with its width doubling in 19 days.

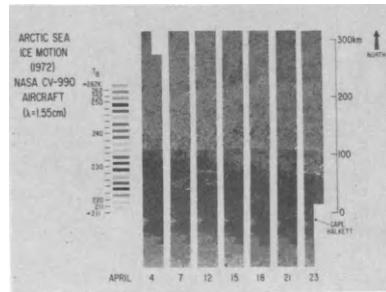


Fig. 6. Successive false-colour ESMR image strips collected between 4 and 23 April 1972 of sea ice of the western Beaufort Sea from Harrison Bay, Alaska to 74° N.

During the U.S./U.S.S.R. BESEX experiment in April 1973, the "Galileo I" obtained four mesoscale microwave images ($10,000 \text{ km}^2$) of parts of the Bering Sea ice pack at the time of its maximum extent. The pack ice in this region is quite unlike that of the Beaufort Sea in that it is all first-year ice, with thicknesses ranging from a few centimetres to a metre. In addition, this ice is subject to rapid dynamic changes because it is unbounded on its southern edge. Actually the Bering Sea ice pack more closely resembles the Antarctic ice pack than the ice pack of the Arctic Ocean. Figure 7 shows the 19 GHz image for 20 February 1973 (Gloersen et al., 1974b) with a N-S orientation of polynyas that had formed as the pack drifted rapidly southward during the preceding six days. The amount of open water indicated by this image is 16%.

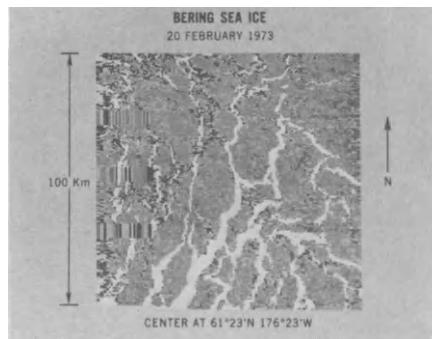


Fig. 7. False-colour ESMR image collected on 20 February 1973 on NASA CV-990 aircraft over the BESEX area, Bering Sea. The image is centred near 61° 24'N, 176° 24'W.

During April and August 1975, the "Galileo II" performed a variety of missions in conjunction with the Main AIDJEX experiment, which involves an array of four manned drifting stations and eight unmanned data buoys drifting from March 1975 until May 1976 in the Beaufort Sea. Similar flights will also take place in October 1975 and April 1976. The data from these missions is being used to help interpret satellite imagery of the ice pack and as input into numerical models of the pack.

Microwave images of lake ice acquired from aircraft (Ramseier and Gloersen, personal communication) in most respects resemble those of young sea ice in that the ice emits a warm signature and can clearly be distinguished from open water. The exception is that very thin lake ice emits various signatures because penetration is occurring and the Fabry-Perot effect takes place. For ice thicknesses less than five times the wavelength, the emissivity oscillates from close to zero to unity as thickness changes by $\lambda/4$, therefore a wide range of signatures can be emitted by thin ice.

3.3.2 ESMR. The Electronically Scanning Microwave Radiometer (ESMR) on NIMBUS-5 has provided the first large-scale synoptic views of the polar sea ice packs. These have been interpreted in light of the AIDJEX-NASA and BESEX findings revealing that both the morphology and dynamics of these major ice masses are more complex than had hitherto been assumed. Since its launch on 11 December 1973 NIMBUS-5 has provided a wealth of data on the variations of these packs at time scales ranging from several days to seasons. Figure 8 shows a pair of ESMR images of the Arctic at different seasons. The winter image of 11 January 1973 shows the ice canopy to be composed of two general types of ice: the principally multiyear ice covering the main portion of the Arctic Ocean, with brightness temperatures ranging from 209 to 223 K, and either first-year or first-year/multiyear mixtures with higher brightness temperatures covering the southern portion of the marginal seas (i.e. the Beaufort, Chukchi, East Siberian, Laptev, Kara and Barents Seas). Note that the multiyear ice extends far south into the eastern Beaufort Sea. By using a combination of ESMR and LANDSAT-1 images combined with the AIDJEX-NASA data, Campbell *et al.*, (1974) have shown that the Beaufort Sea ice cover is made up of large multiyear floes in its eastern sector while its western sector consists of small, fragmented first-year and multiyear floes.

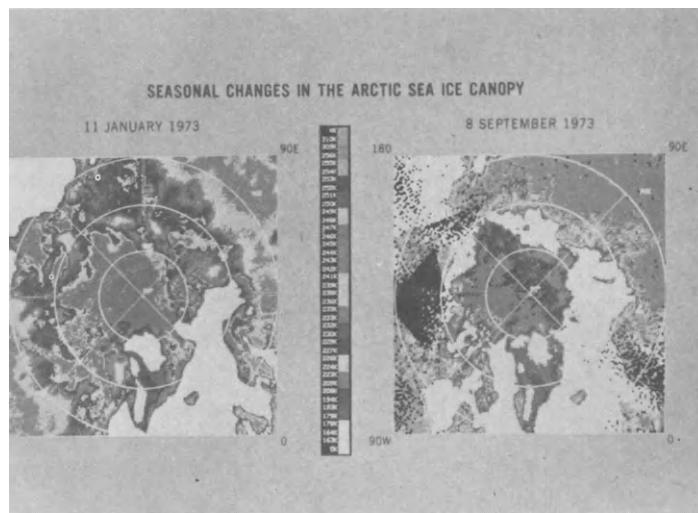


Fig. 8. False-colour ESMR mosaics collected over the northern hemisphere on 11 January 1973 and 8 September 1973 depicting seasonal changes in the Arctic sea ice canopy.

By the end of the summer melt season the ESMR image for 8 September 1973 (Fig. 8) shows that most of the ice that covered the marginal seas had melted. The area of the end-of-summer pack is approximately half of that of the winter pack. Analysis of these and other ESMR images by Gloersen et al., (1974a) reveals that not all of the first-year/multiyear ice mixture melted. Figure 9 shows an analysis of the near-maximum ice extent (10 February 1973) and near-minimum ice extent (9 September 1973) based on the ESMR images. The areas of the pack in which the first-year/multiyear ice mixture survived the summer melt are source areas for multiyear ice which is advected into the multiyear pack. This process thereby allows the ice pack to maintain a steady-state mass by balancing the volume of ice that is advected out of the Arctic Ocean by the East Greenland current. The analysis also shows that the summer pack extended closer to Svalbard than did the winter pack, suggesting a strong summer ocean surface current flowing southward in this region. Gloersen et al., (1974a) also show how the ice concentration can be inferred from ESMR images using a linear interpolation between the emissivity for open water and that of first-year ice. Note that the ESMR image for the Arctic in the winter of 1973 (Fig. 8) shows that the ice cover of the Bering Sea, which we have shown was made up of first-year and younger ice, has a cold radiometric signature similar to that of the multiyear ice of the Arctic Ocean. The reason for this is that the large (32 km) footprint of ESMR is covering a mixture of ice and open water, and the average brightness temperature appears cold because the water signature is much colder than that of the first-year ice.

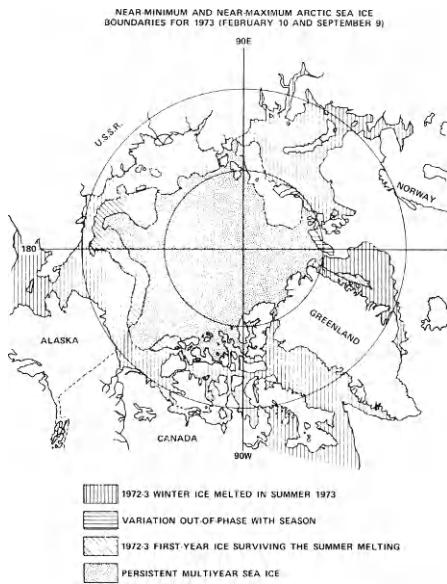


Fig. 9. Seasonal variations of sea ice in the northern hemisphere.

Another parameter that can be inferred from a series of ice concentration measurements is the state of divergence of the ice, since increasing concentration with time implies a converging ice pack and decreasing concentration implies divergence.

The ESMR images of the Antarctic are equally useful as those of the Arctic for studying large-scale ice morphology and dynamics. Figure 10 shows both a winter and summer image of the Antarctic. The winter image indicates that the ice cover is made up exclusively of first-year ice. By late in the following melt season (10 February 1973), the ESMR image indicates that approximately 80% of the ice cover has melted. This is shown in the analysis given in Fig. 11 in which the near-minimum (10 February 1973) and near-maximum (16 July 1973) ice extents are delineated. The observations at first seem to indicate the formation of multiyear ice where ice remains at the end of the melt season. Analysis of ESMR images in Austral Spring 1972 and Austral Winter 1973 by Gloersen *et al.*, (1974a) indicate that no multiyear ice was present, thus the regions where ice remains at the end of the melt season are accumulation regions in a system of steady depletion and replenishment.

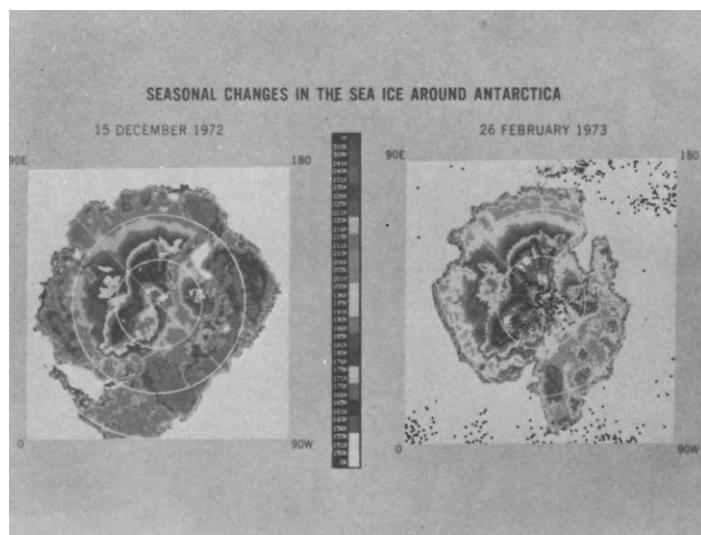


Fig. 10. False-colour ESMR mosaics collected over the southern hemisphere on 15 December 1972 and 26 February 1973 depicting seasonal changes in the Antarctic Sea ice canopy.

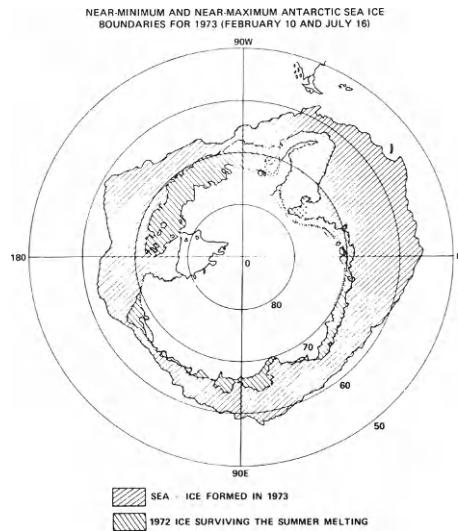


Fig. 11. Seasonal variations of sea ice in the southern hemisphere.

One striking feature appearing on all the ESMR images is that the edge of an ice pack is quite irregular, unlike the ice-pack boundaries shown in atlases. The same disparity between the images and atlases exists in regards to ice concentration, especially for the Antarctic where the images show very small concentrations in many areas. Of course when we realize that an ESMR view is essentially synoptic while the curves shown in atlases result from statistical smoothing of many seasons of picemeal data, we should not expect them to

agree. The ESMR images allow us to observe sea ice behaviour over large space scales and will thereby permit us to test dynamic rather than statistical models.

ESMR images of lake ice show the onset of freezing and melting but are of limited use for ice morphology and dynamic studies because of the large (32 km) footprint. The same can be said for river ice.

ESMR images of shelf ice, such as the Ross Ice Shelf, have been shown by Chang et al., (1975) to give valuable data on the accumulation rate of the surface snow cover. These measurements, when coupled to similar sequential ones of the ice cap from which the shelves flow, should give very valuable data on variations of snow accumulation patterns and amounts that are impossible to obtain by *in situ* surface measurements due to the immense logistical costs.

3.3.3 SMMR. The Scanning Multichannel Microwave Radiometer (SMMR) to be flown on SEASAT-A and NIMBUS-G has great potential application to floating ice research. The five imaging radiometers - 6.6 GHz, 10.69 GHz, 18.0 GHz, 22.05 GHz, and 37.0 GHz - will perform all the measurements of the 19.35 GHz ESMR on NIMBUS-5 and the 37.0 GHz ESMR in NIMBUS-6.

For sea ice observation the multifrequency capability of SMMR will help considerably in distinguishing between multiyear ice and first-year ice with unresolvable amounts of open water, both of which have similar radiometric signatures in the 19 GHz ESMR images. Multifrequency images of the same area will permit the observer to distinguish between these ice types because the brightness temperature differences between water and ice and between new ice and old ice vary as a function of frequency.

All forms of floating ice commonly have snow covers. The amount of snow is a prime determinant of the energy balance in the ocean-ice-atmosphere system. By comparing multifrequency, simultaneous images of floating ice the snow accumulation rates can be measured. Successive measurements from the onset of the accumulation season can give the total snow cover. Neither of these measurements have been made over most of the earth's floating ice.

3.4 Active Microwave Sensors

The active microwave sensors, such as side-looking airborne radars and scatterometers, operate normally at wavelengths between 0.08 and 0.25 meters. This permits the electromagnetic radiation to penetrate fog, rain, snow, darkness, and cloudy weather, i.e. they are all-weather, day and night systems.

The parameters that the radars sense on the ice surface are its dielectric properties and roughness. A certain amount of electromagnetic radiation is reflected from the surface, and the remainder penetrates the ice. The penetration of the radar wave into the ice is limited by the salinity of the ice and by discontinuities such as air bubbles, brine pockets and channels near the surface.

3.4.1 Side-Looking Radar (SLR). The SLR transmits electromagnetic radiation from an antenna mounted on an aircraft, with its beam looking in a fixed sidewise direction at right angles to the aircraft track. Scanning is achieved by the forward motion of the aircraft so that the radar image obtained is in a cartesian format relative to the aircraft flight path. The intensities of the received radar echoes are usually recorded on a continuous strip of film.

There are two basic SLR systems which differ markedly in the image resolution obtained. The simplest type is the non-coherent or real-aperture SLR, in which the azimuthal, or along-track, resolution is determined by $\lambda R/L$ where R is the range, L is the aperture of antenna and λ is the wavelength. The range resolution, or across-track resolution, is determined by the pulse duration or band width. To obtain a higher azimuthal resolution the antenna beam width would have to be decreased, resulting in a much longer antenna, which would make it impractical for mounting on an aircraft, or the wavelength or range could also be changed.

It is also possible to use the motion of the aircraft itself to generate a long "synthetic" antenna. This second, more complex and costly type of radar is the coherent synthetic aperture system (SAR). By illuminating a small ground element by many successive radar pulses and recording the phase and amplitude of the returns, one may, after suitable processing, obtain an image of the ground element with a much higher resolution.

The best possible along-track resolution for a focused SAR is independent of range and decreases as the real antenna is made shorter. For an unfocused SAR the resolution has a quadratic dependence on λ and R ($\approx \sqrt{\lambda R}$). Even though the theory predicts that the resolution is half the real antenna length, this resolution is limited in practice by unavoidable perturbations in the flight path of the aircraft which restrict the time over which the signal phase history can be processed to produce the image.

Several studies of the application of such radars to the investigation of floating ice that have been reported in the literature have been reviewed by Page and Ramseier, (1975) and Campbell et al., (1975a). Since these papers have been written, the main AIDJEX experiment has commenced, yielding new and interesting results. Figure 12 shows a real aperture SLR image of the AIDJEX 1975 Big Bear area. The camp can be identified as the black area near the centre of the image. Rough features appear dark in this image while smooth areas are light. The light linear features represent refrozen leads. All leads are covered with varying thicknesses of grey ice. Just north of and adjacent to the camp, the white linear feature running east-west represents the runway, consisting of 1.6 m thick first-year ice. The camp is located on a small multiyear ice floe having an average thickness of 3.5 m. The dark lines within this floe represent ridges having a height of approximately 2 m and a width of 3 m. The large lead which runs from northwest to southeast consists of grey ice with a thickness of 0.7 m. One aspect which is striking in this image concerns the backscattering coefficient σ . For example, the radar returns (grey scale) for the multiyear ice on which the camp is located and the large lead to the west of the camp are very much alike. If it would not be for the geometry of the ice, one could not distinguish first-year from multiyear ice. This is of great concern, since one of the future goals is to be able to process digitized SLR data automatically to obtain percentages of first-year and multiyear ice. The wavelike pattern on the east side of the camp running north-south is due to the aircraft motion. There is no question, however, that the information content of this SLR image is significant. For example, there was a relatively thick snow cover (≈ 0.20 m) over the first and multiyear ice. The grey ice was snow-free. The geometrical features of ice floes, leads, ridges, rubble ice fields, and rafting in grey ice, are very distinct.

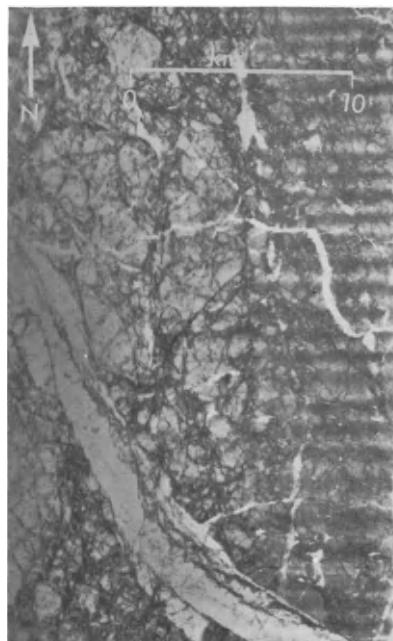


Fig. 12. SLR image of AIDJEX area taken by DND Argus Aircraft with a Motorola AN/APS real aperture radar during April 1975. The dark spot in the center of the image is the AIDJEX camp "Big Bear". Light areas indicate low returns and dark areas indicate high returns. Image was taken by the port antenna with the aircraft flying in a northerly direction.

Figure 13 shows a SLR image of the Mackenzie Delta area. The light tone on the south side represents shore-fast ice, followed in the north with the distinct shore-fast ice-pack ice boundary. The darker, rounded areas within the shore-fast ice represent stranded pieces of multiyear ice which are snow-covered. These pieces of multiyear ice are not visible in the aerial photography taken over the same area. The large floe in the centre of the image is multiyear ice. A lead running from the southwest to the northeast contains some grey-white ice (the darker area) and newly formed thin grey ice (light area) both of which are of varying thicknesses due to rafting. The northern part of the pack ice consists almost entirely of ridged first-year ice.

Figure 15 shows a number of frozen fresh water lakes in the vicinity of Barrow, Alaska. Sellmann et al., (1975) have shown that the difference in the backscatter from the tundra lakes indicates whether or not the lake is frozen completely to the bottom. The "dark" lakes (which here indicates low backscatter) do freeze completely to the bottom, whereas the "light" lakes (high backscatter) have an ice-water interface. The ability to differentiate via SLR between lakes that are fresh and not frozen to the bottom and lakes that are either frozen to their bottom and/or are brackish is of considerable economic, scientific and engineering interest. For example, one could determine if: a given lake was suitable as a year-round source of water; should be considered for possible fishing or for stocking with fish; or should be considered as a site where permafrost may be deeply thawed due to the thermal effect of the year-round presence of a water body.

SLR imagery obtained as part of the Canadian Beaufort Sea Project in the Mackenzie Delta area shows similar results to those obtained by Sellmann et al., (1975). Figure 16 represents two SLR images taken at 90° to each other showing the channels and lakes in the Delta. The dark areas in the main channel in Fig. 16a represent high backscatter, meaning the channel is not frozen to the bottom as compared to the lighter areas which are (note that high backscatter is indicated by a dark image in Fig. 16 and by a light image in Fig. 15). Where the channel meanders, the ice on the shallow, inside portion of the meander freezes all the way to the bottom, whereas on the outside of the curve where the water is much deeper, an ice-water interface exists. The thickness of the ice is of the order of 2 m. The high backscatter from the central part of the river does not vary very much since the incidence angle stays nearly constant. However, in Fig. 16b, which shows part of the same portion of the channel the backscatter varies considerably as a function of incidence angle. This becomes even more pronounced if one compares the backscatter of the river with the one of the lakes located at high incidence angles. It should also be noted that the large lake in Fig. 16b shows variations in backscatter of the floating part of the ice cover which presently cannot be explained satisfactorily.

An operational application of SLR to fresh-water ice reconnaissance is shown in Fig. 14. This imagery was obtained by the NASA Lewis Research Centre during February 1974 over the western end of Lake Superior as part of a demonstration program for the extension of the navigation season in the St. Lawrence River and the Great Lakes. Above the SLR image is an ice chart indicating the various ice types as interpreted from the radar image. It is interesting to note the large number of ship tracks extending from Two Harbours in the most direct line to the THIN-MED ice. Note that open water, which is indicated by the dashed lines in the bottom of the ice chart, cannot be distinguished in the SLR image from uniform thin ice in the Point Detour region. After processing of the SLR images and preparation of the interpretive chart, the composite was sent to a number of radio stations for transmission by radio link to vessels on the lakes (Vickers et al., 1973).

Figure 17 shows a SLR image of the lower part of the Peterman Glacier in northwest Greenland. Even though the glacier is covered with snow the lateral moraines are very distinct and can be traced to the tributary glaciers. There is also some evidence that the radar can detect subsurface structures in this type of ice, although it would not be possible to establish this without extensive ground truth observations. Additional subsurface information can be obtained by using a 25 cm radar operating at near vertical incidence as reported by Elachi and Brown (1975).

Another area which shows great promise for the use of SLR is the detection and tracking of icebergs. Figure 18 taken off Cape Atholl near Thule, Greenland indicates a large number of icebergs which appear as light spots surrounded by darker first-year ice and open water. Here the light areas indicate high backscatter. Accurate counts of icebergs can readily be made. The minimum size of iceberg which can be detected is of the order of 10 m. The height of large icebergs could be obtained based on the length of the radar



Fig. 13. SLR image of McKenzie Delta area taken by the DND Argus aircraft with a Motorola AN/APS 94D real aperture radar during April 1975. Image was taken by the port antenna with the aircraft flying in a northerly direction. Light areas indicate low return and dark areas high returns. Note the man made disturbances of the snow cover on the shore fast ice at the beginning of the image. Black lines across the image are time reference marks.

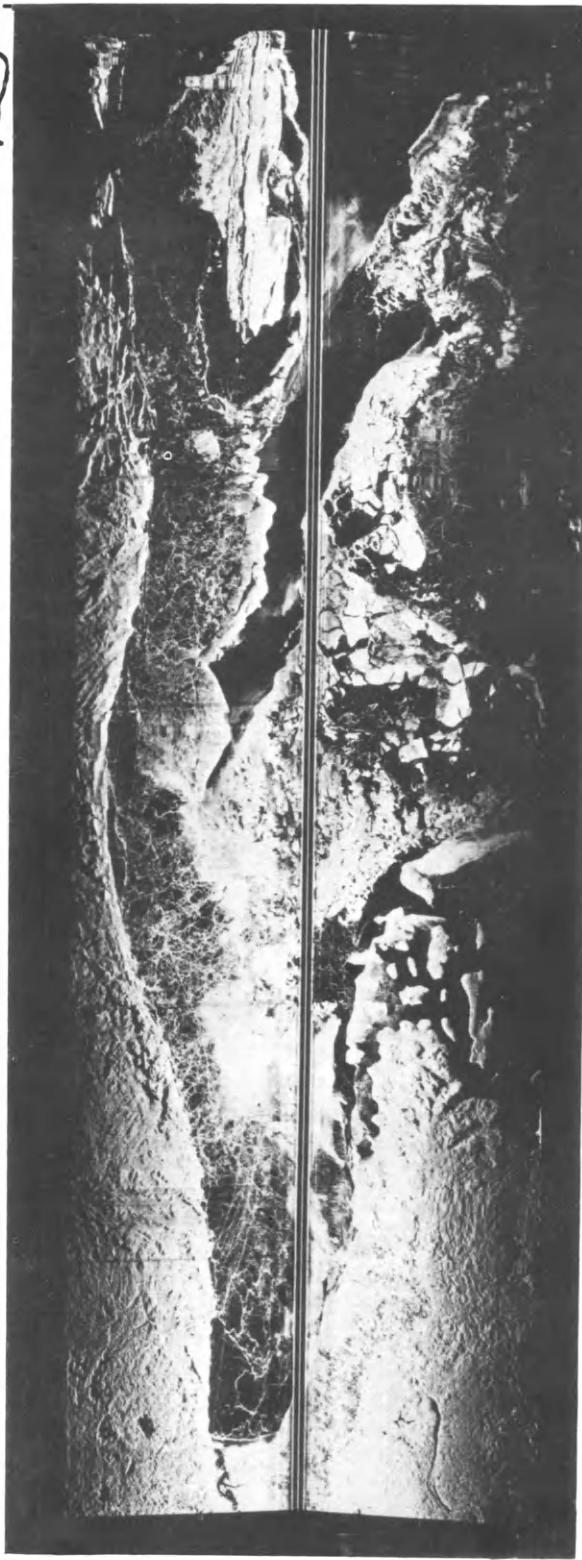
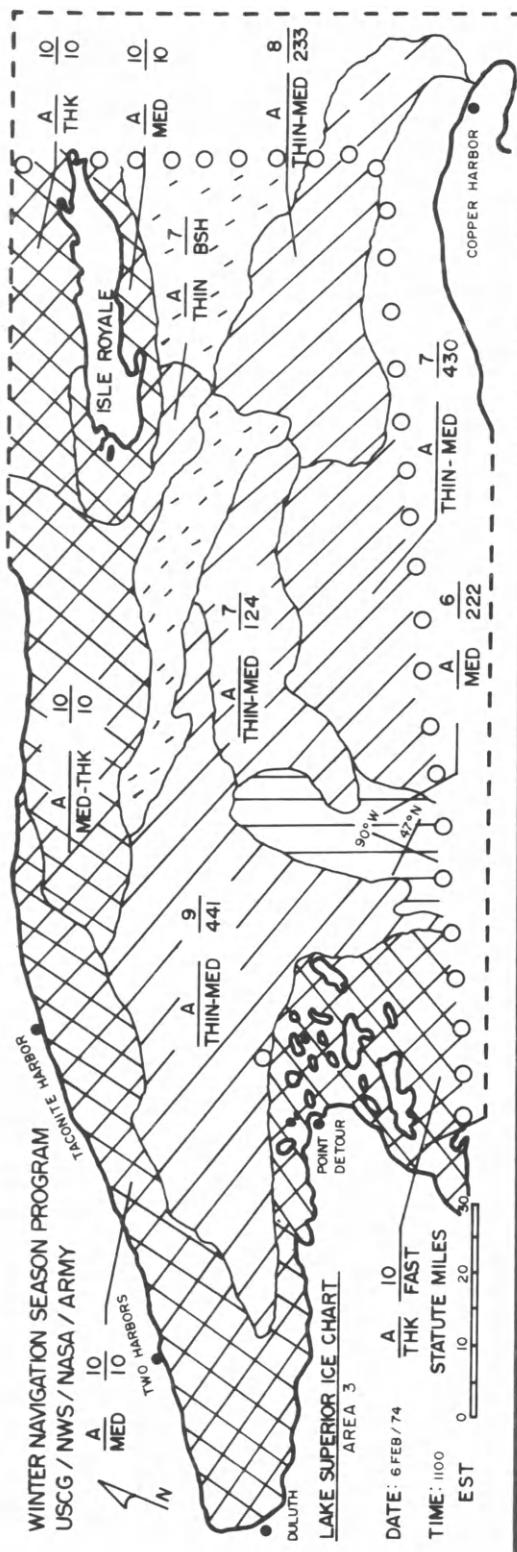


Fig. 14. SLR image of the western part of Lake Superior taken by a Mohawk aircraft with the Motorola AN/APG-94C real aperture radar on 6 February 1974. The dark areas indicate low returns and the light areas indicate high returns (Vickers et al., 1973).

Fig. 15. SLR image of area just west of Harrison Bay, Alaska taken by a USGS Mohawk aircraft with a Motorola AN/APS 94XE1 real aperture radar during April 1974. The image includes Naluakruk, Okalik and part of Teshekpuk Lakes. Note the differences in radar return from lakes which at the time the image was made were covered with 2m of ice (Sellman et al., 1975).

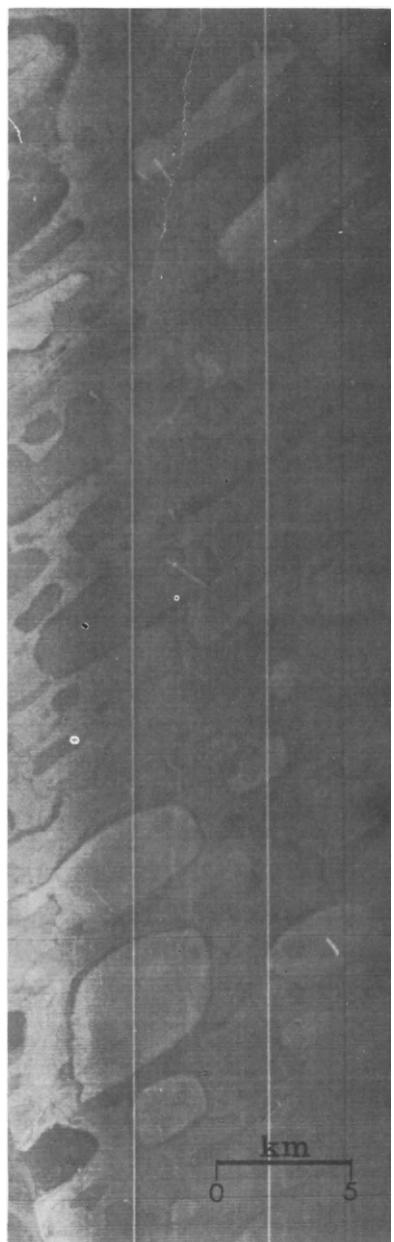


Fig. 16. SLR image with McKenzie Delta channels and lakes taken by the DND Argus aircraft with a Motorola AN/APS 94D real aperture radar in April 1975. Image was taken by the port antenna which was used to image the main channel twice at 90° to each other. The light areas indicate low returns and the dark areas indicate high returns.



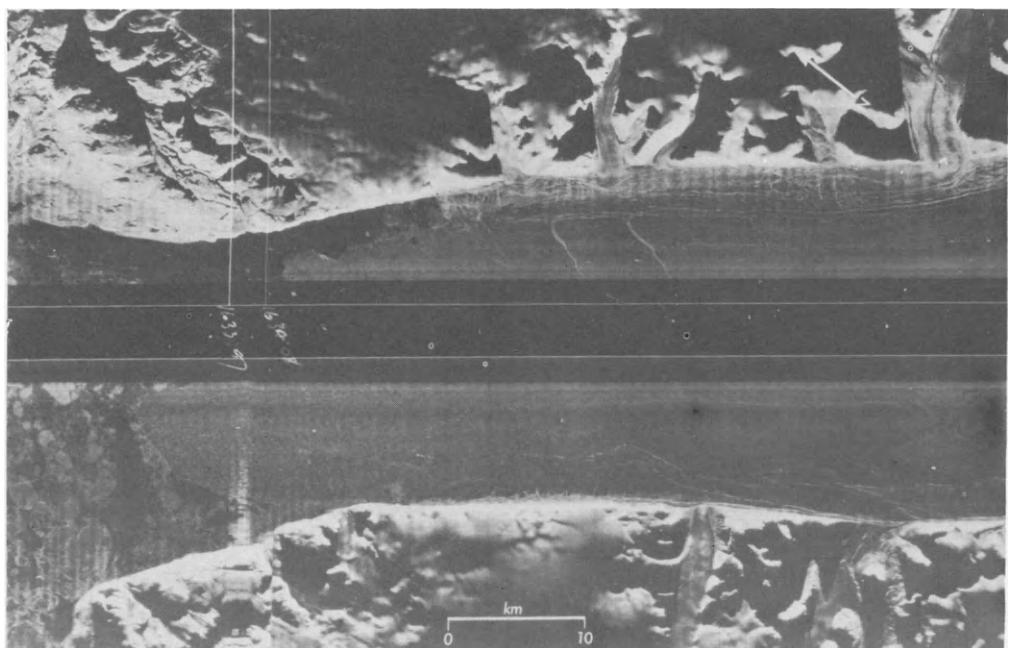


Fig. 17. SLR image of Peterman Glacier, northwest Greenland taken by the DND Argus aircraft with a Motorola AN/APS 94D real aperture radar in March 1975. Dark areas indicate high returns and light areas low returns. The aircraft was flying in an easterly direction.

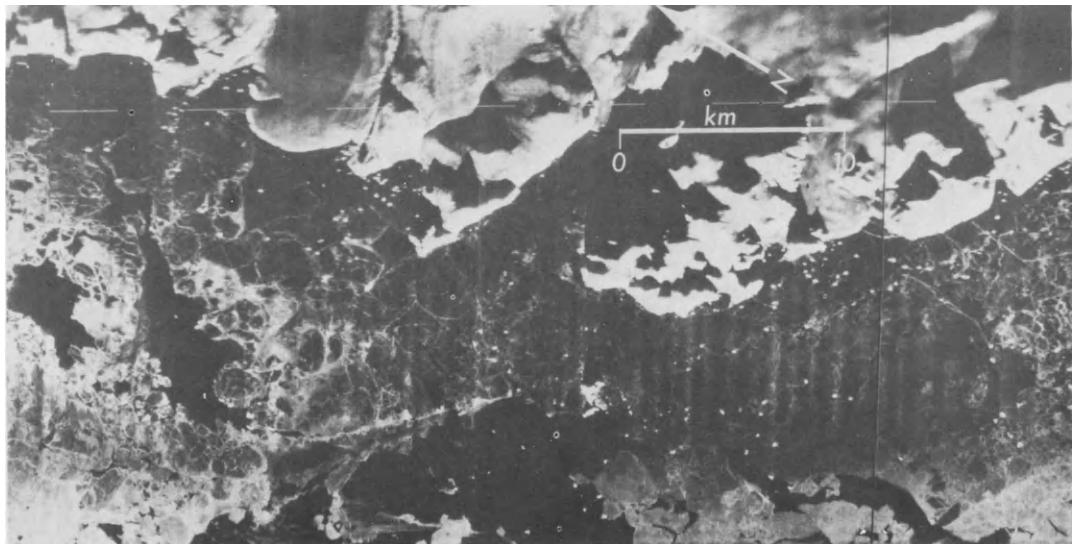


Fig. 18. SLR image of icebergs off the coast of Cape Atholl, near Thule, Greenland taken by the DND Argus aircraft with a Motorola AN/APS 94D real aperture radar in March 1975. Dark areas indicate low returns and light areas high returns. The image was taken by the port antenna while the aircraft was flying in a southerly direction.

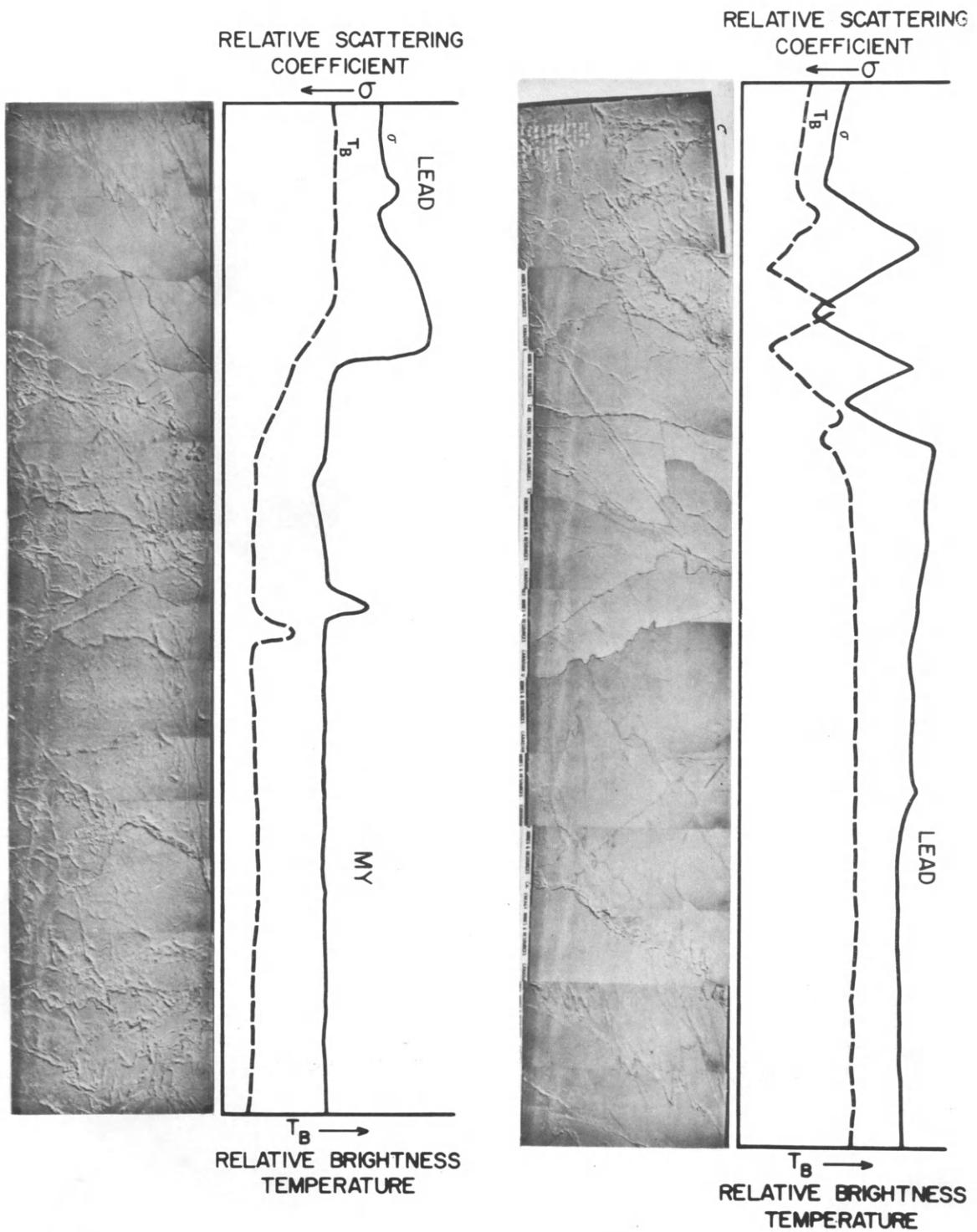


Figure 19A

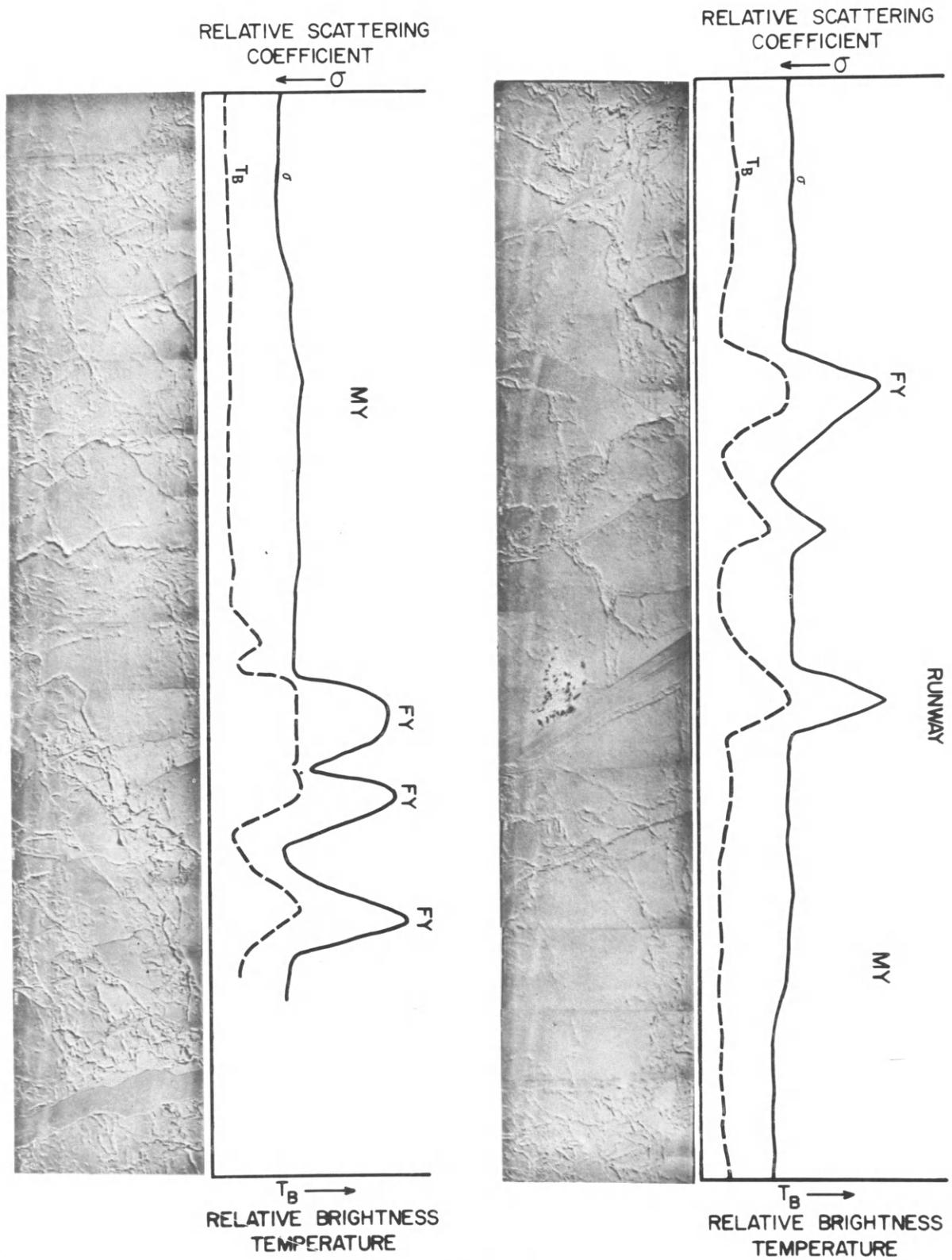


Figure 19B

shadow. The movement of these kinds of icebergs could easily be monitored to determine if they will become potential hazards for offshore drilling in the Labrador Sea or shipping and fishing in the North Atlantic.

3.4.2 Scatterometer. The microwave scatterometer is emerging as a powerful tool for determining ice type. The scatterometer is a calibrated down-looking airborne radar which can be used to measure the backscatter coefficient σ , as a function of aspect angle along a strip of terrain under an aircraft or spacecraft. Most scatterometers in existence today operate at roughly 13 GHz. Earlier work by Rouse (1969) and Parashar et al., (1975) demonstrated the potential of scatterometry for use in determining different ice types. Recent, preliminary analysis of scatterometer data from the AIDJEX experiment has confirmed these earlier findings. In fact, the discrimination between first-year and multiyear ice is much greater than previously reported. The proposed scatterometer for SEASAT will undoubtedly yield useful information concerning ice type distribution on the mesoscale in the Arctic Ocean.

The scatterometer forms a very powerful tool for use in determination of sea ice type in combination with SLR and passive microwave radiometer data. Figures 19a and 19b show a photo mosaic of the test line A at the 1975 AIDJEX camp taken during April. The corresponding profiles from the dual-polarised 13.3 GHz scatterometer (solid line) and the 37 GHz passive microwave radiometer (dashed line) are shown above. In this example the passive microwave radiometer was used to "ground truth" the scatterometer. As one can see there is an excellent correlation between the two sets of data. There is about a 60' difference in brightness temperature between first-year ice and multiyear ice and about 20 db difference in backscattering cross-section as observed by the scatterometer.

In Fig. 19a the first broad plateau corresponds to a large lead, identified earlier in the SLR image of Fig. 12, followed by multiyear ice, which continues in Fig. 19b. A small piece of first-year ice located just before the camp and another piece of first-year ice, which constitutes the runway, follow. Both are also clearly distinguished in Fig. 12. The profile continues with a large segment of multiyear ice and ends with three small areas of first-year ice.

3.4.3 Radar Altimeter. This sensor which is scheduled to be operated on SEASAT-A will determine the distance between the spacecraft and the ice-ocean surface to an rms precision of roughly ± 10 cm. The results are expected to be particularly useful in geodesy and in studies of significant wave heights. Inasmuch as the footprint of the radar is a 2×7 km spot, the results will not resolve individual roughness elements in the sea ice such as ridges and leads as is possible via an aircraft-borne laser profilometer. Nevertheless it may be possible to discern both spatial and temporal variations in the average (2×7 km) elevation of the surface of ice covered seas. These variations would presumably correlate with the intensity of ice deformation which has been shown to vary significantly between different provinces in the Arctic Ocean (Hibler, 1975). Other interesting studies that will be possible once SEASAT is launched concern the inter-relations between the results of the radar altimeter, the scatterometer, and the synthetic aperture radar.

3.4.4 SAR (Synthetic Aperture Radar). A number of countries (U.S.A., Canada, and countries of the European Space Agency) are carrying out programs with the aim of putting SAR's into space. The U.S. SEASAT-A program, the most advanced of these programs, is aimed at launching a SAR, complemented by a scatterometer, radar altimeter, multifrequency microwave radiometer and a visible/IR radiometer in early 1978. Canada plans to put a SAR with complementary sensors (payload not defined yet) into space by spring 1984.

The SEASAT-SAR will operate in the L-band range and is expected to have a resolution as low as 25 m and a swath width of 100 km. At the present time the proposed circular orbit at an altitude of 800 km will allow SAR imagery to be obtained as far north as 75° N with repeat coverage every 36 hours.

If current orbit proposals are not altered, a large part of the Arctic Ocean will not be imaged by the SAR. It would be highly desirable to change the orbit enabling the SAR system to image well into the Arctic Ocean. It is very likely that future SEASAT-type missions will be launched into a polar orbit. For example, the Canadian POLESAT, as it may be named, is intended for a polar orbit.

The quality of the SAR imagery is expected to be similar to that of the L-band airborne SAR obtained by Ketcham and Tooma, (1973) and Elachi and Brown, (1975). Even though the SAR resolution is higher, the information content pertinent to ice (surface and near-surface properties) in the X-band imagery appears to be greater. Because of the longer wavelength of the L-band SAR, the penetration (skin depth) is greater resulting in less scattering from the surface as compared with the X-band SLR. This is in accord with multifrequency passive microwave measurements made during AIDJEX 1975.

From the work accomplished so far and including the forthcoming results of the main AIDJEX experiment, some projections can be made of the kind of SAR product one may expect from SEASAT-A. Leads, ice boundaries, as well as a variety of ice features such as ridges and rubble fields should be readily identifiable. Major ridging of first-year ice should be clearly seen, but multiyear ice ridges will not be as distinguishable. What would be difficult to determine are ice types, particularly the difference between first-year and multiyear ice. The successive tracking of the identifiable features will however permit quantitative information on ice drift and deformation to be obtained on an all-weather basis. We believe that the successful operation of the SEASAT-SAR system will prove to be the first step in implementing an all-weather, operational sea ice reconnaissance system that is useful to marine operators during the times of nearly continuous cloud that are encountered during the summer shipping season in the Arctic.

4 SENSOR-ICE PHENOMENA COUPLING

4.1 Distribution and Boundaries

Excellent pictures have been obtained from space showing the development and the dynamics of the boundaries between pack ice and the open ocean. For instance Fig. 20a, which was obtained by the SKYLAB astronauts over the northeastern part of the Gulf of St. Lawrence on January 11, 1974 (Campbell et al., 1975b) shows a striking example of the ice plumes that occur to the east of Newfoundland and Labrador. The ice here is unconsolidated frazil which is kept from congealing by the continuous passage of waves through the plumes. These features appear to form characteristically at the edge of the pack during periods of high winds. The plumes can best be described as possessing a vortex type of structure. The small plumes that are composed of slush are elongated parallel to the surface wind. Figure 20b was taken on January 19 over the same general area. Again the most striking features off Newfoundland are the strongly developed plumes. Note the marked difference in the structure of these plumes compared to the ones on January 11, in that these plumes are elongated perpendicular to the surface wind direction (as deduced from the stratocumulus bands). The earlier plumes were composed of slush, while this orientation change indicates that the present plumes are composed of small, congealed pack ice floes with sizes below the resolution of the photograph.

A good example of the type of detailed information that is available, from a LANDSAT image on a clear day, on the distribution of open leads and thin ice areas within the pack is shown in Fig. 21. This image is interesting in that it contains a variety of thin ice as revealed by the varied grey tones in the image. This should be contrasted with the LANDSAT spring image shown in Fig. 1(b) which contains only thick ice (> 1 m) and open water. Because of the striking albedo changes that are observed between open water or thin ice, and ice thicker than say 50 cm, the structure of open and newly refrozen leads can many times be observed through thin cloud layers.

Lead structure as obtained from a high resolution scanning microwave radiometer on board NASA's Convair 990 over the Bering Sea is shown in Fig. 7 (Gloersen et al., 1974c). The leads are located near the edge of the southerly advancing Bering Sea ice cover on February 20, 1973. The ice concentration is 84%. The boundary between the ice (grey ice) and the water is very distinct and sharp in this imagery. In many cases along the edges of the leads one can see thin blue lines. These lines indicate that the brightness temperature is lower due to the averaging effect of water and first-year ice within the beam of the antenna. This becomes very distinct in the ESMR imagery obtained from NIMBUS-5 shown in Figs. 8 and 10.

From the previously presented examples of SLR imagery it is quite evident that boundaries within the pack are often very distinct, e.g. the leads around the AIDJEX Big Bear shown in Fig. 12. Unfortunately, in many cases it is not possible to distinguish open water from thin ice in ice infested waters. An example of this is seen in Fig. 18. The bottom centre of the SLR image represents open water and thin ice. The presence of

Fig. 20. SKYLAB-4 images taken by the astronauts over the Gulf of St. Lawrence.



(a) 11 January 1974, NE part of Gulf of St. Lawrence, Straits of Belle Isle. Note the frazil ice plumes parallel to the wind direction.



(b) 19 January 1974, NE part of Gulf of St. Lawrence, Straits of Belle Isle. Note the consolidated ice pack and ice plumes perpendicular to the wind direction.

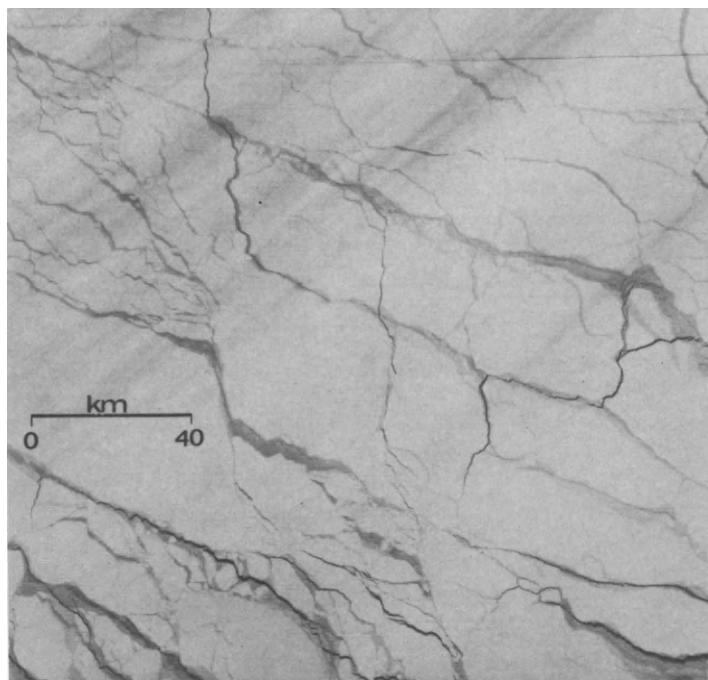


Fig. 21. LANDSAT (MSS-7) image of pack ice in the central Beaufort Sea ($77^{\circ} 52'N$, $138^{\circ} 04'W$, 30 March 1975). The image is 185 km on a side.

open water was deduced due to the motion of the icebergs, determined from two SLR images taken at 20 min. intervals. If thin ice had been present, the icebergs would have been either stationary or two thin lines would have been visible due to the rough edge caused by the iceberg moving through thin ice. It should be remembered also that SLR primarily senses surface roughness as opposed to ice thickness. Since, in principle, thick undeformed first-year may be as flat as skim ice in a newly formed lead, the identification of newly formed leads in SLR imagery is, to a considerable extent, a matter of pattern recognition.

4.2 Morphology

The ability to distinguish between open water, thin ice, first-year ice and multiyear ice is, of course, very important. For example, both thickness and physical properties differ greatly between first-year and multiyear ice. Generally speaking first-year ice is smooth on the surface while the edges of first-year floes and ice fields are rather angular. In contrast, multiyear ice is generally rough on the surface while the edges of the floes are rounded. This change in the geometric characteristics of the ice is due to the fact that the melt which takes place during the summer modifies the surface, and also that the general "grinding" interaction between floes which with time tends to round-off corners. These events are coupled with the formation of melt ponds and hummocks on the originally flat first-year ice surfaces. If this process is repeated over a number of years, the surface relief may become very pronounced.

Visual observations, aerial photography and to some extent SLR reveal these features permitting, for example, the preparation of maps which show the concentration of first-year and multiyear ice. As one can see, this is done more by circumstantial evidence than by direct observations of the material properties. The difficulty in distinguishing between such geometrical features increases depending on the amount of snow on the ice. A large amount of snow often will smooth out many of the surface features of multiyear ice making it difficult to distinguish multiyear from first-year ice.

IR scanners are not very useful in distinguishing thick first-year ice from multiyear ice. However, they can readily distinguish different varieties of thin ice, as well as thin ice from thick ice (Gloersen *et al.*, 1974c).

Passive microwave radiometers integrate the received electromagnetic radiation emitted from the surface over a relatively large area, therefore the resulting brightness temperature is generally independent of surface roughness.

The melt process during the summer not only modifies the surface features as mentioned earlier but also extensively changes the material properties themselves. Consequently the dielectric properties of first-year and multiyear ice differ significantly and the distinction between these two important ice types using passive microwave radiometers becomes quite clear. Figures 5, 6, and 8 represent good examples of this capability.

The ability of SLR and passive microwave radiometers to penetrate snow in most cases eliminates the camouflaging effect snow has on ice.

In theory SLR should be as effective in distinguishing first from multiyear ice as the results of the scatterometer shown in Fig. 19. Unfortunately, the art of processing SLR signals obtained over ice has not progressed to the point which will permit this. In most cases, SLR imagery is still being used in a manner similar to aerial photography.

To obtain an idea of the distribution of surface roughness in the Arctic Ocean, Hibler (1975) has determined, by using laser profilometer data, some regional ridge height distributions. Figure 22 shows three height distributions for the Northern Canadian Archipelago, the area near the North Pole and the Beaufort Sea. The distributions indicate that the trend is from heavily to lightly ridged ice respectively. It is feasible that the altimeter on SEASAT-A may also give quantitative broad scale information on this subject.

4.3 Thickness

Thickness of sea ice is probably the most important parameter for many transportation, engineering and climatological applications. To date, there is no remote sensing instrument available which can measure directly the thickness of grey, first and multiyear ice. However, there are some promising developments in

radar probes taking place (Bogorodskii and Tripol'nikov, 1974); Campbell and Orange, 1974; Chudobiak et al., 1974) which will be able to measure the time delay in an ice cover from which, combined with the dielectric constant, the thickness can be determined from a helicopter or a fixed wing aircraft.

As mentioned in the previous section IR techniques provide some possibilities for determining the thickness of thin ice up to 0.80 m depending on how cold the air temperature is (Gloersen et al., 1974c).

SLR imagery can be used to infer ice thickness via the identification of ice types but such a method suffers for the reasons given in the morphology discussion. On the other hand, passive microwave radiometers and scatterometers can identify ice types from which the thickness can be inferred. The accuracy of the ensuing result depends on the sophistication of the assumed ice growth model. At best it can give a representative average value over a large area (Meeks, et al., 1974).

In the case of fresh water ice and glacier ice the situation is somewhat different. Operational systems are currently available which will measure directly the thickness of fresh water ice. These have come to use primarily in the St. Lawrence Seaway and Great Lakes (Vickers et al., 1973; Chudobiak, et al., 1974; Page and Ramseier, 1975).

4.4 Properties

Figure 23 illustrates a typical example of the salinity, density and temperature profiles of first and multiyear ice. In first-year ice the salinity is usually high both near the surface and at the bottom of the ice sheet. The density is nearly constant, except possibly near the upper surface where in some cases, it is lower. The temperature profile is generally linear during most of the year. In multiyear ice the salinity increases with depth, starting with a salinity of near zero in the ice that is above sea level. The density is also lower in the top part of multiyear ice covers. The temperature behaves in a similar fashion as in first-year ice. In addition, there appears to be a pronounced change in the internal structure of the upper (above water level) portions of sea ice that occurs during the first melt season. At this time the columnar, crystallographically oriented structure produced by initial crystal growth changes over to an equi-axed structure showing a random crystal orientation.

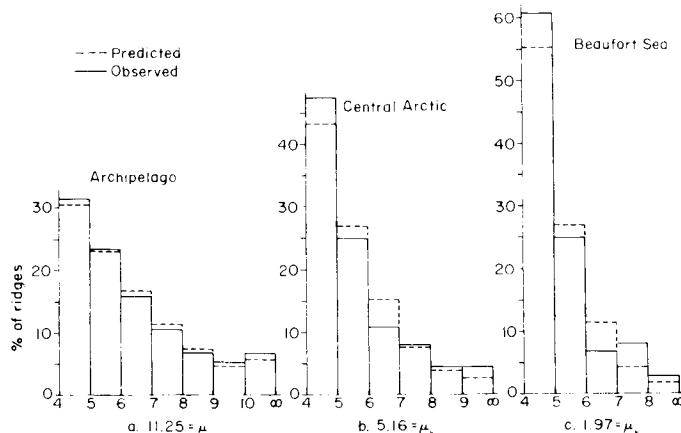


Fig. 22. Ridge height distributions taken in February 1973.

Each distribution was taken from a laser track approximately 40 km in length and the average number of ridges per kilometer above 4 ft (1.22 m) is denoted by μ_n for each distribution. The two parameter fit is indicated by dashed lines with the actual data being solid. Distribution a, was observed at approximately lat. 83° N, long. 85° W., b at lat. 87° N., long. 162° W, c at lat. 70° N., long. 139° W. The mean ridge heights for distributions a, b and c were 1.93 m, 1.70 m, and 1.57 m respectively (Hibler, 1975).

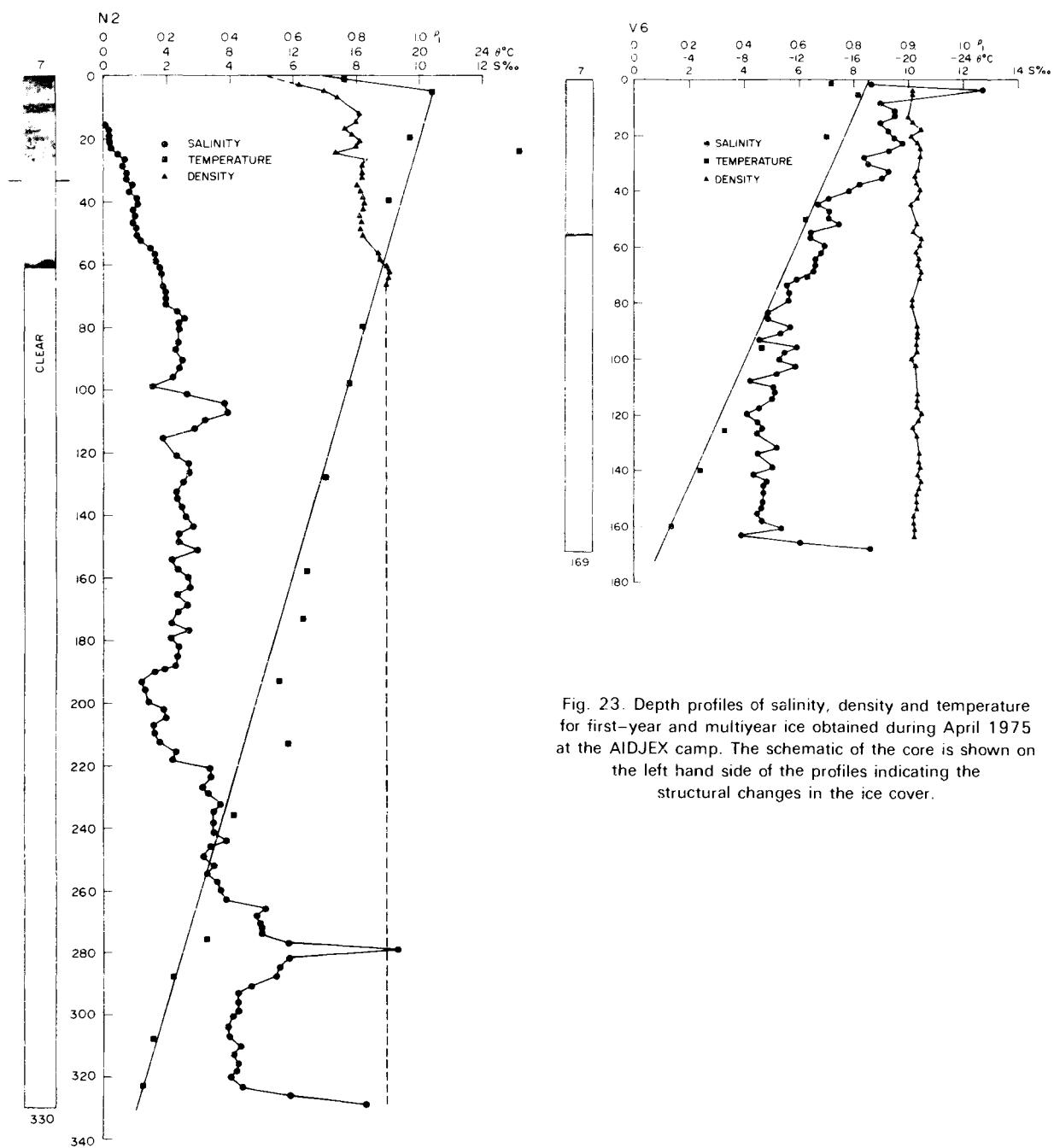


Fig. 23. Depth profiles of salinity, density and temperature for first-year and multiyear ice obtained during April 1975 at the AIDJEX camp. The schematic of the core is shown on the left hand side of the profiles indicating the structural changes in the ice cover.

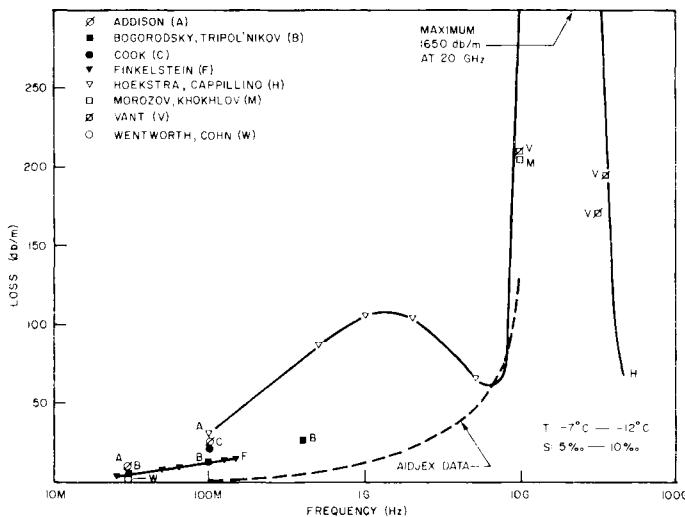


Fig. 24. Summary of the dielectric loss data for ice of a standard salinity and of a standard temperature. The dashed line indicates the trend of the 1975 AIDJEX results obtained by Vant (personal communication).

The dielectric properties also vary substantially from first-year to multiyear ice (Vant et al., 1974). Figure 24 summarizes the dielectric loss estimates obtained by a variety of investigators (Hoekstra and Cappillino, 1971; Bogorodskii and Tripol'nikov, 1973; Vant et al., 1974 and Vant et al., 1975 to name a few). The dashed curve summarizes the preliminary results obtained during the 1975 spring AIDJEX program (Vant, personal discussion). These new results indicate that the losses in first-year ice are much less than previously believed. This is a significant result because it will aid in the interpretation of microwave signatures significantly and suggests that it should be possible to design radar probes that measure the sea ice thickness directly.

4.5 Dynamics

Visible sensing systems such as LANDSAT have produced some highly useful results when applied over a large geographical area during an entire ice season (spring to fall) (Ramseier et al., 1975). However, the operational usefulness of such sensors is very restricted because of the limitations in daily coverage associated with the satellite orbit coupled with the cloud cover restriction. Passive and active microwave sensors on the other hand, are not limited by such restrictions and are capable of providing synoptic coverage. For instance, in Fig. 6 a number of features can be traced over several days. The ice motion averages about 4 km a day to the west. The overall motions are complicated by the alternating anticyclonic-cyclonic activity over the area during this period (Campbell et al., 1975a).

Due to the higher resolution of SLR imagery, similar but more detailed dynamic studies can be made. V.V. Bogorodskii and V.S. Lohschilov (personal communication), during the spring of 1973, made a SLR map of sea ice along the entire shipping lane north of the Eurasian continent. This map was made with the Toroz 16 GHz, real aperture system, and it appears that this system is now being used operationally for ship routing along the northern sea route. On-going studies using SLR imagery as part of the Canadian and USA continental shelf programs and AIDJEX are being pursued by the authors.

4.6 Snow Cover

In most cases, snow cover acts as a camouflaging layer over the ice. The usefulness of visible sensors becomes very limited once the ice is covered by snow cover. If the snow cover is not too thick, IR imagery is still useful in locating thin ice.

The advantage of active and passive microwave sensors lies in the fact that as long as the snow is dry, it is transparent to microwave radiation. All areas of thick ice shown in the SLR Fig.s 12-18 were snow covered. Similarly, the passive microwave images shown in Fig.s 5, 8 and 10 had snow on the ice. For the amount of snow encountered in the Arctic, the effective emissivity is not much affected at the shorter wavelengths. Preliminary results obtained during AIDJEX 1975 indicate that vertically polarized microwave measurements are nearly independent of the amount of snowpack (at -18°C) encountered compared to horizontally polarized measurements which are strongly affected. The effect of liquid water in the snow on the microwave emission from snow is striking. The emissivity at a wavelength of 8 mm increases from 0.78 to near unity as the liquid water content of the surface layer increases from 0 to 1% (Edgerton *et al.*, 1971). These effects are similar in both the vertical and horizontal polarizations. As the liquid water content of snow on multiyear ice increases during the melt season, the emissivity of the multiyear ice will approach the value for first-year ice. This has been observed in the ESMR images (Gloersen and Salomonson, 1975, Fig. 19). Fortunately, this does not represent a serious problem since ice type differences can be observed easily on images taken during the dry snow period. Also, it should be possible to map the progress of the melt season in the Arctic Ocean by studying microwave imagery.

5. SENSORS/TASK COUPLING

5.1 Research

The basic reason for acquiring remote sensing data on all forms of floating ice is to apply it to theoretical models, so as to achieve a cause and effect understanding that will ultimately lead to a predictive capability. The models may be either highly elaborate numerical ones, requiring frequent observations of numerous ice parameters or simple conceptual ones, requiring occasional observations of one parameter.

In discussing models of floating ice, two scale distinctions are necessary. The time scale is of fundamental importance in determining the types of sensor to be used in a given model. On a time scale basis, models can be divided into two types: dynamic models requiring observations on an hourly or daily basis and climate models requiring them on a weekly, monthly, or seasonal basis. The second scale distinction is the spatial scale. Small scale models, say of a given estuary or segment of coastline, require distinctly different sensor resolutions and areal coverage than do oceanic scale models.

5.1.1 Ice Dynamics Models. Large scale, dynamic ice models of pack ice, such as the AIDJEX model, require short time scale, high resolution images covering large spatial areas. Since the ice pack exists in regions of prolonged darkness and frequent cloud cover, the need for all-weather, day or night sensors is essential. In short, these sea ice models are the most demanding ones of all floating ice models in their remote sensing requirements. Therefore, all of the sensors discussed above have some applications to these models.

During periods of darkness or extensive cloud cover, the remote sensing needed as input to the dynamic models can be acquired by satellite-borne active and passive microwave sensors. The SAR would provide the high spatial resolution and synoptic sequential images necessary to map the ice velocity vector fields, the changing distribution of leads and polynyas and the variations in the edges of the pack. The passive microwave sensors would provide data on the ice type, concentration, and divergence. Multi-frequency passive microwave sensors can also provide, on an all-weather basis, data on surface snow accumulation rate, total snow amount, and the ice (not snow) surface temperature.

During periods of sunlight and minimal cloudiness, pack ice can be observed by visible and infrared sensors as well as by the microwave. Under these conditions, these sensors provide additional air/surface interface temperatures and the familiar visual ice signatures. The latter are invaluable because of the vast experience with visual observations by the community of ice scientists.

The sensor coupling given above for sea ice applies as well to lake and river ice, with the important difference that the microwave sensors can be used to determine the thickness of the ice. Such observations are also less demanding on the observational system, since only mesoscale areas are involved.

Active microwave systems also appear to be the preferred sensors for acquiring data for dynamic models of ice shelves. Just as in the ice caps in which they originate, dynamic waves occur as a result of the flow produced by variations in mass balance. These waves have amplitudes ranging from several to tens of meters and wavelengths in the order of one to ten kilometers, easily resolvable by satellite-borne altimeters. Also, subsurface ice features can be observed with the satellite-borne SAR. A key term in the mass balance of ice shelves is their calving rate, which could be accurately measured for the first time by using repetitive SAR observations.

5.1.2 Climate Models. The full spectrum of remote sensing needs required by polar climate models is described in detail in the POLEX I and POLEX II reports published by the National Academy of Sciences, 1974. In this section, we wish to address the floating ice satellite observations pertaining to ocean-atmosphere interactions in the polar regions.

Climate models require ice remote sensing data that are both of lower spatial resolution and are obtained at larger time intervals than for ice dynamics models (see Table 1). The two primary forms of floating ice that relate to climatic variations are sea ice and shelf ice. Sea ice is a fundamental part of large-scale climatic processes because it undergoes great annual and seasonal variations in areal extent, its presence causes the significant alteration of the albedo of large oceanic areas, and it serves as an effective insulator in the ocean-atmosphere heat exchange. Ice shelves are also an important aspect of climate remote sensing, not because they are ordinarily prime movers in changes in climatic processes but because their variations are related to climatic change.

Because lower spatial resolutions are required for climate studies and because these data must be obtained on an all-weather day or night basis, passive microwave sensors are again ideally suited. The frequency of coverage required by climate models is assured since ESMR and SMMR systems will be flown on satellites that are designed for more frequent observations for other purposes. In short, satellite systems designed for acquiring data for floating ice dynamic models will collect, as part of their mission, all data required by the climate models.

5.2 Operations

5.2.1 Transportation. Sea ice remote sensing from satellites has the capability of assisting surface shipping in ice-infested waters in two main classes of problems. The first of these is the problem of prediction of ice conditions for either the "long" - term or the "short" - term future. Here by "long" we mean 1 to 10 years in advance (information that might be used for planning) and by "short" we refer to 1 day to 2 weeks (useful in short term routing). The second problem is the one of specifying the ice conditions as they exist at any instant.

In long term prediction, rapid assessment of current conditions is not essential. Almost any sensor could conceivably be of use if it assists in either calibrating by measuring such parameters as the ice thickness distribution or in validating operational numerical models for ice drift and deformation. Therefore information collected by visual, IR and active and passive microwave sensors would be useful in providing documentation on the state of the pack as a function of time. The exact parameters that would be measured from the imagery would, of course, depend upon the requirements of the specific model being used. These requirements would also change with time as the science of sea ice modeling changes. One important point about long term problems is that sensors that cannot penetrate clouds can still be extremely useful in that you always have enough time to await a day when conditions permit your favourite sensor to see the ice. .

The problems of specifying the state of the ice at any instant in time and of short term forecasting are quite similar from a remote sensing point of view, in that both require real-time or near real-time information. The ship captain needs to know where the line of least resistance is located in the vicinity of his ship. He cannot wait for a clear day to make his decisions. The same is true of operational short term forecasts; to predict the

state of the ice tomorrow, you need to know its state today. The most promising sensor in this area is SAR which with its map-like format, ability to discern leads and highly deformed ice areas, its usefulness in studying ice drift and deformations and its all-weather capability make it extremely attractive. To be able to provide near real-time SAR processing and transmittal to ships is a large order that would, however, seem to be possible within the next few years. It should not be hard to convince maritime operators who are used to utilizing ship radar of the utility of SAR imagery. Passive microwave imagery would also be useful but only if a significant increase in resolution can be achieved. The problems of rapid processing of microwave data are also severe.

5.2.2 Coastal Engineering. The coastal engineering problem is rather different to the transportation problem in that it is primarily a design problem. As such one is more interested in extremes than in averages; what are the largest ice pile-ups, the highest drift rates, the thickest ice, the largest ice forces. Many of these problems cannot at present be addressed from satellites inasmuch as the items of interest (ice thickness, pressure ridge characteristics, ice strength) are below the resolution of the sensor.

Satellite systems that would be of interest are SEASAT-SAR, which could provide valuable information on ice drift rates, deformation patterns and location of near coastal ridging and LANDSAT, which permits the determination of floe size distributions, presence of brash ice and (in conjunction with SEASAT-SAR) identification of ice islands.

Aircraft-based systems are very valuable, such as laser profilometers, ice-thickness radar and standard photography.

5.2.3 Surveillance. Non-military surveillance is becoming increasingly important in areas of endeavour such as oil pollution control and fisheries surveillance. This is due to the stepped up exploration of non-removable resources offshore and in the arctic, and the strong interest coastal states are taking in managing a 200 mile economic limit, as proposed at the Law of the Sea conference. For example, Canada in 1972 passed the Arctic Waters Pollution Prevention Act which extends sovereignty (or control) 100 miles offshore above 60° N, half way between Greenland and Canada in the NE and along longitude 141° W. To be able to control these zones established by national and international laws, all-weather remote sensing is required due to the large areas involved many of which have ice infested water for part, or most, of the year.

Another aspect concerns the surveillance of man's activity in remote areas. A good example is shown in Fig. 13. A drill rig on an artificial island is located near the southern end of the SLR image. The regular patterns just west of the drill rig represent plowed trenches in the snow cover on top of shore fast ice. The dark lines south of the drill rig are plowed roads on ice, whereas the curved lines north of the rig are pressure ridges. The type of sensors and systems required for monitoring such activities are of the type planned for SEASAT. It is envisaged that SAR would provide the basic information with support of passive microwave radiometry.

5.2.4 Biological Aspects. Ice infested oceans are, because of the logistical difficulties involved, places of meager biological observations. Yet in terms of biomass productivity the Bering Sea and the Antarctic Seas are among the most prolific areas on Earth. Remote sensing offers the only means to assess accurately the productivity of these regions and to monitor the populations of critical species. These observations are especially important now and in the next few years since many of the large sea animals, such as the mammals, are at the point of extinction due to oceanic pollution and over-hunting.

Although most sea life can be sensed only indirectly by observing the way in which certain species affect their surroundings, such as the colour change of the sea surface due to plankton blooms, recent work shows that large sea mammals can be sensed directly (Ray and Wartzok, 1975). In the Bering Sea Mammal Experiment (BESMEX) walrus herds have been observed by airborne infrared images and their population estimated. In BESMEX the walrus is used as an indicator species, and by observing its distribution and relating its variations to sea ice dynamics it is hoped that the biological health of the walrus and other associated species can be monitored.

Whether or not sea mammals can be observed directly by satellite is yet to be seen. Shapiro and Burns (personal communication) are using LANDSAT imagery to map the distribution of various ice types in the Bering and Beaufort Seas that are associated with seals and walrus. By coupling these data with usual population estimates of these species from aircraft, it is hoped that accurate population estimates and migration patterns can be obtained.

Again the dynamics and morphology of sea ice must be observed on an all-weather, day or night basis if the techniques now being explored are to be made truly useful. Therefore high-resolution active microwave sensors may prove to have great potential in biological studies of animals associated with floating ice.

6. VIEW TO THE FUTURE

6.1 Future Ice Programs

6.1.1 U.S.-Canadian Continental Shelf Ice Research. Both Canada and the U.S. are focusing a number of their research programs on the development of the natural resources of the arctic continental shelves and the arctic islands. In the U.S. the main current activity concerns itself with the NOAA (BLM) Outer Continental Shelf Arctic Program for the northern Bering Sea and the Chukchi Sea. In Canada the Beaufort Sea Project is terminating, the Arctic Islands Pipeline Program has just commenced and a high arctic islands offshore program is on the drawing board.

It is clear that these programs should be concerned to a major extent with remote sensing from satellites. Unfortunately this is not the case. Instead, aircraft are being employed with a variety of sensors as a stop-gap measure. When SLR is used it is in aircraft which have limited range or can fly only occasionally. This kind of approach is very expensive and ultimately does not provide the necessary data base for comprehensive understanding of the environmental factors being sensed.

6.1.2 POLEX. The Polar Experiment (POLEX) is a polar research program for both hemispheres proposed by the Soviet Union (Treshnikov et al., 1968; Borisenkov and Treshnikov, 1971). The U.S. contribution to POLEX (National Academy of Sciences, 1974) discusses the experiment in detail. The objectives as defined for GARP (Global Atmospheric Research Program) of which POLEX is a part are twofold:

First GARP objective: An understanding of the transient behaviour of the atmosphere as manifested in the large-scale fluctuations that control changes of the weather; this would lead to increasing the accuracy of forecasting over periods from one day to several weeks.

Second GARP objective: An understanding of the factors that determine the statistical properties of the general circulation of the atmosphere, which would lead to better understanding of the physical basis of climate. In order to fulfill these objectives in the frame work of POLEX the POLEX panel recommends that maximum use be made of remote-sensing and unmanned data buoys. It recognizes also that some data is available from satellites, but no systematic analysis program exists that would be useful for climate studies. Furthermore, it does not seem practical to have long term observational programs based on manned stations. The emphasis should be towards creating a useful meteorological and oceanographic data set from satellite observations.

NIMBUS-G, SEASAT-A and its descendants will be ideally suited for acquisition of the kind of information required to better understand the influence of polar regions on world climate.

6.1.3 Nansen Drift Station. The Fridtjof Nansen Drift Station project proposes to freeze a decommissioned U.S. Coast Guard icebreaker into the ice pack of the Arctic Ocean at the edge of the continental shelf in the area north of the Lena Delta. The ship would then serve as a scientific platform as it drifts with the ice over the North Pole. It would then emerge from the pack after a time period of 2 to 3 years in the vicinity of Svalbard. Because the drift of the station will follow the general path of the Trans-Polar Drift Stream, (the principal drift feature in the Arctic Ocean) there will be a chance to observe the metamorphosis and dynamics of the ice as it changes from first to multiyear ice. Both satellite-based and aircraft-based remote sensing techniques will undoubtedly be useful in this program. Presumably radar ice thickness measurements, microwave and IR radiometer observations and laser profilometry would be performed from the helicopters

based on the icebreaker. It might be possible to supplement these observations with SLR and microwave imagery obtained during the resupply flights. Finally, an overview of the general ice area around the ship platform could be provided by DMSP and NOAA imagery. SAR measurements from SEASAT would also be extremely useful during the start and finish of the drift if the satellite orbit was changed so that imagery could be obtained up to 80° N. If the northern-most limit of SEASAT coverage remains at 75° N the Nansen drift track will be almost completely out of range.

6.1.4 Operational Ice Forecasting. Operational ice forecasting is probably the most important and costly service governments provide to industry and agency operations. It has to be general, yet precise enough, to be useful to a multitude of users such as in marine operations, coastal engineering, and oil and gas industry to name a few. Current practise is to have ice observers on aircraft, who map the ice conditions by eye, aided in some cases by IR scanners, laser profilometers and search radars. These maps include, besides open water and fast ice, four different concentration levels of ice for five different ice conditions such as thickness and floe size. The parameters given above are the principal ones used in the Gulf of St. Lawrence and Bering Sea. The number of parameters increases considerably if one includes ice conditions in areas containing multiyear and fully developed first-year ice. This requires the ice observer to be very judicial, and to have an exceptional ability for the synthesis and assimilation of synoptical information. The limited areal coverage severely reduces the general usefulness of this kind of operation except if a large number of aircraft and observers are at hand.

In order to optimize this type of operation to yield the economic benefits which have been identified by McQuillan and Clough (1975), and Nagler and McCandless (1975), SEASAT must involve not only the acquisition of the necessary information, but also its reduction, analysis, the integration of various sensor results, interpretation and the near real-time distribution of the final product. The format or formats of the final product should be designed in such a way to enable: the operator to use it in tactical situations; and the forecaster to use it for a series of short to long term forecasts. The climatologist requires the same type of information for his research.

6.2 The Ideal Operational Ice Satellite System

Increasing scientific and commercial interest in the remote sensing of floating ice indicates the need for an operational satellite specifically designed for these purposes. It is clear that an operational satellite has to have an all-weather, day or night capability. Furthermore, to achieve the desired frequency of coverage and to image the areas of interest a number of satellites in a polar orbit will be required.

In order to achieve the required high resolution imagery it will be necessary to utilize active microwave systems. The contemplated L-band SAR for SEASAT-A may not give all the details required but with the addition of an X-band frequency system could provide a complete data set. A multifrequency passive microwave scanner of higher resolution than the current SMMR planned for NIMBUS-G and SEASAT-A is another basic requirement for an operational satellite. Combining the data products of MSAR (Multifrequency Synthetic Aperture Radar) with SMMR would provide the necessary basic information for an operational ice satellite.

As pointed out earlier the methods and procedures employed to deliver the information in an integrated form to the user is a formidable task in itself. The basic requirement is to have near-real time imagery in the user's hand, at great distances from the receiving station, within a few hours of obtaining it. This will require the transmission of the various sensor data to some sort of command and control centre (Morley and Clough, 1975) where the reduction, analysis, integration and interpretation takes place. The final product should go via communication satellite, or in some cases over telephone lines, back to the user. If the final product is an image it has to be of very high quality at a price the user can afford.

It is encouraging to note that some progress is being made in the implementation of similar information systems. This progress may prove very beneficial to the planners of the future operational satellite systems. For example Shaw (1975) and Gedney *et al.*, (1975) have shown that near-real time transmission of satellite imagery and real-time transmission of SLR imagery respectively, can be accomplished successfully with a resulting high-quality user product.

Finally the proposed SEASAT-A satellite scheduled for launch in April 1978 will provide an excellent data base for the type of information required in an operational ice satellite system. It is hoped that international cooperation will prevail in putting an operational SEASAT-type system together.

7. ACKNOWLEDGEMENT

R.O. Ramseier would like to acknowledge the support of the Polar Continental Shelf Project, the Department of National Defence and the Canadian Centre for Remote Sensing. W.F. Weeks would like to acknowledge the support of the Arctic Program, Office of Naval Research.

8. REFERENCES

- Ackley, S.F., and Hibler, W.D. (1975). Measurements of Arctic Ocean ice deformation and fracture patterns from satellite imagery. In "**SCAR/SCAR Polar Oceans Conference**", Montreal, May 1974 (M. Dunbar, ed.), Cambridge, Scott Polar Research Institute.
- Barnes, J.C., Chang, D.T., and Willard, J.H. (1972). Application of ITOS and NIMBUS infrared measurements to mapping sea ice. **Final Report Contract No. 1-36025 to NOAA (NESS)** by Allied Research Associates. 87 pp.
- Barnes, J.C., and Bowley, C.J. (1974). The application of ERTS imagery to monitoring Arctic sea ice. Lexington, Mass., Environmental Research and Technology, Inc. (**ERT Document 0408-F**).
- Barnes, J.C., and Bowley, C.J., and Smallwood, M.D. (1975). The application of imagery to monitoring Arctic sea ice: Supplemental Report. Lexington, Mass., Environmental Research and Technology, Inc. (**ERT Document 0408-S**).
- Bogorodskii, V.V., and Tripol'nikov, V.P. (1973). Electro magnetic characteristics of sea ice in the range of 30-400 MHz. **Akademii Nauk SSSR Doklady**, Tom 213, No. 3 p. 577-579.
- Bogorodskii, V.V., and Tripol'nikov, V.P. (1974). Radar sounding of sea ice. **Zhurnal Tekhnicheskoy Fiziki** 44(3), p. 660-662.
- Borisenkov, E.P., and Treshnikov, A.F. (1971). On the role of polar regions in the problem of the global investigation of the circulation of the atmosphere and ocean. **Trudy of the Arctic and Antarctic Institute**, 296, Leningrad.
- Campbell, K.J., and Orange A.S. (1974). A continuous profile of sea ice and freshwater ice thickness by impulse radar. **Polar Record** 17(106), p. 31-41.
- Campbell, W.J., Gloersen, P., Nordberg, W., and Wilheit, W.W. (1974). Dynamics and morphology of Beaufort Sea ice determined from satellites, aircraft, and drifting stations. In "**Symposium on Approaches to Earth Sciences through Use of Space Technology**" (P. Bock, ed.), Konstanz, Federal Republic of Germany (23-25 May 1973), p. 311-27, Akademie-Verlag.
- Campbell, W.J., Weeks, W.F., Ramseier, R.O., and Gloersen, P. (1975a). Geophysical studies of floating ice by remote sensing. **Journal of Glaciology** 15 (73) p. 305-328.
- Campbell, W.J., Ramseier, R.O., Weeks, W.F., and Wayenberg, J.A. (1975b). Visual observation of floating ice from SKYLAB. **SKYLAB Visual Observa..on Project Rept.** No. 2, Johnson Space Center, Houston, Texas.
- Campbell, W.J., Gloersen, P., Ramseier, R.O., Webster, W.J., and Wilheit, T.T. (in press) Beaufort Sea ice zones by means of microwave imagery. **Journal of Geophysical Research**.
- Chang, T.C., Gloersen, P., Schmugge, T., Wilheit, T.T., and Zwally, H.J. (1975). Microwave emmission from glacier ice, NASA - Goddard Space Flight Center, Greenbelt, Maryland. **Preprint X-910-75-36**.
- Chase, P. (1972). Using ITOS vidicon and scanning radiometric data for Great Lakes ice analysis and prediction. **NOAA/NESS Contract 1-36012**, The Bendix Corporation, Aerospace Systems Division, Ann Arbor, MI.

- Chudobiak, W.J., Gray, R.B., Ramseier, R.O., Makios, V., Vant, M., Davies, J.L., and Katsube, J. (1974). Radar remote sensors for ice thickness and soil moisture measurements. **Proceedings Second Canadian Symposium on Remote Sensing**. (Guelph, Ontario), p. 418-424.
- DeRycke, R.J. (1973). Sea ice motions off Antarctica in the vicinity of the eastern Ross Sea as observed by satellite. **Journal of Geophysical Research** 73(36), 8873-8879.
- Dickinson, L.J., Boselly, S.E., and Burgmann, W.S. (1974). Defense Meteorological Satellite Program (DMSP) Users' Guide. Air Weather Service (MAC), **United States Air Force AWS TR-74-250**, Hq. AWS, Scott AFB, Illinois 62225, 109 pp.
- Dunbar, M., and Wittmann, W. (1963). Some features of ice movement in the Arctic Basin. In "**Proceedings Arctic Basin Symposium (Hershey, Pennsylvania**, October 1962)", p. 90-104, Arctic Institute of North America.
- Dunbar, M., and Weeks, W.F. (1975). Interpretation of young ice forms in the Gulf of St. Lawrence using side-looking airborne radar and infrared imagery. **Cold Regions Research and Engineering Laboratory Research Report** 337, 41 pp.
- Edgerton, A.T., Stogryn, A., and Poe, G. (1971). Microwave radiometric investigations of snow pack. **Final Report No. 1285R-4 for U.S.G.S. contract no. 1A-08-001-11828**, Aerojet-General Corporation Microwave Division, El Monte, California.
- Elachi, J., and Brown, W.E., Jr., (1975). Imaging and soundings of ice fields with airborne coherent radars. **Journal of Geophysical Research** 80(8), p. 1113-1119.
- Gedney, R.T., Schertler, R.J., Mueller, R., Jirberg, R.J., and Mark, H., (1975). An operational all-weather Great Lakes ice information system. **Proceedings, Third Canadian Remote Sensing Meeting, Edmonton, Alberta**.
- Gloersen, P., Nordberg, W., Schmugge, T.J., Wilheit, T.T., and Campbell, W.J. (1973). Microwave signatures of first-year and multiyear sea ice. **Journal of Geophysical Research** 78(18), p. 3564-72.
- Gloersen, P., Chang, T.C., Wilheit, T.T., and Campbell, W.J. (1974a). Polar sea ice observations by means of microwave radiometry. **Proceedings Inter-disciplinary Symposium on Advanced Concepts and Techniques in the Study of Ice and Snow Resources, Monterey, California**, p. 541-550, National Academy of Science.
- Gloersen, P., Ramseier, R.O., Campbell, W.J., Chang, T.C., and Wilheit, T.T. (1974b). Variation of the morphology of selected micro-scale test areas during the Bering Sea Experiment. In "**Results of the U.S. Contribution to the Joint U.S./U.S.S.R. Bering Sea Experiment**". Greenbelt, Maryland, Goddard Space Flight Center, p. 104-21. (NASA X-910-74-141).
- Gloersen, P., Ramseier, R.O., Campbell, W.J., Kuhn, P.M., and Webster, W.J., Jr. (1974c). Ice thickness distribution as inferred from infrared and microwave remote sensing during the Bering Sea Experiment. In "**Results of the U.S. Contribution to the Joint U.S./U.S.S.R. Bering Sea Experiment**". Greenbelt, Maryland, Goddard Space Flight Center, p. 104-21. (NASA X-910-74-141).
- Gloersen, P., and Salomonson, V.V. (1975). Satellites - New global observing techniques for ice and snow. **Journal of Glaciology**, 15(73) p. 373-389.

- Hibler, W.D., Ackley, S.F., Crowder, W.K., McKim, H.L., and Anderson, D.M. (1974). Analysis of shear zone ice deformation in the Beaufort Sea using satellite imagery. In "**The Coast and Shelf of the Beaufort Sea**", (J.C. Reed and J.E. Sater, eds.), Arlington, Virginia, p. 285-296 Arctic Institute of North America.
- Hibler, W.D., Tucker, W.B., and Weeks, W.F. (1975). Techniques for studying sea ice drift and deformation at sites far from land using LANDSAT imagery. **Proceedings of the 10th International Symposium on Remote Sensing of Environment**, 1975 Ann Arbor, Willow Run Laboratories, Environmental Research Institute of Michigan.
- Hibler, W.D. (1975). Characterization of cold regions terrain using airborne laser profilometry. **Journal of Glaciology** 15(73) p.329-347.
- Hoekstra, P., and Capillino, P. (1971). Dielectric properties of sea and sodium chloride ice at UHF and microwave frequencies. **Journal of Geophysical Research**, 76(20), p. 4922-4931.
- Ketchum, R.D., Jr., and Tooma, S.G., Jr. (1973). Analysis and interpretation of airborne multi-frequency - Side Looking Radar sea ice imagery. **Journal of Geophysical Research**, 78(3) p. 520-538.
- McClain, E.P. (1974a). Earth satellite measurements as related to sea ice studies. In "**COSPAR Symposium on Approaches to Earth Survey Problems Through Use of Space Techniques, Constance, F.R.G., May 1973**" (P. Bock, ed.), Akademie-Verlag, Berlin.
- McClain, E.P. (1974b). Some new satellite measurements and their application to sea ice analysis in the Arctic and Antarctic. In "**Advanced Concepts and Techniques in the Study of Snow and Ice Resources**" (H.S. Santeford and J.L. Smith, eds.), Washington, National Academy of Sciences, p. 457-466.
- McQuillan, A.K., and Clough, D.J. (1975). Benefits of remote sensing systems to petroleum operations in Canadian ice infested waters. **Proceedings, Third Symposium on Remote Sensing, Edmonton, Alberta**.
- Meeks, D.C., Ramseier, R.O., and Campbell, W.J. (1974). A study of microwave emission properties of sea ice - AIDJEX 1972. **Proceedings Ninth International Symposium on Remote Sensing of Environment**, University of Michigan, Ann Arbor.
- Morley, L.W., and Clough, D.J. (1975). Proposal for an arctic operations centre. **Proceedings, Third Canadian Symposium on Remote Sensing, Edmonton, Alberta**.
- Nagler, R.G., and McCandless, S.W. (1975). Operational Oceanographic Satellite - SEASAT potential for oceanography, climatology, coastal processes and ice. **NASA - Headquarters, Washington, D.C.** (draft report).
- National Academy of Sciences, (1974). U.S. contribution to the Polar Experiment (POLEX). Part 1 POLEX-GARP (North); and Southern Ocean Dynamics, a strategy for Scientific Exploration 1973-1983.**
- Nye, J.F. (1975). The use of ERTS photographs to measure the movement and deformation of sea ice. **Journal of Glaciology** 15(73).
- Page, D.F., and Ramseier, R.O. (1975). Application of active radar techniques to the study of ice and snow. **Journal of Glaciology** 15(73).
- Parashar, S.K., Biggs, A.W., Fung, A.K., and Moore, R.K. (1975). Investigation of radar discrimination of sea ice. **Proceedings of the Ninth Symposium on Remote Sensing of Environment, University of Michigan, Ann Arbor**.

- Poulin, A.O. (1973). On the thermal nature and sensing of snow-covered Arctic terrain. **U.S. Army Topographic Laboratories Research Note ETL-RN-73-4**, 173 pp.
- Poulin, A.O. (1975). Significance of surface temperature in the thermal infrared sensing of sea and lake ice. **Journal of Glaciology** 15(73).
- Ramseier, R.O., Campbell, W.J., Weeks, W.F., Arsenault, L.D., and Wilson, K.L. (1975). Ice dynamics in the Canadian Archipelago and adjacent Arctic Basin as determined by ERTS-1 observations. In "**International Symposium on Canada's Continental Margins and Offshore Petroleum Exploration (Calgary, Sept. 1974)**," Canadian Society of Petroleum Geologists p. 853-877.
- Ramseier, R.O., Gloersen, P., and Campbell, W.J. (1974). Variation in the microwave emissivity of sea ice in the Beaufort and Bering Sea. In "**Specialist Meeting on Microwave Scattering and Emission from the Earth (E. Schanda, ed.)**", Proceedings of the URSI, Commission II, p. 87-93, Institute of Applied Physics, University of Bern, Switzerland.
- Ray, G.C. and Wartzok, D. (1974). BESMEX, Bering Sea Marine Mammal Experiment. **NASA Ames Research Centre Report TM X-62.399** p. IV + 51.
- Rothrock, D.A., and Hall, R.T. (1975). Testing the redistribution of sea ice thickness from ERTS photographs. **AIDJEX Bulletin** No. 29, p. 1-19.
- Rouse, J.W., Jr. (1969). Arctic ice type identification by radar. **Proceedings of the IEEE**, 57(4), p. 605-611.
- Sellmann, P., Weeks, W.F., and Campbell, W.J. (1975). Use of side-looking airborne radar to determine lake depth on the Alaskan North Slope. **Cold Regions Research and Engineering Laboratory Special Report 230**, 6 pp.
- Shaw, E. (1975). Near real-time transmission of sea ice satellite imagery. **Proceedings, Third Canadian Symposium on Remote Sensing, Edmonton, Alberta**.
- Streten, N.A. (1974). Large-scale sea ice features in the western Arctic Basin and the Bering Sea as viewed by the NOAA-2 satellite. **Arctic and Alpine Research** 6(4), p. 333-345.
- Swithinbank, C.W.M., and Zumberge, J.H. (1965). The ice shelves. In "**Antarctica (T. Hatherton, ed.)**", London Methuen and Co. Ltd., p. 199-220.
- Swithinbank, C.W.M. (1972). Arctic pack ice from below. In "**Sea Ice, Proceedings of an International Conference** (T. Karlsson, ed.), p. 246-254. National Research Council, Reykjavik, 1972.
- Thorndike, A.S., Rothrock, D.A., Maykut, G.A. and Colony, R. (1975). The thickness distribution of sea ice. **Journal of Geophysical Research**.
- Treshnikov, A.F., Borisenkov, Ye.P., Nikiforov, E.G., Mustafin, N.V., Chaplygin, E.D., and Shpaikher, A.O. (1968). An ocean-atmosphere interaction experiment for the Arctic. **Probl. Arkt. Antarkt. No. 28**. Translated by the RAND Corporation, November 1969.
- Vant, M.R. Gray, R.B., Ramseier, R.O., and Makios, V. (1974). Dielectric properties of fresh and sea ice at 10 GHz. **Journal of Applied Physics**, 45(11) p. 4712-4717 and 46(5) p. 2339 (1975).
- Vickers, R.S., Heighway, J., and Gedney, R. (1973). Airborne profiling of ice thickness using a short pulse radar. In "**Advanced Concepts and Techniques in the Study of Snow and Ice Resources**", (H.S. Santeford and J.L. Smith, eds.). National Academy of Sciences, p. 422-431.
- Weeks, W.F., and Assur, A. (1969). Fracture of lake and sea ice. **Cold Regions Research and Engineering Laboratory Research Report 269**, p. 77.

Weeks, W.F., Kovacs, A., and Hibler, W.D. (1971). Pressure ridge characteristics in the Arctic coastal environment. In "**Proceedings First International Conference on Port and Ocean Engineering under Arctic Conditions**", 23-30 August 1971", Vol. I, p. 152-183, Department of Port and Ocean Engineering, Technical University of Norway, Trondheim.

Weeks, W.F. and Campbell, W.J. (1973). Icebergs as a fresh water source: an appraisal. **Journal of Glaciology** 12(65), p. 207-233.

Weeks, W.F. and Campbell, W.J. (1975). The remote sensing plan for the AIDJEX main experiment. **AIDJEX Bulletin** No. 29, p. 21-48.

Wilheit, T.T., Nordberg, W., Blinn, J., Campbell, W.J., and Edgerton, A. (1972). Aircraft measurements of microwave emission from Arctic sea ice. **Remote Sensing of Environment** 2(3), p. 129-139.