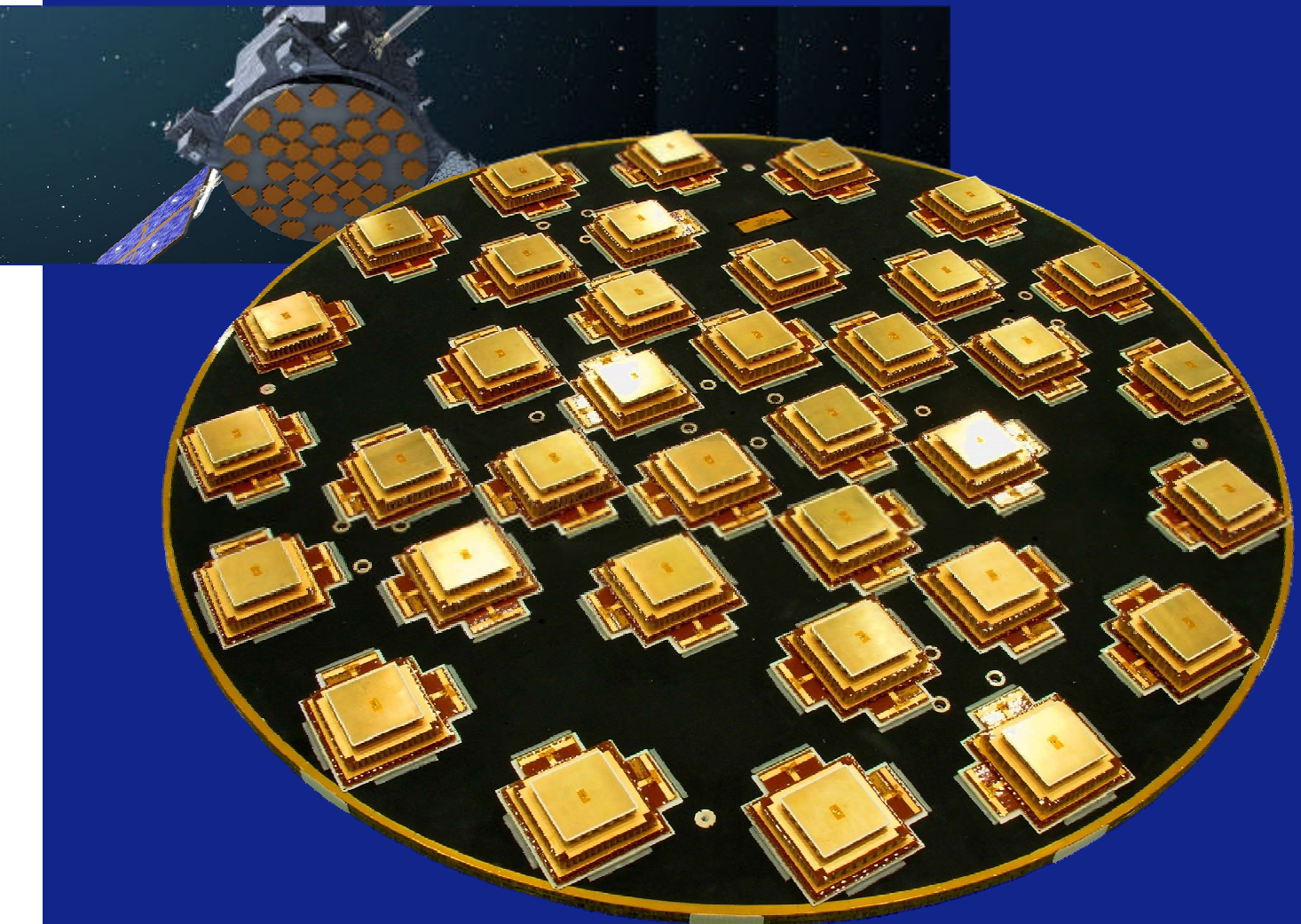


Introduction to the design of

# ***ANTENNAS FOR SPACE APPLICATIONS***

Marco Sabbadini





# ***Introduction to the design of ANTENNAS FOR SPACE APPLICATIONS***

**ESTEC Working Paper No. 1963**

**2<sup>nd</sup> Edition**

Checked by: C. Mangenot

Approved by: A.Roederer

**Note:** ESTEC Working Papers are non-official documents intended for the presentation of material for discussion purposes only. The contents of this document do NOT represent approved ESA procedures, practices, or standards. This document or any part of it may not be issued in any form other than the present without the approval of the European Space Agency.

## Introduction to the Design of Antennas for Space Applications

Copyright © 1998, 2006 ESA

Written by Marco Sabbadini

Graphics by Marco Sabbadini

Printed by ESTEC Reproduction Services

First English edition January 1998

Second English edition March 2006

Italian editions 1990, 1991, copyright © UniTor

Cover: Giove A Navigation antenna (courtesy of Alcatel Alenia Space)  
Artist view of the Giove A satellite

Cover art: Marco Sabbadini

All images for which credits are not specified are from ESA photo archives

# ***Introduction to the design of ANTENNAS FOR SPACE APPLICATIONS***

**M a r c o   S a b b a d i n i**

March 2006



## **Preface**

Antennas of all types are used for space applications in frequency bands scattered across about 4 decades, from a few hundreds of MHz to a few THz. Wire antennas (e.g. dipole and helix), aperture antennas (e.g. waveguide horns), array antennas, reflector antennas and lens antennas have been embarked on several missions over the past 50 years.

A complete and exhaustive coverage of the whole topic would be a daunting task and it is incompatible with the introductory intent of these notes. Among others, the discussion would have to include theoretical, technological and production aspects, covering electromagnetic, mechanical, thermal and material science, as well as mathematical modelling and computer simulation, which are major tool in antenna engineering. While concentrating on the electromagnetic aspects, an attempt has been made in the following to provide links to the other disciplines by identifying the contact points.

To maintain the exposition at a manageable depth and length, the discussion is limited to the most commonly used antennas and to the basic concepts applied in their design, while trying to give a sufficiently general view of the possible alternatives applicable to each area. The material is organised application by application, as an overview of design problems and possible solutions. Specific theoretical or technological topics are purposely introduced in relation with the application for which they are more relevant, rather than in a systematic way that could be more appropriate for an antenna theory book. Occasionally the discussion is broadened to introduce aspects relevant to different applications and some specific topics are discussed in more than one place, with the objective of stressing the many links among apparently very different solutions underpinning this engineering discipline.

An attempt has been made to define all terms and quantities when they are first encountered, trying to find a balance between mathematical precision and practical use. In the same way most concepts of antenna theory are briefly recalled when first applied. Nevertheless it is necessary to have a good background in electromagnetic theory and antenna theory to easily follow and fully understand the discussion.

The first and second versions of this book were written, in Italian, as lectures notes for a course on "The Design of Antennas for Space Applications", which I lectured in the years 1990-1997, as part of the course on Antennas and Propagation at the Faculty of Engineering of the University of Rome, Tor Vergata, under the co-ordination of Prof. F. Bardati, to whom I remain in debt for this excellent opportunity to try organising my knowledge in this field.

The first English version, dating January 1998, was the result of further revisions and of a translation kindly produced by M. Taillefer, at the time with the Translation Division of ESA at ESTEC, Noordwijk, The Netherlands, who I wish to warmly thank for this great help, without which this text would never have been.

The current edition has been expanded covering the most recent developments and introducing pictures of actual hardware at the occasion of the Space Antenna Design course held at ESTEC in March 2006, organised in cooperation with the European School of Antennas (part of the Antenna Network of Excellence funded by EU under the 6<sup>th</sup> Framework Programme).

## ***Table of Contents***

|                                                        |    |
|--------------------------------------------------------|----|
| Preface                                                | i  |
| Introduction                                           | 1  |
| CHAPTER 1                                              |    |
| Types of antennas                                      | 5  |
| 1.1 Applications                                       | 5  |
| 1.2 An antenna, what is it exactly?                    | 8  |
| 1.3 Characteristic parameters                          | 10 |
| 1.4 Categories of antennas                             | 14 |
| CHAPTER 2                                              |    |
| Fixed communication systems                            | 21 |
| 2.1 Multiple-beam antennas                             | 25 |
| 2.2 The transform-chain model                          | 26 |
| 2.3 Design parameters for multiple-beam antennas       | 28 |
| 2.4 Reconfigurable antennas                            | 30 |
| 2.5 Beam-forming networks                              | 33 |
| 2.6 Selective surfaces                                 | 37 |
| 2.7 Frequency reuse in multiple-beam coverages         | 39 |
| 2.8 Passive, semi-active and active antennas           | 41 |
| CHAPTER 3                                              |    |
| Mobile communication, multimedia and broadband systems | 43 |
| 3.1 Satellite constellations                           | 45 |
| 3.2 Folding reflectors                                 | 45 |
| 3.3 Passive intermodulation products                   | 48 |
| 3.4 High-efficiency feeds                              | 50 |
| 3.5 Array antennas and magnified array antennas        | 52 |
| 3.6 Magnification and non-focusing antennas            | 54 |
| 3.7 Multiple-layer planar antennas                     | 55 |
| 3.8 Meteorological attenuation and reconfigurability   | 58 |
| 3.9 Large communication satellites                     | 59 |
| 3.8 Antennas for mobile terminals                      | 60 |

|                                                            |     |
|------------------------------------------------------------|-----|
| CHAPTER 4                                                  |     |
| Direct satellite broadcast systems                         | 63  |
| 4.1 Contoured-beam antennas                                | 64  |
| 4.2 Reflector shaping                                      | 67  |
| 4.3 Double reflector antennas                              | 71  |
| 4.4 Degrees of freedom of an antenna                       | 72  |
| 4.5 The resonant discharge in vacuum                       | 76  |
| 4.6 Small receiving antennas                               | 78  |
| CHAPTER 5                                                  |     |
| Remote sensing                                             | 81  |
| 5.1 Radar systems                                          | 83  |
| 5.2 Degrees of freedom of arrays                           | 83  |
| 5.3 Synthetic aperture radars                              | 84  |
| 5.4 Array synthesis                                        | 85  |
| 5.5 Radiating elements for arrays                          | 88  |
| 5.6 Active array antennas                                  | 89  |
| 5.7 Measurement of the radiation pattern of large antennas | 91  |
| 5.8 Radiometers                                            | 93  |
| 5.9 Synthetic aperture radiometer antennas                 | 97  |
| 5.10 Millimetre and sub-millimetre waves                   | 98  |
| CHAPTER 6                                                  |     |
| Other applications                                         | 101 |
| 6.1 Scientific applications                                | 101 |
| 6.2 Navigation                                             | 102 |
| 6.3 Search and Rescue                                      | 103 |
| 6.4 Data relay                                             | 103 |
| 6.5 Avionic antennas                                       | 107 |
| 6.6 Ground station antennas                                | 110 |



# Introduction



Sputnik 1, the first artificial satellite,  
launched the 4 October 1957

The first artificial satellite, Sputnik I, was launched by the USSR on October 4<sup>th</sup>, 1957 [1] marking the beginning of the development of a new branch of technology and of an era of completely new opportunities for science as well as daily life. From the early 1960's, the number of satellites put into orbit around the Earth or sent to explore the Solar System and beyond has increased rapidly and tens of spacecrafts are now launched every year. In the late 1980's, after some 30 years of development, satellite technology entered a phase of intensive commercial exploitation, mostly in the field of communications and more recently navigation.

In the following years commercial exploitation has extended to the area of remote sensing, where the primary product are the data provided by the satellite instrument -possibly after processing-, rather than the services provided from the satellite itself as it happens in the communications and navigation area. An attempt has also been made to exploit commercially the very low-gravity environment on-board manned and unmanned crafts in low-earth orbit, like the International Space Station, which offers, for instance, the possibility to produce crystals with a very low content of defects and to produce chemicals difficult to obtain on earth. The cost, the risks and the very new way of operating have kept the interest of potential users below initial expectations.

Since the very early days of space technology, the number of scientific missions to increase the knowledge about the Earth, the Solar System and the whole universe has been and still is also quite significant. Recent successful missions, like those to Mars, Saturn and Titan, provide invaluable knowledge about the present and the past of the Solar System and of the whole Universe as well as the incentive for new missions with their scientific and technological challenges. New engineering solutions are required either to meet new scientific objectives or to be able to survive the harsh environment found on planets and moons across the Solar System. Finding these solutions stimulates basic research in the space field, which often produces results applicable to daily life.

After the landing on the Moon in 1969, the presence of mankind in space –close to Earth for the time being- has become a daily reality. The USA Space Shuttle and the Russian space station MIR have been for years the key instruments. Today, despite the financial and political difficulties, the International Space Station and the Soyuz continue to ensure this presence and there are NASA plans to come back to the Moon and possible reach Mars. At the same time China is also quickly moving toward having a stable human presence in space. In Europe, beyond the participation to several manned missions and to the International Space Station program with the Columbus and Cupola modules and several experiments, various other lines are being pursued with the same objectives.



Photo of the International Space Station taken from the Space Shuttle Discovery in July 2005 (courtesy of NASA)

Also space antenna technology has been in constant development over the past 50 years, following the continuous evolution of space missions and their demand for increased performance. The available transmission capacity of communication satellites and the amount of data produced by remote sensing and scientific ones are constantly increased and better and better use has to be made of the radio spectrum to cope with this growth. Antennas have a major role in this endeavour. For instance, new technologies are continuously developed to make use of higher frequencies, so as to widen the available bandwidth increasing the symbol rate of transmissions as well as the accuracy and resolution of instruments.

In the communications sector, the growing commercial interest and industrial competition have been giving further impetus to technological development over the past 20 years. Meanwhile the emphasis has changed from technical excellence to the overall economic return, involving industrial process and investment aspects, and resulting in large changes in the organisation of the whole space sector in Europe and worldwide. Not surprisingly, under the pressure of the “globalisation of economy” in this turn of millennia, a similar change is taking place also in the scientific and manned space-flight areas, where cost-effectiveness and solution reuse are becoming more and more important. As a result technological research and development is increasingly organised around very short term goals and toward long term innovation objectives. This pattern is typical of all industrial sectors having reached a good degree of maturity.

The most visible effect of technology evolution is the growth antenna complexity, which is essentially linked to two factors. On one hand, the demand for more concentrated beams, to increase the density of power radiated per unit of solid angle or, which is the same, the spatial resolution, leads to a growth in antenna dimensions and requires the use of advanced composite materials and of structures that can be opened in orbit. On the other hand, the tendency to incorporate in the antenna functions previously carried out by other components of the system, such as signal amplification or discrimination between different frequency bands, implies an extension of the concept of antenna from a passive component to an active system. An interesting aspect of the new commercial applications in the communication sector is that, despite the economic factors driving toward simple and affordable solutions, antennas are getting even more sophisticated because of the specific needs of their applications. Another consequence is the tendency to pack as many antennas as possible on the same satellite to reduce the cost of each service, which implies that antennas have to be designed taking into due account the presence of neighbouring ones.

Space antennas operate in a very adverse environment, which is the source of a significant amount of added complexity. The amount of Sun radiation impinging on the satellite is very high ( $\sim 1.4 \text{ Wm}^{-2}$ ) and its direction changes while the satellite moves along the orbit, at the same time its dark side is exposed to cold space at around 4K. Much worse than standing close to a fireplace in a very cold room... The absence in vacuum of heat exchange by convection and the relatively small thermal inertia of the satellite expose it to wide temperature gradients. This effect is even more pronounced on

antennas that are located externally and are in general thermally isolated from the satellite body. The antenna temperature range reaches and exceeds 300°C, from around -160°C to around 150°C typically. Large thermal gradients develop within elements having a large extension and low conductivity (such as composite material reflectors) whenever the Sun illuminates one part or side of them, an important effect for all satellites orbiting around the earth periodically enter its shadow (eclipse), cooling down very quickly, and warming up as quickly when coming out of it. The effects of the distortions due to differential thermal expansion in such conditions are often too large to be considered as perturbations and must be taken directly into account in the antenna design or corrected dynamically. Unfortunately, this is only one part of the picture. In the absence of the screening effect of the atmosphere, satellites are exposed to a continuous bombardment of high-energy radiations, from UV and Gamma rays to high-energy proton and electrons, and to high velocity dust particles and conglomerates of various sizes (micro-meteoroids). Low-orbiting satellites are also exposed to the highly erosive and corrosive effects of atomic oxygen. Furthermore during the initial part of the launch phase satellites are exposed to very high mechanical stresses, caused by the intense vibrations generated by the rocket engines and the high velocity airflow around the launcher. Local accelerations of tens of G's are common and every component and connection has to withstand the resulting mechanical loads.

The design of antenna systems requires therefore the participation of experts from various sectors of engineering. Some thirty years ago, space antennas were simple enough to make their design an essentially electromagnetic problem. Mechanical, thermal and materials technology aspects were usually addressed in a second phase of the design. Today, their design is feasible only with the concurrent work of a team of experts in electromagnetics, thermal and mechanical engineering, material technology as well as production and testing experts. The growing use of active antennas is also making the participation of radio frequency and digital circuit engineers increasingly important.

Finally, some considerations about complexity are in order as a warning to ourselves, as antenna engineers. Despite the appeal that such idea may have, good technology is not the same thing as complicated solutions, quite the opposite actually. Simple solutions are often more robust and flexible and certainly more elegant and affordable, if difficult to elaborate on a short time scale. Letting the search for better solutions be mistaken for the use of the most immediate and often more complex one, possibly as showcase for technological capabilities, is often the cause of failures, with their toll in wasted efforts and resources, a luxury human kind should increasingly abandon for its common wellbeing.

Natural evolution comes in waves of great differentiation, to fill new ecological niches, followed by consolidation of a few successful species, when a niche reaches saturation. The process is inherently slow but quite effective on a long time scale. The study of natural systems, including human society, has shown how this behaviour is strictly linked with the inherent properties of complex systems, i.e. systems that owing to the large number of inputs and internal connections exhibit hardly predictable behaviours. Nevertheless the "solutions" that emerge as successful are very often beautifully simple and efficient.



Artist impression of the Rosetta lander. The first man made object to be touching a comet

Satellites are certainly among the most complex (partially) automated systems that modern technology has produced. It is today necessary to make them as simple and efficient as possible to affordably support the needs of human societies.

**References**

- [1] <http://www.hq.nasa.gov/office/pao/History/sputnik/sputorig.html>

## CHAPTER 1

# *Types of antennas*



The Venus Express probe ready for shipping to the Baikonur launch base

Radio signals are the only means to transmit commands to satellites and receive their status information (telemetry) back on earth. Antennas are therefore absolutely necessary for their operations and each satellite has at least one.

They are also used in the majority of on-board system operating with electromagnetic fields. The only exceptions are those operating at very low frequency (e.g. magnetometers) or very high frequency (e.g. optical).

Launchers and other space vehicle, like the NASA Space Shuttle or the ESA Automated Transfer Vehicle also need antennas, for communication with the ground control stations and to receive navigation signals. In all manned spacecraft, including orbiting stations, antennas are part of the vital radio link allowing the crew to be in constant contact with ground.

In conclusion antennas are one of the most widely used components in space applications.

### **1.1 Applications**

Antennas are generally specialised to the requirements of each application and can be quite different one from another. On the other hand, there are much less antenna types than applications so that many similarities can be found, making it possible to speak in general about antennas for space applications. In this area they are traditionally used in communications, remote sensing and science, including radio telescopes, which are to all effects antennas with an incorporated receiver (since it is usually quite smaller than the antenna itself). Navigation was added more recently together with the Search and Rescue services, constituted by a repeater able to receive the faint signals generated by transmitters placed in buoyancy aids and send them to surveillance centres.

A quick review of these applications is presented in the following and forms the basis for the subdivision of subsequent chapters.

#### *Communications*

The civilian use of satellites, limited to intercontinental telephony and television signal relay, began with the Intelsat satellites, the first of which, known as Early Bird, was launched in 1965 [1]. New applications have gradually appeared, including radio and

television broadcasting and mobile telecommunication systems. In the late 1990's various plans were made to deploy global satellite networks for mobile communications, data transmission (e.g. for multimedia applications) and other special services, but only two satellite constellations were launched Iridium and Globalstar and have not been commercially successful. Since then only more traditional missions have been considered in this area.

In particular communications satellites are used to provide the following services:

- Communications for fixed systems (telephony and data transmission)
- Television signal relay (satellite TV link)
- Communications with and among mobile units (ships, aircrafts and ground vehicle)
- Direct radio and television broadcasting by satellite
- Personal communications (satellite phones)
- Wideband multimedia and computer networks.

### *Remote sensing*

The second well established application of satellites is remote sensing, for military purposes (spy satellites) or civilian purposes (meteorological and earth observation satellites), such as the well-known European Meteosat (first and second generation), ERS1 and ERS2 and Envisat. In this case, the antennas are the sensors of the measurement instruments.

The data produced by remote sensing satellites are used for many scientific and non-scientific applications, including:

- Weather forecast
- Meteorology (detection and measurement of cloud masses, winds, precipitation, sea and land temperatures)
- Earth sciences (oceanography, geology, etc.)
- Measurement and planning of earth resources
- Environmental monitoring
- Observation in emergency conditions (fires, earthquakes, volcanic eruptions and other disasters).

It is worth noting that remote sensing instruments are also a major element of planetary exploration missions (probes).

### *Scientific applications*

The field of scientific applications is vast and differentiated. Each satellite has its own characteristics and mission so it is quite difficult to produce a complete list. However the most typical applications are:

- Radio astronomy (especially in frequency bands absorbed by the atmosphere)
- Ultra-long base radio interferometry (with orbiting and/or earth stations)
- Exploration of the solar system
- Astrophysics (e.g. the study of comets).

### *Navigation*

A more recent application is the use of satellites for navigation. These systems are based on constellations of beacon satellites that transmit coded signals containing information about time and position. The first system, the USA Global Positioning System (GPS), was

originally deployed for military purposes. The use of its public channel became widespread in the 90's for all sorts of civilian applications, from geodesy to air-traffic control, prompting the open availability of more GPS channels for commercial use. A European system, called Galileo, is being established. The first test satellite, Giove A, was launched on December 28<sup>th</sup>, 2005 and the full constellation is planned to be fully operational around 2010. Although it might seem contradictory to place another navigation system in orbit, it should be remembered that a position can be determined more rapidly and accurately when more satellites can be used, i.e. are visible from the user location. Furthermore the existence of more than one system reduces the risk associated to unavailability for failures of other reasons. As the plan is to use, among others, the navigation signals to ease air traffic over Europe it is clearly critical to have a system controlled by European authorities.

#### *Data relay*

Although later than initially predicted and at a lower pace and both for technical problems –the tragic failures of the Space Shuttle- and political reasons –mostly lack of resources-, the so-called space infrastructure is being established. Unfortunately the future of the International Space Station (ISS) is today uncertain, but it is well known that a similar facility is necessary to support the continuing presence of mankind in space, even just on the Moon.

The construction and servicing of ISS requires a number of "space ferries" carrying crew and cargo to and from the station, including the Space Shuttle, the Russian Soyuz and the European Automated Transfer Vehicle. The station itself and all the vehicles, especially manned ones, require a continuous link to ground, which is hard to achieve since they move quite fast in the sky and a large number of ground stations would be required to track them. Relay satellites are used for this purpose: the USA TDRSS (Tracking and Data Relay Satellite System) and the European Artemis.

For the same reasons relay satellites are also used to relay to ground the data produced by low orbiting scientific or remote sensing satellites. The advantage in this case is related to the increased transmission capacity, rather than to link continuity.

#### *Inter-orbit link*

Space vehicles and orbiting stations also require communication links among them, for instance for approach and docking operations. These links usually operate at low frequencies to easily achieve omni-directional coverage.

Satellite constellations may also use direct links among satellites to route the signals among a "gateway" satellite visible to a ground station and all the others. In communication systems, two or more "gateway" satellites may be used to connect terminals on the ground via the "satellite network". These links can be operated at radio frequency, typically around the 60GHz to profit of the atmospheric attenuation that completely mask the signal from ground noise, or via optical links (laser beams).



Artist view of the ESA Automated Transfer Vehicle approaching the International Space Station

### *Spacecraft avionics systems*

As already mentioned, antennas are used for two vital satellite functions. The Telemetry, Tracking and Command (TT&C) system receives commands from ground and transmits telemetry data, i.e. the operating status of the satellite. A failure of this system results in the loss of the mission and it must have a very low probability to occur. Earth observation missions also require the transmission to ground of the measured data and the same applies to scientific ones, be that from a low earth orbit, i.e. some 800 km above the surface, or nearby Saturn. Clearly the antenna requirements for the two latter cases are quite different.

Another service function, typical of space vehicles, is the use of antennas for communication between the vehicle and crew operating outside it (Extra Vehicular Activities).

Finally, as the use of navigation signals for timing, navigation and attitude determination is becoming common also in space, satellites embark antennas dedicated to these functions.

To summarise, antennas are used in the following service sub-systems:

- Telemetry, tracking and command system
- Transmission to ground of data from remote sensing and scientific spacecraft
- Reception of navigation and timing signals (for scientific and earth observation missions, space vehicles and orbiting stations)
- Attitude control using navigation signals
- Radio link for Extra Vehicular Activities

## **1.2 An antenna, what is it exactly?**

Antennas can be seen in many different ways when looking from the engineer's perspective. In the first place, antennas are mechanical pieces, mostly composed by metallic and composite polymer parts. Their shape is determined by the required electromagnetic behaviour, and subject to mechanical, environmental, material technology and manufacturing requirements.

From the electromagnetic point of view, which is by necessity the dominant one in antenna design, an antenna can be seen in three or four different ways.

### *An antenna is a transformer*

A transformer is a passive two-port device that changes the characteristic of energy flow, for instance the voltage and intensity of an AC current. Antennas transform guided electromagnetic waves into free propagating waves (figure 1.1) and translate the impedance of the energy flow to match the characteristic impedance of the "lines" connected to their "ports": the antenna terminals, true port connected to true lines, and the "radiation ports", looking into free space.

In this perspective, the main design objective is to maximise the power transfer, which does not only imply the minimisation of losses but also the maximisation of the power flux in the desired directions. Dealing with travelling waves the minimisation of power loss has two sides: the reduction of ohmic losses in the conductors and dielectrics forming the antenna and the proper matching of the antenna ports to the feeding line(s), on one side, and to free space, on the other. Opposite to non-radiating passive devices, it is not strictly

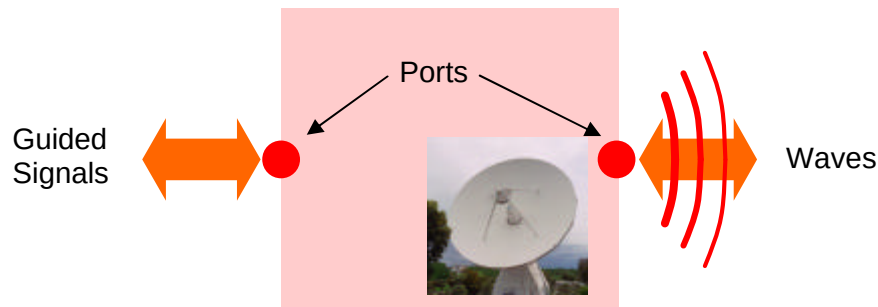


Figure 1.1 – The antenna as a transformer

necessary to minimise these two to maximise power transfer, as it may actually be better to compromise on them to achieve a higher advantage in the focalisation of power flux in the main antenna beam.

#### *An antenna is a filter*

As any other (electrical) system antennas can operate only over a finite frequency band. Antennas have two distinct filtering characteristic one related to the time and frequency domains, the other to the space and spectral domains, i.e. far-field directions and angular frequencies (figure 1.2). Such duality is a consequence of the relations between time and space variations of the electromagnetic field described by Maxwell's equations. In most cases space antennas operate in harmonic regime. Therefore they are typically designed in the frequency domain and considering far-field directions (angles) for their radiation. Design objectives are usually described in terms of performance masks, with a “pass-band” region and one or more “rejection” regions, corresponding to the frequency band and angular region over which the antenna has to operate and those over which signals should not be transmitted through it. Often also the ripple and the derivatives of the response are important, since they have an impact on the antenna behaviour as part of the transmission channel.

#### *An antenna is a transducer*

Both the transformer and the filter view consider the antenna as a black box. An alternative is to see the antenna as an electric current to electromagnetic field transducer. The composition and the shape of the transducer are the means available to the engineer to obtain the desired response. Currents in the antenna must be distributed in such a way

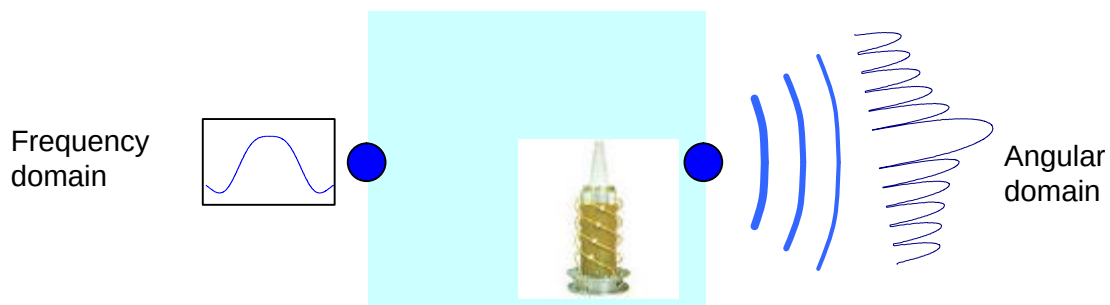


Figure 1.2 – The antenna as a filter

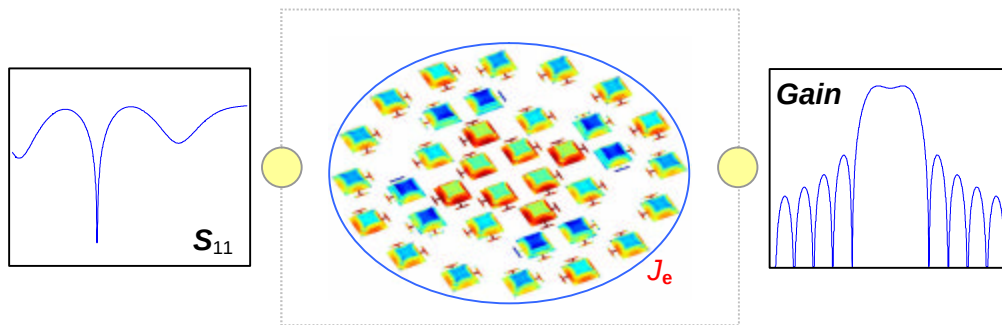


Figure 1.3 – The antenna as a transducer

to generate the desired reactive and radiated field distribution and, at the same time, properly flow in or out the antenna ports (figure 1.3). Clearly this view requires the knowledge of how antennas are actually made and it is the most central to their design.

#### *An antenna is a boundary condition*

Finally, an antenna can be seen, when looking at it from the mathematical modelling point of view, as a boundary condition to be used in solving Maxwell's equations to predict its behaviour. Such perspective may appear rather odd, but it is very important to gain a physical understanding of how antennas work. Clearly this aspect is also important in the development of the computer-based antenna design tool as well as for their proper use. Thus this peculiar perspective, which may at first appear to be a bit remote from the engineer mind, should actually be very present.

### **1.3 Characteristic parameters**

Antennas, as any other multi-port device, can be characterised using a transfer matrix. Focussing on a two-port case, the matrix comprises a forward and a reverse transfer term and two reflection (self) terms, one for each port. For space applications antennas are usually characterised in the frequency domain, a fact that further enhance this decoupling. As a consequence, in the following it is assumed that the antenna is excited by a continuous monochromatic RF signal or electromagnetic wave, i.e. to operate in continuous wave (CW) conditions.

Studying antenna characteristics the transfer matrix terms are seldom, if ever, considered all together. The main reason lies in the very different nature of the input and output ports. The output antenna “port” is free space and has the quite peculiar characteristics of extending all around the antenna itself. While the input port, which supports guided propagation, can be characterised using the quantities normally used for lines (input impedance, reflection coefficient, return loss, etc.), for the “free propagation” port this is impossible. The wave reflection characteristic of an antenna could be quantified by using the (bi-static) radar cross-section, but this quantity is of little or no interest for many applications and certainly for space ones. For the sake of completeness, it has also to be noted that the radar cross-section of an antenna does not provide a complete description of its field reflection characteristic, since it does not include phase information and does

not quantify the near-field effects related to the non-radiating (quasi-static) field components present in the immediate vicinity of the antenna and decaying as  $r^{-2}$  and  $r^{-3}$ .

A complete characterisation is indeed possible using spectral domain representations, i.e. expressing the field as a summation or integral of “modal” components, for example using the Plane Wave Spectrum or the Spherical Wave Expansion. Both forward and backward wave as well as evanescent ones can be included, thus providing a complete description of the “output characteristic” of the antenna. Unfortunately these representations, which are very useful for computer modelling, are of little use to quantify antenna performances with simple parameters. As most space antennas are reciprocal devices and their forward and reverse transfer functions are identical. Directivity and gain, both measuring the ability of the antenna to concentrate its radiation in a limited angular region, are used to characterise these functions. These two quantities are derived from the antenna radiation pattern, i.e. the distribution of power associated to each electric field component radiated by the antenna. *Gain* provides a measure of the ability of the antenna to transfer power from the input port to a point in space [2]. *Directivity* is normalised to take into account only the power actually radiated by the antenna, i.e. to ignore the effect of losses within the antenna itself.

It is quite clear that, as the signal applied to the antenna input port is spread over the whole space surrounding the antenna, a complete characterisation of the transfer function would require the quantification of the field in the whole space, which would be quite difficult to achieve both via measurement and via analysis. The equivalence principle allows the reduction of this requirement to the knowledge of the tangential electric and magnetic field components over any surface enclosing the antenna. This is not a minor step but it is not quite sufficient to obtain a practical solution. For instance it is very difficult to completely measure the field distribution all around the antenna, since it needs to be supported in some way.

Fortunately in the vast majority of cases the radiation is significant only over a limited angular region and it is sufficient to measure or predict the field distribution in this region. In other words, for most purpose, it is sufficient to know the transfer characteristic of the antenna between its input port and a set of points covering a portion of a surface surrounding it. Since antennas are normally placed at a large distance from each other or from the objects from which they receive the signal and this is certainly the case for space applications, the measurement surface is chosen to be a sphere at very a large distance from the antenna (usually called the far-field sphere). The radiation boundary condition (or far-field plane wave condition) applicable at a very large distance from the source of the field (ideally at infinity) imposes a fixed relation between the E and H components of the radiated field:

$$\mathbf{H} = \frac{\hat{\mathbf{r}} \times \mathbf{E}}{z_0}$$

where  $\hat{\mathbf{r}}$  is the propagation direction and  $z_0$  is the free-space impedance. It is then possible to ignore the quasi-static components of the field and to limit the quantification to the sole E (or H) component of the electromagnetic field.

The tangential (and only) components of the electric field over the far-field sphere are fully quantified with two complex scalar functions of two coordinates. Furthermore the absence of the quasi-static components ensure the possibility of sampling their

distribution with no loss of information, making it feasible to characterise the antenna transfer function, at least for what is relevant for practical purposes.

The tangential electric field can be decomposed in various ways into its two complex components, giving rise to a number of different alternatives for describing the transfer function, i.e. the antenna radiation pattern and the gain and directivity derived from it. The most used decompositions are based on the reference system used for the far-field sphere, i.e.  $(q, f)$ , or else make reference to the polarisation (best received field component) of a second antenna, i.e. linear (e.g. vertical and horizontal) and circular (left-hand and right-hand). These descriptions are used interchangeably and can be converted one into another. The circular polarisation components decomposition uses the phase to represent also the time dependence of the orientation of the field vector; therefore its conversion requires full knowledge of both amplitude and phase of the field. Such knowledge is not necessary for all the other conversions, since amplitude and phase maintain their meanings.

In most cases interest is focused on the energy flow in one particular polarisation within a region of space (co-polar component). The antenna gain is then to be as high as possible for the chosen polarisation and within the desired angular range. For the other polarisation (cross-polar component) and outside the angular range the gain is normally required to be below a certain level to ensure that interference and wasteful energy dispersion (in transmission) are avoided.

The antenna gain, other than being function of the position over the far-field sphere, is also a function of the operational frequency, and possibly of other physical parameters, e.g. the temperature. Fortunately all these variations are usually of second order in a good antenna, at least over a reasonable range.

Several other quality figures are used in the design of antennas and are described in the following in combination with examples of their relevance. However there is one commonly applied to all antenna types: *efficiency*. It is defined as the ratio between the gain of a real antenna in given angular sector and the directivity obtainable with a planar aperture with an optimal illumination law and having the same equivalent aperture surface. In most cases the uniformly illuminated aperture is used as reference.

Antenna efficiency is the product of several factors contributing to reduce its gain compared to an ideal antenna of the same dimensions. The first contribution comes, quite obviously, from ohmic losses. These can be further distinguished in conduction or feeding losses, along feeding lines and waveguides as well as within the radiating structure itself, including those in the dielectric if present, reflection losses on reflectors, and transmission losses in lenses or other RF transparent devices. Clearly the latter two contributors may be absent, depending on the type of antenna.

A second contribution comes from the aperture efficiency, which is generally defined only for aperture antennas. However the same concept can be applied to all other types of antennas. The maximum possible gain theoretically achievable by an aperture antenna is directly related to its area,

$$G_{\max} = 4\pi \frac{A}{\lambda^2}$$

and it is obtained for a uniformly illuminated aperture. However, the boundary conditions

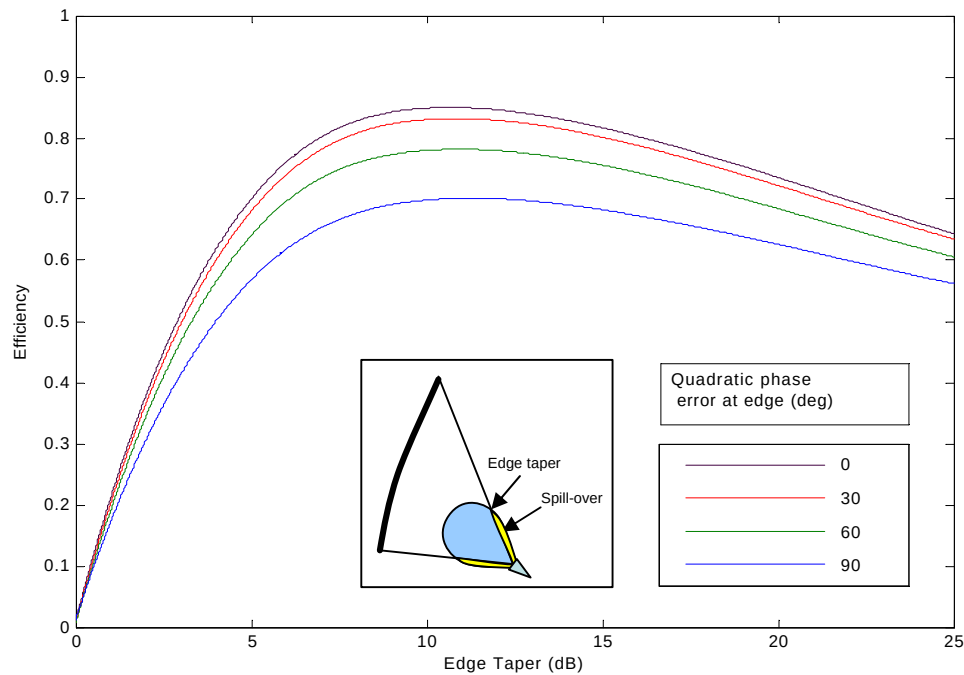


Figure 1.4 – Combined aperture and spill-over efficiency of a reflector antenna

at the aperture edge make it physically impossible to obtain a perfectly uniform field distribution, both in amplitude and phase, and therefore to reach the maximum theoretical gain. For example, a roughly quadratic behaviour of both amplitude and phase is typical of the currents generated by the feed on reflector antennas. For similar reasons the (equivalent) current distribution in a real non-aperture antenna will be different from the best possible one and therefore its efficiency will be always smaller than 1 (also not counting the ohmic losses).

A third category of losses, only present in reflector and lens antennas, comes from the amount of RF radiation produced by the feeder that not being focused by the reflector or lens does not actually contribute to the antenna gain. This is not a true power loss, in the sense that the corresponding amount of power is indeed radiated by the antenna, but in unwanted directions. Often this loss is called spill-over loss, referring to the fact that part of the radiation does not encounter the reflector or lens. Figure 1.4 shows the combined effect of aperture and spill-over losses for a reflector antenna. The optimum efficiency, beside ohmic losses, is achieved for an illumination taper at the reflector edge of about 12dB, for lower edge taper values the spill-over loss dominates, while for higher values the aperture efficiency drops.

A fourth element in the efficiency of reflector antennas is the loss due to the diffusion generated by the small deviation of the reflecting surface from its ideal profile. Lenses and other RF transparent devices also add a contribution due the reflection of part of the incident RF power at both interfaces with free space.

A polarisation efficiency term accounts for the amount of power radiated into the undesired polarisation and a blockage term may account for the effects of part of the

antenna, e.g. a sub-reflector and its supporting structure or a feed, which partly obstruct the antenna field of view.

Finally, it would also be necessary to account for the reflection losses at the antenna input. However for space applications the return loss is (quite) better than  $-20\text{dB}$ , corresponding to an efficiency loss of less than 1% and it is therefore ignored.

In conclusion, the efficiency of antennas can be expressed as follows

$$h = h_{\text{ohmic}} h_{\text{aperture}} h_{\text{spill-over}} h_{\text{reflector-roughness}} h_{\text{lens-mismatch}} h_{\text{polarisation}} h_{\text{blockage}} h_{\text{input-mismatch}} .$$

Usually  $h$  is at least 0.5 for reflector antennas and can reach 0.85 for arrays.

As mentioned above, it is usually required that the antenna radiates most of the power in one polarisation, therefore another rather common figure of merit is *polarisation purity*, which measures the ability of the antenna to radiate power in one desired polarisation and can be derived from the ratio of the directivity for the two polarisations. The definition of polarisation and polarisation purity figure is not as straightforward as it may seem. Polarisation can only be defined in a strict sense for a plane wave, i.e. at a large distance from the field source, and locally, i.e. its definition is not univocally determined over the full far-field sphere. The first consequence of this is that different ways of measuring the antenna radiation pattern result in different definitions of the polarisation [3]. The same applies, clearly, also to the use of different coordinate systems in computations.

In general for space applications the, so called, Ludwig3 linear decomposition is used as reference and circular polarisation is considered as built starting from it. Considering a spherical reference system  $(r, q, f)$  and a decomposition of the electric field vector in  $q$  and  $f$  components, the Ludwig3 polarisation components are expressed as [4]:

$$\begin{aligned} E_a &= E_q \cos j - E_f \sin j \\ E_b &= E_q \sin j + E_f \cos j \end{aligned}$$

Clearly the definition of which one is the desired (co-polar) component is case dependent. Circular polarisations can be defined as follows:

$$\begin{aligned} E_{RHC} &= \frac{1}{\sqrt{2}} (E_a - jE_b) \\ E_{LHC} &= \frac{1}{\sqrt{2}} (E_a + jE_b) \end{aligned}$$

Given the nature of the circularly polarised field component any other right-handed pair of linear orthogonal polarisation, e.g.  $E_q$  and  $E_f$ , can be equally used to construct it.

### 1.4 Categories of antennas

The antenna gain parameter is not used in the same way for all applications. For point-to-point signal transmission, typical of communications and radars, the point values of the gain functions and, possibly, of its derivatives are the most relevant. While using antennas as sensor, i.e. to measure the electromagnetic field present in a region of space, which is typical of remote sensing applications, it is the integral of gain over an angular region to be of importance.

This difference has a significant impact on the solutions adopted in the two cases. Designing an antenna to meet point value specifications implies that great attention has to be paid to all the effects that tend to alter the shape of its pattern, i.e. the details of the distribution of the field over the sphere, for instance optical aberrations are a major limiting factor in reflector antennas as they have a direct impact on the beam shape. On the other hand, meeting integral value requirements tends to stress other characteristics, like the amount of energy flowing outside the reflector, which has a direct impact on the total power contained in the main beam.

Many other parameters and characteristics can be used to distinguish among different antenna types, leaving to many different ways to classify antennas. It is therefore useful to briefly examine the various categories of antennas found in space applications.

A first traditional division, which has been used extensively in the following, groups reflector antennas on one side, and array antennas on the other. This partition is well rooted in antenna technology history and it is justified by the very different physical configuration of the two types of antennas, at least when taken in their simplest and oldest implementations. A parabolic reflector illuminated by a waveguide horn is clearly quite different from an array of slotted waveguide and the same applies to their performances. In space antenna design the distinction, which is routinely used, is more relevant to implementation aspects rather than to performances. As it will become apparent in the following chapters, there are also many situations in which this distinction is rather blurred. For example, the radiation characteristics of a non-focusing single-reflector antenna, e.g. one constituted by an array placed outside the focal region of a parabolic reflector, are intermediate between array and conventional reflector antennas.

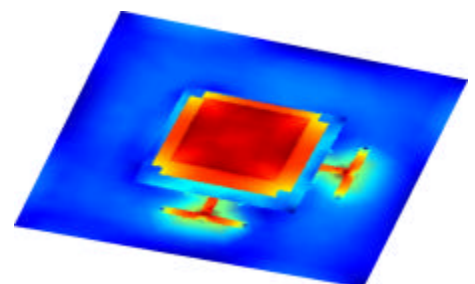
A second classical subdivision is among wire antennas and aperture antennas. It originates from the difference in their electromagnetic behaviour, mainly associated to the geometry of the distribution of the field sources (real or equivalent currents). Another widespread categorisation relates to the way in which antennas operate: resonant and travelling wave antennas. Finally different manufacturing technologies are used as classifiers, e.g. waveguide antennas, printed antennas or wire antennas.

The criterion used in the classification given below is instead essentially empirical. It is mainly based on the complexity of the antenna and on the amount of design flexibility it offers to the engineer, both aspects being intimately linked to the size of the antenna with respect to the field wavelength.

### *Elementary radiators*

Elementary radiators are the simplest radiating structures, e.g. the ones obtained by allowing a line to radiate, typically by modifying its geometry toward the open end. They may be considered to be the building blocks for the construction of more complex antennas. Monopoles, dipoles, Vivaldi antennas and truncated waveguides belong to this category. Other types of elementary radiators are obtained perturbing an otherwise non-radiating transmission line, for instance a slot in a waveguide, a periodically loaded microstrip or a grating (metallic strips or groves) on a dielectric waveguide.

All these radiators are characterised by a rather wide



Electromagnetic model of a "stacked patch" antenna with computed current amplitude distribution

radiation pattern, a low gain ( $\leq 10\text{dBi}^1$ ) and the impossibility of considering them as isolated from the environment. Any diffusing object present in their immediate vicinity alters their behaviour. Therefore, in all practical situations the evaluation of the performance of an elementary radiator must account for the presence of external objects. For example, the feeding line affects the radiation of a dipole; the currents flowing on the outer surface of an open-ended waveguide significantly alter its radiation pattern; also the microstrip feeding patch antennas are often the cause of spurious radiation.

#### *Omni-directional antennas*



TT&C antennas (courtesy of Saab Ericsson Space)

These antennas are used to obtain the widest possible spread of the radiated field (antenna coverage). They are generally rather simple antennas, with a gain lower than  $0\text{dBi}$ , as a result of ohmic losses and input reflections.

These antennas are normally used for service systems (TT&C, data transmission and the like) and the few design parameters available are typically used to maximise the coverage and the polarisation purity. In general, they are wire antennas, single or multiple helices, or slot antennas, with the addition, where necessary, of conical or cylindrical reflectors or directors. Their principal design characteristic is the use of simple but refined solutions to obtain the desired radiation pattern and good input matching without compromising their reliability.

#### *Low-gain antennas*

This class comprises antennas with dimensions comparable to the wavelength and with a directional radiation pattern. They are capable of generating approximately circular beam, possibly in two different polarisations. Their coverage has the shape of a spherical cup with an aperture angle between  $120^\circ$  and  $40^\circ$ . Their gain is below  $10\text{dBi}$ . Note that a circular uniformly illuminated aperture with a diameter equal to the wavelength radiates a beam covering an angle of about  $60^\circ$  at the  $-3\text{dB}$  level and has a peak directivity of about  $10\text{dBi}$ .



X-band corrugated horn for ground station use (courtesy of Tilab)

In general, the few design parameters of these structures are used to control the most important features of the radiation pattern: gain, beamwidth and polarisation purity and to ensure good input matching. These antennas are typically used to form arrays or to feed reflectors, but they can also be used separately, for example to generate a beam covering all the area visible from low orbit satellites.

This class of antennas includes a vast range of shapes. For example, waveguide horns can have a pyramidal, conical or stepped profile; they can have rectangular, circular, elliptical, hexagonal and octagonal sections; they can be smooth or have corrugated walls, and a linear or otherwise shaped

<sup>1</sup> The  $\text{dBi}$  unit refers to the radiation intensity obtained by distributing uniformly over a sphere a unit power. A lossless isotropic radiator fed with  $1\text{W}$  has a gain and a directivity of  $0\text{dBi}$ .

flare. Finally they can be modified inserting dielectric or metal grids, adding irises, pins or steps and be fed at one or more ports. All these variations are used to better control the aperture field distribution and the input impedance over frequency, exciting or suppressing the higher modes.

Individual printed elements also belong to this class. Here, too, the variety is enormous: dipoles, patches of various shapes, slots and small spirals. The dielectric layering and the feeding mechanism -direct, capacitive coupling, inductive coupling- offer other variations as does the possibility to combine a few elements in a stacked configuration to enlarge bandwidth and (slightly) increase the gain.

Other low-gain antenna configurations yet are obtained from wire structures, possibly, backed by a reflecting element (plate or cup) or associated to one or more a directing element. Some examples are helices, cup-dipoles, cross-dipoles, short-backfire antennas and conical spirals. Also here several variations are possible, e.g. replacing the wire with strips, adding passive elements and so on.

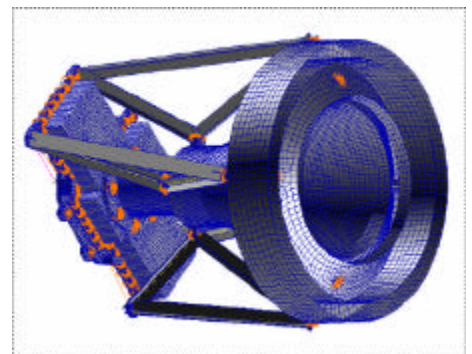
### *Medium gain antennas*

There is a range of intermediate sizes in which the behaviour of wire and aperture antennas is difficult to control, while the use of lenses and reflector is not practical. It extends from about 3 to 10 times the wavelength. At the lower end of the range, reflector and lens antennas tend to be affected by resonance phenomena, i.e. stationary oscillations of the current due to the strong edge diffraction; while at the other end the effect of edge diffraction is still too strong to obtain the desired performances. Horn and wire antennas, instead, are affected by the existence of numerous higher order modes, which are generated by aperture diffraction and other perturbations and are very difficult to control.

In this range of dimensions, i.e. to obtain a gain of about 10 to 20dBi, array antennas are generally preferred, since the consistent number of elements allows a better control of their characteristics. Design problems are however also encountered for arrays, since the behaviour of the peripheral elements needs to be correctly predicted as they form a significant percentage of the total. For example, in a square array with 36 elements, 20 are at the periphery and have neighbours only on one side. As a consequence the equivalent current distribution on their aperture is asymmetrical and the antenna pattern is significantly affected by the lack of translation symmetry on the whole array aperture. In some cases a further ring of passive elements is used to overcome this problem.

Obviously, arrays of this size may also be constructed from a hundred or so small elements such as slots or dipoles, thus reducing the percentage of peripheral elements, although these would still account for 40% of the total, in typical cases. However, this solution complicates the antenna to a degree, which is frequently excessive compared to the performance improvement.

Other solutions are also possible, for example a waveguide horn may be surrounded by a number of chokes to enlarge the effective aperture. A gain of up to about 12-13dBi can also be obtained using stacked patches arranged in a travelling wave structure. The same is possible with wire elements (e.g. the Yagi-Uda antenna).



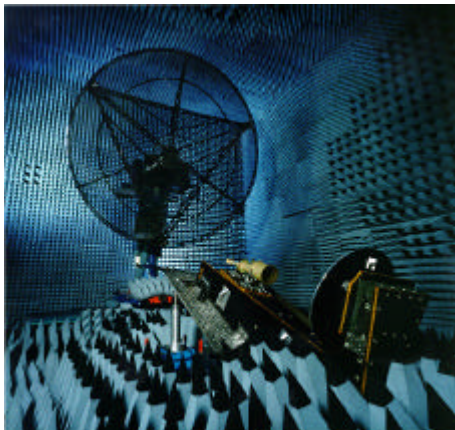
Mechanical model of a horn with beam shaped for global earth coverage

This class of antennas is used for applications requiring a relatively wide angular coverage and limited gain. Examples are the global beams of satellite communications systems, which are used to collect the signals originating from users scattered over large geographical areas (e.g. the whole visible earth surface).

Very often also the arrays used to feed reflector antennas fit in this category. However, they are normally considered as a collection of distinct elements, i.e. their collective radiation pattern is not considered in the design, unless a non-focusing configuration is used.

### *High-gain antennas*

For gain levels ranging from 20dBi to 45dBi, it is possible to use reflector or array antennas; the latter are more flexible, while the former are simpler. The choice between the two is made according to the specific requirements: reflector antennas are typically advantageous when multiple coverages are required or when antenna efficiency is of paramount importance. Arrays are preferred when beam scanning, high reconfigurability of the beam shape or strict control of minor lobes (or sidelobes) are required. However, there are no overriding general arguments for preferring one solution to another and sometimes the choice is made on the basis of mechanical, thermal or manufacturing consideration rather than purely electromagnetic ones.



Ultra-light reflector antenna under RF test (courtesy of EADS-CASA)

In other cases, it is the placement on the satellite, which causes one solution to be preferred to the other. In the overwhelming majority of cases, antennas placed on the sides of the satellite body are of the reflector type, since the absence of RF cables reaching the reflector considerably simplifies the installation as compared to an array having similar performance. For antennas to be placed on the top face of the body, i.e. the one facing the earth, the choice is less restricted, and arrays are more frequently used.

The number of parameters available for these antenna configurations is sufficient to allow a very detailed design optimisation, seeking the optimum for all the performance parameters.

The great majority of communications antennas belong to this class. Their typical requirements, further than a gain of between 20 and 45dBi, are a sidelobe level 20-25dB below the peak gain and polarisation purity better than 20-30dB.

Antennas for many remote sensing instruments, including altimeters, scatterometers, Synthetic Aperture Radar and Synthetic Aperture Radiometer antennas, belong to this category. They are typically (large) array antennas, with the exception of altimeters that most often use reflector antennas.

### *Large antennas*

For some applications, including the most advanced communication systems, antennas with even higher gain are used. These antennas normally use reflector systems to increase the size of the equivalent radiating aperture. There are two main reasons for this. First, the large physical dimensions: an antenna of 100 wavelengths has a maximum directivity of 50dBi and a diameter of about 3 metres at 10GHz. Second, the production problems associated

with the manufacturing and testing of large arrays: the above 100 wavelength antenna would have from 15000 to 20000 elements.

These antennas are often at the forefront of current technology and their typical performances are quite lower than the theoretical maximum. At frequencies of a few GHz, it is the large size that creates problems, while at frequencies above 10GHz the limit is due to the deformations of the reflecting surface caused by thermal expansion. Around 12GHz, antennas of this size are used for direct television broadcasting. At higher frequencies, 20 and 30GHz, they are used for communications, radiometers and radio-telescopes.



15m L-band ground station antenna at the ESA Villafranca (Spain) facilities

The use of antennas of all these categories for specific applications is the subject of the remaining chapters of this book.

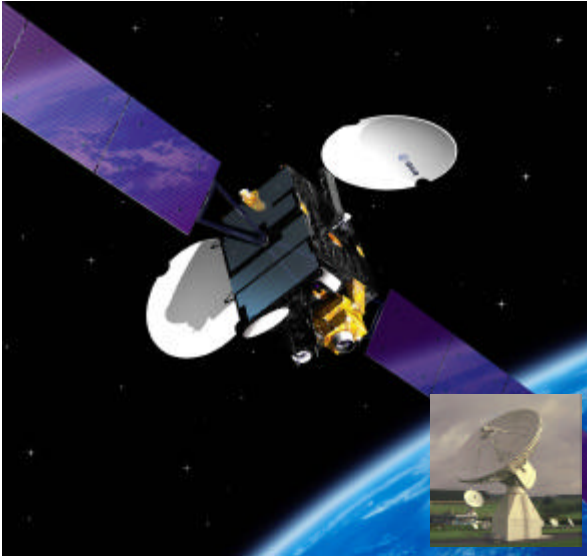
### **References**

- [1] <http://www.intelsat.com/aboutus/ourhistory/yr1960s.aspx>
- [2] "IEEE Standard definitions of Terms for Antennas", IEEE Std 145-1993, March 1993
- [3] See for instance: J.S. Hollis, T.J. Lyon, L. Clayton, "Microwave Antenna Measurements", Scientific-Atlanta, Inc., Atlanta, GA (USA), July 1970
- [4] A.C. Ludwig, "The Definition of Cross Polarization", IEEE AP, January 1973, pp. 116-119



## CHAPTER 2

# *Fixed communication systems*



Artist view of the ESA Artemis satellite with the ESA Redu ground station

The first use of satellites in the field of communications has been for intercontinental telephony and TV relay with the Intelsat fleet. Later on, the traffic growth made it necessary to use satellites to increase the capacity of networks also at the continental, regional and national levels. More recently the diffusion of computers has created an increasing demand for the transmission of data, which is currently met using fixed communication systems.

For all these applications, satellite systems provide a convenient way to serve isolated areas with limited traffic and offer a valuable backup to ground network when they are made inoperative by exceptional events. Moreover for medium and long distance links in new service areas, e.g. developing countries, it is faster to deploy a satellite system rather than to create a terrestrial network.

Fixed satellite links use two ground stations, or terminals, on for each end of the link (figure 2.1). A repeater on the satellite amplifies the signal received from one terminal, translates it to a different frequency band to prevent interference between the ascending and descending paths and transmits it toward the other ground station. Clearly a symmetrical link exists to guarantee the return path. In more recent systems, the satellite repeater is also capable of traffic routing, i.e. to act as a switch to dynamically route the signals coming from one station to several others and vice-versa.

For links between fixed points, it is best to use a geo-synchronous satellite, which orbits around the Earth at such a speed to complete one orbit in 24 hours and therefore remain fixed with respect to the Earth's surface. Such condition is attained by a satellite placed on the equatorial plane at an altitude of 35786km, i.e. on a circular orbit with a radius of 42164km known as the geostationary orbit (figure 2.2). The only limitation is that, in this way, the satellite appears quite low over the horizon at our latitudes.

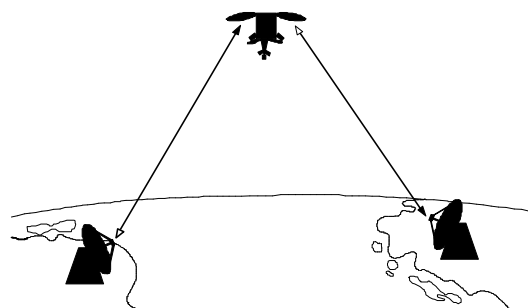


Figure 2.1 – Architecture of a fixed satellite communication link

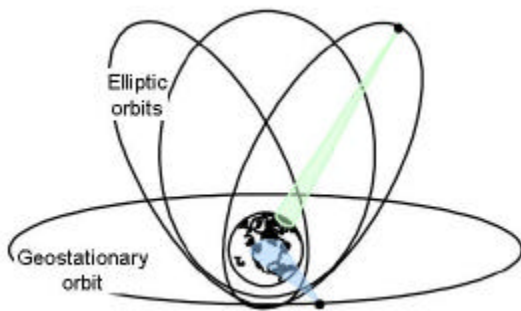


Figure 2.2 – Geometry of geostationary and inclined elliptic orbits

A possible alternative is to place satellites in a sun-synchronous elliptic orbit inclined with respect to the equatorial plane. These are orbits with a fixed position with respect to the Earth-Sun vector, so that the Earth revolves once a day under them. Service is normally provided in the period in which the satellite is near the apogee so that, according to Kepler's second law (the law of equal areas), it moves more slowly along the orbit. The satellite is visible from different points on earth during the 24 hours and a number of orbits placed at

different angles around the Earth axis need to be used to provide coverage of any location for the whole day. For example, the three Molniya orbits, used by the Russians to obtain good visibility at the high latitudes (where most of Siberia is located), are inclined at  $60^\circ$  on the equatorial plane, spaced  $120^\circ$  apart, and each orbit is used for 8 hours.

As for any transmission channel, the number of simultaneous links that can be established between ground and the satellite is limited by the bandwidth of the system and by its signal-to-noise ratio. Assuming that the noise level at the input of the receivers, related to the noise in the two paths and to the noise figure of the receiving system on board and on ground, is invariable, then an increase in the number of available circuits (independent signal paths) requires an increase in the power captured by the receiving antennas. This increase may be obtained, theoretically at least, by using more powerful transmitters and antennas with higher gain. Such solution is viable, up to a certain point, for ground stations, while on board the satellite the situation is totally different. The available electrical power is drastically limited by the solar panel capacity, the need of storing energy to prevent interruptions of service during eclipses, and, last but not least, the output power of the microwave amplifiers. The possibility to increase the gain is considerably limited by the maximum dimension of the antennas that can be accommodated on the satellite. With the evolution of space technology the available power and the maximum size of antennas have increased significantly over the past 50 years, but the limitations above remain applicable at any given point in time, although their thresholds change.

To significantly increase the gain of on-board antennas it is necessary to increase their directivity<sup>2</sup>, so as to concentrate the radiation in a smaller angular sector and thus increase the power received by a second antenna located within it. A directivity increase requires an increase in dimensions, giving rise to a number of problems. On one hand, there is an increase in mass, which is aggravated by the increase in the fuel required for launch, which also increases the launch cost. On the other hand, larger dimensions entail additional mechanical design problems related to the stability and to the vibration modes of the antenna. Clearly larger structures tend to be more flexible and temperature variations will produce larger deformations. At the same they tend to have lower mechanical resonance frequencies and this may cause serious problem during the launch phase. The propulsion system and the air pressure on the launcher produce a very high

<sup>2</sup> Minor gain increases can also be obtained by reducing losses.

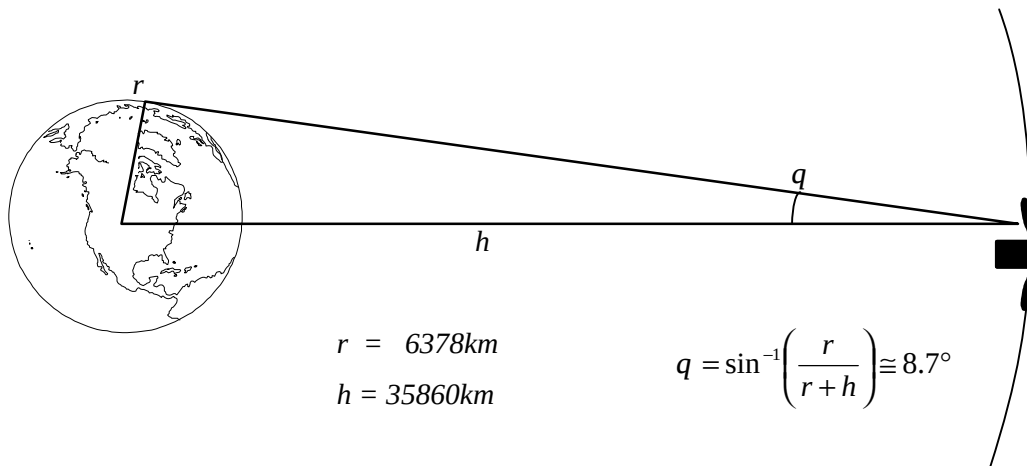


Figure 2.3 – Geostationary satellite coverage geometry

level of vibrations at low frequencies (roughly up to 2kHz), which if coupled with a self-resonance of the antenna (or any other satellite part) may result in very high amplification of displacements and in very high stress levels, leading to possibly catastrophic damage. All these problems are quite challenging, but the most important side effect of gain increase, from the point of view of communication systems, is the inevitable reduction of the beam width associated with an increase of the directivity. For a given altitude of the satellite, higher directivity means a reduction of the area in which the transmitted signal can be satisfactorily received, leading to the need of great changes in the architecture of satellite repeaters and antennas.

Seen from a satellite in geostationary orbit, Earth appears as a disc with a diameter of approximately  $17.4^\circ$  (figure 2.3). A circular aperture antenna, covering the visible part of the Earth and provide 3dB less gain at the horizon (edge of coverage) than at the peak, has a maximum theoretical directivity of approximately 26.5dB and an area of approximately 45 square wavelengths. Assuming, as a first approximation, a uniform illumination of the antenna aperture, placed on the x-y plane, the radiation pattern may be expressed as:

$$D(q, j) = \frac{p d}{l} \frac{J_1(u)}{(u)}$$

where  $q$  and  $j$  are the angular co-ordinates on a sphere surrounding the antenna at a large distance,  $l$  is the wavelength,  $d$  is the diameter of the aperture and  $J_1(x)$  is the Bessel function of the first kind and order 1 and  $u = \frac{d}{l} \sin q$ .

The desired beam shape is obtained if

$$D(8.75^\circ, j) = \frac{1}{\sqrt{2}} D(0, j),$$

which implies

$$\frac{d}{l} \cong 6.7$$

and therefore

$$D(0, j) \cong 21.2 \cong 26.5 \text{ dB}.$$

The gain of a real antenna will be reduced of about 3dB by several factors, compounded by the antenna efficiency.

If the number of circuits is to be multiplied by four, only modifying the antenna, it is necessary to increase the gain by 6dB. The antenna beam will have half of the original diameter, i.e. approximately  $8.7^\circ$ , which is roughly equal to the size of Africa as seen from geostationary orbit. If all the visible portion of the Earth is to be covered, it will be necessary to use an antenna that generates at least seven beams so as to cover it. The resulting antenna will be about twice as large as the first and significantly more complex.

In recent communication satellites, the gain required is such that the antenna main lobe has an angular aperture of a few degrees and sometimes even less than one (figure 2.4). The number of beams necessary to cover a geographically and economically significant area is therefore rather large, i.e. from ten to above one hundred. Since both generating more power and making an antenna with many beams entail increased complexity, weight, risk and cost, one of the most important aspects of the design of a communication satellite is the choice of the best compromise between these two alternatives to increase the number of available circuits. Reflector antennas are used as a rule in this case, since they offer the best mass to diameter ratio.

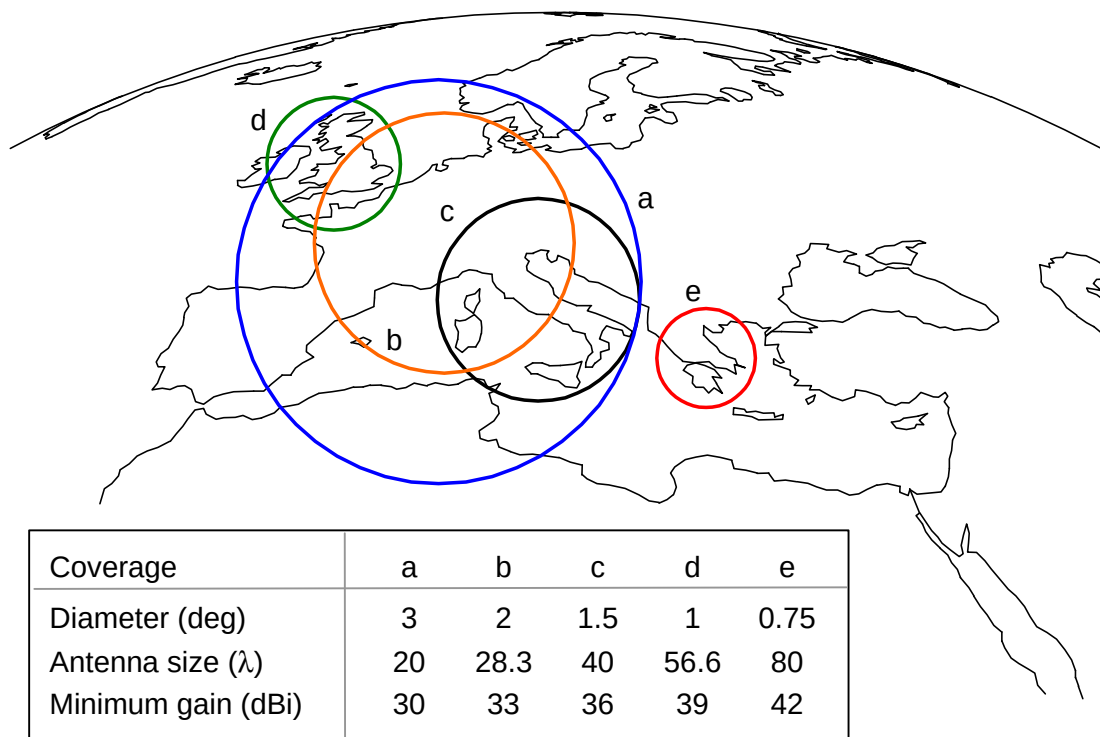


Figure 2.4 – Relationship among coverage diameter, antenna diameter and gain

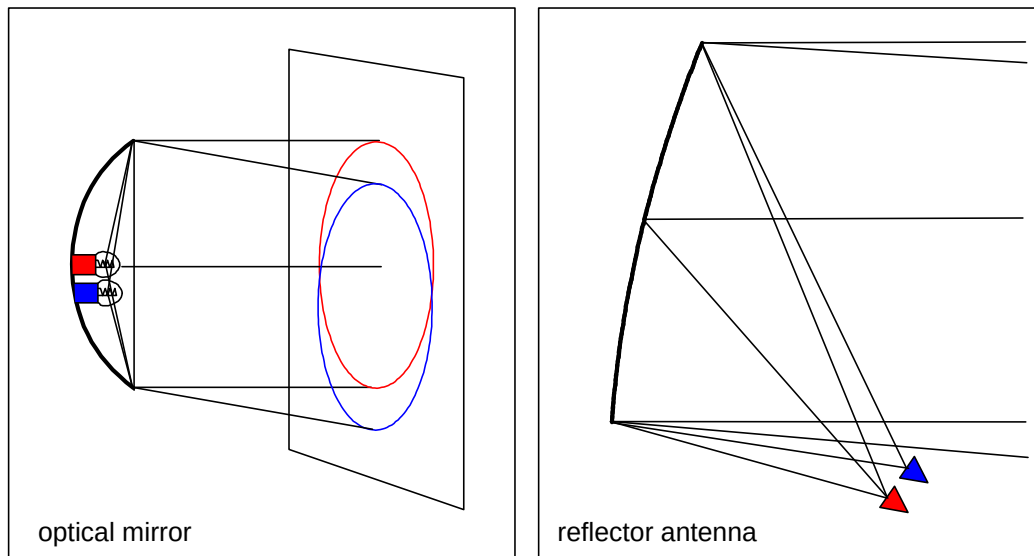


Figure 2.5 - The similarity between an optical mirror and a microwave reflector antenna

### 2.1 Multiple-beam antennas

As a first approximation, the reflector of a parabolic antenna acts as an optical mirror (figure 2.5) and the area of the Earth over which the gain is larger than a given value, i.e. the antenna beam footprint, can be seen as a (distorted) image of the field source illuminating the reflector. Using the laws of geometrical optics (paraxial approximation), a small displacement of the source in a plane perpendicular to the optical axis, usually called the focal plane, will cause the antenna beam to change its direction accordingly.

Therefore placing a number of sources in the focal plane of a reflector antenna produces a number of beams directed in different directions. In this way it is possible to illuminate a number of areas of the Earth of approximately equal size (figure 2.6). Since the integral over the whole sphere of the power flux must be equal to the radiated power, it follows that when two beams of identical angular dimension are used with the same signal their gain is reduced by 3dB. To avoid this it is necessary for the signals to be orthogonal, for instance to have different frequency.

Antennas using several feeds to generate each a separate beam are called multiple-beam antennas and are commonly used in communication satellites because they provide coverage of large geographical areas while maintaining a high power flux. Sometimes each beam is generated by a small group of feeds and each feed may contribute to generate more than one

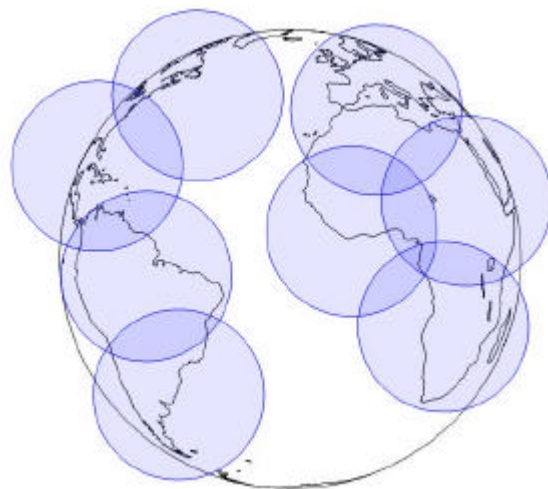


Figure 2.6 – Example of multiple-beam Earth coverage

beam, always at different frequencies. In all cases a separate input port is necessary for each beam, resulting in antennas with several input ports (and several “beam” output ports).

A multiple-beam reflector antenna has typically two or three components: a signal (power) distribution network, also called a beam-forming network, needed in case multiple feeds are used for each beam (and possibly multiple beams use the same feed), an array of feeds and an optical system to focus the radiated power increasing the antenna gain.

The most important limitation in the performances of multiple-beam antennas is caused by the need for the feeds to be relatively close to the focus, in terms of wavelengths, to avoid excessive distortion of the beam generated by them. Clearly this limits the total number of beams than can be produced for a given antenna.

## ***2.2 The transform-chain model***

The principal difficulty in the design of multiple-beam antennas is the large number of degrees of freedom to be controlled to achieve optimum performance. A good design requires the comparison of many different antenna configurations using quantitative criteria based on measurable parameters common to all of them. The conventional characteristic parameters, such as gain, beamwidth, level of the side lobes, etc., are obviously applicable, but they do not provide sufficient information for a complete assessment of the quality of the design. For example, for an equal edge-of-coverage gain and beamwidth, two different antennas may have different peak gains and different gain slope within the coverage.

A way to derive suitable comparison criteria can be formulated using the analogy between aperture antennas and optical systems mentioned before.

A single-reflector antenna may be assimilated to an optical system consisting of a single lens in which an object located in one focal plane is reproduced, more or less undistorted, on the other one (figure 2.7). Similarly a parabolic reflector produces an image of an object placed in its focus at infinity, where its second focus is located. The analogy easily can be extended to more complex cases, for instance a multiple-reflector antenna corresponds to a system consisting of several of lenses. A number of optical elements may be used to reduce the dimensions of the optical system and to correct its aberrations, as it is frequently done in camera lenses (aspheric groups of lenses) and in telescopes (correcting mirrors). An optical system consisting of a number of lenses can be approximated by a single equivalent lens and it is therefore possible to define an equivalent reflector for any multiple-reflector system. If the last reflector is parabolic then also the equivalent reflector is and this is by far the most common case in space antenna applications.

A single lens can be modelled with good approximation by a pair of Fourier transforms, a direct one between the focus and the lens and an inverse one between the lens and the image plane. This approximation is fairly satisfactory for common optical systems, but in microwave antennas the smaller ratio of the aperture diameter to the wavelength, which is associated to more significant diffraction effects, makes it inadequate for the calculation of the antenna performance. However, it still provides a suitable conceptual model for design purposes, as explained in the following.

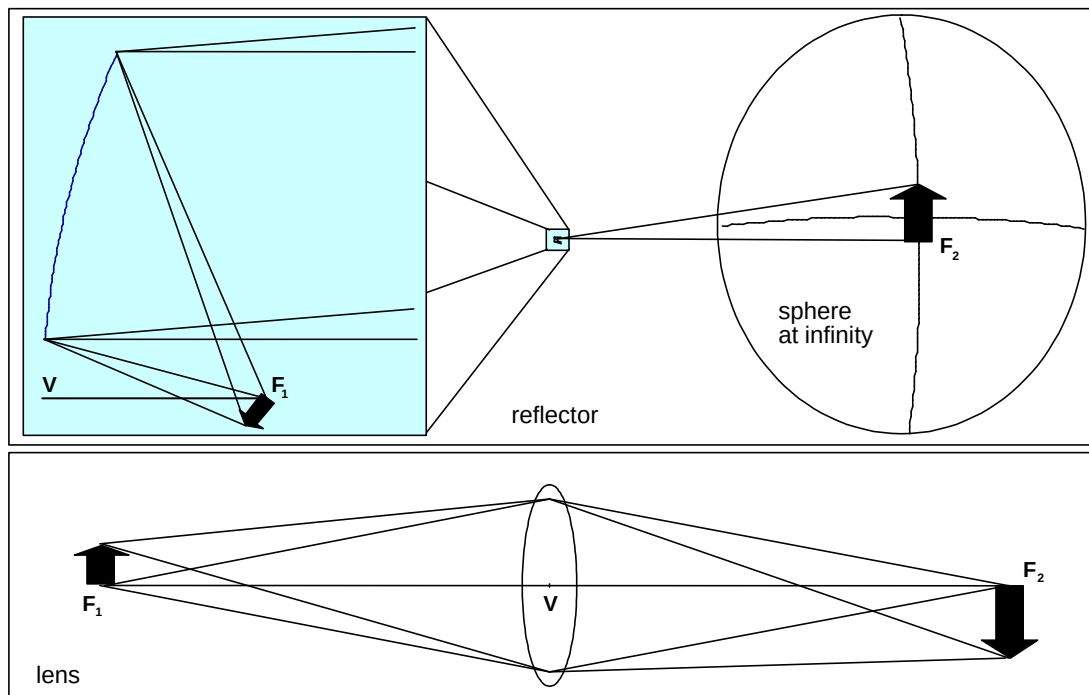


Figure 2.7 - The analogy between a parabolic reflector and a lens

The function of a multiple-beam antenna is to convey the signals present at its input ports to a number of angular regions (beams), which may be considered as output ports. Following the equivalence principle, the field concentrated in each beam may be thought of as radiated by an equivalent current distribution on a surface enclosing the antenna. For large antennas, these currents may be limited, with negligible error, to those present on the antenna aperture, which are related by a Fourier transform to the radiation intensity (or directivity) distributions in the far field. To complete the model, it is necessary to include another transform, inverse of the preceding one, which relates the signals present at the input ports to the current distributions over the aperture. Considering the simple case in which each feed is fed with the signal for one beam, the distribution of equivalent currents present on the feed aperture generates a distribution of currents on the reflector that, in turn, gives rise to an equivalent distribution over the antenna aperture. The overall result is a Fourier-transform pair relating the input ports to the output ports (beams).

Of course, in this simple case the use of the Fourier transform pair adds little to the knowledge of the antenna behaviour. However many variations are possible and the model helps understanding their behaviour. Consider for instance the fact that the radiating elements provide a transition from a discrete space (the input signals) to a continuous one (their radiated field distribution) giving rise to the possibility of combining continuous and discrete Fourier transforms pairs to obtain various effects. Furthermore the beam-forming networks may implement a large class of different discrete transforms, of which the DFT (Discrete Fourier Transform) is only a special case. The first transform (input port to aperture) may also be split in two portions, realising part of it by means of the power distribution network and part through free space propagation. In this way each

beam is generated by more than one feed and each feed is shared among several beams. Obviously, such arrangement may also be described as the cascade of two independent (non-Fourier) transformations, as already mentioned.

Other pairs of direct and inverse transforms may also be added, e.g. by adding a second reflector (sub-reflector). A good reason for such modification is to increase in the number of degrees of freedom of the antenna and therefore its flexibility.

### 2.3 Design parameters for multiple-beam antennas

The transform-chain model is useful to gain a better understanding of how multi-beam antennas work and to draw comparison among different solutions, but also to obtain the means to quantify their respective advantages and limitations.

The inverse proportionality law between the size of the beam footprint and the minimum gain obtainable within it makes it possible to define a first figure of merit for multiple-beam reflector antennas: the area-gain product. For a given antenna geometry and a specified level of illumination at the reflector edge, the product of the solid angle occupied by the beam called for simplicity the coverage area, and the minimum gain is, to a first approximation, constant and gives a measure of the ability of the antenna to concentrate power within the coverage. It is therefore possible to compare different solutions on the basis of this value, which increases with the efficiency of the antenna. It is worth noting here that, except in the case of a circular beam, the antenna having the best area-gain product is not necessarily the one providing the highest peak gain.

The variations of the area-gain product are essentially related to the aberrations of parabolic reflector antennas and are of second order. Still they may be significant, as for instance a gain variation of 0.5dB implies a RF power loss of about 11%, which becomes at best a 22% DC power loss due to the RF amplifier efficiency. Aberrations are

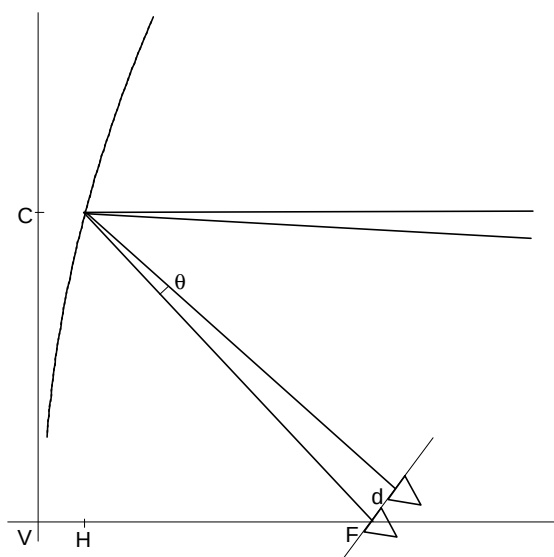
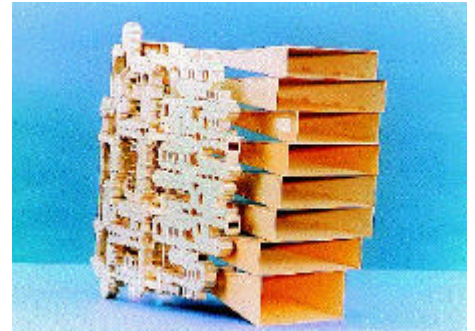


Figure 2.8 – Relation between spacing of feeds and beam pointing

responsible for the deviation from the ideal behaviour described by the Fourier transform model and cause an imperfect reproduction of the feed array geometry in the shape and arrangement of the antenna beams. First, the angular distance among beams generated by equally spaced feeds decreases moving away from the most central one. Further aberrations are typical of offset reflector antenna, i.e. those in which the reflector rim is not centred on the paraboloid axis (figure 2.8). Large displacements of the feed in a direction perpendicular to the plane of symmetry of the antenna generate a tilt of the beam also along this plane. Instead, large displacements of the feed parallel to the symmetry plane result in different tilts of the beam, depending on the direction of feed displacement. This behaviour is affected by the focal length of

the reflector, the larger the better, and it is measured by the Beam Deviation Factor, a quality figure that for paraboloid reflectors is slightly smaller than unit and clearly depends on the beam pointing (or feed displacement).

Another factor having considerable influence on the design of multiple-beam antennas is the ratio between the angular separation of the beams, seen from the centre of the reflector, and the angular separation of the corresponding feeds, which is evidently related to their spacing and the focal length of the reflector (figure 2.8). This ratio, which usually included in the beam deviation factor (as its minimum value or its average, depending on preferences), gives an indication of the amount of beam aberrations due to the reflector geometry present within a given angular coverage.



Feed array for contoured beam coverage (courtesy of Alcatel Alenia Space)

The spacing of the feeds, or their diameter if they are circular and adjacent, can not be chosen freely. In the first place, it is not possible to make feeds much smaller than a wavelength –the exact limit depends on the aperture shape-. Second, higher order modes can easily appear in large feeds. They are difficult to control and adversely affect the performances. Furthermore, the size of a feed has a direct effect on the width of the major lobe of its radiation pattern, and therefore on the level of radiation at the edge of the reflector. In turn, the level of radiation at the edge has a direct impact on the antenna efficiency, a second-order effect on the gain and a first order one on the sidelobe level. Having a feed spacing much larger than their diameter also reduces the antenna efficiency. To better understand this behaviour it is useful consider an antenna operating in receive mode. The diffraction figure generated by the reflector on the focal plane for each beam determines the size of the feed required to capture the maximum amount of energy. Since the beam deviation factor is close to one and the beams need to be slightly overlapped to ensure a continuous coverage with an acceptable gain ripple (usually between 2 and 3dB) it is clear that the feeds need to be closely packed.

Frequently, to simplify manufacturing, the feeds are positioned with their axes parallel to each other. A feed placed at the focus is usually pointed at the reflector centre so as to make the currents on the reflector as symmetrical as possible. In a large array with parallel axes the peripheral feeds will instead point towards reflector points very far from its centre, generating non-symmetrical currents on the reflector with a consequent loss of gain, usually called scan loss, since it is associated with the movement of the beam from its axial position. This loss is due to two factors, the asymmetry itself, which causes a spreading of the beam, and the increase in the amount of power not intercepted by the reflector (spill-over), due to the lateral movement of the feed, which causes an unbalance in the illumination levels at the edge. The asymmetry of currents also causes a rise in the crosspolar of linearly polarised antennas and a squint in circularly polarised ones, which changes with the polarisation. This latter effect, which also present in the axial beam of an offset reflector, is due to the non-uniformity of the relative phase of the two orthogonal (in space and time) current components forming the circular polarisation across the reflector surface.

To minimise these effects the feeds of a multiple-beam reflector antenna are best placed in a portion of the focal plane that is a small fraction of the reflector area. The extension of this area varies with the ratio between the focal length and the diameter of the reflector, called the  $f/D$  ratio. An increase in the  $f/D$  ratio is accompanied by an increase in the area subtended in the focal plane by a given solid angle and a decrease of the aberrations. Unfortunately, at the same time there is an increase in the size of the feed aperture required to keep the radiation level constant at the reflector edge so also their size and mass increases.

When the antenna gain, which is related to the reflector diameter, has been determined, and an approximate focal length has been chosen, the spacing of the feeds may be varied within narrow limits, and consequently the distance of the beam footprints is fixed. A solution to this problem is based on the use of more than one feed to generate each beam already mentioned before. In this way, by choosing an appropriate number of feeds for each beam with a suitable value of the  $f/D$  ratio, it is possible to meet the conflicting requirements of a wide angular coverage obtained with a large number of beams.

The use of antennas with large  $f/D$  has the disadvantage of requiring very elongated antenna geometries that are not easy to accommodate on the spacecraft (with the exception of the new generation of large platforms). Typical values of the  $f/D$  for offset reflector antennas are between 0.5 and 1.5; while for centred ones the value is usually lower (0.25 to 1).

## **2.4 Reconfigurable antennas**

The next step in the evolution of the multiple-beam antenna arises from the idea of making the number of circuits allocated to each beam flexible. The requirement for such flexibility comes from several factors. First, the communication traffic tends to fluctuate significantly and peaks may not normally occur at the same time in all places over areas of the size of Europe or even smaller. Thus power consumption can be reduced by dynamically allocating channels (i.e. groups of circuits) to beams so as to supply more bandwidth and power where required, taking them from areas where they are not used. This solution also enables a more efficient operation of power amplifiers, since they can operate close to maximum capacity for a higher average time. Second, some satellite systems, such as the Intelsat and Eutelsat ones, are designed to have different satellites in different orbital positions and a different coverage for each of them. If there is the possibility to change the shape of the coverage and the channel-to-beam allocation then the same antenna can then be used for all orbital positions, with significant saving in the development and operational cost (a single spare satellite can be used to replace any element of the fleet). Third, it is often difficult to reliably predict the geographical distribution of demand over the satellite life-time with the advance required for its development and launch; therefore the possibility to change the allocation after launch is quite interesting from an operational point of view. Finally, to provide communication services with adequate availability and quality a system must be designed to operate in all meteorological and load conditions. If the number of circuits associated with each beam is fixed, each repeater chains must be sized to operate in extreme conditions and therefore operates in non-optimal conditions for most of the time.

Various methods are used to obtain a flexible allocation of circuits or channels to beams. One extreme solution is to use base-band switching matrices similar to those used in

terrestrial switched networks, so that the antenna has in fact fixed beams and signal routes among them are dynamically modified. The other extreme is to the use of variable beam-forming networks that change the routing of signals among feeds and beams. These are designed using the properties of the transform chains described above.

Let us consider the simple network illustrated in the lower part of figure 2.9. It can be used, in combination with a reflector antenna fed by two horns, to generate two separate coverages on earth, as illustrated in the upper part of the figure. The part of the network inside the box at the bottom is a single waveguide component, known as 3dB directional coupler or 3dB-90° hybrid coupler. It splits the signal entering one (input) port between two other (output) ports, while the power reaching the fourth port is ideally zero. A signal entering this second and independent (de-coupled) input port is treated equally, since the device is symmetric. The only change is in the relative phase of the signal pairs.

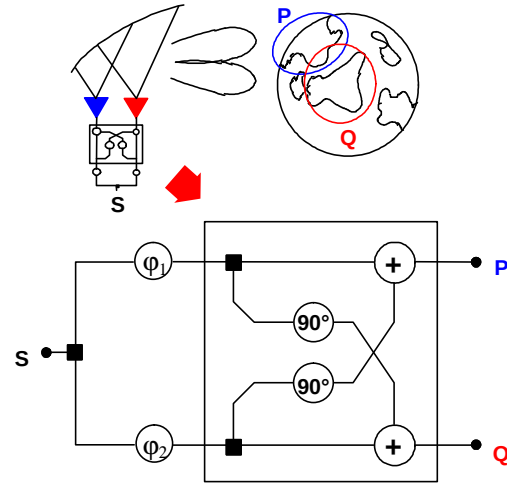


Figure 2.9 – A basic variable beam-forming network (variable power divider)

The complete network shown in figure 2.9 adds two phase shifters, one at each input port of the 3dB coupler and a power divider in front of them (which could actually be realised by a second 3dB coupler with one input port closed on a load). When a signal  $S$  is applied to the input, the output signals  $P$  and  $Q$  are determined by

$$P = \frac{1}{2} S (e^{jj_1} + e^{j(\frac{P}{2} + j_2)})$$

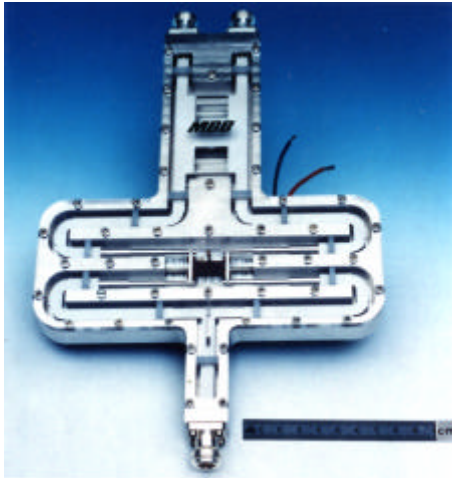
$$Q = \frac{1}{2} S (e^{j(\frac{P}{2} + j_1)} + e^{jj_2})$$

If  $j_1 = 0, j_2 = -\frac{P}{2}$ , then  $P = S$  and  $Q = 0$ .

If  $j_1 = -\frac{P}{2}, j_2 = 0$ , then  $P = 0$  and  $Q = S$ .

Therefore an input signal can be routed to one or the other beam of the antenna by simply switching the phases of the two phase shifters between two values. Actually, since the relative phase of  $P$  and  $Q$ , i.e. of the two beams, is usually irrelevant, a single phase shifter is sufficient, for example by letting

$$j_1 = 0, j_2 = \pm \frac{P}{2}.$$



Variable power divider (courtesy of EADS-Astrium)

This relatively simple network can actually do more. For intermediate values of  $\phi_2$  the signal  $S$  is distributed between the two outputs  $P$  and  $Q$  with a continuously varying ratio. As the same signal is routed to the two beams they combine into a single beam one, which shape is uniformly morphed between beams  $P$  and  $Q$  by changing the phase.

Duplicating input section of the network so as to have two separate dividers and phase shifters in front of the 3dB coupler, it is possible to route two independent signals, i.e. signals at different frequencies, to the two moveable beams in a completely independent way. Note that such modification requires the introduction of multiplexing filters to combine the signals at the input of the hybrid combiner, to ensure that input signals are only routed toward the combiner and not into the adjacent dividers as well.

These concepts can be easily extended to a beam-forming network with  $N$  ports, for example, using a network that implements a discrete Fourier transform. Applying signals with a suitable linear phase progression at the inputs of such network results in the selection of a particular output port and of a particular beam. It is interesting to note that such behaviour is similar to the one of a phased array, as it can be seen describing the two antennas with the transform-chain model discussed previously.

Up to this point only reflector antennas illuminated by an array of feeds placed on the focal plane have been discussed. However there are other four possible solutions to obtain multiple-beam antennas with reconfiguration capability, each offering different advantages, mainly by concentrating complexity in different components. In the first place, array antennas can generate multiple independent beams by using a separate power division network for each beam. This removes the need of a reflector and offers much higher flexibility in the position of the beams. Each beam is formed by properly phasing the signals that feed the individual elements. A linear phase progression across the array aperture points the beam in the corresponding direction, while higher order variations change the shape of the beam (to a certain extent). Each beam can be moved to any direction independently from the others, attaining the maximum possible flexibility. The obvious drawback is in the complexity of the beam-forming network, which increases as the product of the number of array elements by the number of beams.

An alternative choice, which partly overcomes this problem, is the use of digital beam forming, in which the phase shifts are generated digitally at intermediate frequency, using the properties of complex signals. After phase shifting and, possibly amplitude weighting, the signals for all beams are combined, frequency converted and amplified, separately for each radiator. In this way the multiple beam-forming networks are completely eliminated and complexity is moved to the transmit-receive modules, placed behind each radiating element of the array, and to the digital processor. A third solution, still aimed at partially avoiding the complexity of an array, while retaining some of the extra flexibility, is to use magnifying (or hybrid or imaging) antennas, in which two reflectors are used to produce an enlarged virtual image of an array (figure 2.10). Since the optics is equivalent to a chain of two back-to-back Fourier transform pairs, such antenna operates, in a first

approximation, in the same way as the isolated array (see also section 3.6). But it has a higher gain and fewer elements are required for the same coverage and number of beams. The reduction of complexity is paid with a potential loss of gain due to the presence of grating lobes, i.e. additional lobes approximately equal in gain and shape to the main beam arising from the poor sampling of the antenna aperture. The virtual magnified elementary radiator is by necessity much larger than one wavelength (else the gain in complexity with respect to a normal array would be marginal), therefore the continuous aperture distribution required to generate a given beam will be sampled with a discontinuous distribution obtained combining the enlarged images of the aperture distributions of each element of the feed-array. These discontinuities combined with the large size of the virtual elements (in terms of wavelength) result in the presence of aliases of the main beam at the directions where the contributions from the virtual array elements add in phase. As a result the gain for any beam pointing away from the antenna axis will be reduced. Nevertheless this solution is interesting when only a small scan angle is required to cover all the desired area.

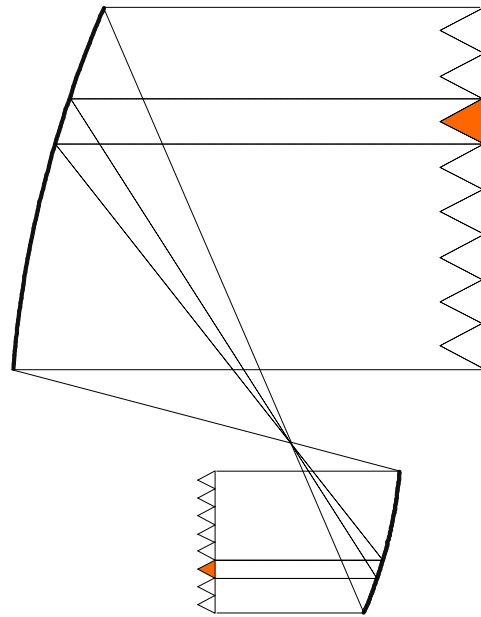


Figure 2.10 – Magnifying antenna

Finally another way to reduce the complexity of the beam forming in an array is to group elements. This solution can be ideally derived from the imaging antenna by absorbing the transform chain, which is equivalent to its optics, into the transformation operated by the beam-forming network. In other words, instead of creating a virtual enlarged image of a single radiating element, several elements of a larger array are grouped into blocks fed by a fixed distribution network (sub-arrays) creating physically larger “elements”. The limitations are clearly the same and the advantage is the removal of the reflector. The increased number of radiating elements need not be a problem in itself since they may, for instance, be realised using printed technology. The major drawback, with respect to the reflector based solutions is that the array will be quite large and long run of cables or other feeding lines are necessary to feed the sub-arrays, introducing significant ohmic losses and mechanical complexity.

## 2.5 Beam-forming networks

All types of multiple-beam antennas require a more or less complex power distribution network to route each channel to the radiators that generate the associated beam. The only exceptions are fixed-beam (non-reconfigurable) antennas with a single radiator for each beam.

A beam-forming network can be defined as a device having  $N$  input ports, associated with the same number of channel groups, and  $M$  output ports, associated with the same

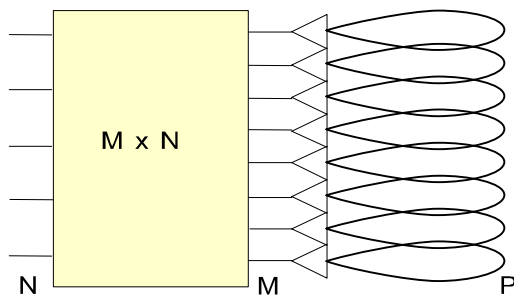


Figure 2.11 – Scheme of a generic beam-forming network

number of radiators. The device is characterised by the capacity of generating a number  $P$  of distinct beams, which are independent of (orthogonal to) each other so that each channel can be independently routed to one, and only one, of these beams. In the simplest case (Figure 2.11),  $P$  and  $M$  are equal, each feed generates one beam and each beam can be generated independently of the others. More generally, each beam can be generated by means of 1 to  $M$  radiators, and the beams may share one or more

feeds. A beam-forming network may be modelled by means of its scattering matrix  $S$  or, provided that it is reciprocal, by its transmission matrix  $T$  formed with the sole forward terms of the scattering matrix  $S$ . The matrix  $T$  provides the information on the input-output transformation of the signals realised by the network. In a reconfigurable antenna, the signal routing and the  $T$  matrix are modified by adjusting variable power dividers, phase shifters and attenuators (or amplifiers), the latter being avoided in transmit sections.

In a multiple-beam system, it is essential to minimise the crosstalk (interference) among different beams. It is therefore important to ensure that the signals from each antenna port are routed only to the outputs associated to the proper beam, and conversely that the signal from a given beam is routed only to the port associated with it. Moreover, to avoid power losses other than the inevitable ohmic losses, input signals must all reach the desired output and they must not be reflected toward any of the input ports.

While the presence of an undesired signal in any beam is simply related to the non-ideal behaviour of the beam-forming network, internal reflections have more interesting consequences on the characteristics of the matrix  $T$ , which are worth a closer examination. Consider a generic ideal passive network, reciprocal and free of ohmic losses (figure 2.12). The network is fed with two signals  $I_1(t)$  and  $I_2(t)$ , applied to input ports 1 and 2 respectively. When applied separately they produce the outputs  $O_1(t)$  and  $O_2(t)$  respectively, having powers  $P_{O1}(t)$  and  $P_{O2}(t)$ . If the signals are considered as two particular input vectors of the network:  $I_1 = [I_1(t), 0]$  and  $I_2 = [0, I_2(t)]$ , it is evident that they are orthogonal with respect to the normal internal product for complex vectors ( $\langle x, x^* \rangle$ ),

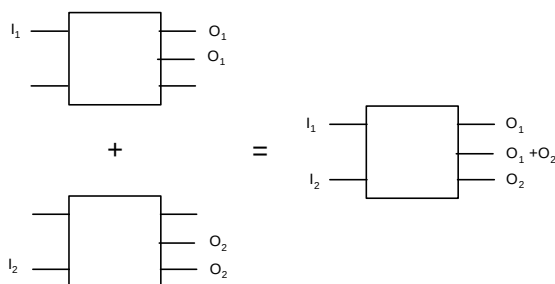
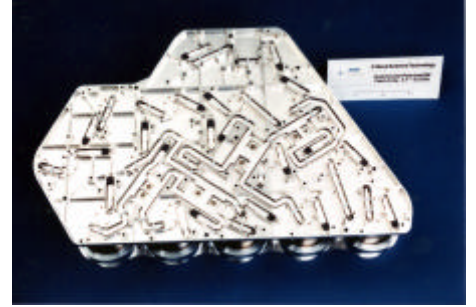


Figure 2.12 - Combination of signals in a beam-forming network

i.e. they have no coupled energy content. The total output power of the network  $P_O(t)$ , found by applying the two inputs simultaneously, may be expressed as the sum of three terms, generally all different from zero, the powers  $P'_{O1}(t)$  and  $P'_{O2}(t)$  associated with each components of the combined signal and equal, up to a multiplying factor, to  $O_1(t)$  and  $O_2(t)$ , plus the correlation term  $P'_{12}(t)$ . If this last term is not zero, the powers  $P'_{O1}(t)$  and  $P'_{O2}(t)$ , actually present at the output of the network, are smaller than  $P_{O1}(t)$  and  $P_{O2}(t)$ .

In other words, part of the input power “disappears” and it is in fact reflected at the input or dissipated in the network. The network is therefore not orthogonal, since it does not preserve the orthogonality of the input signals. Such (undesirable) property is directly reflected in the properties of the matrix  $T$ .

To ensure that a network has no losses other than the ohmic ones, the network must convert orthogonal inputs to orthogonal outputs and preserve their norm, i.e. its matrix  $T$  (purged of ohmic losses) must be unitary or Hermitian. The number of orthogonal input and output vectors, i.e. the number of independent beams, that a network can handle, is equal to the rank of its matrix  $T$ . One particular set of orthogonal inputs are the eigenvectors of  $T$ .



C-band beam-forming network in bar-line technology (courtesy of EADS-Astrium)

An example is useful to illustrate the great flexibility obtained with an appropriate use of these concepts. Consider a network  $R$  with 4 inputs and 4 outputs, passive, reciprocal, symmetrical, perfectly adapted and free from ohmic losses, implementing the following transmission matrix  $T$ , which is an extension to 4 ports of the 3dB-90 hybrid coupler:

$$T = \frac{1}{2} \begin{bmatrix} 1 & j & j & -1 \\ j & 1 & -1 & j \\ j & -1 & 1 & j \\ -1 & j & j & 1 \end{bmatrix}$$

The inverse path, from the output to the input ports, is characterised by the matrix  $T^{-1} = T^*$ , i.e. by the Hermitian adjoint (conjugate transpose) of  $T$ . The vector

$$\mathbf{U}_1 = \frac{1}{2} \{-j \quad 1 \quad -1 \quad -j\},$$

which corresponds to signals of equal amplitude and different phases at the feed array ports, is generated by the vector

$$\mathbf{I}_1 = \frac{1}{2} \{1 \quad 0 \quad 0 \quad 0\},$$

while the vector

$$\mathbf{U}_2 = \frac{1}{2} \{-j \quad -1 \quad 1 \quad -j\}$$

is associated to

$$\mathbf{I}_2 = \frac{1}{2} \{0 \quad 1 \quad 0 \quad 0\}.$$

Thus feeding one port produces an equal-power excitation of the four feeds. The output vectors  $\mathbf{U}_1$  and  $\mathbf{U}_2$  are orthogonal with respect to the internal product  $\langle x, x^* \rangle$  and are generated by two orthogonal input vectors  $\mathbf{I}_1$  and  $\mathbf{I}_2$ . Simple linear algebra considerations on  $T$  lead to the conclusion that the network  $R$  can generate four orthogonal outputs at

the same time and that sets of four orthogonal inputs form the basis of the space of all possible inputs. For example, in addition to  $\mathbf{I}_1$  and  $\mathbf{I}_2$ , we may consider the pair

$$\mathbf{I}_3 = \frac{1}{2} \begin{Bmatrix} 0 & 0 & 1 & 0 \end{Bmatrix}, \mathbf{I}_4 = \frac{1}{2} \begin{Bmatrix} 0 & 0 & 0 & 1 \end{Bmatrix}.$$

With the simple choice of input vectors made above each port corresponds to a different excitation law of the feeds and to a different beam. But many other choices are possible using orthogonal sets of linear combinations of the vectors  $\mathbf{I}_1$ - $\mathbf{I}_4$ . For example, if the vectors  $\mathbf{U}$  are used at the input of the network R the vectors  $\mathbf{I}$  are obtained at the output, and it follows that other four orthogonal beams may be generated, each using a single feed. The beams of the first set, generated by all four feeds, will have a larger footprint (coverage) than those of the second set, so that the same antenna can be used to generate coverages of different sizes. In fact, using the network R it is also possible to generate beams using only two feed, for example

$$\mathbf{I} = \frac{1}{4} \begin{Bmatrix} 1-j & 1-j & -1-j & -1-j \end{Bmatrix} \Rightarrow \mathbf{U} = \frac{1}{2} \begin{Bmatrix} 1 & 1 & 0 & 0 \end{Bmatrix}.$$

Not all pairs of feeds can be used in this way if it is desired to have input signals of equal amplitude, which allows the use of equal power amplifiers at the optimum operating point. Obviously beams from different sets are not orthogonal and must be used with different frequencies or coding to avoid their recombination in a single beam.

If the orthogonal network R is connected to an antenna such that orthogonal vectors correspond to beams with orthogonal field distributions, i.e. such that the integral of the mixed terms of the combined Pointing vector is zero, then the antenna will generate N orthogonal beams. Unfortunately, the orthogonality of beams is a purely theoretical concept, as the diffraction patterns produced by an antenna do not constitute an orthogonal basis on a spherical surface. However using the same level of approximation of the transform-chain model, it is possible to assume that individual beams are orthogonal to each other on a portion of the sphere that intercept almost all the power radiated by the antenna for each of them. Such approximation is usually quite good for satellite antennas since they need to have low sidelobes to minimise interferences and maximise their efficiency. Often an even more crude approximation is used for rectangular apertures, e.g. arrays, by assuming the beam pattern is given by

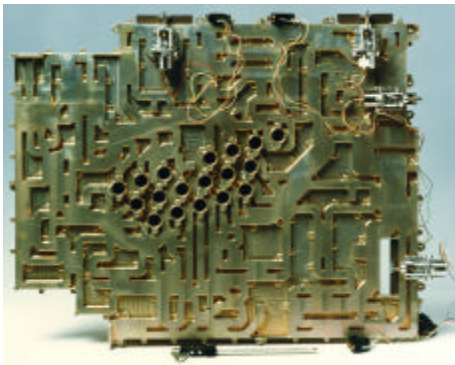
$$\frac{\sin[p(u-k)]}{p(u-k)} \frac{\sin[p(v-k)]}{p(v-k)}, k=1,2,\dots,$$

where

$$u = \sin(q) \cos(j), v = \sin(q) \sin(j).$$

Such beams are, strictly speaking, orthogonal only over the complete u-v plane, i.e. including values of k corresponding to complex angles (i.e. wave numbers of evanescent waves on the aperture).

In practice, the use of such approximations in the design of antenna is helped by the fact that the beams usually operate at different frequencies or with different polarisation, thus their orthogonality is ensured by other



Beam-forming network with motorised variable power dividers (courtesy of EADS-CASA)

means, at least for the adjacent beams. This choice would in all cases be necessary to counteract the rather large lack of orthogonality among adjacent beams that occurs to maintain the gain ripple across the coverage below 3dB, as it is usually the case, as well as non-ideal behaviour of the beam-forming network. The same frequency and polarisation are only used for beams whose angular separation is sufficient to ensure that a good isolation, i.e. orthogonality, exists. The orthogonality of the signals generated by a beam-forming network can not usually be stated in such a simple way: the networks may be very complex, and the number of radiators may vary from beam to beam, as may the number of elements common to various beams. Still the transmission matrix model of the beam-forming network can be extended to these cases and used as a tool for the initial design of the antenna. For instance, it allows an efficient first-order evaluation of the impact of amplitude and phase errors induced by losses, manufacturing errors, frequency dependence and mismatches within the network on the coupling among input and output ports by just checking the deviation from orthogonality of the output vectors.

From a technological point of view, beam-forming networks can be realised in various ways, e.g. using waveguides, coaxial lines (typically with square section, for simplicity of construction), microstrips or striplines. In all cases, particular attention has to be paid to the minimisation of losses, to the mechanical stability and manufacturing tolerances to reduce to a minimum the reflections inside the networks, and finally to the design of components (phase shifters and couplers, fixed and variable) to obtain the maximum bandwidth. Further attention is required when dealing with high power signals.

An aspect that needs to be carefully considered in the design of beam-forming networks for space application is thermal expansion. Ohmic losses inevitably produce heat that in turn induces non-uniform thermal expansion. The differential expansion among different portion of the network causes distortions, which result in an alteration of its electrical parameters. In network based on “coaxial” or quasi-TEM lines in which the centre conductor is electrically and thermally isolated and has the higher current densities and losses are especially sensitive to this problem. Therefore quarter-wavelength high-impedance shorted stubs are often used to overcome this problem by thermally connecting the centre conductor to the network housing with a minimal impact on its electrical behaviour.

## 2.6 Selective surfaces

Very frequently to best use the available spectrum, fixed communication systems transmit signal on two orthogonal polarisations of the electromagnetic field over a same frequency band, usually two linear or two circular polarisations. A dual polarisation system can be easily realised using two different antennas with overlapped beams and to achieve the high polarisation purity required for dual polarisation operation it is possible to use particular double reflector solutions, having the characteristic of being equivalent to a rotationally symmetric single reflector system and therefore enjoying high polarisation purity. The drawback is clearly in the increased size and mass. An interesting and quite common alternative is the use of reflecting surfaces able to discriminate between polarisations, which unfortunately are known today only for linear polarisations. Still on-board a



Dual-gridded reflector (courtesy of EADS-CASA)

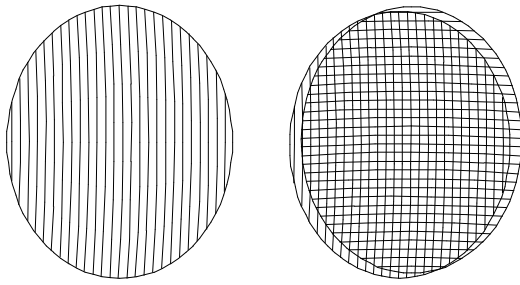


Figure 2.13 - Gridded reflectors. Single grid (left), dual orthogonal grids (right)

satellite the room available is fairly limited and having to accommodate two antennas, one for each polarisation, is undesirable. The use of polarisation selective surfaces helps also in this respect.

Consider the grid generated by the intersection of a parabolic surface with a stack of planes parallel to its axis (figure 2.13, left). A metallic wire grid with this geometry reflects only the component of the field parallel to the generating planes,

while the orthogonal component passes through it almost undisturbed and the reflected field has very high polarisation purity. A second reflecting surface of the same type, but with the grid generated by a stack of planes orthogonal to the preceding ones, placed behind the first surface, will reflect the other field component. By properly arranging the reflectors and feeds is possible to obtain that the two beams, one for each polarisation, overlap. The added benefit is that, if the geometry is properly chosen, the two gridded reflectors will also “filter” the cross-polarisation generated by the feeds reflecting it outside the coverage area.

An antenna based on the above principle (figure 2.14) can be built using two dielectric reflectors with conducting strips etched from a metallic layer deposited on the dielectric. The width of the metal strips causes a small reflection of the undesired field component and a slight loss of power for a field with the same polarisation passing through it to reach the second reflector. However both perturbations are sufficiently small to be acceptable. Another drawback is the increase in the losses due to the lower reflectivity of the surfaces

and to the additional attenuation of the signal transmitted through the front surface and reflected by the rear surface. Often to simplify manufacturing the rear reflector is not gridded and reflects both polarisations, however since the reflected field passes again through the front reflector grid its polarisation purity is not severely impaired.

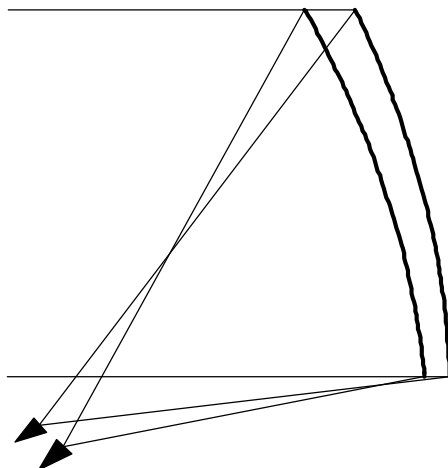


Figure 2.14 - Geometry of an antenna with gridded reflectors

A second type of selective surfaces has a reflection and transmission characteristic that is a function of frequency, just like a filter. These structures, also known as dichroics, consist of one or more layers of arrays of metallic elements, mostly printed on a dielectric substrate (figure 2.15), which reflect the electromagnetic radiation at their resonant frequencies and are transparent elsewhere. Alternatively, it is possible to use holes in a metal screen, which have a dual behaviour, transmitting at the resonant frequencies and reflecting at others.

Frequency selective surfaces are used to reduce the complexity of multiple-beam antennas. A higher number of beams imply a higher number of feeds and a higher complexity of the beam-forming network. Since in communication systems reception and transmission take place in two separate frequency bands, using two separate antennas would reduce the complexity of the beam-forming network by four and reduce the overall complexity by two. Introducing a frequency selective sub-reflector allows the recombination of the two antennas into a single one, with two simpler and smaller beam-forming network and feed arrays, a single main reflector and a, relatively small and possibly flat, sub-reflector (figure 2.16). The group of feeds operating in the frequency band transmitted by the dichroic is located at the main reflector focus and the other, operating in the reflected band, is located at the second focus of the sub-reflector.

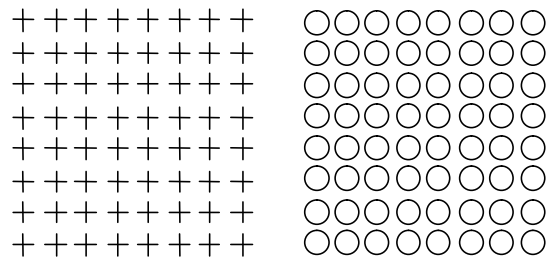


Figure 2.15 - Example of planar dichroic structures: cross (left) and ring (right)

Compared to a single reflector antenna with dual frequency feeds the above solution has also the advantage of limiting the bandwidth of the beam-forming networks and of the radiators permitting further simplifications to the design.

The limitations of this solution are similar to those mentioned for gridded reflectors, i.e. increased losses in both the reflected and transmitted fields. In addition the limited dimensions and curvature, if any, of the sub-reflector adversely affects its frequency discrimination and it is not possible to use dichroics when the two frequency bands are too close.

More complex structures with multiple dichroic sub-reflectors may be used to combine (or separate) the signals in multiple-frequency systems, normally antennas for remote sensing instruments, radio-telescopes and ground-stations.

## 2.7 Frequency reuse in multiple-beam coverages

In communication applications it is often necessary to generate an earth coverage constituted by a large number of beams, which may be of the order of 100 or more, while making an efficient use of the available frequency allocations. Then it is necessary to devise schemes to use the same portion of the frequency spectrum in several beams. The major issue is that the sub-bands allocated for each channel and beam need to be separated by gaps so as to achieve the desired isolation taking into account the finite slope of channel filters, therefore the higher the number of sub-

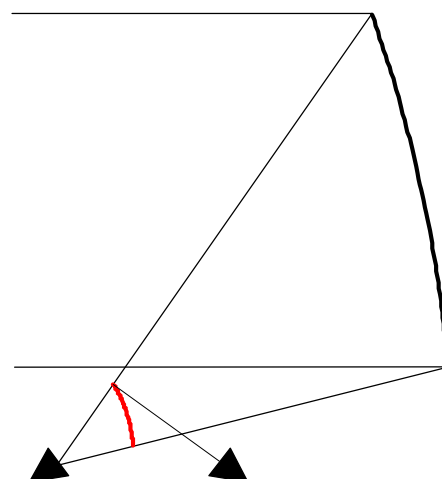


Figure 2.16 - Geometry of an antenna with dichroic sub-reflector

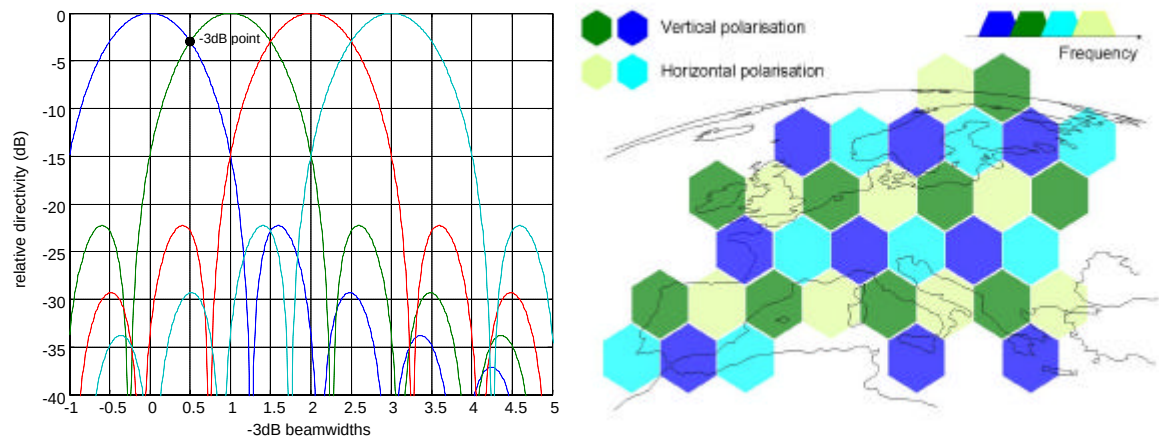


Figure 2.17 – Multiple beam coverage. Left: beam positions in a reflector antenna with -10dB aperture taper and for -3dB beam crossover. Right: example of multi-beam coverage using a four-band schema for frequency and polarisation reuse.

bands the higher the percentage of the overall frequency band that is not actively used. The consequent reduction of the effective bandwidth is clearly undesirable since it has a direct impact on the overall transmission capacity.

The alternative is to seek isolation among the signals routed through each beam at antenna level. This is relatively easy, at effector illuminated by a feed generating a -10dB level at the edge with respect to the centre, the first sidelobe will be at about -22dB with respect to the peak and the second at about -30dB (figure 2.17, left). Selecting the beam width determined by the -3dB contour of the beam, the peak of the first sidelobe is at about 1.6 beamwidths from the beam centre and the second at about 2.5. Therefore in a grid of adjacent beams, with 3dB crossover, the second next beam will have about 22dB isolation and the third-next about 30db. These decoupling levels may not be sufficient in

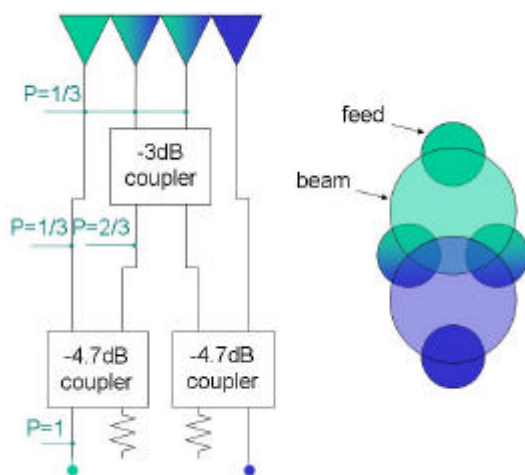


Figure 2.18 – Feed sharing among beams. Left: example of feeding scheme. Right: feed and beam

some case, but constitute a sufficiently good starting point for the use of the same frequency band in several beams whenever their number is reasonably high. When an even wider reuse of frequency sub-bands is needed polarisation discrimination can be used to improve isolation (figure 17, right). The limit for frequency reuse is defined by the four colours theorem of topology, stating that any partitioning of the plane can be coloured with four colours never having the same colour in adjacent cells, but not less.

The spacing among adjacent beams is especially important for frequency reuse schemes as it determines the isolation among them. In order to achieve the optimum, while illuminating the reflector rim at low level so as to have low

sidelobes, it would often be needed to use feeds with a diameter larger than the required spacing, which is quite impossible. Each beam is then generated using group of feeds and one feed contributes to more than one beam, using arrangements similar to the one shown in figure 2.18. Alternatively it is possible to use double reflector antennas in which the sub-reflector is seen under a wider angle from the feeds making them smaller.

## 2.8 Passive, semi-active and active antennas

Following the trend toward higher number of beams and increased flexibility in the channel-to-beam allocation and coverage layout, the beam-forming network of antennas has become more and more complex. The first noticeable effect of the increased complexity is an increase of losses; then it makes sense to modify the architecture of the RF high-power section of the transponder to reduce these losses.

Traditional communication payloads are equipped with high-power amplifier followed by filters and then the signal is routed to the antenna by cables or waveguides. The antenna is entirely passive (figure 2.19 - left). If the antenna operates in both transmit and receive bands, the filters are replaced by diplexers, i.e. a couple of filters sharing one port, so that the signals in one frequency band are routed between the common port and one of the other two, while those in the other frequency band will be routed to the opposite one.

An increased level of flexibility and a reduction of losses can be obtained if a part of the beam-forming network is placed before the amplifiers. In particular all variable components, which are usually the most lossy, can be placed before the amplifiers reducing their impact on the overall power budget. In most cases the network up and including amplifiers and filters is placed on the spacecraft and cables or waveguides route the signals to the antenna. If the antenna is used for both transmit and receive functions since the architectures of the receive and transmit paths do not need to be the same, as the power levels are much smaller in the first. As a consequence the two chains are

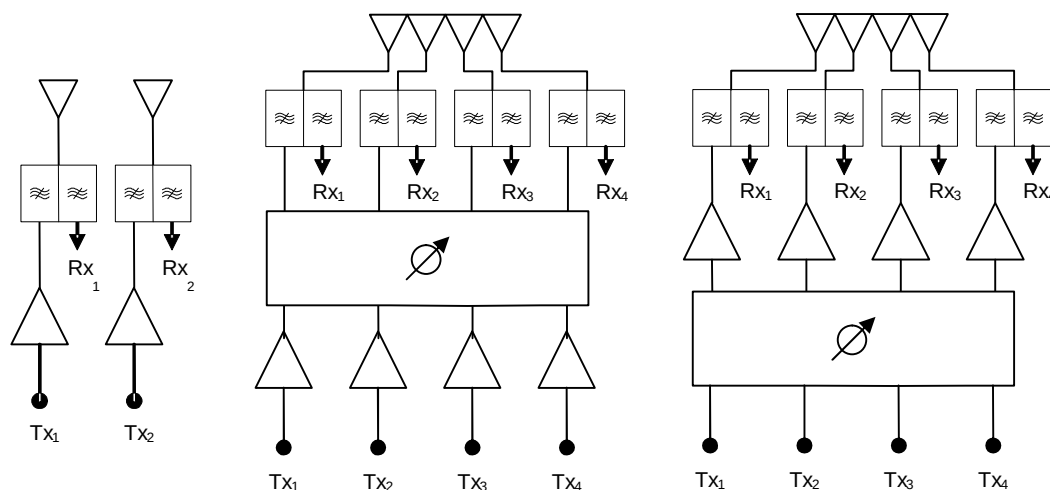


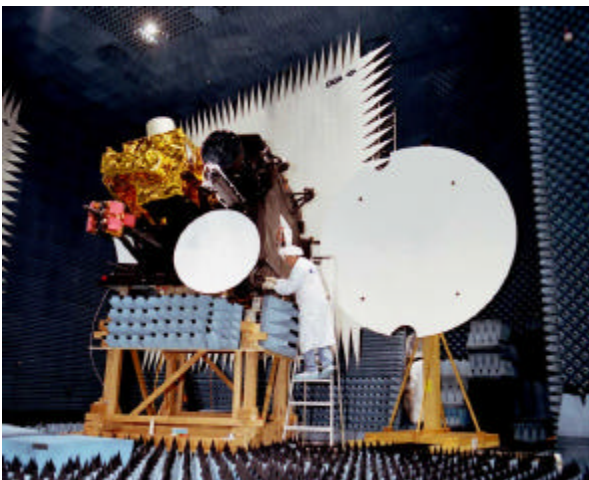
Figure 2.19 – Examples of antenna front-end architectures.  
Left: passive; centre: semi-active; right: active.

separated by diplexers located just after the feeds (figure 2.19 – centre). This type of antenna is usually called semi-active.

Finally when the flexibility required is high, as it is the case for current Ka band payloads for high data-rate point-to-point communications, it may be convenient to realise the full beam forming function at low level by placing the amplifiers, one per feed, as last element of the transmit chain, or in the case of combined transmit-receive operation just before the diplexers. In particular this becomes necessary to individually control the pointing of the beams generated by the antenna (figure 2.19 – right). This latter architecture, usually called active antenna, can also be combined with digital beam forming processors, that implement the low-level signal combination in digital form and produce the RF signal to be amplified by a digital-to-analog conversion possibly followed by a frequency conversion.

## CHAPTER 3

# ***Mobile communication, multimedia and broadband systems***



The ESA Artemis satellite that embarks an L-band mobile communication payload

Mobile communication systems are based on essentially the same operating principle as fixed systems, but at least one of the two terminals is mobile and posing some additional and quite specific problems. The system also provides one or more gateways toward fixed terrestrial systems, for example telephone networks to allow users of the two networks to communicate among them.

Multimedia and broadband systems, which are devoted to the communication of large amount of data for many different uses, have very similar architectures, although they operate on different frequency band and with much higher data rates.

The part of the system connecting the satellite to the gateways (uplink) has the same characteristics of a fixed system (figure 3.1) and uses relatively simple single-beam antennas with a simple coverage shape (e.g. elliptical). The other half of the system (downlink) handles communication with the mobile units and uses separate frequency bands and antennas, which have rather different characteristics. Circular polarisation is generally used so as to improve the link with the mobile units, which may have any orientation with respect to the satellite. It is worth noting, in passing, that the use of circular polarisation was introduced in the first satellite links and despite the additional complexity for antennas, because a circular polarisation wave is insensitive to the polarisation rotation induced by the ionosphere (Faraday rotation), as well as to an axial rotation of the receiving antenna, making the link much easier to establish.

Initially, the main users of mobile services were ships and aircraft and it was possible to use relatively large terminals. The early maritime terminals -generally installed on cargo, cruise and ocean fishing vessels- used reflector antennas with diameters around one meter. The possibility of fitting terminals on offshore fishing boats and private yachts, significantly increasing the market, stimulated to the reduction of size

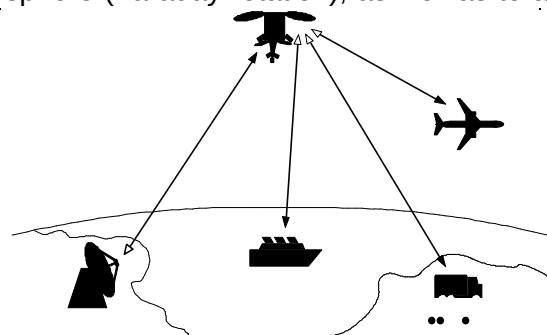


Figure 3.1 – Architecture of mobile satellite communication links

leading to the developed of terminals suitable for medium and small vehicles (e.g. trucks) and of portable terminals similar to cellular phones, but larger.

Two characteristics of the transmission channel of these systems determine the main features of their downlink antennas. First, the transmission channel is almost totally uncontrollable. For a fixed system it is possible, given the geographical and meteorological characteristics of the terminal locations, to select the best position and design to ensure the desired percentage of availability of the link. For mobile units, which can be anywhere this is obviously impossible, apart from some very global geographical consideration, e.g. the range of elevations of the satellite over the horizon. Therefore the design must be based on an estimate of the average and worst-case conditions and on the availability of sufficient margins. At the same time, the desire to use low-power terminals with small, low cost antennas easily installed on vehicles or hand held units, implies that the satellite has to contribute for a large part to the overall system performances. High satellite antenna gains are therefore essential, leading to the use of multiple-beam antennas with large aperture diameters.

Since the traffic generated by mobile users can not be concentrated, e.g. by multiplexing, so as to optimise the exploitation of the available frequency band, it is necessary to re-use the spectrum extensively by means of spatial and polarisation diversity or using time multiplexing or orthogonal modulation codes. The use of both polarisations combined with the use of circular polarisation creates an additional requirement. First order scattering in the environment of the mobile unit inverts the sense of the circular polarisation creating the conditions for additional interference among beams using the same frequencies in different polarisations. In most cases, the necessary reuse of the spectrum is obtained combining several types of diversity among space, frequency, time, polarisation, modulation space, to properly balance the system requirements.

A further problem, due to historic choices rather than to intrinsic system characteristics, arises from the use of relatively low frequencies for the link with the mobile, i.e. L band (1.5-1.6GHz) and S band (around 2GHz). This makes the mobile terminal antennas simpler and more economical, but entails the use of very large antennas on the satellite. At present, the state of congestion of the portion of spectrum allocated to these service in the L band stimulates the investigation of systems operating in frequencies up to twenty times higher, which are still free, especially to offer services dedicated to data transmission among personal computers.

Multimedia and broadband systems share many of these problems as they are intended for a combination of fixed and mobile user terminals (e.g. portable digital satellite TV receivers and laptop computers). They operate at much higher frequencies, i.e. in Ku band (11.7-12.7 and 14-17.8GHz) and Ka band (18-31GHz), respectively. As the data rate of these systems is much higher, antennas with significant higher gain are needed, particularly at the higher frequencies at which the attenuation due to rain and other meteors is significant. However, the shorter wavelength helps to maintain the antenna sizes within manageable limits. Still the problem of surface stability and accuracy remains, since the degradation of the antenna performances with distortions and surface error depend upon the ration between their amplitude and the wavelength.

### **3.1 Satellite constellations**

Before the turn of the century, the explosive growth of the market for mobile telephony led to the development of systems based on constellations of satellites placed in non-geostationary orbits, which move with respect to the Earth's surface. The satellite fleet is distributed along a number of orbits, on different orbital planes around the Earth so as to be able to provide communication services all over the world. Many systems of this kind were studied and two are operational, Iridium [1] and Globalstar [2], but neither one has been as successful as initially expected.

Since at least the user terminals would be moving anyhow, the satellite movement does not add significant complexity and there are two ways of taking care of it. On one side, it is possible to use fixed beams, which move with respect to the Earth surface forming a system of translating cells requiring the transfer of circuits from one satellite to the next in a virtually continuous way as users enter and leave their respective coverages. On the other side, it is possible to use moveable beam, so as to keep the cells fixed associating each of them to a visible satellite and transferring all the circuits together to the next satellite when the first is about to lose visibility. Other intermediate solutions are possible, e.g. to dynamically associate group of users to a satellite and a beam using the information about the signal they receive best.

A further difficulty would arise from the elliptical nature of some of the orbits that were chosen for some of the systems. A change in the satellite altitude results in a change of both the size and the power flux density of the coverage. For some types of orbit, the two effects cancel each other out. In others, it may be necessary to use antennas capable of dynamically adjusting the angular diameter of the beams (zooming antennas) to compensate, at least in part, these variations.

The satellite constellations studied for mobile communications used low or elliptical orbits with 10-60 satellites over a number of orbital planes. Several solutions were proposed for multimedia and broadband applications, ranging from a few satellites on intermediate orbits (at about 20.000Km) to up to about 300 satellites in low orbit (400-900Km), and including mixed system formed by a few geostationary satellites and a number (50-100) of low orbiting ones. Today geostationary solutions serving specific geographical regions are prevailing.

Array antennas were generally considered for satellite constellations, especially those in low orbit that require lower antenna gain, owing to their greater flexibility and to the possibility of producing them in large numbers at low cost, a particularly important aspect given the relatively large number of identical satellites that form each constellation.

### **3.2 Folding reflectors**

For medium-orbit and geostationary satellites the use of small ground terminals operating at L and S band can only be accommodated resorting to large antennas on the satellite. The request of aperture diameters from 4 to 10 metres is met using unfurlable and unfoldable reflectors, i.e. having a flexible or segmented structure. Various factors have to be considered in the development of these solutions: the limited space available in the launch vehicles, the intense vibrations caused by the engines and the air flow during the first part of the launch, the lifting cost increase with the mass of the satellite due to higher fuel consumption.

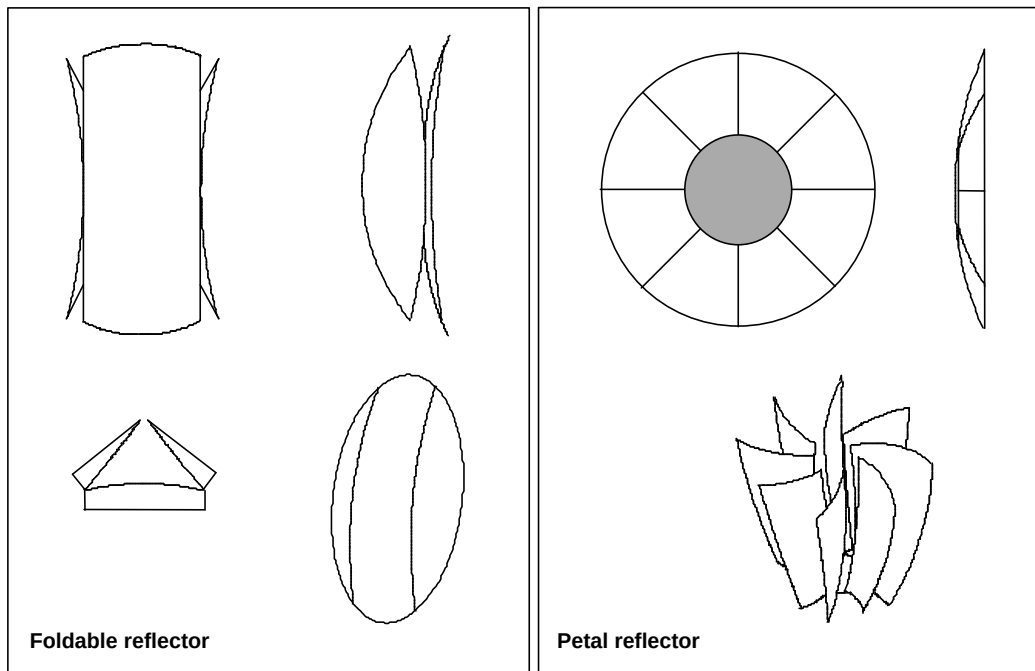


Figure 3.2 – Examples of reflectors with opening panels

Several solutions have been investigated over the past 20 years. A factor that significantly limits the number of possibilities is the fact that the folded structure (i.e. the reflector in the launch configuration) can only be attached to the satellite body at a few points that need to be rigidly connected to the “fixed” part of the reflector. The moving parts need instead to be supported and usually can not carry mechanical loads either to avoid damage to the joints or because they are too flexible, making it quite challenging to guarantee their integrity in the launch phase. The solutions developed up to now fall into four categories: panel reflectors, mesh reflectors, inflatable reflectors and flexible reflectors.

Leaving aside the differences in the segmentation of the surface and the opening mechanism and cinematic, panel reflectors (figure 3.2) are made by splitting the reflecting surface into segments of relatively small size connected by means of a system of hinges and possibly rods. In the folded configuration the panels occupy a limited space and form a concentrated mass with good dynamic behaviour. They are opened when the satellite having reached the nominal orbit conditions, experiences only the small vibrations produced by the attitude and orbit control thrusters, which are fired at more or less regular intervals to correct the satellite position and orientation. One of the major unknowns in the design of these structures is their deployment behaviour, which can not be easily studied experimentally and tested on earth in conditions similar to the operating ones, i.e. vacuum and micro-gravity. The only possibility to better understand and verify the deployment dynamics is the use of the expensive, difficult to repeat and very short lived, free fall (zero-G) condition obtained with special aircrafts during parabolic flight [3]. As a best-resort alternative, mechanical off-loading, using hoists and weights, is used where possible, but it adds several unknowns that are hard to quantify and control.

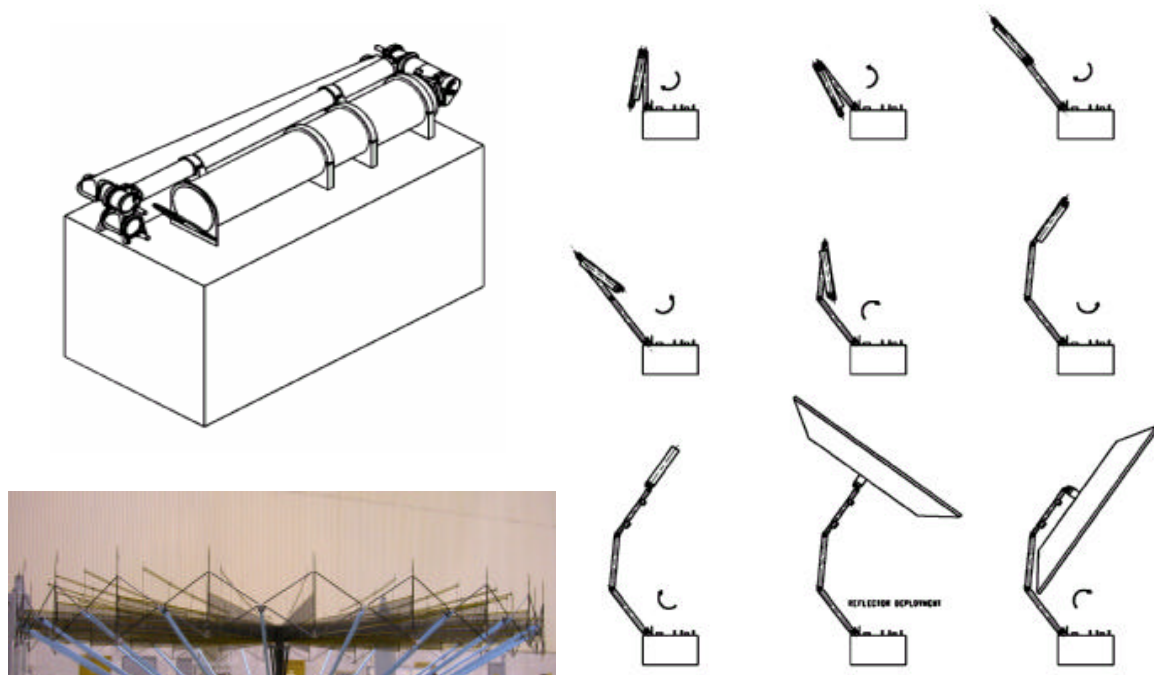


Figure 3.3 - Mesh reflector (12m). Upper left: stowed configuration. Lower left: deployed structure without reflective mesh. Right: deployment sequence. (Courtesy of Alcatel Alenia Space)

Mesh reflectors (figure 3.3) are based on the idea of a reflecting surface free of bending resistance, such as a net or a tricot, placed under tension and kept in the desired shape by a system of brackets and rods resembling an umbrella or other similar arrangements. The mechanical problems faced by these structures are similar to those presented by panel antennas, but since they are lighter it is possible to reduce the total number of joints and hinges by using the elasticity of the rods and the characteristics of the mesh. On the other hand, it is not easy to ensure that the mesh can not become entangled during the deployment and either be damaged or jam the mechanisms. Nevertheless this currently is the most commonly adopted technology and reflectors with diameters up to at least 12x16m have been flown [4].

Inflatable reflectors are based on the same principle of pneumatic boats (figure 3.4). They may consist of one or more chambers, for instance a main one with a toroidal shape, which is pressurised to achieve deployment and supports a pair of membranes that, once inflated, make the whole assembly resemble a lens. One membrane is cut to give it a parabolic shape and its inner side of one membrane is coated with metal; the other one is transparent to microwaves. The whole structure is made of layers of synthetic fibre fabric impregnated with a polymer resin having the characteristics of becoming rigid when exposed to UV radiation. Once in orbit, the reflector assembly is deployed and then inflated with low-pressure inert gas (e.g. nitrogen). Next it is oriented so as to maximise the incidence of sunlight on it and, after sufficient exposure, the gas is evacuated. Finally the reflector is placed in its operating position. This complicated pre-operational cycle and the uncertain quantification of the resistance to the adverse environment of space makes this technology less attractive than the other ones.

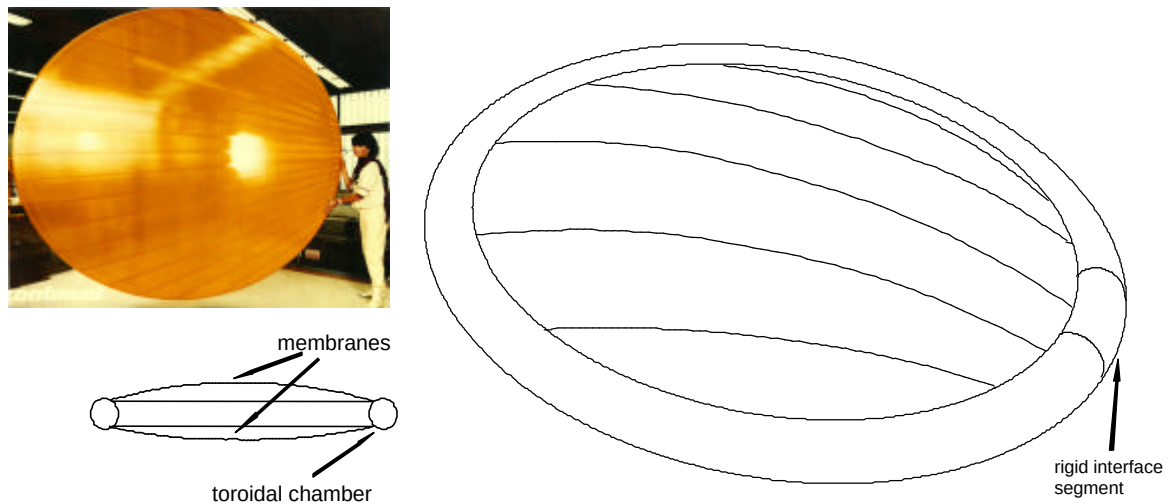


Figure 3.4 – Inflatable reflector: structure and prototype (courtesy of Oerlikon-Contraves)

A further alternative for reflector of sizes up to 4-5m is to use a highly flexible surface made by a thin membrane, for example a skin of Kevlar fibre fabric, woven along tree axes placed at  $60^\circ$  degrees to each other (like the sails of 19<sup>th</sup> century clippers) and impregnated with epoxy resins. During launch the membrane is kept folded by wires, connecting rods or similar arrangements and, once the satellite is in orbit, it is released and tensioned by a set of flexile ribs (that may be an integral part of the membrane structure). The ribs also ensure the mechanical stability of the parabolic shape.

All these types of reflectors are affected, to a varying degree, by significant deviations of the reflecting surface from an ideal parabolic shape. The main characteristic of these perturbations is their well-defined geometry. In the first type of structures they arise from the segmentation, in the second are linked to the connections of the mesh to the support and tensioning structures (pillowing), to the fabric cutting in the third and to the ribs in the fourth. Given the regular -essentially periodic- distribution of these perturbations, their impact on the sidelobe level is significant as it may be easily understood considering the Fourier transform relationship existing between the field distribution on the antenna aperture and its radiation pattern.

### 3.3 Passive intermodulation products

Mobile communications antennas are subject the undesirable occurrence of passive intermodulation products, i.e. spurious harmonic signals generated by the non-linear behaviour of objects exposed to the electromagnetic field or subject to an RF current flow. The relatively high transmitted power, the large number of carriers and the relative position of the transmit and receive bands are the factors concurring to the appearance of this phenomenon, which is very dangerous in these system with very low received power.

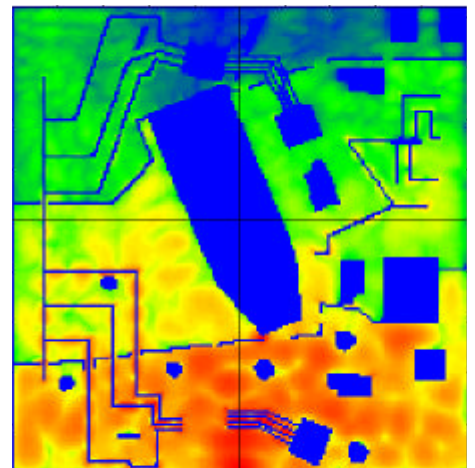
The harmonics can originate from various sources: the non-linear voltage-current characteristic of metal junctions and metal-oxide layers, the hysteresis of ferromagnetic materials, the electrostriction of dielectrics or the non-linearity of resistive materials at high current densities. Also other factors concur to the phenomenon like the presence of impurities or the roughness at microscopic level in metal junctions. In other words, almost

any material or combination may be the cause of passive intermodulation, provided that the power is sufficiently high. For instance, rust is a well-known source of intermodulation products on military ships equipped with high-power radars, where they are known as the “rusty bolt problem”.

The phenomenon requires the presence of at least two signals at different but close frequencies so that intermodulation products are generated by the non-linearity, like it occurs in semiconductor components, for example diodes. Clearly the power of the incoming signals must be high enough to produce on the non-linearity a signal with sufficient amplitude to create detectable harmonics. The conditions exist in mobile communication, multimedia and broadband satellites. In the first ones, the situation is made even worst by the fact that low-order (third to ninth) products, which have the higher power, may fall in the receive band and, unfortunately, even the most careful selection of transmit carriers can not prevent seventh and ninth order products from falling in the receive band.

The importance of the problem may be fully understood considering that, owing to the great distance of the satellites from ground, the free space loss combined with the limited power and low gain of the mobile terminal creates a very large difference between the transmit and receive power levels on the satellite. For the proper operation of the system it is often necessary to transmit signals with a total power of several hundreds of Watts, i.e. 170 to 200dB higher than signals in the receive band. Therefore passive intermodulation interferences have to be suppressed to -140/-160dBm to have a marginal impact on the system and even a minor non-linearity in the areas exposed to the transmit antenna fields becomes dangerous.

The most obvious solution to the problem is to spatially separate the transmitting and receiving sections, by using two different antennas, possibly located on the opposite sides of the satellite body. However, such precaution does not completely guarantee immunity to passive intermodulation products. Reducing the level of a signal by 200dB means reducing its power by a factor of  $10^{20}$  and the simple propagation between the feeds of the transmitting and receiving antennas through a path consisting of a combination of reflections and one or two diffractions is a potential threat. An isolation of more 90dB between two antennas operating at 1-2GHz is quite difficult to achieve and typical values are of the order of 70-80dB. Even signals in the milliwatts range, i.e. four orders of magnitude weaker than the transmitted signals, interacting with generators having a “high” conversion factor may easily create harmonic products with a power of 11 to 13 orders of magnitude lower. The only reliable solution to the problem is to avoid the generation of intermodulation products, or at least to make sure they never exceed levels below the thermal noise floor of the system. To fulfil this requirement it is necessary to remove all potential sources starting from a detailed study of the most critical elements, i.e. the transmitting chain up to the feeds, the antenna as a whole and the areas surrounding it. As metallic junctions abound on the exterior of the satellite, e.g. in the thermal protection layers (thermal blankets) that are made of



Simulation of the RF power distribution at L-band on the Earth-facing panel of Italsat-F2 to evaluate the risk of generating Passive Intermodulation (courtesy of IDS).

metallised plastic sheets connected by metallic fasteners among them and to the satellite body to ensure their grounding, removing the causes of passive intermodulation is not an easy task.

Since the harmonic conversion depends on the signal power, threshold phenomena are often encountered and appear to be related to quantum-physics mechanisms (for example, electron tunnelling) or to local thermodynamic phenomena (the dynamic formation of adiabatic hot spots due to the rapid and localised decrease of the contact resistance associated to a self-sustaining local increase of the current flow). Moreover, experience has shown that these phenomena tend to be intermittent, appearing for example only in particular conditions of temperature or temperature distribution (thermal gradients) often occurring only during thermal transients. This behaviour makes the experimental verification of the "cleanliness" of the satellite very difficult. In the first place it is necessary to measure signals at a level very close to the thermal noise of the instrumentation (-150/-160dBm) and with intermittent characteristics. Second, small non-linear objects in the anechoic chamber, like a miniature light bulb or a rusty screw behind an absorbing panel, can be a problem for radiated test. Third, it normally required to reproduce as accurately as possible the characteristics of the space environment encountered by the various antenna (or satellite) components, for example temperatures variations between -120°C and +100°C with a rate of 4-5°C per minute. In conclusion, any object present in the immediate vicinity of the antennas and capable of generating intermodulation constitutes a risk for the system.

The only currently available method to control this risk are the use of robust protective measures in design and manufacturing, including dedicated electromagnetic analysis with suitable modelling tools, execution of extensive and expensive test campaigns from the initial design phases to establish the levels of criticality of the various materials and components exposed to the higher power densities, and the final tests of the whole assembled satellite to ensure that there are no problems and, possibly, correct the ones found.

### **3.4 High-efficiency feeds**

Another characteristic of reflector antennas for mobile, multimedia and broadband systems and, more in general, of all multiple-beam antennas is the use of small feeds with high aperture efficiency that reduce to the minimum the power not radiated toward the reflector (spill-over loss). Maximum efficiency is required both to minimise the transmit power and to minimise the noise figure of the receiver. A specific quality factor is used to measure the latter performance parameter, the ration between antenna gain and antenna noise temperature, usually called G/T. For all these systems, the feeds must also have highly pure polarisation, circular or linear, often in both polarisations at the same time. The specific problem for mobile application is that, owing to the relatively long wavelength, it is not feasible to use conventional waveguide horn, which would be too large and heavy. While for multimedia and broadband the major issue is sometimes to obtain feeds able to operate at both transmit and receive frequencies. The following discussion concentrates on the solutions applied for mobile applications.

A circular horn fed with the fundamental mode ( $TE_{11}$ ) and a percentage of  $TM_{11}$  determined by its diameter combines a high polarisation purity and good efficiency. A way to achieve the desired mode combination at the horn aperture is to use a step in its

profile, which is sized so as to excite the right amount of  $TM_{11}$ . The length of portion of the horn following the step is used to achieve the right phasing of the two modes at its aperture. Sometimes, to achieve a broader operating bandwidth, the inner waveguide protrudes into the large one to create a short coaxial line closed on short-circuit or irises are added after the step. The major drawback of these radiators, usually called Potter horns from the name of the inventor, is their size that makes them too bulky for low-frequency applications.

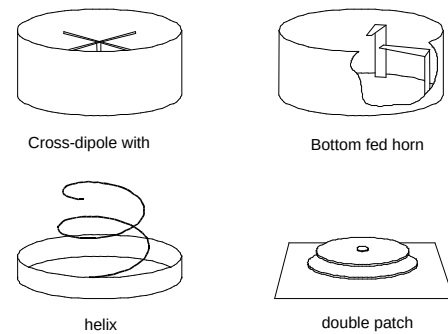


Figure 3.5 – Examples of high-efficiency feeds

There are several types of high-efficiency feeds (some examples are presented in figure 3.5) and the following considerations apply to them all. The common objective is to reduce the radiating structure to the minimum acceptable dimensions for good operations, using the smallest possible volume to convert the field (or current) present at the input port into an aperture distribution as uniform as possible and as purely polarised as possible. Furthermore, to minimise the risk of passive intermodulation, the feeds need to have the minimum number of contacting metal parts, welded joints and must not contain any unsuitable materials, e.g. ferromagnetic ones. These objectives are achievable, but entail a significant amount of work to find the solution.

The best structure can be searched in various directions. A first class of solutions is constituted by circular waveguide excited with a very high modal purity, typically using a symmetric arrangement of feeding probes, so as to reduce to a minimum the length of the portion of guide necessary to ensure the required polarisation purity and bandwidth. The two elements in the upper part of figure 3.5 belong to this class. A possible alternative is the use of cavity radiators, consisting of a cylindrical cavity resonating in the fundamental mode and partially open at one end, e.g. having an annular slot on the radiating face. Also in this case good performances are ensured by the modal purity.

A third class of structures is based travelling-wave radiators, including helices, spirals and conical spirals. These antennas are shaped to adapt the characteristic impedance of the travelling wave to that of free space, gradually matching it to a free-space wave. For example, a helix generates a circular polarisation wave by means of the constructive interaction between the currents flowing in the conductor and the wave propagated in the surrounding space and a suitable choice of its dimensions and, possibly, the shaping of its tip ensure a wide operating bandwidth.

A fourth possibility is to use almost planar structures, like the "patch radiator" illustrated in figure 3.5. Their operation is essentially based on the excitation of currents having a suitable distribution over one or more radiating elements placed above a ground plane. These structures are theoretically the smallest and are also the easiest to build, using low-permittivity dielectric substrates and metallic



L-band feed array (Courtesy of Saab Ericsson Space)

parts etched from a copper film deposited on a Kapton foil. They have the disadvantage of being rather difficult to excite in a satisfactory way and, since the operating bandwidth of an antenna is roughly related to its volume, they have a relatively limited operating band. Finally they have some limitations for high-power operations. To avoid that charged particles impinging on dielectric and electrically isolated metallic parts create over time large potential differences among different objects that would lead to spark discharge (known as ESD, electrostatic discharge), it is necessary to ground all the elements of a patch antenna, but soldering cannot be used to avoid passive intermodulation. At the same time the very small thickness of the printed feeding lines (a few tens of microns usually) makes them not suited for high-density currents. Nevertheless engineers ingenuity and the availability of powerful modelling tools enabling a better understanding of their behaviour has allowed the development of very efficient radiating elements of this class, e.g. by realising them using a completely metallic construction with no dielectric substrates.

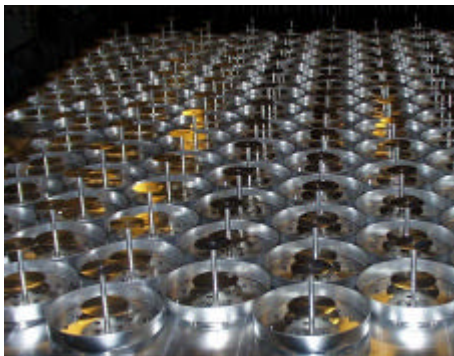
### **3.5 Array antennas and magnified array antennas**

Reflector antennas are not the only way to satisfy the requirements of mobile communications, even for satellites in high orbit: use can be made of large arrays. One reason for such choice is the high radiation efficiency of arrays, typically in excess of 80% and to be compared to the 50-60% usually achieved by reflector antennas. Another reason is that whenever a very high number of beams are necessary to guarantee sufficient gain over an extended geographical area the feed array of a reflector antenna may become so large that an array alone becomes attractive. For multimedia and broadband applications, mostly because of the higher frequency and of the much higher gain requirements, the use of arrays is, at least today, not practical. Still the discussion that follows about magnified array antennas is, at least in principle, applicable.

There are also other advantages in the use of arrays. In the first place, they are flat, which makes their manufacturing and accommodation on the satellite easier, especially when they are large. Second, the presence of a relatively large number of radiators simplifies the designer's task of optimising performances. Third, arrays tend to have better behaviour than reflector antennas when the beam is pointed off-axis, i.e. they have lower scan losses, provided that the radiation pattern of their elements is chosen appropriately. Moreover, array antennas are compact, relatively lightweight and have a large number of identical components, making them particularly suited for small series

production. Finally, the use of "patch" radiators makes it possible to build them with techniques very similar to that used for printed circuits, making the manufacturing of small series much more economical than their metallic counterparts.

Arrays also have several disadvantages with respect to reflector antennas. The large number of radiators entails the use of complex feed networks, which suffer from high ohmic losses, frequency dependent behaviour and manufacturing errors. As individual beams are generated using the whole aperture, each one requires a separate and complete distribution network, leading to very large network assemblies. Even if all the beams have the same



L-band array for mobile applications  
(courtesy of Saab Ericsson Space)

shape and therefore all elements can be excited with the same amplitude, it is necessary to change the phases from beam to beam to obtain the desired beam pointing. Beam forming matrices, like those described for multiple-beam reflector antennas can be used to generate multiple beams from a single network, but the high number of output ports makes them quite complex anyhow. Finally, the pointing and shape of beams, as well as the level of the sidelobes, are very sensitive to amplitude and phase errors in the feeding of the individual elements and therefore to the thermal deformation and frequency response of the feeding network.

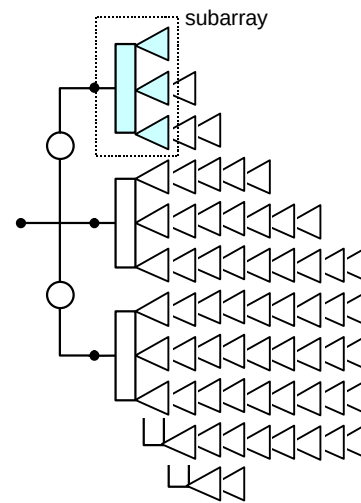


Figure 3.6 - Array with sub-arrays

As already mentioned previously, there are two ways of overcoming these problems, at least in part. The first consists in dividing the array into sub-arrays (figure 3.6), i.e. groups of elements connected to a fixed power distribution network. In practice, this solution is equivalent to an array of large radiators, but offers a considerably better control over the element pattern, which can be shaped using the fixed beam-forming network of the sub-array to help equalising the array beams, i.e. to reduce the scan losses. The overall array excitations can instead be used to shape and point each beam as desired. The main disadvantage is that the radiation pattern of an array of elements larger than the wavelength has grating lobes, which considerably reduce the efficiency and the gain of the antenna (figure 3.7). However, it is sometimes possible to keep this effect within acceptable limits by a correctly spacing the radiators in function of the maximum angular scanning required or by suitably shaping the element pattern or by a pseudo-randomisation of the array layout and sub-array excitation phases, so as to avoid the constructive interference that produces the grating lobes.

The second way to overcome the disadvantages of array antennas is to blend them with reflector antennas using a reflector to magnify the array, i.e. using the concept of “magnifying” antennas introduced earlier. Since magnifying the array also magnifies the individual elements, such an approach suffers like the previous one from the presence of grating-lobes. The main advantage lies in the reduced dimensions of the array. In terms of mass and overall dimensions the advantage is roughly counterbalanced by the presence of the reflector, but not in

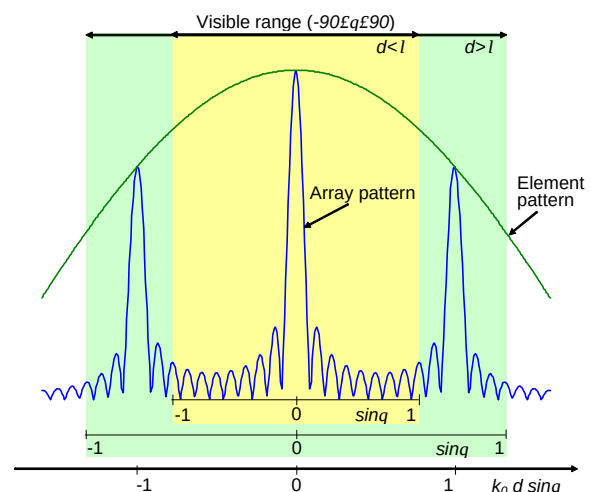


Figure 3.7 – Position of the array grating lobes with respect to the visible range

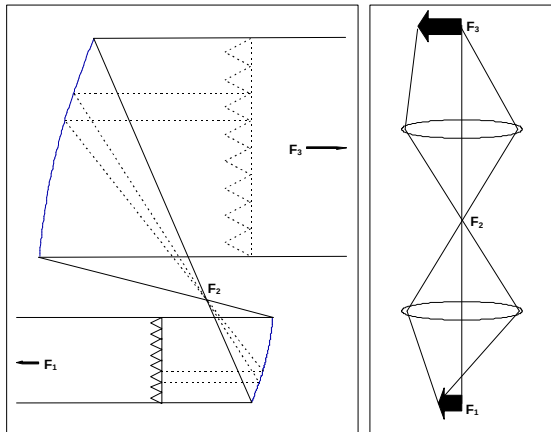


Figure 3.8 - Magnifying antenna with double parabolic reflector

terms of simplicity and reliability: it is much easier to build a reflector than an array of elements with the associated feeding networks.

### 3.6 Magnification and non-focusing antennas

The principle of magnification is entirely analogous to that for lenses in optics, and opens the door to interesting variations on the conventional theory of reflector antennas. Optical magnification is usually intended as a geometrical scaling essentially free of aberrations. However, in microwave antennas the relatively small

size of the reflector with respect to the wavelength causes diffraction to significantly contribute to the radiated field and makes magnification much more approximate.

Therefore the analogy with optical systems is again useful for understanding the antenna behaviour and for its initial design. It is instead quite inadequate for a correct evaluation of performances, as discussed in Section 2.2 about the transformation-chain model.

In optical instruments, like the telescope, two lenses are coupled to create a magnified image of objects. The same technique can be used for antennas, using two confocal paraboloids (figure 3.8). The characteristics of such an antenna are radically different from those of a conventional focusing antenna having the feeds in the focal plane of a parabolic (equivalent) reflector. The essential difference is that in a focusing antenna a movement in the focal plane of the feed produces a change of direction of the beam, while in a magnifying antenna a lateral movement has no effect except for a change in the phase reference of the beam, as it happens translating an array. To obtain angular scanning, it is necessary to create a linear phase variation across the aperture field and therefore across the field radiated by the feed array, for instance varying the phases of its excitations. This implies that the beam generated by the array has a different direction and to maintain a good efficiency the reflectors must be large enough to intercept most of the power contained in the beam under maximum scanning conditions.

As the total effect of the transformation-chain is only the creation of a magnified virtual image of the feed array on the antenna aperture, there are other possibilities than a cascade of two back-to-back Fourier transforms. First, it is possible to use non-parabolic reflector pairs, for instance elliptical or hyperbolic ones, having two foci in real space, to create real or virtual images of the feeds over the equivalent aperture. Extending this idea, it is possible to incorporate the final (aperture-to-coverage) Fourier transform in the transformation chain and use elliptical reflectors to directly focus the feed array image on the coverage area. For satellites in geostationary orbit the difference between such reflectors and parabolic ones is insignificant, while for satellites in low orbit it is sufficiently large to allow some further optimisation.

Another interesting variation is the use of single-reflector magnifying antennas, often called "imaging" antennas, obtained by placing the feeds out of the focal plane. The term "imaging antenna" is somewhat inappropriate in this case as the corresponding transform chain is not

even approximately equivalent to a simple scaling. However, the term "hybrid" has already been used for many years in the literature for all antennas consisting of combinations of arrays and reflector pairs (magnified arrays). It would be more appropriate to refer to them as non-focusing antennas, since their behaviour is intermediate between that of a conventional focusing antenna and a magnifying one. Thus the term imaging can be meant to indicate the intermediate characteristics of these antennas. The non-ideal behaviour of reflectors at microwave frequencies is in this case an advantage as it makes this intermediate behaviour possible. Considering a single beam, true scanning requires both a movement of the feed array and a linear progression in its excitation phase, or equivalently a change of both the amplitude and phase distributions across a larger array. However, also phase only control can be used to modify the beam pointing at the expenses of the overall efficiency. The major drawback of this antenna configuration remains the need of a larger reflector, to avoid excessive spill-over losses, and of a larger array, to accommodate for the "movement" of the primary source distribution.

In practice non-focusing antennas can be considered sub-class of multiple-beam antennas, where the feed array is displaced from the focal plane to improve the efficiency or the beam scanning performance. From a transform-chain point of view they just represent a different way to split the chain between free-space and guided propagation, selecting the partition that produces the best conjugate match between the feed array illumination and the field that would be obtained on the same plane if the antenna was illuminated with a wave having the same angular distribution of the desired beam.

### 3.7 Multiple-layer planar antennas

The use of multiple-layer planar antennas is becoming increasingly widespread for the production of array antennas for space applications. They are often called microstrip antennas because the first ones were realised using printed elements fed by a microstrip, i.e. a planar line printed on a dielectric placed on a ground plane. These antennas consist in a series of dielectric layers with planar metallic shapes and paths placed at the interface between them. Metallic connections among metallic parts on different layers are also often present. The basic reason for the success of this type of antenna in all kinds of applications is the simplicity of its construction. Unfortunately, in space applications, most of this simplicity is sacrificed to achieve the required high performances. The manufacturing advantages are however maintained to a large extent, especially when several antennas have to be built or a single antenna can be split in several equal panels.

The structure of the radiators used for space applications is generally more complicated than that of typical microstrip antennas for ground applications, in particular microstrip feeding is certainly not the only one used. These antennas are therefore preferably referred as "multiple-layer" or "multi-layer" to underline the variety of configurations. Since microstrip antennas are often built on conformal supports, e.g. to fit the shape of aircraft fuselages, the word "planar" is usually added to specify that



Cavity-backed annular-slot planar multiple-layer array (courtesy of Alcatel Alenia Space)

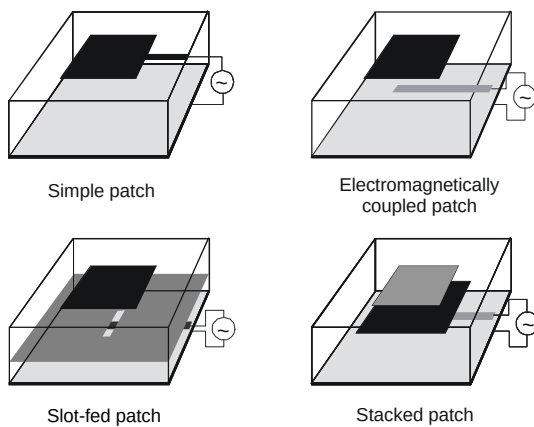


Figure 3.9 - Various types of multiple-layer antennas

a flat dielectric substrate is used. Another name commonly used for these antennas is “patch antennas” or simply “patch”, owing to the fact that the radiating portion of the antenna appears to be the metallic patch on the radiating face, although in fact radiation occurs mostly from the two edges parallel to the feeding line.

Conventional microstrip antennas have several performance limitations: small bandwidth, low radiation efficiency, presence of spurious radiation originating from the feeding lines, low polarisation purity, low power handling, need for high-quality dielectrics.

Various factors limit the bandwidth of multiple-layer antennas. The operating frequency band of a single radiator depends on the input impedance, the directivity, the cross-polarisation level, the ohmic losses and the spurious radiation losses (i.e. power radiated into the substrate, giving rise to leaky waves). The input impedance of the array is further affected by coupling among elements, among elements and feeding lines (the undesired part of it) and among feeding lines. The radiation pattern of the array is also affected by these factors and by the spurious radiation of the feeding lines (beam forming network).

The useful band is found as the intersection of the frequency ranges in which the values of each parameter are compatible with antenna requirements. It is therefore necessary to adjust all factors simultaneously in order to obtain maximum superimposition of the corresponding useful bands and to widen each of them as much as possible. Unfortunately, these factors are not totally independent of each other. For instance in the simpler microstrip antenna designs, i.e. a patch fed by a coplanar stripline, trying to increase the radiation bandwidth by increasing the distance between the patch and the ground plane also increases the spurious radiation of the feeding line. Therefore more complex solutions with more degrees of freedom have been developed to be able to control the different factors independently (figure 3.9).

Considering the bandwidth in terms of input impedance, we may observe that since the patch radiates more efficiently at its first resonance, its characteristic is that of a loaded resonant circuit. Therefore its 3dB bandwidth is approximately equal to  $1/Q$ , where  $Q$  is the quality factor of the equivalent circuit. If the antenna port is matched to the feeding line at the resonant frequency, then outside the quasi-resonance range the input will be mismatched.

The impedance bandwidth,  $B$ , is a function of the acceptable standing wave ratio  $S$ , according to the formula

$$B = \frac{(S - 1)}{Q\sqrt{S}}.$$

Therefore the bandwidth can be enlarged by decreasing the factor  $Q$ , e.g. by using thicker substrates and a lower value of the dielectric constant. However, this entails an increase in

the spurious radiation from the feeding lines, if these are printed on the same substrate. The bandwidth that can be obtained in this way is of the order of 5-8%, and other solutions must be found to achieve the values of 15 to 25%, typically required in satellite communications.

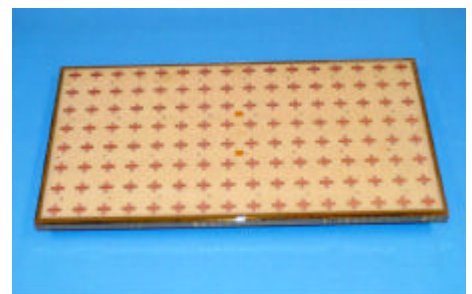
One possibility is to modify the radiator to have multiple resonances at close frequencies, e.g. using shapes other than square or circular, or adding parasitic elements of slightly different sizes. For instance it is possible to cut one or more slots in the patch that resonate at a different frequency from the patch itself or to add a parasitic patch of a slightly different size at a small distance above or below the first.

Another solution is to introduce a second resonance circuit by decoupling the feeding line from the patch, e.g. placing it at a lower height with respect to the ground plane. The gap between the patch and the feeding line forms a capacitance, while the length of feeding line beyond the point on the patch where the impedance is the same of the line introduces an inductance. An additional resonant circuit can also be obtained placing the feeding line on the opposite side of the ground plane and coupling it to the patch through a slot. In this way, input impedance bandwidths in excess of 25% can be achieved.

If we now examine the radiation efficiency, we find that it is closely related to the ohmic losses, in conductors and dielectric, and to losses due to spurious radiation. In general, an increase in the thickness of the dielectric results in a decrease of the losses in the conductor and a decrease in the dielectric constant causes a decrease of the losses in the dielectric. Glues, which most often contain carbon particles, may contribute significantly to ohmic losses. For frequencies up to a few GHz, the ohmic losses are generally fairly low and contribute to a decrease of 2-3% in the radiation efficiency, while, as already noted, an increase in the substrate thickness tends to increase the spurious radiation losses due to the excitation of surface waves. While for space applications 3% power losses in the antenna are significant, it must also be noted that the beam-forming network may have much higher losses and can be by far the dominant element in this respect.

One way to improve radiation efficiency would be to decrease the substrate thickness and the dielectric constant at the same time. However, this would have a negative effect on the radiation bandwidth. This impasse can be overcome adopting structures in which the feeding line is capacitively coupled to the radiator, so that their distance from the ground plane can be optimised separately, reducing ohmic and surface wave losses without compromising the bandwidth. In some cases, to reduce to a minimum the ohmic losses and the excitation of surface waves the multiple-layer antennas used for space applications have no dielectric in the regions of high field density, i.e. close to the feeding lines.

Placing the feeding lines on the opposite side of the ground plane with respect to the radiating part is also an effective way to remove spurious radiation. A second ground plane below the lines can be used to completely remove radiation from them, but it introduces the risk of higher coupling since waves can propagate in the resulting parallel-plate waveguide. Then it is quite common in space applications to separate the lines by vertical walls, creating a strip-line network. In some cases to minimise ohmic losses the lines are realised using “thick” metallic strips or bars and possibly a fully metallic housing, creating in this way a variant of coaxial lines. Clearly such approach complicates manufacturing and increases the



C-band multiple-layer antenna  
(courtesy of Saab Ericsson Space)

mass of the beam-forming network, but it can be necessary to achieve the desired performances.

Even better performances in terms of radiation efficiency and polarisation purity can be obtained by combining the qualities of planar structures with those of cavities. The bandwidth of an open cavity is much larger than that of a planar structure of equal dimensions and coupling a patch with a simple geometry to a cavity produces a radiator with an operating bandwidth of the order of 20% and very good polarisation purity at the cost of an increased structural and manufacturing complexity.

Another way to obtain very good polarisation purity is to use symmetric feeding, i.e. to have two lines one at each side of the patch. When circular polarisation is required four lines can be used fed with a 90° phase shift between orthogonal pairs.

Beyond with the purely electromagnetic problems mentioned so far, the design of multiple-layer antennas for space applications poses a series of technological and mechanical challenges. The adverse space environment imposes several constraints having a negative effect on electromagnetic performance. On one side, there are a number of problems typical of space hardware such as resistance to radiation and dimensional stability under the action of wide thermal variations, which have a major effect on the choice of materials often at the expenses of electrical characteristics. Also the high vibration levels experienced at launch create difficulties in multiple-layer antenna design. On the other side, there are problems associated with particular phenomena, which need to be solved case by case. For instance, the electrostatic discharges mentioned before. To avoid the discharge is necessary to ensure the draining of charges with resistive coating of the dielectric and the grounding of all metallic parts; solutions that both are not really desirable in an antenna, but nevertheless necessary. Furthermore, in high-power applications the intrinsic non-linearity of the materials and metal joints generate passive intermodulation products whose level can easily exceed the noise threshold of low-noise receivers creating further constraints in the antenna design and manufacturing.

Despite these complications multiple-layer antenna technology has reached a high degree of maturity and offers a valid alternative to more traditional technologies, at least for applications in the frequency range from a few hundreds MHz to a few GHz.

### ***3.8 Meteorological attenuation and reconfigurability***

The attenuation caused by the absorption of radio frequency signals due to clouds and precipitation (hydrometeors) has considerable repercussions on all systems operating around 20-30GHz, like broadband systems. It is therefore useful to design antennas, which help to decrease these effects.

When geographical areas of continental proportions are examined, it is intuitively evident that regions covered with thick cloud and regions of intense precipitation will normally alternate with regions having clear or almost clear sky conditions. Statistical analysis of the weather has shown that the typical correlation distance in the temperate regions is about 700Km, i.e. the regions affected by heavy hydrometeors have a mean diameter of approximately 700Km and beyond this distance two receiving stations have statistically independent weather conditions.

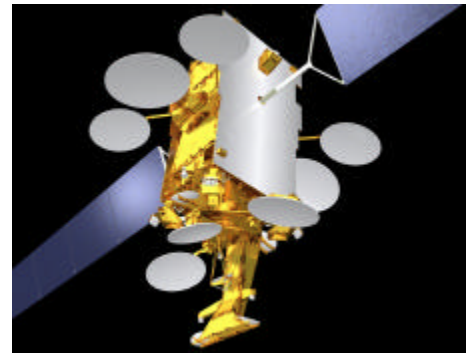
Consider a system that ensures good reception for, say, 95% of the time. Such a system must radiate sufficient power to compensate for the attenuation of a significant portion of

the hydrometeors, which may reach 10dB in the worst cases. In clear sky conditions, the additional power is practically wasted. Given the amount of power used, the question arises if it might be possible to find a more energy-efficient solution, so as to reduce complexity, mass and cost.

The solution exists, and lies in the use of multiple-beam antennas in which the power distribution can be varied between the various beams in such a way to compensate for the temporary localised meteorological attenuation present in some of them. Power is removed from the beams operating in clear sky conditions and transferred to those affected by adverse propagation conditions. The same approach can be applied for a single beam antenna if a feed array or a variable-shape surface is used to vary the local gain within the beam. This solution clearly reduces the total power required to provide the same quality of service.

### 3.9 Large communication satellites

The steady growth of digital communications, in particular Internet-based services, with their huge bandwidth demand has spawned the study of a new generation of satellite platforms. Both multimedia and broadband systems require very high power and very high gain antenna to provide high data rate services to terminals with sufficiently small antennas (like a satellite TV dish or smaller) and medium data rate connections for mobile terminals (e.g. laptops).



A possible configuration for the AlphaBus large platform

Given the operational frequencies of the services (i.e. Ku band for multimedia and Ka for broadband) antennas with gain of the order of 40dBi and 50dBi are feasible. However their beams are quite narrow (i.e. about  $1^\circ$  or less) and therefore a large number of them are required to cover geographical areas of the size of Europe, which are the minimum target area to justify the cost of the system. The first problem encountered with such antennas is that the beam is so narrow that the inevitable oscillations of the satellite, of the order of a few tenths of a degree, need to be compensated to maintain the beam correctly pointed over the service area. This is achieved by equipping the antenna with a pointing mechanism and a closed-loop control system, normally based on a beacon transmitted from ground, which keeps the antenna pointed in the right direction (RF tracking). The scan performance limitations, combined with the need of a pointing system, which by necessity can be accurate only within a relatively small angular region around the beacon direction, force the use of several antennas. Furthermore, the large bandwidth of each link, combined with the need of frequency and polarisation reuse to optimise frequency spectrum usage, requires the use of separate antennas for transmit and receive in both up- and downlink. As a result the number of antennas to be accommodated on a single satellite is large and the RF power needed to achieve the desired power flux density on ground is proportionally elevated.

Large platforms are being developed to satisfy this need. While average size communication satellite platforms can provide about 3kW of DC power to the payload and accommodate 4 to 6 antennas, these platforms are sized to host up to 10-12 antennas, a

total payload mass of about 1000kg and to support a DC power consumption of 12-15kW from payloads.

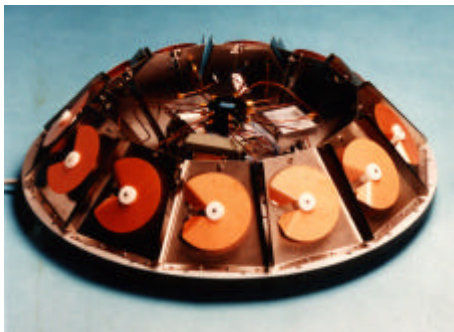
A very challenging aspect of such large antenna systems (antenna farms), beyond the already rather complex design of each antenna, is to make sure that they will operate properly when placed all together in the relatively small space available. First, in the design of each antenna it is necessary to take into account the effects of the electromagnetic interactions with the other ones and with the surrounding structures. Second, also in this case the large difference between transmitted and received power calls for accurate and extensive analysis of the interferences generated by each transmit antenna on all the receive ones, including passive intermodulation effects. Considering the fact that exhaustive testing of all the possible combinations in all possible payload operating modes would be a daunting task with incredibly high costs, also involving significant measurement and safety problems linked to the very high power densities to be handled, computer-based predictions are in fact the only viable solution.

### **3.10 Antennas for mobile terminals**

To complete the discussion of antennas for mobile, multimedia and broadband communication systems, it is necessary to discuss the antennas used for the mobile terminals. In the following the attention is concentrated on antennas to be installed on vehicles, rather than those for portable units, that have a very specific design essentially dominated by their very small size, and those for fixed ones, which are much alike satellite TV ones. These antennas must be suitable for operation at any point in the area in which the service is provided and therefore they must have a rather extended coverage in elevation, in addition to the obvious requirement of 360° coverage in azimuth. A satellite in geostationary orbit is located in the equatorial plane and is seen from ground in a position closer and closer to the horizon as the latitude (north or south) of the observation point increases and as the difference between the observer's longitude and that of the satellite increases. Conversely, satellites in low orbit may be in any position, and are generally considered "active" for mobile units when they are at elevations of more than 5-10° on the horizon. Therefore from all practical points of view the coverage required to mobile terminal antennas is approximately hemispherical. In this coverage it is also important to maximise the polarisation purity so as to reduce the interference due to reflections to a minimum. Since a reflection changes the direction of the electrical field component parallel to the reflecting surface, an odd number of reflections inverts the

direction of rotation of circular polarisation, therefore the unavoidable presence of cross-polarised signals generated by the satellite antenna creates an interfering co-polarised signal with a different time of arrival, due to the longer path travelled with respect to the direct one. For similar reasons, sidelobes must be reduced to a minimum to filter multiple reflections of even order.

From the mechanical viewpoint, in addition to small size and low weight, an antenna for a mobile terminal mounted on a vehicle must also have minimal aerodynamic resistance and must not interfere excessively with its ergonomics and appearance. The problems arising in its



Prototype of an electronically pointed L-band terminal antenna

design are therefore, on the whole, very different from those described previously.

A first possibility is to use small antennas with a hemispherical coverage. If higher gain is needed a larger antenna can be equipped with mechanical pointing systems, covered by a suitably shaped housing (radome). This type of solution is suitable for ships, but not for aircraft or road vehicles (except large trucks). Small antennas for mobile communications with the desired characteristics can be produced by variations on helices, monopole and loop antennas. In recent years, the use of active array antennas, planar or non-planar, with electronically steerable beams has given a considerable impulse to the attempt to maximise the gain while maintaining small sizes and avoiding the expensive and bulky mechanical pointing devices. A planar array, which offers the minimum dimensions in the vertical direction, suffers from a considerable decrease in gain and polarisation purity when the beam points at directions close to the horizon and much effort has been devoted to develop methods to avoid this problem. They are mostly based on the use of structures having a limited vertical extension, for example truncated cones with a very high aperture angle ( $120^\circ$  or more). Clearly electronically steerable antennas are the only solution for continuous communication between vehicles and multimedia or broadband satellites.

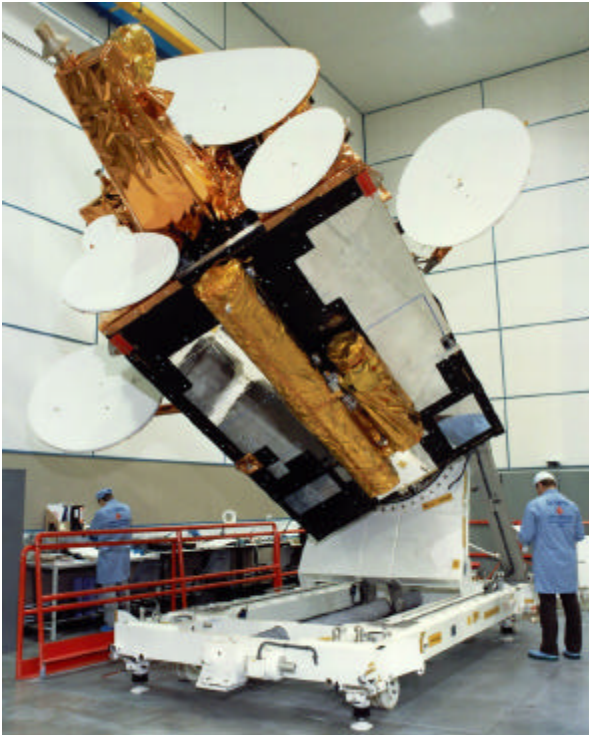
#### References

- [1] <http://www.iridium.com/>
- [2] <http://www.globalstar.com/>
- [3] [http://www.estec.esa.nl/outreach/parabolic/flight\\_technique\\_frame.htm](http://www.estec.esa.nl/outreach/parabolic/flight_technique_frame.htm)
- [4] [http://www.boeing.com/defense-space/space/bss/factsheets/geomobile/thuraya2\\_3/thuraya2\\_3.html](http://www.boeing.com/defense-space/space/bss/factsheets/geomobile/thuraya2_3/thuraya2_3.html)



## CHAPTER 4

# *Direct satellite broadcast systems*



Olympus, an ESA experimental satellite that embarked a TV broadcasting payload

Direct broadcasting systems for television (analog or digital) and digital radio (CD quality) differ from all the preceding systems in that transmission is strictly one-way. A ground station transmits (uplink) the television or radio signal to the satellite that spreads it over the coverage zone in a separate frequency band (downlink). The uplink antennas do not have very special performance requirements. They are typically to single beam antennas similar to the ones used for fixed communication systems, the only difference being the frequency of operation (Ku band in this case). The downlink or broadcasting antennas have instead rather different requirements from the ones discussed in previous chapters (figure 4.1).

For radio frequency systems the maximum power flux allowed outside their coverage zone and frequency band is regulated, while broadcasting operators have an obvious interest to maximise the reception area. Consequently downlink antennas need to be designed to meet regulations as closely as possible, leading to beams that follow the regional geography and country boundaries.

The need to use a single frequency over the coverage, so as to simplify the user terminals, prevents the use of multiple-beam systems limiting the available gain. To minimise the dimensions of the receive antennas, so as to contain their cost and simplify their installation, is then necessary to increase the transmitted power and to ensure that the antenna beam fits as closely as possible the coverage shape.

The above factors lead to two distinctive characteristics of downlink antennas for direct broadcast systems: a high transmission power (around 1-2 kW) and contoured beams, usually obtained with shaped reflectors.

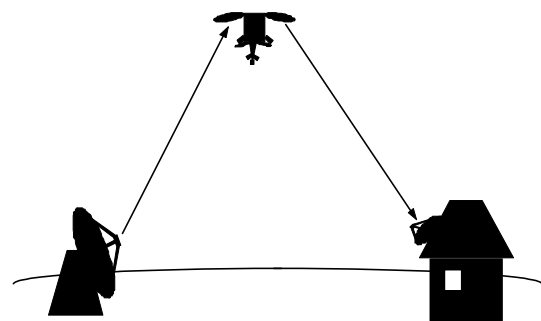


Figure 4.1 – Architecture of a direct broadcast link

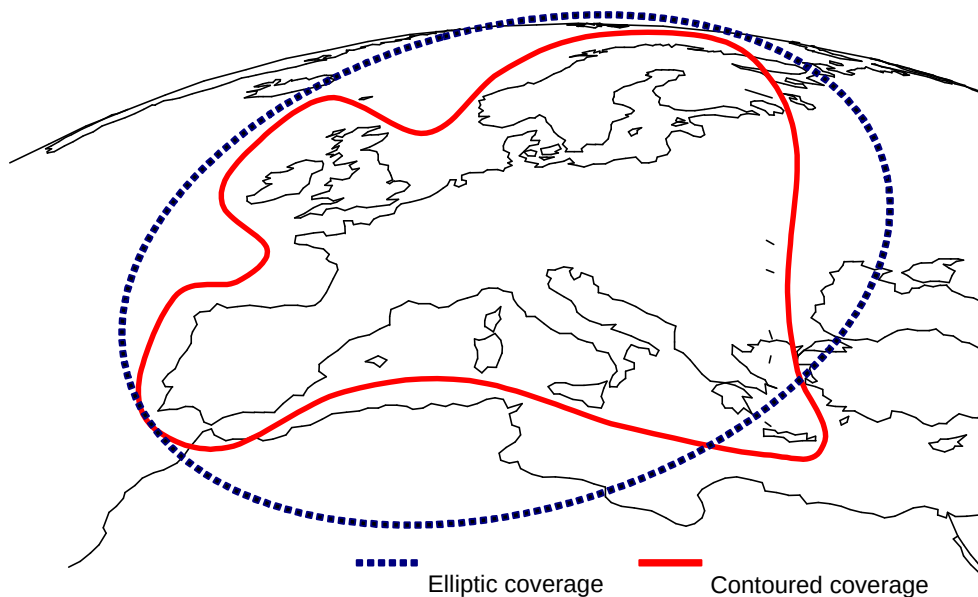


Figure 4.2 - Examples of elliptical and shaped coverage

#### 4.1 Contoured-beam antennas

Radio and television users are easily separated in linguistic groups. Over Europe this is achieved roughly following national or regional (figure 4.2). This objective is attainable with an antenna that instead of producing a circular or elliptic beam produces a beam with the desired shape, for example a reflector antenna fed by an array having the mirrored shape (to compensate for the inversion made by the reflector). The extent to which it is possible to follow the geographical contours is limited by the size of the image of a single feed. Using the analogy with optical systems, it is easy to recognise that the dimension of the (equivalent) aperture determines the diffraction pattern and therefore the maximum resolution, i.e. the minimum angular dimension of the beam obtained with an optimal illumination of the reflector.

The frequency band allocated for satellite broadcasting downlink is around 12GHz therefore the typical ratio between antenna diameter and wavelength is 50 to 100 and the angular resolution is between half and one degree, which from a geostationary orbit corresponds to diameters of 300 to 600 kilometres. It is therefore impossible to closely follow geography, but the efficiency increase obtained is nevertheless significant. Even in the geographically extreme case of countries like Chile, Japan or Norway, whose shape is anything but round, it is still possible to reduce the radiated power by 30-40%. Considering that the average efficiency of high-power RF amplifiers (travelling wave tubes) is around 50% and that the required output power is of the order of 1kW, the DC power requirements are reduced by 600-800W, not a minor improvement.

Another limit to the use of coverages that follow coastlines and country boundaries with minimum deviation arises from satellite motion. A satellite in geostationary orbit appears approximately fixed with respect to ground, but its motion along the orbit is affected by several sources of disturbance, like the solar wind pressure on large surfaces (solar panels), the tidal effects of the Moon and the Sun as well as the non-uniformity of Earth's

gravitational field. In other words, the satellite fluctuates about its nominal position with respect to ground and rotates about its centre of gravity.

To compensate for these changes of position and attitude, the satellite is daily manoeuvred by means of small chemical or electrical (ion) thrusters. Left uncontrolled, it would rapidly start an inverse-precession motion and slowly lose height, until the air drag of the outer atmosphere suddenly slows it down making it fall at high speed and burn in the lower and denser layers of the atmosphere. The manoeuvres to correct the attitude and maintain position inevitably cause further but smaller perturbations. Taking all the combined effects into account, the pointing direction of an antenna rigidly attached to satellite body oscillates within a cone having an aperture of about  $0.2\text{--}0.5^\circ$ .

To continuously cover the service zone and comply with radio frequency regulations, it may thus be necessary to dynamically adjust the antenna pointing so as to compensate for the satellite oscillations. In all cases the antenna beam is usually slightly larger than necessary to allow for some residual misalignment.

Contoured-beam antennas are typically quite inefficient when compared to their multiple-beam siblings. Conservation of energy requires that the gain achieved by an antenna over an assigned solid angle be inversely proportional to its extent. To achieve the necessary resolution, a contoured-beam antenna must be considerably larger than an antenna having a circular beam covering an equal area and its efficiency is correspondingly lower. As a result contoured-beam antennas tend to be rather large and heavy. It is therefore very important that the shaping is made in the best possible way using the smallest possible reflector compatible with the required coverage. The figure of merit for this purpose is the gain-area product. The area of the coverage zone is usually measured in normalised units obtained by approximate the angles with their tangents or sines, thanks to their small angular sizes.

Contoured beams can be obtained in several ways. The Fourier transform-chain model can be used to derive a few solutions. A first way to shape the coverage is to use a multiple-beam antenna transmitting the same signal in all beams, thus producing a single beam that is the combination of the beams generated by the individual feeds and has approximately the same shape of the feed array. Further shaping can be obtained adjusting the amplitude and phase of the signals fed to the individual array elements. In particular these additional free parameters are often used to obtain the best power flux distribution (usually the most uniform).

The practical distinction between multiple-beam and contoured-beam antennas of this kind, which is already rather small, is made even smaller by the fact that a beam-forming network can often generate simultaneously multiple beams, shaped if necessary, as well as a single beam, also shaped, which uses all the radiators available to provide global coverage. As already shown previously, a simple example of such a network has the transfer matrix

$$T = \frac{1}{2} \begin{bmatrix} 1 & j & j & -1 \\ j & 1 & -1 & j \\ j & -1 & 1 & j \\ -1 & j & j & 1 \end{bmatrix},$$

and is able to generate beams using all the radiators

$$\mathbf{I} = -\frac{1}{4}\{j \quad j \quad j \quad j\} \Rightarrow \mathbf{O} = \frac{1}{2}\{1 \quad 1 \quad 1 \quad 1\}$$

as well as only two of them

$$\mathbf{I} = \frac{1}{4}\{1-j \quad 1-j \quad -1-j \quad -1-j\} \Rightarrow \mathbf{O} = \frac{1}{2}\{1 \quad 1 \quad 0 \quad 0\}.$$

A second way to produce a single contoured beam, often preferred for broadcasting downlink antennas is to modify the antenna optics (reflecting surfaces) so as to change the equivalent current distribution on the aperture and thus the beam shape. This is equivalent to the use of a distorted lens or mirror in an optical system, i.e. to the introduction of a special transformation in the Fourier transform-chain, before the last one. The main advantage of this approach lies mainly in its simplicity: the antenna has only one feed and one or two reflectors whose surfaces are deformed with respect to the ideal parabolic profile. Another advantage is that the feed can be easily designed to have very low ohmic losses and produce a pattern with a high degree of polarisation purity. Typically circular corrugated horns are used for this purpose, i.e. circular horns having their wall replaced by a series of adjacent narrow annular cavities with a depth of about a quarter wavelengths (figure 4.3). Each cavity acts as a shorted radial line and presents an open circuit condition at its input; as a result the inner “wall” of the horn acts as a magnetic boundary. This structure supports the propagation of “hybrid” modes, formed by the superposition of TE and TM ones with the same index and propagation constant. The virtual absence of currents flowing on the walls significantly reduces ohmic losses and the almost straight E-field flux lines of the fundamental hybrid mode (HE<sub>11</sub>) produce a radiation pattern with very high polarisation purity. Furthermore the presence of a single species of modes offers a very large operating bandwidth (up to an octave), since it is easy to avoid the generation of high order ones by carefully choosing the longitudinal profile of the horn, and makes it possible to design horns with an aperture diameter up to about 10 wavelengths. The major drawback of corrugated horns is their weight, which is much higher than that of a similar smooth-walled one.

Of course, nothing prevents the combined use of multiple feeds and shaped reflectors, a solution that is indeed adopted in some cases. Having multiple feeds it is possible to generate multiple beams, leading to the interesting problem of shaping the reflector in the best way for the different beams, which usually have different shapes.

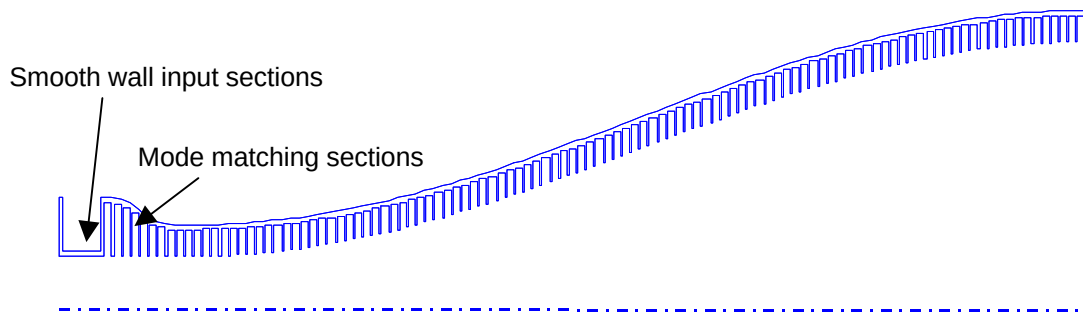


Figure 4.3 – Example of the profile of a corrugated horn

Finally, a third alternative would be the use of an array excited in such a way to generate the desired beam shape. This is possibly the solution with the highest antenna efficiency and would lead to the smallest antennas. However the very large number of elements and the complex, lossy beam-forming network combined with the high transmit power makes such solution impractical.

#### **4.2 Reflector shaping**

Shaped-reflector antennas are very different from all kind of antennas discussed previously. Parabolic and other focusing antennas operate, from an optical point of view, in stable conditions, i.e. small perturbations of their shape do not affect significantly their radiation patterns. Shaped-reflector antennas, instead, operate in the opposite conditions and small changes of the reflector shape may have a strong impact on the antenna beam. The main beam of a focusing antenna is obtained by in-phase combination of the contribution from the whole reflector surface and local departures from the ideal profile have a second order effect, since they modify the contributions as the cosine of the phase error. Instead the beam of a shaped-reflector antenna is formed by the phased superposition of the contributions, with continuously changing phases across the beam, so even minor changes may significantly alter the beam shape. On the other hand, since reflector shaping alters both the amplitude and the phase of the equivalent aperture currents and only the amplitude distribution of the field is relevant for the shape of the beam, there are infinitely many possible shapes that produce the same contoured beam and the problem is to find one that is sufficiently stable. Clearly the control of the surface deformations induced by thermal variations and gradients plays a fundamental role in the design of these antennas. Therefore the optimum (stable) solution is searched using the nominal and the deformed profiles in an optimisation loop.

The number of parameters required to describe a finite portion of any surface is infinite, but two factors contribute to reduce the design freedom. First, the surface must be continuous, if only for manufacturing reasons. Second, the effects of undulations with a period smaller than a wavelength are filtered by the electromagnetic field, because their Fourier transform corresponds to imaginary angular frequencies (in other words, they affect only the configuration of the quasi-static component of the electromagnetic field close to the reflector surface). Owing to these facts, it is possible to reconstruct the equivalent aperture currents within the approximation required to correctly compute the far field from samples taken with half a wavelength over the whole surface. As a consequence the number of free parameters is approximately equal to four times the area of the aperture in square wavelength units. This kind of approximation is strictly valid only if the surface is plane and infinite, but can be safely used for reflectors having a continuous surface characterised by a minimum curvature radius and a diameter that are both large with respect to the wavelength. These conditions are usually met in all practical applications also because they are the same required to apply the physical optics approximation, which is normally used in computer modelling tools for the calculation of the field radiated by large reflectors.

For reflectors with diameters of a few tens to a few hundreds of wavelength the number of free parameters is in any case very high and it is therefore necessary to choose a proper way to define their shape to avoid making the problem unsolvable (i.e. badly conditioned). The question is essentially similar to that of the selection of the best orthogonal basis for the expansion of a given function, i.e. the one approximating the function within a given



Double reflector antenna shaped to produce an elliptical beam (courtesy of Alcatel Alenia Space)

error with the smallest number of terms. In this case, however, the function is not known in advance and therefore the basis must be selected according to considerations of other kind. Since the solutions that can be obtained are those and only those expressed as a linear combination of the basis functions its choice has a significant impact on the final solution. For any selection of the basis many of the possible solutions can not be represented with a finite (small) number of terms, therefore it is impossible to guarantee that the solution found is globally optimal.

Suitable sets of continuous and differentiable functions are typically used as bases to represent the reflector surface

or the aperture field, rather than point values (sampling), since they guarantee surface continuity and they offer the possibility to control the minimum period and the maximum gradient of surface oscillations. Two classes of functions are commonly used having, respectively global and local support. For example, the surface deformations can be expanded in Zernike polynomials, defined as  $Z_{k,m}(r,\phi) = r^k(\cos \phi)^m$ , each one spanning the entire surface, or described by means of splines, i.e. by segmenting the surface in cells and using interpolating polynomials defined on groups of two or three cells in the two directions. Recently it has been proposed to use a third class of functional bases, which include terms with diminishing size of the support, so that there are a global terms as well as ones with higher and higher resolution. Using these bases is possible to control at the same time both the local and global features of the surface, potentially leading to better solutions. They called multi-resolution bases or wavelets (although the latter term would not be strictly applicable from a mathematical viewpoint).

Surface shaping is also used to maximise the efficiency of radio telescopes or radiometers operating at millimetre and sub-millimetre wave frequencies. As they have very large reflectors (in wavelength units) the number of basis functions would be extremely high and the approach just described would be impractical, therefore use is made of methods based on geometric or Gaussian optics (which describes the field reflected from a surface by a superimposition of modes, having a concentrated field distributions somewhat alike a beam of light). The difference from the previous case is radical, as in both cases the field is described locally using the geometrical properties of the optical paths and corresponding power flux tubes. The reflecting surface is generated by progressively modifying a paraboloid, so that the rays from its focus, instead of being all reflected parallel to the focal axis, are distributed to form a shaped beam. By calculating the power associated with each flow tube it is possible to define not only the shape but also the power distribution within the beam. This also allows the calculation of the power distribution over equivalent aperture of the antenna that is then integrated to compute the radiation pattern with better accuracy. These methods, mainly the first, are frequently used as the first step in the shaping of smaller reflectors.

The use of geometrical optics is also helpful for the visualisation of some important characteristics of shaped reflectors. Referring for simplicity to a two-dimensional geometry (figure 4.4), it is possible to represent the feed radiation with a bundle of rays having an inversely proportional spacing with respect to the power flux density, so that all the flux tubes carry approximately the same total power. In this way it is possible to

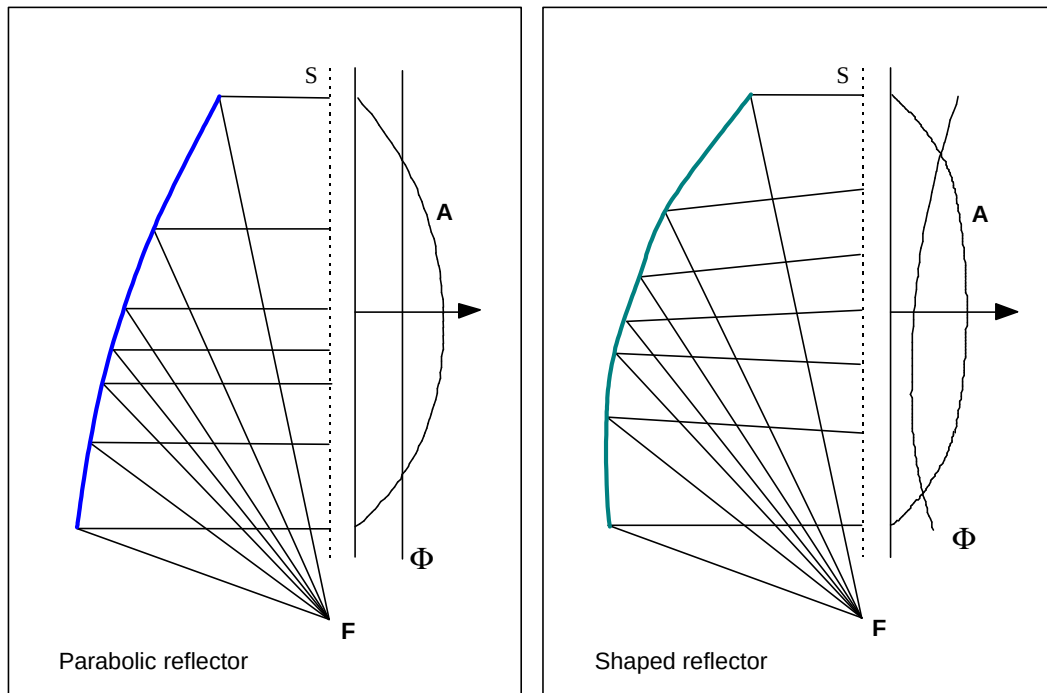


Figure 4.4 – The effect of reflector shape on the aperture distribution

visualise the power flow up to the antenna aperture. Reflection on an infinite parabolic surface of the spherical wavefront generated by a point source would give rise to an aperture field distribution with a uniform phase (equal path length) on a plane normal to the paraboloid axis (aperture plane), although the aperture would be at infinite distance. A normal reflector antenna approximates this theoretical behaviour. Instead a shaped surface generates an amplitude distribution that depends on its shape. For example, it is possible to shape it to obtain a field with more uniform amplitude in the central area of the aperture, as shown on the right side of the figure 4.4. However the phase will no longer be uniform, owing to the differences in the path travelled along different rays from the focus to the aperture plane. The beam generated by the reflector will therefore be wider and have lower directivity. The more uniform illumination tends to increase the antenna gain, but the phase error acts in the opposite direction, setting a limit on the advantage that can be obtained. To circumvent this problem, it is possible to use a two-reflector system (figure 4.5). The first reflector is used to generate the desired power distribution, and the second is shaped in such a way to equalize the lengths of the optical paths, thus restoring an approximately uniform phase distribution over the antenna aperture.

It is important to underline that, although it would be sufficient to alter the reflector surface by at most one wavelength to obtain any desired phase on the aperture, the need to redistribute also amplitude combined with the continuity of the surface and its first and second gradients, result in much larger variations (figure 4.6).

The transform-chain model can be applied also to understand the behaviour of shaped-reflector antennas. Starting from the desired shaped coverage and applying an inverse Fourier transform, it is possible to obtain an equivalent distribution over the antenna aperture. The reflecting surface has to be modified to concentrate the radiated power in a

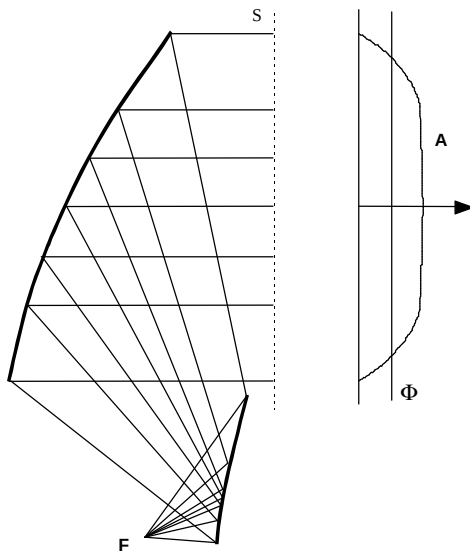


Figure 4.5 - Double shaped reflector antenna

phase over the coverage. Considering that the phase distribution over coverage is irrelevant, it is clear that the above procedure can not be applied without further refinements as the solution obtained is not likely to be optimal. Various schemes have been proposed to overcome this limitation. The main problem with all of them is that the reconstruction of the source distribution from the modulus of the field leads to non-linear

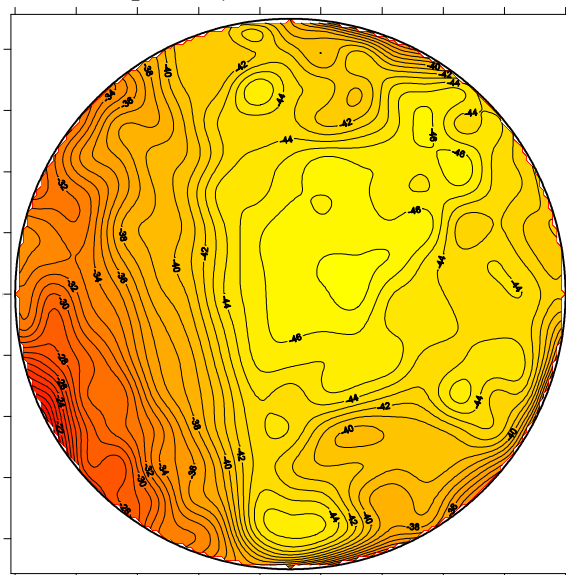


Figure 4.6 - Difference from the ideal parabolic profile (mm) in a shaped reflector for 30GHz operations.

region of the aperture plane so as to match this distribution (in a conjugate sense) and the problem is equivalent to that of finding the transformation necessary to map the field radiated by the feed onto the desired aperture distribution.

Unfortunately this procedure introduces some indeterminacy in the problem. The shape of the coverage on earth defines only the amplitude of the field, or rather the square of its modulus, since a power density is generally specified. The coverage shape is typically asymmetric, and assuming that the distribution is purely real its Fourier transform is a complex function. The phase of this function, i.e. the phase distribution over the aperture, is not necessarily the one leading to the best overall result, as there could be other phase distributions that produce a beam with higher gain, but also a non-uniform

system of equations, which are rather hard to solve. Usually the solution is found starting from an initial guess of the reflector shape, e.g. one obtained with geometrical optics, which then refined using an optimisation technique working on the differences between the desired and actual shape of the antenna beam. The price to pay is that at each step it is necessary to recompute the antenna pattern, which may be quite costly if the reflector is very large.

As said before, sometimes it is useful to be able to change the antenna coverage during the satellite life. Variable-shape reflectors are then an attractive possibility. They can be realised using techniques similar to those used for foldable reflectors. A first solution is to use a highly flexible reflecting surface, modifying its shape by means of motorised control

points. However, while it is relatively easy to achieve good results in the laboratory, it is much more difficult to ensure good performances for an antenna in orbit, where many uncontrollable factors are present, while movements of several wavelengths are required together with a maximum acceptable shape error of the order of a few hundredths of wavelength.

### 4.3 Double reflector antennas

Double reflector antennas are used in direct broadcasting and other applications, whenever a single, possibly shaped, beam is required. They can also be used for multiple-beam antennas, but the presence of two reflectors further limits the scan capabilities, unless they are made larger than it would strictly be necessary from a gain point of view. Their most evident drawback is clearly in the added mass and mechanical complexity.

The basic reason for introducing an additional reflector is to make the antenna system more compact. The added reflector, typically a portion of an ellipsoid or hyperboloid of revolution, provides two advantages in this respect. First it folds the propagation path from the feed to the main reflector reducing the length of the antenna. Second it can provide some angular magnification making the complete system work as if it had a longer (or shorter) focal length. This is exactly the same principle applied in telescopes and camera lenses.

Placing a flat mirror between the feed and the reflector creates a new focus at the mirror image of the original one, where the feed would have to be placed. Rotating the flat mirror would change the position of the focus and thus that of the feed. This is actually a way to create a moveable beam with a single feed antenna, although the main reflector has to be made larger still to capture most of the power radiated by the feed and reflected by the mirror at different angles. Replacing the flat mirror with an elliptic or hyperbolic one, having one focus coincident with the main reflector focus, a feed placed at its other focus will correctly illuminate the main reflector and generate the desired beam. The beam reflected by the secondary reflector (sub-reflector) will have an angular width that depends on the choice of the focal distance and ellipticity. Consequently the (virtual) image of the feed aperture present at the main reflector focus will be of a different size from the actual one.

An elliptical reflector is normally used with the feed in the outer focus so that the reflected beam is (much) wider than the feed beam. A hyperbolic reflector is normally used in the same way and with the same result. Looking at the transform-chain model it is clear that the system, comprising now four transforms, is equivalent to one with two, but with different characteristics from the ones of the main reflector alone. The equivalent reflector will therefore have a different focal length.

As already mentioned in Chapter 2, the behaviour of any double (or multiple) reflector antenna can be described in the first approximation by an equivalent single reflector system. The diameter of the reflector will be the same of the main reflector of the double reflector system; its focal length will instead have to be such that the feed will see the



Ku-band double reflector antenna  
(courtesy of EADS-CASA)

reflector under the same angle at it sees the sub-reflector. In both the cases described above the result is a focal length larger than the one of the main reflector. Such increase of the (equivalent)  $f/D$  ratio decreases most of the aberration typical of microwave reflector systems discussed earlier. The price to be paid is an increase in the power loss due to the additional spill-over and ohmic losses from the sub-reflectors.

#### **4.4 Degrees of freedom of an antenna**

Dealing with shaped reflector antennas, where the number of available design parameters is hard to count, it is useful to introduce the concept of degrees of freedom of an antenna, which is however applicable to any type of antennas. Whichever is the solution adopted for coverage shaping, it is clear that there is a limit beyond which an increase in the number of feeds, surface shape coefficients or array size brings no improvement, despite the increased number of free parameters. This fact can be formulated more rigorously by defining the degrees of freedom of an antenna, similarly to what is done in systems theory and in mechanics, for example to quantify the freedom of movement of an object.

Consider a closed surface enclosing an antenna at such a distance that the static and quasi-static components of the field are negligible, or explicitly neglecting them, so as to keep only the radiated field, which is the one of interest. The equivalence theorem ensures that there is one and only one equivalent source distribution over that surface that radiates the same field. Maxwell's equations ensure its continuity and require that the maximum variation of the amplitude and phase between any two points be finite and compatible with the wavelength. Applying the sampling theorem, it is possible to define the number of samples necessary to completely describe the field, which is also the number of degrees of freedom of the antenna.

This property can also be recognised with the following reasoning. The field extends indefinitely throughout space while for obvious physical reasons its energy must be finite. Therefore the Parseval theorem ensures that its angular spectrum, i.e. the two-dimensional Fourier transform of its distribution over a spherical surface, be square-integrable, thus the field distribution is band-limited. As a consequence the number samples required to describe it completely is finite.

The surface chosen to apply the equivalence theorem must enclose a volume of space greater than or equal to that occupied by the antenna, and the number of degrees of freedom of the distribution is an approximation by excess to the number of degrees of freedom of the antenna. However, it can be shown that the number of degrees of freedom of an "optimal" antenna is approximately equal to its peak gain, expressed in natural numbers. Considering an expansion in spherical harmonics of the field over the minimum enclosing sphere, it is found that the highest significant harmonic required to express the field is

$$N = p \frac{d}{l},$$

while the peak gain  $G$  is approximately

$$G = N^2 + 2N$$

where the right hand is the total number of spherical harmonics with order less than  $N$ .

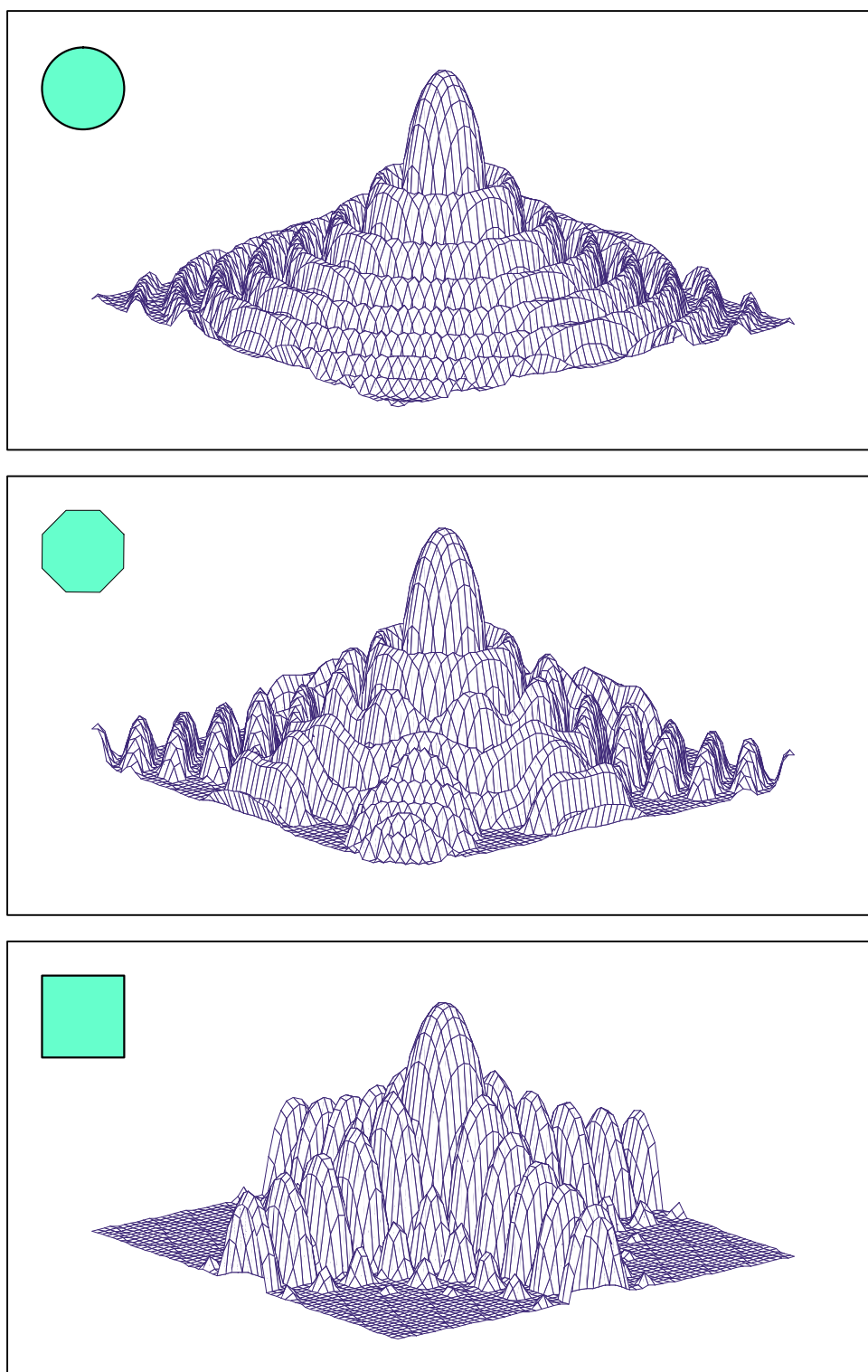


Figure 4.7 - Effect of the change of shape of the aperture

Substituting the value of  $N$  the following expression is obtained for the gain

$$G = \left( p \frac{d}{l} \right)^2 + 2p \frac{d}{l},$$

which is very similar to the expression for the gain of a uniformly illuminated circular aperture. The additional linear term accounts for the increase in gain obtainable by a spherical distribution of sources, showing that a tri-dimensional structure can have higher gain than an aperture of the same diameter. Note that the relative increase of gain rapidly diminishes with growing antenna sizes (with respect to the wavelength) and it is, in fact, significant only for relatively small antennas.

This formulation, unfortunately, is not very helpful for improving the estimation of the number of degrees of freedom, since it requires that the antenna be "optimal", in the sense that its peak gain can not be improved further - a purely theoretic condition. Figure 4.7 illustrates a good example of the related uncertainty: the variation in amplitude and structure of the sidelobes in the radiation pattern of apertures with different shapes, but equal illumination and peak gain. The question may arise for instance whether the optimal antenna is a non-circular one with uniform illumination or a circular one with a different illumination but still enclosed by the same minimum sphere and having a rather similar lobe structure.

The degrees of freedom of the antenna are associated with its radiation throughout the whole space, whereas the designer frequently aims to control only a very limited region of space within the major lobe of the radiation pattern. It is therefore necessary to determine whether it is possible to select from the many available parameters the ones having a more direct effect on the desired characteristics of the antenna. Nevertheless having a way to know the total number of free parameters is good to fix a starting point and a limit beyond which it is useless to go.

An antenna with a diameter equal to a few wavelengths already has a strongly directive radiation pattern. Therefore a surface, for example a sphere, enclosing it is strongly illuminated only in a limited area. The width of this area roughly corresponds to the angular aperture of the major lobe and of the first sidelobe. The ratio of the area of this region to the area of the whole surface can then be used to roughly estimate the minimum number of parameters required to control effectively the antenna radiation pattern. The remaining degrees of freedom can be thought as being associated with the rest of the sphere, where the field level is low, as they have a predominant effect only in this large but unimportant portion of the radiation pattern. It is interesting to note that this procedure is essentially equivalent to counting the number of elementary beams, i.e. beams radiated by a uniformly illuminated aperture of the same area of the antenna, which would fit into its main beam. However the former approach is more rigorous than the latter, since the elementary beams are not orthogonal. An even more rigorous estimate can be made resorting to a multi-resolution expansion, which allows the definition of the number of independent functions available for a selected aperture size over a selected portion of the sphere.

The different weights of the free parameters may be illustrated further with another simple example. We shall consider a circular aperture of unitary radius with illumination of the polynomial type, which is a good first order approximation of a reflector antenna.

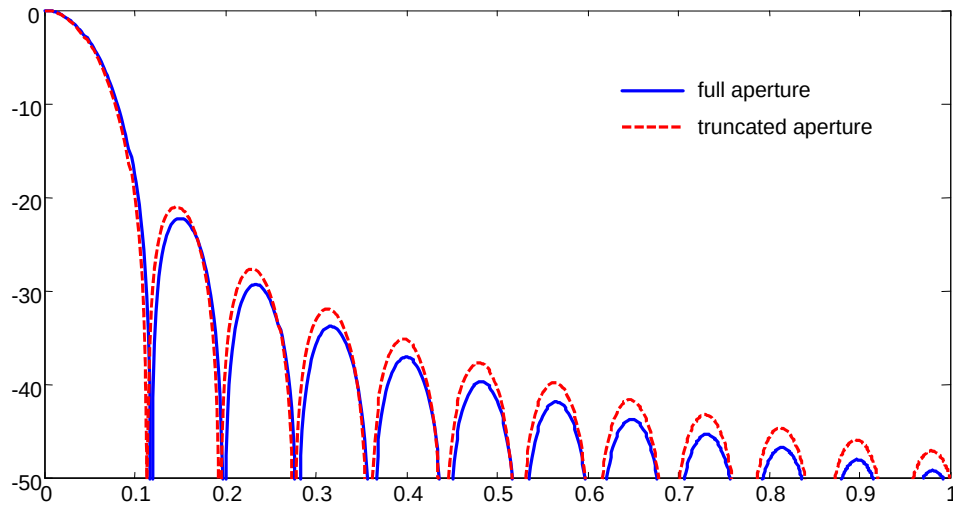


Figure 4.8 – Structure of the sidelobes of a circular aperture illuminated with a 10dB taper and of the same aperture truncated at a radius of 0.9 of the initial value.

The field distribution over the aperture, or illumination, is expressed as:

$$I(r) = I_0 \left[ \sqrt{0.1} + (1 - \sqrt{0.1})(1 - r^2) \right]$$

where  $r$  is the radius of the aperture and  $I_0$  is the maximum value, attained at the centre and  $\sqrt{0.1}$  (-10dB) is the aperture taper, i.e. the level at the aperture edge. The peripheral annular region for  $0.9 < r \leq 1$  covers 19% of the aperture and contributes to the radiation with only 0.23% of the power. Removing it, i.e. setting the field to zero for  $0.9 < r \leq 1$ , reduces the gain by only 0.01dB without significantly altering the main beam, but changes noticeably the levels of all sidelobes and especially of the most distant ones (figure 4.8). In other words, 19% of the degrees of freedom are linked to just 0.23% of the total radiated power and have a visible impact only on the low level regions of the pattern.

Reversing the point of view, the above fact gives a measure of the difficulty in controlling the sidelobes, since very small changes in the antenna will result in large changes in their level and even in their structure.

An important consequence of these considerations is that antennas are normally designed using a number of parameters greater than the theoretical minimum to be able to compensate for spurious effect originating from the "uncontrolled" degrees of freedom. Obviously, the introduction of redundant free parameters can not ensure better performance but, if the parameters are carefully chosen, can provide considerable help in the design optimisation process. For instance, it facilitates the isolation of groups of configuration parameters related only to one performance figure and therefore allows a more systematic approach to design. The structures obtained in this way, however, are inevitably more complicated and less efficient than strictly necessary according to theory. A perfect example of this situation are array antennas, the number of available amplitude and phase controls is  $2(N-1)$ , if  $N$  is the number of elements, and they are seldom completely used, not to mention the free parameters linked to the element alone. The

obvious consequence is that arrays are typically more complex and less efficient than other antennas that could at least theoretically achieve the same performances, on the other side they are often much easier to design.

#### 4.5 The resonant discharge in vacuum

The design of antennas for direct television broadcasting, and of all other antennas in which the transmitted power is high, faces problems related to the flow of high power waves through the antenna components. There are three key issues: thermal dissipation due to ohmic losses, resonant discharges in vacuum and passive intermodulation. The first aspect is quite straightforward and requires careful thermal design to make sure the thermal power is correctly radiated into space; the third has been discussed previously and only the second is left for closer examination.

Resonant discharges in vacuum are a complex phenomenon, some aspects of which are still under discussion. In particular, some uncertainty exists about the safety factors to be used in the design of microwave components. They are related to the quantum characteristics of the materials used, which determine their sensitivity to the phenomenon, and to the spectrum of the signal carried by the device, which determines the ratio between the peak value and the mean power and consequently the probability that a discharge will be initiated.

The structure illustrated in figure 4.9 will be considered to explain the phenomenon. The two walls (thick horizontal lines) are made of metal, as for example in the longitudinal section of a waveguide, and the interior is in high vacuum conditions. A wave with a frequency  $f$  propagates in the structure. The difference of potential  $V(t)$  between the two opposite walls in a section  $z=z_0$  will have a sinusoidal variation with period  $f^{-1}$ . Assume that an electron, thanks to its thermal energy, escapes the potential barrier at the surface

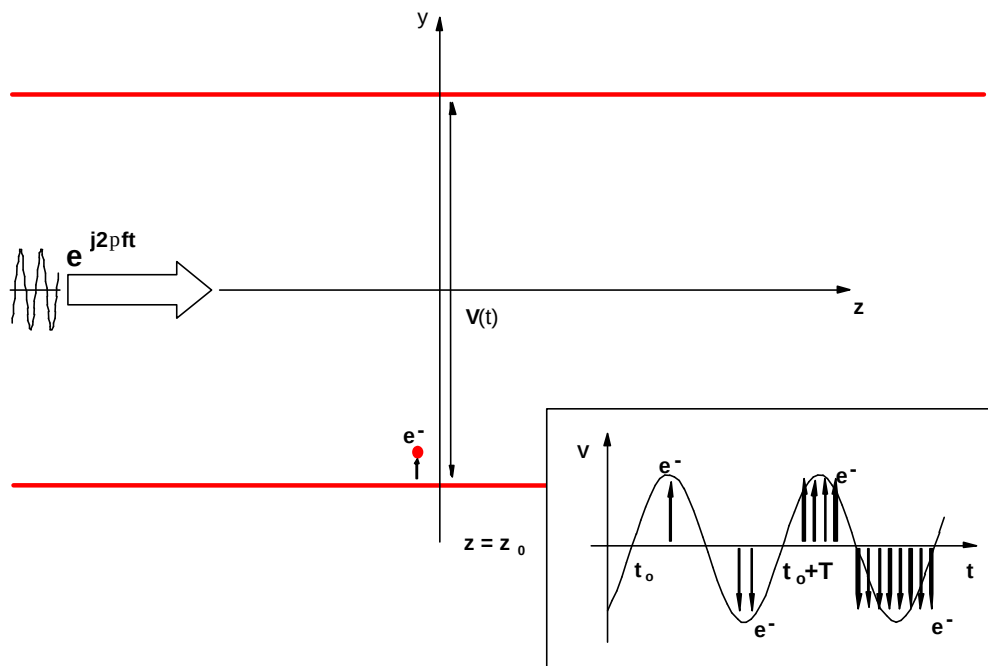


Figure 4.9 - Resonant discharge in vacuum

of the lower plate at the instant  $t_0$  at which the positive half-period of  $V(t)$  begins. The electron is then accelerated by the electric field in the direction of the opposite wall. If the electric field is inverted with the opposite half-wave before it reaches the opposite wall it will be slowed down or even stopped and accelerated in the opposite direction. But if the signal frequency is low enough it will instead reach the opposite wall at the instant  $t_1$ . If the energy acquired from the electric field is sufficient, i.e. the difference of potential sufficiently high, the electron may extract two or more electrons from the opposite wall when it strikes. If additionally, the time  $t_1$  coincides with, or is very close to, the start of the opposite half-wave of the potential  $V(t)$ , these electrons will be in turn accelerated in the direction of the first wall thus initiating an avalanche phenomenon, which is called resonant discharge in vacuum or multipaction. The difference of potential is directly related to the square root of the power carried by the wave, which is therefore an important parameter. Furthermore the synchronisation of the voltage oscillation with the electron flow is linked to the product of the frequency times the distance between the two walls, making the  $fd$  product another important parameter for establishing the possible occurrence of this phenomenon.

The electron cloud created after a few impact cycles oscillates at the frequency of the RF signal between the two walls and grows with each successive impact. Thus it creates a current flowing between the two walls of the waveguide, i.e. a short circuit. Over a short time, the density of the electron cloud increases making the repulsive force among electrons dominant and therefore limits their velocity and spreads the cloud. In other words, the field associated with the electrically charged cloud tends to neutralise the incident RF field, masking its accelerating effect, and the discharge is extinguished in a few milliseconds.

Unfortunately, the energy of the cloud may become large enough to extract some ions from the wall material with the successive impacts. The ions limit the velocity of the cloud because of the impacts with the electrons, but in limit conditions their density may increase to the point of initiating a gas discharge, with destructive effects due to the increased spark energy and the ablative effect of the impact of the ions on the material.

The effect of the resonant discharge is double: on the one hand, it causes physical damage, albeit small, to the walls of the structure, and on the other hand it causes a brief but consistent peak in the reflection coefficient of the section of guide affected by the short circuit. This second phenomenon induces serious disturbance to transmissions and may, in the long term, cause damage to the RF amplifiers or generators powering the system, which generally can not withstand a high reflected power. Since as soon as the cloud is dispersed the conditions for a new discharge are created again, the phenomenon is repetitive, and even in the less harmful cases permanently compromises the operation of the system.

The resonant characteristics of the phenomenon make avoiding it by design reasonably simple. Essentially it is a matter of selecting suitable values for the parameters governing the (absence of) discharge: the distance between the opposite walls, the frequency of the signal and the power carried by the wave. In practice, however, conflicting design requirements may enter into play complicating the implementation of conceptually simple solutions. For instance, since frequency and power level are usually fixed, it may not be easy to change the geometry of a device without compromising its performances.

The minimum peak voltage required for the initiation of the discharge may be expressed as:

$$V_p = k f d$$

where  $f$  is the frequency and  $d$  is the distance between two points at opposite potential, while  $k$  is a constant that depends on the material, its surface finish and the geometry of the structure, and is roughly equal to  $0.02 \text{ [sm}^{-1}\text{V}^{-1}\text{]}$ . The minimum average power propagating along a guiding structure of impedance  $Z_0$  and necessary for the discharge is therefore given by

$$P_e = \frac{1}{2} \frac{(k f d)^2}{Z_0}$$

so that to avoid a discharge it must be

$$d \geq \frac{1}{k f} \sqrt{2 Z_0 P_e}.$$

This simple expression does not account for two thresholds: the voltage (power) threshold and the  $fd$  threshold, both associated with the accelerations required to overweight the initial random escape velocity of electron and impart a coherent motion to the first free electrons. These thresholds are normally quite low and strongly dependent on the material and finishing of the structures, so that experimental data are necessary to fix them.

Sharp edges and any marked discontinuities tend to increase the density of the electric field and constitute higher risk points. The same applies to all areas in which the surface finish or contacts between different metals facilitate the emission of electrons. When for some design reasons, it is not possible to achieve the required safety factors, generally 6dB, testing is used to verify the absence of discharge. Alternatively the components may be pressurised or a solid dielectric with low losses and a low dielectric constant may be inserted to fill the gap.

Since a condition of very high vacuum is necessary for the initiation of the phenomenon, which in general takes place inside structures with a rather complex geometry, experimental verification is not easy and requires special and rather expensive equipment. However the problem can be easily addressed in the design using the formula given above as well as other means, thus limiting the use of testing to the unavoidable cases. Typical examples are the waveguide sections just before or within the radiating elements, like filters, combiners and polarisers, which have the most complex geometries and where the power of several channels may be combined.

#### **4.6 Small receiving antennas**

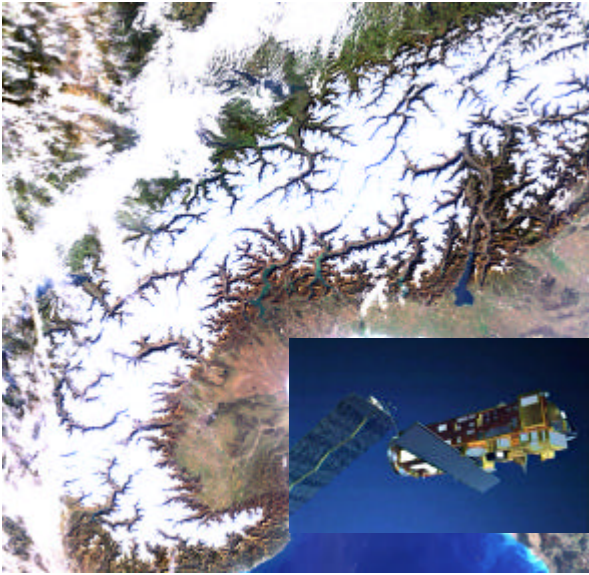
To conclude the discussion of direct television broadcasting, mention should be made of the aspects relating to receiving antennas on ground. These antennas, usually of the reflector type, with a diameter of 60cm to one metre, are not particularly sophisticated. Attention is focused on the reduction of their size and ease of installation and use, e.g. antennas that can simultaneously receive a number of satellites located in different directions, avoiding the use of rotors.

Microwave antennas are normally produced in small runs, as in the case of antennas for radio links, or in a few items, such as those to be installed on satellites. In this case instead antennas providing good performance at low cost have to be produced in large runs, similarly to what happens for mobile phones and base station antennas. From this point of view, array antennas made from plastic with printed feeding networks and radiators are an ideal solution, at least in theory. If the antennas normally available on the market at present are of the reflector type is because the solution most rapidly available when the market was first established was to use the existing production capacity for parabolic dishes for radio links and only the feeds had to be built on purpose. The marketing of array antennas would instead have required the development of suitable production processes. As it often happens, once a solution is on the market it is by far easier to improve upon it than to launch a completely different alternative, so array antennas are seldom seen in satellite TV receivers.



## CHAPTER 5

# Remote sensing



The Alps seen by Envisat with an artist view of this ESA remote sensing satellite

The use of microwave systems for remote sensing from satellites, begun in the '60s for military purposes, has grown remarkably in recent years, also as a result of the general attention for the environment. Peaceful applications are many, ranging from the most traditional ones as meteorology, geology and oceanography, to more recent ones, like environment management, agriculture control and disaster prevention and relief.

A special feature of remote sensing missions is that, although they are designed with some specific applications in mind, the data gathered by them are frequently used for very different purposes. Therefore the objective is usually to obtain the maximum amount of data of the best possible quality from each instrument. In other words, each instrument is treated as scientific equipment, rather than as payload of a commercial mission. In fact the use of the term

"instrument" which is in common with scientific missions, rather than transponder or payload, typical of communication satellites, is a clear sign of this attitude.

Remote sensing systems operate over a much wider frequency range than communication ones, i.e. from a few hundreds MHz to hundreds of GHz and beyond. The upper limit is conventionally placed at the lower bound of the infrared region of the electromagnetic radiation spectrum, marking the boundary between the "microwave" and "optical" frequency bands, a traditional distinction becoming increasingly blurred.

Although the characteristics of remote sensing instruments may differ widely, they can be grouped into a few broad categories.

*Altimeters*: pulsed radar systems, used to measure the distance between the satellite and the surface of the Earth or other planets, and therefore suitable for plotting altitude profiles of the surface.

*Scatterometers*: pulsed or continuous-wave radar systems, whose principal application is

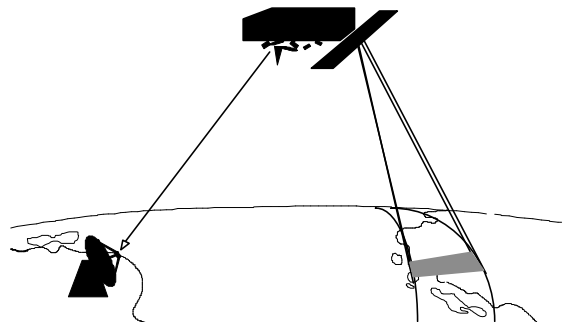
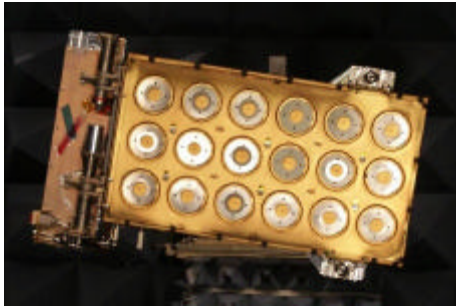


Figure 5.1 – Architecture of remote sensing systems

in the measurement of the waving of rough surfaces, using the fact that echoes vary with its amplitude; over the sea they are used to determine by further processing the wind strength and direction.

*Synthetic aperture radar (SAR)*: radar systems whose resolution is increased by the processing of the echoes obtained from different angles and used to generate high-definition images of the Earth's (or planetary) surface.



The antenna of the GRAS sounder embarked on METOP (courtesy of Saab Ericsson Space)

*Sounders*: instruments used to determine the volumetric characteristics of an RF transparent medium, typically Earth's atmosphere, recording the variations of a signal travelling through it. The signal may be originated by thermal radiation, astronomic radio sources (stars) or artificial ones (e.g. GPS satellites). Also active sounders are possible, the signal is generated by the instrument itself and reflection rather transmission is measured.

*Radiometers*: passive instruments that capture the thermal radiation of objects. They are used either as volumetric instruments to determine, for example, the temperature or vapour density of the atmosphere, or to measure surface characteristics like soil humidity, soil temperature, salinity of the sea, etc.

*Synthetic aperture radiometers*: radiometers with increased resolution obtained by processing the data measured on the same scene from a number of receivers.

In most cases the data generated by these instruments are not used in isolation, but are combined (correlated) so that more complete information is obtained on the object under examination. This data processing work is frequently long and complex, and suffers from the combination of the inevitable measurement error, random or intrinsic, of each instrument as well as from the inevitable misalignment among them, for example, the imprecise time matching among the data produced by different satellites. Therefore the search continues for more powerful and sophisticated instruments, so as to obtain more precise measurements.

All the above systems share the basic characteristic of being measuring instruments. The main design objective is therefore to maximise their sensitivity and resolution, to reduce the acquisition time and to extend the operating limits, for example into increasingly high frequencies or improving the instrument hardware. The antenna is the sensor and is therefore a key element of the system. As discussed at the beginning of this book, the radical differences between remote sensing instruments and communication transponders entails the use of quite different parameters to evaluate their characteristics and those of antennas in particular.

Obviously, since the parameters that describe the beam radiated by an antenna are related with each other, those used for remote sensing antennas may easily be derived from those discussed till now, or viceversa. The major difference lies in the fact that integral quantities, such as the ratio between the power flux in the major lobe and the total power, are important, rather than point values, such as the peak gain. This is not surprising as in communication antennas it is the quality of a point-to-point link that matters, whereas in remote sensing it is required is to measure global values like sensitivity or resolution.

In antenna design it is possible to further restrict the categories listed above to four, on the basis of the similarity of the requirements: radar antennas (altimeters, scatterometers and active sounders), radiometer antennas (including passive sounders) and synthetic aperture radars and radiometers, which have rather different requirements but both based on the use of large arrays.

All instruments have been traditionally used on satellites placed in low non-equatorial orbits (300-1000km altitude). Owing to Earth rotation, the satellite passes over different points in each orbit. In almost all cases, remote sensing satellites use polar orbits, which are the only ones enabling an almost complete scan of Earth's surface, leaving out only a small part of the Polar Regions. A notable exception the rule are meteorological satellites like Meteosat [1], which operate from geostationary orbit, to allow continuous monitoring of a specific region. More recently geostationary remote sensing missions are being studied also for other types of applications owing to the increased interest in the continuous monitoring of other phenomena.

### 5.1 Radar systems

The antennas for systems grouped in the radar category are mainly characterised by their resolution, as there is no processing of the return signal (echo) to enhance it. The antenna has to generate a beam that projects a specified footprint on ground, determining the resolution, and which sidelobes are contained within amplitude masks so as to minimise the interference from spurious echoes or other signals coming from areas outside the footprint. The actual variation of the gain within the footprint, generally defined by another mask, is not very critical, at least as long as it is known with sufficient accuracy and can be taken into account in the use of the data produced by the instrument.



SAR/Interferometric Radar Altimeter antenna for the Cryosat satellite (courtesy of Saab Ericsson Space)

The antennas used for these instruments are either of the reflector type, for large apertures, or of the array type, that combines ease of beam pointing with high efficiency, essential to minimise the transmitted power, and provides good flexibility in the control of sidelobes. Moreover, since there is no need for multiple beams or special shaping of the footprint, the complexity of an array is relatively low. Reflector antennas, which may be illuminated by arrays of radiators, are typically used for special applications; for example, when it is important that the beam shape remains unchanged with a change of pointing.

### 5.2 Degrees of freedom of arrays

Also for arrays, it is possible to analyse the synthesis problem through the study of the number of degrees of freedom, leading to results closer to practical applicability than for reflector antennas as it is demonstrated by the existence of numerous array synthesis methods. There are two main reasons for such difference, first the Fourier transform relation between aperture field and radiation pattern offers a much better approximation for arrays than for reflector antennas and, second, the number of degrees of freedom directly related to accessible control points, i.e. the element input ports, is much larger than in reflector antennas.

Once the type of radiating element has been selected, the amplitude and phase of each of them remains to be assigned, i.e.  $2N$  real values for  $N$  elements. The total power, which has no effect on the shape of the radiation pattern, is normally assumed to be unitary resulting in a constraint on the amplitudes. Phases, being intrinsically relative values, are defined up to a constant, and it is possible to fix the phase of one element to zero via a further constraint. Therefore an  $N$  element array has  $2(N-1)$  free parameters.

The mutual coupling between adjacent elements introduces further dependency among the parameters, which has unfortunately a very complicate mathematical formulation. Nevertheless the actual number of available degrees of freedom is smaller than the number of parameters to be determined, leading to an over-determined problem, which can only be solved by minimising the norm of the error, e.g. in a least square sense. Furthermore, as mentioned previously, the conditions expressing the desired performances are not independent, making the problem quite difficult to be dealt rigorously in the most general case. To overcome this problem numerical optimisation methods are routinely used also in array antenna design.

### 5.3 Synthetic aperture radars

Various types of system are given the generic name of synthetic aperture radar, or SAR, because they are based on similar operating principles. The common feature justifying the single name is the fact that a coherent system, which provides amplitude and phase information, is used to excite and measure the echoes of a scene, for example the Earth's surface, which are processed to reconstruct the details of the scene. The reconstruction process, based on signal transformations (filtering), yields a result similar to the one obtained with a radar system with a much wider aperture, hence the name "synthetic aperture radar". The domain in which the filters are applied, time, frequency or space, separately or in combination, for the two image scanning axes, gives rise to different types of SAR instruments.



The ERS-1 SAR antenna in the RF test chamber (courtesy of Saab Ericsson Space)

The transformations are carried out digitally, and the use of synthetic aperture radar has received a considerable stimulus in the last decades from the great reduction of the resources needed for digital processing, in terms of cost, energy and time.

Without going into the details of the synthesis systems, it is important to note that SAR antennas are generally characterised by a marked elongation in the direction of the orbit, so as to obtain footprints on earth, which are very elongated in the direction perpendicular to the projection of

the orbit on its surface. Typically, the illumination is uniform along the major axis, providing a beam which is as narrow as possible in the direction of movement, at the price of higher sidelobes, while in the perpendicular direction the illumination is chosen to limit the sidelobe level and a obtain shaped beam that compensate for the variation of attenuation with the distance of the points on ground from the antenna (range attenuation).

In terms of design, the same considerations as for normal radar systems apply, with the additional point that the large longitudinal dimension of the antenna implies the use of separated panels connected by folding mechanisms.

A major problem in SAR antennas, except active ones, is the high power flowing in the distribution network feeding the antenna. In this case, the problem is increased by the size and complexity of the foldable network and it is especially marked in the first sections in which the power is concentrated in a very small number of branches. Special arrangements are necessary to route the signals across panel boundaries with minimum losses, for instance contact-less waveguide connections, which are realised exploiting the fact that a quarter-wavelength open line acts as a short circuit.

The most recent developments in SAR systems are directed towards the use of multiple-beam and double linear polarisation antennas, which are able to gather much more information on the target scene. In particular, each of the four possible operating modes (table 5.1) permits the identification of objects, which are virtually invisible in some other modes. Consider for example grass, which has a roughly vertical extension and therefore a very low reflectivity in horizontal polarisation and, by contrast, the foliage of broad-leaved trees, which tends to be horizontal and therefore scarcely visible in vertical polarisation. Further, consider the tree trunks “hidden” by the foliage, which are however visible in vertical polarisation, scarcely reflected by leaves, or the possibility to distinguish conifers from broad-leaved trees.

| <b><i>Mode</i></b> | <b><i>Polarisation transmitted</i></b> | <b><i>Polarisation received</i></b> |
|--------------------|----------------------------------------|-------------------------------------|
| <b>H-H</b>         | Horizontal                             | Horizontal                          |
| <b>H-V</b>         | Horizontal                             | Vertical                            |
| <b>V-H</b>         | Vertical                               | Horizontal                          |
| <b>V-V</b>         | Vertical                               | Vertical                            |

Table 5-1 - Operating modes of a double-polarisation SAR

#### 5.4 Array synthesis

The synthesis of beams of a specified shape by means of arrays is a subject that has been studied in great detail for several decades, since the invention of radar and thanks to their extensive use. The literature describes numerous methods of synthesis, many of them derived from similar techniques applied in filter theory. The existence of a relation equivalent to a Fourier transform between the distribution of (equivalent) currents on the array aperture and the far-field radiation pattern provides an excellent starting point for the design. As discussed previously, however, the desired pattern is defined only in terms of amplitude and the absence of phase information makes it impossible to derive the distribution over the aperture by simple inversion of the transform.

As a first step, it would be possible to associate a phase variation to the far-field amplitude profile, for example a uniform phase in the major lobe and on each sidelobe with jumps of  $180^\circ$  at each zero, and calculate the inverse transform of the resulting (real) function. With this method, however, there is no guarantee that the distribution obtained will be achievable and have minimum phase, i.e. require the minimum possible number of complete phase rotations of the excitations between two opposite sides of the array. One way of alleviating this problem is to filter (smooth), e.g. by windowed averaging, the phase of the distribution obtained on the array aperture, anti-transform the new distribution and

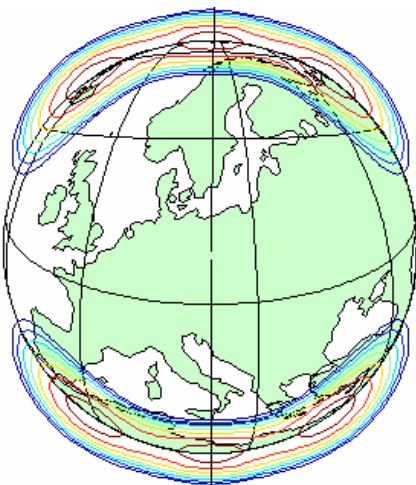
associate the new far-field phase profile to the original amplitude profile. The process is then repeated iteratively and if the smoothing operator has been selected properly it converges, however there is still no guarantee that the final solution has minimum phase.

This aspect is particularly important for array antennas, since the reciprocal influence (mutual coupling) of nearby radiating elements makes it difficult or impossible to produce distributions with large phase differences between any two adjacent elements. The situation is further complicated by the fact that an array is formed, by definition, from discrete elements and therefore only a discontinuous approximation of the continuous aperture distribution generated by the inverse transform can be obtained. Thus, if the first does not have minimum phase, the phase discontinuities introduced by the sampling process have a more pronounced effects on the radiation pattern. To remove this obstacle, the excitations can be obtained by convolution of the desired distribution with the conjugate of the aperture field of each element, possibly finding the sample values by minimising a suitable error function, like the square sum of the errors corresponding to each element (least squares). Alternatively, the diagram to be synthesised can be weighted with the transform of the element pattern in order to isolate the array factor, which can be associated with a discrete sequence. A set of excitations compatible with the assigned element pattern can then be obtained from the discrete inverse transform, e.g. using an FFT. Using the FFT implies the assumption that a geometric ( $L^2$ ) norm is the proper choice. Different norms can also be used leading to other inversion algorithms.

As it is always the case in design, the synthesis of the array radiation pattern requires a compromise between conflicting requirements. The objective and the constraints to be satisfied, for example, the gain, the aperture of the major lobe (size of the footprint on earth) and the level of the sidelobes, are not independent of each other, i.e. changing one parameter to maximise the gain will also affect the others. For instance the gain and the aperture of the beam are inversely related and any attempt to reduce the level of the sidelobes, beyond a certain value, while keeping the aperture fixed, entails a widening of the main beam and a reduction of gain. More in general, there are only a finite number of parameters available to obtain a radiation pattern that satisfies the assigned constraints

and it is easily seen that an increase in the number of parameters used to satisfy one constraint makes it more difficult to satisfy the remaining ones, unless they are automatically satisfied.

The selection of the number and type of parameters to be used to satisfy each constraint is normally made implicitly, i.e. selecting the antenna configuration considered most suitable as starting point for the design. For example, if a uniform distribution of excitations for the radiators is selected, it is possible to modify the gain by varying the number of elements, while the level of the sidelobes is fixed by the element radiation pattern. In arrays with a large number of elements, the array beam is much narrower than the element beam, so that the element radiation pattern mainly affects the level of the lower lobes, while that of the first few lobes is only dependent on the array excitations.



The coverage of the two antennas of the GRAS sounder embarked on the Eumetsat METOP satellite

In the initial synthesis aimed at the selection of the best candidate configuration, it is usually possible to reduce the problem to simple canonical cases for which design methods are available. Mutual coupling is usually disregarded in this phase and the initial design is further optimised using computer programs that accurately simulate the array behaviour, compare it with the assigned objectives and constraints and then adjust the available parameters, like array excitations, element positions and possibly orientation. At each iteration the selection of the new set of parameters is made according to the strategy used by the specific optimisation algorithm implemented by the computer tool.

Unfortunately, these algorithms do not generally ensure the identification of the optimal solution, i.e. the global optimum, but only of a sufficiently good one, which may still not completely meet the requirements. The engineer must therefore study the solution, evaluate its correctness, and if necessary carry out further optimisation runs starting from different initial conditions in order to check the possibility of obtaining better results. The only alternative, which is affordable only for a limited number of parameters and yet does not guarantee the result, is a parametric study.

An increasing use of genetic algorithms has been made recently, also for array design, trying to overcome the above limitations of classical optimisation methods. These algorithms emulate, in a rather simplified way, natural evolution, i.e. the mixing of the genetic heritage of two parents to create diversity and the survival of the fittest to ensure propagation of the best combinations of traits. This overly crude view of reality, which ignores the strong influence of collective interactions, e.g. co-operation for the survival, works exceptionally well on problems with a relatively small number of parameters (i.e. less than hundred) but more sophisticated solutions are needed beyond this point. Compared with conventional methods, these algorithms have the advantage of exploring wide regions of the solution space, leading to a high probability of finding the location of the global optimum. On the other side, since they are not deterministic, they do not guarantee the attainment of the optimum, whether local or global. Therefore they are best used in the initial search for promising solutions to be subsequently refined using conventional algorithms, which ensure that the optimum is finally attained.

As the initial synthesis typically disregard the effects of mutual coupling, they may be accounted for in the final optimisation by evaluating the radiation pattern of the element in the presence of all the others. As an alternative the radiation pattern of an element radiating in the presence of a limited number of neighbours can be used in the synthesis. Both approaches entail some approximations. The first, commonly used for small arrays, completely ignores the coupling in the optimisation phase, and the second, normally used for large arrays, ignores the fact that the coupling also affects the input impedance of each element. Moreover, equal conditions are frequently assumed for elements in the central area of the array as for peripheral ones, although the latter are not symmetrically surrounded by other elements and therefore have a different radiation pattern and impedance. Clearly it is also possible to use different patterns for elements placed in different positions, e.g. in the inner portion, on the side and on the corner, somewhat improving the results, but the only proper solution is to predict the array performances using a computational model of the array that accounts for all the mutual coupling effects. Such complete and accurate approach is perfectly feasible with the computing power and electromagnetic modelling tools currently available.

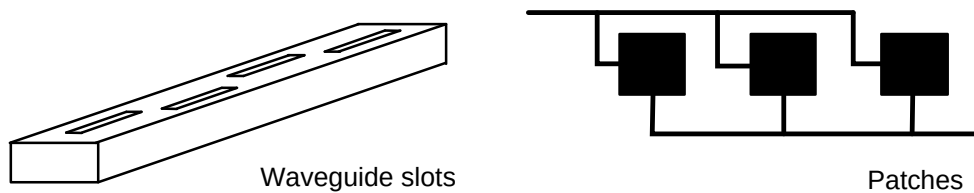


Figure 5.2 - Some typical radiators for array antennas

### 5.5 Radiating elements for arrays

From a technological point of view, there are many possible ways of building arrays for space applications with good characteristics and the choice depends on several and diverse factors, like the size of the array, the type of polarisation and the operating frequency. First of all, a distinction must be drawn between arrays based on travelling wave structures and those based on resonant structures. The former are constituted by a transmission line, terminated into a load, with radiating elements attached, at more or less regular intervals. They are not very used, owing to the frequency dependence of the beam pointing and to the dissipation of a small but significant part of the input power into the end load. Resonant arrays use instead feeding lines that are either terminated in a short circuit or in a radiating element. They are much more common and can be realised using different types of radiators (figure 5.2).

Waveguide slots are frequently used for single linear polarisation antennas with high input power (i.e. radars), mainly because of the very low losses of the waveguide and the simplicity of their construction. Unfortunately, while dual-polarisation radars are more and more used it is difficult to realise double-polarisation antennas with these elements while maintaining a high efficiency.

Planar elements (multiple-layer antennas), despite the limitation already discussed, have the advantage of being relatively light and structurally stable. This is important, particularly for SAR antennas, which are normally very large and have to be made of several equal panels. The high ohmic losses are their major drawback, which is especially relevant in radar system given the high transmitted power. To drastically reduce the impact of the ohmic losses it is possible to place the transmit and receive circuitry in the immediate vicinity of the antenna itself (active antennas), e.g. on the rear.

The amplifier and phase shifters necessary for this purpose are typically realised with solid-state components, using microwave integrated circuit (MMIC) technology, and are relatively compact and light. The major problem is to radiate into space the thermal power developed by them.

In large active antennas, or even passive beam scanning ones, the number of control points and thus of variable or active components can be very high. Therefore use is often made of sub-arrays, resulting in building blocks than make the series-production features of multiple-layer antennas even more attractive.



Prototype of Ka-band active array with high-efficiency hexagonal elements (courtesy of Saab Ericsson Space)

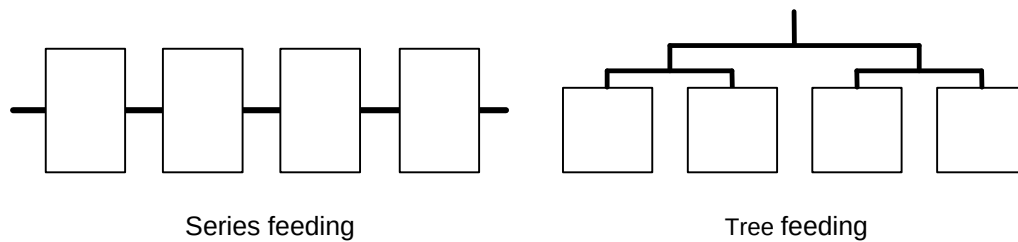


Figure 5.3 - Feeding structures for printed antennas

Two different arrangements are used in the non-resonant feeding of arrays or sub-arrays (figure 5.3). The first solution consists in feeding the radiators in series, ensuring that each element sees the same phase at the input port, thus limiting the total length of the lines. However, this creates a cascade of resonating elements and narrows the pass band. Furthermore, the excitation phase of the individual elements varies (slightly) with frequency, creating a linear phase progression and therefore a (slight) variation of the beam pointing. The second solution consists in using a mixed tree-like feeding architecture, also known as corporate feeding. The combination of serial and parallel distribution keeps line length within acceptable margins and avoids phase dispersion, which can be very close to zero if all paths are equal. Since this solution requires much more space for the feeding lines, the two approaches are often used in combination, with only small groups of radiators fed in series.

### 5.6 Active array antennas

The increased efficiency and reliability and the decreased cost of MMIC components have favoured an increased use of active array antennas in the space sector and elsewhere. Each active element of the antenna, radiating element or sub-array, is associated with a transmit-receive module (TR module), i.e. each radiator or group of a few radiators is fed by a separate RF amplifier and connected to a separate low-noise amplifier (figure 5.4). All amplitude and phase controls are realised on the other side of the amplifiers, which is less critical, as the signal power is much lower in transmit, eliminating power handling and ohmic loss problems, and much higher in receive, improving the system noise floor level.

One of the major problems to be overcome for the use of solid-state amplifiers in space is their still relatively low efficiency, when compared for instance to RF tubes, which rises DC power consumption and thermal dissipation. The latter aspect is particularly critical for active arrays, in which the active components have to be kept within the working temperature limits even though the antenna is completely exposed to the space environment. The major problem is

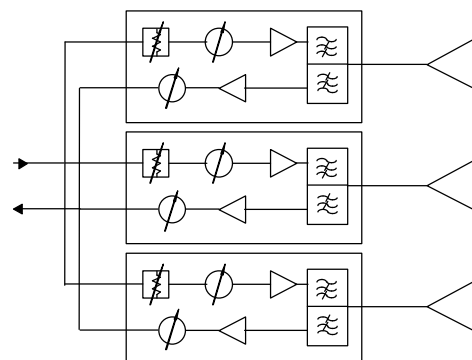


Figure 5.4 – Block diagram of an active array antenna

constituted by direct solar radiation, with a power flux density of  $1.2\text{kW/m}^2$ , which needs to be reflected back or radiated in addition to the thermal flux generated by the TR modules, while thermally reflective materials are by necessity poor thermal radiators.



The active antenna of the Envisat SAR instrument (courtesy of EADS-Astrium)

In the field of SAR antennas, the use of active arrays is favoured by the fact that the special geometry of the antenna results in very long paths for the feeding network and consequently very high losses, so that the inefficiency of the solid state RF power amplifiers is compensated by the elimination of the high-power distribution network. Unfortunately, there are still some open problems related to the gain, phase and time delay stability of the active elements with the variation of temperature, which combine with those related to the phase stability of the low-power feeding lines. By measuring the phase difference between the input and output of the TR modules, it is possible to

correct their errors by adjusting the phase shifters. Attenuators may instead be used to correct variations of gain between adjacent TR modules. Alternatively, it is possible to use a special calibration mode to sequentially measure the relative amplitude and phase of the transfer functions of each active path (feeding line, module and radiating element), compare these data with the nominal values and calibrate the TR modules accordingly.

In the space sector, a distinction is generally drawn between active arrays, in which the TR modules are located on the rear of the radiating surface, and semi-active ones, in which the modules are located on or in the satellite and are connected to the antenna by transmission lines, generally very low-loss coaxial cables. This second solution has some mechanical advantages owing to the lower weight of the antenna, some thermal advantages, since the modules can be located in a more suitable thermal environment, and some manufacturing advantages, since the radiating panel of the antenna does not have to carry the TR modules mass or accommodate for their thermal stability and it is therefore simpler to design and build.

Coupling a passive and an active array back to back, it is in principle possible to produce active lenses, in which the desired amplitude and phase changes between the input and output surfaces are obtained by properly adjusting the phase and amplitude settings of each TR module. Obviously, a passive fixed-beam version is also possible, in which the phase shift function is carried out with different line lengths.

Beyond the “active lens” solution described above, there are other two types of “special” arrays: “smart skins” and “reflectarrays”. The first ones are active arrays in which microwave circuits and digital control electronic are integrated with the radiating surface, by etching on a same semi-conductor wafer or by integration of the antenna elements, MMIC and ASIC on the same substrate, forming a thin structure that is also easily adapted to curved surfaces, for example the fuselage of an aircraft. Reflectarrays are passive or active arrays, sometimes with phase shifters or reactive elements only, which are illuminated by an external source and re-radiate the RF power with a different phase and/or amplitude distribution with respect to the incident wave, so that the beam can be focused and shaped although the antenna is flat. A typical use proposed for space is to replace a parabolic reflector with a planar reflectarray, easier to manufacture and

accommodate, which has the same focusing power, but also higher losses and a much smaller operational bandwidth.

Passive (fixed) reflectarrays can be easily realised with elements located on a regular grid and with different dimensions, so that they exhibit a capacitive or inductive behaviour at the operating frequency and the signal is consequently phase shifted. Such an approach may be seen as an adaptation of the dichroic structures examined before. Using the same principle, it is also possible to produce planar lenses, in practice forming an artificial dielectric with non-uniform dielectric constant.

### 5.7 Measurement of the radiation pattern of large antennas

The verification of the radiation pattern of SAR antennas encounters a problem common to all large antennas, namely the great distance between the transmitting and receiving antennas required to perform the measurements in a condition sufficiently close to the operating one.

The only way of measuring the radiation pattern of an antenna is to use a second antenna, of known suitable characteristics, to probe the field radiated by it. According to the reciprocity theorem, inverting the functions of transmitter and receiver does not change the measured results and this is the solution adopted in the majority of cases, with the test antenna transmitting and the antenna under test receiving. The most obvious exceptions are active antennas, for which the measurement must clearly be made in both modes, unless there is a way to bypass the TR modules and test them separately.

The field radiated by any source has characteristics that vary with the distance from it (figure 5.5). In the vicinity of the source, the field has a distribution that reproduces that of

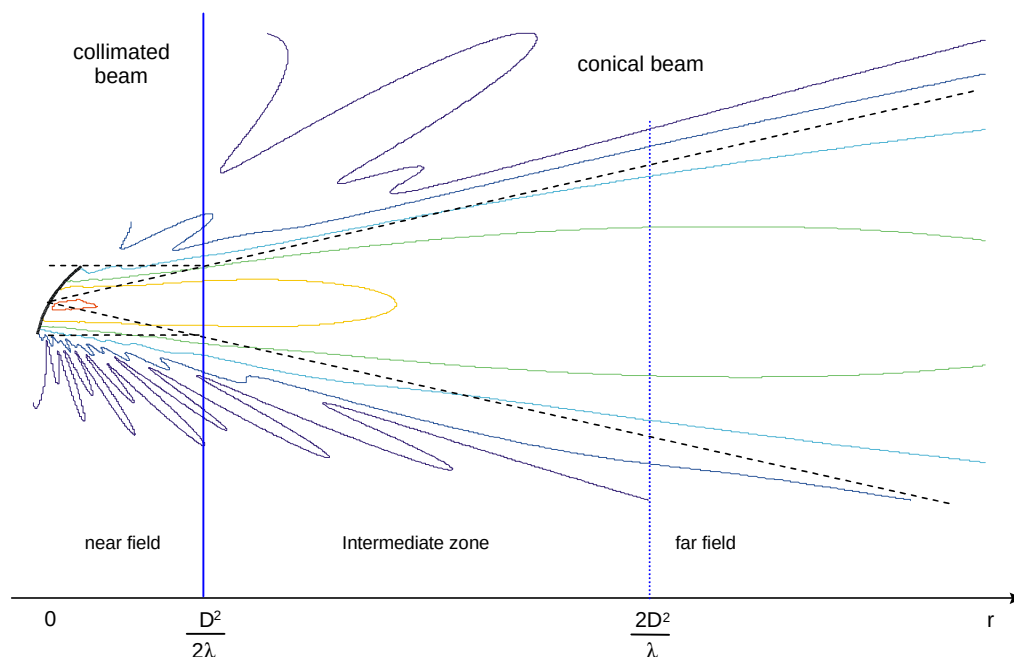


Figure 5.5 - Structural modification in the radiated field distribution with distance

the real or displacement currents generating it. Moving along a straight line perpendicular to the aperture plane, the distribution does not vary significantly with distance, while the field amplitude oscillates around a fixed value. In this region of space, called the near field or Fresnel region, all the field components generally have non-zero values. Moving away from the antenna, there is an intermediate region in which the field distribution gradually changes its structure as the distance from the source increases. Finally, at even greater distances the field tends to assume a stable distribution with respect to a spherical system of co-ordinates centred on the antenna; i.e., its distribution on a sphere centred on the antenna is fixed except for an amplitude factor, which takes into account the fact that the radiated power is distributed over a surface whose area increases with the square of distance, and a phase variation linked to the distance. The latter region, together with the preceding one, is called the Fraunhofer region. At large distances from the source, within the third region, the radiated field has the local characteristics of a plane wave, therefore only the components perpendicular to the direction of propagation exist, so the last region, which extends to infinity, is called the far field region.

It can be concluded from the above that in order to obtain correct information on the distribution of the radiated field at a long distance it is necessary to measure it at positions falling within the far field region. An approximate formula, known as the far field condition, is frequently used to determine the distance  $d$  beyond, which it may be assumed that the condition of absence of components in the direction of propagation is satisfied:

$$d = \frac{2D^2}{l}$$

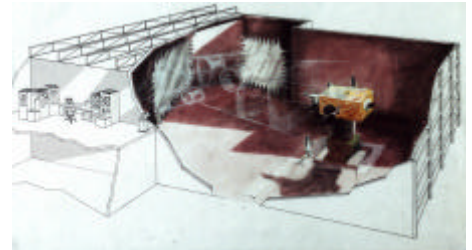
where  $D$  is the largest dimension of the antenna and  $l$  is the wavelength. This simple formula is based on considerations about the maximum path differences path between the observation point and different points of the antenna, which are also related to the maximum phase difference between the field distributions at the measurement distance and at infinity. The difference in direction of the various paths is instead related to the variation of the field in the radial direction with distance and therefore to the level of the residual radial components.

For SAR antennas, which have lengths of the order of 10 metres and typically operate at frequencies between 1 and 10GHz,  $d$  is of the order of a few kilometres, making it practically impossible to directly measure the far-field radiation pattern of the antenna, at least while the antenna is on ground.

By using the equivalence principle, however, it is possible to reconstruct the radiated field at large distance from measurements made at short distance from the antenna (near field). The amplitude and phase of the field are measured on a closed or almost closed surface, i.e. one on which the field is very low in the missing portion and can therefore be assumed to be null in the subsequent processing. The radiated field at long distance is then calculated as the radiation of a distribution of sources equivalent to the field distribution measured on the surface. The calculation involved in this transformation are quite delicate, since the integral relation existing between samples measured in the near field and the distribution in the far-field causes a small relative error in the samples of larger amplitude to give rise to large errors in low-field regions at large distance, such as nulls and far out sidelobes as well as on the orthogonal polarisation.

The large number of samples and the accuracy required in the measurement of the field and in the position of the sampling points to ensure sufficient accuracy result in very long measurement times. The subsequent processing is also laborious, making these measurements difficult to carry out.

A possible alternative is to reconstruct the distant field conditions at shorter distances from the antenna. As well known, at a long distance from the source a wave is locally planar. Then applying the reciprocity theorem to a plane wave with an inverse direction of propagation it is evident that the source will react to it receiving a power proportional to the percentage of power that would be radiated by it in the same direction. The way in which the plane wave is obtained has no effect. Additionally, by applying the reciprocity theorem again, it is evident that the signal received from any system capable of generating a plane wave at finite distances will still be proportional to the percentage of the total power radiated by the source in the wave direction.



Sketch of the ESA Compact Payload  
Test Range at ESTEC, Noordwijk  
(The Netherlands)

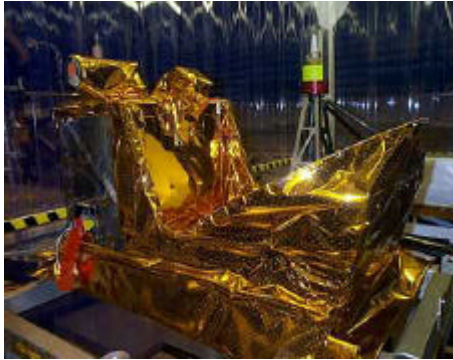
The noteworthy aspect is that the system required to generate a plane wave in the near field region is a simple parabolic antenna with a point source. Unfortunately, the diffraction from the edges of the reflector and from everything in its vicinity considerably perturbs the planarity and uniformity of the wave. Nevertheless the approach is valid and is widely used, applying special measures to minimise the scattering from the reflector edges and surrounding environment.

Antenna measurement systems based on the latter approach are known as "compact ranges". One or more reflectors are used to generate a locally planar wave within a volume, called the quiet zone. The dimensions of the chamber, and of the reflectors, must be at least twice the size of this volume in all directions, so as to minimise the effects of residual reflections on the RF absorbing materials covering all the surfaces. To measure the radiation pattern of antennas larger than 3-4 metres it is necessary to use chambers of such size that their cost would not be justifiable given the very small number of such antennas produced. The main purpose of such large measurement chambers is in fact to allow the testing of complete systems, for example whole satellites, in a protected and clean environment.

### **5.8 Radiometers**

Radiometers are totally different instruments from the previous ones. They are completely passive: in other words, they only receive, and their sensitivity needs to be extremely high. A radiometer measures the RF power emitted as part of the thermally induced radiation from objects placed within its visibility sector. All the energy reaching the antenna from other sources, especially very hot ones such as the Sun, contributes to a decrease in the sensitivity and accuracy of the instrument.

Spatial resolution is therefore an important parameter, but beam efficiency is even more important. It is given by the ratio between the integral of the power flow extended to the major lobe and its integral extended to the whole sphere and measures the ability of the antenna to discriminate between the "signal" coming from the scene and the "noise" coming from all sources. For radiometer antennas, the shape of the major lobe is of



The flight model of the Envisat radiometer

secondary importance, provided that it is known, whereas it is essential that radiation outside the major lobe be at very low levels, to minimize the reception of unwanted energy and thus reduce the noise level at the receiver input. Generally the radiometers used for space applications are very sensitive, i.e. capable of measuring very low noise levels and extremely small variations of the noise power, equivalent to temperature variations of the order of 0.2 to 0.5 Kelvin. Obviously, this makes it necessary to protect the sensitive element (generally a bolometer) from excessively strong fields, i.e. by preventing the antenna from receiving large unwanted signals from the rest of the satellite or other sources.

The noise originating externally is beyond the control of the engineer, who can only attempt to reduce the sensitivity of the instrument to this noise, by using screens, for example. On the other hand, the internally generated noise must be reduced to a minimum, and therefore it is important that the ohmic losses and the temperature of the antenna and of the whole receiving chain up to the first low-noise amplifier are kept as low as possible. To this end, the whole system is protected from direct solar radiation by thermo-optical screens, the receiver is mounted as closely as possible to the feed and it is housed in a chamber held at a very low temperature (a few Kelvin) by a cryostat. The chamber usually extends to incorporate a part of the feed, sometimes leaving outside only its aperture. Radiometers generally use reflector antennas and the reflectors have special coatings to increase their reflectivity on the exposed face and the thermal radiation capacity on the opposite face, so as to keep it at the lowest possible temperature.

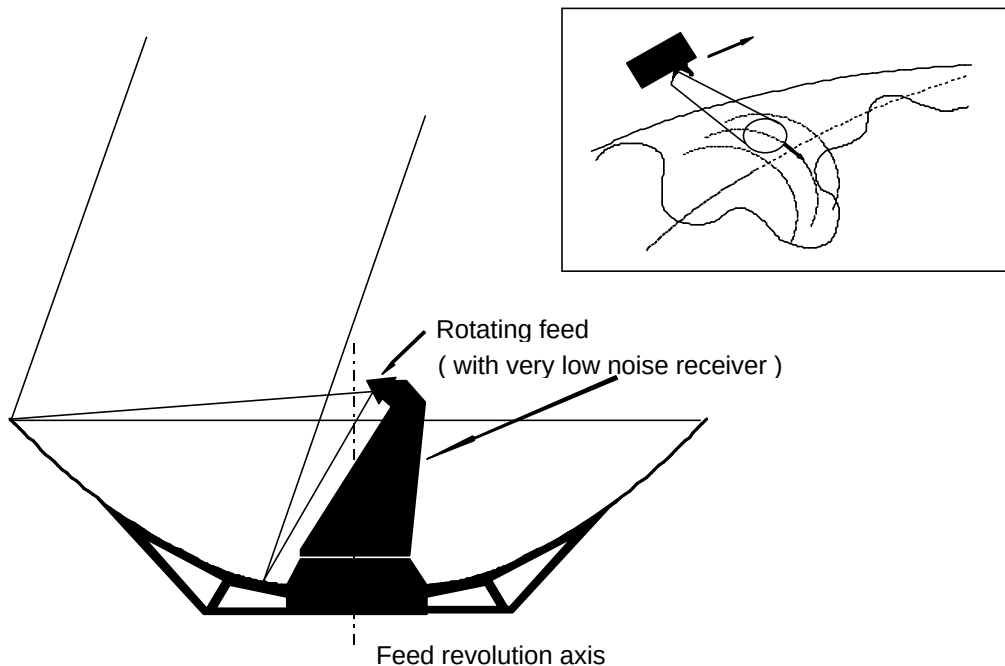


Figure 5.6 - Architecture of scanning radiometer with fixed reflector

Radiometers mounted on satellites may have fixed or moving beams. For the second type, the use of electronic scanning arrays is attractive, but the variation of the beam with scanning, together with the high noise temperatures of such antennas, discourages their use. The beam scanning is frequently carried out mechanically, moving the whole instrument, the reflector or only the radiator. When a large scanning angle is required, small arrays of radiators can be used so that scanning is obtained by switching the receiver to the different radiators. Scanning radiometers generally use reflectors whose shape is designed to reduce beam deformations to a minimum. Portions of surfaces of revolution are generally used, for example a torus with a parabolic section, and are arranged in such a way that the feeds rotate about the axis of revolution (figure 5.6).

The control of the minor lobes is one of the most critical problems in radiometer antenna design. The particularly low levels required mean that small errors of alignment or shape of the surface, due for example to thermal deformation caused when part of the antenna is directly exposed to the Sun and part is in shadow, would be sufficient to make the level of the lobes increase significantly. The only solution to this problem is to design the antenna with sufficient margins, so as to prevent that the required levels be exceeded even in the worst conditions of deformation, and at the same time protect it from direct Sun radiation with thermally insulated shrouds.

To obtain more accurate data and to increase the quantity of information obtained from each mission, use is made of multi-channel radiometers operating in multiple frequency bands. To obtain the maximum amount of information, the beam efficiency and the pointing must be virtually identical for all frequencies, often distributed over a range of tens or hundreds of GHz. Such requirement gives rise to an interesting design problem due to the inevitable difference in the relative dimension of the antenna with respect to the wavelength at the different frequencies.

In the proximity of the peak, and up to approximately 10dB below it, the angular aperture  $q$  of the main lobe is, with a good degree of approximation, inversely proportional to the ratio between the wavelength and the diameter of the antenna, i.e.

$$q = k \frac{\lambda}{D}$$

where  $k$  is a constant that depends on the level at which the beamwidth is measured, the shape of the aperture and its illumination law. For example, for uniform illumination the beamwidth at -3dB is determined by  $k=51$  for rectangular apertures and  $k=58.4$  for circular apertures. Given the Fourier transform relation existing between the illumination on the aperture and the antenna radiation pattern, a change of the illumination to reduce the level at the edges, while maintaining a constant integral over the aperture, is accompanied by a widening of the beam. Therefore, decreasing the illumination at the aperture edge (figure 5.7) increases the value of  $k$  and in a reflector antenna  $k \approx 70$ .

To obtain beams with constant angular width at different frequencies from a same aperture, it is possible to adjust its illumination. In the limit, the main lobe of the source may illuminate only small portion of it. If the distribution is smooth, for example parabolic or cosine, its details have only a small effect on the main lobe. However, because of the direct link existing between the level of the lobes and the content of higher-order harmonics of the aperture distribution, the sidelobe level is closely related to the level of illumination at the edges, to its gradient and to the ripple of the illumination across the aperture.

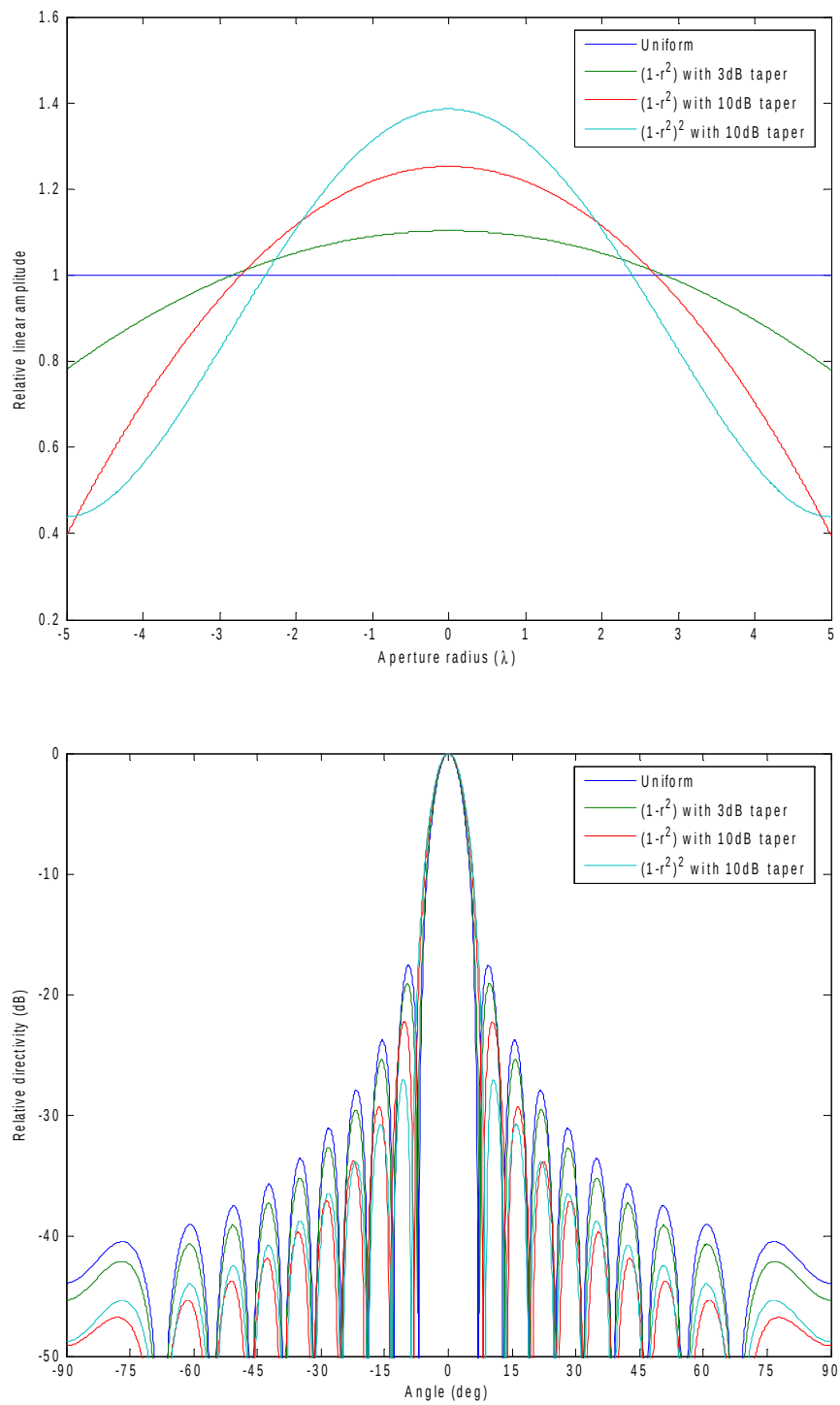


Figure 5.7 - Effect of illumination law changes on the radiation pattern of a  $10\lambda$  circular aperture

A synthesis of first approximation of the illumination required to obtain the desired beam may be achieved by using orthogonal basis functions, like harmonic series or polynomials, to expand the illumination of an equivalent flat aperture or of the currents on the reflector. As seen previously, the lack of information on the phase of the far field distribution creates problems, which are not easy to solve.



Artist view of the synthetic aperture radiometer on the SMOS satellite

### 5.9 Synthetic aperture radiometer antennas

A rather new concept in the world of radiometry is the Synthetic Aperture Radiometer. The basic idea of these systems is to increase the spatial resolution by using all possible combinations of the signals received from a scene by a set of antennas arranged in a suitable way. For example, considering a satellite in geostationary orbit and a frequency of 50GHz a resolution of 40Km requires an antenna aperture of about 4m ( $\sim 667\lambda$ ), which is not easy to accommodate on a satellite, while a number of small ones placed over the same geometrical area is.

A sparse 1D or 2D array is used for these instruments. The signals received from each element, possibly in two polarisations, are fed to low-noise amplifiers located on the back of the element itself and then routed to the signal processor. Combining the in-phase and quadrature components (I and Q signals) of the signals coming from pairs of elements with different distances it is possible to synthesize the signal that would be received by elements placed in a different position (figure 5.8). This interferometric process allows the synthesis of element positions obtained by subtracting the two signals with both orders, which correspond to positive and negative phase shifts with respect to a reference positions. In this way it is possible to fill missing positions within the element distribution as well as to create virtual elements in positions placed outside it, i.e. on the mirror image of the real distribution with respect to the reference element. The antenna is therefore constituted by an array of elements arranged over a grid (or line) with several empty positions.

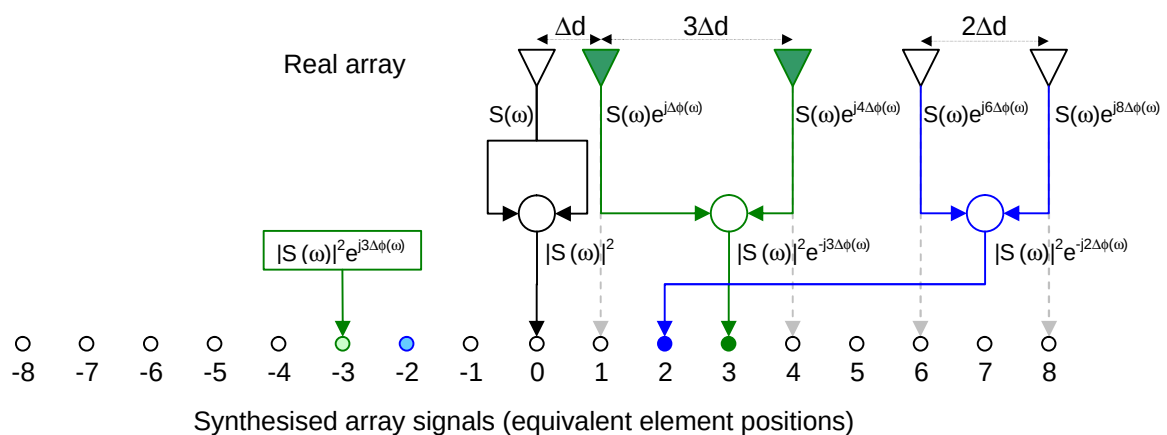


Figure 5.8 – Working principle of a Synthetic Aperture Radiometer



an increase in frequency, from transparency to opacity. The skin effect confines the currents to very thin layers (a few  $\mu\text{m}$ ), and therefore the intrinsic roughness of materials, related to their physical structure, has a considerable effect on conductivity. Finally, in all semiconductor devices the wavelength is comparable to their dimension and, in particular, to that of their active region. For example, the air-bridges commonly used in electronic packaging between a component and the leads may easily radiate.

Nevertheless there are also some advantages. The most significant one for antenna design is that at these frequencies it is easy to integrate active elements and antennas on the same substrate, thanks to their comparable size. This simplifies the manufacturing of integrated receivers, but it also raises some interesting design problems, as shown for example by the variation of the radiation pattern of a resonating dipole formed on a dielectric substrate (figure 5.9). The behaviour would be the same at lower frequencies, but the substrate thickness in wavelength unit and the losses would be much smaller, so that a ground plane below the substrate would result in a working antenna instead that in a very lossy one.

Recently, to reduce the losses in dielectric substrates, which are typically very high at these frequencies, there has been considerable interest in the use of special dielectric substrates, called photonic band-gap materials, behaving quite similarly to dichroic structures. The basic idea is to use an artificial dielectric having periodic variations of its permittivity to obtain a resonant or anti-resonant behaviour of the bulk material, rather than of a planar structure. At optical frequencies, some materials have these characteristics as a result of their molecular or crystal structure, while in the case of millimetre waves it is normally necessary to obtain the desired periodicity artificially, e.g. by means of regularly spaced holes or a matrix of fibres, in two or three dimensions. The main application of these materials is the control of the radiation pattern of the antennas formed on them; for example, it is evident that, if the periodicity of the material is such that it becomes reflecting at the operating frequency of the antenna, it will act as a ground plane in spite of its dielectric nature. Conversely, if it is transparent, it will tend to "absorb" the whole field generated by the antenna.

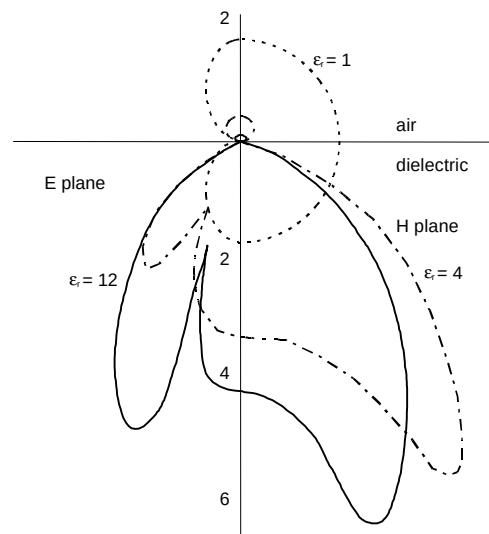


Figure 5.9 - Effect of the dielectric substrate on the radiation pattern of a printed dipole operating at 250GHz

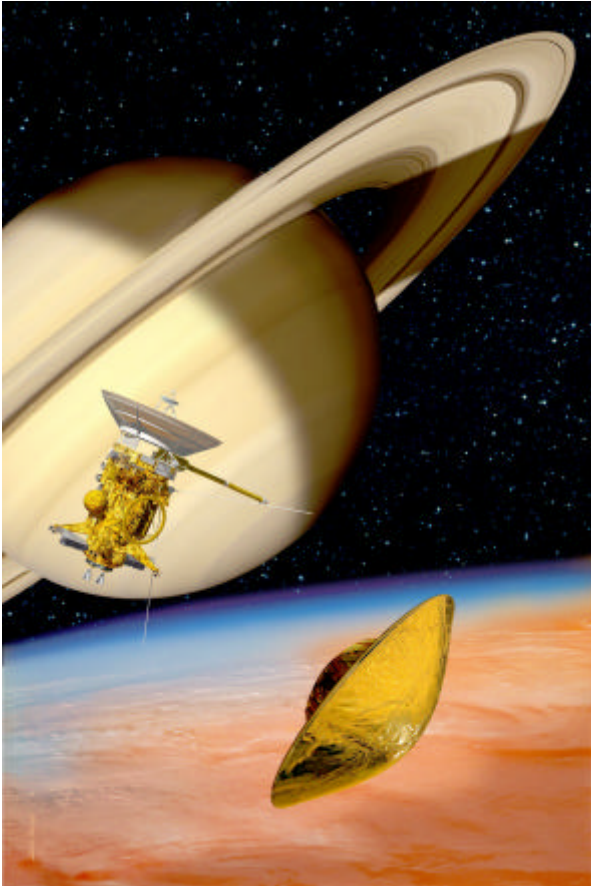
## References

- [1] <http://www.eumetsat.int/>



## CHAPTER 6

# *Other applications*



The Cassini satellite with the Huygens probe (artist impression)

There are several other systems on board satellites, space probes and space vehicles that use antennas. However, since the way in which they are used and designed is not radically different from the ones described so far, this chapter will only provide a brief description of these applications. The major exception to this is constituted by the telemetry, tracking and control antennas, which have some special characteristics.

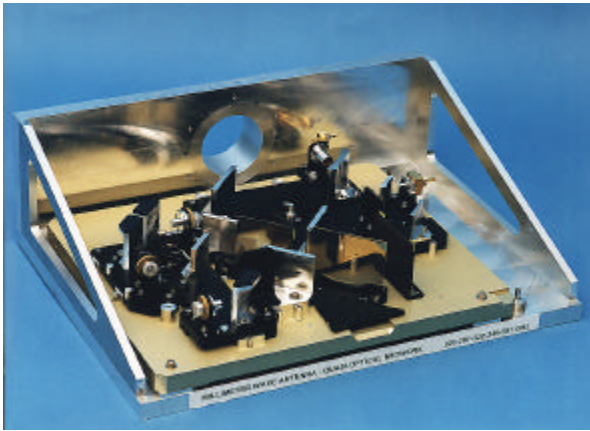
### **6.1 Scientific applications**

The use of antennas for scientific instruments is quite differentiated, from ultra-low frequency antennas to detect waves in interplanetary space to radio-telescopes operating at frequencies up to 3THz. Each scientific mission has its special characteristics, and it is therefore difficult to summarise the problems involved in designing antennas for these applications. As a general outline, it may be said that there are three main areas: radio-telescopes, planetary probes and probes for particular experiments.

Antennas for radio-telescopes and planetary probes are quite similar to radiometers. These instruments are used to measure the power of the waves arriving from deep space or radiated

from a planet surface. The basic parameters of their antennas are therefore the same of radiometers: beam efficiency, angular resolution and sidelobe level.

In the case of radio telescopes, the primary aim is to benefit from the lower level of thermal and artificial radiation present in orbit and to be able to operate in frequency bands at which atmospheric absorption is too high. The angular resolution is the main parameter for radio telescope antennas, which have considerable apertures (antennas measuring up to ten metres and beyond are under investigation) and are very complex in terms of mechanical and technological design. The radiation captured from the sidelobes is usually called “stray-light” with a term typical of optical telescopes and in some applications, e.g. probing the outer limit of the known universe or measuring the fluctuations of the cosmic microwave background, sidelobe levels 90dB below the peak



Quasi-optical beam multiplexer for radiometry applications (courtesy of EADS-Astrium)

are not unusual. To meet such requirements the antenna is surrounded by carefully designed screens to further reduce the power received from unwanted direction by a low sidelobe reflector antenna, isolating it from the thermal radiation of the satellite, of the Earth and of the Sun.

Planetary probes use antennas of various kinds, from monopoles to reflector antennas and arrays, according to the specific application, the characteristics of the planetary atmosphere and the frequency range used.

An impressive component of some space radiometers operating at very high frequencies is the quasi-optical beam multiplexer. Their

antennas operate over a number of frequency bands spread over several decades and having to provide overlapped beams at all these frequencies, with very low losses (noise level) and high polarisation purity. The only way to combine the beams generated by separate feeds, usually corrugate horns, while maintaining high performances, is to use a chain of mirrors, polarisers, dichroics and beam splitters forming rather intricate structures that owing to the very high frequency need to be designed and built to have very high accuracy (figure 6.1).

## 6.2 Navigation

The use of navigation signals is today commonplace. Navigation satellite antennas operate in L band and C band in circular polarisation and have the peculiar characteristic of having extremely flat and stable phase pattern and very stable group delay. Both these requirements derive directly from the way in which a navigation system operates. A time

reference signal is used to determine the position of the user from the knowledge of the time of arrival and of the orbital position of three or more beacon satellites. Any phase and group delay variation are seen as a change in the time of arrival and therefore induce an error in the position determination.

The antenna needs to cover the visible Earth surface and to have a beam shaped to compensate the increase of free space attenuation toward the horizon. In the case of the European Galileo system, orbiting at about 24000km of altitude, the beamwidth is about  $24^\circ$  and the gain required at the coverage edge is about 2.5dB higher than that at the centre, i.e. in the nadir direction. A relatively small array

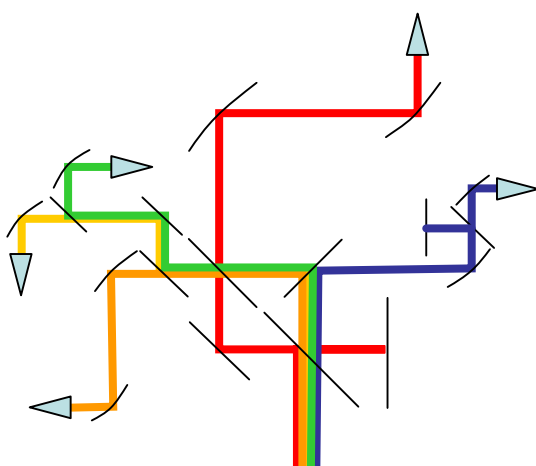
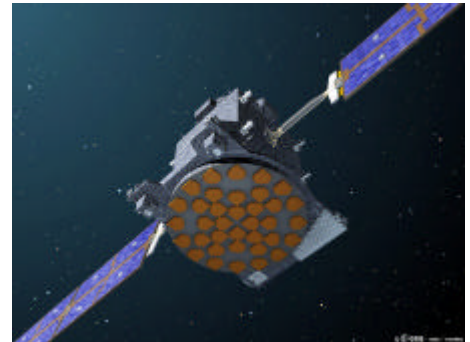


Figure 6.1 – Layout example of a quasi-optical beam multiplexer

is sufficient to fulfil these requirements. To obtain a virtually constant phase within the coverage, i.e. a purely real circularly symmetric pattern distribution, it is necessary that the aperture distribution be rotationally symmetric and real. Therefore to obtain a higher gain at the edge of coverage it is necessary to produce an out-of-phase contribution in the central part of the coverage, which is produced by feeding the external ring of array elements with an inverted signal with an amplitude about 1/10 of the one feeding the central elements of the array.

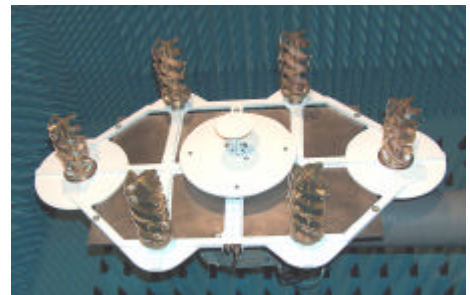
The challenging aspect of these antennas is the fact that the phase and time delay stability requirements need to be met over the full operational temperature range (about 250°C, typically), imposing severe restrictions on the characteristics of the beam forming network, which have to use very wideband components and be carefully balanced to minimise the spreading among the paths to different elements.



Artist impression of GIOVE A, the first experimental satellite of the Galileo European navigation system

### 6.3 Search and Rescue

The antennas for the worldwide Search and Rescue system receive distress signals from the emergency beacon, typically attached to buoyancy aids, in the UHF band (around 406MHz) and transmit them to the surveillance centres in L band (around 1545MHz), using circular polarisation for both links. The main characteristic of the UHF antennas is that they have to provide about 12dB gain at the coverage edge (i.e. at the horizon) while being as compact and lightweight as possible. This leads to the use of small arrays (6-7 elements) of high efficiency radiators, e.g. helices.



The UHF and L-band Search and Rescue antenna of the Galileo satellites (courtesy of Rymssa)

### 6.4 Data relay

All scientific missions require antennas to transmit the measured data to ground. Obviously their gain and complexity depends on the mission. Interplanetary probes required to operate at a very long distance, for example to explore the outer planets of the Solar System, require high-gain antennas and use large reflectors. Missions operating near the Earth's surface, on the other hand, need antennas that maximise the coverage of the visible surface, so as to reduce the number of ground stations required to receive the data. At the same time data need to be stored on board, in between successive downloads to ground.

An alternative possibility for low earth orbit satellites is to relay data to ground via another satellite. This way of operating has appeared with the proliferation of remote sensing (e.g. military) and scientific satellites and to support long-duration manned missions. The basic idea is to use geostationary satellites to provide an easier link with a spacecraft in low orbit, which moves very rapidly with respect to the ground (figure 6.1).

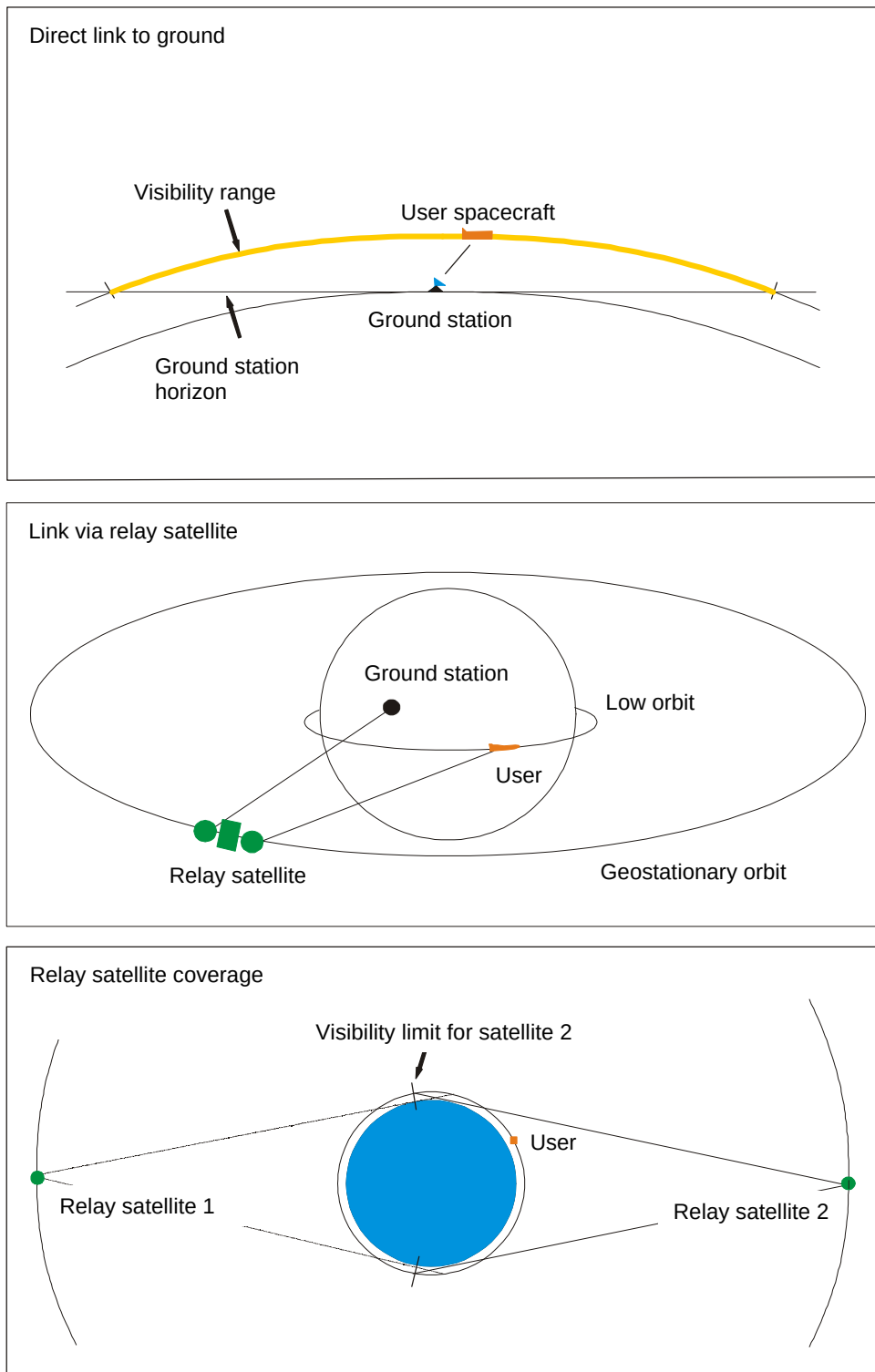


Figure 6.1 - Direct transmission to Earth and relay via geostationary satellite

There are two fundamental reasons for the use of data relay systems. The first, of an operational nature, arises from the fact that most of the users mentioned above orbit the Earth at heights between 300 and 1000Km, and therefore complete one orbit in 80-110 minutes. The transit time in the region of visibility from a ground station is therefore very short and the data gathered during each orbit must be stored on board for hours and transferred to earth in a brief lapse of time. The use of multiple stations may alleviate the problem for satellites in equatorial orbit, but it is not very effective for polar orbits. On the contrary, the use of relay satellites enables an almost continuous transfer of data.

The second reason arises from the increase in the amount of measured data, which makes it necessary to transmit data at very high speed. The increase in information transmitted per unit of time requires, for the same channel characteristics, a proportional increase in the received power to keep reception quality constant. The power available on board satellites is very limited and the only alternative is to increase the antenna gain. Higher gain antennas are larger, heavier and have to be constantly pointed towards the receiving station during the passage of the satellite, causing considerable complications. Instead, if the transmission is towards a satellite in geostationary orbit, the visibility time from the low orbit are much longer (about half an orbit) and the transmission channel is less disturbed, so that, in spite of the greater distance, the overall balance is in favour of this solution. The satellite in low orbit transmits data to the relay satellite, which in turn sends them to ground, using high-gain antennas, benefiting from its fixed position in the sky. The relay satellite may serve a number of users at the same time, owing to the greater capacity of its link with the ground stations, with the only limitation that each user needs to be tracked independently.

The antennas on board the user satellites, in low orbit, are typically of small size. The possibility of transmitting at a high speed, or the necessity of doing so in the case of a manned spacecraft, has promoted the use of antennas with higher gain, which however require a tracking system to maintain the link with the relay satellite. Such choice adds significant complexity to these antennas, but it has to be weighted against the need of a complex ground network to route the signal received using direct transmission to ground to the control centre.

Reflector antennas or small arrays are used for these applications. Obviously, it is possible to consider the use of electronically steered array antennas, where the necessity of achieving at least hemispherical scanning would lead to the use of non-planar arrays. For example, it is possible to locate the radiating elements on a spherical cap or a cone frustum. Only a few of them are excited at a time, this leads to the use of a switching matrix or, alternatively, a beam-forming matrix can be used to achieve the same result by phase only control.

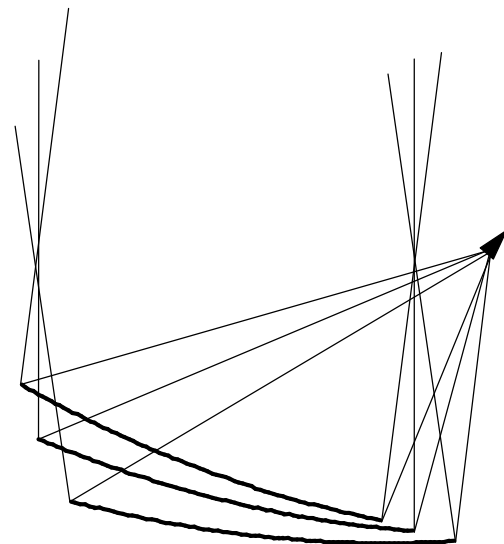


Figure 6.2 - Scanning reflector with rotation about the focus

On this subject, it should be noted that for these systems, especially the manned ones, data transmission to earth is a vital function and it is preferable to use a proven and reliable technology, possibly forgoing potentially higher performance to avoid their inevitably higher cost and risk. The use of simpler more reliable antennas results in a reduction of the total quantity of information that can be transmitted, but the alternative entails the much greater risk of losing the whole mission or reducing its scope as a result of the failure of an unreliable system.

Antennas on board relay satellites are quite alike those for mobile communications. However the antennas for the link to the user satellite have some special characteristics. In order to provide a high-performance service and because of the limitations of the user terminals, it is required to use antennas with a high gain, limited minor lobes and high polarisation purity (the polarisation is circular, for reasons similar to those mentioned for mobile systems). Unfortunately the use of high-gain antennas makes it necessary to continuously move the beam to follow the user satellite along its orbit. The scanning angle is relatively small, owing to the significant distance of the geostationary orbit from the low orbits, i.e.  $\pm 10^\circ$  for orbits up to 1000Km of altitude, and does not constitute a problem in itself. However the considerable increase of the distance between the two antennas with the increase in the scanning angle, i.e. when the user moves toward the horizon, creates an interesting design problem. Ideally, to maintain the same link budget it would be required to progressively increase the gain when the beam moves from the centre toward the edge of the coverage area. As a second best solution it is at least necessary to reduce the reduction of gain caused by the scan losses to a minimum.

A possible way to achieve this objective is to use a reflector antenna with mechanical pointing, which features simplicity of construction, proven technology and limited scan losses (less than 0.5dB). Moreover, by a suitable arrangement it is possible to use the same antenna to operate in a number of frequency bands, e.g. S and Ka band, with different bandwidth and transmission capacity, so as to meet the requirements of different user satellites. Studies carried out in preparation for a European relay satellite system (of which the Artemis satellite is the only component at present) have shown that the best way to minimise scan losses is to make the reflector rotate about its focus while keeping the feed fixed (figure 6.2). This movement causes the orientation of the beam to change with respect to the satellite, without modifying the position of the feed with respect to the focus of the reflector, so that the resulting beam remains focused and directed along the axis of the paraboloid. The only effect impacting on the antenna radiation pattern is a slight change in the edge illumination of the reflector, which has two main consequences. On one hand, there is an inevitable increase in the reflector spill-over when it is not in its centred position. On the other hand, there is an alteration of the sidelobes due to the change in the level and derivative illumination at reflector edge.

A good alternative solution would be to use an array antenna with an electronically scanned beam. The particularly interesting aspect of such a solution is that the same physical aperture can generate several beams, each of which can independently follow a different user. Obviously a number of beam-forming networks equal to the number of beams would be required. The problems entailed have already been outlined in the discussion of antennas for mobile communications systems, but here they are aggravated by the larger scan requirement, at least  $20^\circ$  against a maximum of  $17^\circ$  for a mobile system with global coverage. Among other considerations, this rules out the use of magnifying antennas, at least in the present state of knowledge of their behaviour, owing

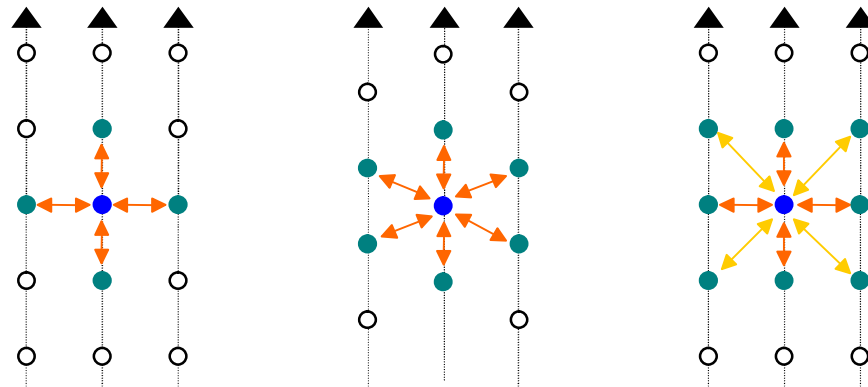


Figure 6.3 - Inter-satellite link diagrams for different constellation geometries

to their higher scanning losses, combined with the high percentage of energy that the feeds radiate outside the reflector. Furthermore, with current technology, such a solution would only be viable for S band.

It is worth noting that an alternative solution to the use of microwaves in the relay-satellite to user-satellite link, namely the use of laser transmissions, has been successfully experimented with the Artemis mission. Both the transmission speed and the immunity to interferences of these systems are very large as a result of the use of optical frequencies. However, one problem which can not be ignored arises from the necessity of keeping a beam with very small angular aperture (generally a few milliradians) aimed at a target that may be more than 45,000Km away and despite the fact that both terminals are placed on spacecrafts which attitude variations, albeit slow, are two or three orders of magnitude larger than the necessary pointing accuracy.

Finally, the application of inter-orbit links studied for satellite constellations should be mentioned. As two interacting users will seldom be connected to a same satellite, signals need to be routed between the satellites to which the users are connected. This can be done in various ways, one being the use of inter-satellite links. The increase of delay due to the use of inter-satellite links is tolerable except for voice or real-time video transmission and it is possible to conceive a system that operates as a switching network or as a packet network. Such solution provides a considerable increase in total capacity and fault-tolerance, owing to the multiple possibilities of routing the signals or packets. For example, each satellite may be linked to the four, six or eight adjacent satellites (figure 6.3) depending on the constellation geometry. Each of these links normally uses antennas with fixed pointing, operating at frequencies for which the Earth's atmosphere has a very high absorption, typically around 60GHz, so as to avoid all problems of interference with ground. Additionally, by operating at such frequencies high gains are obtained with simple, small and low cost reflector antennas, an essential advantage for systems that could include tens or hundreds of satellites.

### 6.5 Avionic antennas

The antennas used by the platform avionic sub-system, i.e. the one handling all the platform operations in any satellite, are of various kinds. Each satellite has to be provided with antennas for its control from ground, known as telemetry, tracking and command

(TT&C) antennas. Some missions also need to tag the measured data with time and position and therefore require antennas for navigation signals, e.g. the USA GPS or the European Galileo. Space vehicles and orbiting stations, whether manned or automatic, require navigation signals for approach and docking manoeuvres and for atmospheric re-entry. They also require dedicated radio links to ensure direct connection during docking manoeuvres and for extra vehicular operations performed by the crew.

The TT&C antennas of the early satellites operated in the VHF band (140-150MHz) and were simple monopoles. Their radiation pattern was strongly influenced by the satellite, which had dimensions comparable to the wavelength. Antennas in the S and X band with circular polarisation have been used afterwards in Europe. Today the use of the S band is virtually impossible, due to congestion and the use of TT&C systems operating in Ka band is being experimented. Since these antennas have the vital function of enabling satellite control from ground, their coverage must be as close as possible to the whole sphere, while the overriding importance of their correct operation makes it necessary to use simple, virtually fault-free antennas. Owing to the physical impossibility of producing omni-directional antennas and to the obstruction caused by the body of the satellite itself, at least two antennas, placed in opposite positions, are used.

A first solution is to have two antennas each covering slightly more than a hemi-sphere (e.g. having a beamwidth of  $200^\circ$ ), possibly operating in opposite circular polarisations to minimise interferences. A second typical solution is the use of one antenna with a wider beam (about  $240^\circ$ ) pointing to earth in nominal orbit conditions and a “fill-in” antenna with a relatively narrow beam ( $140^\circ$ ), which is used during manoeuvres and in emergency. Finally, especially at the higher frequencies, it is possible to use a  $140^\circ+70^\circ+140^\circ$  antenna arrangement, where the mid antenna has a toroidal beam with  $70^\circ$  degrees of aperture and it is typically located with one of the others.



The 4m Cassini antenna working at S, X, Ku and Ka bands for TT&C and other functions (courtesy of Alcatel Alenia Space)

TT&C antennas are quite simple, mainly because they must be very reliable, at S and X band helices are very often used. Other solutions are based on the combination of circular slotted waveguides with reflecting skirts to obtain the desired angular aperture of the beam. Patch antennas have also been used, mainly for the “fill-in” antennas. At Ka band waveguides are the only solution.

Scientific satellites operating far from Earth require high gain TT&C antennas, which are generally also used for the transmission of the data collected. These are mostly reflector antennas, often rigidly attached to the body of the satellite and kept pointing toward the Earth by rotating the satellite itself. Owing to the great distance and the high directivity, an incorrect manoeuvre causing the TT&C antenna to point in the wrong direction would result in the satellite loss. To reduce such risk, if the TT&C link is lost, the satellite control system automatically changes the satellite attitude placing it into a known orientation, such as to have Earth within the antenna field of view, typically determined using the Sun and some bright stars as reference, from which control can be regained.

The transmission of data to Earth has undergone the same development as TT&C systems and is now performed in X band, often using the same antennas. However dedicated antennas are used when the amount of data to be transmitted is large, as it is often the case for remote sensing satellites. It has already been mentioned that these antennas must guarantee the link for the longest possible time while the ground station is within the visibility cone. The use of wide-beam antennas is not generally sufficient, since the distance of the ground station varies with the movement of the satellite along the orbit, so that the free space attenuation of the signal varies significantly within the footprint. The variation is proportional to the ratio between the minimum distance, equal to the height of the orbit, and the maximum distance, when the station is at the horizon. Depending on the height of the orbit, this variation is 4 to 7 times, and the signal level varies by 12 to 17dB. To compensate for this variation, the antenna radiation pattern is shaped in the opposite way, with maximum gain in the direction of the horizon (figure 6.4) and a minimum in the nadir direction. This kind of beam shaping is also used for 240° TT&C antennas and for other applications in which the extension of the footprint produces large variations in the free space attenuation between the point closest to the satellite and the most distant points. For example, as mentioned before, the antennas transmitting the satellite navigation signals of both GPS and Galileo have this characteristic.

Another service sub-system that can make use of antennas is the spacecraft attitude determination and orbit control system (AOCS). Navigation signals, their carriers actually, can be used to determine the satellite orientation through interferometry. Three or four antennas, located at different points of the satellite, are used to measure the phase differences among the signals received from a number of navigation satellites. Knowing the relative position of these beacons it is possible to derive the orientation of the satellite with sufficient precision to control its attitude. Such an approach is especially interesting for small satellites, for which attitude control systems of the usual inertial type have the

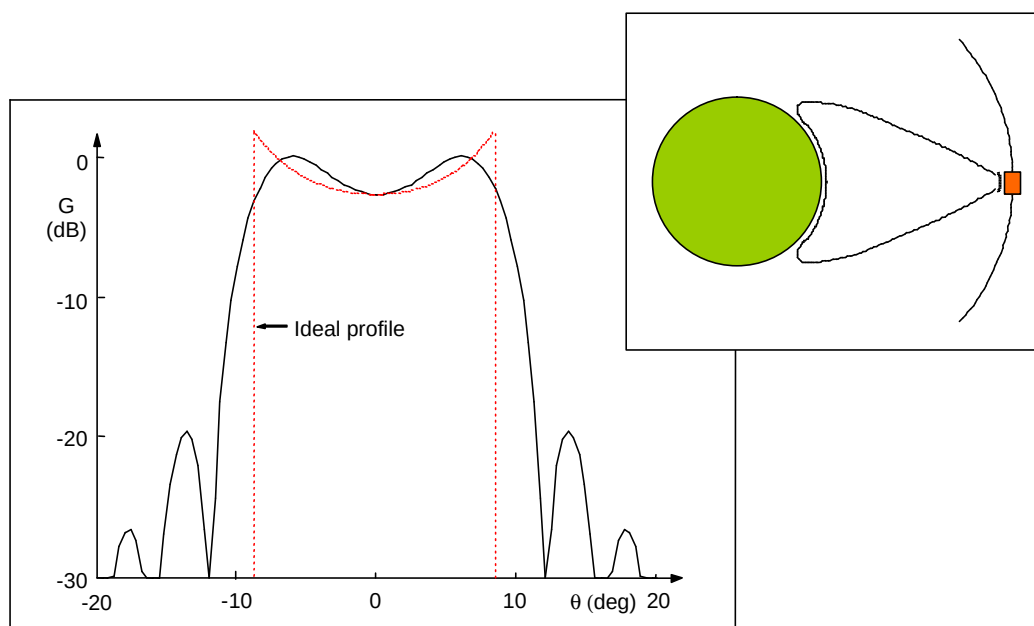


Figure 6.4 - Example of a global beam shaped to compensate the range attenuation



An X-band antenna with a shaped global beam (courtesy of Rymsa)

disadvantage of being bulky, heavy and to require a significant amount of power. The basic problem in the use of navigation signals is the accuracy, which is strongly limited by the interference due to the reflection and diffraction of the incoming wave by the structure of the satellite itself. This causes a variation of the phase with the satellite motion that is hardly distinguishable from the one due to the change of satellite orientation and therefore constitutes a systematic error. Unfortunately, this error is the dominant factor in the overall error budget and can only be minimised through careful positioning of the antennas on the satellite and by properly selecting their radiation pattern. However, since to obtain sufficient accuracy it is necessary to have four or five beacon satellite visible at any given time, the antenna beam can not be too narrow, at least with the current number of navigation satellites. Therefore development in this area continues.

## 6.6 Ground station antennas

Finally it is necessary to mention ground station antennas. From a conceptual point of view there is no difference between a terminal and a ground station: they both are at one end of a radio link. Practically the difference can be enormous. Terminals are intended for users of the satellite system, while ground stations are part of it.

Originally ground station provided for both satellite control and access to its services. Gradually these functions have been separated into control stations and access ones. Large satellite systems, like the European Galileo, actually have several types of ground stations, some devoted to the control of the satellite system and some to the control of the various services, plus those for the access to services. As communication systems become more and more devoted to direct user-to-user link the access stations tend to disappear, but there will always be the need of gateways to connect satellite systems to equivalent terrestrial ones. Furthermore the modern ground communication networks have allowed to physically separate the control centres from the location of the antennas, at least for normal operations.

Control stations, which are normally devoted to the management of several satellites, are the most complex of them and have some of the most impressive antenna systems. The stations devoted to the control of scientific probes in the outer region of the solar systems have antennas of up to 70m of diameter. They operate in the TT&C and data transmission bands, i.e. S, X and Ka band.

Beyond the sheer size of antennas, the very high transmitted power (several kW are normal) and the very high sensitivity of the cryogenic receivers, the most interesting antenna design aspect of control station technologies is the fact that these huge antennas often have to operate at several frequencies.

A first notable effect of the antenna large antenna sizes typical of ground station antennas is that their gain is limited by the surface accuracy. In other words the diffusion caused by surface irregularities, e.g. the tolerances in the alignment of different segments making up

the main reflector, produces a significant broadening of the beam. An estimate of the efficiency loss induced by surface irregularities, valid as long as these are much smaller than the wavelength  $\lambda$ , can be obtained using the Ruze formula

$$h = e^{-\left(\frac{4\pi s}{\lambda}\right)^2}$$

where  $s$  is the RMS deviation of the surface from the ideal profile. Assuming X-band operations, e.g. for the link with deep space missions around 8GHz, and a surface error of 2mm RMS, which over a reflector of 15 or 35 meters is not a minor achievement, the estimated gain loss is of the order of 2dB.

Owing to their very high gain and very narrow beams, ground station antennas need to accurately track satellites. The high power levels, the cryogenic receivers and the multi-frequency operations make it impossible to move the whole antenna, therefore a chain of mirrors is used to make it possible to point the beam while keeping the feed system fixed (or at least have limited motion, e.g. only around the vertical axis). These chains of mirrors, usually called beam waveguides, include also the means to separate the path of the different frequency bands. The beam splitting is achieved using dichroic plates and additional mirrors, as in the quasi-optical multiplexer used in radiometers. Given the very high power levels only transmission-type dichroics constituted by apertures in a thick metallic plate are used and their thermal design is still not a minor issue. Carefully designed high-power and high efficiency corrugated horns are placed in the foci of the beam waveguide system and their radiation reaches the main reflector through a series of conjugate mirror to minimise and correct optical aberrations and maximise polarisation purity. The main reflector and the sub-reflector are usually a shaped pair dedicated to maximise aperture efficiency.

A very dangerous phenomenon related to very high transmit power, beyond the obvious possibility of spark discharges, is the corona effect, i.e. the ionisation of gas within the guiding structure that leads to the creation of plasma, which damages the materials of the RF components, and to a reduction of the threshold for spark discharge.

Clearly there would be a lot more to say about these antennas to properly describe the related design problems and technology, here we will conclude on this subject observing that most of the relevant topics have been discussed in the previous sections.



The framework of the new ESA 35m antenna being built at New Norcia (Australia)





