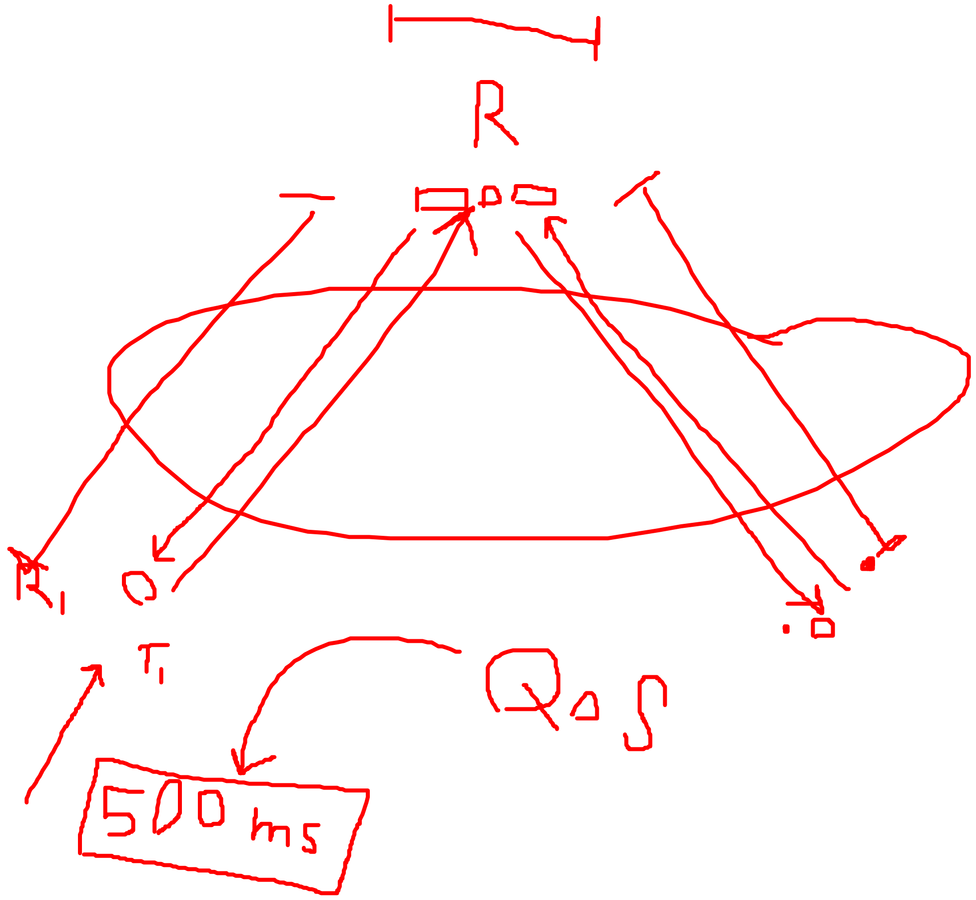
Satellite Communications

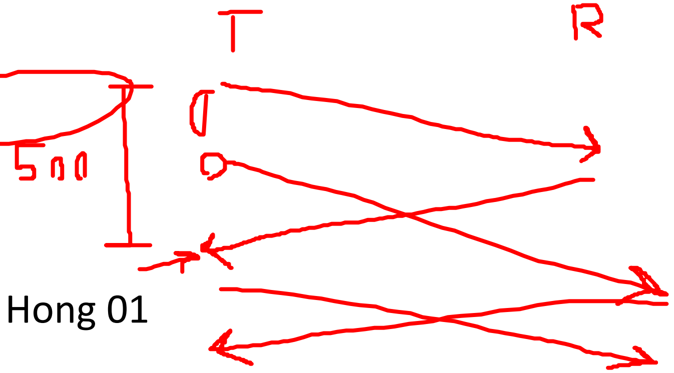
Dr Ashraf Ali
Hashemite University
Electrical Engineering Department
Second Semester 2020

BIRTH OF SATELLITE COMMUNICATIONS

- The Second World War stimulated the expansion of two very distinct technologies—missiles and microwaves
- The expertise eventually gained in the combined use of these two techniques opened up the era of satellite communications.
- **1957**
  - October 4, 1957: - First satellite - the Russian Sputnik 01
  - First living creature in space: Sputnik 02
- **1958**
  - First American satellite: Explorer 01
  - First telecommunication satellite: This satellite broadcast a taped message: Score
- **1959**
  - First meteorology satellite: Explorer 07
- **1960**
  - First successful passive satellite: Echo 1
  - First successful active satellite: Courier 1B
  - First NASA satellite: Explorer 08
- April 12, 1961: - First man in space

- 1962
  - First telephone communication & TV broadcast via satellite: Echo 1
  - First telecommunication satellite, first real-time active, AT&T: Telstar 1
  - First Canadian satellite: Alouette 1
  - On 7th June 1962 at 7:53p the two-stage rocket; Rehbar-I was successfully launched from Sonmiani Rocket Range.
  - Rehbar-II followed a successful launch on 9th June 1962
- **1963**
  - Real-time active: Telstar 2
- **1964**
  - Creation of Intelsat
  - First geostationary satellite, second satellite in stationary orbit: Syncom 3
  - First Italian satellite: San Marco 1
- 1965
  - Intelsat 1 becomes first commercial comsat: Early Bird
  - First real-time active for USSR: Molniya 1A
- 1967
  - First geostationary meteorology payload: ATS 3
- 1968
  - First European satellite: ESRO 2B
- July 21, 1969: - First man on the moon
- 1970
  - First Japanese satellite: Ohsumi
  - First Chinese satellite: Dong Fang Hong 01





- 1971
  - First UK launched satellite: Prospero
  - ITU-WARC for Space Telecommunications
  - INTELSAT IV Launched
  - INTERSPUTNIK - Soviet Union equivalent of INTELSAT formed
- 1974
  - First direct broadcasting satellite: ATS 6
- 1976
  - MARISAT - First civil maritime communications satellite service started
- 1977
  - EUTELSAT - European regional satellite
  - ITU-WARC for Space Telecommunications in the Satellite Service
- 1979
  - Creation of Inmarsat
- 1980
  - INTELSAT V launched - 3 axis stabilized satellite built by Ford Aerospace
- 1983
  - ECS (EUTELSAT 1) launched - built by European consortium supervised by ESA
- 1984
  - UK's UNISAT TV DBS satellite project abandoned
  - First satellite repaired in orbit by the shuttle: SMM

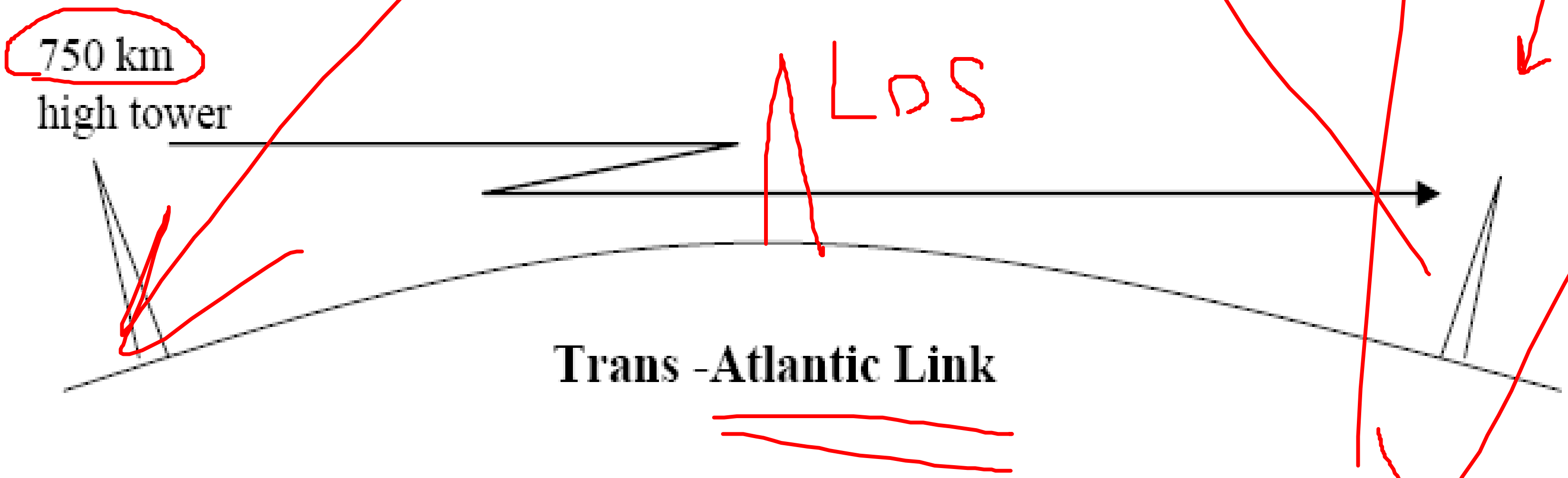
- 1985
  - First Brazilian satellite: Brazilsat A1
  - First Mexican satellite: Morelos 1
- 1988
  - First Luxemburg satellite: Astra 1A
- 1989
  - INTELSAT VI - one of the last big "spinners" built by Hughes
  - Creation of Panamsat - Begins Service
  - On 16 July 1990, Pakistan launched its first experimental satellite, BADR-I from China
- 1990
  - IRIDIUM, TRITIUM, ODYSSEY and GLOBALSTAR S-PCN projects proposed - CDMA designs more popular
  - EUTELSAT II
- 1992
  - OLYMPUS finally launched - large European development satellite with Ka-band, DBTV and Ku-band SS/TDMA payloads - fails within 3 years
- 1993
  - INMARSAT II - 39 dBW EIRP global beam mobile satellite - built by Hughes/British Aerospace
- 1994
  - INTELSAT VIII launched - first INTELSAT satellite built to a contractor's design
  - Hughes describe SPACEWAY design
  - DirecTV begins Direct Broadcast to Home
- 1995
  - Panamsat - First private company to provide global satellite services.
- 1996
  - INMARSAT III launched - first of the multibeam mobile satellites (built by GE/Marconi)
  - Echostar begins Diresct Broadcast Service

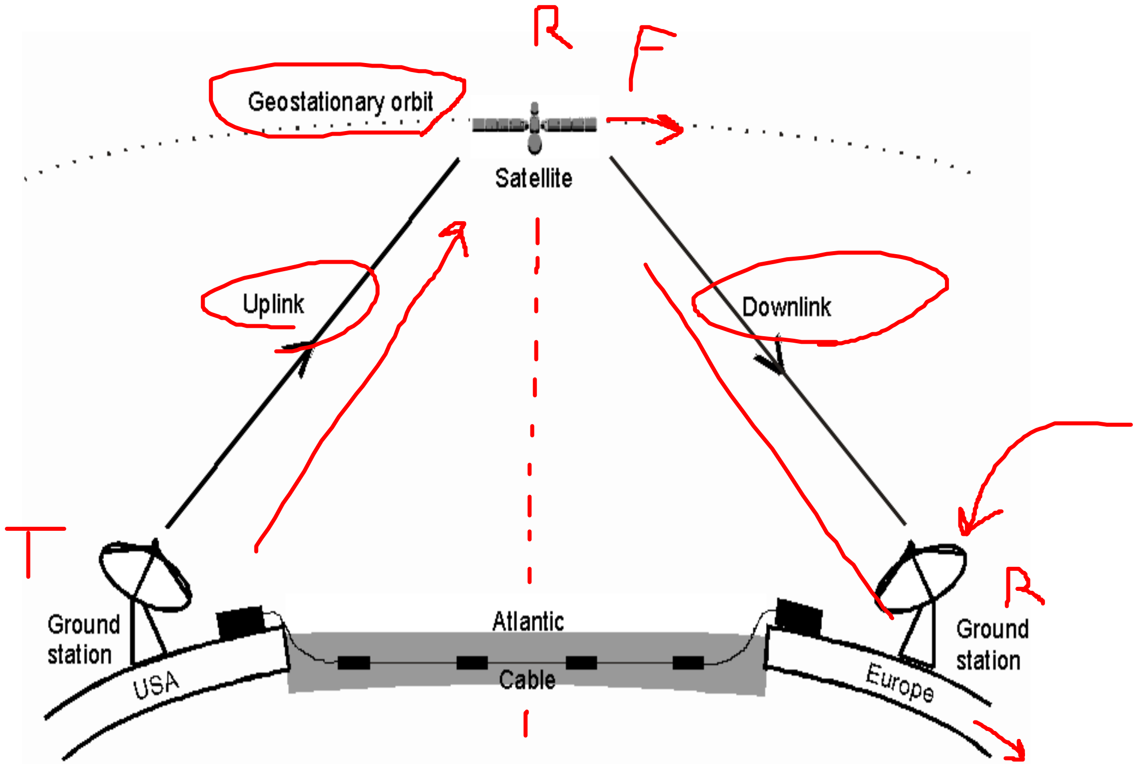
- 1997
  - IRIDIUM launches first test satellites
  - ITU-WRC'97
- 1999
  - AceS launch first of the L-band MSS Super-GSOs - built by Lockheed Martin
  - Iridium Bankruptcy - the first major failure?
- 2000
  - Globalstar begins service
  - Thuraya launch L-band MSS Super-GSO
- 2001
  - XM Satellite Radio begins service
  - Pakistan's 2nd Satellite, BADR-B was launched on 10 Dec 2001 at 9:15a from Baikonour Cosmodrome, Kazakistan
- 2002
  - Sirius Satellite Radio begins service
  - Paksat-1, was deployed at 38 degrees E orbital slot in December 2002, Paksat-1, was deployed at 38 degrees E orbital slot in December 2002
- 2004
  - Teledesic network planned to start operation
- 2005
  - Intelsat and Panamsat Merge
  - VUSat OSCAR-52 (HAMSAT) Launched
- 2006
  - CubeSat-OSCAR 56 (Cute-1.7) Launched
  - K7RR-Sat launched by California Politechnic University
- 2007
  - Prism was launched by University of Tokyo
- 2008
  - COMPASS-1; a project of Aachen University was launched from Satish Dawan Space Center, India. It failed to achieve orbit.

Motivation to use Satellites

➔ Satellites can relay signals over a long distance
Geostationary Satellites

➔ Remain above the equator at a height of about 22300 miles (geosynchronous orbits)
Travel around the earth in exactly the same time, the earth takes to rotate





- When using a satellite for long distance communications, the satellite acts as a repeater.
- An earth station transmits the signal up to the satellite (uplink), which in turn retransmits it to the receiving earth station (downlink).
- Different frequencies are used for uplink/downlink.

as a consequence of the limited performance of the satellite, it was necessary to use earth stations equipped with large antennas and therefore of high cost (around $10 million for a station equipped with a 30m diameter antenna).

DEVELOPMENT OF SATELLITE COMMUNICATIONS

- The first satellites provided a low capacity at a relatively high cost; for example, INTELSAT I weighed 68 kg at launch for a capacity of 480 telephone channels and an annual cost of $32 500 per channel at the time.
- Launch cost , short life time , Low capacity

- **How to reduce cost?**

1) production of reliable launchers which can put heavier and heavier satellites into orbit (typically 5900 kg at launch in 1975, reaching 10 500 kg by Ariane 5 ECA and 13 000 kg by Delta IV in 2008)

2) increasing expertise in microwave techniques has enabled realisation of contoured multibeam antennas whose beams adapt to the shape of continents

3) frequency re-use from one beam to the other and incorporation of higher power transmission amplifiers

4) increased satellite capacity has led to a reduced cost per telephone channel.

What has Elon Musk Done ??

DEVELOPMENT OF SATELLITE COMMUNICATIONS (Continued)

- The increasing **size and power** of satellites has permitted a consequent reduction in the **size of earth stations**, and hence their cost, leading to an increase in number.

➔ability to collect or broadcast signals from or to several locations

➔multipoint data transmission networks and data collection networks have been developed under the name of VSAT(very small aperture terminals) networks For TV services, satellites are of paramount

➔Satellite news gathering (SNG), for the exchange of programmes between broadcasters, for distributing programmes to terrestrial broadcasting stations and cable heads.

➔Direct broadcasting by satellite (DBS) systems, or direct-to-home (DTH) systems. A rapidly growing service is digital

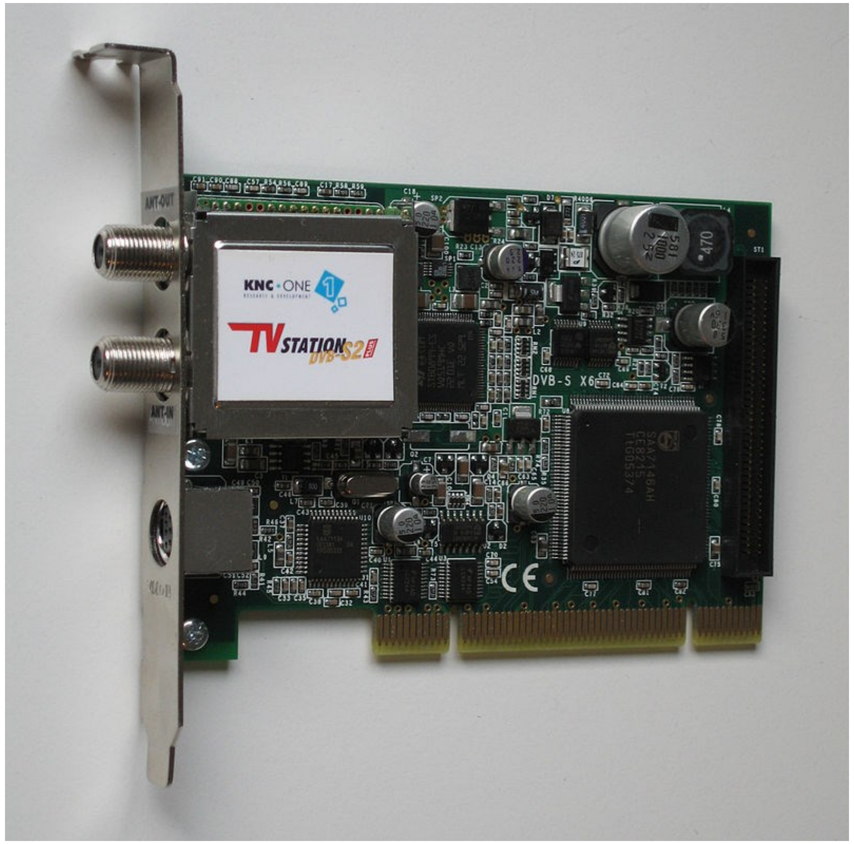
➔video broadcasting by satellite (DVB-S), developed in early 1991; the standard for the second generation (DVB-S2) has been standardised by the European Telecommunication Standard Institute (ETSI). These DBS systems operate with small earth stations having antennas with a diameter from 0.5 to 1 m.

DEVELOPMENT OF SATELLITE COMMUNICATIONS (Continued)

➔ Feedback systems

➔ With the introduction of two-way communications stations, satellites are a key component in providing interactive TV and broadband Internet services thanks to the implementation of the DVB satellite return channel (DVB-RCS) standard to the service provider’s facilities.

➔ This uses TCP/IP to support Internet, multicast and web-page caching services over satellite with forward channel operating at several Mbit/s and enables satellites to provide broadband service applications for the end user, such as direct access and distribution services. IP-based triple-play services (telephony, Internet and TV) are more and more popular.

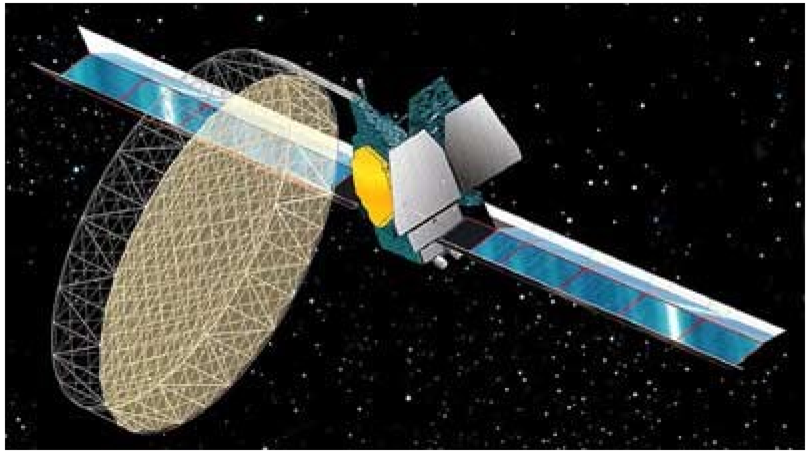


Communication applications in satellites

➔ Since the end of the 1970s,INMARSAT satellites have been providing distress signal services along with telephone and data communications services to ships and planes and, more recently, communications to portable earth stations (Mini M or Satphone).

➔ Personal mobile communication using small handsets is available from constellations of non-geostationary satellites (such as Iridium and Globalstar) and geostationary satellites equipped with very large deployable antennas (typically 10 to 15 m) as with the THURAYA, ACES, and INMARSAT 4 satellites.

➔ The next step in bridging the gaps between fixed, mobile and broadcasting radiocommunications services concerns satellite multimedia broadcast to fixed and mobile users. Satellite digital mobile broadcasting (SDMB) is based on hybrid integrated satellite–terrestrial systems to serve small hand-held terminals with interactivity.



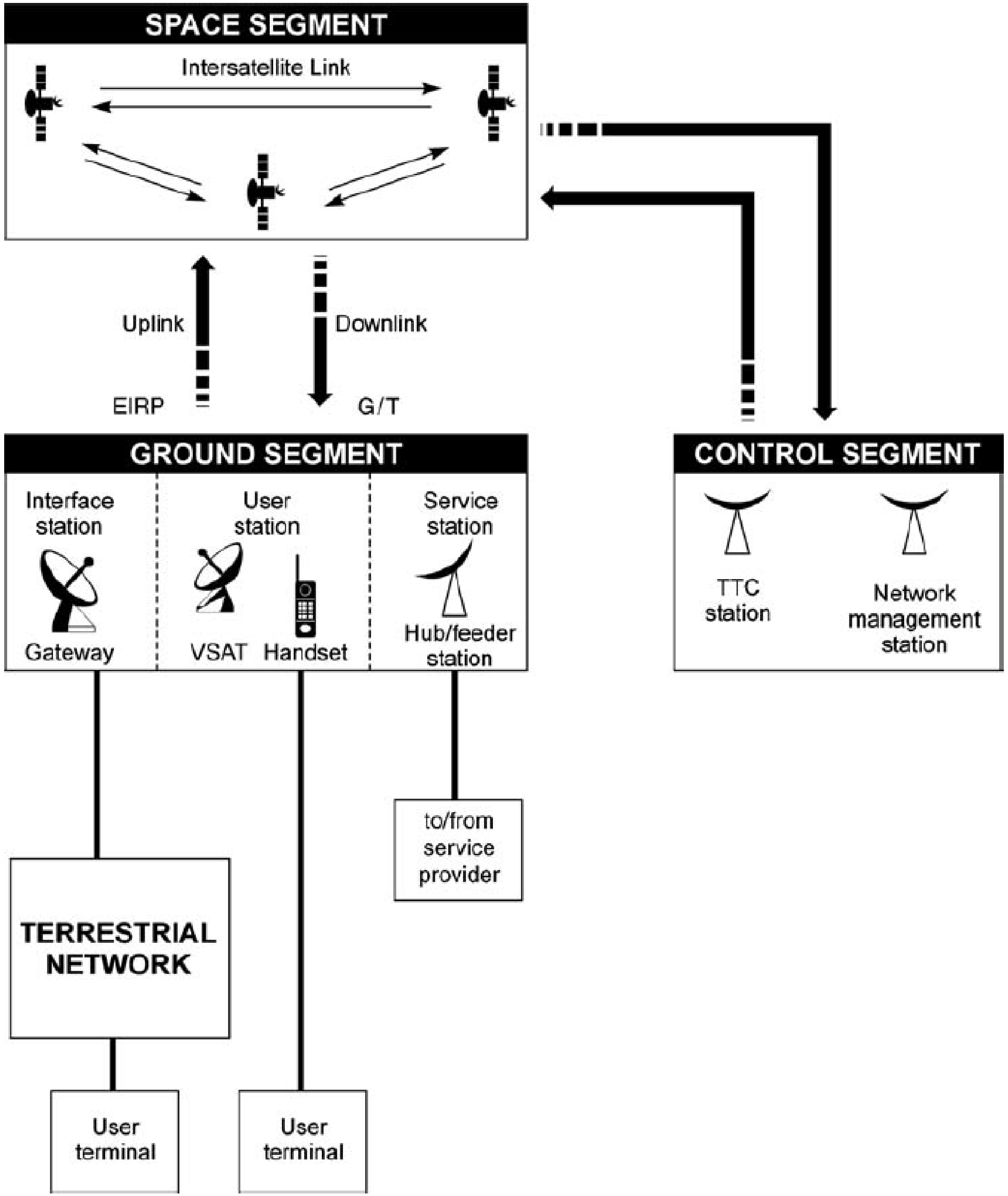
CONFIGURATION OF A SATELLITE COMMUNICATIONS SYSTEM

- The satellite system is composed of a space segment, a control segment and a ground segment:

—The space segment contains one or several active and spare satellites organised into a constellation.

—The control segment consists of all ground facilities for the control and monitoring of the satellites, also named TTC (tracking, telemetry and command) stations, and for the management of the traffic and the associated resources on-board the satellite

—The ground segment consists of all the traffic earth stations. Depending on the type of service considered, these stations can be of different size, from a few centimetres to tens of metres.



Communications links

Definition: A link between transmitting equipment and receiving equipment consists of a radio or optical modulated carrier.

➔ The performance of the transmitting equipment is measured by its effective isotropic radiated power (EIRP), which is the power fed to the antenna multiplied by the gain of the antenna in the considered direction.

➔The performance of the receiving equipment is measured by G/T, the ratio of the antenna receive gain, G, in the considered direction and the system noise temperature, T; (G/T) is called the receiver’s figure of merit.

➔The types of link are:

- —the uplinks from the earth stations to the satellites;
- —the downlinks from the satellites to the earth stations;
- —the intersatellite links, between the satellites.

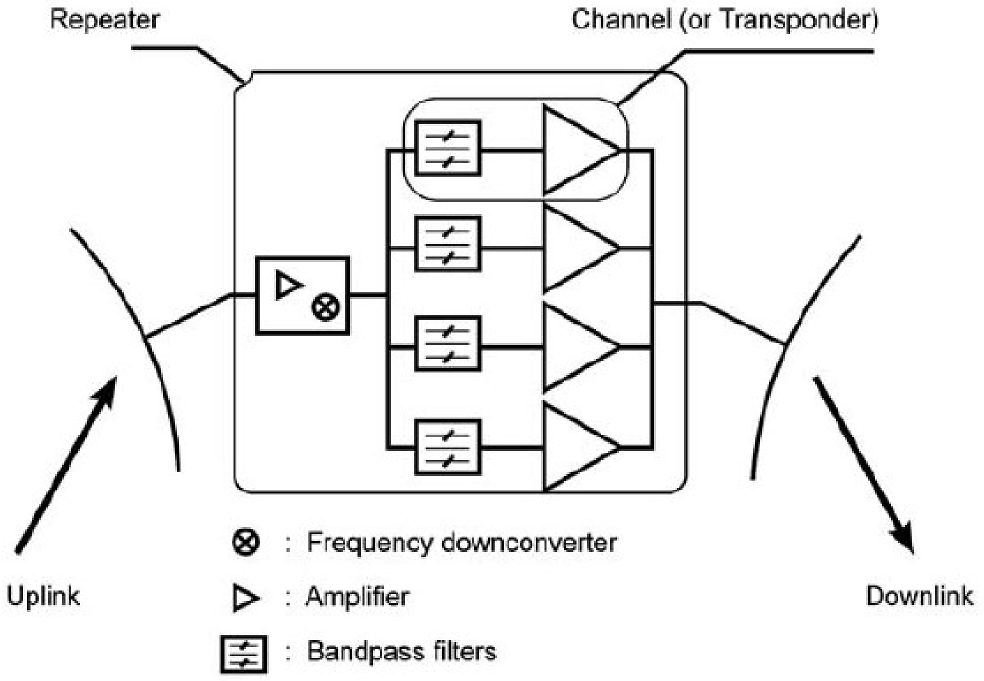
The link performance can be measured by the ratio of the received carrier power, C, to the noise power spectral density, N_0, and is denoted as the C/N_0 ratio, expressed in hertz (Hz).

The values of C/N_0, for the links which participate in the connection between the end terminals, determine the quality of service, specified in terms of bit error rate (BER) for digital communications.
Another parameter of importance for the design of a link is the bandwidth, B, occupied by the carrier.

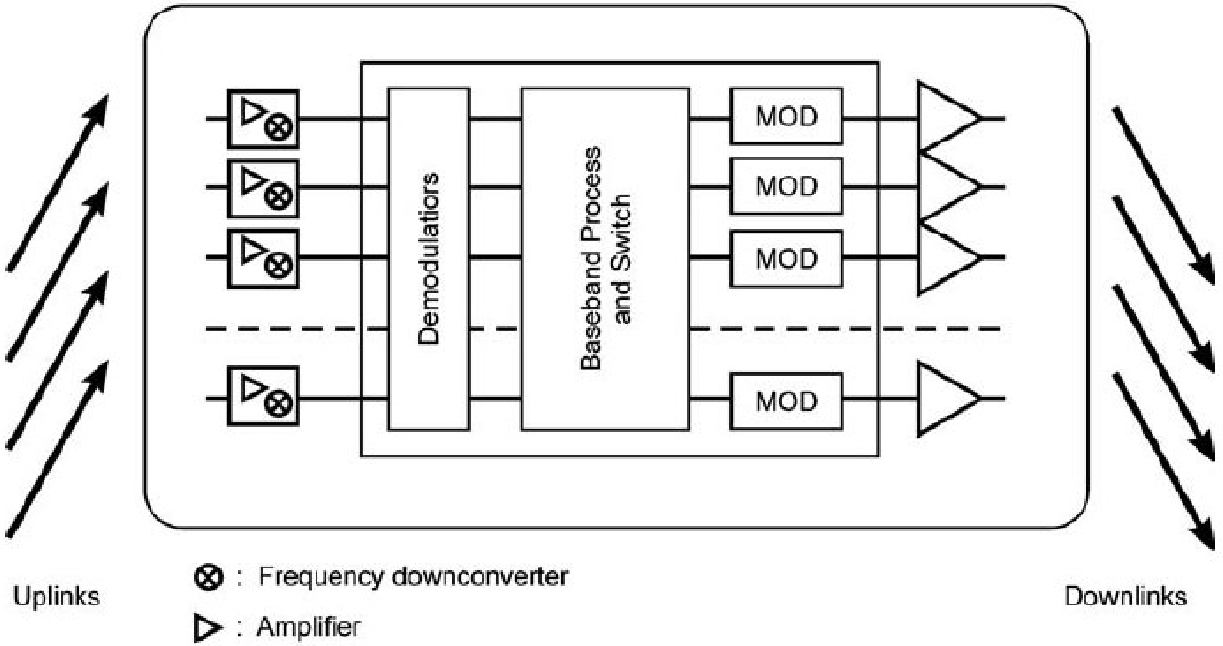
This bandwidth depends on the information data rate, the channel coding rate (forward error correction) and the type of modulation used to modulate the carrier. For satellite links, the trade-off between required carrier power and occupied bandwidth is paramount to the cost-effective design of the link.

The space segment

- The satellite consists of the payload and the platform. The payload consists of the receiving and transmitting antennas and all the electronic equipment which supports the transmission of the carriers. The two types of payload organisation are illustrated next figure.
- Figure (a) shows a transparent payload (sometimes called a 'bent pipe' type) where carrier power is amplified and frequency is downconverted. Power gain is of the order of 100–130 dB, required to raise the power level of the received carrier from a few tens of picowatts to the power level of the carrier fed to the transmit antenna of a few watts to a few tens of watts.
- Frequency conversion is required to increase isolation between the receiving input and the transmitting output. Due to technology power limitations, the overall satellite payload bandwidth is split into several sub-bands, the carriers in each sub-band being amplified by a dedicated power amplifier.
- The amplifying chain associated with each sub-band is called a satellite channel, or transponder. The bandwidth splitting is achieved using a set of filters called the input multiplexer (IMUX). The amplified carriers are recombined in the output multiplexer (OMUX).



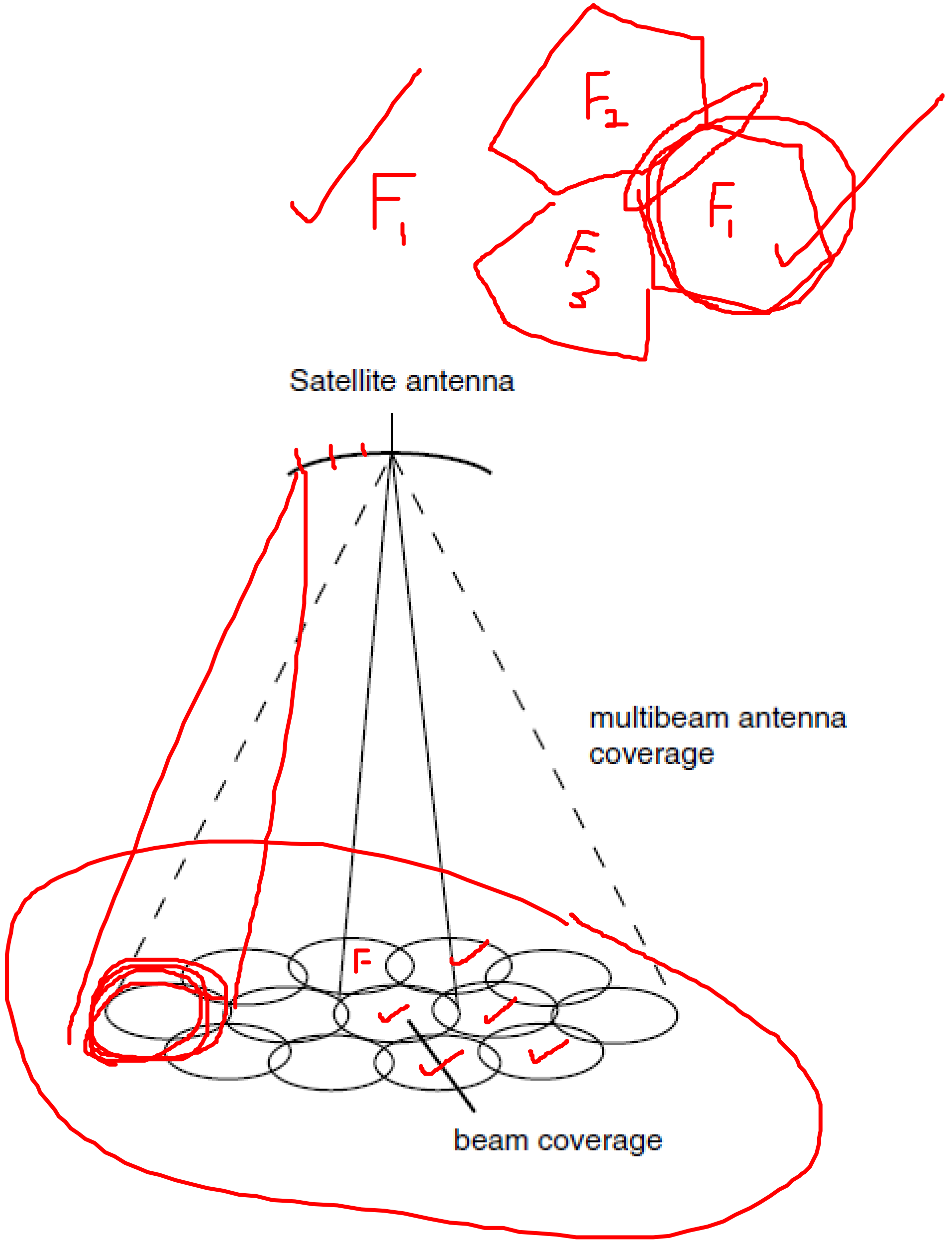
(a)



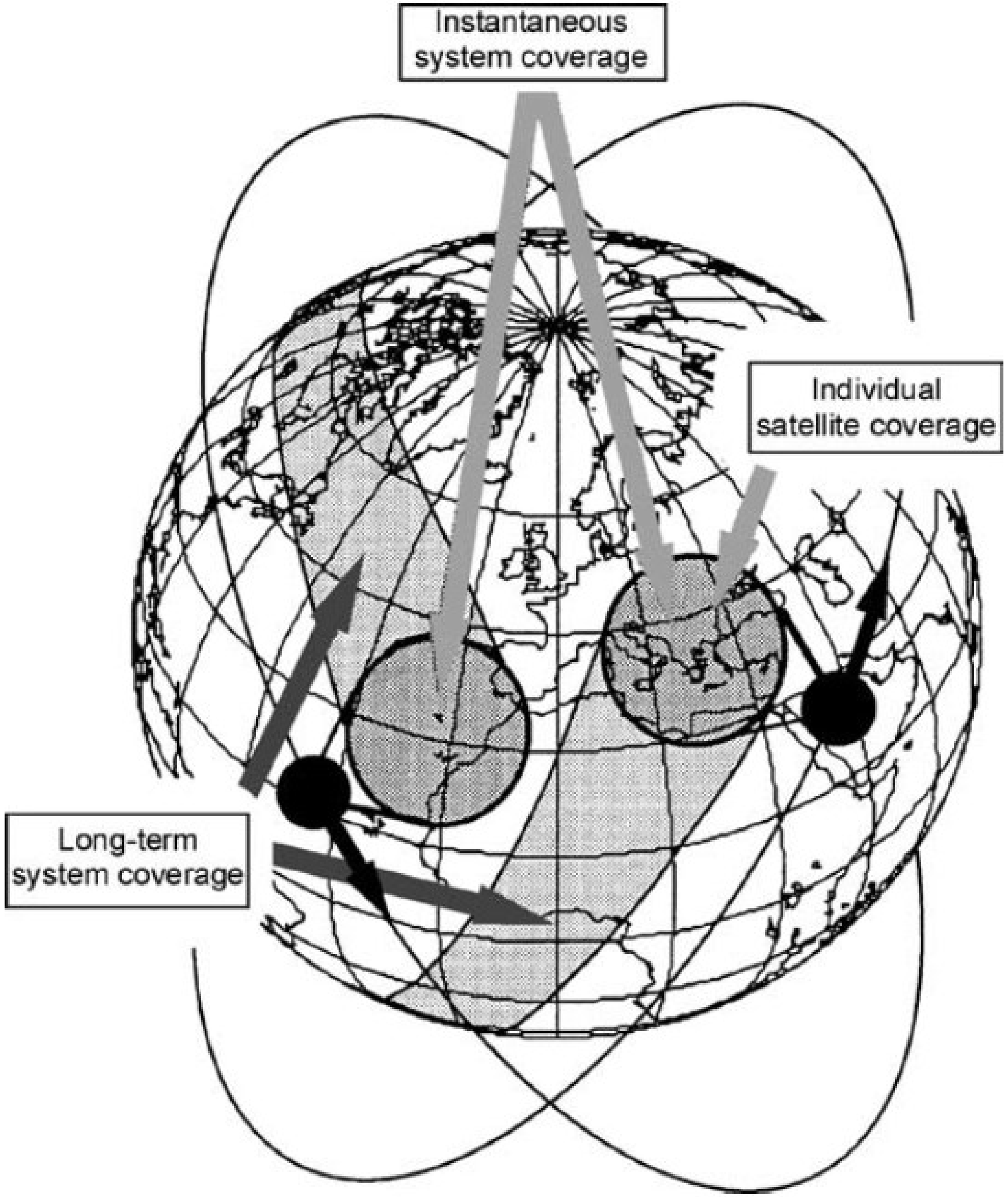
- The transparent payload in Figure 1.2a belongs to a single beam satellite where each transmit and receive antenna generates one beam only. One could also consider multiple beam antennas. The payload would then have as many inputs/outputs as upbeams/downbeams.
- Routing of carriers from one upbeam to a given downbeam implies either routing through different satellite channels, transponder hopping, depending on the selected uplink frequency or on-board switching with transparent on-board processing.

Figure (b) shows a multiple beam regenerative payload where the uplink carriers are demodulated.
The availability of the baseband signals allows on-board processing and routing of information from upbeam to downbeam through on-board switching at baseband. The frequency conversion is achieved by modulating on-board-generated carriers at downlink frequency. The modulated carriers are then amplified and delivered to the destination downbeam.

- Figure 1.3 illustrates a multiple beam satellite antenna and its associated coverage areas. Each beam defines a beam coverage area, also called footprint, on the earth surface. The aggregate beam coverage areas define the multibeam antenna coverage area. A given satellite may have several multiple beam antennas, and their combined coverage defines the satellite coverage area.

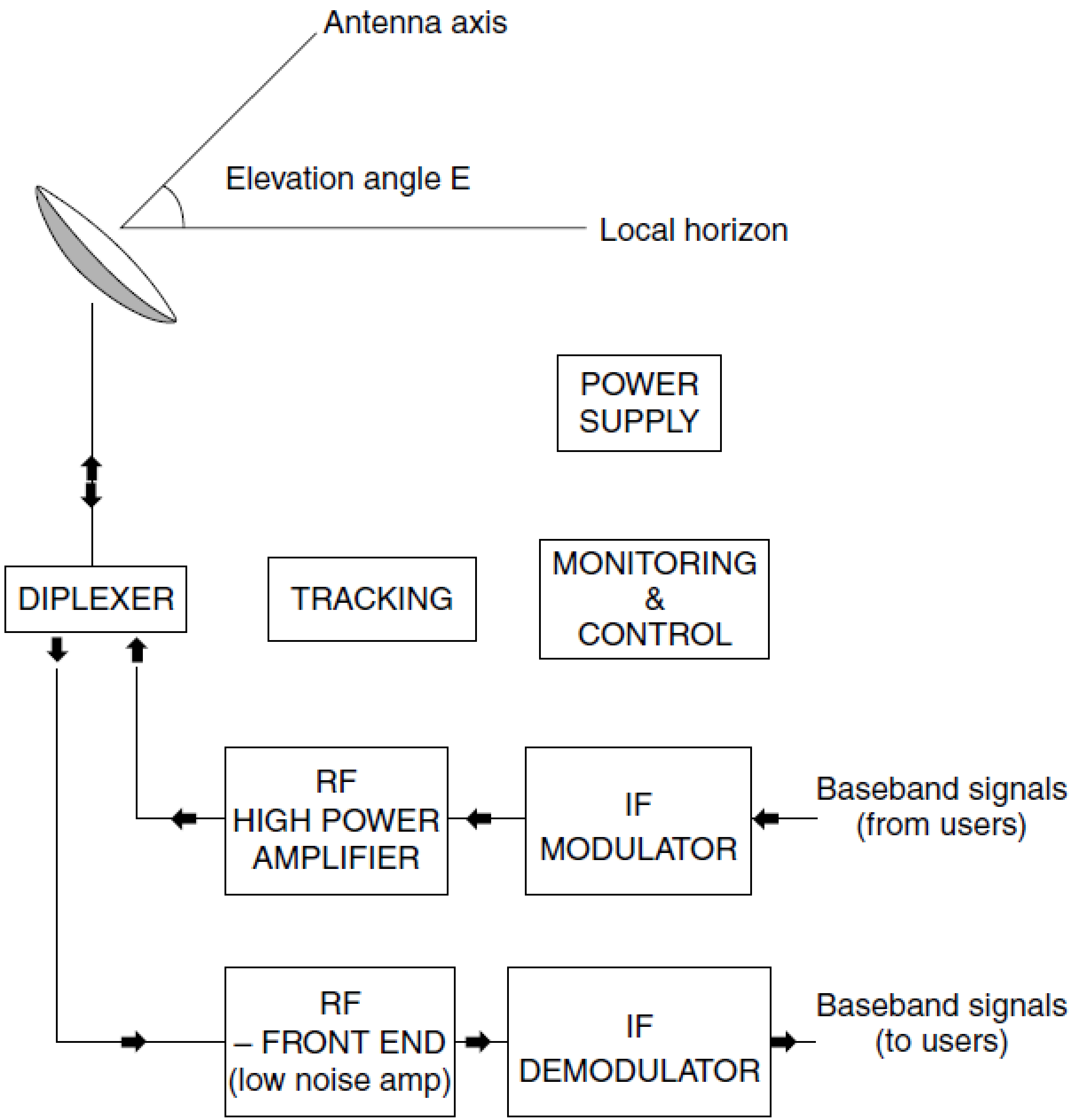


- Next figure illustrates the concept of instantaneous system coverage and long-term coverage. The
- instantaneous system coverage consists of the aggregation at a given time of the coverage areas of the individual satellites participating in the constellation. The long-term coverage is the area on the earth scanned over time by the antennas of the satellites in the constellation.
- The coverage area should encompass the service zone, which corresponds to the geographical region where the stations are installed. For **real-time services**, the instantaneous system coverage should at any time have a footprint covering the service zone, while for **non-real-time** (store-and-forward) services, it should have long-term coverage of the service zone.



The ground segment

- The ground segment consists of all the earth stations; these are most often connected to the end user's terminal by a terrestrial network or, in the case of small stations (Very Small Aperture Terminal, VSAT), directly connected to the end-user's terminal.
- Stations are distinguished by their size which varies according to the volume of traffic to be carried on the satellite link and the type of traffic (telephone, television or data).
- In the past, the largest were equipped with antennas of 30 m diameter (Standard A of the INTELSAT network). The smallest have 0.6 m antennas (receiving stations from direct broadcasting satellites) or even smaller (0.1 m) antennas (mobile stations, portable stations or handsets).Some stations both transmit and receive. Others are receive only (RCVO) stations; this is the case, for example, with receiving stations for a broadcasting satellite system or a distribution system for television or data signals.
- Next Figure shows the typical architecture of an earth station for both transmission and reception.



The organisation of an earth station. RF = radio frequency, IF = intermediate frequency.

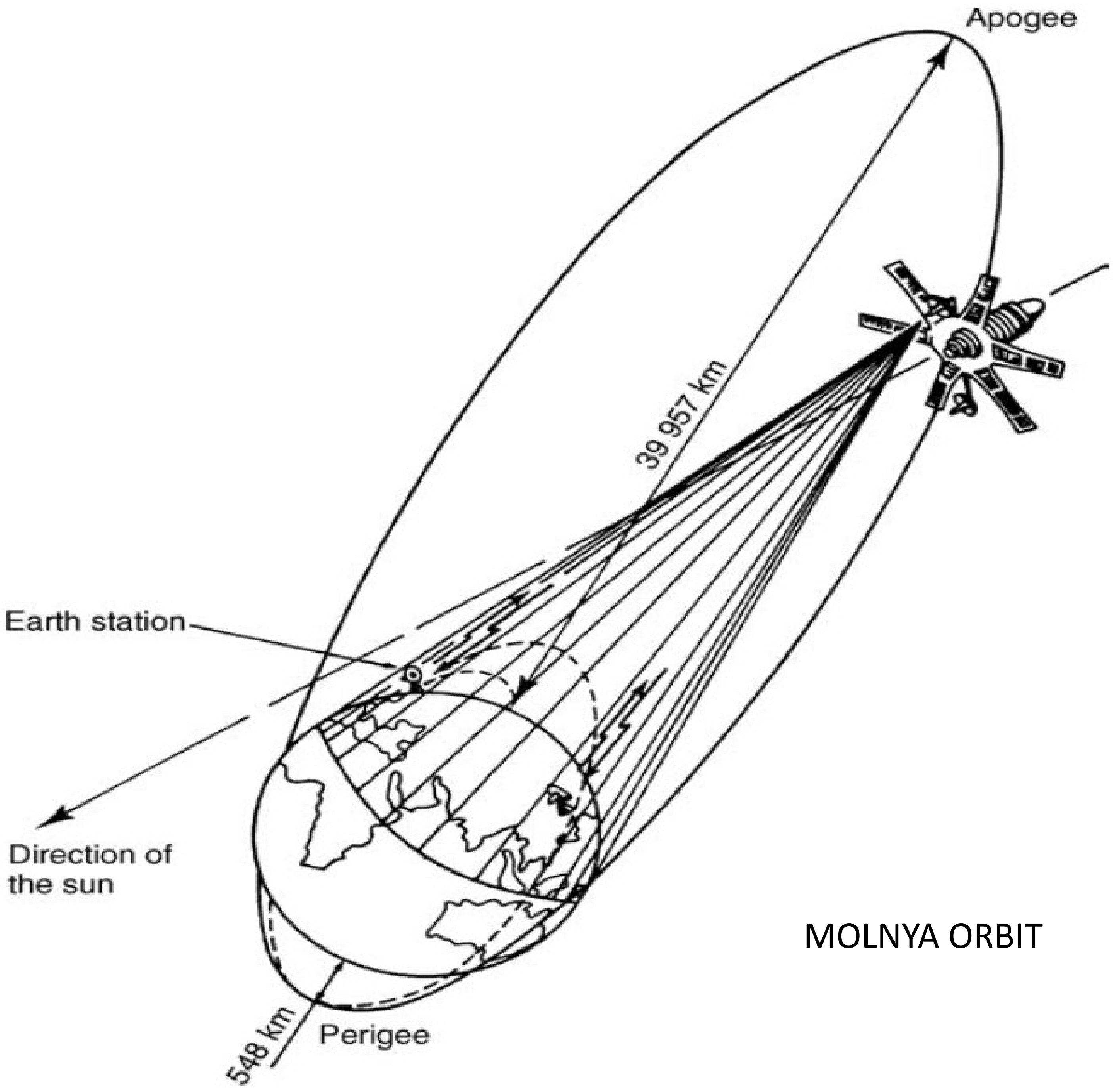
TYPES OF ORBIT

- The orbit is the trajectory followed by the satellite. The trajectory is within a plane and shaped as an ellipse with a maximum extension at the apogee and a minimum at the perigee.
- Types of Orbits:

1) **Elliptical orbits inclined at an angle of 64 with respect to the equatorial plane**. This type of orbit is particularly stable with respect to irregularities in terrestrial gravitational potential and, owing to its inclination, enables the satellite to cover regions of high latitude for a large fraction of the orbital period as it passes to the apogee.

➔ This type of orbit has been adopted by the USSR for the satellites of the MOLNYA system with period of 12 hours. Geometry of the orbit is shown next figure. The satellite remains above the regions located under the apogee for a time interval of the order of **8 hours**. Continuous coverage can be ensured with three phased satellites on different orbits. Several studies relate to elliptical orbits with a period of **24 h** (TUNDRA orbits) or a multiple of 24 h. These orbits are particularly useful for satellite systems for communication with mobiles where the masking effects caused by surrounding obstacles such as buildings and trees and multiple path effects are pronounced at low elevation angles (say less than 30). **(Remember the STK simulation showed before)**

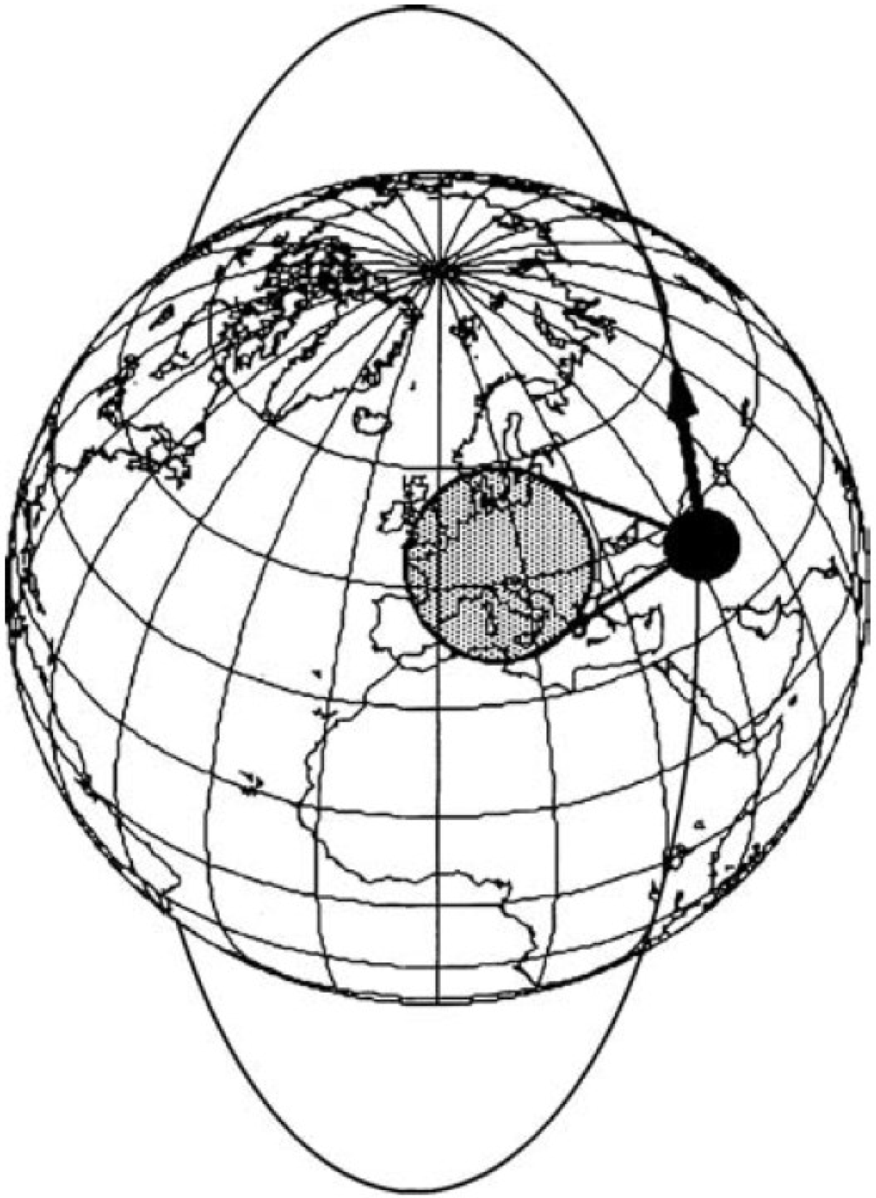
➔ In fact, inclined elliptic orbits can provide the possibility of links at medium latitudes when the satellite is close to the apogee with elevation angles close to 90; these favourable conditions cannot be provided at the same latitudes by geostationary satellites. In the late 1980s, the European Space Agency (ESA) studied the use of elliptical highly inclined orbits (HEO) for digital audio broadcasting (DAB) and mobile communications in the framework of its Archimedes programme. The concept became reality at the end of the 1990s with the Sirius system delivering satellite digital audio radio services to millions of subscribers (mainly automobiles) in the United States using three satellites on HEO Tundra-like orbits.



MOLNYA ORBIT

2) **Circular low earth orbits (LEO).** The altitude of the satellite is constant and equal to several hundreds of kilometres. The period is of the order of one and a half hours. With near 90 inclination, this type of orbit guarantees worldwide long term coverage as a result of the combined motion of the satellite and earth rotation, as shown in next figure. This is the reason for choosing this type of orbit for observation satellites (for example, the SPOT satellite: altitude 830 km, orbit inclination 98.7, period 101 minutes). One can envisage the establishment of store-and-forward communications if the satellite is equipped with a means of storing information.

➔A constellation of several tens of satellites in low altitude (e.g. IRIDIUM with 66 satellites at 780 km) circular orbits can provide worldwide real-time communication. Non-polar orbits with less than 90 inclination, can also be envisaged. For instance the GLOBALSTAR constellation incorporates 48 satellites at 1414km with 52 orbit inclination.



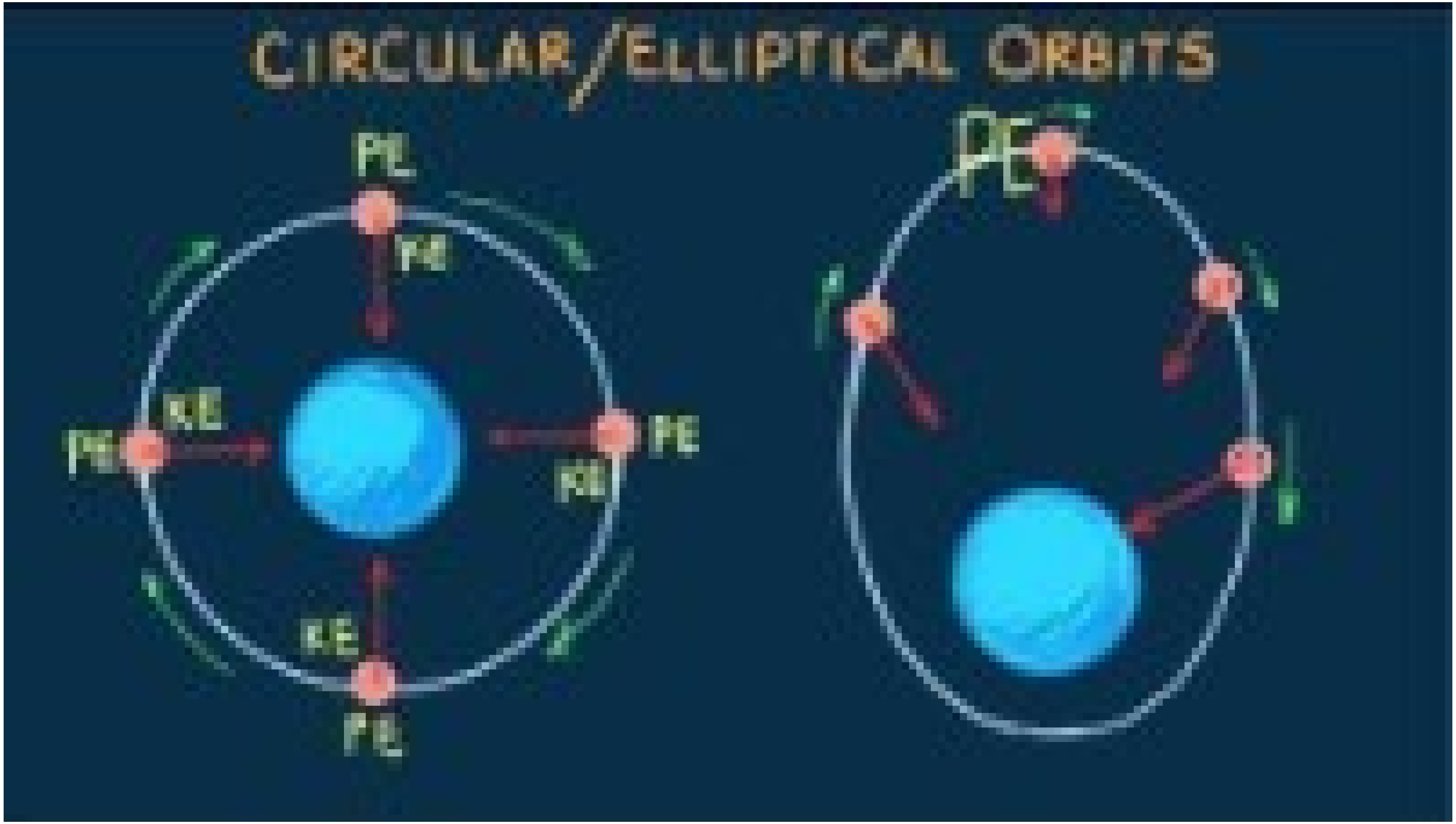
3) **Circular medium earth orbits (MEO),** also called intermediate circular orbits (ICO), have an altitude of about 10 000km and an inclination of about 50. The period is 6 hours. With constellations of about 10 to 15 satellites, continuous coverage of the world is guaranteed, allowing worldwide real-time communications.

A planned system of this kind was the ICO system (which emerged from Project 21 of INMARSAT but was not implemented) with a constellation of 10 satellites in two planes at 45 inclination.

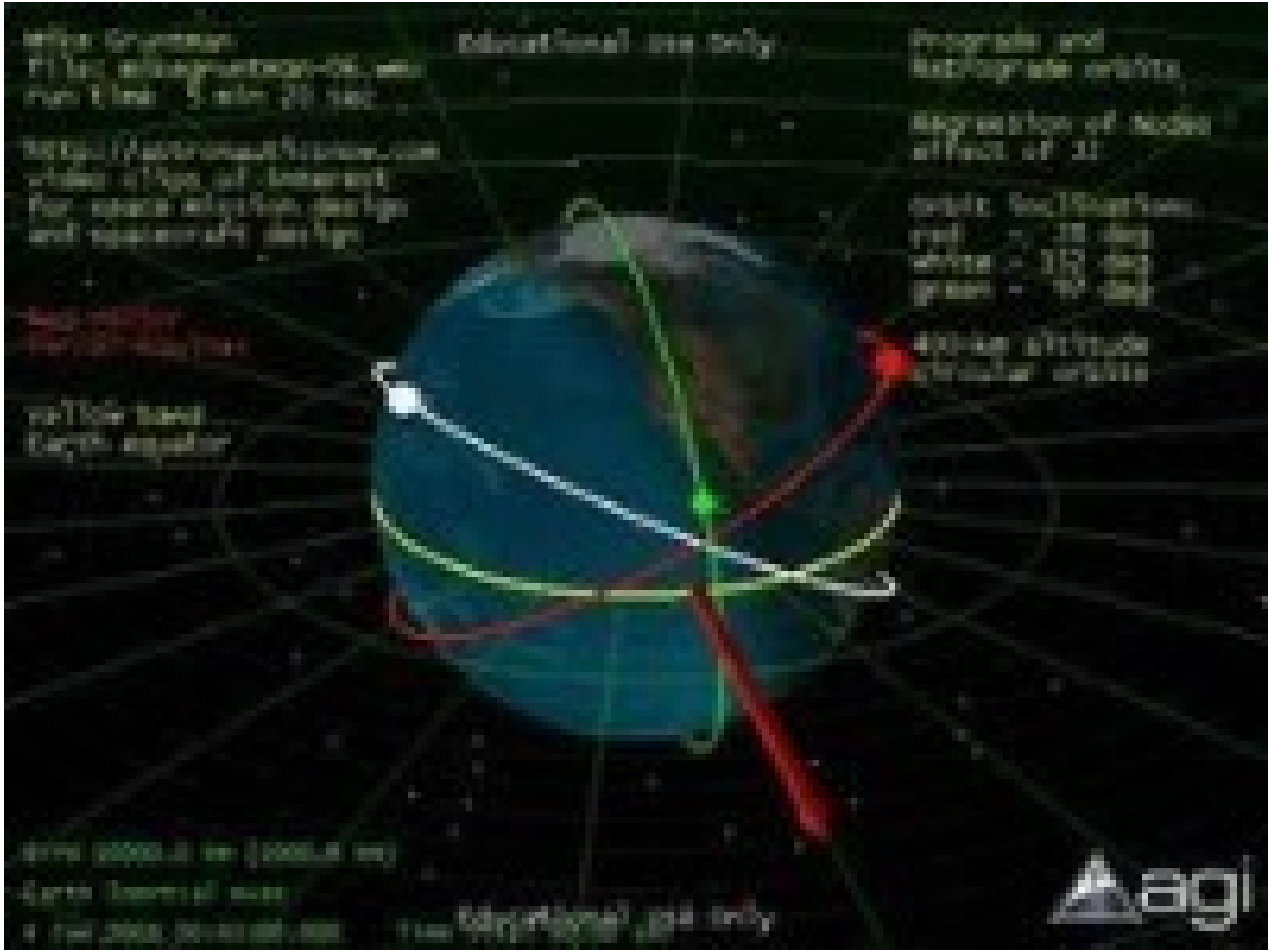
4) **Circular orbits with zero inclination (equatorial orbits).** The most popular is the geostationary satellite orbit; the satellite orbits around the earth in the equatorial plane according to the earth rotation at an altitude of 35 786 km. The period is equal to that of the rotation of the earth. The satellite thus appears as a point fixed in the sky and ensures continuous operation as a radio relay in real time for the area of visibility of the satellite (43% of the earth's surface).

5) **Hybrid systems.** Some systems may include combinations of orbits with circular and elliptical orbits. Such a design was envisaged for the ELLIPSO system.

Circular Vs Elliptical Orbits



Sun Synchronous Orbit



Orbit Choice Factors

1) **The extent and latitude of the area to be covered**

➔ Altitude Vs. Link Budget

➔ Satellite Antenna Gain Vs. Angle of Coverage

➔ Low gain omni-directional antennas Vs Satellite Tracking Devices

➔ Geostationary Satellites for extensive regions

2) **The Elevation Angle**

➔ A satellite in an inclined or polar elliptical orbit can appear overhead at certain times which enables communication to be established in urban areas without encountering the obstacles which large buildings constitute for elevation angles between 0 and approximately 70. With a geostationary satellite, the angle of elevation decreases as the difference in latitude or longitude between the earth station and the satellite increases.

Orbit Choice Factors (Continued)

3) **Transmission duration and delay**

➔the geostationary satellite provides a continuous relay for stations within visibility but the propagation time of the waves from one station to the other is of the order of 250 ms. This requires the use of echo control devices on telephone channels or special protocols for data transmission.

➔ Satellite Altitude Vs Propagation time.

4) **Interference**

➔ RF planning of geostationary satellites

➔The small orbital spacing between adjacent satellites operating at the same frequencies leads to an increase in the level of interference and this impedes the installation of new satellites.

➔Different systems could use different frequencies but this is restricted by the limited number of frequency bands assigned for space radiocommunications by the Radiocommunication Regulations.

➔ With orbiting satellites interference is even more complicated, **Why?**

5) **The performance of launchers**

➔ the mass which can be launched decreases as the altitude increases.

Space Telecommunication Services and Frequency allocation

- **Fixed Satellite Service** (FSS); **Mobile Satellite Service** (MSS); **Broadcasting Satellite Service** (BSS); **Earth Exploration Satellite** Service (EES); **Space Research** Service (SRS); **Space Operation** Service (SOS); **Radiodetermination Satellite** Service (RSS); **Inter-Satellite Service** (ISS); **Amateur Satellite Service** (ASS).

The **fixed satellite service** makes use of the following bands:
—Around 6GHz for the uplink and around 4GHz for the downlink (systems described as **6/4GHz or C band**). These bands are occupied by the oldest systems (such as INTELSAT, American domestic systems etc.) and tend to be saturated.
—Around 8GHz for the uplink and around 7GHz for the downlink (systems described as **8/7GHz or X band**). These bands are reserved, by agreement between administrations, for government use.
—Around 14GHz for the uplink and around 12GHz for the downlink (systems described as **14/12GHz or Ku band**). This corresponds to current operational developments (such as EUTELSAT, etc.).
—Around 30GHz for the uplink and around 20GHz for the downlink (systems described as **30/20GHz or Ka band**). These bands are raising interest due to large available bandwidth and little interference due to present rather limited use.
—VHF (**very high frequency, 137–138MHz downlink and 148–150MHz uplink**) and UHF (**ultra high frequency, 400–401MHz downlink and 454–460MHz uplink**). These bands are for nongeostationary systems only.
—About **1.6GHz for uplinks and 1.5GHz for downlinks**, mostly used by geostationary systems such as INMARSAT; and 1610–1626.5MHz for the uplink of non-geostationary systems such as GLOBALSTAR.
—About 2.2GHz for downlinks and 2GHz for uplinks for the satellite component of IMT2000 (International Mobile Telecommunications).
—About 2.6GHz for uplinks and 2.5GHz for downlinks.
—Frequency bands have also been allocated at higher frequencies such as Ka band.

Radiocommunications service	Typical frequency bands for uplink/downlink	Usual terminology
Fixed satellite service (FSS)	6/4 GHz	C band
	8/7 GHz	X band
	14/12–11 GHz	Ku band
	30/20 GHz	Ka band
	50/40 GHz	V band
Mobile satellite service (MSS)	1.6/1.5 GHz	L band
	30/20 GHz	Ka band
Broadcasting satellite service (BSS)	2/2.2 GHz	S band
	12 GHz	Ku band
	2.6/2.5 GHz	S band

Orbits and Kepler’s Laws

- Keplerian Orbits: named after Kepler who first found that planets trajectories around the sun are eclipses rather than circles.
- Kepler Laws:

(a) The planets move in a plane; the orbits described are ellipses with the sun at one focus (1602).

(b) The vector from the sun to the planet sweeps equal areas in equal times (the law of areas, 1605).

(c) The ratio of the square of the period T of revolution of a planet around the sun to the cube of the semi-major axis a of the ellipse is the same for all planets (1618).

Newton’s Universal Law:

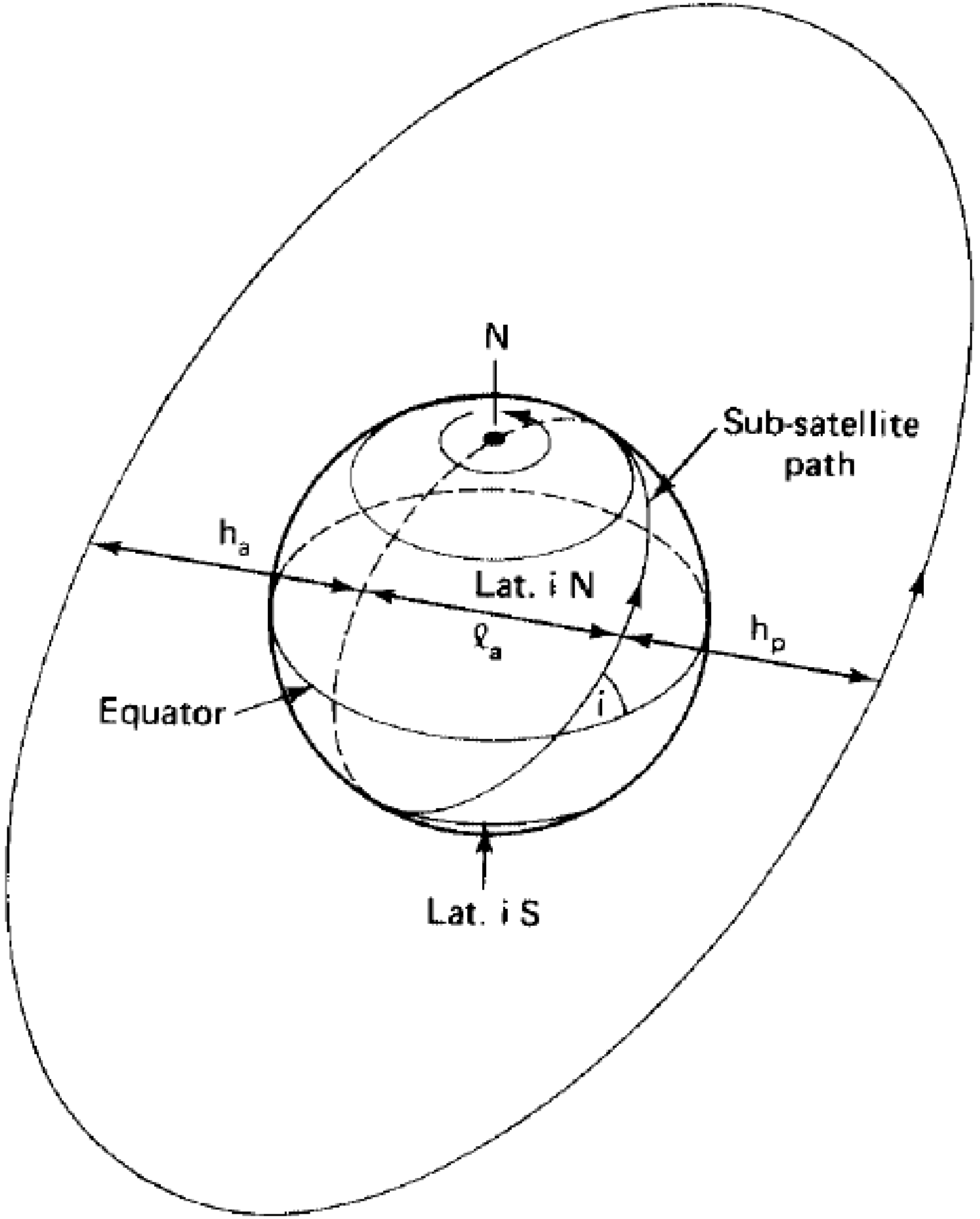
- Newton extended the work of Kepler and, in 1667, discovered the universal law of gravitation. This law states that two bodies of mass (m) and (M) attract each other with a force which is proportional to their masses and inversely proportional to the square of the distance (r) between them:

$$F = GM\, m/r^2$$

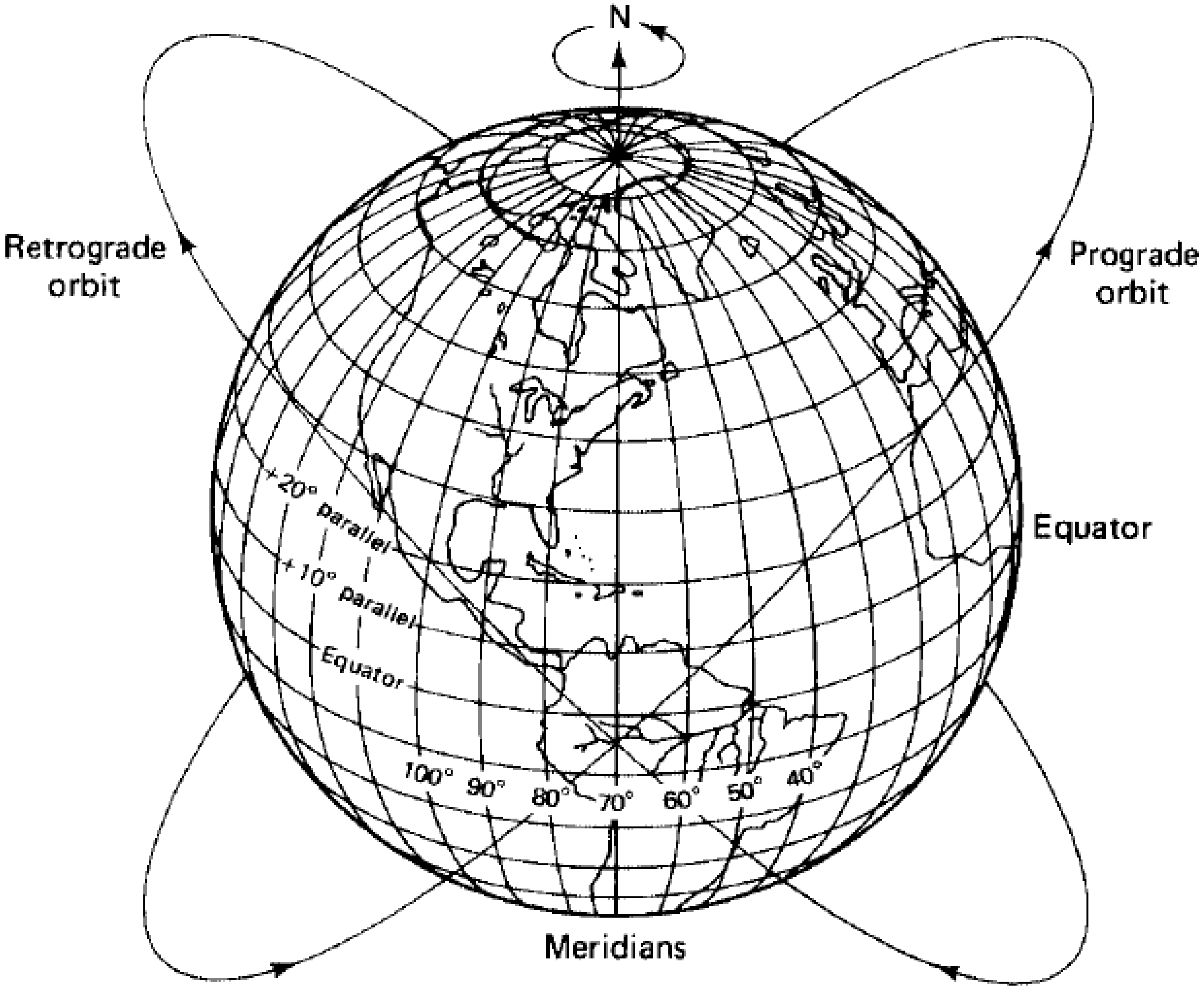
Where G is a constant, called the universal gravitation constant $G = 6.672 \times 10^{-11}\ \mathrm{m^3\,kg^{-1}\,s^{-2}}$.

Definitions of Terms for Earth-Orbiting Satellites

- **Apogee** The point farthest from earth. Apogee height is shown as *ha* in next figure
- **Perigee** The point of closest approach to earth. The perigee height is shown as *hp* in next slide
- **Line of apsides** The line joining the perigee and apogee through the center of the earth.
- **Ascending node** The point where the orbit crosses the equatorial plane going from south to north.
- **Descending node** The point where the orbit crosses the equatorial plane going from north to south.
- **Line of nodes** The line joining the ascending and descending nodes through the center of the earth.
- **Inclination** The angle between the orbital plane and the earth's equatorial plane. It is measured at the ascending node from the equator to the orbit, going from east to north. The inclination is shown as *i* in next slide, It will be seen that the greatest latitude, north or south, is equal to the inclination.



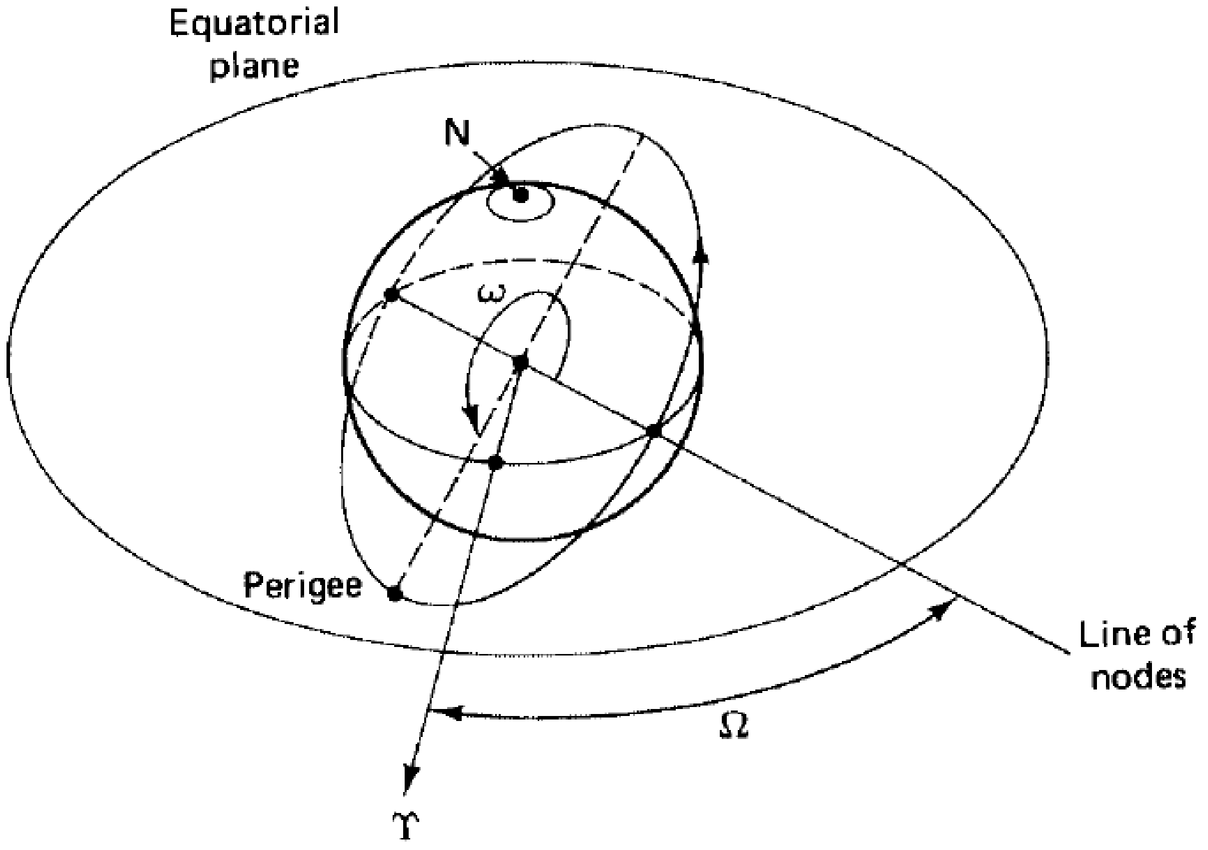
- **Prograde orbit** An orbit in which the satellite moves in the same direction as the earth's rotation. The prograde orbit is also known as a *direct orbit.* The inclination of a prograde orbit always lies between 0 and 90°. Most satellites are launched in a prograde orbit because the earth's rotational velocity provides part of the orbital velocity with a consequent saving in launch energy.
- **Retrograde orbit** An orbit in which the satellite moves in a direction counter to the earth's rotation. The inclination of a retrograde orbit always lies between 90 and 180°.



Argument of perigee The angle from ascending node to perigee, measured in the orbital plane at the earth's centre, in the direction of satellite motion (w in next figure).

Right ascension of the ascending node (RAAN) To define completely the position of the orbit in space, the position of the ascending node is specified. However, because the earth spins, while the orbital plane remains stationary (slow drifts which do occur are discussed later), the longitude of the ascending node is not fixed, and it cannot be used as an absolute reference.

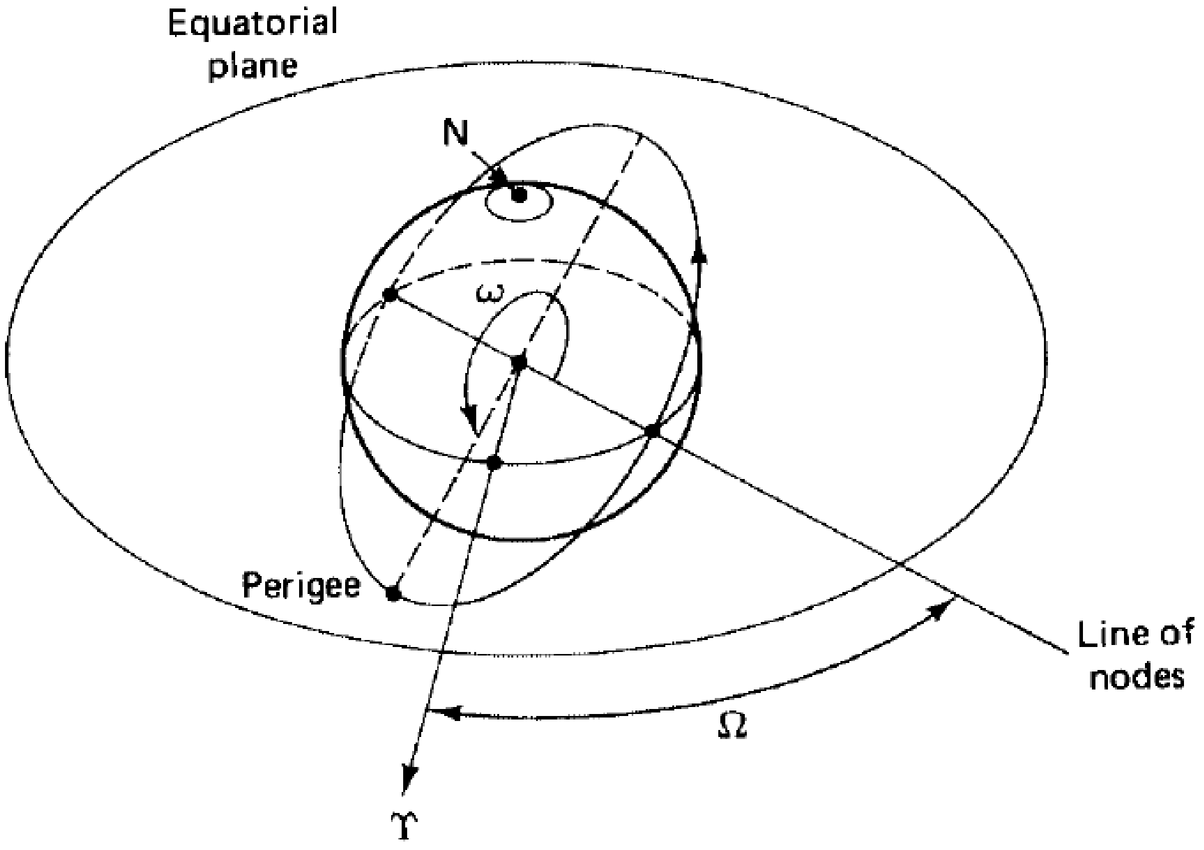
- For the practical determination of an orbit, the longitude and time of crossing of the ascending node are frequently used. However, for an absolute measurement, a fixed reference in space is required. The reference chosen is the *first point of Aries,* otherwise known as the vernal, or spring, equinox.



Reference position of the satellite

The vernal equinox occurs when the sun crosses the equator going from south to north, and an imaginary line drawn from this equatorial crossing through the center of the sun points to the first point of Aries (symbol ϒ); This is the *line of Aries*.

- The right ascension of the ascending node is then the angle measured eastward, in the equatorial plane, from the line to the ascending node, shown as Ωin next figure



Reference position of the satellite

- **Mean anomaly** Mean anomaly *M* gives an average value of the angular position of the satellite with reference to the perigee.

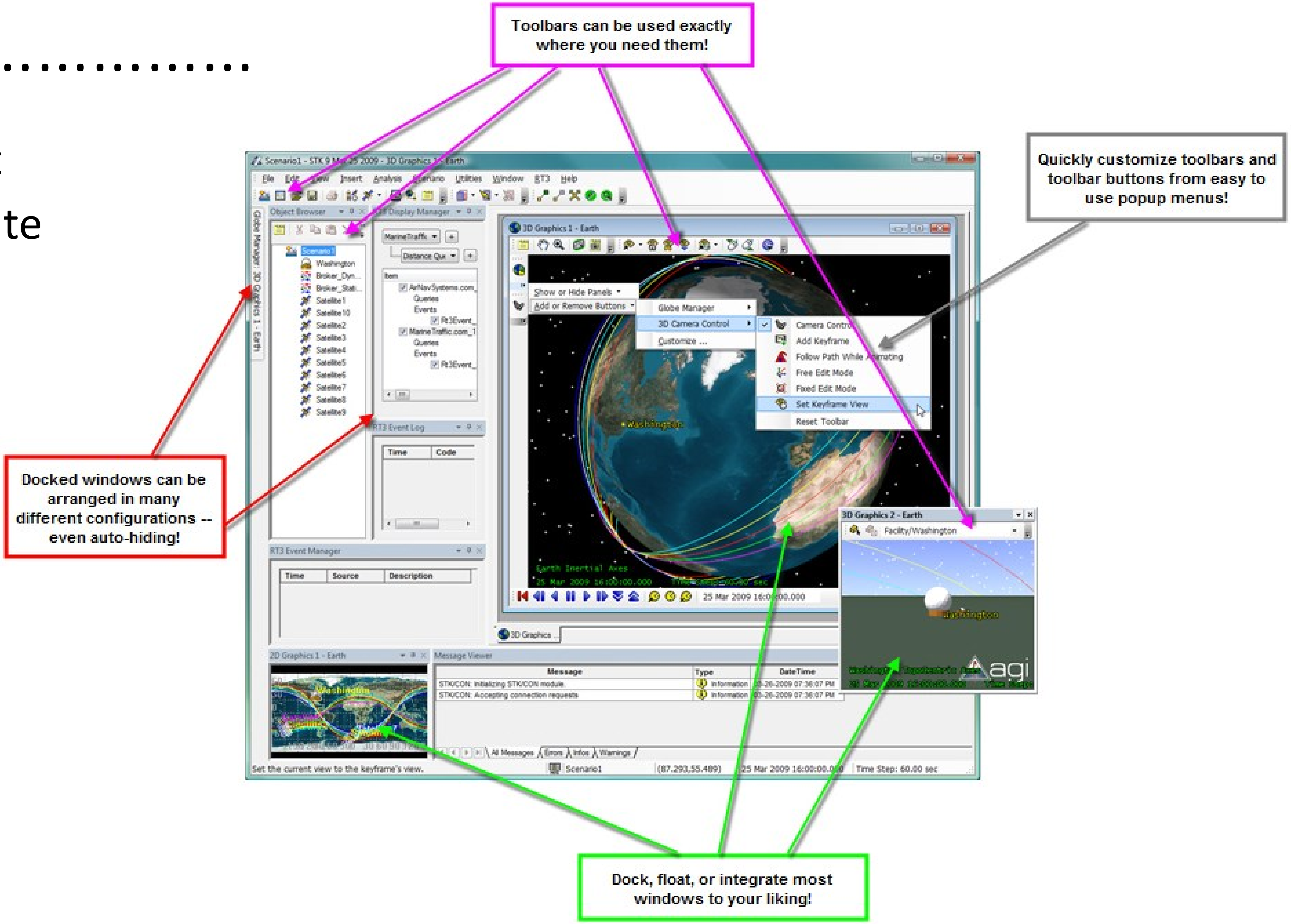
➔ For a circular orbit, *M* gives the angular position of the satellite in the orbit.

➔For elliptical orbit, the position is much more difficult to calculate, and *M* is used as an intermediate step in the calculation

- **True anomaly** The true anomaly is the angle from perigee to the satellite position, measured at the earth's center. This gives the true angular position of the satellite in the orbit as a function of time.

STK Tutorial...............

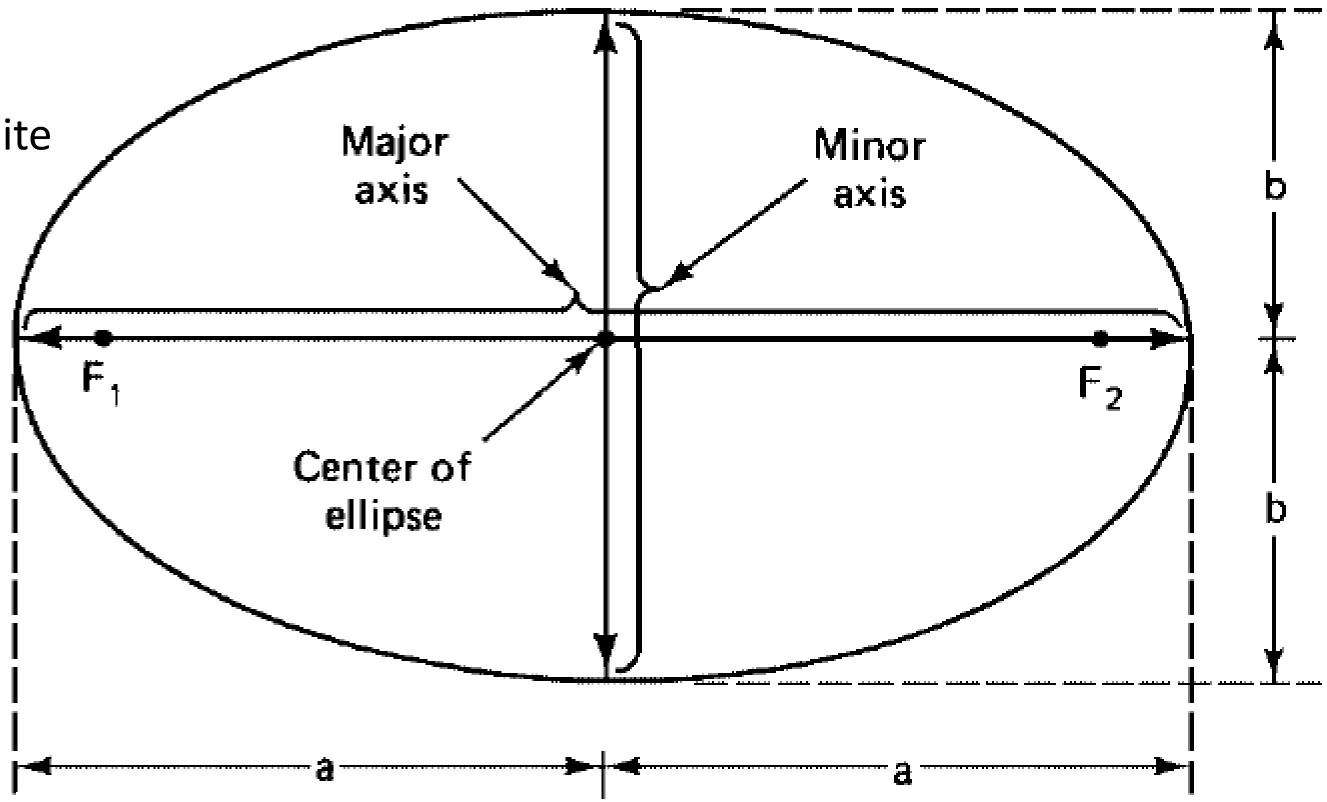
- Simulation of Orbit
- Definition of Satellite
- Access time
- Molnya coverage



Kepler's First Law

- *Kepler's first law* states that the path followed by a satellite around the primary will be an ellipse. An ellipse has two focal points shown as *F*1 and *F*2

- The center of mass of the two-body system, termed the *barycenter,* is always centered on one of the foci. In our specific case, because of the enormous difference between the masses of the earth and the satellite, the center of mass coincides with the center of the earth, which is therefore always at one of the foci.



The semimajor axis of the ellipse is denoted by *a,* and the semiminor axis, by *b.* The eccentricity *e* is given by:

$$\sqrt{1 - \frac{b^2}{a^2}}$$

The eccentricity and the semimajor axis are two of the orbital parameters specified for satellites (spacecraft) orbiting the earth. For an elliptical orbit, $0 < e < 1$. When $e = 0$, the orbit becomes circular.

Question:
Find the eccentricity of the ellipse with the given equation $9x^2 + 25y^2 = 225$
Solution:
Given :

$9x^2 + 25y^2 = 225$

The general form of ellipse is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

To make it in general form, divide both sides by 225, we get: $\frac{x^2}{25} + \frac{y^2}{9} = 1$

So, the value of a = 5 and b = 3
From the formula of the eccentricity of an ellipse $e = \sqrt{1 - \frac{b^2}{a^2}}$
Substituting a = 5 and b = 3,

$$e = \sqrt{1 - \frac{3^2}{5^2}} = \sqrt{\frac{25-9}{25}} = \sqrt{\frac{16}{25}}$$

e = 4/ 5

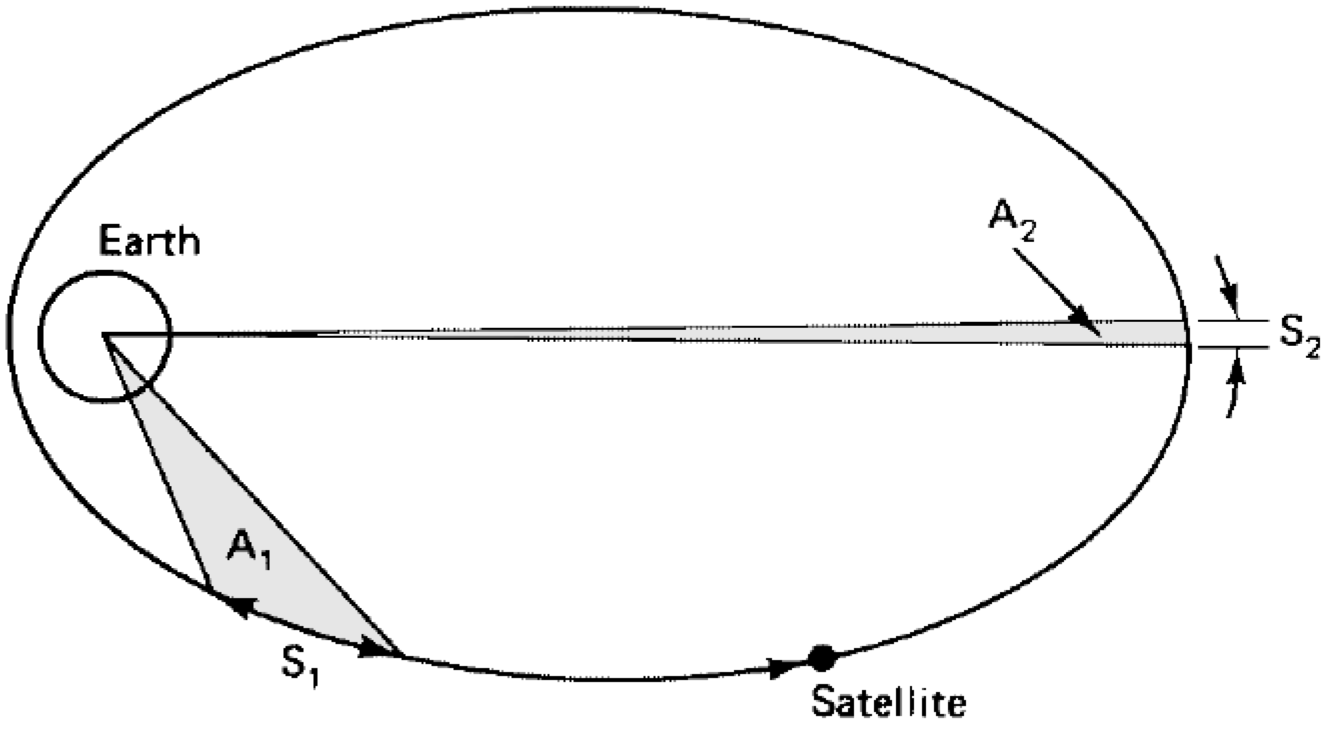
Therefore, the eccentricity of the given ellipse is 4/5.

Kepler’s Second Law

- *Kepler’s second law* states that, for equal time intervals, a satellite will sweep out equal areas in its orbital plane, focused at the barycenter.
- assuming the satellite travels distances *S*1 and *S*2 meters in 1 s, then the areas *A*1 and *A*2 will be equal. The average velocity in each case is *S*1 and *S*2 meters per second, and because of the equal area law, it follows that the velocity at *S*2 is less than that at *S*1.

➔ An important consequence of this is that the satellit takes longer to travel a given distance when it is farther away from earth.

➔ Use is made of this property to increase the length of time a satellite can be seen from particular geographic regions of the earth.



Kepler’s Third Law

- *Kepler’s third law* states that the square of the periodic time of orbit is proportional to the cube of the mean distance between the two bodies. The mean distance is equal to the semimajor axis *a.* For the artificial satellites orbiting the earth, Kepler’s third law can be written in the form:

$$a^3 = \frac{\mu}{n^2}$$

where *n* is the mean motion of the satellite in radians per second and is the earth’s geocentric gravitational constant. With *a* in meters, its value is:

$$\mu = 3.986005 \times 10^{14}\ \text{m}^3/\text{sec}^2$$

➔This applies only to the ideal situation of a satellite orbiting a perfectly spherical earth of uniform mass, with no perturbing forces acting, such as atmospheric drag

➔With *n* in radians per second, the orbital period in seconds is given by:

$$P = \frac{2\pi}{n}$$

The importance of Kepler's third law is that it shows there is a fixed relationship between period and size. One very important orbit in particular, known as the *geostationary orbit,* is determined by the rotational period of the earth. In anticipation of this, the approximate radius of the geostationary orbit is determined in the following example.

Example 2.1 Calculate the radius of a circular orbit for which the period is 1-day.

solution The mean motion, in rad/day, is:

$$n := \frac{2 \cdot \pi}{1\ \text{day}}$$

$$n = 7.272 \cdot 10^{-5} \cdot \frac{\text{rad}}{\text{sec}}$$

The earth's gravitational constant is $\mu := 3.986005 \cdot 10^{14} \cdot \text{m}^3 \cdot \text{sec}^{-2}$

Kepler's third law gives $a := \left(\frac{\mu}{n^2}\right)^{1/3}$

$$a = 42241 \cdot \text{km}$$

Orbital Elements

- Earth-orbiting artificial satellites are defined by six orbital elements referred to as the *keplerian element set:*

1) the semimajor axis *a*

2) the eccentricity *e* ***(Both a and e*** **give the shape of the ellipse*)***

3) the mean anomaly M_0, gives the position of the satellite in its orbit at a reference time known as the *epoch*

4) A fourth, the argument of perigee ω , gives the rotation of the orbit's perigee point relative to the orbit's line of nodes in the earth's equatorial plane.

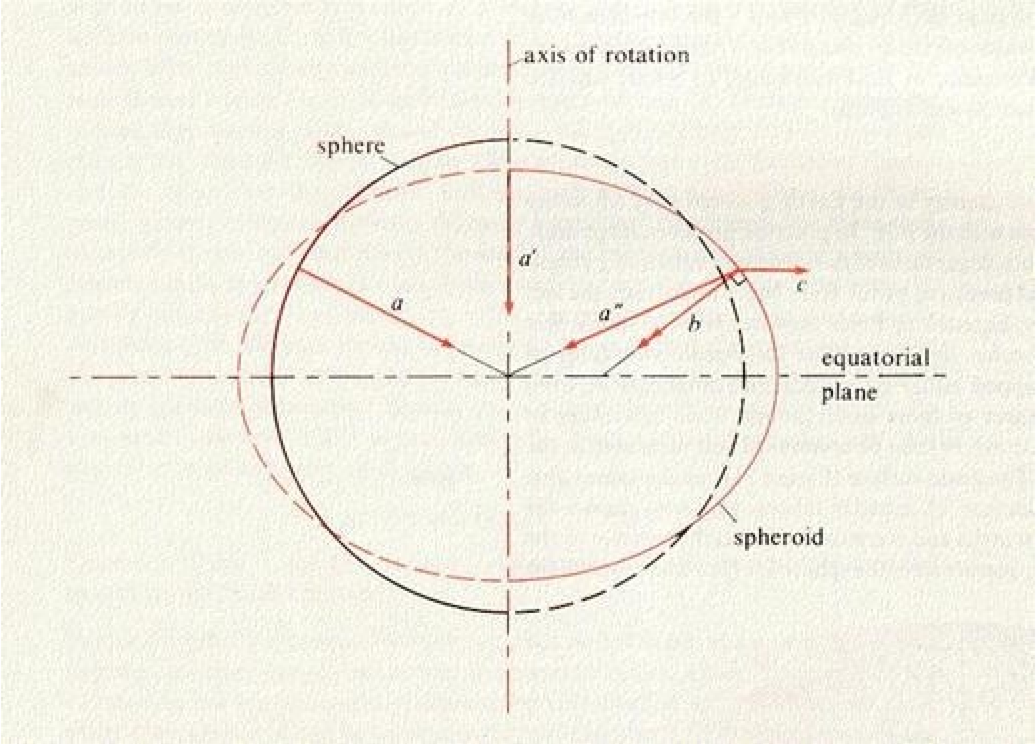
5) the inclination *i*

6) the right ascension of the ascending node Ω , **(both i and Ω relate the orbital plane's position to the earth)**

➔ Because the equatorial bulge causes slow variations in and , and because other perturbing forces may alter the orbital elements slightly, the values are specified for the reference time or epoch, and thus the epoch also must be specified.

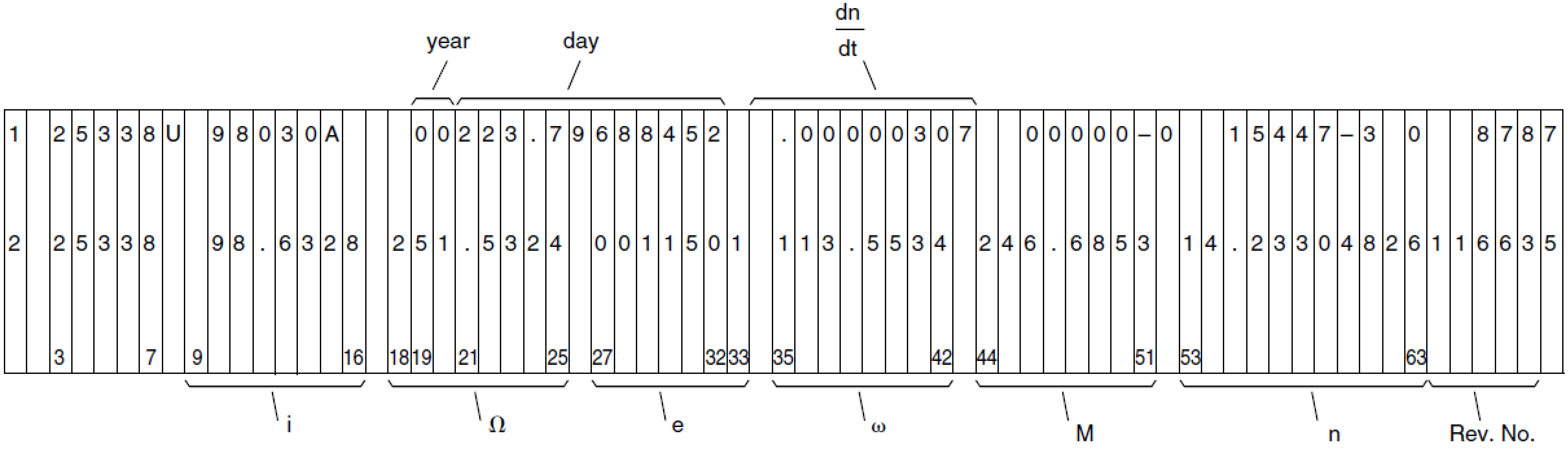
equatorial bulge

a'' is pure gravity, ***c*** is centrifugal force and ***b*** is their vector sum. The ellipsoidic surface is then perpendicular to the vector ***b***.



NOAA SAT. Two Line Element

Line no.	Columns	Description
1	3–7	*Satellite number:* 25338
1	19–20	*Epoch year* (last two digits of the year): 00
1	21–32	*Epoch day* (day and fractional day of the year): 223.79688452 (this is discussed further in Sec. 2.9.2).
1	34–43	*First time derivative of the mean motion* (rev/day^2): 0.00000307
2	9–16	*Inclination* (degrees): 98.6328
2	18–25	*Right ascension of the ascending node* (degrees): 251.5324
2	27–33	*Eccentricity* (leading decimal point assumed): 0011501
2	35–42	*Argument of perigee* (degrees): 113.5534
2	44–51	*Mean anomaly* (degrees): 246.6853
2	53–63	*Mean motion* (rev/day): 14.23304826
2	64–68	*Revolution number at epoch* (rev/day): 11,663



Using NOAA tow lines element to calculate semi-major axis

Example Calculate the semimajor axis for the satellite parameters

solution The mean motion is given in as

$$NN := 14.22296917 \cdot \mathrm{day}^{-1}$$

This can be converted to rad/sec as

$$n_0 := NN \cdot 2 \cdot \pi$$

$$\mu := 3.986005 \cdot 10^{14} \cdot \mathrm{m}^3 \cdot \mathrm{sec}^{-2}$$

Kepler's 3rd law gives

$$a := \left(\frac{\mu}{{n_o}^2}\right)^{1/3}$$

$$\underline{\underline{a = 7192.3 \cdot \mathrm{km}}}$$

Using NOAA tow lines element to calculate Apogee and Perigee Heights

➔Although not specified as orbital elements, the apogee height and perigee height are often required, the length of the radius vectors at apogee and perigee can be obtained from the geometry of the ellipse:

$$r_a = a\,(1 + e)$$

$$r_p = a\,(1 - e)$$

➔In order to find the apogee and perigee heights, the radius of the earth must be subtracted from the radii lengths, as shown in the following example.

Using NOAA tow lines element to calculate Apogee and Perigee Heights

Example 2.3 Calculate the apogee and perigee heights for the orbital parameters given in Table 2.1. Assume a mean earth radius of 6371 km.

solution The required data from Table 2.1 are: $e := .0011501$ $a := 7192.3 \cdot \text{km}$. (Note that the value for a was determined in Example 2.2.)

Given data:

$$R := 6371 \cdot \text{km}$$

$r_a := a \cdot (1 + e)$ ...Eq. (2.5) $r_a = 7200.6 \cdot \text{km}$

$r_p := a \cdot (1 - e)$ ...Eq. (2.6) $r_p = 7184.1 \cdot \text{km}$

$h_a := r_a - R$ $h_a = 829.6 \cdot \text{km}$
= = = = = = =

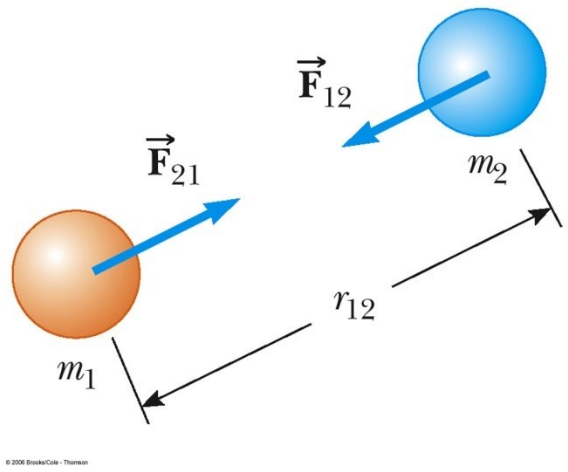
$h_p := r_p - R$ $h_p = 813.1 \cdot \text{km}$
= = = = = = =

Kepler’s and Newton Laws

Newton's Law of Universal Gravitation

- Every particle in the Universe attracts every other particle with a force that is directly proportional to the product of the masses and inversely proportional to the square of the distance between them.

$$F = G\frac{m_1 m_2}{r^2}$$



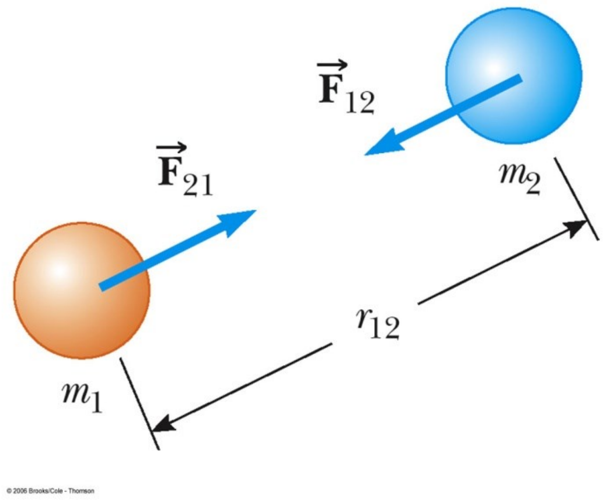
Universal Gravitation

- G is the constant of universal gravitation
- G = 6.673×10^{-11} N m² /kg²
- This is an example of an *inverse square law*
- Determined experimentally
- Henry Cavendish in 1798

$$F = G\frac{m_1 m_2}{r^2}$$

Universal Gravitation

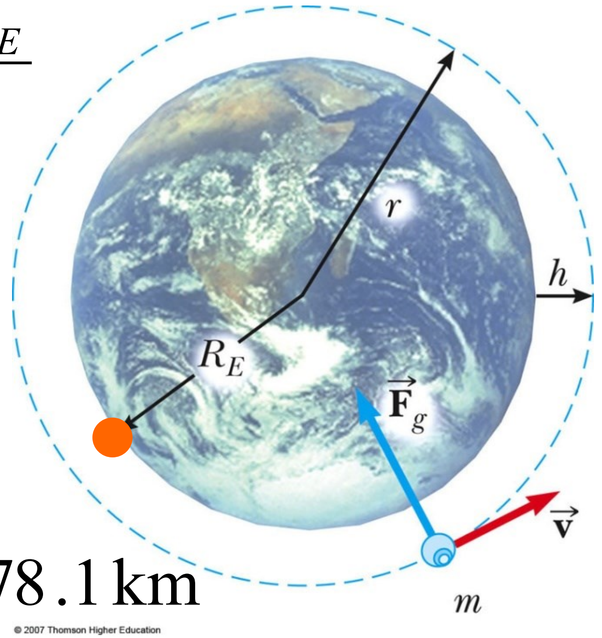
- The force that mass 1 exerts on mass 2 is equal and opposite to the force mass 2 exerts on mass 1
- The forces form a Newton's third law action-reaction



Free-Fall Acceleration and the Gravitational Force

- Consider an object of mass m near the Earth's surface

$$F = G\frac{m_1 m_2}{r^2} = G\frac{mM_E}{{R_E}^2}$$



- Acceleration a_g due to gravity

$$F = G\frac{mM_E}{{R_E}^2} = ma_g$$

- Since $M_E = 5.9742 \times 10^{23}\ \text{kg} \quad R_E = 6378.1\ \text{km}$

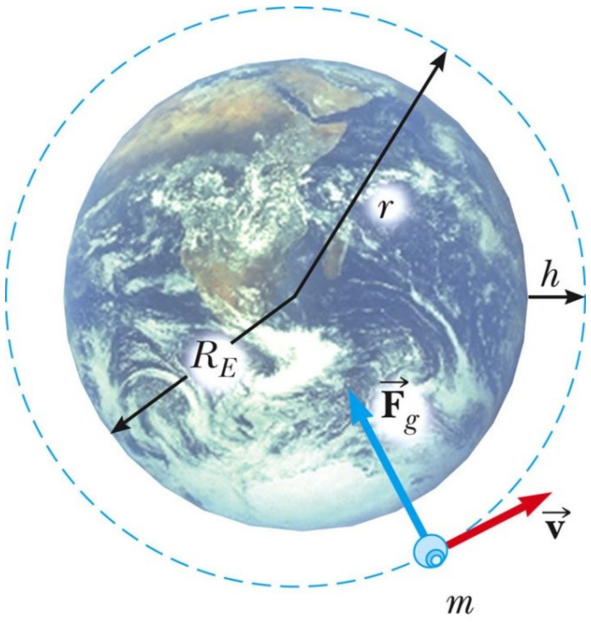
  we find at the Earth's surface

$$a_g = G\frac{M_E}{{R_E}^2} = 9.8\ \text{m/s}^2$$

Free-Fall Acceleration and the Gravitational Force

- Consider an object of mass m at a height h above the Earth's surface

$$F = G\frac{m_1 m_2}{r^2} = G\frac{mM_E}{(R_E + h)^2}$$



- Acceleration a_g due to gravity

$$F = G\frac{mM_E}{R_E^{\ 2}} = ma_g$$

- a_g will vary with altitude

$$a_g = G\frac{M_E}{(R_E + h)^2}$$

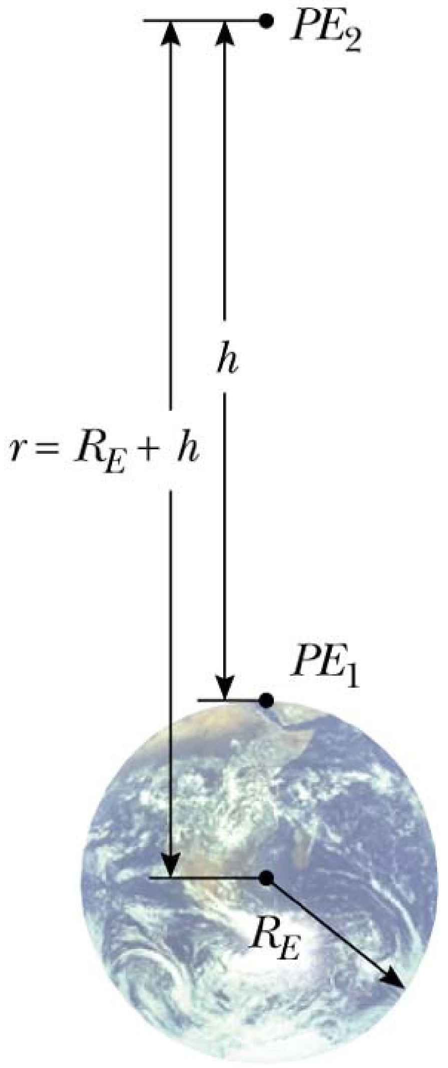
Gravitational Potential Energy

- $U = mgy$ is valid only near the earth's surface
- For objects high above the earth's surface, an alternate expression is needed

$$U = -G\frac{M_E m}{r}$$

  - Zero reference level is infinitely far from the earth, so potential energy is everywhere negative!
- Energy conservation

$$E = K + U = \frac{1}{2}mv^2 - G\frac{M_E m}{r}$$



Energy of an Orbit

- Consider a circular orbit of a planet around the Sun. What keeps the planet moving in its circle?
- It is the centripetal force produced by the gravitational force, i.e.

$$F = \frac{mv^2}{r} = G\frac{Mm}{r^2}$$

- That implies that: $\tfrac{1}{2}mv^2 = \dfrac{GMm}{2r}$

- Making this substitution in the expression for total energy:

$$E = \frac{1}{2}mv^2 - \frac{GMm}{r} = \frac{GMm}{2r} - \frac{GMm}{r}$$

$$\boxed{E = -\frac{GMm}{2r} \text{ (circular orbits)}}$$

- Note the total energy is negative, and is half the (negative) potential energy.
- For an elliptical orbit, r is replaced by a:

$$\boxed{E = -\frac{GMm}{2a} \text{ (elliptical orbits)}}$$

Escape Speed

- The escape speed is the speed needed for an object to soar off into space and not return

$$E = K + U = \frac{1}{2}mv^2 - G\frac{M_E m}{r} = 0$$

- For the earth, v_{esc} is about 11.2 km/s
- Note, v is independent of the mass of the object

$$v_{esc} = \sqrt{\frac{2GM_E}{R_E}}$$

Escape Speeds for the Planets and the Moon

Planet	v_e (km/s)
Mercury	4.3
Venus	10.3
Earth	11.2
Moon	2.3
Mars	5.0
Jupiter	60.0
Saturn	36.0
Uranus	22.0
Neptune	24.0
Pluto	1.1

Kepler's Third Law applied for planets around the Sun

- The square of the orbital period of any planet is proportional to cube of the average distance from the Sun to the planet.

$$T^2 = Ka^3$$

  - T is the period of the planet
  - a is the average distance from the Sun. Or a is the length of the semi-major axis
  - For orbit around the Sun, $K = K_S = 2.97 \times 10^{-19}\ s^2/m^3$
  - K is independent of the mass of the planet

$$K_s = \frac{4\pi^2}{GM_s}$$

The Mass of the Sun

- Calculate the mass of the Sun noting that the period of the
- Earth's orbit around the Sun is $3.156\ (10)^7$ s and its distance from the Sun is $1.496\ (10)^{11}$ m.

$$T^2 = \frac{4\pi^2}{GM} a^3$$

$$M = \frac{4\pi^2}{GT^2} a^3 = 1.99 \times 10^{30}\ \text{kg}$$

Geosynchronous Orbit

- From a telecommunications point of view, it’s advantageous for satellites to remain at the same location relative to a location on the Earth. This can occur only if the satellite’s orbital period is the same as the Earth’s period of rotation, 24 h. (a) At what distance from the center of the Earth can this geosynchronous orbit be found? (b) What’s the orbital speed of the satellite?

$$T = \sqrt{\frac{4\pi^2}{GM_E} a^3} = 24\,\mathrm{h} = 86400\ \mathrm{s}$$

$$a = \left(GM_E T^2 / 4\pi^2\right)^{1/3} = \left[(6.67\mathrm{e}-11)(5.97\mathrm{e}24)(86400\ \mathrm{s})^2 / 4\pi^2\right]^{1/3} = 41500\,\mathrm{km}$$

Orbit Perturbations

- The type of orbit described so far, referred to as a *keplerian orbit,* is elliptical for the special case of an artificial satellite orbiting the earth.
- keplerian orbit is ideal in the sense that it assumes that the earth is a uniform spherical mass and that the only force acting is the centrifugal force resulting from satellite motion balancing the gravitational pull of the earth.
- In practice, other forces which can be significant are the **gravitational forces of the sun and the moon** and **atmospheric drag**.

➔ The gravitational pulls of sun and moon have negligible effect on low-orbiting satellites, but they do affect satellites in the geostationary orbit.

➔ Atmospheric drag, on the other hand, has negligible effect on geostationary satellites but does affect low-orbiting earth satellites below about 1000 km.

Effects of a nonspherical earth

- For a spherical earth of uniform mass, Kepler's third law gives the nominal mean motion *n*0 as:

$$n_0 = \sqrt{\frac{\mu}{a^3}}$$

- The 0 subscript is included as a reminder that this result applies for a perfectly spherical earth of uniform mass. However, it is known that the earth is not perfectly spherical.
- there being an equatorial bulge and a flattening at the poles, a shape described as an *oblate spheroid.* When the earth's oblateness is taken into account, the mean motion, denoted in this case by symbol *n,* is modified to:

$$n = n_0 \left[\frac{1 + K_1 (1 - 1.5 \sin^2 i)}{a^2 (1 - e^2)^{1.5}} \right]$$

K_1 is a constant which evaluates to 66,063.1704 km^2

Effects of a nonspherical earth over the semi-major axis

The earth's oblateness has negligible effect on the semimajor axis *a,* and if *a* is known, the mean motion is readily calculated. The orbital period taking into account the earth's oblateness is termed the *anomalistic period* (e.g., from perigee to perigee). The mean motion specified in the NASA bulletins is the reciprocal of the anomalistic period. The anomalistic period is:

$$P_A = \frac{2\pi}{n} \text{ sec}$$

where n is in radians per second.

If the known quantity is n (as is given in the NASA bulletins, for example), one can solve Eq. (2.8) for a, keeping in mind that n_0 is also a function of a. Equation (2.8) may be solved for a by finding the root of the following equation:

$$n - \sqrt{\frac{\mu}{a^3}}\left[1 + \frac{K_1 (1 - 1.5 \sin^2 i)}{a^2 (1 - e^2)^{1.5}}\right] = 0$$

Effects of a nonspherical earth over the semi-major axis

The earth's oblateness has negligible effect on the semimajor axis *a,* and if *a* is known, the mean motion is readily calculated. The orbital period taking into account the earth's oblateness is termed the *anomalistic period* (e.g., from perigee to perigee). The mean motion specified in the NASA bulletins is the reciprocal of the anomalistic period. The anomalistic period is:

$$P_A = \frac{2\pi}{n} \text{ sec}$$

where n is in radians per second.

If the known quantity is n (as is given in the NASA bulletins, for example), one can solve Eq. (2.8) for a, keeping in mind that n_0 is also a function of a. Equation (2.8) may be solved for a by finding the root of the following equation:

$$n - \sqrt{\frac{\mu}{a^3}}\left[1 + \frac{K_1 (1 - 1.5 \sin^2 i)}{a^2 (1 - e^2)^{1.5}}\right] = 0$$

Example 2.4 A satellite is orbiting in the equatorial plane with a period from perigee to perigee of 12 h. Given that the eccentricity is 0.002, calculate the semimajor axis. The earth's equatorial radius is 6378.1414 km.

solution Given data:

$$e := .002 \qquad i := 0 \cdot \text{deg} \qquad P := 12 \cdot \text{hr}$$

$$K_1 := 66063.1704 \cdot \text{km}^2 \qquad a_E := 6378.1414 \cdot \text{km}$$

$$\mu := 3.986005 \cdot 10^{14} \cdot \text{m}^3 \cdot \text{sec}^{-2}$$

The mean motion is

$$n := \frac{2 \cdot \pi}{P}$$

Kepler's third law gives

$$a := \left(\frac{\mu}{n^2}\right)^{1/3}$$

$a = 26597 \cdot \text{km}$ ...This is the nonperturbed value which can be used as a guess value for the root function.

= = = = = =

Perturbed value:

$$a := \text{root}\left[n - \left(\sqrt{\frac{\mu}{a^3}}\right) \cdot \left[1 + \frac{K_1\,(1 - 1.5 \cdot \sin(i)^2)}{a^2 \cdot (1 - e^2)^{1.5}}\right], a\right]$$

$$a = 26598.6 \cdot \text{km}$$

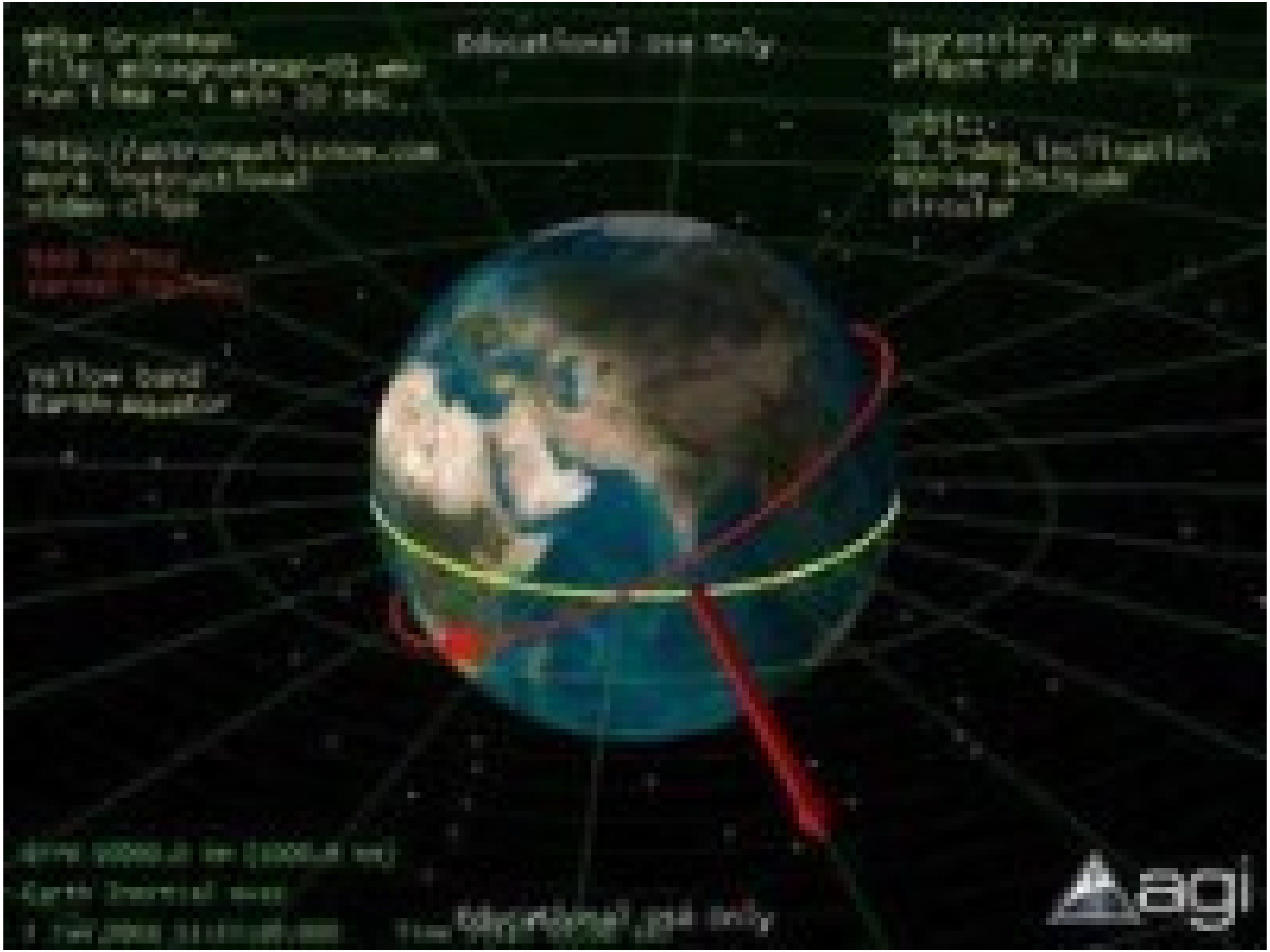
Effects of a nonspherical earth over RAAN

The oblateness of the earth also produces two rotations of the orbital plane.

1) The first of these, known as ***regression of the nodes,*** is where the nodes appear to slide along the equator.

➔ In effect, the line of nodes, which is in the equatorial plane, rotates about the center of the earth. Thus , the right ascension of the ascending node, shifts its position.

If the orbit is prograde, the nodes slide westward, and if retrograde, they slide eastward. As seen from the ascending node, a satellite in prograde orbit moves eastward, and in a retrograde orbit, westward. The nodes therefore move in a direction opposite to the direction of satellite motion, hence the term *regression of the nodes.* **For a polar orbit (i 90°), the regression is zero.**


Educational Use Only
Yellow band
Earth equator

Effects of a nonspherical earth over RAAN (Continued)

2) The second effect is **rotation of apsides in the orbital plane**, described next slide. Both effects depend on the mean motion *n,* the semimajor axis *a,* and the eccentricity *e.* These factors can be grouped into one factor *K* given by:

$$K := \frac{n \cdot K_1}{a^2 \cdot (1 - e^2)^2}$$

K will have the same units as n. Thus, with n in rad / day, K will be in rad / day, and with n in °/day, K will be in °/day. An approximate expression for the rate of change of with respect to time:

Regression of nodes rate change equation:

$$\frac{d\Omega}{dt} = -K \cos i$$

The rate of regression of the nodes will have the same units as *n.* When the rate of change is negative, the regression is westward, and when the rate is positive, the regression is eastward. It will be seen, therefore, that for eastward regression, *i* must be greater than 90°, or the orbit must be retrograde. It is possible to choose values of *a, e,* and *i* such that the rate of rotation is 0.9856°/day eastward.

Effects of a nonspherical earth over Angle of Perigee

In the other major effect produced by the equatorial bulge, **rotation of the line of apsides** in the orbital plane, the argument of perigee changes with time, in effect, the rate of change being given by:

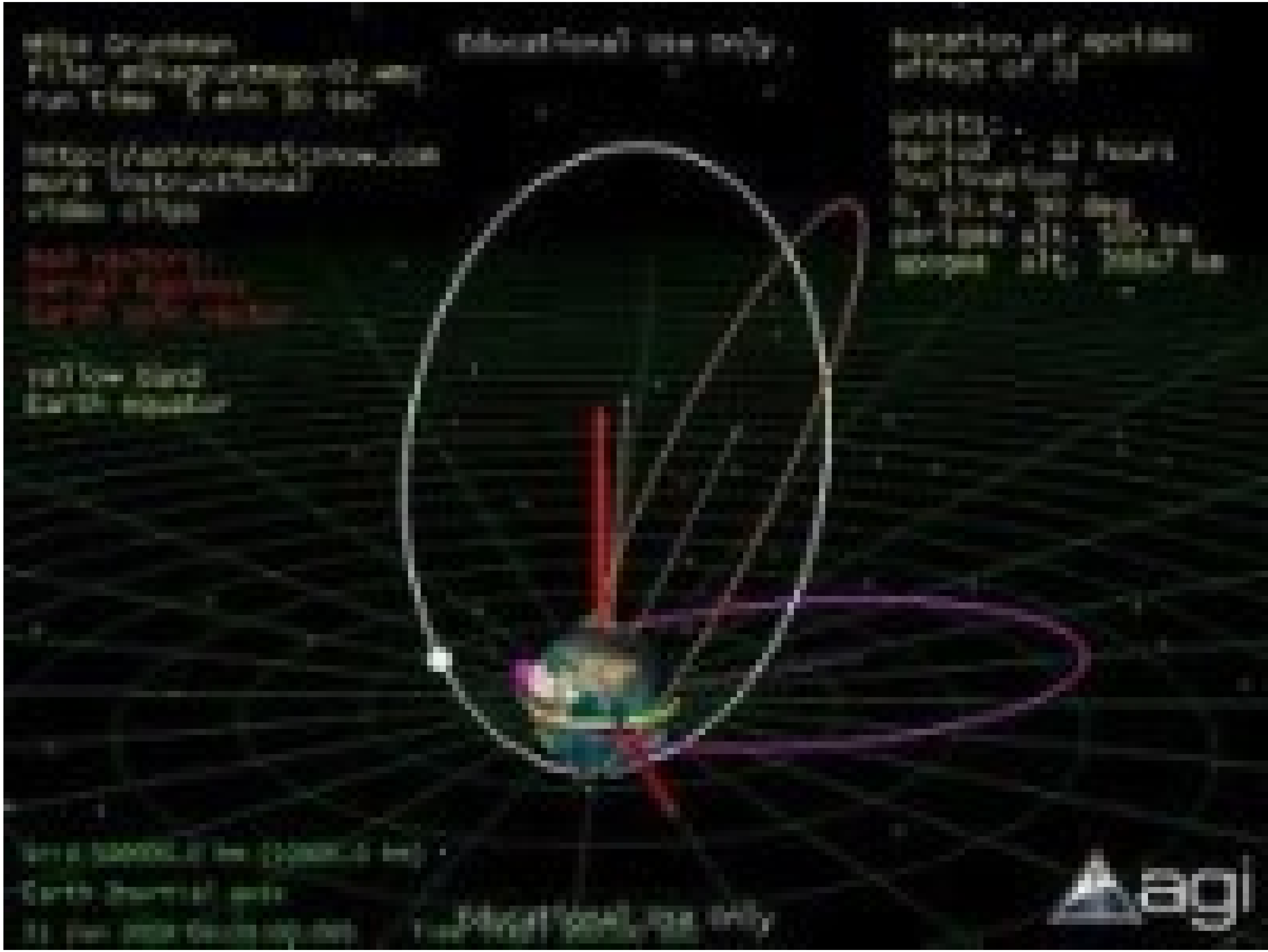
$$\frac{d\omega}{dt} = K\,(2 - 2.5\,\sin^2 i)$$

When the inclination *i* is equal to 63.435°, the term within the parentheses is equal to zero, and hence no rotation takes place. Use is made of this fact in the orbit chosen for the Russian Molniya satellites

Denoting the epoch time by t_0, the right ascension of the ascending node by Ω_0, and the argument of perigee by ω_0 at epoch gives the new values for Ω and ω at time *t* as:

$$\Omega = \Omega_0 + \frac{d\Omega}{dt}\,(t - t_0)$$

$$\omega = \omega_0 + \frac{d\omega}{dt}\,(t - t_0$$


Educational Use Only
Yellow band
Earth equator
agi

Homework 1

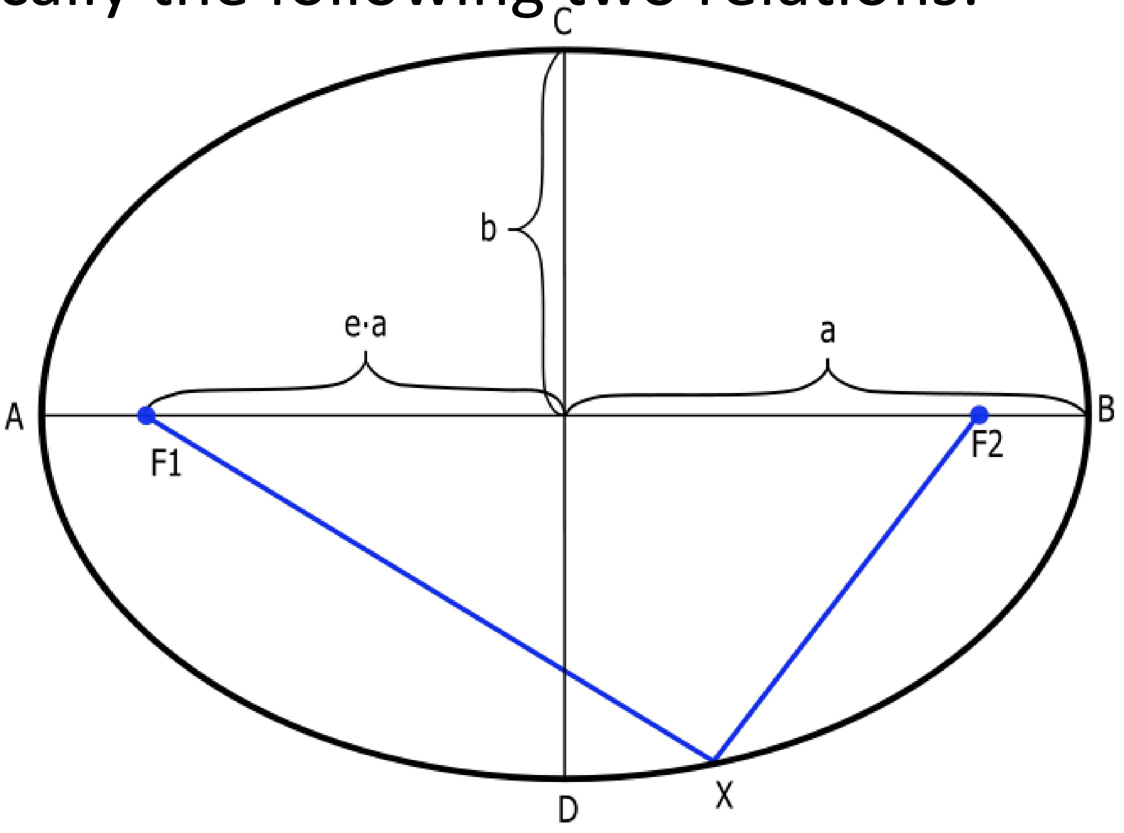
- For a given ellipse , prove mathematically the following two relations:

$$r_a = a\,(1 + e)$$

$$r_p = a\,(1 - e)$$

r_a and r_p are the apogee and perigee hights respectively

Homework hint:
➔ Distance from one focus to center of the ellipse is $e.a$, where e is the **eccentricity**.



Homework 2

- Using Sun-Synchrounous satellite orbit simulation in STK, find the line of nodes regression rate change in time and line of apsides rotation rate in rad/s. (make any realistic assumptions you want)

- Explain the results…
- Repeat for inclined satellite orbit i=70 degrees
- Do the simulations in STK and send the project file using the assignment submission link, and submit a word file for the answers.

Example 2.5 Determine the rate of regression of the nodes and the rate of rotation of the line of apsides for the NOAA Satellite

Given that :

from Table 2.1 and Example 2.2:

$i := 98.6328 \cdot deg \qquad e := .0011501$

$n := 14.23304826 \cdot day^{-1} \qquad a := 7192.3 \cdot km$

$K_1 := 66063.1704 \cdot km^2$

solution

$n := 2 \cdot \pi \cdot n$ ...Converts n to SI units of rad/sec.

$$K := \frac{n \cdot K_1}{a^2 \cdot (1 - e^2)^2} \qquad K = 6.544 \cdot \frac{deg}{day}$$

$$\Omega' := -K \cdot \cos(i) \qquad \Omega' = 0.982 \cdot \frac{deg}{day}$$

$= = = = = = =$

$$\omega' := K \cdot (2 - 2.5 \cdot \sin(i)^2) \qquad \omega' = -2.903 \cdot \frac{deg}{day}$$

$= = = = = = = =$

Example 2.6 Calculate, for the satellite in Example 2.5, the new values for ω and Ω one period after epoch.

solution From Example 2.5:

$$\Omega' := .982 \cdot \frac{\text{deg}}{\text{day}} \qquad \omega' := -2.903 \cdot \frac{\text{deg}}{\text{day}}$$

From NOAA two lines element information

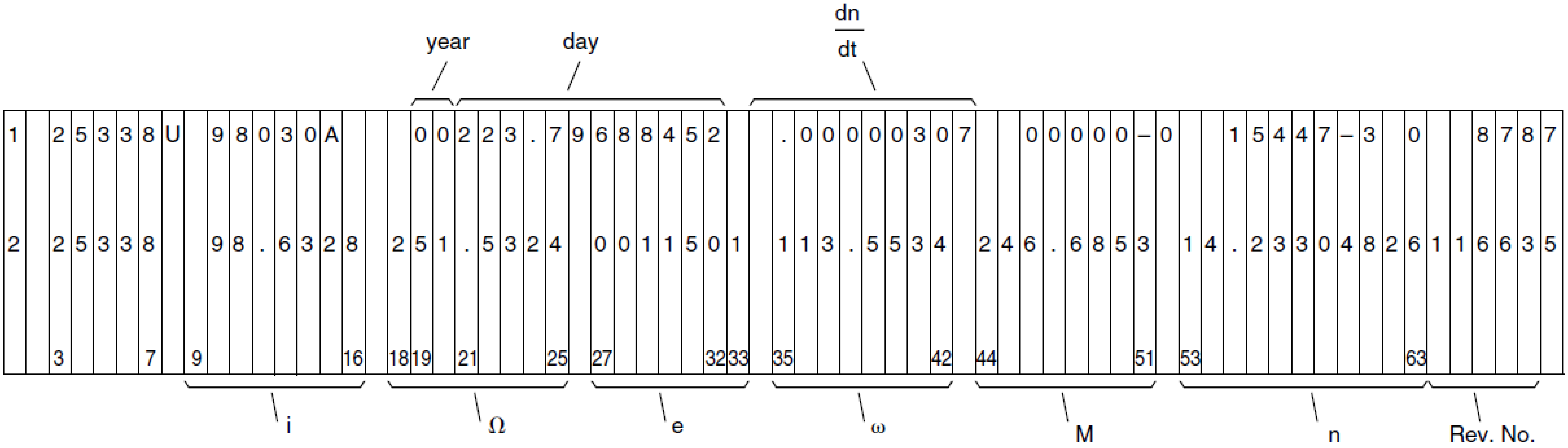
$$n := 14.23304826 \cdot \text{day}^{-1} \qquad \omega_O := 113.5534 \cdot \text{deg} \qquad \Omega_0 := 251.5324 \cdot \text{deg}$$

The period is

$$P_A = \frac{1}{n}$$

$$\Omega := \Omega_O + \Omega' \cdot P_A \qquad \underline{\underline{\Omega = 251.601 \cdot \text{deg}}}$$

$$\omega := \omega_O + \omega' \cdot P_A \qquad \underline{\underline{\omega = 113.349 \cdot \text{deg}}}$$



year
day
dn/dt
1 25338U 98030A 00223.79688452 .00000307 00000-0 15447-3 0 8787
2 25338 98.6328 251.5324 0011501 113.5534 246.6853 14.23304826116635
3 7 9 16 18 19 21 25 27 32 33 35 42 44 51 53 63
i
Ω
e
ω
M
n
Rev. No.

Atmospheric Drag

- ➔ For near-earth satellites, below about 1000 km, the effects of atmospheric drag are significant.
- ➔ Because the drag is greatest at the perigee, the drag acts to reduce the velocity at this point, with the result that the satellite does not reach the same apogee height on successive revolutions.
- ➔ The result is that the semimajor axis and the eccentricity are both reduced.

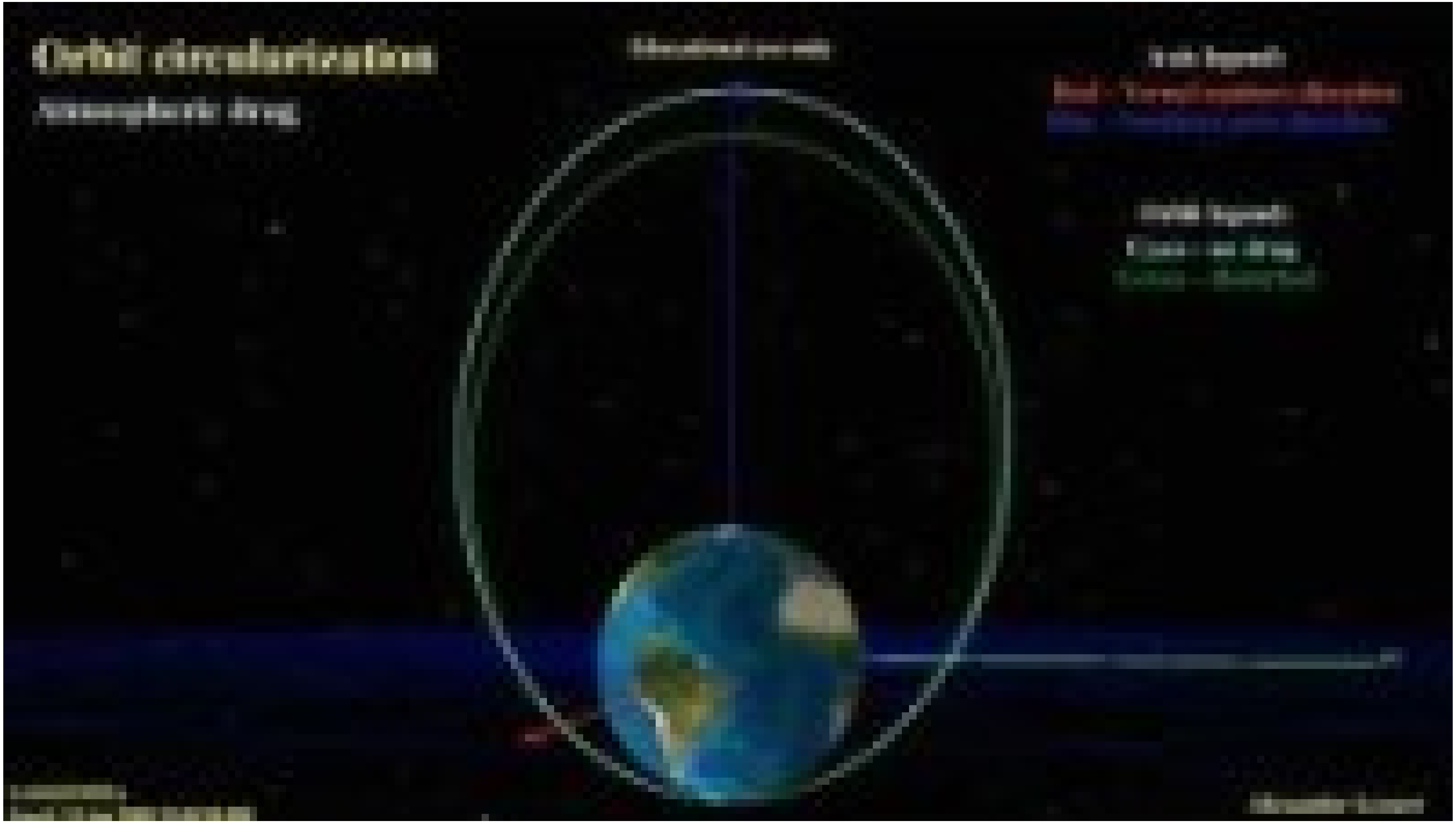
An approximate expression for the change of major axis is:

$$a \cong a_0 \left[\frac{n_0}{n_0 + n_0{}' (t - t_0)} \right]^{2/3}$$

The mean anomaly is also changed. An approximate expression for the amount by which it changes is:

$$\delta M = \frac{n_0{}'}{2} (t - t_0)^2$$

it is seen that the first time derivative of the mean motion is listed in columns 34–43 of line 1 of the NASA bulletin. For the example shown in the previous slide, this is 0.00000307 rev/day[2]


Orbit circularization
Atmospheric drag

Satellite Waves Transmission Equation

➔ The transmission equation relates the received power level at the destination, which could be the Earth station or the satellite in the case of a satellite communication link, to the transmitted RF power, the operating frequency and the transmitter–receiver distance.

➔ The quality of the information delivered to the destination is governed by the level of the signal power received.

POWER FLUX DENSITY: It is assumed that the transmitter radiates a power P_T watts with an antenna having a gain G_T as compared to the isotropic radiation level. The power flux density (P_{RD} in W/m^2) due to the radiated power in the direction of the antenna bore sight at a distance d metres is given by:

$$P_{RD} = \frac{P_T G_T}{4\pi d^2}$$

The product $P_T G_T$ is the effective isotropic radiated power (EIRP)

Also, if the radiating aperture A_T of the transmitting antenna is large as compared to λ^2, where λ is the operating wavelength, then G_T equals $(4\pi A_T/\lambda^2)$. If A_R is the aperture of the receiving antenna, then the received power P_R at the receiver at a distance d from the transmitter can be expressed as:

$$P_R = \left(\frac{P_T G_T}{4\pi d^2}\right) A_R$$

where A_R is related to the receiver antenna gain by: $G_R = 4\pi A_R/\lambda^2$.

The expression for the received power is modified to: $P_R = \frac{P_T G_T G_R \lambda^2}{(4\pi d)^2}$

or

$$P_R = \frac{P_T G_T G_R}{(4\pi d/\lambda)^2} = \frac{P_T G_T G_R}{L_P}$$

The term $(4\pi d/\lambda)^2$ represents the free space path loss L_P. The above expression is also known as the Friis transmission equation. The received power can be expressed in decibels as:

$$10\log P_R = 10\log P_T + 10\log G_T + 10\log G_R - 10\log L_P$$

$$P_R(\text{in dBW}) = \text{EIRP}(\text{in dBW}) + G_R(\text{in dB}) - L_P(\text{in dB})$$

The above equation can be modified to include other losses, if any, such as losses due to atmospheric attenuation, antenna losses, etc. For example, if L_A, L_{TX} and L_{RX} are the losses due to atmospheric attenuation, transmitting antenna and receiving antenna respectively, then the above equation can be rewritten as:

$$P_R = \text{EIRP} + G_R - L_P - L_A - L_{TX} - L_{RX}$$

Example:

A geostationary satellite at a distance of 36 000 km from the surface of the Earth radiates a power of 10 watts in the desired direction through an antenna having a gain of 20 dB. What would be the power density at a receiving site on the surface of the Earth and also the power received by an antenna having an effective aperture of 10 m^2?

Solution: The power density can be computed from

$$\text{Power flux density} = \frac{P_T G_T}{4\pi d^2}$$

where the terms have their usual meaning. Here,

$$G_T = 20\,\text{dB} = 100, \qquad P_T = 10\,\text{watts}, \qquad d = 36\,000\,\text{km} = 36 \times 10^6\,\text{m}$$

This gives

Power flux density $= (10 \times 100)/[4 \times \pi \times (36 \times 10^6)^2] = 0.0614 \times 10^{-12}\,\text{W/m}^2$

Power received by the receiving antenna $= 0.0614 \times 10^{-12} \times 10 = 0.614\,\text{pW}$

Satellite Link Parameters

Important parameters that influence the design of a satellite communication link include the following:

1. Choice of operating frequency
2. Propagation considerations
3. Noise considerations
4. Interference-related problems

1.Choice of operating frequency:

The choice of frequency band from those allocated by ITU for satellite communication services such as (FSS), (BSS) and (MSS) is mostly governed by factors like **propagation considerations**, **coexistence with other services**, **interference-related issues**, **technology status**, **economic considerations**.

➔ While it may be more economic to use lower frequency bands, there would be interference-related problems as a large number of terrestrial microwave links use frequencies within these bands. Also, lower frequency bands would offer lower bandwidths and hence a reduced transmission capacity.

➔ Higher frequency bands offer higher bandwidths but suffer from the disadvantage of severe rain-induced attenuation, particularly above 10 GHz. Also, above 10 GHz, rain can have the effect of reducing isolation between orthogonally polarized signals in a frequency re-use system. It may be mentioned here that for frequencies less than 10 GHz and elevation angles greater than 5^0, atmospheric attenuation is more or less insignificant.

➔ The bands of interest for satellite communications lie above 100MHz and include the VHF, UHF, L, S, C, X, Ku and Ka bands.

➔ Higher frequencies are employed for satellite communication as the frequencies below 100 MHz are either reflected by the ionosphere or they suffer varying degrees of bending from their original paths due to refraction by the ionosphere.

➔ Initially, the satellite communication was mainly concentrated in the C band (6/4 GHz) as it offered fewest propagation as well as attenuation problems.

Band	Frequency (GHz)
L band	1–2
S band	2–4
C band	4–8
X band	8–12
Ku band	12–18
K band	18–27
Ka band	27–40
V band	40–75
W band	75–110

2. Propagation Considerations:

The nature of propagation of electromagnetic waves or signals through the atmospheric portion of an Earth station–satellite link has a significant bearing on the link design.

➔ From the viewpoint of a transmitted or received signal, it is mainly the **operating frequency** and to a lesser extent the **polarization** that would decide how severe the effect of atmosphere is going to be.

➔ From the viewpoint of atmosphere, it is the first few tens of kilometres constituting the troposphere and then the ionosphere extending from about 80 km to 1000 km that do the damage.

➔ The effect of atmosphere on the signal is mainly in the form of attenuation caused by atmospheric scattering and scintillation and depolarization caused by rain in the troposphere and Faraday rotation in the ionosphere.

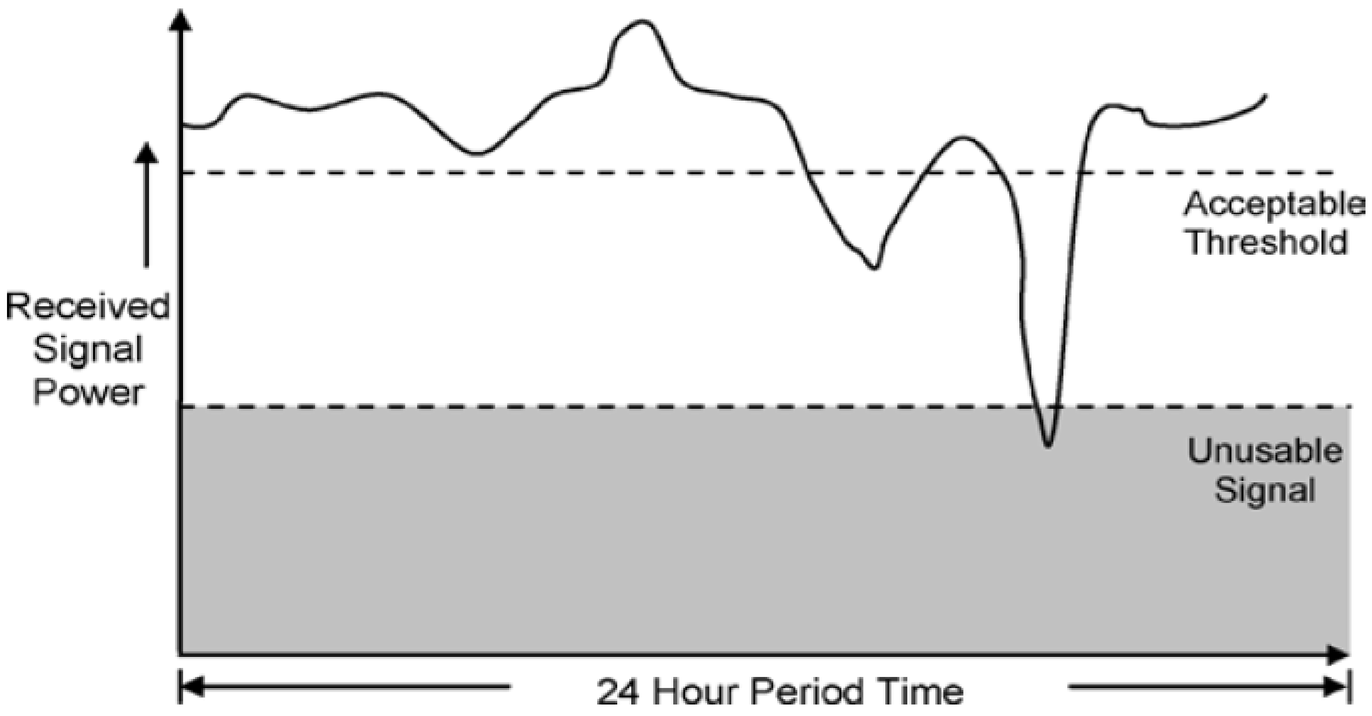
➔ While rain-induced attenuation is very severe for frequencies above 10 GHz, polarization changes due to Faraday rotation are severe at lower frequencies and are almost insignificant beyond 10 GHz. In fact, atmospheric attenuation is the least in the 3 to 10 GHz window. That is why it is the preferred and most widely used one for satellite communications.

Attenuation is defined as the difference between the power that would have been received under ideal conditions and the actual power received at a given time.

$$A(t) = P_{rideal}(t) - P_{ractual}(t)$$

Where,

$A(t)$ is the attenuation at any given time t
$P_{rideal}(t)$ is the received power under ideal conditions at time t
$P_{ractual}(t)$ is the actual received power at time t



➔ Free-space loss is the loss of signal strength only due to distance from the transmitter. In the absence of any material source of attenuation of electromagnetic signals, therefore, the radiated electromagnetic power diminishes as the inverse square of the distance from the transmitter, which implies that the power received by an antenna of 1m^2 cross-section will be $P_t/(4\pi R_2)$ where P_t is the transmitted power and R is the distance of the receiving antenna from the transmitter. In the case of uplink, the Earth station antenna becomes the transmitter and the satellite transponder is the receiver. It is the opposite in the case of downlink.

The free-space path loss component can be computed from

$$L_{\mathrm{FS}} = \left(\frac{4\pi R}{\lambda}\right)^2 = 20\log\left(\frac{4\pi R}{\lambda}\right) \quad \mathrm{dB}$$

where L_{FS} is the free space loss and λ = operating wavelength. Also, $\lambda = c/f$, where

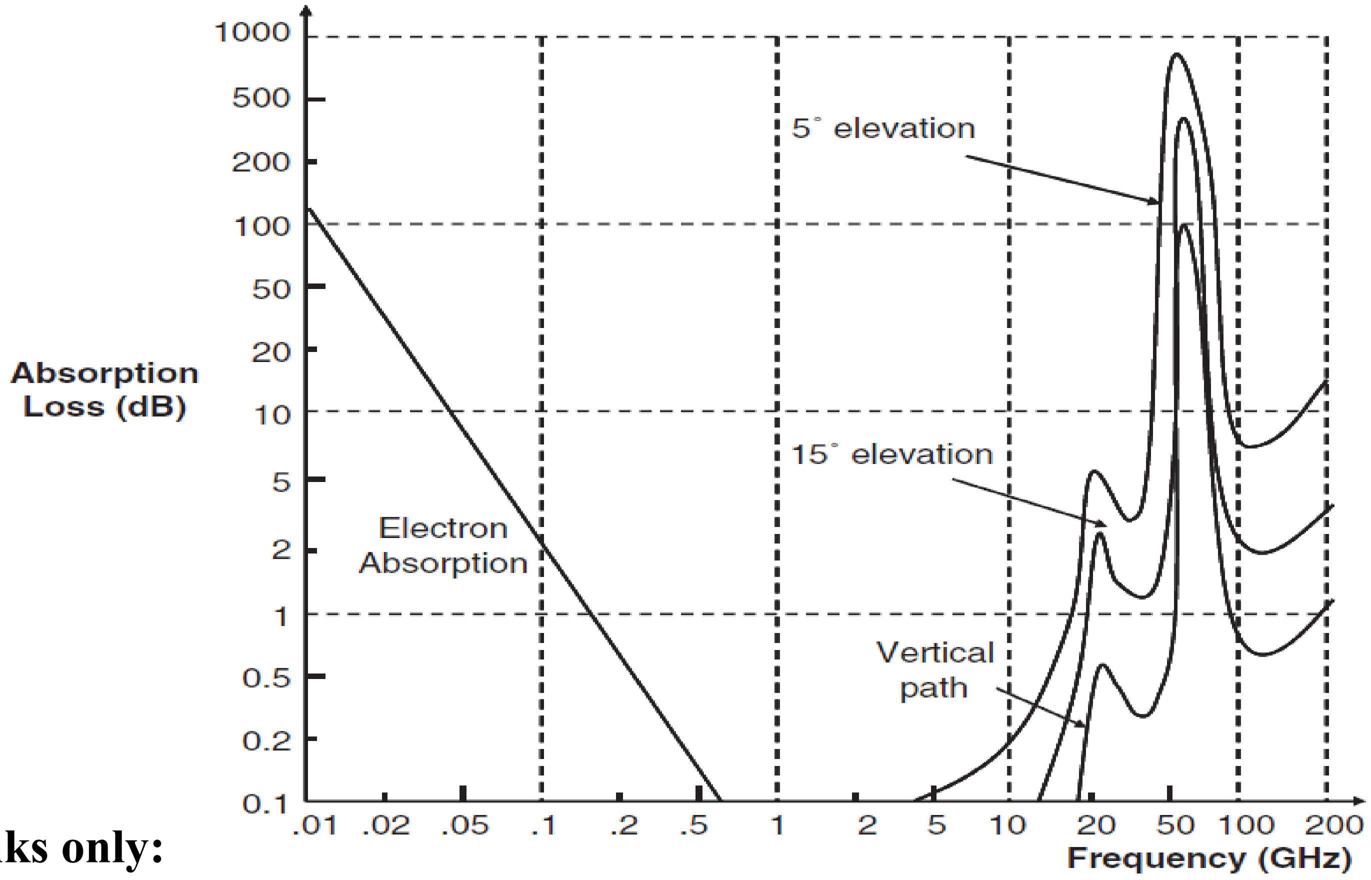
c = velocity of electromagnetic waves in free space
f = operating frequency

If c is taken in km/s and f in MHz, then the free-space path loss can also be computed from

$$L_{\mathrm{FS}}(\mathrm{dB}) = (32.4 + 20\log R + 20\log f)$$

➔ *Gaseous Absorption:*

Electromagnetic energy gets absorbed and converted into heat due to gaseous absorption as it passes through the troposphere. The absorption is primarily due to the presence of **molecular oxygen** and uncondensed water **vapour** and has been observed to be not so significant as to cause problems in the frequency range of 1 to 15 GHz. The presence of **free electrons** in the atmosphere also causes absorption due to collision of electromagnetic waves with these electrons. However, electron absorption is significant only at frequencies less than 500 MHz.



The following two bands are used for inter-satellite links only:

- The first absorption band is caused due to the resonance phenomenon in water vapour and occurs at 22.2 GHz
- The second band is caused by a similar phenomenon in oxygen and occurs around 60 GHz.

➔ It should also be mentioned that absorption at any frequency is a function of temperature, pressure, relative humidity (RH) and elevation angle.

➔ Absorption increases with a decrease in elevation angle E due to an increase in the transmission path. Absorption at any elevation angle E less than 90∘ can be computed by multiplying the absorption figure for 90∘ degree elevation by **1/sin E**. Applying this correction, it is observed that the one-way absorption figure range in the 1–15 GHz frequency band increases from (0.03–0.2 dB) for 90^0 elevation to (0.35–2.3 dB) for 5^0elevation.

➔ Absorption is also observed to increase with **humidity**. In the case of resonance absorption of water vapours around 22.2 GHz, it varies from as low as 0.05 dB (at 0% RH) to about 1.8 dB (at 100% RH). It may also be mentioned here that the data in the previous figure applies to the Earth station at sea level and the losses would reduce with an increase in height of the Earth station.

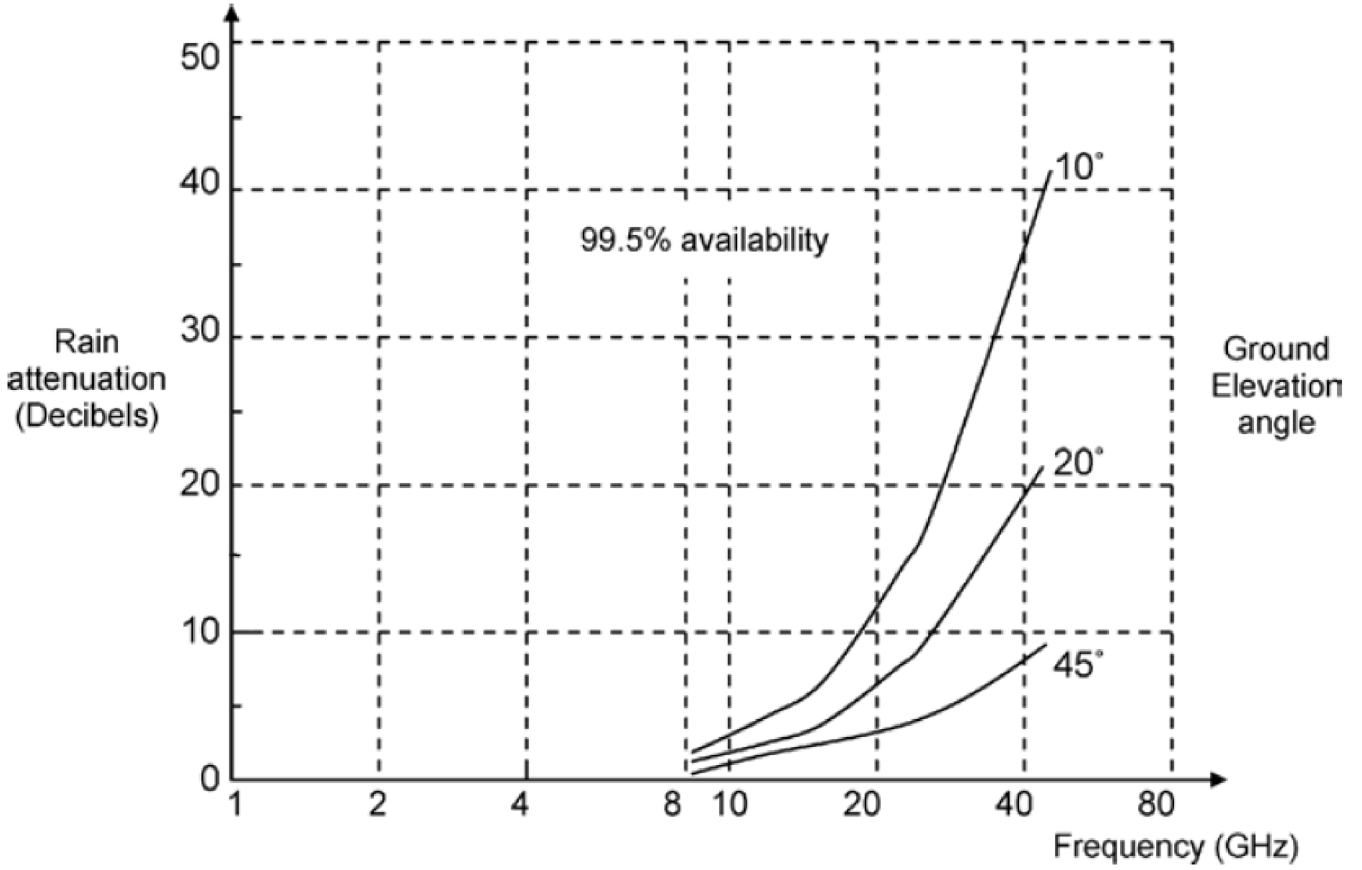
➔ It is evident from the plots of the previous figure that there are two transmission windows in which absorption is either insignificant or has a local minimum. The first window is in the frequency range of 500 MHz to 10 GHz and the second is around 30 GHz. This explains the wide use of the 6/4 GHz band. The increasing interest in the 30/20 GHz band is due to the second window, which shows a local minimum around 30 GHz. Losses at the 14/11 GHz satellite band are within acceptable limits with values of about 0.8 dB for 5^0 elevation and 0.2 dB for 15^0 elevation.

Attenuation due to Rain:

➔ After the free-space path loss, rain is the next major factor contributing to loss of electromagnetic energy caused by absorption and scattering of electromagnetic energy by rain drops.

➔ It may be mentioned here that the loss of electromagnetic energy due to gaseous absorption discussed in the earlier slide tends to remain reasonably constant and predictable. On the other hand, the losses due to precipitation in the form of rain, fog, clouds, snow, etc., are variable and far less predictable.

Losses due to rain increases with an increase in frequency and reduction in the elevation angle. It is evident from the family of curves that there is not much to worry about from rain attenuation for C band satellite links. Attenuation becomes significant above 10 GHz and therefore when a satellite link is planned to operate beyond 10 GHz, an estimate of rain caused attenuation is made by making extensive measurements at several locations in the coverage area of the satellite system.



Attenuation of electromagnetic waves due to rain (A_{rain}) extended over a path length of L can be computed from

$$A_{rain} = \int_0^L \alpha \, d\alpha$$

where $\alpha =$ specific attenuation of rain in dB/km. Specific attenuation again depends upon various factors like rain drop size, drop size distribution, operating wavelength and the refractive index. In practice, rain attenuation is estimated from

$$\alpha = aR^b$$

where a and b are frequency and temperature-dependent constants and R is the surface rain rate at the location of interest.

Cloud Attenuation

- Attenuation due to clouds is more or less irrelevant for lower frequency bands (L, S, C and Ku bands), but is largely relevant for satellite systems employing Ka and V band frequencies.

- The attenuation figure for water-filled clouds is much larger than the attenuation figure for clouds made from ice crystals. It may be mentioned here that the attenuation figure for ice clouds is negligible for the frequency bands of interest.

- The typical figure of attenuation for water filled clouds is of the order of 1 to 3 dB for frequency bands around 30 GHz at elevation angle of 30° in temperate latitude locations. However, the attenuation figures increase with increase in the thickness of the clouds and its probability of occurrence. In addition, the cloud attenuation increases for lower elevation angles.

Signal Fading due to Refraction

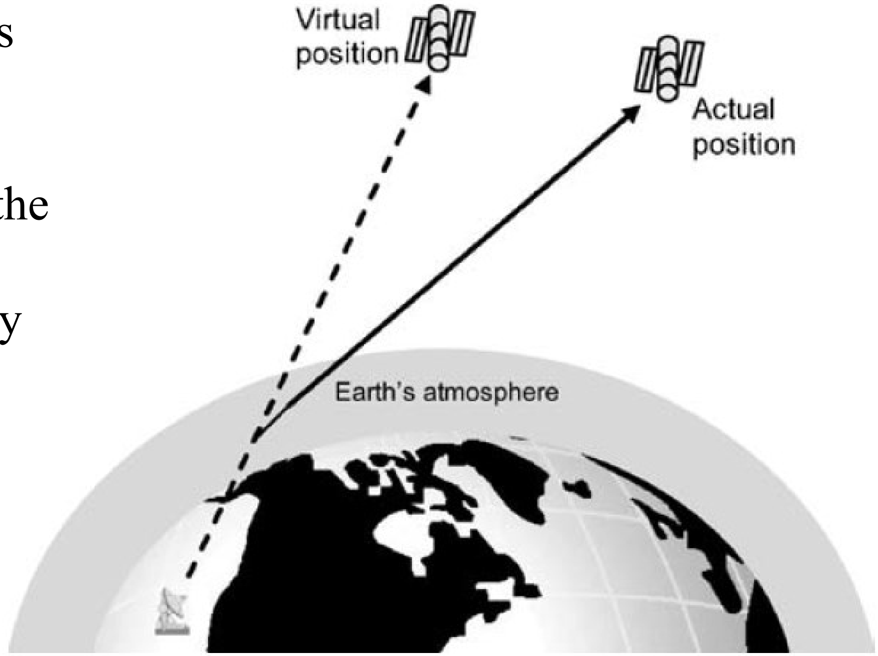
Refraction is the phenomenon of bending of electromagnetic waves as they pass through the different layers of the atmosphere. Refraction of the satellite beams occur in the troposphere (lower layer of the atmosphere from the Earth's surface to a height of 15 km approximately) due to the variations in the refractive index of the air column.

The variations in the refractive index is the result of the turbulent mixing of the different columns of the air due to the agitative convective activity in the troposphere caused by heating of the Earth's surface by the sun.

➔ The variations in the refractive index lead to bending of electromagnetic waves resulting in fluctuations in the received signal levels, also referred to as scintillations.

➔ It leads to a virtual position for the satellite slightly above the true position of the satellite.

➔ The random nature of bending due to discontinuities and fluctuations caused by unstable atmospheric conditions like temperature inversions, clouds and fog produces signal fading, which leads to loss of signal strength. Fading is the phenomenon wherein the Earth station receiving antenna receives the signal transmitted by the satellite via different paths with different phase shifts. The fading phenomenon is more adverse at lower elevation angles.

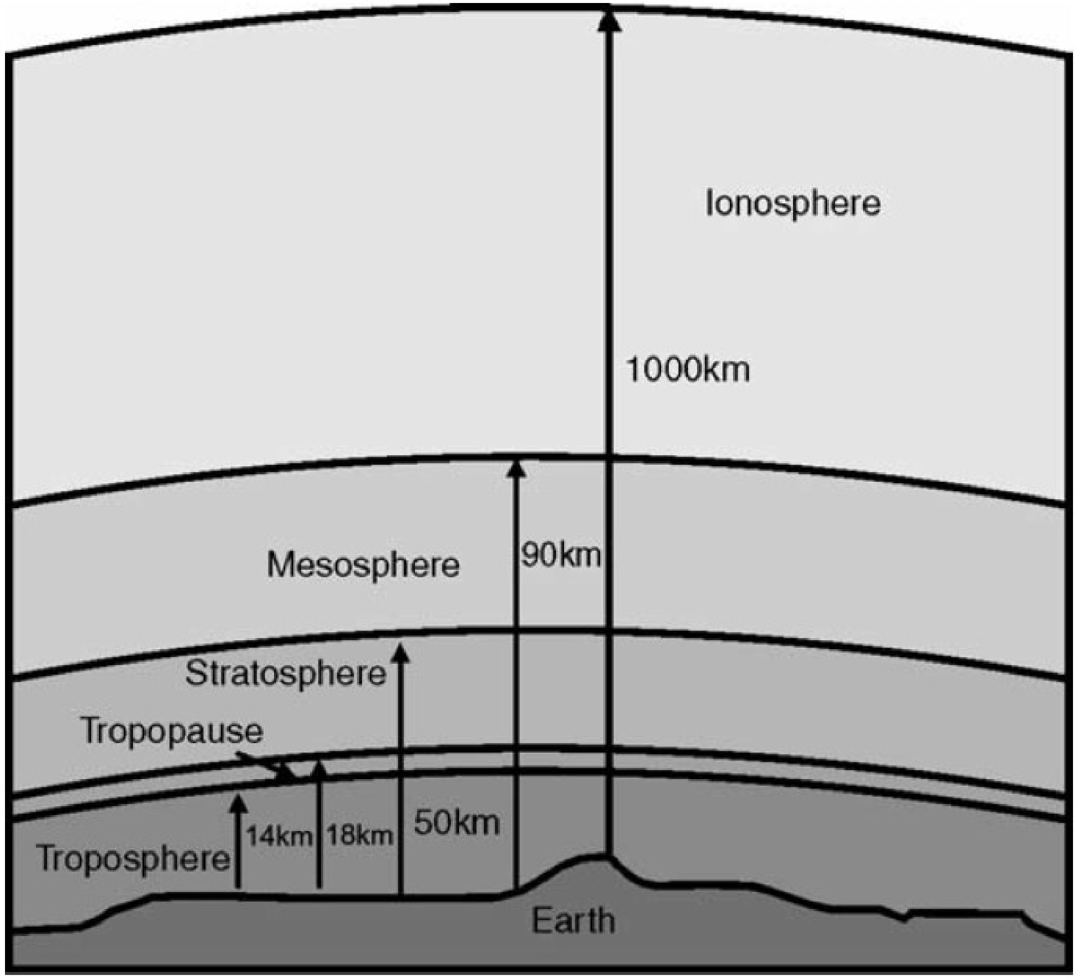


Ionosphere-related Effects

- ➔ The ionosphere is an ionized region in space, extending from about 80–90 km to 1000 km formed by interaction of solar radiation with different constituent gases of the atmosphere.

- ➔ The effects that are of concern and need attention include **polarization rotation**, also called the Faraday effect, and **scintillation**, which is simply rapid fluctuation of the signal amplitude, phase, polarization or angle of arrival.

- ➔ most ionospheric effects including those of primary interest, like polarization rotation and scintillation, decrease with an increase in frequency having a $1/f^2$ dependence.



Polarization Rotation – Faraday Effect

➔ When an electromagnetic wave passes through a region of high electron content like the ionosphere, the plane of polarization of the wave gets rotated due to interaction of the electromagnetic wave with the Earth's magnetic field.

➔ The angle through which the plane of polarization rotates is directly proportional to the total electron content of the ionized region and – inversely proportional to the square of the operating frequency. It also depends upon the state of the ionosphere, time of the day, solar activity, the direction of the incident wave, etc. Directions of polarization rotation are opposite for transmit and receive signals.

➔ The polarization rotation angle ($\Delta\psi$) for a path length through the ionosphere of Z metres is given by:

$$\Delta\Psi = \int\left(\frac{2.36 \times 10^4}{f^2}\right) ZNB_o \cos\theta dz$$

Where,

$\Delta\Psi$ is the rotation angle (radians)
θ is the angle between the geomagnetic field and the direction of propagation of the wave
N is the electron density (electrons/cm^3)
B_o is the geomagnetic flux density (Tesla)
f is the operating frequency (Hz)

For a polarization rotation angle (also referred to as the polarization mismatch angle) of $\Delta\Psi$, the attenuation of the co-polar signal given by

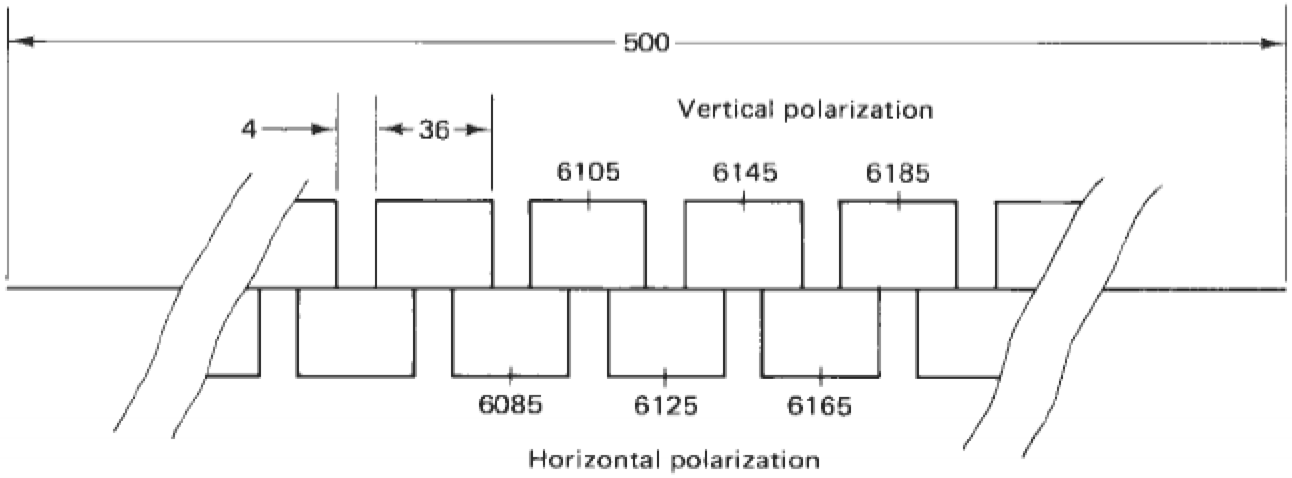
$$A_{\mathrm{PR}} = -20\ \log(\cos \Delta\Psi)$$

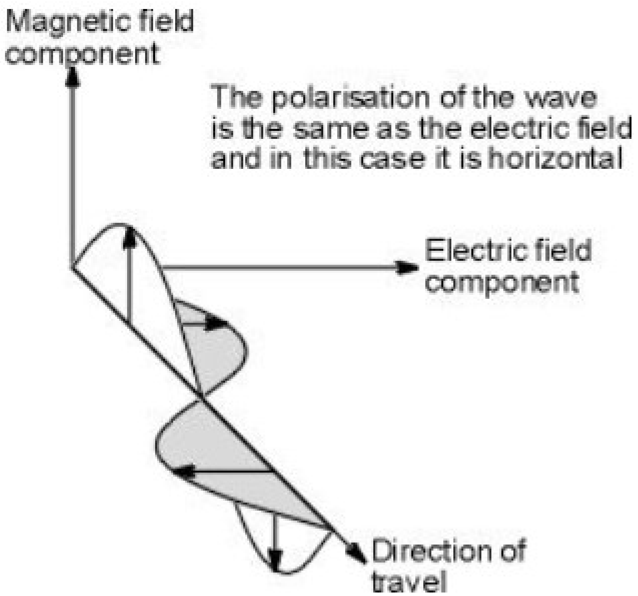
where, A_{PR} is the attenuation due to polarization rotation in dB.
The mismatch also produces a cross-polarized component, which reduces the cross-polarization discrimination ($\mathrm{X_{PD}}$), given by

$$\mathrm{X_{PD}} = -20\ \log(\tan \Delta\Psi)$$

where, X_{PD} is the cross-polarization discrimination in dB.
The magnitudes of attenuation and $\mathrm{X_{PD}}$ due to polarization mismatch will be 0.1 dB and 16 dB respectively at 4 GHz for which $\Delta\Psi = 9^{\circ}$.

- Polarization Discrimination can be obtained by making the carriers having the same frequency but with difference in polarization





➔ An antenna is a transducer that converts radio frequency electric current to electromagnetic waves that are then radiated into space. The electric field or “E” plane determines the polarization or orientation of the radio wave. In general, most antennas radiate either linear or circular polarization.

➔ Polarization is an important design consideration. The polarization of each antenna in a system should be properly aligned. Maximum signal strength between stations occurs when both stations are using identical polarization

➔ Most communications systems use either vertical, horizontal or circular polarization

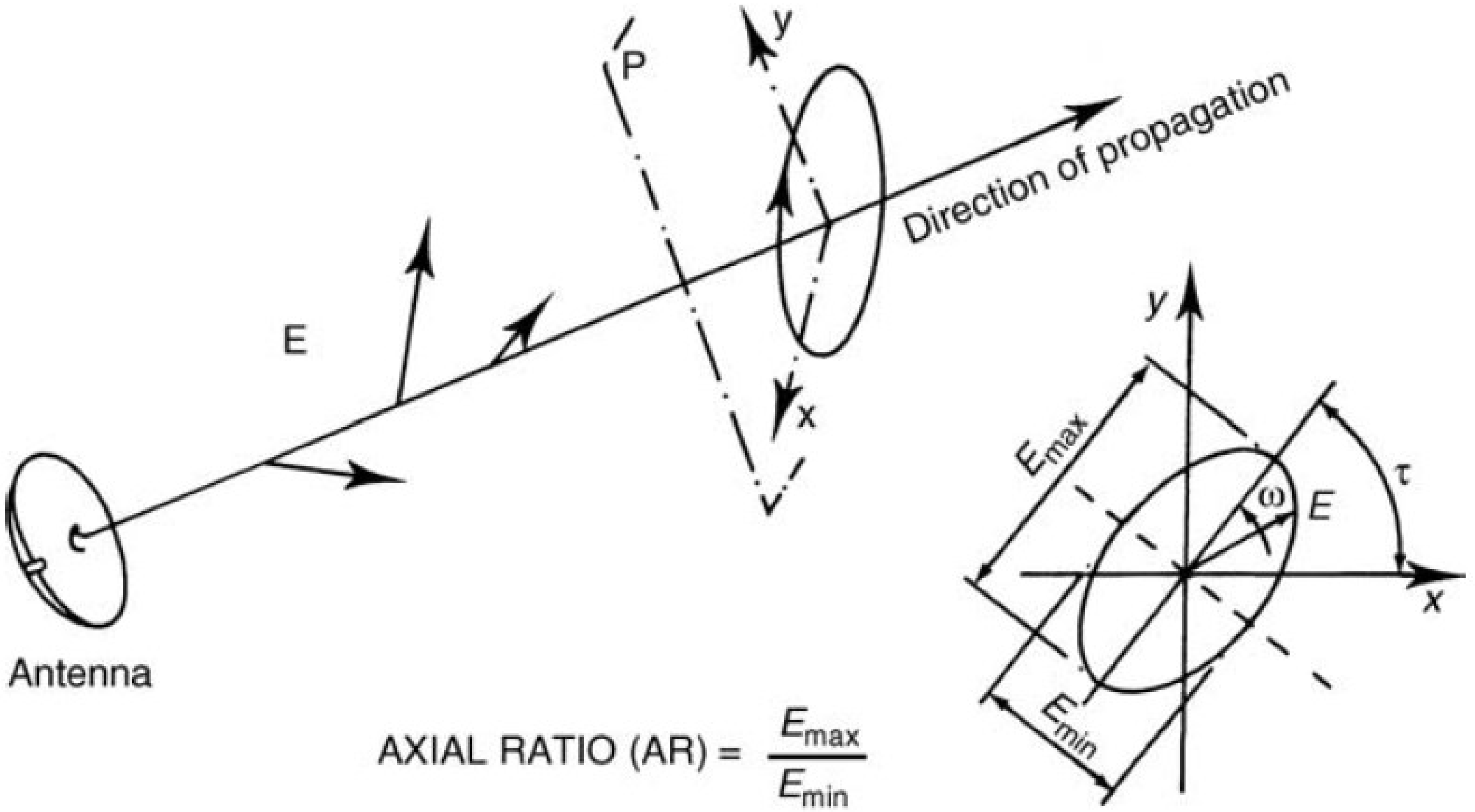
A linear polarized antenna radiates wholly in one plane containing the direction of propagation. In a circular polarized antenna, the plane of polarization rotates in a circle making one complete revolution during one period of the wave.

➔ If the rotation is clockwise looking in the direction of propagation, the sense is called right-hand-circular (RHC). If the rotation is counterclockwise, the sense is called left-hand-circular (LHC).

➔ An antenna is said to be vertically polarized (linear) when its electric field is perpendicular to the Earth's surface. An example of a vertical antenna is a broadcast tower for AM radio

➔ Horizontally polarized (linear) antennas have their electric field parallel to the Earth's surface. Television transmissions in the USA use horizontal polarization.

➔ A circular polarized wave radiates energy in both the horizontal and vertical planes and all planes in between.



Polarisation is characterised by the following parameters:
—direction of rotation
—axial ratio (AR): AR = E_{max}/E_{min}, that is the ratio of the major and minor axes of the ellipse. When the ellipse is a circle (axial ratio = 1 = 0 dB), the polarisation is said to be circular. When the ellipse reduces to one axis (infinite axial ratio: the electric field maintains a fixed direction), the polarisation is said to be linear;
—inclination τ of the ellipse.

Two waves are in orthogonal polarisation if their electric fields describe identical ellipses in opposite directions. In particular, the following can be obtained:

—two orthogonal circular polarisations described as right-hand circular and left-hand circular (the direction of rotation is for an observer looking in the direction of propagation)
—two orthogonal linear polarisations described as horizontal and vertical (relative to a local reference).

NOTE: Antenna designed to transmit or receive a wave of given polarisation can neither transmit nor receive in the orthogonal polarisation. (Frequency Reuse by depolarisation)

—The cross-polarisation isolation (two linear):

$$\mathrm{XPI\ (dB)} = 20\log(a_C/b_X) \text{ or } 20\log(b_C/a_X) \quad \mathrm{(dB)}$$

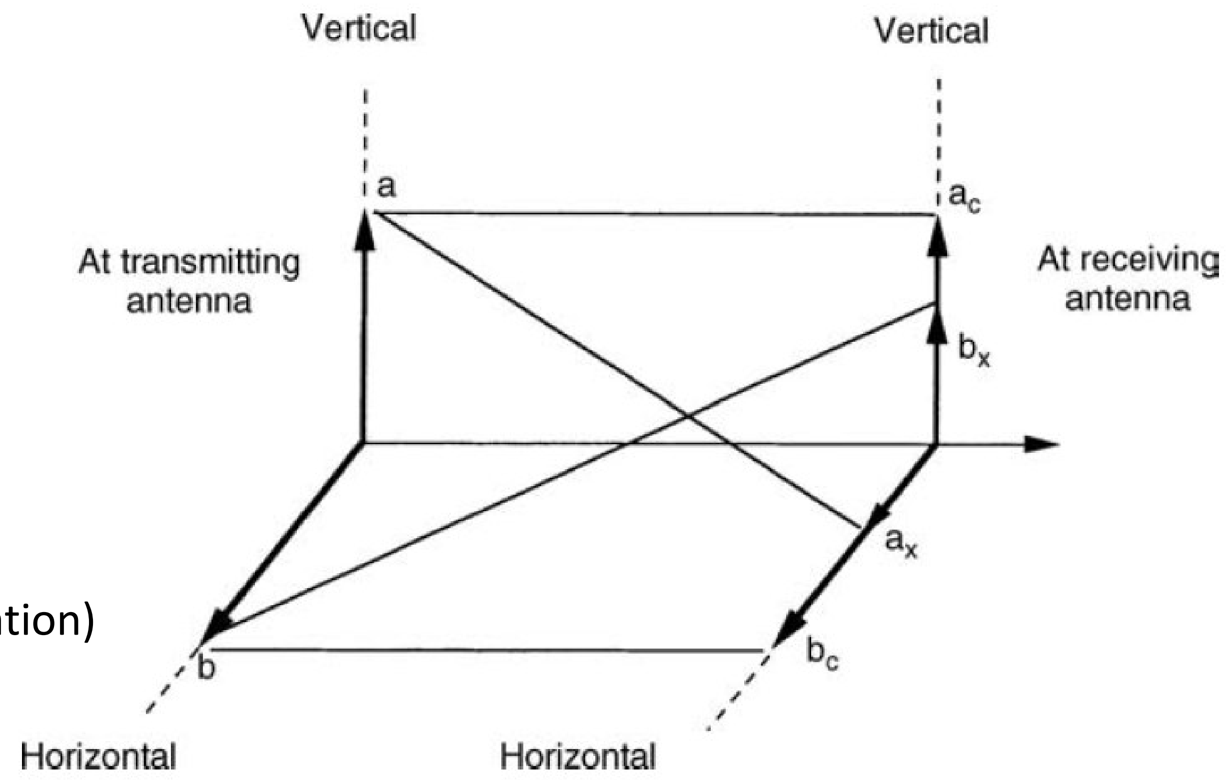
—The cross-polarisation discrimination (one linear)

$$\mathrm{XPD\ (dB)} = 20\log(a_C/a_X) \quad \mathrm{(dB)}$$

—The cross-polarisation isolation (quasi-circular polarisation)

$$\mathrm{XPD} = 20\log[(\mathrm{AR}+1)/(\mathrm{AR}-1)] \quad \mathrm{(dB)}$$

$$\mathrm{AR} = (10^{\mathrm{XPD}/20}+1)/(10^{\mathrm{XPD}/20}-1)$$



EXAMPLE

Compute the free-space path loss in decibels for the following conditions:

1. For a path length of 10 km at 4 GHz operating frequency
2. Earth station transmitting antenna EIRP $= 50\,\text{dBW}$, satellite receiving antenna gain $=$ $20\,\text{dB}$ and received power at satellite $= -120\,\text{dBW}$

Solution:

(a) Path length, $R = 10\,\text{km}$, operating frequency, $f = 4\,\text{GHz}$

$$\text{Operating wavelength, } \lambda = c/f = (3 \times 10^8/4 \times 10^9)\,\text{m} = 0.075\,\text{m}$$

$$\begin{aligned}\text{Path loss (in dB)} &= 20\log(4\pi R/\lambda) \\ &= 20\log(4\pi \times 10\,000/0.075) \\ &= 124.48\,\text{dB}\end{aligned}$$

(b) Path loss can be computed from:

$$\text{Received power} = \text{EIRP} + \text{receiving antenna gain} - \text{path loss}$$

$$\begin{aligned}\text{Therefore, path loss} &= \text{EIRP} + \text{receiving antenna gain} - \text{received power} \\ &= 50 + 20 - (-120) = 50 + 20 + 120 = 190\,\text{dB}\end{aligned}$$

EXAMPLE

Under certain atmospheric conditions, a 2 GHz linearly polarized signal experiences a rotation of its plane of polarization by 75°. How much polarization rotation would have been experienced by a 10 GHz signal under similar atmospheric conditions? Also determine the attenuation (in dB) experienced by the co-polar component due to polarization rotation if it is not corrected for at the receiving antenna.

Solution: Polarization rotation is inversely proportional to the square of the operating frequency. Here, the frequency has increased by a factor of 5. Therefore, the polarization rotation angle will decrease by a factor of 25; i.e.
Polarization rotation experienced $= 75/25 = 3^\circ$
Now, the attenuation due to the polarization loss can be computed from

$$\text{Attenuation (in dB)} = -20\log(\cos\Delta\Psi)$$

where $\Delta\Psi$ = polarization mismatch angle. In the first case, $\Delta\Psi = 75^\circ$. Therefore,

$$\text{Attenuation} = -20\log(\cos 75^\circ) = -20\log(0.2588)$$
$$= -20 \times (-0.587) = 11.74\,\text{dB}$$

In the second case, $\Delta\Psi = 3^\circ$. Therefore,

$$\text{Attenuation} = -20\log(\cos 3^\circ) = -20\log(0.9986)$$
$$= -20 \times (-0.0006) = 0.012\,\text{dB}$$

Ionospheric Scintillation

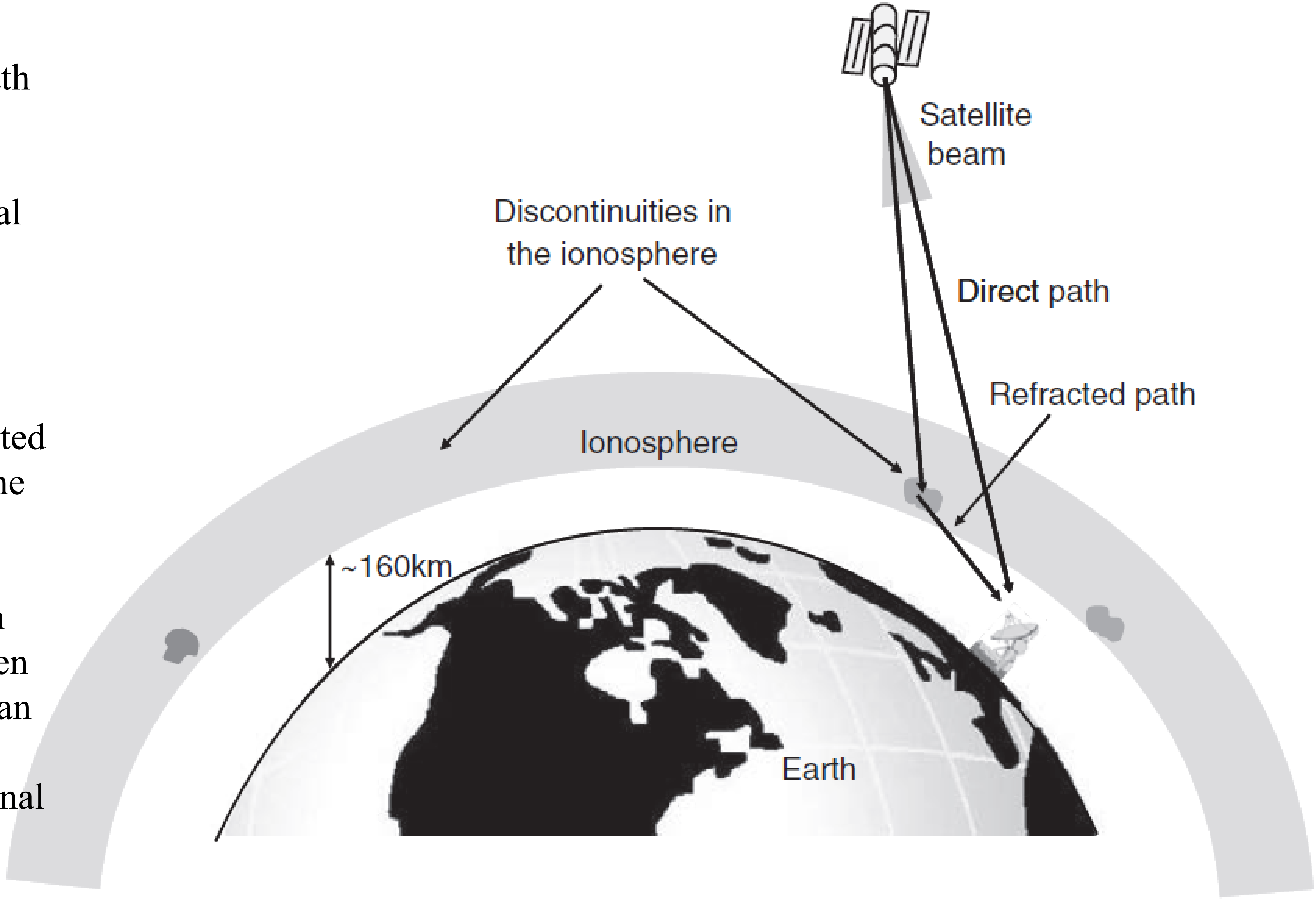
➔ scintillation is nothing but the rapid fluctuations of the signal amplitude, phase, polarization or angle-of-arrival

➔ In the ionosphere, scintillation occurs due to small scale refractive index variations caused by local electron concentration fluctuations. The total electron concentration (total number of electrons existing in a vertical column of 1m$_2$ area) of the ionosphere increases by two orders of magnitude during the day as compared to night due to the energy received from the sun.

➔ This rapid change in the value of total electron concentration from the daytime value to the night-time value gives rise to irregularities in the ionosphere.

➔ the signal reaches the receiving antenna via two paths, the direct path and the refracted path.
Multipath signals can lead to both signal enhancement as well as signal cancellation depending upon the phase relationship with which they arrive at the receiving antenna.

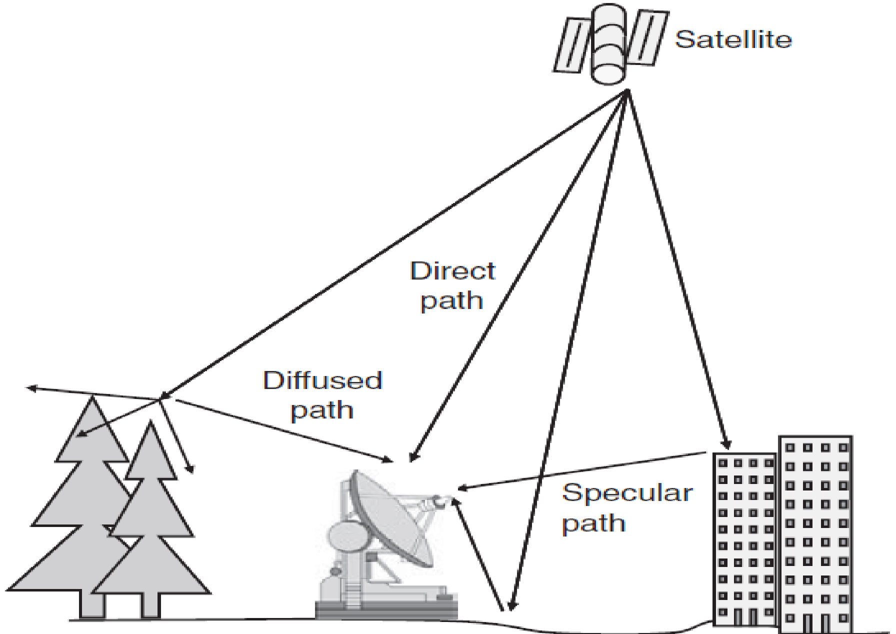
- the signal reaches the receiving antenna via two paths, the direct path and the refracted path.
- Multipath signals can lead to both signal enhancement as well as signal cancellation depending upon the phase relationship with which they arrive at the receiving antenna.
- The resultant signal is a vector addition of the direct and the refracted signal. In the extreme case, when the strength of the refracted signal is comparable to that of the direct signal, cancellation can occur when the relative phase difference between the two is 180°. On the other hand, an instantaneous recombination of the two signals in phase can lead to signal amplification up to 6 dB.

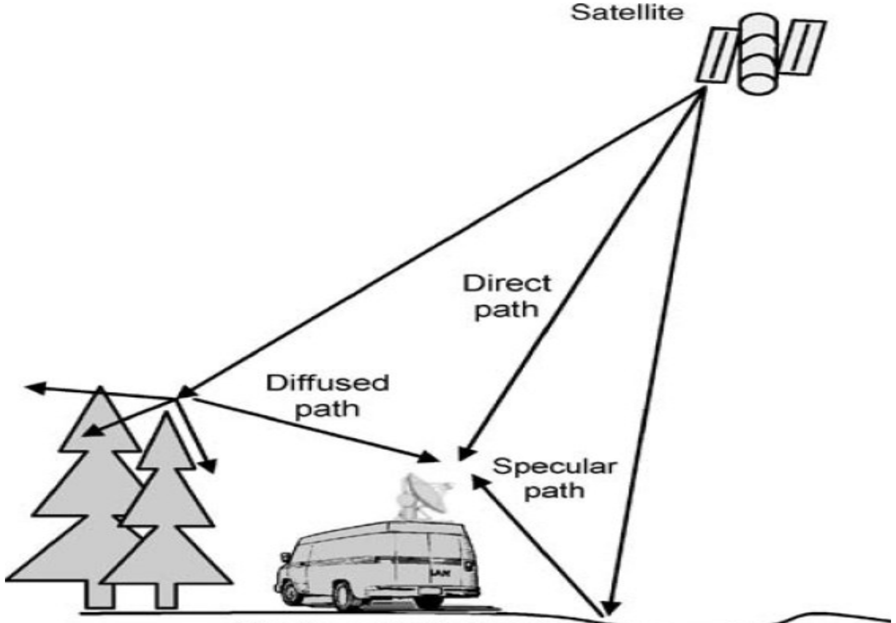


Fading due to Multipath Signals

- In the case of fixed satellite terminals, the situation remains more or less the same and does not change with time as long as the satellite remains in the same position with respect to the satellite terminal.

- In the case of mobile satellite terminals, the situation keeps changing with time. In a typical case, a mobile terminal could receive the direct signal and another signal reflected off the highway, buildings, neighbouring hills or trees. The relative phase difference between the two signals could produce either a signal enhancement or fading. Moreover, the fading signal varies with time as the satellite moves with respect to the points of reflection.





Noise Considerations:

- The quality of signal received at the Earth station is strongly dependent on the carrier-to-noise ratio of the satellite link.

- Satellite communication systems, such as the geostationary satellite communication systems, are susceptible to noise because of their inherent low received power levels.

- The sources of noise include natural and man-made sources, as well as the noise generated in the Earth station and satellite equipment. While the man-made noise mainly arises from **electrical equipment** and is almost insignificant above 1 GHz, the natural sources of noise include **solar radiation**, **sky noise**, **background noise** contributed by Earth, **galactic** noise due to electromagnetic waves emanating from radio stars in the galaxy and the **atmospheric noise** caused by lightening flashes and absorption by oxygen and water vapour molecules followed by re-emission of radiation. Sky noise and solar noise can be avoided by proper orientation and directionality of antennas. Galactic noise is insignificant above 1 GHz. Noise due to lightning flashes is also negligible at satellite frequencies.

- It is mainly the noise generated in the equipment where attention primarily needs to be paid.

Thermal Noise

- Thermal noise is generated in any resistor or resistive component of any impedance due to random motion of molecules, atoms and electrons.

- It is called **thermal noise** as the temperature of a body is the statistical RMS value of the velocity of motion of these particles. It is also called ‘**white’ noise** because, due to randomness of the motion of particles, the noise power is evenly spread over the entire frequency spectrum. It is also known as **Johnson noise**.

According to kinetic theory, the motion of these particles ceases at absolute zero temperature, i.e. *zero degree kelvin*. Therefore, the noise power generated in a resistor or resistive component is directly proportional to its absolute temperature, in addition to the bandwidth over which it is measured;

$$P_{\mathrm{n}} \propto TB = kTB$$

where

$$\begin{aligned} T &= \text{absolute temperature (in K)} \\ B &= \text{bandwidth of interest (in Hz)} \\ k &= \text{Boltzmann's constant} = 1.38 \times 10^{-23}\ \mathrm{J/K} \\ P_{\mathrm{n}} &= \text{noise power output of a resistor (in W)} \end{aligned}$$

Thermal noise power at room temperature (T=290 K) in dBm (decibles relative to a power level of 1mW) can also be computed from:

$$P_n = -174 + 10 \log B$$

If the resistor is considered as a noise generator with an equivalent noise voltage equal to V_n, then this noise generator will transfer maximum noise power P_n to a matched load that is given by:

$$P_n = \frac{V_n^{\;2}}{4R}$$

which gives expression for noise voltage (V_n) as:

$$V_n = \sqrt{}(4kTRB)$$

Expression for noise current (I_n) can be deduced from the expression for noise voltage and is given by:

$$I_n = \sqrt{}(4kTB/R)$$

Another term that is usually defined in this context is the noise power spectral density given by:

$$P_{no} = kT$$

Noise Figure

The noise figure F of a device can be defined as the ratio of the signal-to-noise power at its input to the signal-to-noise power at its output; i.e.

$$F = \frac{S_i/N_i}{S_o/N_o} = \frac{N_o}{(S_o/S_i)N_i} = \left(\frac{N_o}{N_i}\right)\left(\frac{1}{G}\right)$$

where

S_i = available signal power at the input
N_i = available noise power at the input
S_o = available signal power at the output
N_o = available noise power at the output (in a noiseless device)
G = power gain over the specified bandwidth = S_o/S_i

Now, $N_i = kT_iB$, where T_i is the ambient temperature in kelvin. Therefore, the noise figure (F) is expressed as

$$F = \frac{N_o}{GkT_iB}$$

- The actual amplifier, however, introduces some noise, which is added to the output noise power. If the noise power introduced is ΔN, then

$$N_o = GkT_iB + \Delta N$$

$$F = \frac{GkT_iB + \Delta N}{GkT_iB} = 1 + \frac{\Delta N}{GkT_iB}$$

- The noise figure is thus the ratio of the actual output noise to that which would remain if the device itself did not introduce any noise. In the case of a noiseless device, $\Delta N = 0$, which gives $F = 1$.

- Thus the noise figure in the case of an ideal device is unity. Any value of the noise figure greater than unity means a noisy device.

Noise Temperature

equivalent noise temperature T_e. It is the temperature of a resistance that would generate the same noise power at the output of an ideal (i.e. noiseless) device as that produced at the output by an actual device when terminated at its input by a noiseless resistance, i.e. a resistance at absolute zero temperature.

Now, noise generated by the device $\Delta N = GkT_eB$, which when substituted in the expression for noise figure mentioned above gives:

$$F = 1 + \frac{T_e}{T_i} \qquad \text{or} \qquad T_e = T_i(F - 1)$$

If L is the loss factor, then the gain G for this attenuator can be expressed as $G = 1/L$. The expression for the total noise power at the output of the attenuator (N_o) can be written as:

$$N_o = GkT_iB + GkT_eB = \frac{kT_iB + kT_eB}{L}$$

where T_e = effective noise temperature of the attenuator. If the attenuator is considered to be at the same temperature (T_i) as that of the source resistance from which it is fed, then the value of output noise (N_o) is given by:

$$N_o = kT_iB$$

This gives

$$\frac{kT_iB + kT_eB}{L} = kT_iB$$

or

$$\frac{T_i + T_e}{L} = T_i \qquad \text{which gives } T_e = T_i(L - 1)$$

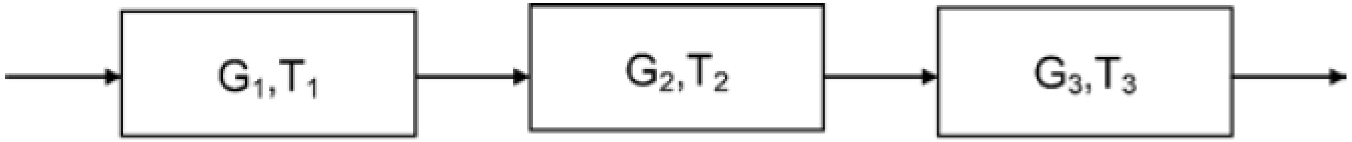
This expression gives the effective noise temperature of a noise source at temperature T_i, followed by a resistive attenuator having a loss factor L.

Noise Figure and Noise Temperature of Cascaded Stages

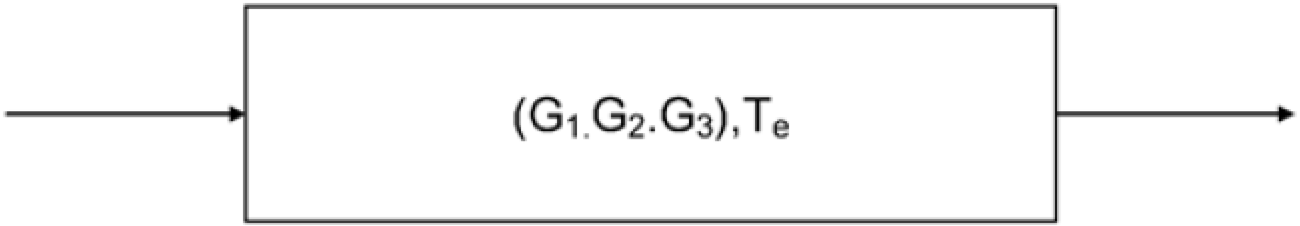
A system such as a receiver would have a large number of individual building blocks connected in series and it is important to determine the overall noise performance of this cascaded arrangement.

Consider a cascaded arrangement of three stages with individual gains given as G_1, G_2 and G_3 and input noise temperature parameters as T_1, T_2 and T_3. In the case of the cascaded arrangement, the total noise power at the output (N_{TO}) can be computed from:

$$N_{TO} = G_3 k T_3 B + G_3 G_2 k T_2 B + G_3 G_2 G_1 k T_1 B$$
$$= G_3 G_2 G_1 k B \left(T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} \right)$$



(a)



(b)

Now if T_e is the effective input noise temperature of the cascaded arrangement of the three stages, which would have a gain of $G_3G_2G_1$, then the total noise power at the output (N_{TO}) can also be computed from:

$$N_{TO} = G_3G_2G_1kT_eB$$

Equating the two expressions for N_{TO} gives

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2}$$

Generalizing the expression for n stages gives

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1G_2} + \frac{T_4}{G_1G_2G_3} + \cdots + \frac{T_n}{G_1G_2G_3 \ldots G_{n-1}}$$

The same expression can also be written in terms of noise figure specifications of individual stages as:

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2} + \frac{F_4 - 1}{G_1G_2G_3} + \frac{F_n - 1}{G_1G_2G_3 \ldots G_{n-1}}$$

where, F_1, F_2, $F_3 \ldots F_n$ are the noise figures for stages 1,2,3 . . . n respectively and G_1, G_2, $G_3 \ldots G_n$ are the gains for stages 1,2,3 . . . n respectively. It may be mentioned here that the gain values in the expression are not in decibels.

Antenna Noise Temperature

As is evident from these expressions, the noise performance of the overall system is largely governed by the noise performance of the first stage. That is why it is important to have the first stage with as low noise as possible.

The antenna noise temperature is a measure of the noise entering the receiver via the antenna.

Noise from these sources could enter the receiver both through the main lobe as well as through the side lobes of the directional pattern of the receiving antenna. Thus, the noise output from a receiving antenna is a function of the direction in which it is pointing, its directional pattern and the state of its environment.

The noise performance of an antenna, as mentioned before, can be expressed in terms of a noise temperature called the antenna noise temperature. If the antenna noise temperature is T_A K, it implies that the noise power output of the antenna is equal to the thermal noise power generated in a resistor at a temperature of T_A K

➔ The noise temperature of the antenna can be computed by integrating the contributions of all the radiating bodies whose radiation lies within its directional pattern. It is given by:

$$T_A = \left(\frac{1}{4\pi}\right) \int \int G(\theta, \phi)\, T_b(\theta, \phi)\, \sin\theta \, d\theta \, d\phi$$

where

$\theta =$ azimuth angle
$\phi =$ elevation angle
$G(\theta, \phi) =$ antenna gain in the θ and ϕ directions
$T_b(\theta, \phi) =$ brightness temperature in the θ and ϕ directions

There are two possible situations to be considered here. One is that of the satellite antenna when the uplink is referred to and the other is that of the Earth station antenna when the downlink is referred to.

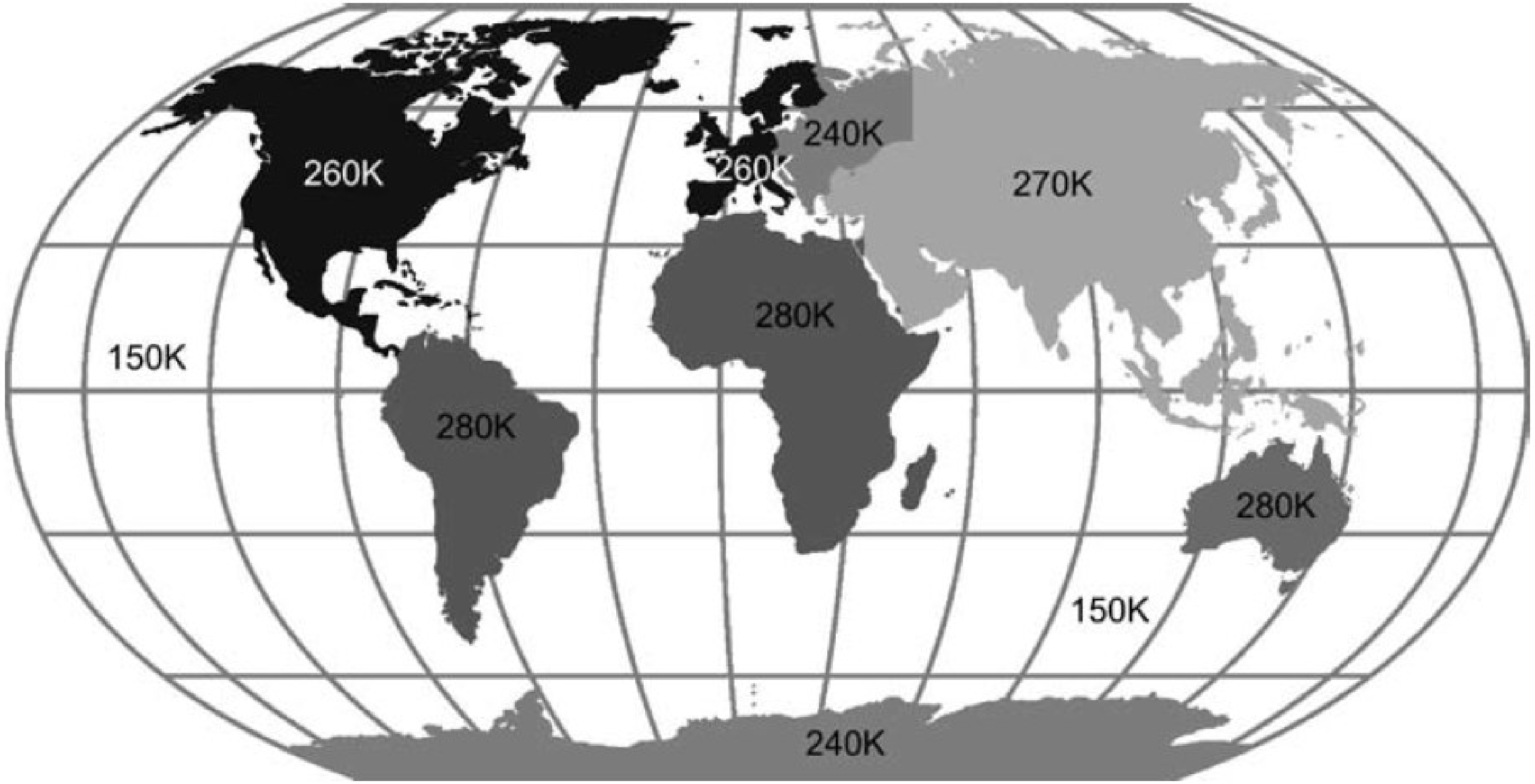
These two cases need to be considered separately because of the different conditions that prevail both in terms of the antenna's directional pattern and the significant sources of noise.

➔ In the case of satellite antenna (uplink scenario), the main sources of noise are Earth and outer space.

➔ the noise contribution of the Earth depends upon the orbital position of the satellite and the antenna beam width. In case, the satellite antenna's beam width is more or less equal to the angle-of-view of Earth from the satellite, which is 17.5° for a geostationary satellite, the antenna noise temperature depends upon the frequency of operation and orbital position.

➔ In case the beam width is smaller as in case of a spot beam, the antenna noise temperature depends upon the frequency of operation and the area being covered on Earth. Different areas radiate different noise levels

brightness temperature model of the Earth in the Ku band.



- In the case of the Earth station antenna (downlink scenario), the main sources of noise that contribute to the antenna noise temperature include the sky noise and the ground noise.

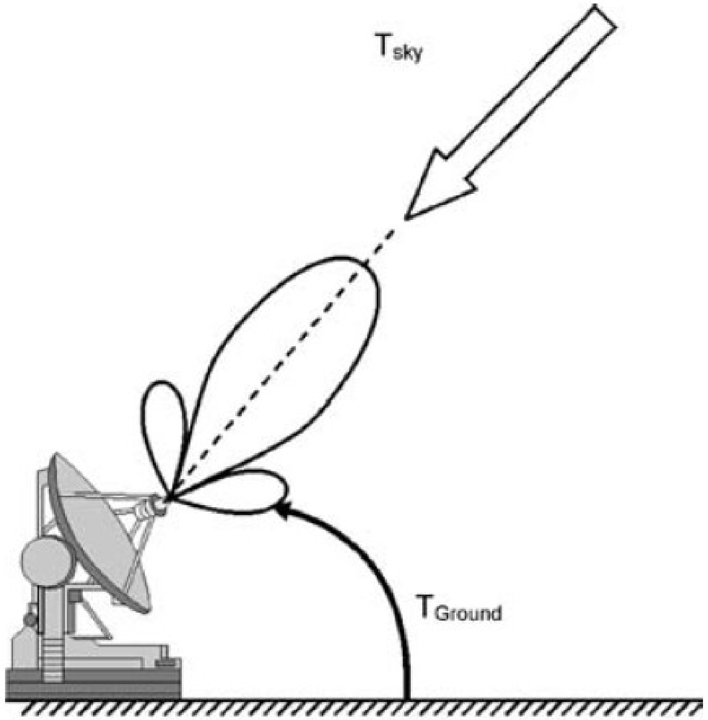
- The sky noise is primarily due to sources such as radiation from the sun and the moon and the absorption by oxygen and water vapour in the atmosphere accompanied by re-emission.

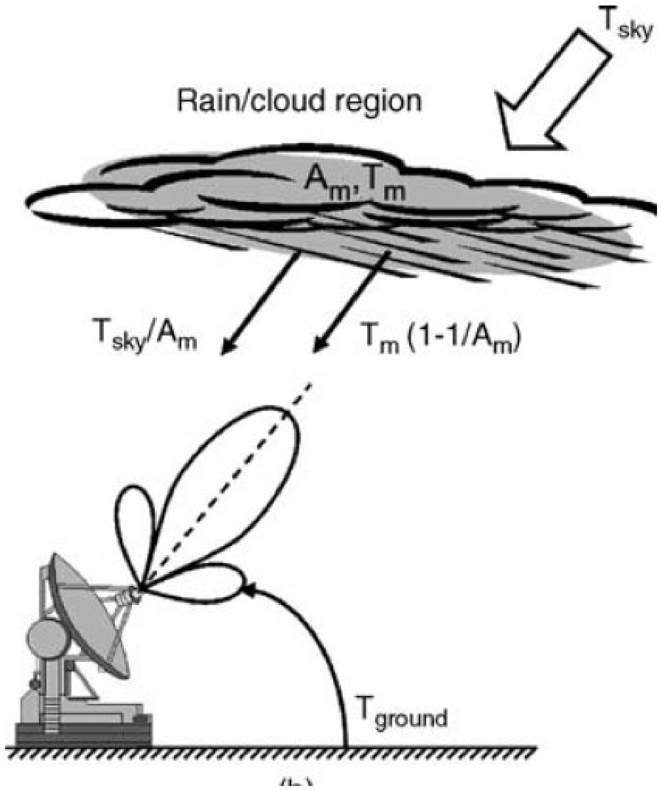
- The noise from other sources such as cosmic noise originating from hot gases of stars and interstellar matter, galactic noise due to electromagnetic waves emanating from radio stars in the galaxy is negligible at frequencies above 1 GHz.

- Here again there are two distinctly different conditions, one of clear sky devoid of any meteorological formations and the other that of sky with meteorological formations such as clouds, rain, etc.

In the clear sky conditions, the noise contribution is from sky noise and ground noise. The sky noise enters the system mainly through the main lobe of the antenna's directional pattern and the ground noise enters the system mainly through the side lobes and only partly through the main lobe, particularly at low elevation angles

if the attenuation and noise temperature figures for the meteorological formation are A_m and T_m respectively, then the noise contribution due to sky noise is T_{sky}/A_m and that due to the meteorological formation is $T_m(1 - 1/A_m)$.

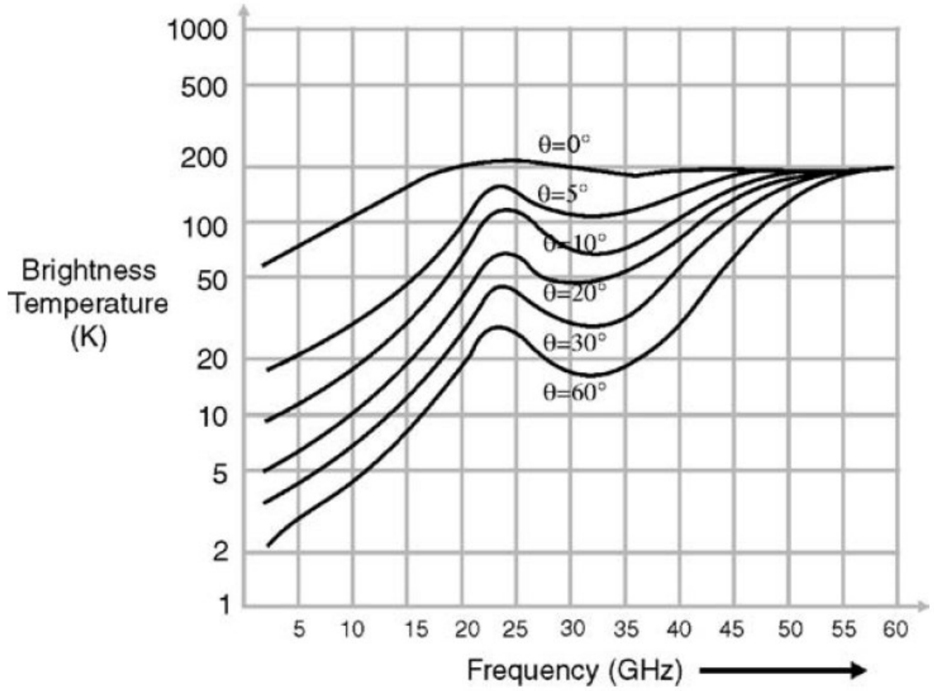


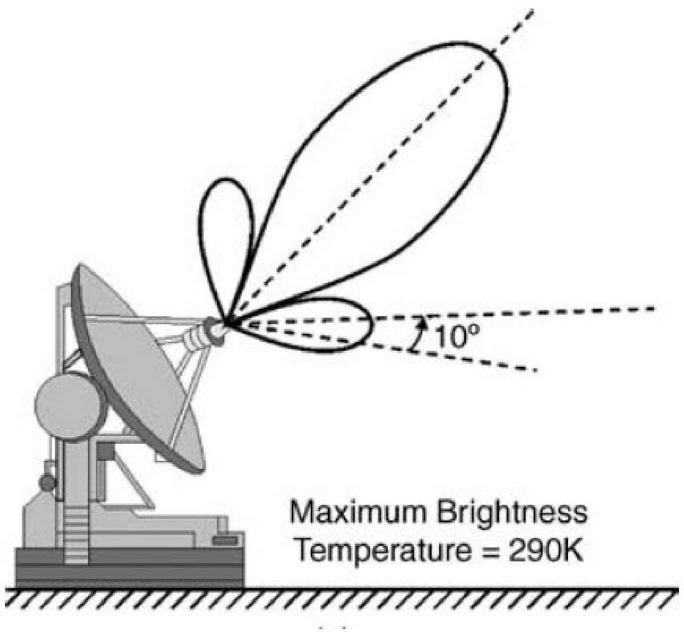


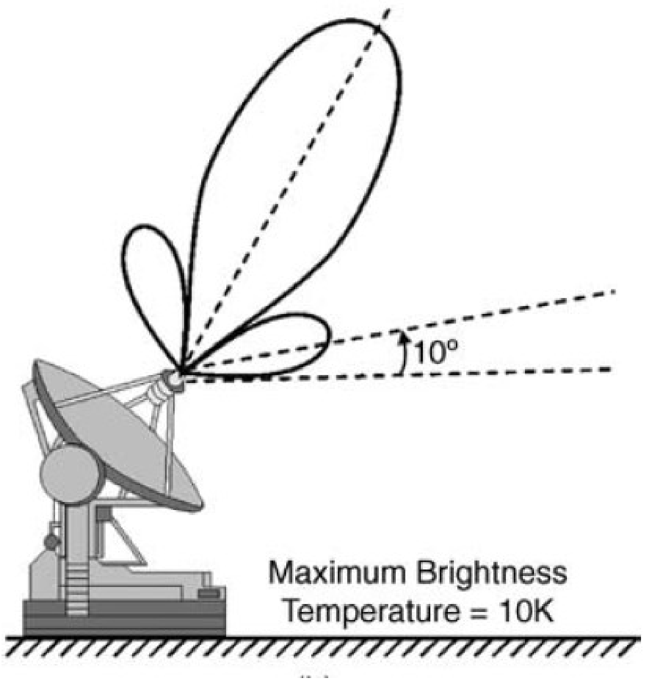
The sky noise contribution can be computed knowing the brightness temperature of the sky in the direction of the antenna within its beam width.

Family of curves showing the brightness temperature of clear sky as a function of operating frequency for various practical elevation angles ranging from 5 to 60 Degrees is shown below.

The contribution of ground noise can also be similarly computed knowing the brightness temperature of the ground. The brightness temperature of the ground may be as large as 290K for a side lobe whose elevation angle is less than −10 degrees and as small as 10K for an elevation angle greater than 10 degrees.







Overall System Noise Temperature

The overall system is a cascaded arrangement of the antenna, the feeder connecting the antenna output to the receiver input and the receiver as shown in Figure 7.14. Expressions can be written for the noise temperature at two points, one at the output of the antenna, i.e. input of the feeder, and second at the input of the receiver. The expression for the system noise temperature with reference to the output of the antenna (T_{SAO}) can be written as:

$$T_{SAO} = T_A + T_F(L_F - 1) + T_e L_F$$

where

T_A = antenna noise temperature
T_F = thermodynamic temperature of the feeder, often taken as the ambient temperature
L_F = attenuation factor of the feeder
T_e = effective input noise temperature of the receiver

The expression for the system noise temperature when referred to the receiver input (T_{SRI}) can be written as

$$T_{SRI} = \frac{T_A}{L_F} + T_F\left(\frac{L_F - 1}{L_F}\right) + T_e \qquad (7.35)$$

The above expression for the noise temperature takes into account the noise generated by the antenna and the feeder together with the receiver noise. The two expressions for the noise temperature given above highlight another very important point that the noise temperature at the antenna output is larger than the noise temperature at the receiver input by a factor of L_F. This underlines the importance of having a minimum losses before the first RF stage of the receiver.

EXAMPLE

A 12 GHz receiver consists of an RF stage with gain $G_1 = 30$ dB and noise temperature $T_1 = 20$ K, a down converter with gain $G_2 = 10$ dB and noise temperature $T_2 = 360$ K and an IF amplifier stage with gain $G_3 = 15$ dB and noise temperature $T_3 = 1000$ K.

Calculate the effective noise temperature and noise figure of the system. Take the reference temperature to be 290 K.

Solution: The effective noise temperature T_e can be computed from

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2}$$

Now,

$$T_1 = 20\,\mathrm{K}, \qquad T_2 = 360\,\mathrm{K} \qquad \text{and} \qquad T_3 = 1000\,\mathrm{K}$$

$$G_1 = 30\,\mathrm{dB} = 1000, \qquad G_2 = 10\,\mathrm{dB} = 10$$

Therefore,

$$T_e = 20 + 360/1000 + 1000/(1000 \times 10) = 20 + 0.36 + 0.1 = 20.46\,\mathrm{K}$$

The system noise figure F can be computed from

$$F = 1 + \frac{T_e}{T_i} \qquad \text{where } T_i = 290\,\mathrm{K}$$

$$= 1 + \frac{20.46}{290} = 1.07$$

EXAMPLE

For the receiver of the previous example, compute the noise figure specifications of the three stages and then compute the overall noise figure from the individual noise figure specifications.

Solution: In problem 7.4, $T_1 = 20\,\mathrm{K}$, $T_2 = 360\,\mathrm{K}$, $T_3 = 1000\,\mathrm{K}$, $T_\mathrm{i} = 290\,\mathrm{K}$, $G_1 = 30\,\mathrm{dB} = 1000$, $G_2 = 10\,\mathrm{dB} = 10$. Let F_1, F_2 and F_3 be the noise figure specifications of the three stages. Then

$$F_1 = 1 + \frac{T_1}{T_\mathrm{i}} = 1 + \frac{20}{290} = 1.069$$

$$F_2 = 1 + \frac{T_2}{T_\mathrm{i}} = 1 + \frac{360}{290} = 2.24$$

$$F_3 = 1 + \frac{T_3}{T_\mathrm{i}} = 1 + \frac{1000}{290} = 4.45$$

The overall noise figure can be computed from

$$\begin{aligned} F &= F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} \\ &= 1.069 + \frac{2.24 - 1}{1000} + \frac{4.45 - 1}{10\,000} \\ &= 1.069 + 0.001\,24 + 0.000\,345 = 1.07 \end{aligned}$$

EXAMPLE

The effective input noise temperature of a satellite receiver is 30 K when the effect of noise contributions from the antenna and feeder are not taken into consideration. If the receiver is fed from an antenna having a noise temperature of 50 K via a feeder with a loss factor of 2.5 dB, determine the effective input noise temperature of the receiver considering the effect of the antenna and the feeder noise contributions. Assume $T_i = 290$ K and also that the feeder is at a temperature T_i. Also compute the noise figure in the two cases in decibels.

Solution: Loss factor L of the feeder = 2.5 dB = 1.778
The contribution of the antenna noise temperature when referred to the input of the receiver is given by

$$\left(\frac{T_A}{L}\right) = \frac{50}{1.778} = 28.1\,\text{K}$$

The contribution of the feeder noise when referred to the input of the receiver is given by

$$\frac{T_F(L-1)}{L} = \frac{T_i(L-1)}{L} = \frac{290(1.778-1)}{1.778} = 290\left(\frac{0.778}{1.778}\right) = 126.9\,\text{K}$$

Therefore, the effective input noise temperature of the receiver taking into account the effect of noise contributions from the antenna and feeder is given by

$$28.1 + 126.9 + 30 = 185\,\mathrm{K}$$

$$\text{Noise figure in the first case} = 1 + \frac{30}{290} = 1.103 = 0.426\,\mathrm{dB}$$

$$\text{Noise figure in the second case} = 1 + \frac{185}{290} = 1.638 = 2.14\,\mathrm{dB}$$

Interference-related Problems:

Major Sources of Interference are :

1. Intermodulation distortion

2. Interference between the satellite and the terrestrial link sharing the same frequency band

3. Interference between two satellites sharing the same frequency band

4. Interference arising out of cross-polarization in frequency re-use systems

5. Adjacent channel interference inherent to FDMA systems

Intermodulation Distortion

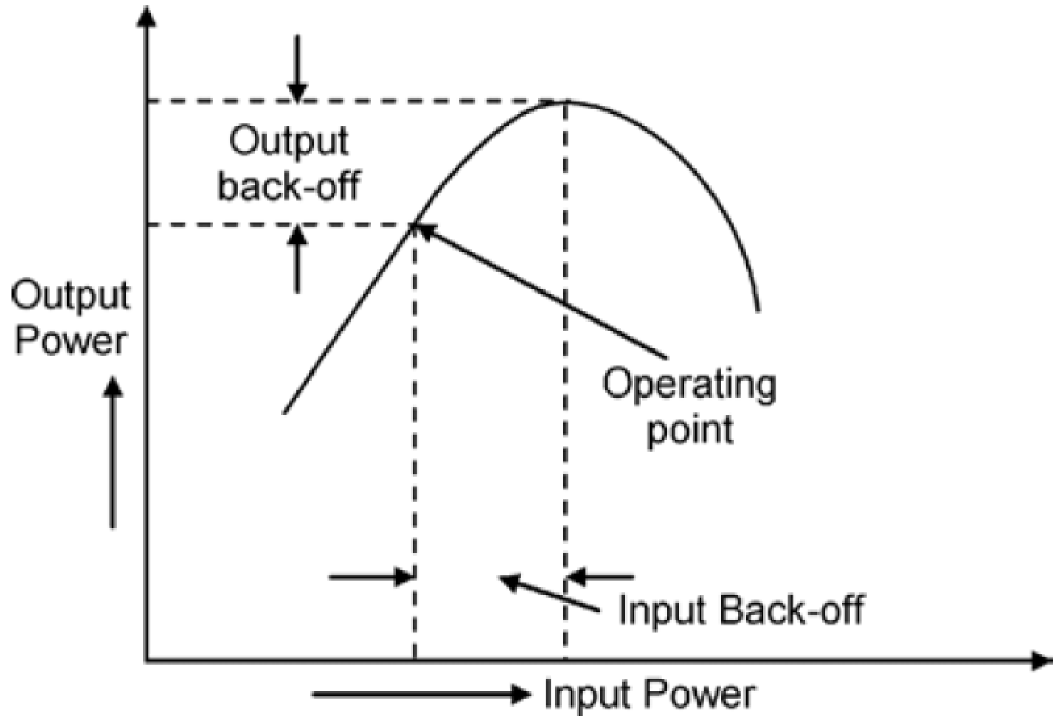
result of the generation of intermodulation products within the satellite transponder as a result of **amplification of multiple carriers in the power amplifier**

Traveling Wave Tube Amplifiers (TWTA) is a specialized vacuum tube that is used in electronics to amplify radio frequency (RF) signals in the microwave range, it is commonly used in satellite communication links.

The generation of intermodulation products is due to both amplitude nonlinearity and phase nonlinearity

Intermodulation products can be avoided by operating the amplifier in the linear region by reducing or backing off the input drive.

But a reduced input drive leads to a reduced output power. This results in a downlink power-limited system that is forced to operate at a reduced capacity.



Problem :Intermodulation products are generated whenever more than one signal is to be amplified by the amplifier with non-linear characteristics.

Solution: Filtering helps to remove the intermodulation products.

Limitations: but when these products are within the bandwidth of the amplifier, filtering is of not much use.

The transfer characteristics of an amplifier can be written as:

$$V_{out} = AV_{in} + B\,(V_{in})^2 + C\,(V_{in})^3$$

Where, A $\gg$ B $\gg$ C

It may be mentioned here that the intermodulation products are mainly generated by the third order component in the equation as the third-order intermodulation products have frequencies close to the input frequencies and hence lie within the transponder bandwidth.

Let us consider that the signal applied to the input of the amplifier is given by

$$V_{in} = V_1 \cos \omega_1 t + V_2 \cos \omega_2 t$$

In other words, two unmodulated carriers at frequencies f_1 and f_2 are applied to the input of the amplifier. Then the output of the amplifier is given by

$$V_{out} = A\,[V_1 \cos \omega_1 t + V_2 \cos \omega_2 t] + B\,[V_1 \cos \omega_1 t + V_2 \cos \omega_2 t]^2$$

$$+C\,[V_1 \cos \omega_1 t + V_2 \cos \omega_2 t]^3$$

The first term is the linear term and it amplifies the input signal by A and represents the desired output of the amplifier (V_{desout}).

$$V_{desout} = A\,[V_1 \cos \omega_1 t + V_2 \cos \omega_2 t]$$

The total desired power output from the amplifier, referenced to a one ohm load is, therefore given by

$$P_{desout} = \frac{1}{2} A^2 V_1^2 + \frac{1}{2} A^2 V_2^2 = A^2 (P_1 + P_2)$$

Where,

$P_1 = 1/2(V_1{}^2)$
$P_2 = 1/2(V_2{}^2)$

The second term of Vout is the second order term and can be expanded as

$$\begin{aligned}
V_{2out} &= B\,[V_1 \cos\omega_1 t + V_2 \cos\omega_2 t]^2 \\
&= B\left[V_1^2 \cos^2\omega_1 t + V_2^2 \cos^2\omega_2 t + 2V_1 V_2 \cos\omega_1 t \cos\omega_2 t\right] \\
&= B\left[V_1^2 \{(\cos 2\omega_1 t + 1)/2\} + V_2^2 \{(\cos 2\omega_2 t + 1)/2\} + V_1 V_2 \{\cos(\omega_1 + \omega_2) t\right. \\
&\quad + \cos(\omega_1 - \omega_2) t\}]
\end{aligned}$$

The term V_{2out} contains frequency components $2f_1$, $2f_2$, $(f_1 + f_2)$ and $(f_1 - f_2)$. All these components can be removed from the amplifier output with the help of band pass filters.

The third term of Vout equation is the third order term and can be expanded as

$$
\begin{aligned}
V_{3out} &= C\,[V_1 \cos \omega_1 t + V_2 \cos \omega_2 t]^3 \\
&= C\left[V_1^3 \cos^3 \omega_1 t + V_2^3 \cos^3 \omega_2 t + 2\left(V_1^2 \cos^2 \omega_1 t\right)(V_2 \cos \omega_2 t)\right. \\
&\quad \left. + 2\left(V_2^2 \cos^2 \omega_2 t\right)(V_1 \cos \omega_1 t)\right] \\
&= C\left[V_1^3 \cos^3 \omega_1 t + V_2^3 \cos^3 \omega_2 t + V_1^2 V_2 (1 + \cos 2\omega_1 t)(\cos \omega_2 t)\right. \\
&\quad \left. + V_2^2 V_1 (1 + \cos 2\omega_2 t)(\cos \omega_1 t)\right] \\
&= C\begin{bmatrix} V_1^3 \cos^3 \omega_1 t + V_2^3 \cos^3 \omega_2 t + V_1^2 V_2 \cos \omega_2 t + V_1^2 V_2/2\,\{\cos(2\omega_1 - \omega_2)\,t \\ + \cos(2\omega_1 + \omega_2)\,t\} \\ + V_2^2 V_1 \cos \omega_1 t + V_2^2 V_1/2\,\{\cos(2\omega_2 - \omega_1)\,t + \cos(2\omega_2 + \omega_1)\,t\} \end{bmatrix}
\end{aligned}
$$

The first two terms contain frequency components f_1, f_2, $3f_1$ and $3f_2$. The triple frequency component can be removed from the amplifier output with the help of band pass filters.

The fifth and the last terms contain the frequency components $(2f_1 + f_2)$ and $(2f_2 + f_1)$ which can again be removed by the band pass filters.

The frequency components $(2f_1 - f_2)$ and $(2f_2 - f_1)$ in the fourth and seventh terms can fall within the bandwidth of the transponder and are referred to as **third-order intermodulation products of the amplifier**. Therefore, the intermodulation products of concern are:

$$V_{3IM} = C\left[V_1^2 V_2 \cos(2\omega_1 - \omega_2)t + V_2^2 V_1 \cos(2\omega_2 - \omega_1)t\right]$$

The power of the intermodulation components is given by:

$$P_{IM} = \frac{1}{2}C^2 V_1^4 V_2^2 + \frac{1}{2}C^2 V_2^4 V_1^2 = 4C^2(P_1^2 P_2 + P_2^2 P_1)$$

It is clear that the ratio of intermodulation power to the desired power increase in proportion to the cubes of the signal power and also to square of (C/A).

The greater the non-linearlity of the amplifier, larger is the value of (C/A) and larger the value of intermodulation products. Also, the intermodulation terms increase rapidly as the amplifier operates near its saturation region.

The easiest way to reduce the intermodulation problems is to reduce the levels of input signals to the amplifier. The output backoff is defined as the difference in decibels between the saturated output power of the amplifier and its actual power.

When the transponder is operated with output backoff, the power level at the input is reduced by input backoff. As the characteristics of the amplifier are non-linear, the value of input backoff is greater than the value of output backoff.

Intermodulation distortion is a serious problem when the transponder is made to handle two or more carrier signals. That is why satellite links that use frequency division multiple access technique are particularly prone to this type of interference.

On the other hand, a single carrier per transponder TDMA system is becoming increasingly popular as in this case the satellite TWTA can be operated at or close to the saturation level without any risk of generating intermodulation products. This maximizes the EIRP for the downlink.

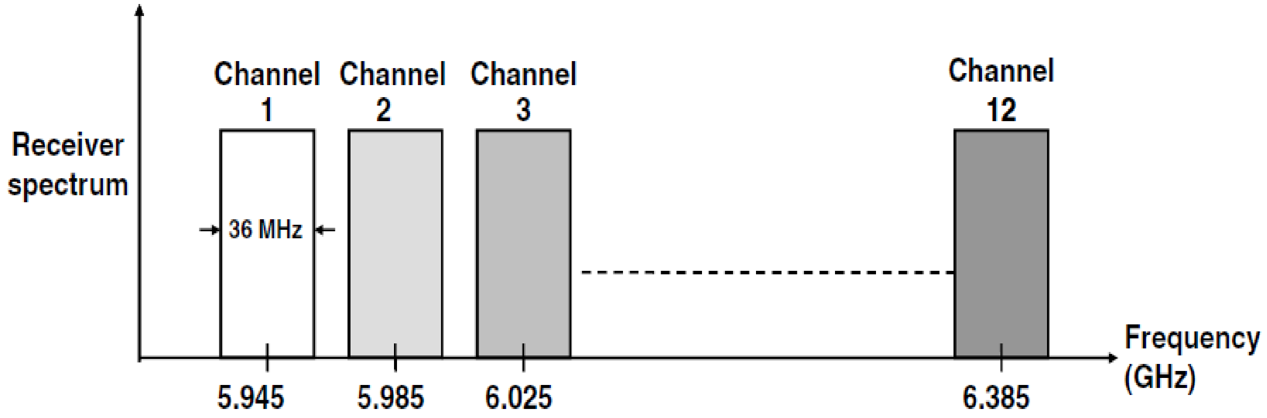
Another intermodulation interference-related problem associated with the FDMA system is that the Earth station needs to exercise a greater control over the transmitted power in order to minimize the overdrive of the satellite transponder and the consequent increase in intermodulation interference. Intermodulation considerations also apply to Earth stations transmitting multiple carriers, which forces the amplifiers at the Earth station to remain underutilized.

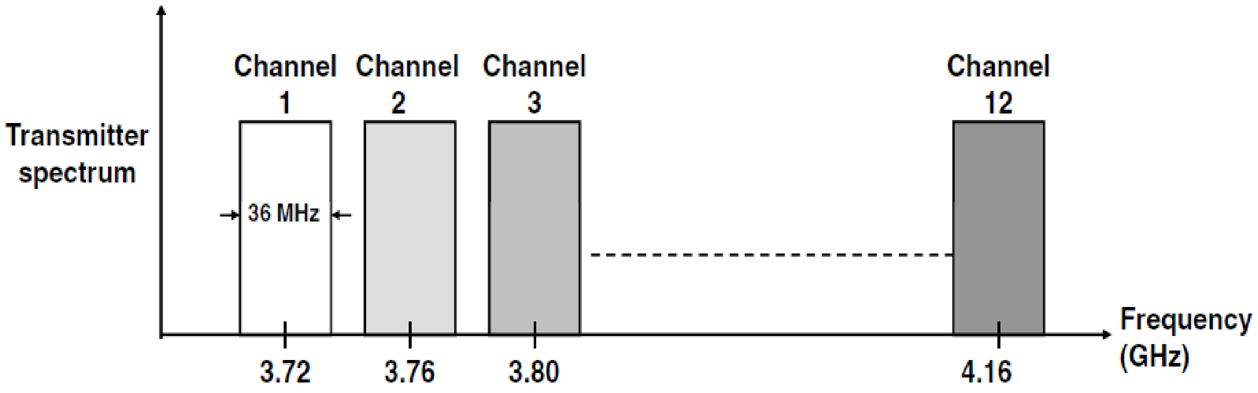
Frequency Division Multiple Access (FDMA)

➔ one of the most commonly employed forms of multiple access techniques for communications via satellite.

➔ In the case of frequency division multiple access (FDMA), different Earth stations are able to access the total available bandwidth of satellite transponder by virtue of their different carrier frequencies, thus avoiding interference among multiple signals.

➔ The figure below shows the typical arrangement for carrier frequencies for a C band transponder for both uplink and downlink. The transponder receives transmissions at around 6 GHz and retransmits them at around 4 GHz, with 12 transponders, with each transponder having a bandwidth of 36MHz and a guard band of 4MHz between adjacent transponders to avoid interference.

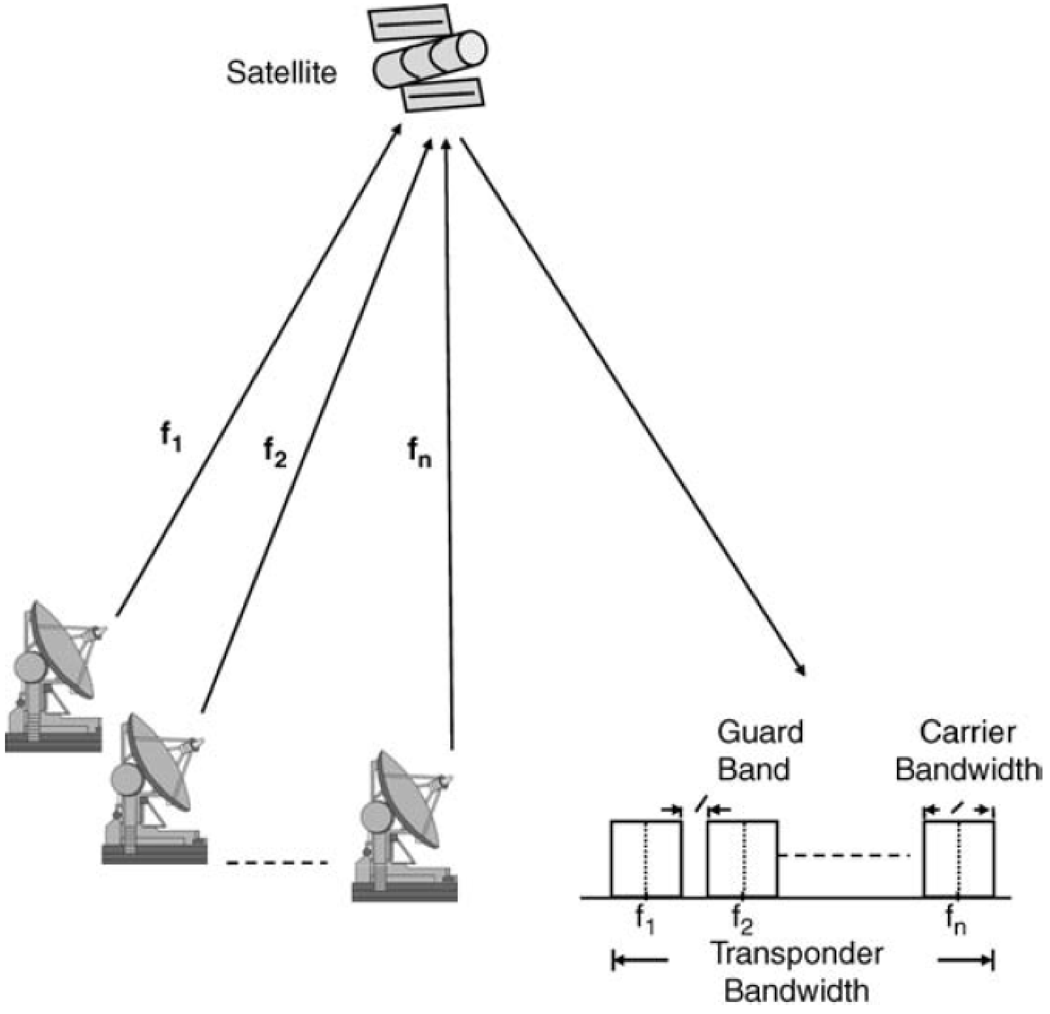




Each of the Earth stations within the satellite's footprint transmits one or more message signals at different carrier frequencies. Each carrier is assigned a small guard band, as mentioned above, to avoid overlapping of adjacent carriers. The satellite transponder receives all carrier frequencies within its bandwidth, does the necessary frequency translation and amplification and then retransmits them back towards Earth.

Two FDMA techniques are in operation today. One of them is the multichannel per carrier (MCPC) technique, where the Earth station frequency multiplexes several channels into one carrier base band assembly, which then frequency modulates an RF carrier and transmits it to an FDMA satellite transponder.

In the other technique, called the single channel per carrier (SCPC), each signal channel modulates a separate RF carrier, which is then transmitted to the FDMA transponder. The modulation technique used here could either be frequency modulation (FM) in case of analogue transmission or phase shift keying (PSK) for digital transmission.



- Major advantages of FDMA include simplicity of Earth station equipment and the fact that no complex timing and synchronizing techniques are required.

- Disadvantages include the likelihood of intermodulation problems with its adverse effect on the signal-to-noise ratio. The intermodulation products result mainly from the non-linear characteristics of the (TWTA) of the transponder, which is required to amplify a large number of carrier frequencies.

- The problem is further compounded when the TWTA is made to operate near saturation so as to be able to supply certain minimum carrier power in order to reduce downlink noise and by the fact that the TWTA when operated near saturation exhibits higher non-linearity.

➔ FDMA system can either be **power-limited** or **bandwidth limited** in terms of the number of carriers that can access the satellite transponder. The maximum number of carriers that can access the transponder is given by ($n = B_{TR}/B_C$), where B_{TR} is the total transponder bandwidth and B_c is the carrier bandwidth.

➔ If the EIRP is sufficient to meet the (C/N) requirements, then the system can support (n) carriers and is said to be bandwidth limited. In case, the EIRP is insufficient to meet the (C/N) requirements, the number of carriers that can access the satellite is less than (n). The system in this case is power-limited.

The overall noise-to-carrier ratio for a satellite link is given by:

$$\left[\frac{N}{C}\right]_{OV} = \left[\frac{N}{C}\right]_{U} + \left[\frac{N}{C}\right]_{D} + \left[\frac{N}{C}\right]_{IM}$$

Where,

$[N/C]_{OV}$ = Overall noise-to-carrier ratio
$[N/C]_{U}$ = Uplink noise-to-carrier ratio
$[N/C]_{D}$ = Downlink noise-to-carrier ratio
$[N/C]_{IM}$ = Inter-modulation noise-to-carrier ratio

The value of noise-to-carrier ratio given by the equation above should be less than the required design value of noise-to-carrier ratio [$(N/C)_{REQ}$], that is:

$$\left[\frac{N}{C}\right]_{OV} \leq \left[\frac{N}{C}\right]_{REQ}$$

$$\left[\frac{N}{C}\right]_{REQ} \geq \left[\frac{N}{C}\right]_{U} + \left[\frac{N}{C}\right]_{D} + \left[\frac{N}{C}\right]_{IM}$$

The overall C/N ratio taking into account the contribution of intermodulation products is given by

$$\left(\frac{C}{N}\right)_o = 1\Big/\left[\left\{1\Big/\left(\frac{C}{N}\right)_{up}\right\} + \left\{1\Big/\left(\frac{C}{N}\right)_{down}\right\} + \left\{1\Big/\left(\frac{C}{N}\right)_{IM}\right\}\right]$$

In an FDMA system, the up-link noise is usually negligible and the inter-modulation noise is brought to an acceptable level by employing back-off in power amplifiers.

It may be mentioned here that the operating point of the power amplifiers mostly TWTA is shifted closer to the linear portion of the curve in order to reduce the inter-modulation distortion.

The output back-off is the corresponding drop in the output power in dB. Therefore, the previous equation can be rewritten as:

$$\left[\frac{N}{C}\right]_{REQ} \geq \left[\frac{N}{C}\right]_D$$

$$\left[\frac{C}{N}\right]_{REQ} \leq \left[\frac{C}{N}\right]_D$$

The downlink carrier-to-noise ($[C/N]_D$) is expressed by

$$\left[\frac{C}{N}\right]_D = [EIRP]_D + \left[\frac{G}{T}\right]_D - [LOSSES]_D - [k] - [B]$$

Where,

$[EIRP]_D =$ Satellite Equivalent Isotropic Radiated Power
$[G/T]_D =$ Earth-station receiver G/T
$[LOSSES]_D =$ Free space and other losses at the downlink frequency
$[k] =$ Boltzmann's constant in dB
$[B] =$ Signal bandwidth and is equal to the noise bandwidth

Therefore,

$$\left[\frac{C}{N}\right]_{REQ} \leq [EIRP]_D + \left[\frac{G}{T}\right]_D - [LOSSES]_D - [k] - [B]$$

In the case of single carrier per channel (SCPC) systems, the satellite will have saturation value of EIRP ($[EIRP]_{sat}$) and transponder bandwidth of B_{TR}, both of which are fixed. In this case, there is no back-off and equality sign applies. Therefore:

$$\left[\frac{C}{N}\right]_{REQ} = [EIRP]_{SAT} + \left[\frac{G}{T}\right]_{D} - [LOSSES]_{D} - [k] - [B_{TR}]$$

$$\left[\frac{C}{N}\right]_{REQ} - [EIRP]_{SAT} - \left[\frac{G}{T}\right]_{D} + [LOSSES]_{D} + [k] + [B_{TR}] = 0$$

Let us now consider the case of multiple carriers per channel (MCPC) systems having N carriers with each carrier sharing the output power equally and having a bandwidth of B. The output back-off is given by $[BO]_o$. The output power for each of the FDMA carriers is given by:

$$[EIRP]_{D} = [EIRP]_{SAT} - [BO]_{O} - [N]$$

The transponder bandwidth (B_{TR}) is shared by all the carriers but due to power limitation imposed by the need of back-off, the whole bandwidth is not utilized. Let us assume that the fraction of the bandwidth utilized is alpha (α). Therefore:

$$NB = \alpha B_{TR}$$

Expressing in terms of decilogs

$$[B] = [\alpha] + [B_{TR}] - [N]$$

Substituting the values of $[EIRP]_D$ and $[B]$ we get:

$$\left[\frac{C}{N}\right]_{REQ} \leq [EIRP]_{SAT} - [BO]_O + \left[\frac{G}{T}\right]_D - [LOSSES]_D - [k] - [\alpha] - [B_{TR}]$$

can be rearranged as:

$$\left[\frac{C}{N}\right]_{REQ} - [EIRP]_{SAT} - \left[\frac{G}{T}\right]_D + [LOSSES]_D + [k] + [B_{TR}] \leq -[BO]_O - [\alpha]$$

in the case of SCPC system LHS of the previous equation is zero. For MCPC systems, it is less than zero. Therefore:

$$0 \le -[BO]_O - [\alpha] \;\; or \; [\alpha] \le -[BO]_O .$$

The best that can be achieved in a MCPC system is to make $[\alpha] \le -[BO]_O$.

Single Channel Per Carrier (SCPC) Systems (SCPC/FM/FDMA system and SCPC/PSK/FDMA system)

SCPC/FM/FDMA System

The transponder bandwidth is subdivided in such a way that each base band signal channel is allocated a separate transponder subdivision and an individual carrier

Though it suffers from the problem of power limitation resulting from the use of multiple carriers and the associated intermodulation problems, it does enable a larger number of Earth stations to access and share the capacity of the transponder using smaller and more economic units as compared to multiple channels per carrier systems

This type of SCPC system also has the advantage that the power of the individual transmitted carriers can be adjusted to the optimum value for given link conditions.

However, this type of SCPC system requires automatic frequency control to maintain spectrum centering for individual channels, which is usually achieved by transmitting a pilot tone in the centre of the transponder bandwidth.

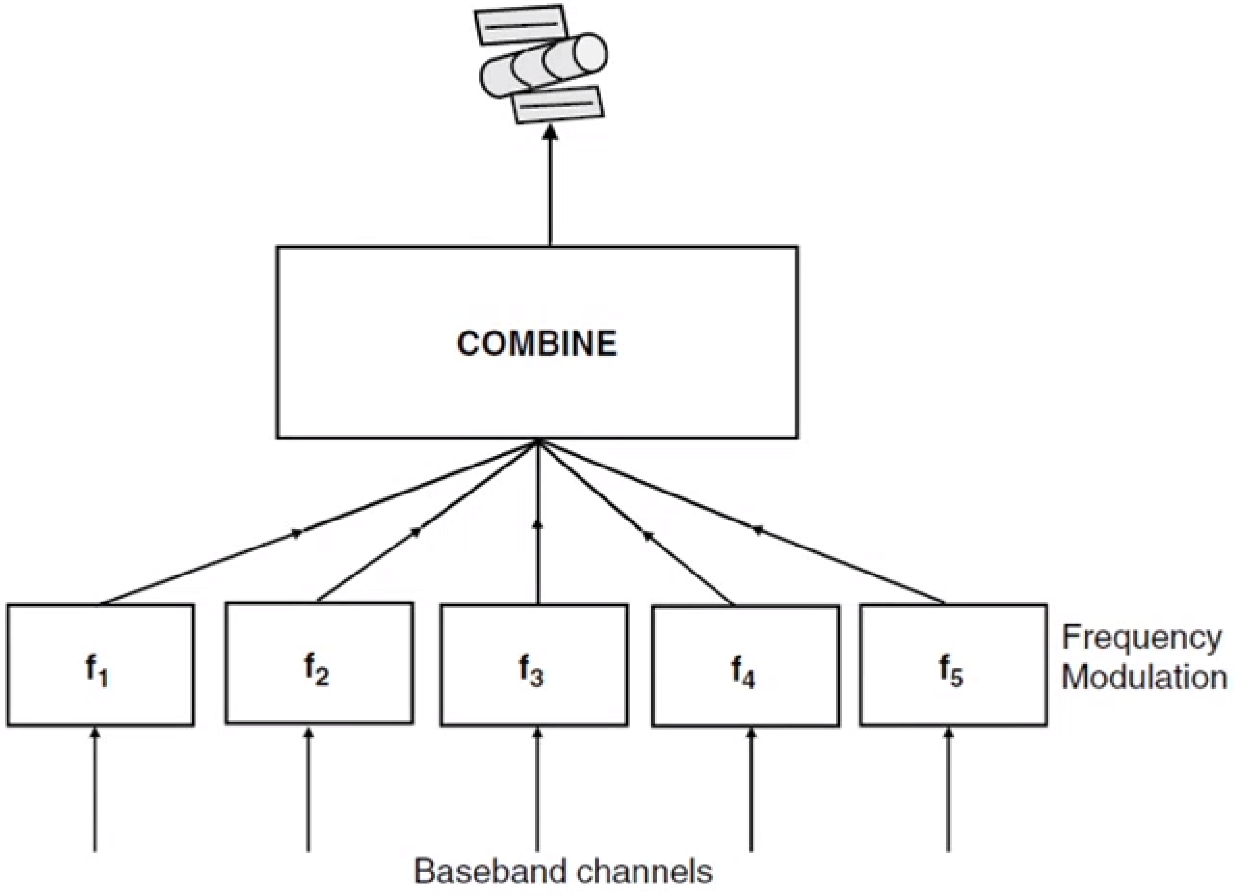
Different base band signals frequency-modulate their respective allocated carriers, which are combined and then transmitted to the satellite over the uplink.

The signal-to-noise power ratio (S/N) at the output of the demodulator for the SCPC/FM/ FDMA system can be compu from:

$$\frac{S}{N} = \left(\frac{C}{N}\right) \times 3B \times \left(\frac{f_d{}^2}{f_2{}^2 - f_1{}^2}\right)$$

where

C = carrier power at the receiver input (in W)
N = noise power (in W) in bandwidth B (in Hz)
B = RF bandwidth (in Hz)
f_d = test tone frequency deviation (in Hz)
f_2 = upper base band frequency (in Hz)
f_1 = lower base band frequency (in Hz)



SCPC/PSK/FDMA System

This is the digital form of the SCPC system in which the modulation technique used is phase shift keying (PSK).

SPADE (single channel per carrier PCM multiple access demand assignment equipment) was the first operational SCPC/PSK/FDMA system. It was designed for use on Intelsat-4 and subsequent Intelsat satellites. This system employs PCM for base band signal encoding and QPSK as the carrier modulation technique.

With this, it is possible to accommodate a 64 kbps voice channel in a bandwidth of 38.4 kHz as compared to the requirement of a full 45 kHz in the case of frequency modulation.

The channel capacity can be determined from the carrier-to-noise density ratio. The carrierto-noise ratio necessary to support each carrier can be computed from:

$$\left[\frac{C}{N}\right]_{\mathrm{th}} = \left[\frac{E_{\mathrm{b}}}{N_0}\right]_{\mathrm{th}} - [B] + [R] + [M]$$

where

$[C/N]_{\mathrm{th}}$ = carrier-to-noise ratio (in dB) at the threshold error rate
$[E_{\mathrm{b}}/N_0]_{\mathrm{th}}$ = bit energy-to-noise density ratio (in dB) at the threshold error rate
$[B]$ = noise bandwidth (in Hz)
$[R]$ = data rate (in bps)
$[M]$ = system margin to allow for impairments (in dB)

The carrier-to-noise density ratio $[C/N_0]$ can be computed from

$$\left[\frac{C}{N_0}\right] = \left[\frac{C}{N}\right]_{\mathrm{th}} + 10 \log B$$

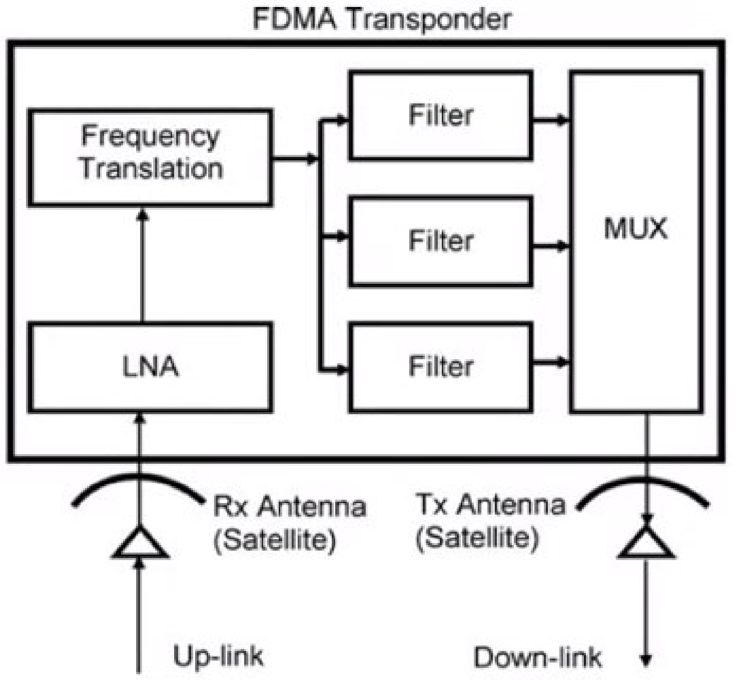
Multiple Channels Per Carrier (MCPC) Systems

MCPC/FDM/FM/FDMA system and MCPC/PCM-TDM/PSK/FDMA system

MCPC/FDM/FM/FDMA System

The FDMA transponder receives multiple carriers, carries out frequency translation and then separates out individual carriers with the help of appropriate filters. Multiple carriers are then multiplexed and transmitted back to Earth over the downlink. The receiving station extracts the channels assigned to that station.

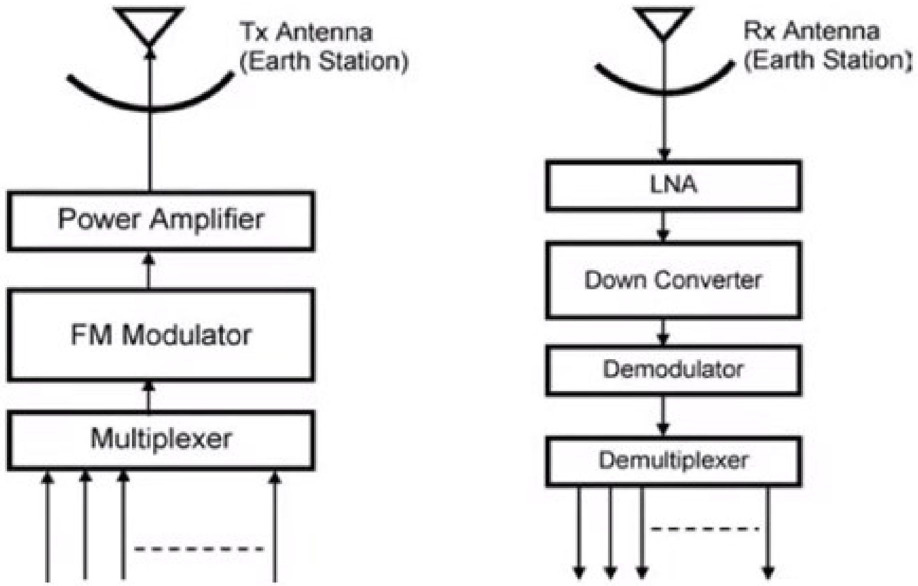
The channel capacity falls with an increase in the number of carriers. Larger number of carriers causes more intermodulation products, with the result that intermodulation-prone frequency ranges cannot be used for traffic. The signal-to-noise ratio at the demodulator output for such a system is given by:



$$\frac{S_b}{N_b} = \left(\frac{f_d}{f_m}\right)^2 \times \left(\frac{B}{b}\right) \times \left(\frac{C}{N}\right)$$

where

f_d = RMS (root mean square) test tone deviation (in Hz)
f_m = highest modulation frequency (in Hz)
B = bandwidth of the modulated signal (in Hz)
b = base band signal bandwidth (in Hz)
C = carrier power at the receiver input (in W)
N = noise power (= kTB) in bandwidth B (in W)



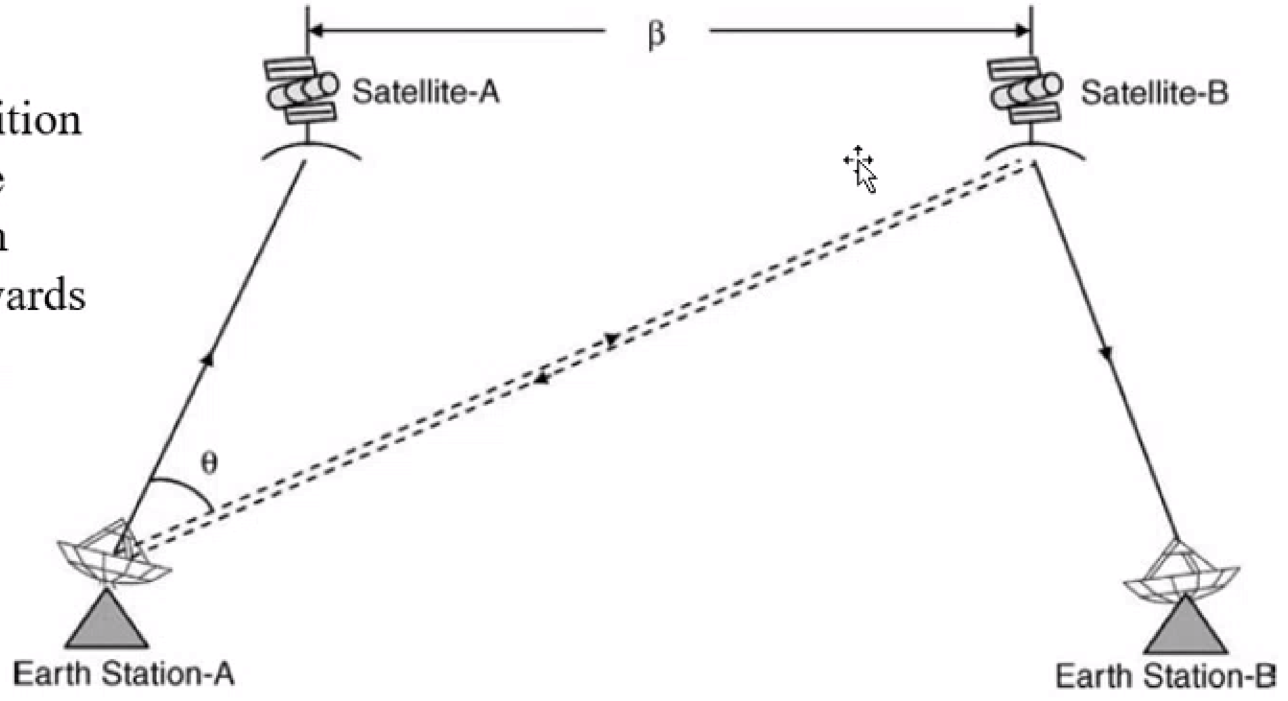
Interference due to Adjacent Satellites

This type of interference is caused by the presence of side lobes in addition to the desired main lobe in the radiation pattern of the Earth station antenna.

If the angular separation between two adjacent satellite systems is not too large, it is quite possible that the power radiated through the side lobes of the antenna's radiation pattern, whose main lobe is directed towards the intended satellite, interferes with the received signal of the adjacent satellite system.

Similarly, transmission from an adjacent satellite can interfere with the reception of an Earth station through the side lobes of its receiving antenna's radiation pattern.

Satellite A and satellite B are two adjacent satellites. The transmitting Earth station of satellite A on its uplink, in addition to directing its radiated power towards the intended satellite through the main lobe of its transmitting antenna's radiation pattern, also sends some power, though unintentionally, towards satellite B through the side lobe. The desired and undesired paths are shown by solid and dotted lines respectively



In the figure, ϑ is the angular separation between two satellites as viewed by the Earth stations and β is the angular separation between the satellites as viewed from the centre of the Earth; i.e. β is simply the difference in longitudinal positions of the two satellites.

transmission from satellite B on its downlink, in addition to being received by its intended Earth station shown by a solid line again, also finds its way to the receiving antenna of the undesired Earth station through the side lobe shown by the dotted line. Quite obviously, this would happen if the off-axis angle of the radiation pattern of the Earth station antenna is equal to or more than the angular separation ϑ between the adjacent satellites. ϑ and β are interrelated by the following expression:

$$\theta = \cos^{-1}\left[\frac{{d_A}^2 + {d_B}^2 - 2r^2(1 - \cos\beta)}{2d_A d_B}\right]$$

where

d_A = slant range of satellite A
d_B = slant range of satellite B
r = geostationary orbit radius

For a known value of ϑ, the worst case acceptable value of the off-axis angle of the antenna's radiation pattern can be computed. Similarly, for a given radiation pattern and known off-axis angle, it is possible to find the minimum required angular separation between the two adjacent satellites for them to coexist without causing interference to each other.

Let us take the case of downlink interference and determine the expression for carrier-to-interference (C/I) ratio. The desired carrier power C_D for the downlink channel in dBW can be expressed as:

$$C_{\mathrm{D}} = \mathrm{EIRP} - L_{\mathrm{D}} + G$$

EIRP = desired EIRP (in dBW, or decibels relative to a power level of 1 W)
L_{D} = downlink path loss for the beam from the desired satellite (in dB)
G = Earth station antenna gain in the direction of the desired satellite (in dB)

The interfering carrier power for the downlink channel (I_D) in dBW is given by:

$$I_{\mathrm{D}} = \mathrm{EIRP}' - L_{\mathrm{D}'} + G'$$

where

EIRP′ = interfering EIRP (in dBW)
$L_{\mathrm{D}'}$ = downlink path loss for the beam from interfering satellite (in dB)
G' = Earth station antenna gain in the direction of the interfering satellite (in dB)

The expression for (C/I) in the case of downlink can then be written as

$$\begin{aligned}(\mathrm{C/I})_{\mathrm{D}} &= (\mathrm{EIRP} - L_{\mathrm{D}} + G) - (\mathrm{EIRP}' - L_{\mathrm{D}'} + G') \\ &= (\mathrm{EIRP} - \mathrm{EIRP}') - (L_{\mathrm{D}} - L_{\mathrm{D}'} + (G - G')\end{aligned}$$

where, $(\mathrm{C/I})_{\mathrm{D}}$ is the C/I for the downlink channel in dB.

If the path losses are considered as identical, then

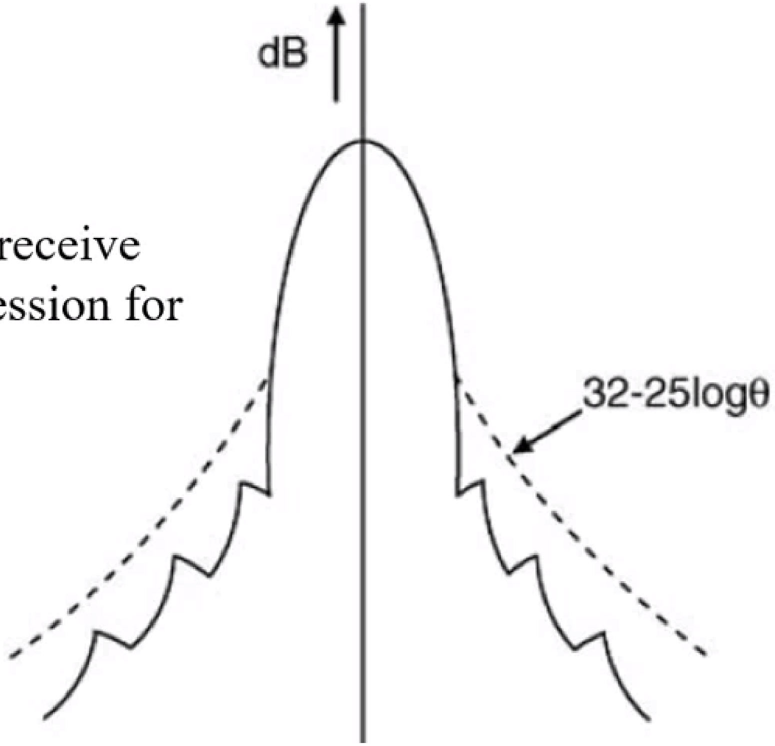
$$(\mathrm{C/I})_{\mathrm{D}} = (\mathrm{EIRP} - \mathrm{EIRP}') + (G - G')$$

Also, the term $(G - G')$ is the receive Earth station antenna discrimination, which is defined as the antenna gain in the direction of the desired satellite minus the antenna gain in the direction of the interfering satellite.

According to CCIR standards, in cases where the ratio of the antenna diameter to the operating wavelength is greater than 100, G' as a function of the off-axis angle ϑ should at the most be equal to $(32 - 25 \log \vartheta)$ dB (forward gain of an antenna compared to an idealized isotropic antenna), where ϑ is in degrees. The requirement of FCC standards for the same is $(29 - 25 \log \vartheta)$ dB.

the typical Earth station antenna pattern, which is a plot of the gain versus the off-axis angle along with CCIR requirements. This gives:

$$(C/I)_D = (EIRP - EIRP') + (G - 32 + 25 \log \theta)$$



A similar calculation can be made for the uplink interference, where a satellite may receive an unwanted signal from an interfering Earth station. In the case of uplink, the expression for C/I can be written as:

$$(C/I)_U = (EIRP - EIRP') + (G - G')$$

where

$(C/I)_U$ = C/I for the uplink channel in dB
EIRP = EIRP of the desired Earth station in dBW
EIRP′ = EIRP of the interfering Earth station in the direction of the satellite in dBW
G = gain of the satellite receiving antenna in the direction of the desired Earth station in dB
G' = gain of the satellite receiving antenna in the direction of the interfering Earth station in dB

EIRP′ is further equal to

$$\text{EIRP}' = \text{EIRP}^* - G_I + (32 - 25\log\theta)$$

where

EIRP* = EIRP of the interfering Earth station in dBW
G_I = on-axis transmit antenna gain of the interfering Earth station in dB
θ = viewing angle of the satellite from the desired and interfering Earth Stations

The overall carrier-to-interference ratio (C/I) for adjacent satellite interference is given by

$$\frac{C}{I} = \left[\left(\frac{C}{I}\right)_U^{-1} + \left(\frac{C}{I}\right)_D^{-1}\right]^{-1}$$

where the subscripts U and D imply uplink and downlink respectively. Where the interference is noise like, it is possible to combine the effects of noise and interference. The combined carrier-to-noise ratio (C/NI) is given by

$$(C/NI) = [(C/N)^{-1} + (C/I)^{-1}]^{-1}$$

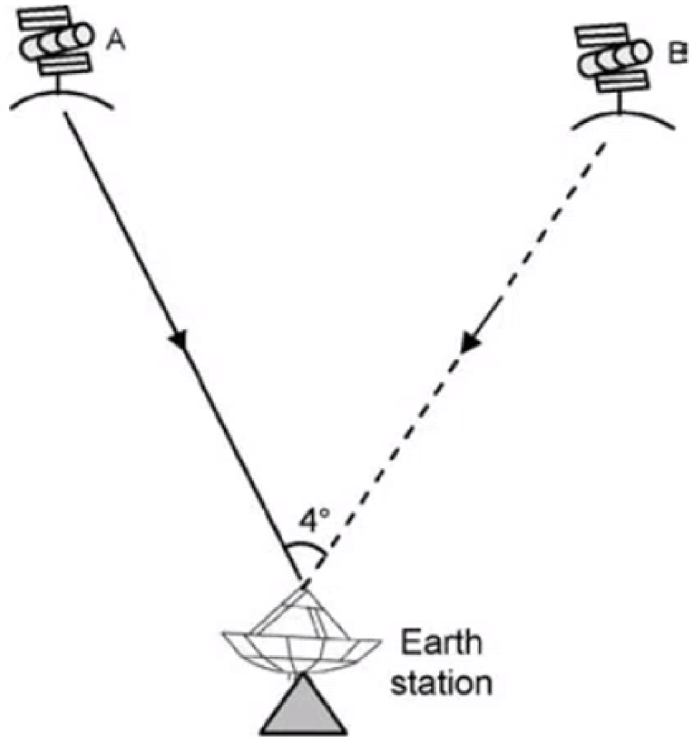
It may be mentioned here that the various terms in equations are not in decibels.

Adjacent Channel Interference

Adjacent channel interference occurs when the transponder bandwidth is simultaneously shared by multiple carriers having closely spaced centre frequencies within the transponder bandwidth. When the satellite transmits to Earth stations lying within its footprint, different carriers are filtered by the receiver so that each Earth station receives its intended signal.

Filtering would have been easier to realize had there been a large guard band between adjacent channels, which is not practically feasible as that would lead to the inefficient use of the transponder bandwidth.

The net result is that a part of the power of the carrier in the channel adjacent to the desired one is also captured by the receiver due to overlapping amplitude characteristics of the channel filters. This becomes a source of noise.



References

- SATELLITE TECHNOLOGY PRINCIPLES AND APPLICATIONS, Second Edition, Anil K. Maini, Varsha Agrawal. Wiley
- Satellite Communication Engineering, Michale O. Kolawole, signal processing and communication series. Copyright © 2002 by Marcel Dekker, Inc.
- Satellite Communications Systems : Systems, techniques and Technologies. 5th edition, Wiley , Gerard Maral
- Satellite Communications, Dennis Roddy, 3rd Edition, McGraw-Hill