Elements of satellite technology and communication

BY ODD GUTTEBERG

621.396.946

1 Introduction and overview

More than 35 years ago, on 4 October 1957, the world's first artificial satellite was launched by the Soviet Union. The satellite, which became known as Sputnik-I, opened up a new era of practical use of the outer space. The satellite's weight was 84 kg and it circled 1400 times around the earth before burning up in the atmosphere after 93 days. The launch of Sputnik-I was followed by the United States' Explorer-I in January 1958. Even though these satellites were primarily not intended for communications, they demonstrated that this was technically and economically feasible.

The use of artificial satellites in earth orbits is now a well established and integrated part of the world's telecommunications network. The evolution of the satellite technology together with more powerful launchers have made the satellites suitable, not only for long distance communications, but also for national communications, television broadcasting and mobile services.

A satellite communication system is divided into two major parts, the earth segment and the space segment. The satellite and its control station form the space segment, while the earth segment comprises the traffic and traffic control stations, see figure 1.

The satellite control station (TT&C station) maintains the satellite in orbit. It keeps control of the satellite status (Telemetry), orbital information (Tracking), and performs orbital attitude manoeuvres and configures the communication system (Command).

The communication part of the satellite system is simply a radio system with only one relay station, the satellite. The signals are transmitted on a carrier frequency, f_A, from an earth station (traffic station), received in the satellite, amplified, shifted in frequency to f'_A and transmitted back to the receiving traffic station. The satellite transmits and receives on radio frequencies mainly in the microwave band, i.e. 3 - 30 GHz. On these frequencies it is feasible to use parabolic reflector antennas. Such antennas concentrate the radio signals in a small cone; thus it is possible to "illuminate" the whole earth or part of the earth, see figure 2. The illuminated part is called the coverage area, and most of the transmitted power from the satellite is concentrated in this area. The larger the transmitting satellite antenna is, the smaller the coverage area becomes.

The up-link signals have only one destination, the orbiting satellite, which means that it is required to transmit energy only to the satellite. The transmitting earth station therefore normally consists of a large antenna and a high power amplifier. Since it is expensive and complicated to generate an equally high power in the satellite, the output power from the satellite is generally much smaller. Due to this, the receiving earth stations have to use larger receiving antennas; up to 30 metres in diameter, in the early days of satellite communications.

However, due to larger satellites, i.e. higher output power, and to better receiving technology, the earth stations can use smaller and cheaper antennas, e.g. satellite TV broadcasting.

2 A brief historical review

The first person to suggest the use of the geostationary orbit for communication satellites, was the English physicist Arthur C Clark (born 1917). He published his article "Extra-Terrestrial Relays" in the Wireless World Magazine in October 1945. In this article a satellite system for broadcasting of television, was described. At that time there was a discussion going on about how to distribute television. The present technology was not mature for the construction of reliable and unmanned satellites. In 1945 no-one could predict the rapid development within electronics, e.g. the invention of the transistor in 1947.

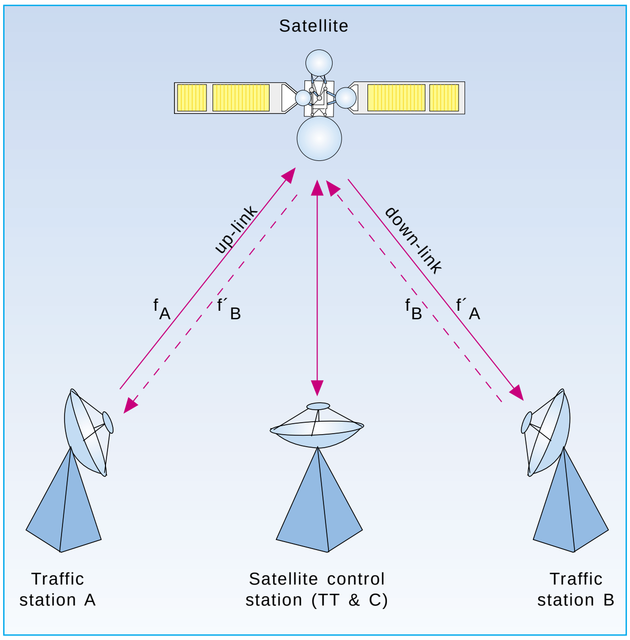


Figure 1 The main building blocks in a two-way satellite communication system

Figure 2 A geostationary satellite "illuminating" the coverage area on earth

The Space Age started with the launching of the first artificial satellite, Sputnik-I, in 1957. It is worth while noticing that the first trans-Atlantic telephone cable was stretched the same year. In the following years several satellites were launched, both for scientific purposes and for other specific applications. Communication experiments with passive reflecting satellites, Echo-I and -II (1960), were also carried out.

An operational system with passive satellites was never realised, but the project made essential contributions to the development of the earth station technology.

The first active telecommunication satellites were launched in the early 1960s, Courier (1960), Telstar-I and -II (1962 and 1963) and Relay-I and -II (1962 and 1964).

They were all put into low earth orbits, the height varying between 1,000 and 8,000 km. Telstar-I was the most important of these satellites, figure 3. This satellite transmitted television directly from USA to England and France for the first time. The Nordic countries started early to make use of satellite communication. In 1964 the Nordic countries built a common receive earth station at Råö, outside Gothenburg. It was an experimental earth station which participated in NASA's experiments with the Relay satellites.

The first satellite in the geostationary orbit was Syncom-II in 1963. The launch of Syncom-I the same year was a failure. This was an important breakthrough for commercial satellite communication. Several important experiments were carried out. One of them

Launch data

Date:	10 July 1962
Vehicle:	Thor-Delta

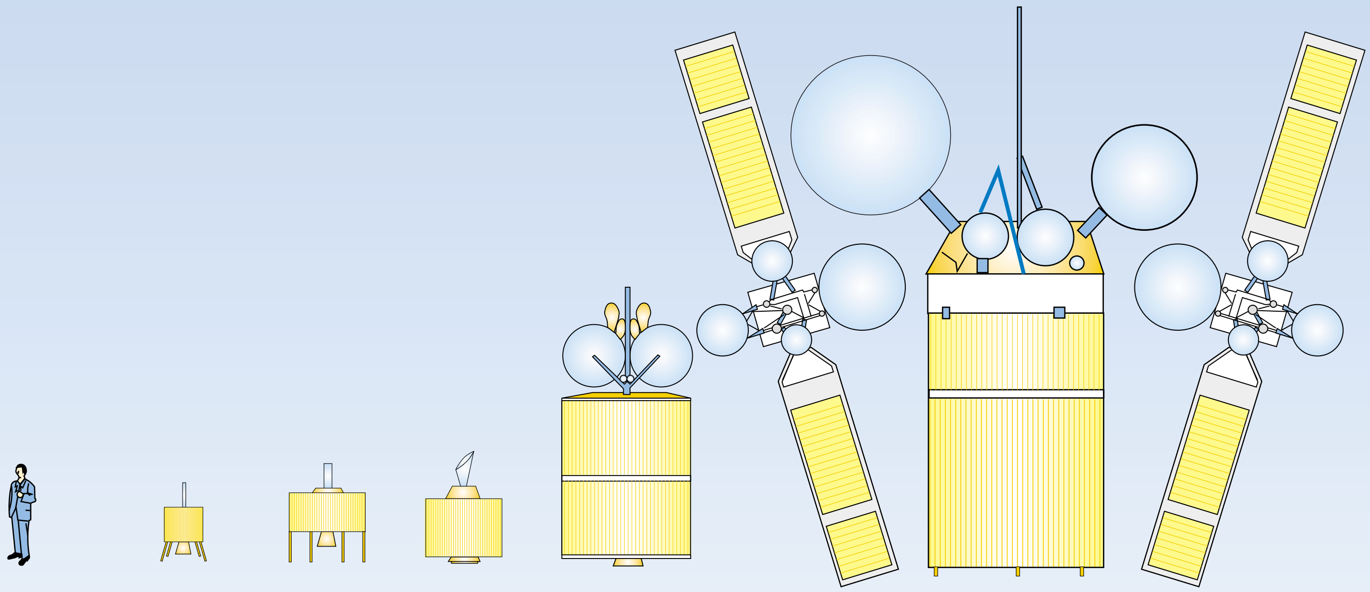
Orbital data

Orbit:	Elliptical
Inclination:	44.8 degrees
Apogee:	5653 km
Perigee:	936 km
Period:	157.8 min

Satellite data

Weight:	77 kg
Diameter:	0,9 meter (spherical)
No of transponders:	1
Bandwidth:	50 MHz
Output power:	3 Watts

Figure 3 The Telstar satellite, main data. The satellite was designed and built by Bell Telephone Laboratories and launched by NASA



	INTELSAT I (Early Bird) 1965	INTELSAT II 1967	INTELSAT III 1968	INTELSAT IV + IVA 1971	INTELSAT V 1980	INTELSAT VI 1989	INTELSAT VII 1993
Telephone channels	240	240	1 200	6 000	12 000	36 000	18 000
Weight in orbit (kg)	40	85	150	790	1 040	2 230	1 400
No. of satellites	1	3	5	7 + 5	13	6	12

Figure 4 The INTELSAT satellites. Technical development

was to examine whether or not geostationary satellites could be a part of the public telephone network. The problem was the time delay of the radio signal. With a mismatch at the other end of the connection, one would hear oneself half a second later. This was thought to be a major obstacle for a two-way telephone conversation. Effective echo cancellers were then developed. This reduced the problem to such a degree that the telephone users accepted a connection via satellite.

In 1965, the first commercial geostationary satellite, Early Bird, was placed in a position over the Atlantic Ocean. This became the first satellite in the international organisation for telecommunication satellites, INTELSAT.
A preliminary agreement was made in 1964 between 11 countries, but the final agreement was not signed until 1973. This organisation has been the leading one in the development of communication satellites. Over 100 countries are members, and there are almost 600 earth stations in 170 countries. INTELSAT has launched 35 geostationary satellites and only six have been unsuccessful.

Figure 4 shows the development of the INTELSAT system.

The first Nordic INTELSAT earth station was built in 1971 at Tanum in Bohuslän. This earth station communicates with geostationary satellites over the Atlantic and Indian Oceans. The first European commercial communication satellite, EUTELSAT-I, was launched in 1983.

3 Applications

The most important application of satellite communication has up till now been inter-continental telephone and TV transmissions. This was the first application and is still the most important. Large and expensive earth stations with antennas up to 20 m in diameter are employed. The INTELSAT-A system is an example of this.

The rapid development in the satellite technology and the use of more powerful launch vehicles has led to the use of satellite systems in more restricted areas. As an example the Norwegian Telecom employs satellites as a part of its domestic telecommunication network and offers satellite connections between the Norwegian mainland and the oil installations in the North Sea, figures 5 and 6, as well as to the Arctic islands of Svalbard (the NORSAT-A system). This system transmits via one of the satellites in the INTELSAT system.

The main station in the NORSAT system is located at Eik in the south western part of Norway. It was put into operation in 1976 and was the first earth station in Europe intended for domestic traffic. Svalbard was connected to the terrestrial public network via this system in 1979 with an earth station at Isfjord Radio, figure 7. The TV broadcasting to Svalbard was established via one of the EUTELSAT satellites in 1984.

Regional systems can make use of satellites with global, semi-global or spot coverage. This can be achieved by larger satellite antennas. The European EUTELSAT system is a satellite system with several application areas. The development of this system started in 1970. The system intended to

- transmit telephony traffic between large national earth stations
- exchange TV programmes (Eurovision)
- cover the need for special services (as business communication)
- distribute TV programmes to cable networks.

The first satellite system for maritime mobile communication, MARISAT, became operative in 1976. This system is now replaced by the INMARSAT system. The first operational European coastal earth station was put into service at Eik in Norway in 1981. This station will now expand to cover satellite communication with aeroplanes

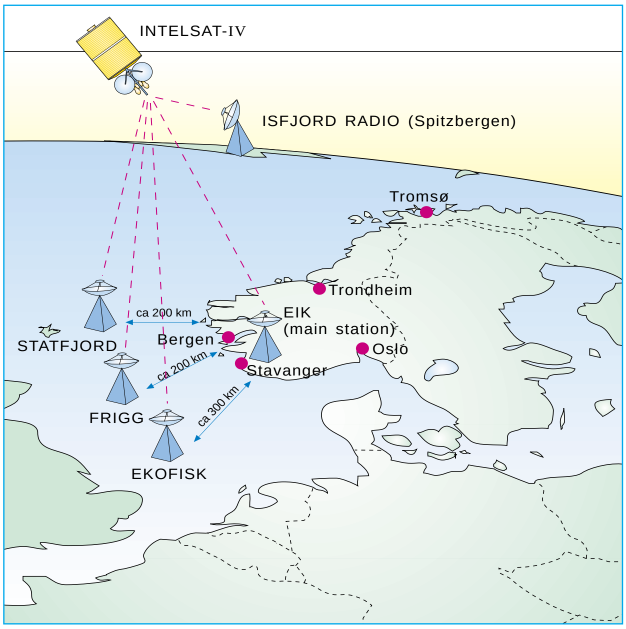


Figure 5 The NORSAT-A system started in 1975. The initial system used leased capacity on an INTELSAT-IV satellite

Figure 6 NORSAT-A earth station on an oil rig in the North Sea

Figure 7 The earth station at Isfjord Radio, Spitzbergen, 78° North. At this station the maximum elevation angle towards a geostationary satellite is 3.1°

Figure 9 Nittedal earth station outside Oslo, Norway

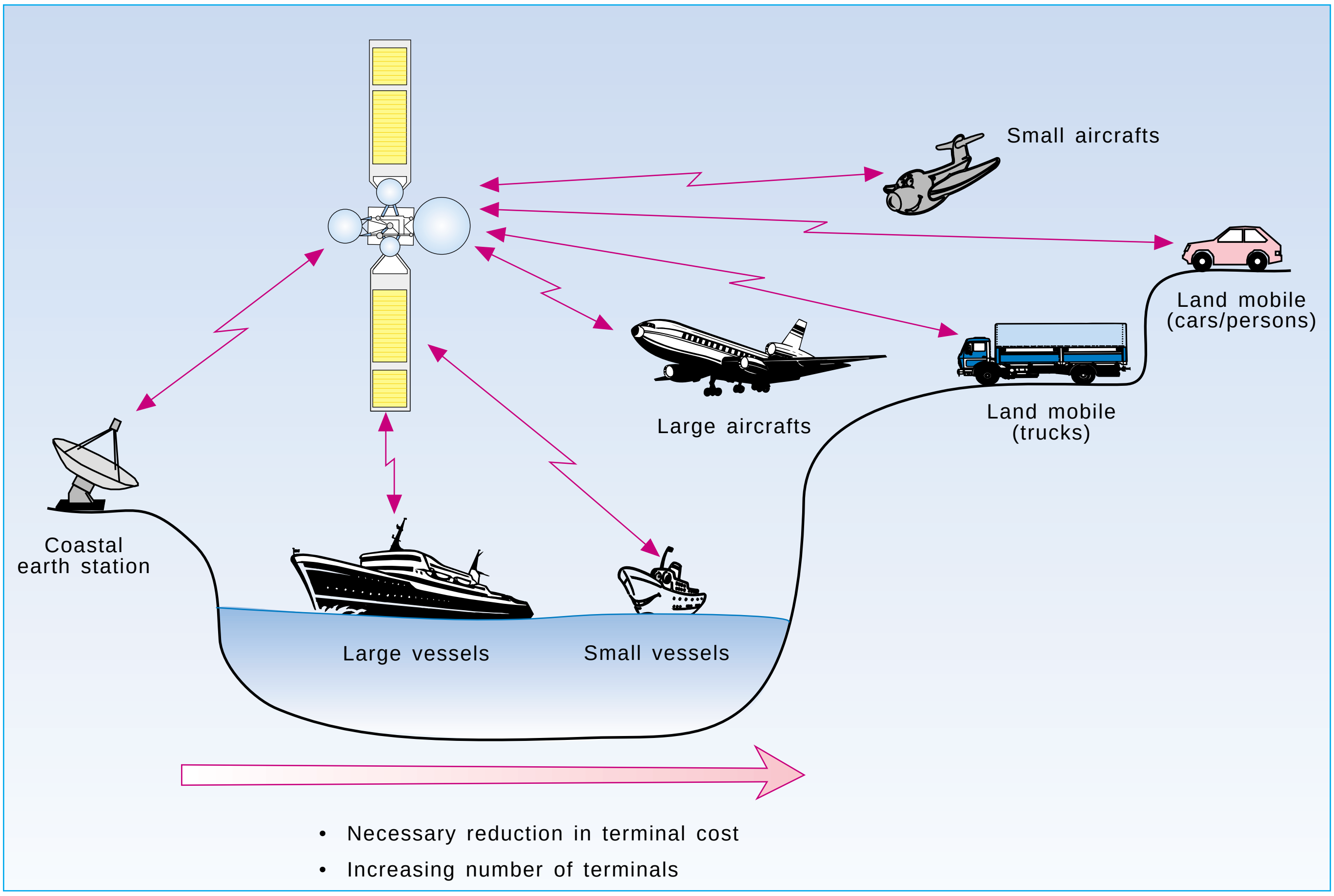


Figure 8 Mobile satellite communication system

(INMARSAT AERO). The evolution points at communication with even smaller units such as trucks, private cars and individuals, figure 8.

As the satellites become more powerful it is possible to make simpler earth station equipment. Then the satellite terminals can be located closer to the user. This is done in the VSAT systems which are mainly used for business communication. The Norwegian Telecom has established an earth station at Nittedal, 1987, for European data communication via the EUTELSAT system, figure 9, as well as a traffic control station at Eik for business communication, the NORSAT-B system. The NORSAT-B system offers connections between unmanned earth stations in Europe.

The last application area is broadcasting of television and sound. This is a point-to-multipoint system where we make full use of the advantage of satellite communication. One satellite in geostationary orbit will cover 97 % of all households in Norway. Today more than 100 television programmes are transmitted via satellite to European countries. Arthur C Clarke's original idea from 1945 has become a reality. Figure 10 shows the total involvement of the Norwegian Telecom in satellite communication.

4 Orbital considerations

A satellite's period of revolution is determined by the distance from the earth. The farther from earth, the longer the period. The nearest satellites, e.g. the space shuttle, has an orbital period of approximately 1.5 hours. An example of a satellite with a long period of revolution is the moon. Due to the long distance from earth (ca 385,000 km) the period is about 27 days.

The orientation of the orbit with respect to the earth's equatorial plane may also be different, i.e. inclined orbits.

The movements of a satellite are determined by the laws of Newton (1642 - 1727) and Kepler (1571 - 1630). Applied to satellites Kepler's laws are:

1 The satellite orbit is an ellipse with the earth at one focus.

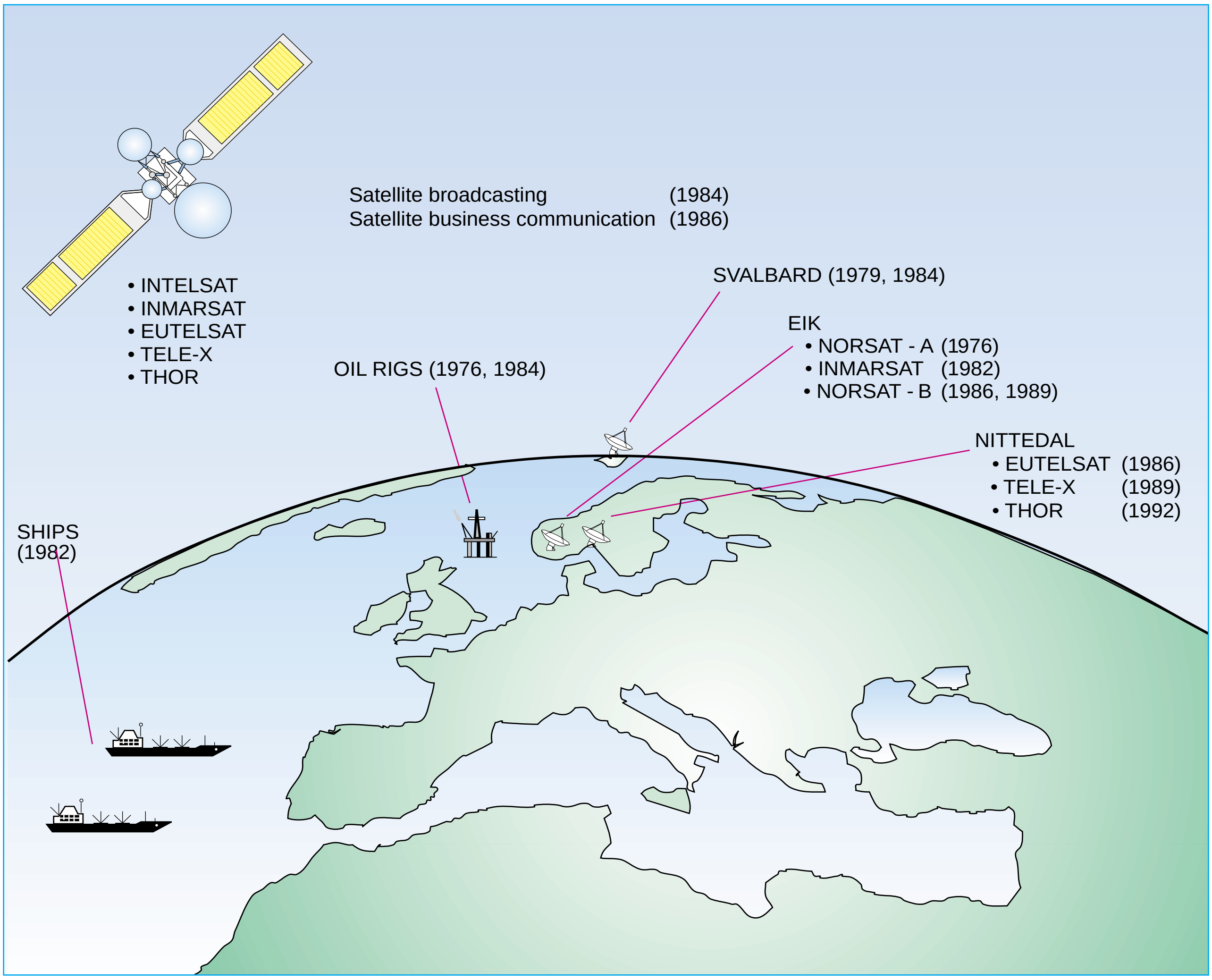


Figure 10 The Norwegian satellite involvement; main earth stations and utilized satellites

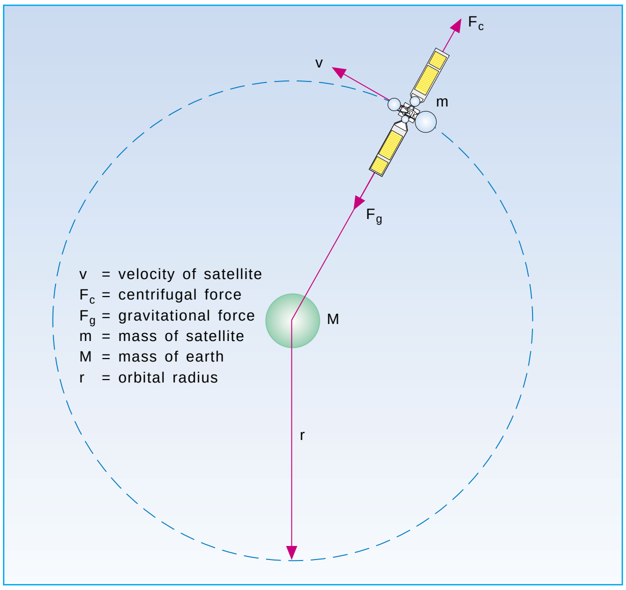


Figure 11 Forces acting on a satellite in a circular orbit

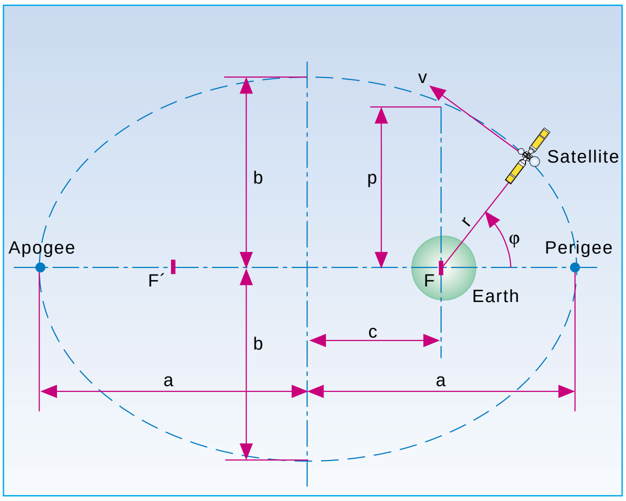


Figure 12 The elliptical satellite orbit with major and minor semiaxis a and b and eccentricity e = c/a. The semilatus rectum $p = h^2/\mu$ where $h = v \cdot r$ (the angular momentum per unit mass)

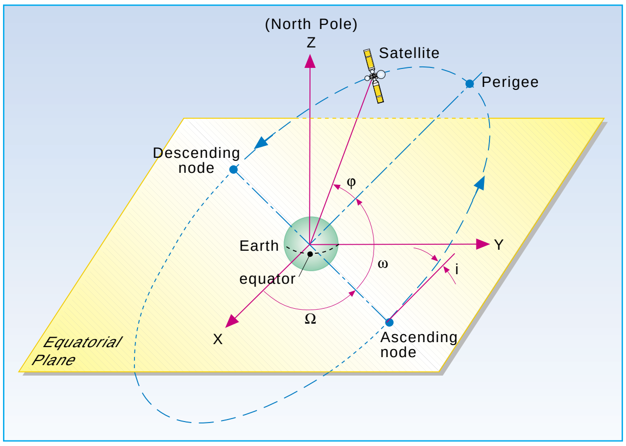


Figure 13 The geocentric equatorial coordinate system. The z-axis coincides with the rotational axis of earth. The x,y-plane cuts through the earth's equator and is called the equatorial plane. The x-axis is the direction from the centre of earth through the centre of sun at the vernal equinox (≈ 21 March), see figure 14

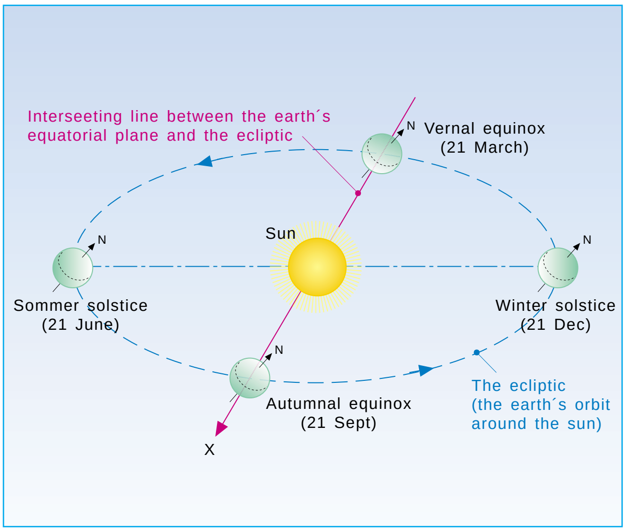


Figure 14 Definition of the x-axis in the geocentric equatorial coordinate system

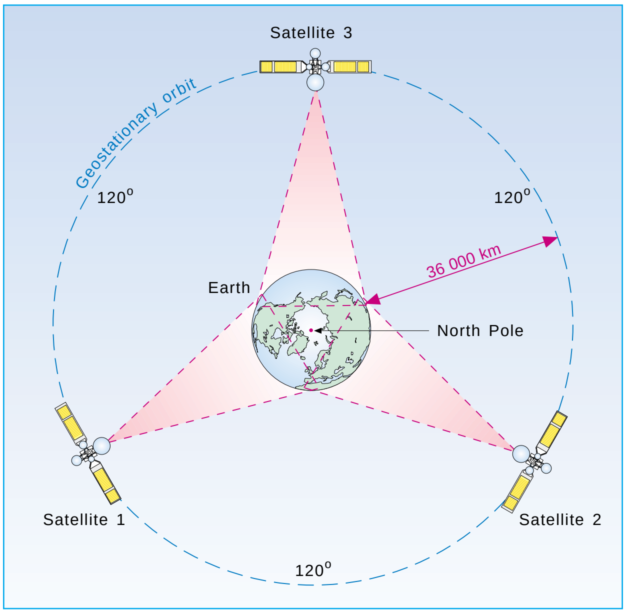


Figure 15 Three satellites in the geostationary orbit will almost cover the whole earth

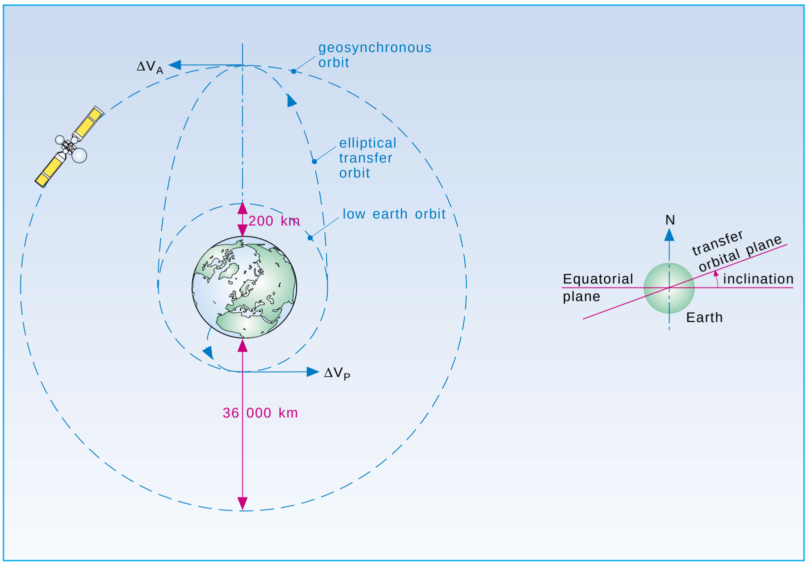


Figure 16 a) and b) Orbits for launching of geosynchronous satellite. The inclination of the transfer orbital plane is equal to or larger than the latitude of the launch site

2 The radius vector to the satellite sweeps out equal areas in equal times.

3 The square of the period of revolution is proportional to the cube of the orbit's semi-major axis.

Considering a circular orbit, the forces acting on the satellite are shown in figure 11.

With reference to figure 11, the gravitational force F_g on the satellite is given by Newton's law:

$$F_g = G\frac{M \cdot m}{r^2} \tag{4.1}$$

where G is the universal gravitational constant.

According to Newton's general law of motion, the centrifugal force is given by

$$F_c = ma = m\frac{v^2}{r} \tag{4.2}$$

Counterbalancing of forces, $F_c = F_g$, gives

$$G\frac{M \cdot m}{r^2} = m\frac{v^2}{r} \tag{4.3}$$

or the satellite velocity

$$v = \sqrt{\frac{G \cdot M}{r}} \tag{4.4}$$

The period of orbit, T, is given by

$$T = \frac{2\pi r}{v} \tag{4.5}$$

Inserting for the satellite velocity, equation (4.4), the period will be

$$T = 2\pi\sqrt{\frac{r^3}{G \cdot M}} = 2\pi\sqrt{\frac{r^3}{\mu}} \tag{4.6}$$

where $\mu = G \cdot M$ = 398 600 km^3/sek^2 (Kepler's constant).

This expression is a mathematical statement of Kepler's third law. This equation applies even if the orbit is elliptical, substituting the radius r with the semi-major axis a.

An elliptical orbital geometry is shown in figure 12.

The equation describing the orbit in polar co-ordinates (r, φ) is

$$r = \frac{p}{1 + e\cos\varphi} \tag{4.7}$$

where

p = semi-latus rectum = $\frac{h^2}{\mu}$

$e = \frac{c}{a}$ = eccentricity (o < e < l)

a = major semi-axis

$h = v \cdot r$ = the angular momentum per unit mass

The earth is located in F, one of the two foci. When the two foci coincide *(e = o)* the orbit will be a circle. The point closest to earth is called "perigee" and the farthest point is called "apogee". The angle *(ϕ)* measured from the perigee to the satellite is called the *true anomaly.*

The true anomaly, eccentricity and length of the semi-major axis determine the satellite's position in the orbital plane. In order to locate the satellite in space, we need information about the orbit's orientation. We introduce a fixed reference co-ordinate system called the geocentric equatorial co-ordinate system, figures 13 and 14. This co-ordinate system moves with the earth around the sun, but it does not rotate.

The intersection of the equatorial plane and the orbital plane is called the line of nodes. The ascending node is the point where the satellite is crossing the equatorial plane from south to north. The angle between the x-axis and ascending node is called the *right ascension (RA) of the ascending node (Ω).* The angle between the ascending node and perigee *(ω)* specifies the orientation of the ellipse, and is called the *argument of perigee.* The *inclination (i)* of the orbital plane is the angle between the equatorial plane and the orbital plane.

The five orbital elements, a, e, Ω, ω, and i, completely define the satellite's orbit around the earth. These elements are time independent. The last element, φ, specifies the satellite's position in its orbit, and is the only time dependent parameter.

For communication satellites there is one orbit which is particularly interesting. This is the near circular equatorial orbit with period of revolution equal to the rotational period of earth (24 hours). A satellite in this orbit will follow the earth's rotation, and for an observer on earth's surface, the satellite will appear to be "fixed in the sky". This orbit is called the geostationary orbit (GEO).

The geostationary satellite has a period of revolution a little under 24 hours, due to the fact that the earth moves around the sun in one year. This means an extra revolution with respect to the fixed stars. The period of a GEO satellite will accordingly be

$$T = \left(\frac{365.25}{365.25+1}\right) \cdot 24 \cdot 60 \cdot 60 = 86\ 164 \text{ sec}$$

Using equation (4.6) we can now calculate the radius of the geostationary orbit:

$$r = \left[\mu\left(\frac{T}{2\pi}\right)^2\right]^{1/3} = 42\ 164 \text{ km}$$

The satellite's height above the earth's surface will be:

$$h = r - R_j = 35\ 786 \text{ km}$$

where R_j = 6 378 km = equator radius of the earth.

The satellite's velocity is:

$$v = \frac{2\pi r}{T} = 3.075 \text{ km/sec}$$

From a satellite in the GEO orbit approximately 42 % of the earth's surface will be visible. Accordingly, with 3 satellites in the GEO orbit, we will have almost full coverage of the earth, except for areas above 81° latitude north and south, figure 15.

Our problem with the geostationary satellites is the transmission delay of the radio signal between the earth and the satellite. Since radio waves propagate with the speed of light (300 000 km/s) the time delay will be about 270 milliseconds to and from the satellite. This was in the early days considered to be a major problem for speech connections via GEO satellites. However, it showed up that the problem was much less than anticipated as mentioned previously.

In the last few years other orbits than the geostationary have been reconsidered with great interest, especially low earth orbits for mobile satellite services, and highly inclined elliptical orbits for broadcasting services, ref Hovstad/Gutteberg: "*Future systems for mobile satellite communications*" (this issue).

Figure 17 Multistage launcher Ariane

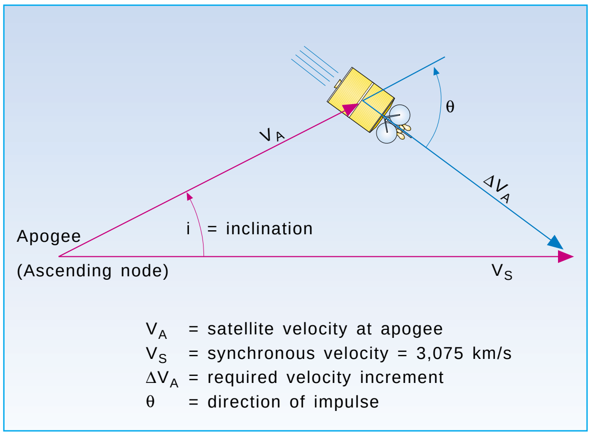


Figure 18 Velocity diagram on a plane normal to the line of nodes

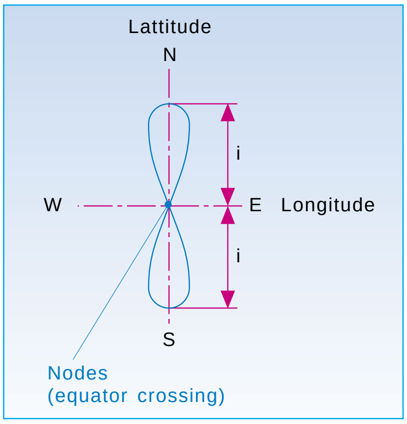


Figure 19 Apparent movement of satellite in an inclined synchronous orbit

5 Launching

For a satellite to achieve the geostationary orbit, it must be accelerated to 3.075 km/s at a height of 35 786 km with zero inclination. Theoretically, one can place the satellite directly into the geostationary orbit in one operation. However, considerations regarding costs and launching capabilities call for a multistage launch vehicle (2 - 4 stages). The launching of the satellite represents a major part of the total investments in a satellite system.

The conventional, and most economical, method of launching a satellite is based on the use of the Hohmann transfer orbit, figure 16. For the following reasons, the launch site should be near equator:

(i) To make maximum use of the surface velocity of the earth

(ii) To achieve minimum inclination. This will give minimum velocity increment for the correction of the inclination. The minimum inclination achievable is equal to the latitude of the launch site.

The launch sequence is as follows: The satellite is placed into low earth orbit with an altitude of about 200 km, e.g. by the space shuttle. There must be two velocity increments; one for injection into the elliptical transfer orbit (ΔV_p) by the perigee kick motor and one for injection into the geosynchronous orbit (ΔV_A) by the apogee kick motor. The transfer orbit has an altitude of the perigee and apogee corresponding to the transfer and geosynchronous orbit altitudes, respectively.

Alternatively, one can put the satellite directly into the elliptical transfer orbit, by a multistage launcher, e.g. Ariane, see figure 17. The idea of using a multistage launcher is to get rid of the unnecessary mass, to achieve the sufficient velocity.

Correction of the inclination has to be done at one of the nodes of the transfer orbit, in order to achieve an equatorial circular synchronous orbit. The velocity increment is depending on the inclination and on the velocity of the satellite. The lower the satellite velocity, the more economical the manoeuvre. The circularization manoeuvre should therefore be carried out at apogee. This implies that the orientation of the transfer orbit has to be such that the line between the apogee and perigee is in the equatorial plane, cfr. figure 13. The altitude of the apogee has to be equal to the altitude of GEO orbit. The magnitude (ΔV_A) and direc-

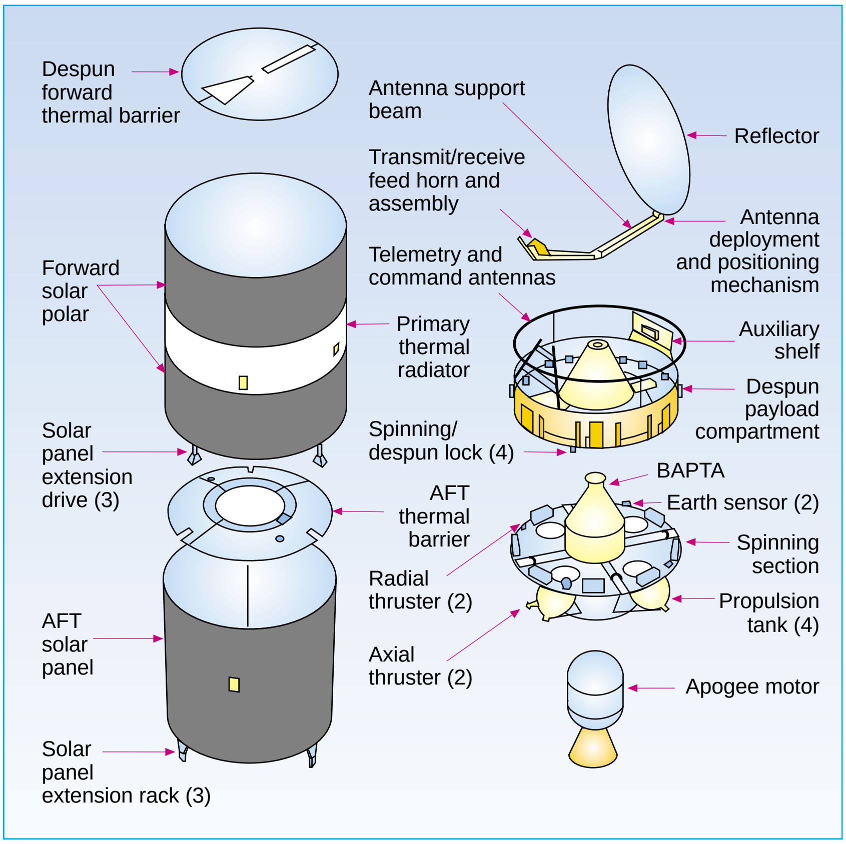


Figure 20 Exploded view of the THOR satellite

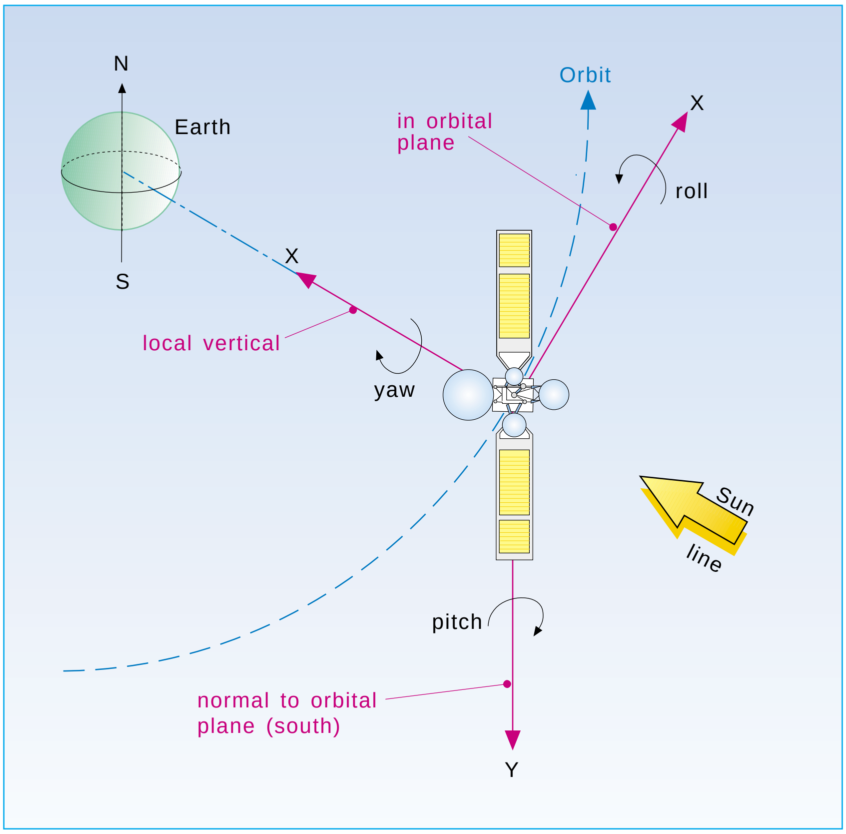


Figure 21 Definition of the reference axes, roll, pitch and yaw

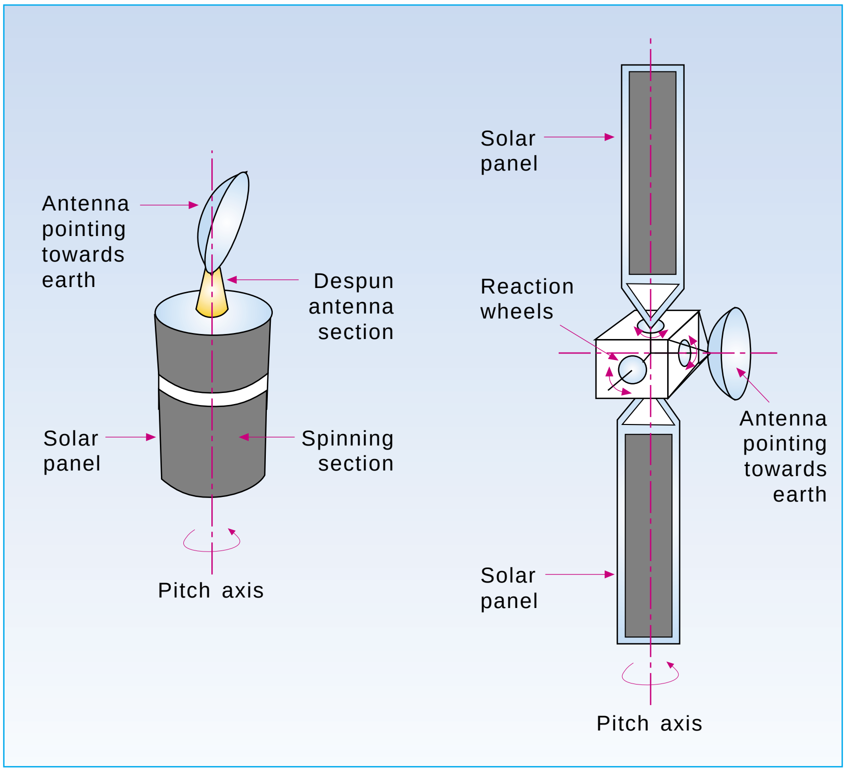


Figure 22 a) and b) The two concepts of stabilization

tion *(θ)* of the velocity vector can be calculated from figure 18. If the geosynchronous orbit has some inclination, the satellite will apparently move in a figure-of-eight with respect to the nodes, figure 19. The maximum excursion from equator in the north or south direction will be equal to the inclination.

6 The spacecraft

The main purpose of a communication satellite is to receive and transmit radio signals within the coverage area on the earth. This means that the spacecraft should be a reliable and stable platform for the communication system. The satellite is generally divided into two main modules:

1 The communication module (or the "payload") which includes
   - repeaters (or transponders)
   - antennas

2 The service module (or the "bus" or "platform") comprising the following subsystems
   - attitude and orbit control (AOCS)
   - telemetry, tracking and command (TT&C)
   - power supply.

Figure 20 shows an exploded view of a typical spacecraft.

6.1 Attitude and orbital control subsystem

The objective of the attitude control subsystem is to maintain the communication antennas correctly pointed towards the earth, and the solar cells correctly pointed towards the sun. The movements of the satellite about its centre of mass can be described by rotations about the three orthogonal reference axes: roll *(x)*, pitch *(y)* and yaw *(z)*, see figure 21.

There are two different methods of controlling the attitude: spin stabilisation and body-fixed stabilisation (3-axis stabilisation), see figure 22. In the first type, the satellite spins around its main axis of symmetry. The rotation is normally about 60 rpm. This will produce an angular momentum in a fixed direction. For a geostationary satellite, the spin axis (pitch) has to be parallel with the earth's axis of rotation. To maintain the pointing of the communication antennas towards the earth, the

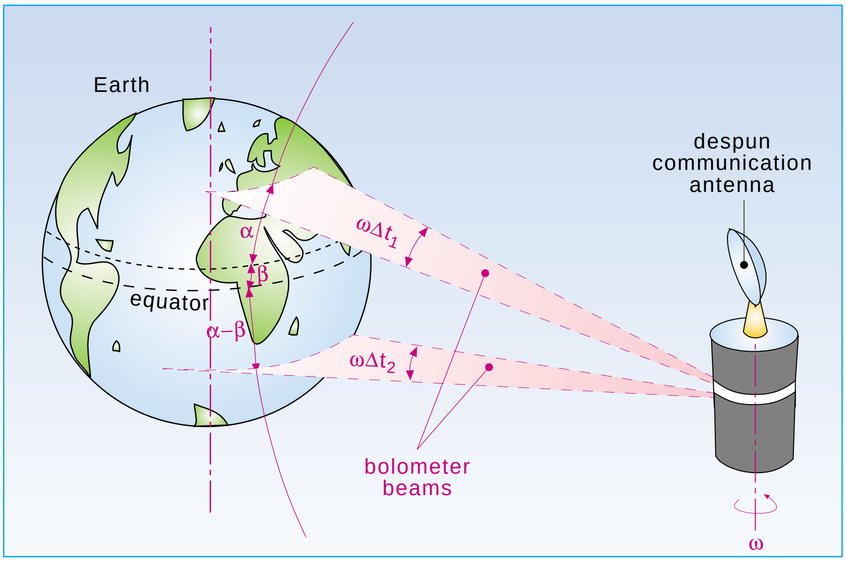


Figure 23 Earth contour detection by scanning infrared sensors (bolometers)

ω =	*spin rate of satellite*
2α =	*angle between the two bolometer beams*
β =	*error angle*
Δt_1 *and* Δt_2 =	*time between the horizon crossings*

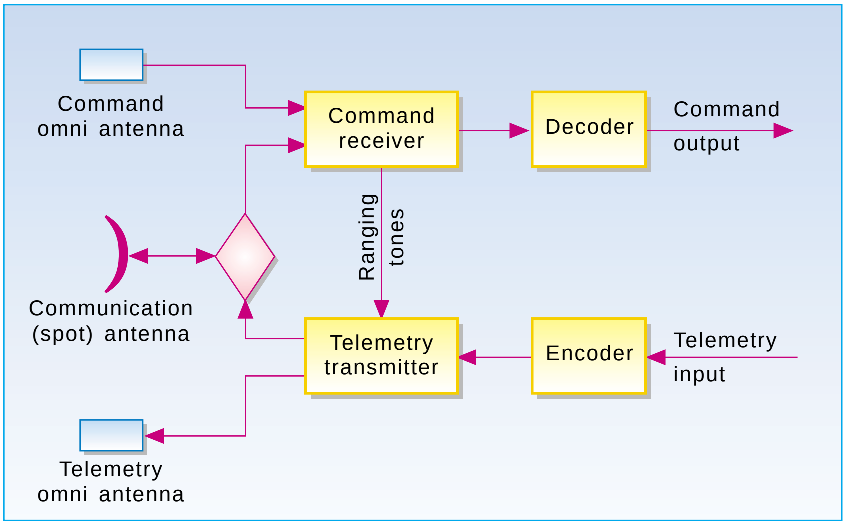


Figure 24 The satellite's TT&C system

platform containing the antennas has to rotate in the opposite direction (despun). Pitch correction is done by varying the angular velocity of the despin motor, since the rate of change of angular momentum is proportional to the torque. Yaw and roll corrections are made by thrusters mounted at appropriate places on the body. A typical spin-stabilised satellite is shown in figure 20.

To keep the gyroscopic stiffness in a body-fixed stabilised satellite, one may use a momentum wheel inside the satellite. This gives rather moderate antenna pointing. Instead, it is usual to use three reaction wheels spinning around the three main axes. While the momentum wheel is spinning at high velocity (about 10 000 rpm), the reaction wheels are spinning in both direction around zero speed. By varying the speed and direction of the reaction wheels, controlling torques can be applied around all axes. When the wheels reach their maximum speed, they must be unloaded by operating thrusters on the satellite's body.
A typical body-stabilised satellite is shown in figure 2. For operating the torque units, it is necessary to obtain the attitude of the satellite. This is done by sensors on board the spacecraft sensing the direction to the sun and earth. As seen from the geostationary orbit the sun subtends only 0.5°. It is therefore a good source for obtaining attitude reference. The sun sensor consists of photo cells which

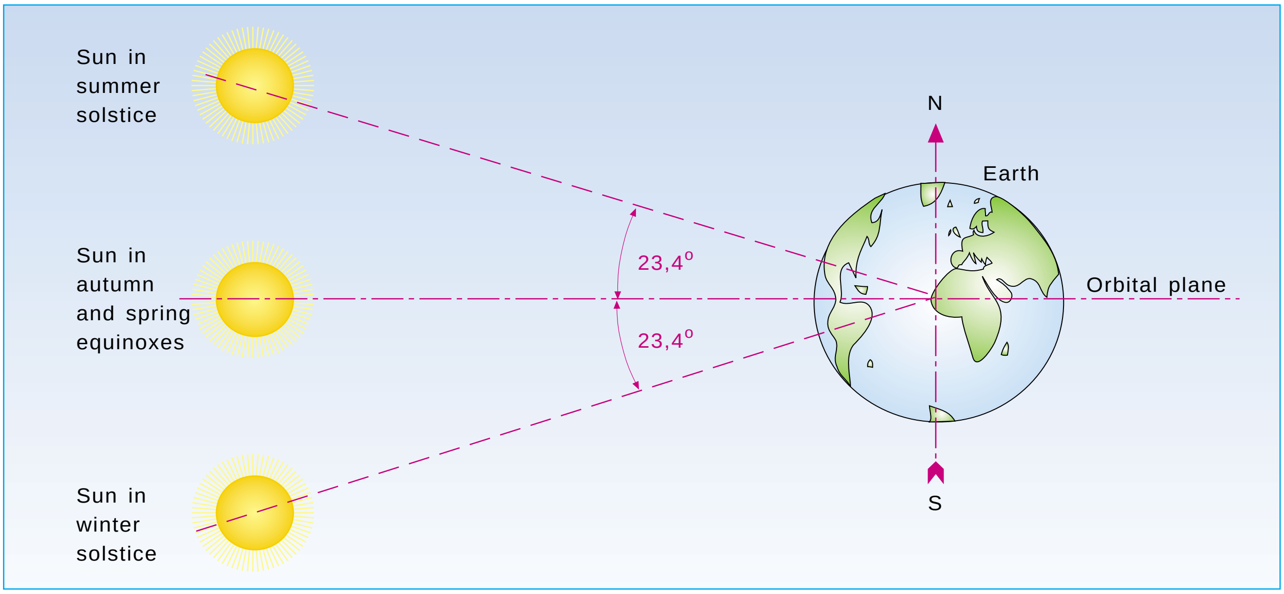


Figure 25 Apparent motion of the sun relative to earth

produce an electric current when illuminated by the sun.

The earth, seen from the satellite, has a much larger view angle (17.3°) and its centre cannot directly be measured. However, using infrared detectors (bolometers), one can measure the edge of the earth, due to the difference in radiation from the cold sky and the warm earth. Infrared detectors can be used both day and night. If two bolometers are mounted on a spinning spacecraft, as shown in figure 23, the pointing error (β) can be calculated measuring the time between the two horizon crossings (t_1 and t_2), knowing the spin rate (ω) and the angle between the bolometer beams (2α).

On a 3-axis stabilised satellite scanning mirrors have to be used, since the satellite itself does not spin.

A higher degree of pointing accuracy may be obtained by using a radio beacon on the earth. In such a system the satellite antenna is directly locked to the transmitted radio signal from the earth beacon.

6.2 Telemetry, tracking and command (TT&C)

The satellite is supervised and controlled via a dedicated earth station, TT&C station, which in turn is connected to the satellite control centre (SCC). The main tasks of the spacecraft management is to control the orbit and attitude of the satellite, monitor the "health" status, remaining propulsion fuel and transponder configuration, together with steering of antennas.

The satellite's TT&C system is shown in figure 24.

The telemetry subsystem collects data from various sensors on board the spacecraft, such as fuel tanks pressure, voltages/currents in different subsystems, sighting from infrared and sunlight detectors, temperature, etc. The information rate is usually less than 1 kbit/s. The data are transmitted on a low-power telemetry carrier using an omnidirectional satellite antenna during the transfer phase and/or a spot beam antenna when the satellite is on station.

The command subsystem is used for remote control of the different functions in the satellite. This may include attitude and orbital manoeuvres, switching transponders on and off, steering antennas, firing the apogee boost motor in transfer phase, etc. When receiving a command, the command subsystem generates a verification signal, which is sent back via the telemetry link. After checking, an execute signal is transmitted to the satellite. This prevents inadvertent commands. As in the telemetry link the omni antenna is used during the transfer phase, while spot-beams are used when the satellite is in its geostationary position.

The tracking or ranging subsystem determines the orbit of the spacecraft. There is a number of techniques that can be used. One method is to measure to pointing of the TT&C earth station antenna towards the satellite, together with the slant range to the satellite, i.e. the range vector. The range to the satellite can be measured by modulating the command carrier with multiple low frequency tones ("tone ranging"). These tones are demodulated and remodulated on the telemetry carrier and received in the TT&C ground station. By comparing the phase between the transmitted and received tones, the range can be calculated. The highest tone gives the best accuracy, but several lower tones resolve the ambiguity.

6.3 Power supply

Solar energy is the only external energy source for orbiting satellites. The solar radiation intensity is about 1.4 kW/m^2. The sun energy is converted to electrical energy by means of photovoltic cells (solar panels). The efficiency of the solar panels is about 10 - 15 %. On the spin-stabilised satellites the solar panels form the exterior of the spacecraft's body, see figures 20 and 22.

One disadvantage of the spin-stabilised satellite is that only 1/3 of total solar array is exposed to the sunlight. On the body-stabilised satellite the solar panels are mounted on two deployable "wings", as shown in figure 2 and 22. The satellite moves around the earth in 24 hours. For intercepting maximum solar flux, the solar panel wings have to make one turn per day. They are driven by two separate motors.

A measure of efficiency is the ratio of produced electrical power to the mass of satellite. For a spinning satellite this ratio is about 10 Watts/kg, whereas for the 3-axis stabilised satellite the ratio is about 50 Watts/kg. This means that if we need a satellite with high RF output power, we should select a body stabilised satellite. The disadvantage is that they are more complex.

On-board batteries are needed to produce power during the launch phase and when the satellite enters the shadow of the earth (eclipse). This happens around the equinoxes, and the maximum duration is approximately 70 minutes. This is due to the fact that the earth's equatorial plane is inclined with an angle of 23.4° with the direction to the sun, see figure 25. During the summer and winter seasons the satellite is out of the earth's shadow, figure 26, whereas during the equinoxes the satellite passes through the earth's shadow once a day for 42 days, figure 27, i.e. the satellite will experience 84 eclipses per year.

6.4 The communication module

The communication subsystem is the primary system in the satellite. All other functions in the satellite can be considered as supporting activities for this system. It consists of the receivers, transmitters and the communications antennas. One receive/transmit chain is called a transponder. A satellite with 5 transponders is shown schematically in figure 28. The building blocks are:

- A wide band receiver and a down converter
- An input multiplexer (IMUX)
- 5 channelised sections including the high power amplifiers
- An output multiplexer (OMUX).

The wide band receiver/converter operates in the whole up-link band, e.g. in one of the common bands

- C-band (5.9 - 6.4 GHz)
- Ku-band (14 - 14.5 GHz)
- Ka-band (17.3 - 18.1 GHz)

or other bands allocated to satellite communication.

The frequency conversion could either be single, as in figure 28, or dual. In the latter case the first mixer converts the frequency down to an intermediate frequency (IF). After the channel amplification on IF the signal is converted up to the down-link frequency, before the last power amplification. This is often done because it is much easier to make filters, amplifiers and equalisers at an intermediate frequency. The intermediate frequency can be chosen independently of the up- and down-link frequencies.

The input multiplexer separates the up-link band into individual channels with a certain bandwidth. For example can a 500 MHz wide up-link band be separated into 12 channels with channel bandwidth of 36 MHz. The channel bandwidths may also be unequal. The channel amplifier/attenuator sets the gain (gain-step) of the transponder in order to control the input back-off of the TWTAs. This can be done from the TT&C station by the up-link command system.

The last high-power amplifier (HPA) is usually a travelling wave tube amplifier (TWTA). This amplifier establishes the level of the output power, which may be in the range of 10 - 200 Watts depending on the application. The TWTA operates near saturation. The input/output characteristic is shown in figure 29. The output multiplexer combines all the channels before signals are transmitted to the earth by the antenna.

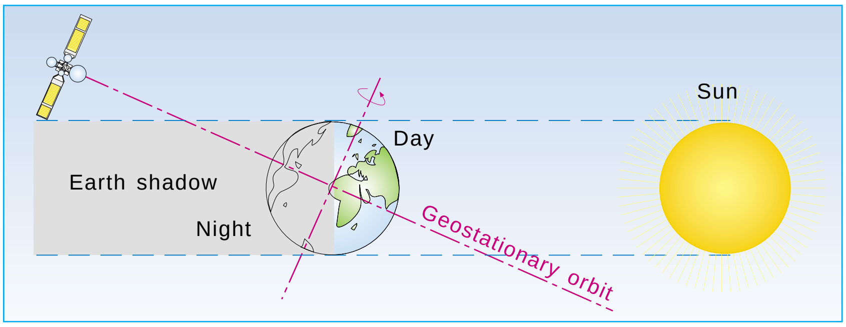


Figure 26 During summer and winter the satellite is always illuminated

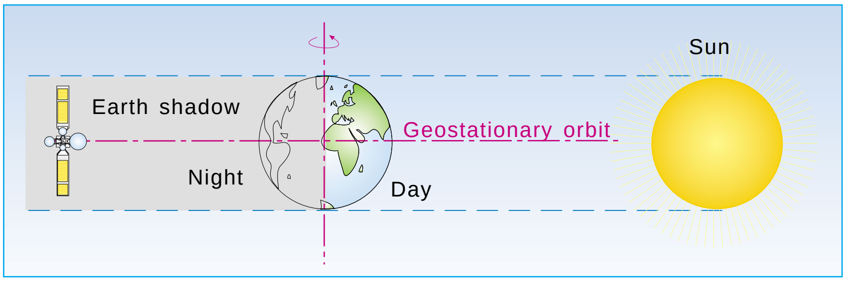


Figure 27 The satellite passes through the earth shadow at the equinoxes once a day

6.5 The satellite communication antenna

Antennas are used for receiving and transmitting modulated radio frequent signals from and to the earth. The most commonly used antenna in the microwave frequency band is the parabolic reflector antenna. The antenna is characterised by its radiation pattern and the maximum gain, that is the antenna's capability to concentrate the energy in one direction.

The design of the satellite antenna is depending on required coverage area,

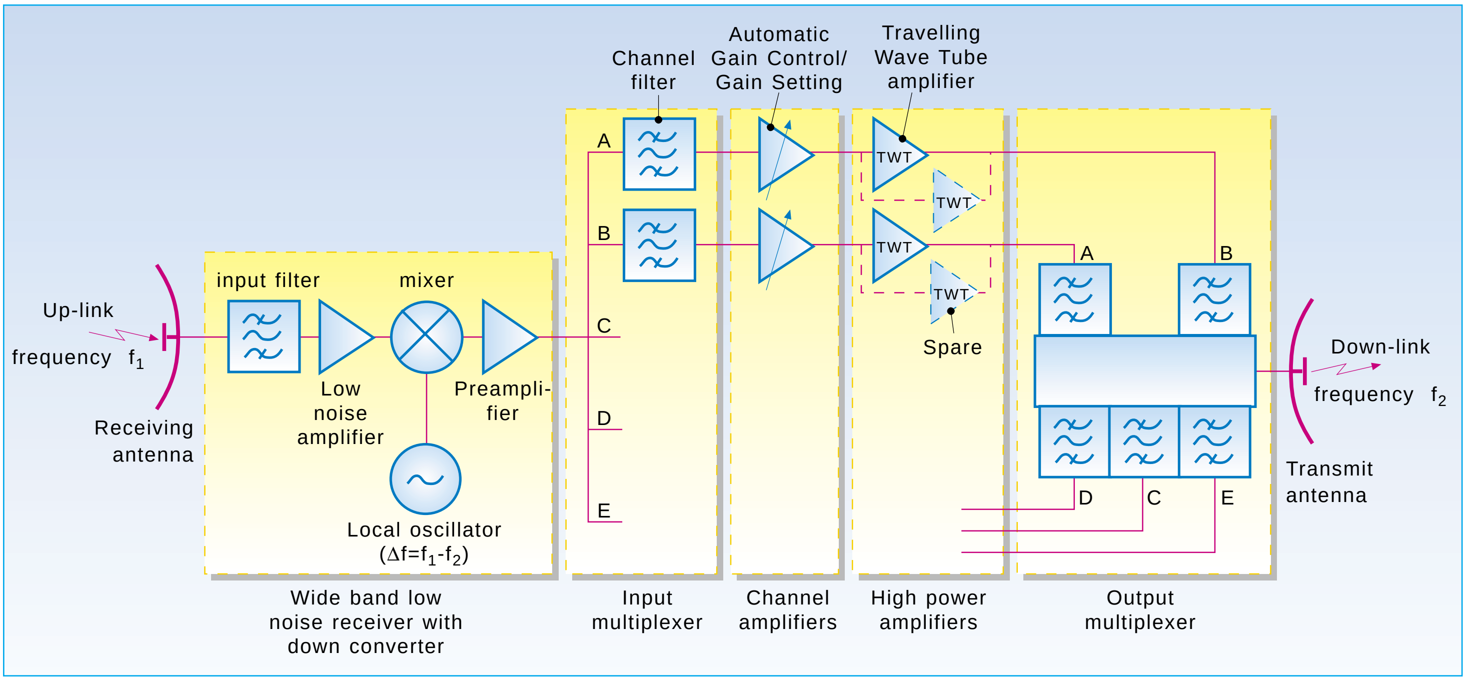


Figure 28 A satellite with 5 transponders and single frequency conversion

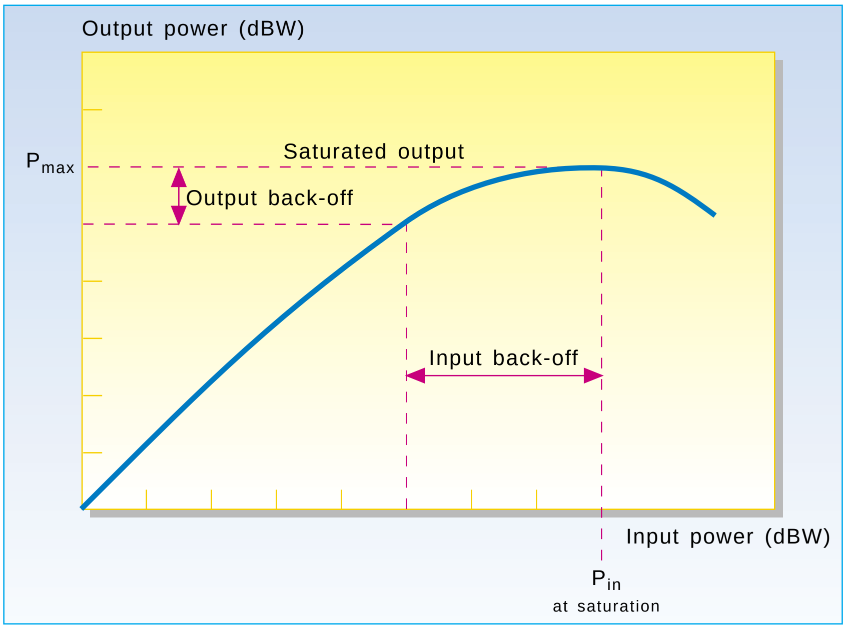


Figure 29 Input/output characteristic of a TWTA

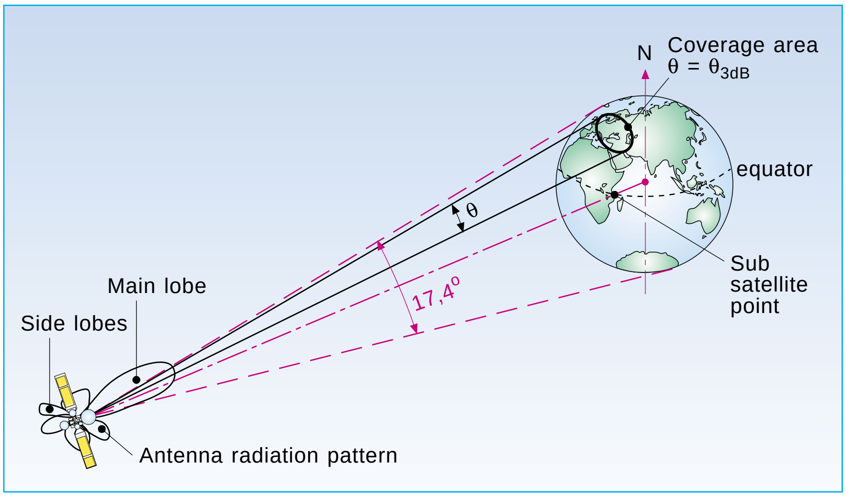


Figure 30 Coverage area of the satellite antenna. For global coverage the required beam width is 17.4°

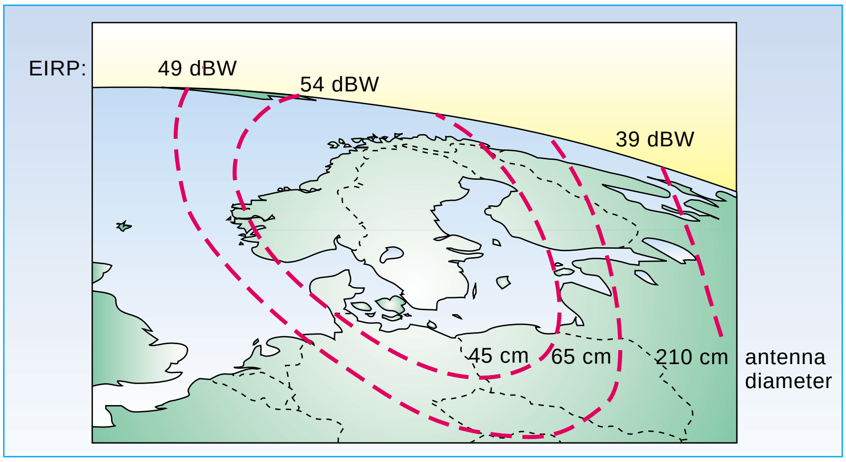


Figure 31 The preliminary EIRP contours of the THOR satellite and the corresponding receiving antenna diameter for TV reception with good quality

see figure 30. This determines the beam width of the antenna, usually taken as the half-power beam width (or 3 dB beam width). The 3 dB beam width is given by

$$\theta_{3dB} = K \cdot \frac{\lambda}{D} \text{ (degrees)} \qquad (6.1)$$

where

K = 65 - 75, depending on the antenna aperture field distribution

λ = wavelength

D = diameter of antenna.

The gain of the antenna is proportional to its area, and is given by

$$G = \eta \frac{4\pi}{\lambda^2} \cdot A = \eta \left(\frac{\pi D}{\lambda} \right)^2 \qquad (6.2)$$

where

η = antenna efficiency (0.5 - 0.8)

A = area of antenna

or in dB:

$$G(dB) = 10 \log \left[\eta \left(\frac{\pi D}{\lambda} \right)^2 \right] \qquad (6.3)$$

Obviously the gain and the beam width of a reflector antenna are related. The gain is approximately given by

$$G = \frac{30000}{\theta_{3dB}^2} \qquad (6.4)$$

where

θ_{3dB} is given in degrees.

Should the service area be the whole earth, the required beam width would be 17.4° as seen from the geostationary orbit. This corresponds to a satellite antenna with 20 dB gain. By increasing the size of the satellite antenna only a small area of the earth is illuminated. These narrow beams are called “spot beams”. If we should cover an area on earth with a diameter of 300 km, the beam width of the antenna should only be 0.5°, with a corresponding gain of 50 dB. This means that we need less power from the HPA in the spot beam to achieve the same flux density on earth as in the global beam.

To express the transmitted power of a satellite, the Equivalent Isotropic Radiated Power *(EIRP)* is used. This is simply the product of output power from $HPA(P_t)$ and satellite antenna gain (G_t):

$$EIRP = P_t \cdot G_t \qquad (6.5)$$

So instead of plotting coverage contours (or gain contours) we can plot EIRP contours. An example is given in figure 31.

A satellite can have more than one antenna. Modern satellites employ several antennas, both with global, hemispheric and spot coverage.

As described, the communication payload overall can be characterised by the following parameters:

- the coverage area
- the maximum EIRP
- the input power flux density for saturation of the output high power amplifier
- the channel arrangement or frequency plan.

7 The satellite link

7.1 The link geometry

In order to communicate with a satellite, it is necessary to know the correct pointing of the earth station antenna. We shall limit ourselves to geostationary satellites. In that case, as previously mentioned, the satellite will not move with respect to an earth station, and the antenna pointing can be fixed.

The pointing of the antenna is defined by two angles, azimuth (α) and elevation (ε). The azimuth is the angle from true north to the projection of the line to the satellite in the local horizontal plane. The elevation is the angle above the local horizontal plane to the satellite.

Knowing the satellite's position, i.e. the longitude, and the earth station position, i.e. longitude and latitude, we can calculate the earth station antenna's pointing angles. The geometry is shown in figure 32.

With the terms given in figure 32, the azimuth, elevation and distance to satellite are given by:

$$\alpha = arctg\left(\pm\frac{tg\gamma}{\sin\phi}\right) \qquad (7.1)$$

$$\varepsilon = arctg\frac{\cos\phi\cos\gamma - \frac{R_e}{R_s}}{\sqrt{1-(\cos\phi\cos\gamma)^2}} \qquad (7.2)$$

$$d = R_s\frac{\sqrt{1-(\cos\phi\cos\gamma)^2}}{\cos\varepsilon} \qquad (7.3)$$

The maximum latitude coverage of a geostationary satellite $(\varepsilon = 0,\ \gamma = 0)$ will be

$$\varphi_{max} = \frac{\pi}{2} - arcsin\frac{R_e}{R_s} = 81.3° \text{ N or S}$$

7.2 Link analysis

The purpose of a satellite system is to provide reliable transmission with a specified quality of the received signal. The transmitted information has to be modulated on an RF carrier. In analogue systems, where frequency modulation (FM) is the dominating modulation method, the signal-to-noise ratio (S/N) after the demodulator is a measure of signal quality. In digital satellite links the measure of quality is the bit error rate (BER). The modulation method most often used in digital system is phase shift keying (PSK).

In both analogue and digital systems there is a unique relationship between the carrier-to-noise ratio (C/N) and the signal-to-noise (S/N) ratio or the bit error rate (BER). Given the modulation method, the performance of a total link is generally specified in terms of a minimum C/N in a certain percentage of time. In order to establish the link quality, we therefore need to calculate the carrier power (C) and the noise power (N) at the receiving station.

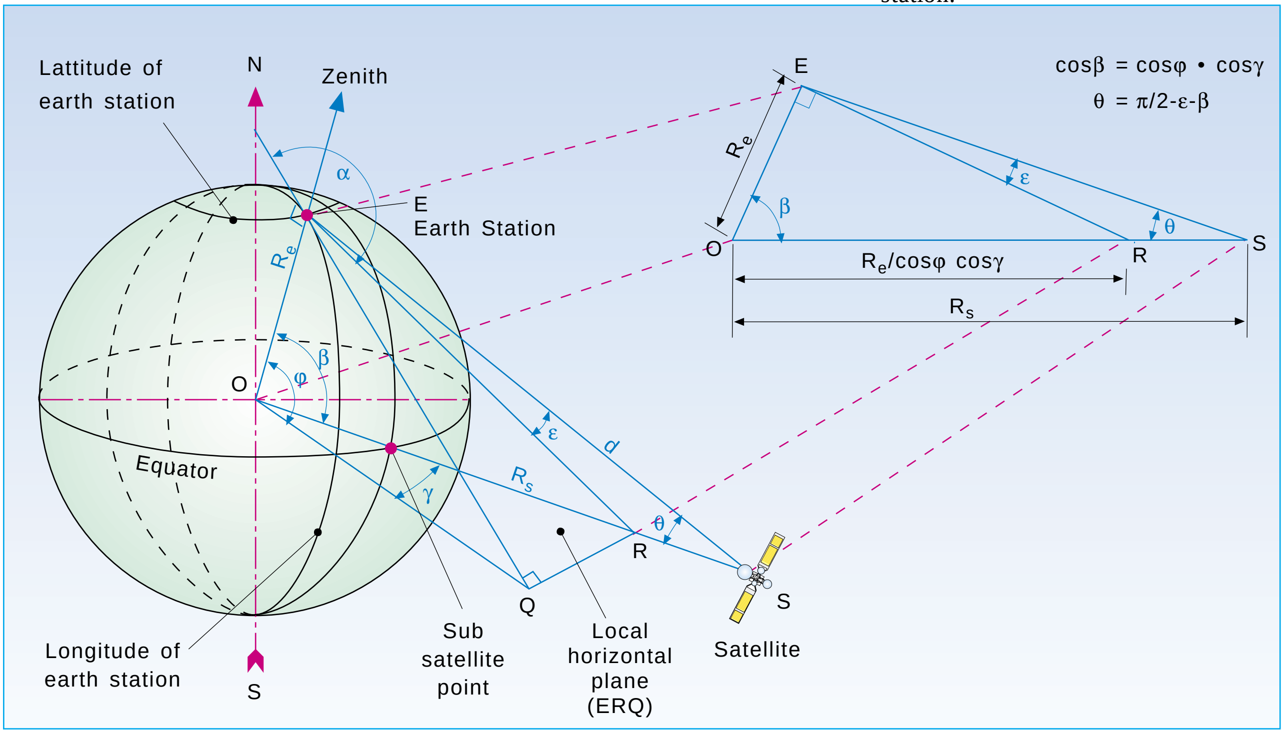


Figure 32 a) and b) Calculations of azimuth, elevation and distance to a GEO satellite

Considering an isotropic antenna, that is an antenna which radiates with uniform intensity in all directions, the power flux density *(PFD)* at a distance, *d*, will be

$$PFD = \frac{P_t}{4\pi d^2} \qquad (7.4)$$

where P_t = the total radiating power. Using a directive antenna with a gain G_t, the power flux density can be increased by G_t:

$$PFD = \frac{P_t \cdot G_t}{4\pi d^2} = \frac{EIRP}{4\pi d^2} \qquad (7.5)$$

If the receiving antenna has an effective area of

$$A_e = \eta\, A_r \qquad (7.6)$$

where

η = efficiency

A_r = physical area

the received power *(C)* will be:

$$C = PFD \cdot A_e = \frac{EIRP}{4\pi d^2} \cdot A_e\,(W) \qquad (7.7)$$

Using the relationship between the effective antenna area (A_e) and gain (G_r), equation (6.2), the received power may be written as

$$C = EIRP \cdot \left(\frac{\lambda}{4\pi d}\right)^2 \cdot G_r \qquad (7.8)$$

The term in the middle is called the free space loss:

$$L_o = \left(\frac{4\pi d}{\lambda}\right)^2 \qquad (7.9)$$

In addition to the free space loss, we have several other sources of attenuation, such as atmospheric absorption, rain attenuation, etc. ref O Gutteberg: *"Effects of atmosphere on earth-space radio propagation"* (this issue).

All components, active and passive, produce electrical noise. This is due to the thermal agitation of electrons. The thermal noise power *(N)* increases with temperature *(T)* and band width *(B)* and is given by

$$N = kTB \qquad (W) \quad (7.10)$$

where

k = Bolzmann's constant $= 1.38 \cdot 10^{-23}$ W / Hz · K

We can now define an effective input noise temperature of a receiver (or a network). This is the temperature of a noise source located at the input of a noiseless receiver giving the same output noise power as the noisy receiver.

Noise power at the earth station receiver input has contributions both from the noisy receiver and the noise picked up by the antenna.

Hence, the total system noise temperature (T_S) is given by

$$T_S = T_A + T_R \qquad (7.11)$$

where

T_A = antenna noise temperature

T_R = effective noise temperature of the receiver.

The total noise power *(N)* is then given by:

$$N = kT_S B \qquad (7.12)$$

If the received carrier power *(C)* is measured at the same point, according to the equations (7.8) and (7.12), the *C/N* ratio will be

$$\frac{C}{N} = EIRP \cdot \frac{1}{L_o} \cdot \frac{G_r}{T_s} \cdot \frac{1}{kB} \qquad (7.13)$$

Equation (7.13) is called the link budget, and the term G_r / T_s is the receiver's "figure-of-merit". It specifies the sensitivity of the receiver, either in the satellite or in the earth station. The reference point for G_r and T_s has to be the same.

A typical receiver "front-end" is shown in figure 33. Referred to the low noise amplifier's input, the overall noise temperature is given by

$$T_s = \frac{T_A}{L} + \left(\frac{L-1}{L}\right) T_0 + T_{LNA} \qquad (7.14)$$

where

L = feeder loss

T_o = ambient temperature = 290 K

T_{LNA} = noise temperature of the low noise amplifier.

Succeeding stages of the receiver will not contribute if the gain of the *LNA* is sufficiently high.

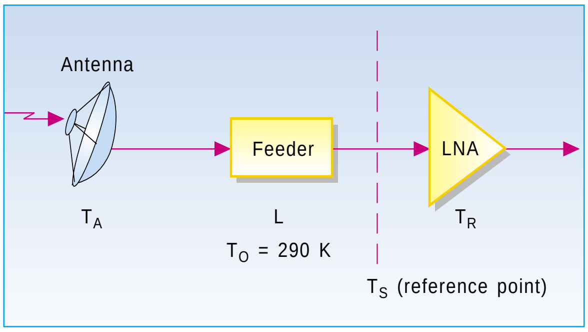


Figure 33 A typical receiver "front-end"

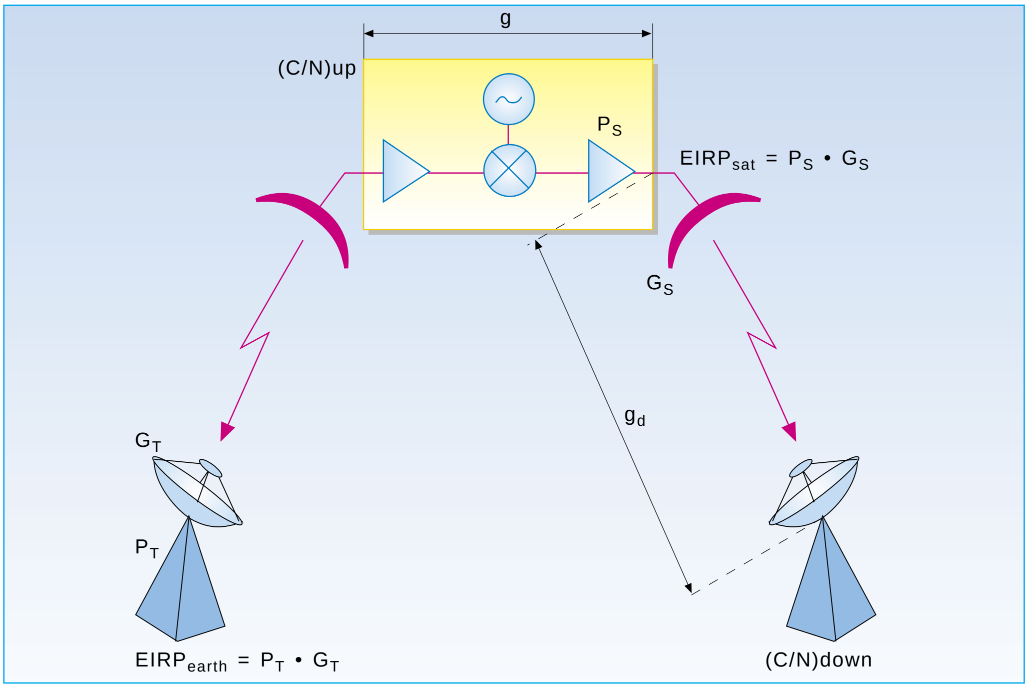


Figure 34 Basic satellite link

Instead of giving the noise temperature of a component, the noise figure *(F)* is often used. This is defined as the signal-to-noise ratio at the input divided by the signal-to-noise ratio at the output:

$$F = \frac{S_i / N_i}{S_o / N_o} \qquad (7.15)$$

If the source noise temperature is at the standard temperature T_o = 290 K, then

$$F = \frac{S_i}{S_o} \cdot \frac{N_o}{N_i} = \frac{1}{g} \cdot \frac{N_o}{kT_o B} \qquad (7.16)$$

where

g = gain of the network.

Using the definition of effective input noise temperature:

$$F = \frac{k(T_o + T)B \cdot g}{kT_o Bg} = 1 + \frac{T}{T_o} \qquad (7.17)$$

or

$$T = (F - 1)\ 290 \qquad (7.18)$$

A satellite system consists of both an up-link and a down-link. Noise on the up-link will also contribute to the overall noise power received at the earth station. The basic satellite link is shown in figure 34.

The total received noise power at the receiving earth station is given by

$$N_t = N_u \cdot g \cdot g_d + N_d$$

where

N_u = noise power on the up-link

N_d = noise power on the down-link

g = total transponder gain

g_d = "section gain" on the down-link (free space loss and antenna gains).

The total noise-to-carrier ratio is then

$$\frac{N}{C} = \frac{N_{up} g g_d + N_{down}}{C} = \frac{N_{up} g g_d}{C} + \frac{N_{down}}{C} = \frac{N_{up}}{C / g \cdot g_d} + \frac{N_{down}}{C}$$

since $\dfrac{C}{g \cdot g_d} = C_{up}$

and $C = C_{down}$

then

$$\frac{C}{N} = \frac{1}{\frac{1}{(C/N)_{up}} + \frac{1}{(C/N)_{down}}} \qquad (7.19)$$

Using equation (7.19) we can calculate the total C/N ratio, knowing C/N ratios on the up- and down-link.

If the $(C/N)_{up}$ is more than 10 dB above $(C/N)_{down}$, the noise contribution from the up-link to the down-link is only a few tenths of a dB. Often it is disregarded in link calculations where there is enough power on the up-link.

Knowing the

- location of satellite and earth station
- *EIRP* and *G/T* of the satellite
- *EIRP* of the earth station
- the required quality *(C/N)*
- the carrier frequencies
- the receiver overall system temperature *T* (or noise figure *F*)
- the equivalent noise bandwidth of the receiver *(B)*

we are now able to calculate the receiving station's figure-of-merit *(G/T)* or receiving station's antenna diameter.

In figure 35 is shown the EIRP contours for INTELSAT-V A satellite in 1° west.

Using the above developed equations, the corresponding receiving antenna diameters have been calculated, figure 35. For the calculations the following parameters have been assumed:

C/N = 12 dB (TV reception with good quality)

B = 27 MHz (receiver noise bandwidth)

F = 1 dB (receiver noise figure)

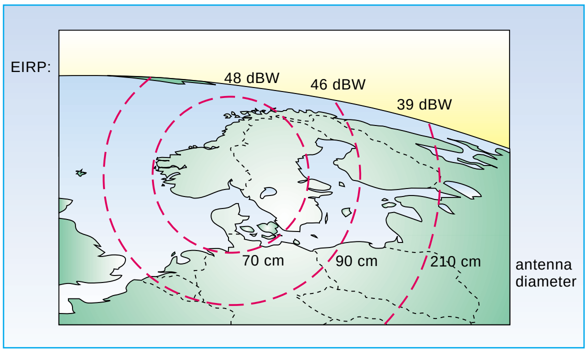


Figure 35 EIRP contours for INTELSAT-V-A at 1° west and the corresponding receiving TV antenna diamters. The different parameters are given in the text

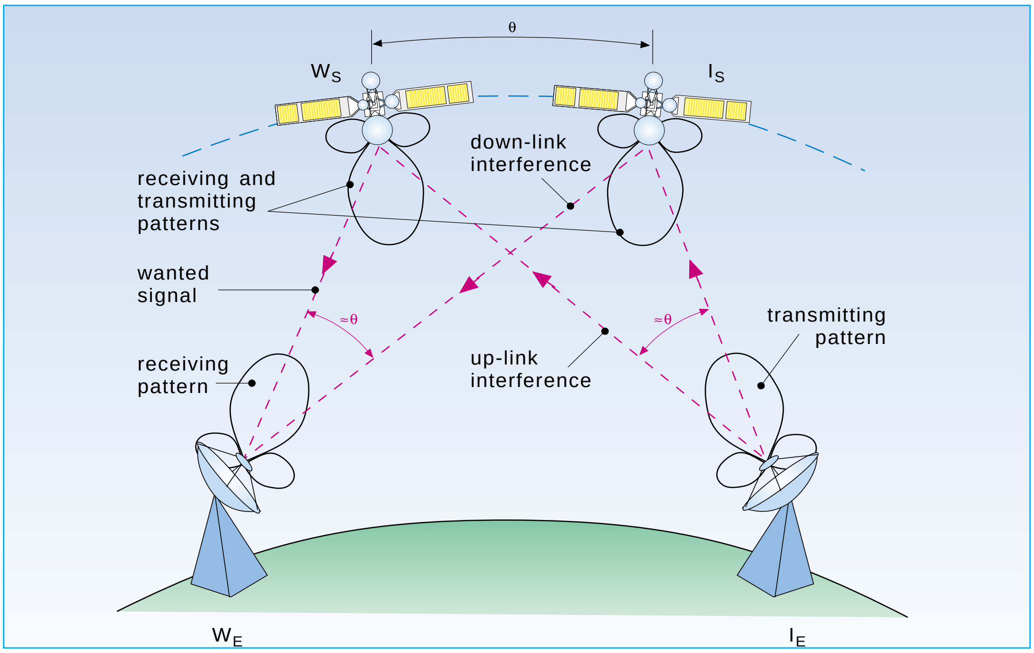


Figure 36 Simplified interference geometry

W_s, I_s = wanted and interfering satellite

W_E, I_E = wanted and interfering earth station

θ = angular separation between satellites

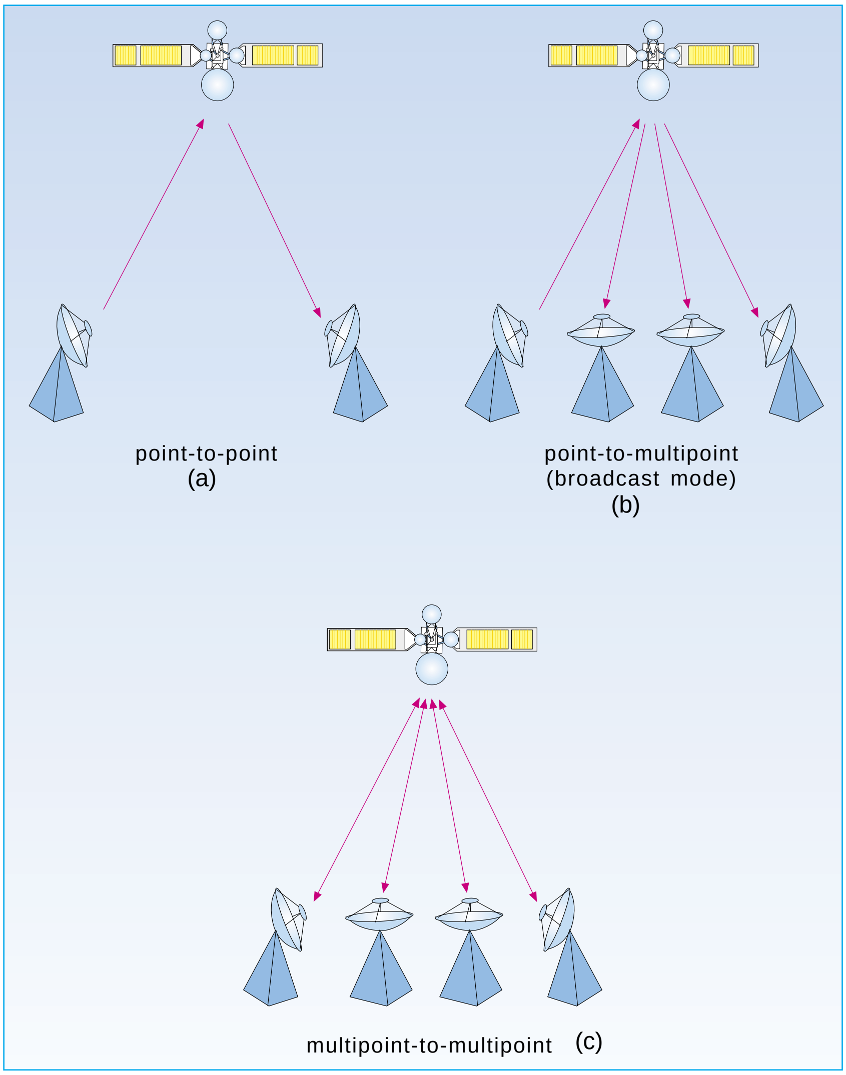


Figure 37 a), b) and c) Satellite network configurations

T_A = 30 K (antenna temperature, clear sky condition)

L = 0.5 dB (feeder loss)

f = 11.7 GHz (carrier frequency).

Corresponding calculations have been done for the diameters given in figure 31. The noise from the up-link has been disregarded.

In the calculations above we have only considered additive thermal noise. There are also other sources of degradation such as interference from other satellites and radio links, intermodulation from the unlinear transponder, etc. The effect of these sources are often also treated as additive thermal noise and the C/I terms should be added in equation (7.19):

$$\frac{C}{N} = \frac{1}{\frac{1}{(C/N)_{up}} + \frac{1}{(C/N)_{down}} + \frac{1}{C/I} + etc.} \tag{7.20}$$

Keeping the interference from other satellites below a certain value will put severe limits on the minimum diameter of the receiving earth station's antenna. Knowing the radiation patterns for the satellites and earth stations involved, one can calculate the interference level. The geometry is shown in figure 36.

In the near future, the interference from other satellites may be the critical factor for the design of earth station antennas. This is already the case for certain positions in the geostationary arc over Europe.

8 Satellite networks

The simplest form of satellite networks is the point-to-point or point-to-multipoint configuration shown in figure 37a) and b). So far, we have only considered these cases. However, there is often a need for several earth stations to be interconnected through the same transponder, multipoint-to-multipoint, shown in figure 37c). The methods for allowing several users to utilise the same transponder are called multiple access techniques. The transponder is a resource which can be characterised by its available power and bandwidth. Efficient use of this common resource is a very important problem in satellite communication. The channels are designed to the users either fixed (pre-assigned) or on demand (demand assignments).

There are basically three methods of multiple access:

- Frequency Division Multiple Access (FDMA)
- Time Division Multiple Access (TDMA)
- Code Division Multiple Access (CDMA).

Random Access (RA) may also be utilised.

In FDMA, each user is permanently allocated a certain frequency band, out of the total bandwidth of the transponder, figure 38a). To reduce the adjacent channel interference, it is necessary to have guard bands between the sub-bands. Frequency drifts of the satellite's and earth station's frequency converters have also to be taken into consideration. FDMA is the traditional technique due to its simple implementation. However, due to the non-linear characteristics of the transponder, figure 29, a certain back-off of the TWT is necessary for multicarrier operation, in order to control the intermodulation products. This reduces the total transponder capacity.

In TDMA all stations use the same carrier frequency, but they are only allowed to transmit in short non-overlapping time slots, figure 38b). Thus, the intermodulation products due to the non-linear transponder are avoided. This means that the high power amplifier in the transponder can be operated near saturation. Accordingly, both the total transponder power and bandwidth are available. However, the TDMA technique obviously requires time synchronisation and buffer storage. This leads to rather complex earth stations.

In CDMA all users occupy the total transponder bandwidth all the time.

The users can be separated because each channel is multiplied by a unique spreading code. The composite signal is then modulated onto a carrier. The information is recovered by multiplying the demodulated signal with the same spreading code. The receiving earth station is accordingly able to recover the transmitted message by a specific user. No frequency or time co-ordination is needed before accessing the transponder. New users can easily be included. CDMA signals are resistant to interfering signals. This property can be utilised in systems with very small aperture terminals (VSAT), where interference may be received from adjacent satellites. The main disadvantages are cost and complexity of the receivers.

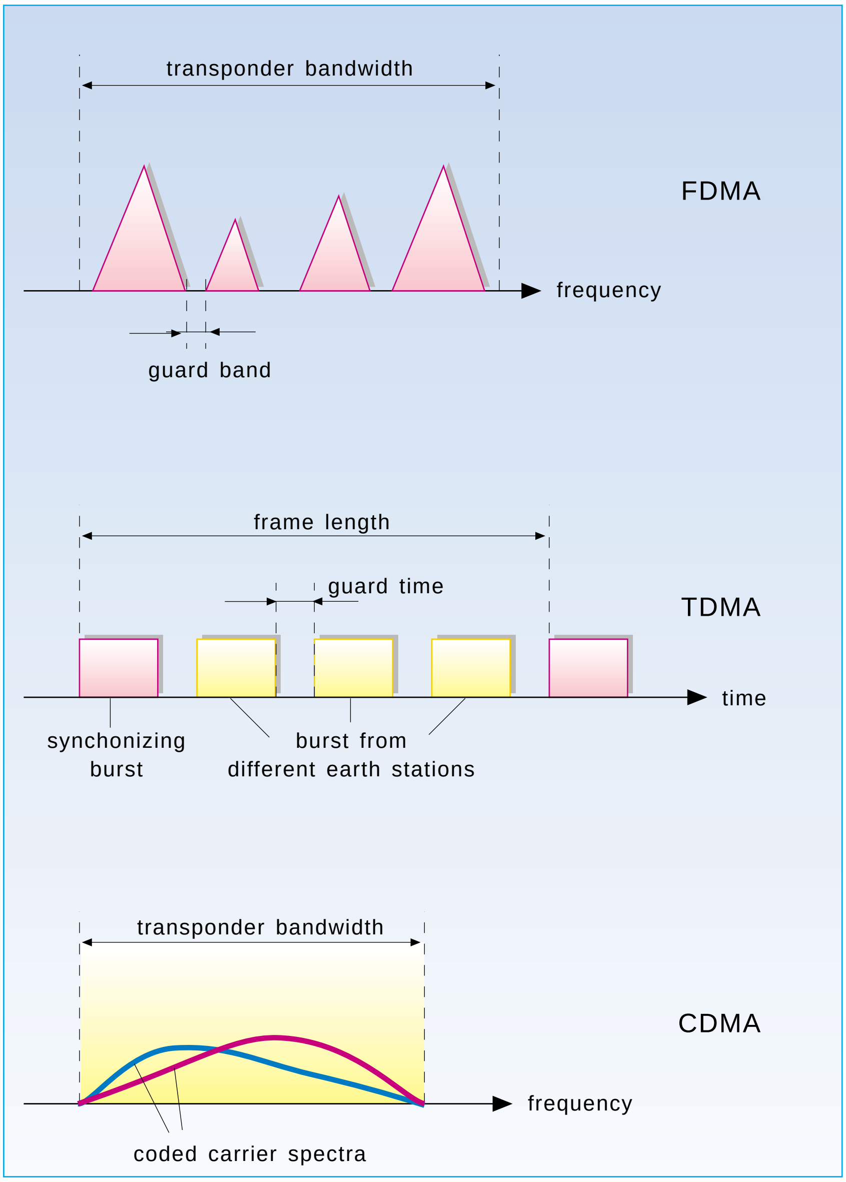


Figure 38 a), b) and c) Basic multiple access methods

9 Literature

1 Berlin, P. *The geostationary applications satellite*, Cambridge University Press, 1988.

2 Maral, G, Bousquet, M. *Satellite communications systems*, John Wiley & Sons, 1986.

3 Ha, T T. *Digital satellite communications*, Macmillan Publishing Co., 1986.

4 Evans, B G. *Satellite communication systems*, Peter Peregrinus Ltd., 2nd edition, 1991.

5 Pratt, T, Bostian, C W. *Satellite communications*, John Wiley & Sons, 1986.

6 Dalgleish, D I. *An introduction to satellite communications*, Peter Peregrinus Ltd., 1989.

7 Feher, K. Digital communications. *Satellite/earth station engineering*, Prentice-Hall Inc., 1983.

8 ITU/CCIR. *Handbook of satellite communications*, ITU, Geneva, 1988.

9 Stette, G. *Forelesninger i satellittkommunikasjon*, NTH.

10 Gutteberg, O, Lundstrøm, L-I, Andersson, L. *Satellittfjernsyn*, Universitetsforlaget, 1986.

11 Gagliardi, R G. *Satellite communications*, Van Nostrand Reinhold, 2nd edition, 1991.