NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Electrical and Electronic Engineering

DIGITAL SATELLITE COMMUNICATION USING TRELLIS CODE

Graduation Project
EE- 400

Student: Amer. A. Abu-Keer (960630)

Supervisor: Prof. Dr. Fakhreddin Mamadov

Lefkoşa - 2000

CONTENTS

ACKNOWLEDGMENT

This project present highly technical subject matter of the technique and the utilization of digital satellite in the communication area showing the details of the structure of the digital satellite, the orbits utilized by the Satellites, the Satellite Access, the Use of Trellis code in the satellite.

This project was not possible to be prepared without the guidance and the support of my supervisor Prof. Dr. Fakherddin Mamedov, the chairman of Electrical and Electronic Engineering Department.

I am indebted to him for his complete support and showing me the guidance throughout all the stages of the preparation, and providing his constructive comments.

So I would like to take this opportunity to thank Prof. Dr. Fakerddin for every single help and support, not just throughout this project but also through the courses which he provides to the students in The Department of Electrical and Electronic Engineering, because through these courses I have gained a lot of knowledge which helped me in the preparation of this project.

ABSTRACT

The future trend in satellite communication is toward DIGITAL techniques. Frequency division multiplexing-frequency modulation –frequency division multi[le access (FDM-FM-FDMA) has been the most popular analog technique used in commercial satellite systems because it has been field-proven and makes it easy to provide quality satellite links at a low cost. As the number of earth stations increases, the transponder capacity decreases markedly in a FDM-FM-FDMA system. In additi0on, FDM-FM-FDMA is inflexible in responding to traffic changes. On the other hand, a digital satellite system such as quarter-nary phase shift keying time division multiple access (QPSK-TDMA) can accommodate a large number of earth stations with only a small loss in transponder capa-city. Furthermore, it can quickly respond to traffic variations. Also associated with a digital satellite commu-nication are techniques such as demand assignment and digital speech int-erpolation to further increase the efficiency. With advanced satellite systems with onboard switching and processing, multiple spot beam, and beam hopping, a digital satellite can serve a mixture of large, medium, and small earth stations with high efficiency. Unlike an analog satellite system, a digital satellite system can employ error-correction coding to trade bandwidth for power. Finally, the use of code-division multiple access (CDMA) for low data rate applic-ation enables users to employ micro earth stations (0.5-m antenna) at an extremely low cost ($3000) to obtain premium quality services. The flexibility of digital satellite systems will make them even mote promising when integrated digital networks become fully impleme-nted.

CHAPTER 1

OVERVIEW OF DIGITAL SATELLITE COMMUNICATION SYSTEM

Overview

Satellite communication has evolved into an every day, commonplace thing. Most television coverage travels by satellite, even reaching directly to home from space. No longer is it a novelty to see a telecast has been carried by satellite. The bulk of transoceanic telephone and data communication also travels by satellite.

A unique benefit has appeared in the area of emergency preparedness and response.

The term " **SATELLITE** " means the actual communication spacecraft in orbit, which relays radio signals between earth stations on the ground.

1.1 Basic Characteristics Of Satellites

A communication satellite permits tow or more points on the ground (earth stations) to send messages to one another over great distance using radio waves.

A satellite in the geostationary earth orbit (GEO) revolves around the earth in the plane of the equator once in 24 hours, maintaining precise synchronization with the earth's rotation, It is well known that a system of three satellites in GEO, each separated by 120 degrees of longitude can receive and send radio signals over the entire globe except for the polar regions.

A given satellite has a coverage region within which earth stations can communicate with and be linked by the satellite.

The GEO (Also referred to as the geostationary satellite orbit (GSO)) is the ideal case of the entire class of geosynchronous (or synchronous) orbits, which all have a 24 hour period of revolution but are typically inclined with respect to the equator, As viewed from the earth, a synchronous satellite in an inclined orbit will appear to drift during a day about its normal position in the sky, The GEO is not a stable arrangement and inclination increases in time. Inclination is controlled during the entire lifetime of the satellite. A synchronous satellite not intended for GEO operation can be launched with considerably less auxiliary fuel for this purpose. Orbit inclination of greater than 0.1 degrees is usually not acceptable for commercial service unless the earth station antennas can automatically repoint toward (track) the satellite as it appears to move.

The key dimension of a geostationary satellite is its ability to provide coverage of an entire hemisphere at one time. A large contiguous land area as well as offshore locations can simultaneously access a single satellite. If the satellite has a specially designed communication beam focused on these areas, then any receiving antennas within the " footprint" of the beam (the area of coverage) will receive precisely the same transmission. Location well outside the footprint will generally not be able to use the satellite effectively.

The term " bypass " is often used to refer to the ability of satellite links to step over the existing terrestrial network and thus avoid the installation problems and service delays associated with local telephone service. Using the satellite in a duplex mode the user can employ earth station at each end, eliminating any connection with the terrestrial network. In a terrestrial microwave system, radio repeaters must be positioned at intermediate points along the route to maintain line of sight contact. This is because microwave energy, including that on terrestrial and satellite radio links, travels in a straight line with a minimum of bending over or around obstacles.

1.1 Satellite Communication System

A communication satellite can be considered to comprise tow main modules:

- The spacecraft bus (or "space platform" or "service module")
- The communication subsystem (or "communication payload" or "module")

The satellite may also include an apogee motor as an integral part if it is to be put into orbit by a multi-stage launcher.

1.2.1 Spacecraft Bus

The spacecraft bus provides all of the support services, such as structural support, power supply, thermal control, and communications module to function port services are required to enable the communication module to function effectively. The subsystem of the bus include the following:

1. ***Structural subsystem***-This comprises a mechanical skeleton on which the equipment The equipment modules are mounted. It also includes akin or shield which protects the sat-ellite from the effects of micro-meteorites and from the extremes. Most communication satellite are either box-shaped or cylindrical. Box-shaped satellte s are stabilized by means of intertia wheels spinning within the body. They are said to be body-stabilized. Cylindrical satellites are spin-stabilized. They are stabilized by spinning the whole body.

2. ***Telemetry, tracking and command*** *(TT&C) subsystem*- this is a system for monitoring the state of the on-board equipment. The telemetry system multiplexes data from many sensors on the spacecraft and transmits them via a digital communication link to the controlling earth station. This station incorpor-ates a satellite tracking system to monitor changes in the satellite orbit. Control functions on the satellite might include firing the apogee motor or using thruster jets to control of certain on-board communication sub-systems, such as operat-ing switching matrices to direct communications signals to specified antennas.

2. ***Power subsystem***- solar arrays provide the primary source of power for a communication satellite. Most of the power is used by the high power amplifiers in the communication module. Back-up batteries are sometimes also provided to cope with those periods when the satellite passes through the earth shadow.

4. ***Thermal control subsystem***-This is a combination of bimetallic louvres, electric heaters and surface finished designed to protect electronic equipment from operating at extreme temperatures.

5. **Attitude and orbit control subsystem**-This includes a system of small rocket motors (thrusters) required to keep the satellite in the correct orbital position and to ensure that its antenna are pointing in the right direction.

1.2.2 Communication Subsystem

The communication subsystem on a communication satellite consists of a number of repeaters which amplify the signals received from the uplink and condition them in preparation for transmission back. Communication subsystem consist of the following types:

1. Transponder repeater (or "non-regenerative" or "bentpipe" repeater).
2. On-board processing repeater (or "regenerative" or "switching regenerative" repeater).

Transponder Repeaters

The communication subsystem in civilian communication satellite are currently of the transparent type. In this type of communication subsystem, the repeaters are referred to as transponders.

Uplink communication signals from an earth station received at a satellite usually consist of multiple frequency division multiplexed (FDM) signals known as carriers. The signals are received at the satellite by a receiver antenna the output of which is connected to the transponders. Each transponder performs the processes of signal amplification, select-ion of one or more received signals (using a bandpass filter), translation of the signals to a new frequency band and amplification of them to a high power level for retransmission.

The transponder consists of the following:

- *A receiver*-which includes a low noise amplifier and a down convertor.
- *An input multiplexer*- in which a bandpass filter selects the channel frequency components assigned to that transponder.
- *A high power amplifier (HPA)*- which consists of a solid state power amplifier (SSPA) or a travelling wave tube amplifier (TWTA).

- *An output multiplexer-* in which a number of transponder output signals are combined before being fed to the satellite's transmitting antennas.

Satellite of the "transparent repeater" type have the advantage that their transponders impose minimal constraints on the characteristics of the communication transmission signal. As it becomes more and more common that satellite system are used to transmit all-digital signals, then it is likely that many future satellites will be of the "on-board processing repeater" type.

On-board processing repeater types of communication subsystem perform signal processing functions, which include:

1. ***Signal regeneration***-coherent reception and regeneration of the digital signal from the received from the uplink signal.

2. ***Switch board in the sky functions*** –a term used to describe a system for circuit switching between electronically hopping spot beam antennas and optical or links between satellites.

3. **Concentrator function**-a number of digital signals at low bit rates received from several VSAT terminals may, after regeneration, be multiplexed in a signal time division multiplex (TDM) signal for transmission at high bit rate to a larger central earth station.

1.2 System Element

1.2.1 Space Segment

Placing satellite into orbit and operating it for ten years involves a great deal Placement in orbit is accomplished by contracting both with a spacecraft manufacturer and with a launch agency, and allowing them the 30 to 40 month period necessary to design, construct, and launch the satellite.

After the satellite is properly positioned at its longitude above the equator, it becomes the responsibility of a satellite operator to control the satellite for the duration of its mission (its lifetime in orbit).

The tracking, telemetry, and command (TT&C) station establishes a control and

Monitoring link with the satellite. Precise tracking data is periodically collected via the tracking antenna to allow the pinpointing of the satellite's position and the planning of on-orbit position correction. This is because the orbit tend to shift with respect to the ground due to irregular gravitational forces from the nonspherical earth, from the sun, and from the moon. The second facility is the satellite control center (SCC) which houses the operator consoles and data processing equipment by which the control and monitoring of the satellite or satellites are accomplished. The SCC could be at the site of the TT&C station, but more commonly is located some distance away, usually at the headquarters of the satellite operator. The actual satellite related data can be passed between the sites over low speed data and voice lines (either terrestrial or satellite).

Routine operations at the SCC and TT&C station are intended to produce continuous and nearly uniform performance from the satellite. Actual communication services via the microwave repeater aboard the satellite do not need to pass through the satellite operator's ground facilities, although the satellite repeater for the purpose of testing and monitoring its performance. One particularly nice feature of a geostationary satellite is that the communication monitoring function can be performed from anywhere within the footprint, the SCC can have its own independent monitoring antenna not connected with the TT&C station. Having several monitoring antennas strategically positioned around the coverage region can be useful when measuring satellite repeater output and trouble shooting complaints and problems.

Another problem area for which monitoring is that of dealing with harmful interference to communication services. Also called "double illumination" it occurs when an errant station operator activates a transmitter on the wrong frequency or even on the wrong satellite.

1.3.2 Ground Segment

The space segment provides a communication repeater at essentially a fixed position in space capable of linking many points on the earth. It is the function of the ground segment to access the satellite repeater from these points in a manner which satisfies the communication needs of users within the structure of the satellite system.

"Earth station " is an internationally accepted term which includes satellite communication stations located on the ground, in the air (on airplanes), or on the sea (on ships). Most commercial applications are through earth stations at fixed locations on the ground, thus the international designation for this arrangement is the fixed satellite service (FSS).

Connection of the satellite network with the outside (terrestrial) world is accomplished through larger stations which access the public switched network (the national telephone system) or an international gateway (allowing communication with foreign countries).

The term very small aperture terminal (VSAT) is used to describe a compact and inexpensive earth station intended for this purpose. The aperture is the surface area of the antenna which radiates or collects the radio signals the satellite link. The ground segment, therefore, is not a single, homogeneous entity, but rather is a diverse collection of facilities, users, and applications. It is constantly changing and evolving, providing service when and where needed.

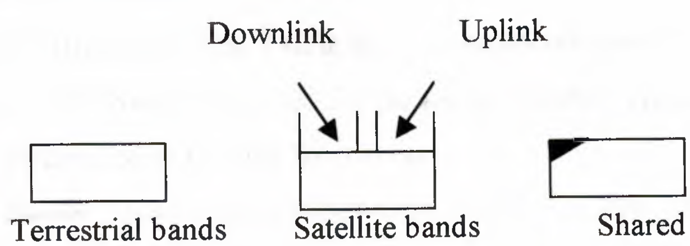
1.4 Frequency Bands

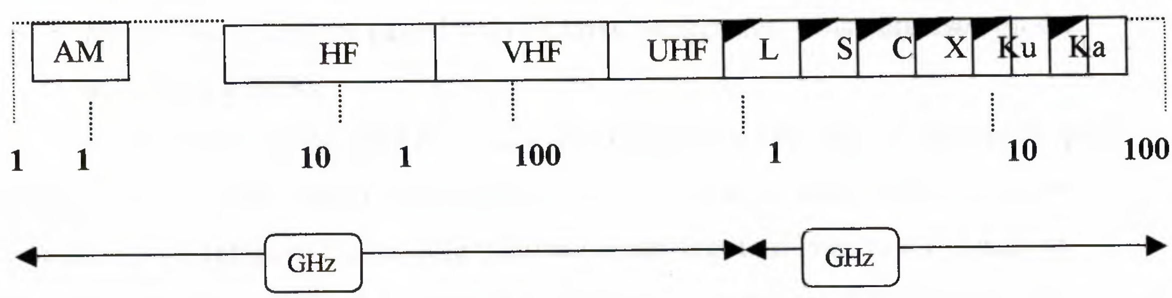
Satellite communication employ electromagnetic waves to carry information between ground and space. The frequency of the electromagnetic wave is the rate of reversal of its polarity in cycles per second. Alternating current in copper wire also has this frequency property, and if the frequency is sufficiently high, the wire will become an antenna, radiating electromagnetic energy at the same frequency.

A particular range of frequency is called a frequency band, while the full extent of all frequencies for zero to infinity is called the spectrum. In particular, the radio frequency (RF) part of the electromagnetic spectrum permits the efficient generation of signal power, its radiation into free space, and reception at a distant point. The most useful RF frequencies lie in the microwave bands (between approximately 300 MHz and 300.000 MHz) although lower frequencies (longer wavelengths) are attractive for certain application.

A RF signal on one frequency is called a carrier and the actual information that it carries (voice, video, or data) is called modulation. A carrier with modulation occupies a certain amount of RF band width within the frequency band of interest. If tow carriers are either on the same frequency or have overlapping bandwidth, then radio frequency interference (RFI) may occur. To the user, RFI can look or sound like background noise (which is neither intelligible nor particularly distressful), or it could produce an annoying effect like herringbone patterns on a TV monitor. When the interfering carrier would be classed as harmful. A condition similar to the jamming encountered in the short wave broadcast band.

The spectrum of RF frequency is depicted in the figure below, which indicates on a logarithmic scale the abbreviations that are in common usage. The bottom end of the spectrum from 0.1 to 100 MHz has been applied to the various radio broadcasting services and is not used for space communication. The frequency bands of interest for satellite lie above 100 MHz, where we find the VHF (very high frequency), UHF (ultra high frequency) and SHF (super high frequency) bands. The SHF range has been broken down further by common usage into sub-bands with letter designations, the familiar C and Ku bands being included. It is interesting to note that these latter designations are of historical interest, since they formerly were classified designations for the microwave bands used for radar and other military or government purposes.





An important consideration in the use of microwave frequency for satellite communication is the matter of *sharing*. The figure above indicates that most of the satellite bands (light shading) are "shared". Which means that the same frequencies are used by terrestrial microwave links. Parts of the Ku and Ka bands, on the other hand, are not shared with terrestrial so that only satellite links are permitted. In most instances, The most services must coexist by virtue of a process called frequency coordination where users who plan to use a given band for a given purpose work with current users to assure that harmful RFI will be avoided. A band which is not shared, therefore, is particularly valuable to satellite communication, since terrestrial microwave systems can be totally ignored. Frequency coordination is often necessary to control interference among satellite systems, which use the same frequency band and operate in adjacent orbit positions.

Of the frequency bands allocated for share use, two pairs of bands are of great importance for commercial fixed satellite service (FSS) operations,(services to fixed earth stations).they are :

1. ***4/6 GHz Bands*:** The bands 3700 - 4200 MHz and 5925 - 6425 MHz are referred to as the 4 and 6 GHz bands. The 4 GHz band is commonly used for downlink satellite services, the 6 GHz band being used for the paired uplinks. These bands are collectively referred to as C-Band frequencies.
2. ***11/14 GHz Bands***: These represent the bands 10,950 - 11,200 MHz, 11,450 - 11,700 MHz and 14,000 – 14,500 MHz, which are referred to as Ku-band frequency. The frequency in the nominal 11 GHz region are commonly used as satellite downlink frequencies. They are paired with 14 GHz frequencies, which are used on the associated uplinks.

A typical satellite band is divided into separate halves, one for ground to space links (the uplink) and one for space to ground links (the downlink). This separation is reflected in the design of the satellite microwave repeater to minimize the chance of downlink signals re-received and thereby jamming the operation of the satellite. By way of contrast, such a division is not provided for terrestrial system, but considerable care must be exercised in assigning frequencies, since links can run in any direction between micro-wave relay towers.

Uplink frequency bands are typically slightly above the corresponding downlink frequency band to take advantage of the fact that it is easier to generate RF power within an earth station than is on-board a satellite, where weight and power are limited. It is a natural characteristic of the types of RF power amplifier used in both locations that the efficiency of conversion from ac power into RF power tends to decrease as frequency is increased. Along with this, the output from the earth station power amplifier is usually greater than that of the satellite by a factor of from 10 to 100. Satellite systems of the future which make extensive use of VSATs will allow less uplink power, so that the cost of the earth station can be minimized.

-Frequencies for land mobile satellite service (MSS)were as follows:

1. GHz band: For satellite to mobile (downlink) transmission, the frequency band 1545–1559 MHz was allocated.
2. For mobile to satellite (uplink) transmission, the band 1646.5 – 1660.5 MHz was allocated.

- For direct broadcasting service (DBS) from satellite to homes:

12 GHz band: The broadcast frequency band of 11.7 – 12.5 Ghz was divided into channel each with a bandwidth of 27 MHz. Each nation was allocated four or five TV channels together with one-position geostationary satellite orbit positions. These are positions designated at spacing of 6° intervals in the geostationary orbit.

1.4.1 C-Band

The C-band was the first part of the microwave spectrum to be used extensively for commercial satellite communication. The C-band had at the outset a principal advantage over bands which are either higher or lower in frequency. C-band lies in a range of frequencies near 1 GHz where the combination of natural and manmade noise sources is a minimum. Hence, all other things being equal, C-band requires less signal level to provide good quality communication. Lower frequencies toward 100 Mhz suffer from a high level of man-made radio noise due to electrical equipment, automobile ignition system, and the like. Another disadvantage of lower frequencies is the meager bandwidth that is available.

The principal factor which affects the performance of satellite links at frequencies above 10 GHz is the absorption of the RF carrier power by the atmosphere. The most detrimental atmospheric effect is rain attenuation, which is a decrease of signal level due to absorption of microwave energy by water droplets in a rainstorm. Due to the relationship between the size of droplets relative to the wavelength of the radio signal, microwave energy at higher frequencies is more heavily absorbed than that at lower frequencies. Rain attenuation is particularly a problem in tropical regions of the world with heavy thunde-rstorm activity, as these storms contain intense rain cells.

Equipment technology and availability were factors in the favor of C-band. In the early years (1965 to 1970), C-band microwave hardware was obtainable from other applic-ations such as terrestrial microwave, tropospheric scatter communication systems (which use high power microwave beams to achieve over the horizon links) and radar. No break-through in contemporary technology was necessary to take advantage of the technical features of C-band.

C-band earth stations were located in remote places where terrestrial microwave signals on the same frequencies would be weak. The potential problem runs in both direc-tions, the terrestrial microwave transmitter can interfere with satellite reception at the earth station. And RF energy from an earth station uplink can leak towards a terrestrial microw-ave receiver and disturb its operation.

The technique by which sharing can be made to work. Is, A natural or man-made obstacle is located near the earth station antenna, but between it and the terrestrial microwave stations existing approximately within a 50-mile radius.

The amount of bending can be predicted and is a function of the distances between the source, obstacle, and receiver, as well as of the height differences. If the height Differ -ences are large, causing all antennas to lie well below the top of the obstacle, then Little signal will reach the receiver and good shielding is therefore achieved. (Shielding is Equal for both direction of propagation.

1.4.2 Ku Band

The frequency band that has done more to interest new users of satellite communication is Ku-band, a part of the spectrum lying just above 10 GHz. Portions of KU-band are not shared with terrestrial radio, which has some advantages over C-band, particularly for direct services using earth stations with small diameter antennas. The precise uplink and downlink frequency ranges allocated by the ITU vary to some degree with the region of the world. There are effectively three sections of KU-band, which have been allocated to different services on an international or domestic basis. The most prevalent is the fixed satellite service (FSS), which is the service intended for one or tow way communication between fixed points on the ground. All of C-band and the bulk of KU-band are allocated to the FSS for wide application in international and domestic communication.

Part of Ku-band is subject to the same coordination and siting difficulties as C-band . The particular part of Ku-band thusly affected is referred to as 14/11 GHz, where the uplink range is 14.00 to 14.500 GHz and the downlink range is 10.95 to 11.7 GHz (minus a gap of 0.25 GHz in the center). Only the downlink part of the allocation is actually subject to sharung.

A portion of Ku allocation for FSS which is not shared with terrestrial services is referred to as 14/12 GHz (uplink range is 14.00 to 14.50 GHz and downlink range is 11.70 to 12.20 GHz).

There is a third segment of Ku-band, referred to as 18/12 GHz, which is allocated strictly to the broadcasting satellite service (BSS). The BSS is not shared with terrestrial services. Its intended purpose is to allow television and other direct to home transmissions from the satellite. There are tow regulatory features of this band, which make direct broadcasting to small antennas feasible. The first is that, without sharing, the satellite power level can be set at the highest possible level. Adjacent satellite interference could be a problem in a common coverage area, but this is precluded by the second feature: BSS satellite are to be spaced a comfortable nine degrees apart. In comparison, while there is no mandated separation between FSS satellites, a two-degree spacing has become the standard in the crowded North American orbit arc.

The operation advantages of 14/12 and 18/12 GHz lie with the simplicity of locating earth station sites (without regard to terrestrial radio stations) and the higher satellite downlink power levels permitted . The latter results in smaller ground antenna diameters than at C-band , Ku-band is subject to higher rain attenuation which can increase the incidents and duration of loss of an acceptable signal.

The amount of margin to overcome a fade is also a strong function of the elevation angle from the earth station to the satellite in orbit. A rain cell exists as an atmospheric volume which is wider than it is high, therefore low elevation angles force the radio signal to pass through a greater thickness of rainfall. Elevation angles of forty degrees or greater are consequently preferred for KU-band frequencies and higher. Another important variable is the local climate, where desert regions are less affected that tropical. In general, the need for greater power margin at Ku-band tends to reduce some of the benefits obtainable by virtue of the higher powers that are permitted by the international regulations.

1.2.2 UHF and L Band

Even though the amount of available bandwidth below C-band is diminished, these frequency bands are effective for providing rapid communication by way of mobile and transportable earth stations. With lower frequency of operation, the receiving antenna can be as simple as small Yagi (TV type antenna) or wire helix, This is because the effective receiving area of the wire or rod antenna is inversely proportional to frequency. The use of relatively high power for each individual channel of communication also helps to reduce the size and cost of the receiving terminal. The tradeoff is in the number of voice channels per satellite: instead of being measured in the thousands for C-band and Ku-band satellites, capacity of each lower frequency satellite ranges from tens to hundreds of channels.

At UHF or L band, ten watts per voice channel provides satisfactory reception by the type of antenna found on a ship or aircraft, but only ten such channels can be supported by this satellite at one time. A C-band satellite can deliver perhaps 10.000 voice channels because 0.01 watts per channel can be received properly by a fixed antenna as large as ten meters in diameter.

The use of such simple antennas on the ground, taking advantage of high power per channel in satellite, also tends to restrict the total capacity of GEO in terms of the number of satellite that can operate at the same time. An earth station antenna has an angular range of operation, measured in azimuth and elevation, over which RF energy passes through at effectively its maximum level.

1.2.3 S, X, AND Ka Bands

The bands identified by S, X, and Ka have been applied to geostationary satellite in varying degrees but generally not for commercial purposes.

S-band, normally centered at 2 GHz, lies just below C-band and was actually the frequency range used for the downlink on the first experimental synchronous satellite, SYNCOM. It is even closer that C-band to the optimum frequency for space commun-ication. The amount of bandwidth is much less than that afforded by C and Ku bands. Sharing with terrestrial services such as industrial and education television and studio to television transmitter links makes it extremely difficult to accomplish frequency coord-ination for earth stations.

Government and military satellite communication systems employ X-band and on experimental basis, Ka band. With an uplink range of 7.90 to 8.40 GHz and a downlink range of 7.25 to 7.75 GHz, X-band is used extensively for military long-haul communi-cation links much like C-band is used on a commercial basis. In highly specialized cases, Ka-band is being applied since very narrow spot beams can be transmitted to and from the satellite.

CHAPTER 2

SATELLITE ORBITS

Overview

While communications satellites perform their missions in many types of orbits, from near-earth constellations like Iridium and Globalstar to the highly inclined, eccentric Molniya orbits, one of the more important classes of orbits for these satellites is the geostationary orbit. In this column, I will examine the unique aspects of this class of orbit, which make it suitable for satellite communications.

Satellites may be launched into orbit, which make a high angle of inclination with the equatorial plane, or a low angle of inclination. High inclination orbits, including fully "polar orbits", pass over nearly all latitudes of earth. Low inclination orbits pass over only tropical latitudes.

Environmental satellites in high inclination orbits are usually launched at relatively low altitudes so that their orbital period is short relative to the 24-hour period of earth's rotation. Thus, the earth rotates slowly underneath the satellite orbit, yielding a remote sensing system, which views all parts of the globe in a 24-hour period.

Environmental satellites in low inclination orbits are launched in either of two altitude categories. Low altitude low inclination satellite orbits yield a remote sensing system, which views all of the tropical latitudes around the earth several times each day. This is the design mode for the Tropical Rainfall Monitor Mission (TRMM) satellites under development. The second category is to launch a satellite with zero angle of inclination (that is, in the equatorial plane itself) and at such a high altitude that its orbital period is precisely one day. In this orbital mode, the satellite orbits in the equatorial plane at exactly the same rate as earth rotates. Such a satellite can observe exactly the same scene over and over (as frequently as every few minutes) providing a remote sensing system which can monitor very rapidly developing weather systems.

2.1 Types of Satellite Orbits

There are several commonly used types of satellite orbits. However the satellite may find itself in any orbit, including non-desired ones after launch vehicle failure. Most frequently used are:

-**Geostationary orbit** - used for telecommunication and meteorology satellites;

-**Geotransfer orbit** - used to get to geostationary orbit;

-**Sun-Sunchronous orbit** - satellite in this orbit always remains over the daylight side of the Earth, used for earth observation;

-**Polar orbit** - satellite in this orbit travels North to South over the Poles of the planet;

-**"Molniya" orbit** - satellite at this orbit travels along tilted ellipse instead of circle.

Some special orbits are used by interplanetary spacecraft's (probes) to take an advantage of Earth (or other planet's) gravity to get more speed.

Some special terminology:

-Orbit's **altitude** is distance between the satellite and the Earth surface (sea level) for circular orbit.

-Orbit's **apogee** is the greatest distance between the satellite and the Earth surface that satellite can have while travelling along its orbit.

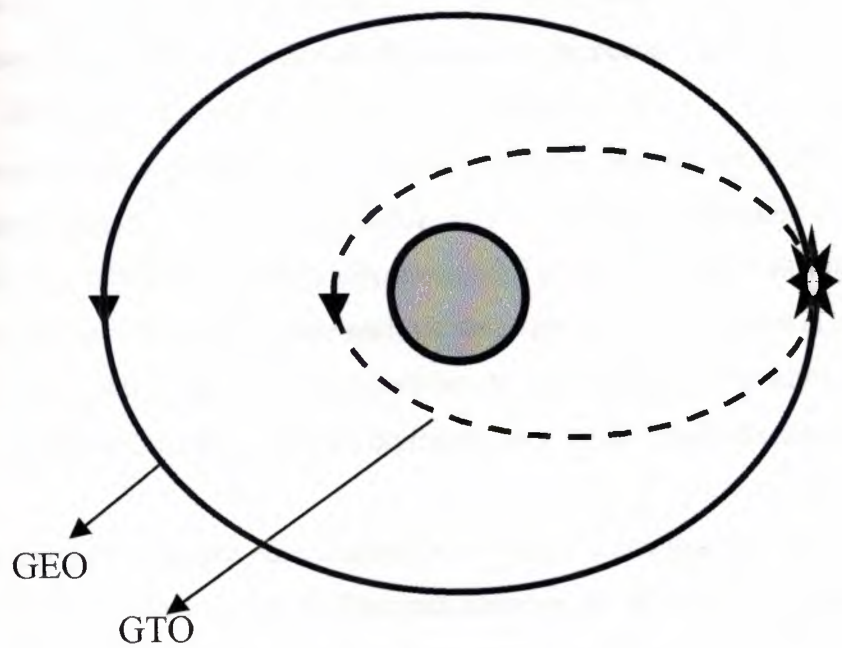
-Orbit's **perigee** is the shortest distance between the satellite and the Earth surface that satellite can have while travelling along its orbit. Apogee is equal to perigee for circular orbit.

-Orbit's **inclination** is an angle between the orbit (or the orbit's plane, to be precise) and Equator of the planet. It is 90 degrees for polar orbits and 0 degrees for equatorial.

2.1.1 Geostationary And Geotransfer Orbits.

Satellite at geostationary orbit (GEO) stay over the same point of the Earth Equator all the time, e.g. it moves on Space fast enough to follow the Earth's rotation and stay over the same spot on the Earth surface all the time. The following picture shows GEO and GTO orbits:

As long as satellite stays in GEO orbit it is viewed in the same spot in the sky from the ground all the time. Accordingly one can build antenna that will be pointed to this spot and use the satellite as a re-translator of the signals from one place on the Earth surface to another. This is a principle of satellite communications using GEO satellites.



What is a geostationary orbit?

In general terms, it is a special orbit for which any satellite in that orbit will appear to hover stationary over a point on the earth's surface. Unlike all other classes of orbits, however, where there can be a family of orbits, there is only one geostationary orbit. For any orbit to be geostationary, it must first be geosynchronous. A geosynchronous orbit is any orbit, which has a period equal to the earth's rotational period. As we shall see, this requirement is not sufficient to ensure a fixed position relative to the earth. While all geostationary orbits must be geosynchronous, not all-geosynchronous orbits are geostationary. Unfortunately, these terms are often used interchangeably. Before continuing, it is necessary to clarify what is meant by "the earth's rotational period." For most timekeeping, we consider the earth's rotation to be measured relative to the sun's mean) position. However, since the sun moves relative to the stars (inertial space) as a result of the earth's orbit, one mean solar day is not the rotational period that we're interested in.

A geosynchronous satellite completes one orbit around the earth in the same time that it takes the earth to make one rotation in inertial (or fixed) space. This time period is known as one sidereal day and is equivalent to $23^h56^m04^s$ of mean solar time. Without any other influences, the earth will be oriented the same way in inertial space each time a satellite with this period returns to a particular point in its orbit.

To ensure that a satellite remains over a particular point on the earth's surface, the orbit must also be circular and have zero inclination. The difference between a geostationary orbit (GSO) and a geosynchronous orbit (GEO) with an inclination of 20 degrees. Both are circular orbits. While each satellite will complete its orbit in the same time it takes the earth to rotate once, it should be obvious that the geosynchronous satellite will move north and south of the equator during its orbit while the geostationary satellite will not.

Orbits with non-zero eccentricity (i.e., elliptical rather than circular orbits) will result in drifts east and west as the satellite goes faster or slower at various points in its orbit. Combinations of non-zero inclination and eccentricity will all result in movement relative to a fixed ground point.

The geostationary satellite (GSO) sits fixed at the crossover point over the equator. If we now give the geosynchronous satellite an eccentricity of 0.10, the slanted teardrop shape results. Typically, eccentric geosynchronous orbits will result in a slanted over the equator —this one just happens to have the crossover point at the northern apex of the ground track

Then we can say that the only satellites which orbit with a period equal to the earth's rotational period and with zero eccentricity and inclination can be geostationary satellites. As such, there is only one geostationary orbit—a belt circling the earth's equator at an altitude of roughly 35,786 kilometers.

It should also be clear that it is not possible to orbit a satellite, which is stationary over a point, which is not on the equator. This limitation is not serious, however, since most of the earth's surface is visible from geostationary orbit. In fact, a single geostationary satellite can see 42 percent of the earth's surface and a constellation of geostationary satellites—like the one Clarke suggested—can see all of the earth's surface between 81° S and 81° N.

Of course, the advantage of a satellite in a geostationary orbit is that it remains stationary relative to the earth's surface. This makes it an ideal orbit for communications since it will not be necessary to track the satellite to determine where to point an antenna. However, there are some disadvantages. Perhaps the first is the long distance between the satellite and the ground. With sufficient power or a large enough antenna, though, this limitation can be overcome.

The fact that there is only one geostationary orbit presents a more serious limitation. Just as in putting beads on a loop of string, there are only so many slots into which geostationary satellites can be placed. The primary limitation here is spacing satellites along the geostationary belt so that the limited frequencies allocated to this purpose don't result in interference between satellites on uplink or downlink. Of course, we also want to make sure the satellites aren't close enough to run into one another since they will have some small movement.

While new communications satellites may be placed in a true geostationary orbit initially, there are several forces, which act to alter their orbits over time. Since the geostationary orbital plane is not coincident with the plane of the earth's orbit (the ecliptic) or that of the moon's orbit, the gravitational attraction of the sun and the moon act to pull the geostationary satellites out of their equatorial orbit, gradually increasing each satellite's orbital inclination. In addition, the non-circular shape of the earth's equator causes these satellites to be slowly drawn to one of two stable equilibrium points along the equator, resulting in an east-west libration (drifting back and forth) about these points.

To counteract these perturbations, sufficient fuel is loaded into all geostationary satellites to periodically correct any changes over the planned lifetime of the satellite. These periodic corrections are known as stationkeeping. North-south stationkeeping corrects the slowly increasing inclination back to zero and east-west stationkeeping keeps the satellite at its assigned position within the geostationary belt. These maneuvers are planned to maintain the geostationary satellite within a small distance of its ideal location (both north south and east west). This tolerance is normally designed to ensure the satellite remains within the ground antenna beamwidth without tracking.

Once the satellite has exhausted its fuel, its inclination will begin to grow and it will begin to drift in longitude and may present a threat to other geostatio-nary satellites. Oftentimes, geostationary satellites are boosted into a slightly higher orbit at the end of their planned lifetime to prevent them causing havoc with other geostationary satellites. This final maneuver assumes that no unplanned failure has occurred which would prevent it (such as a power or communications failure).

2.1.1.1 Geosynchronous Equation Orbit:.

The concept of the geostationary orbit, an orbit at an altitude of 35,900 kilometers whose period exactly matched the earth's rotational period, making it appear to hover over a fixed point on the earth's equator.

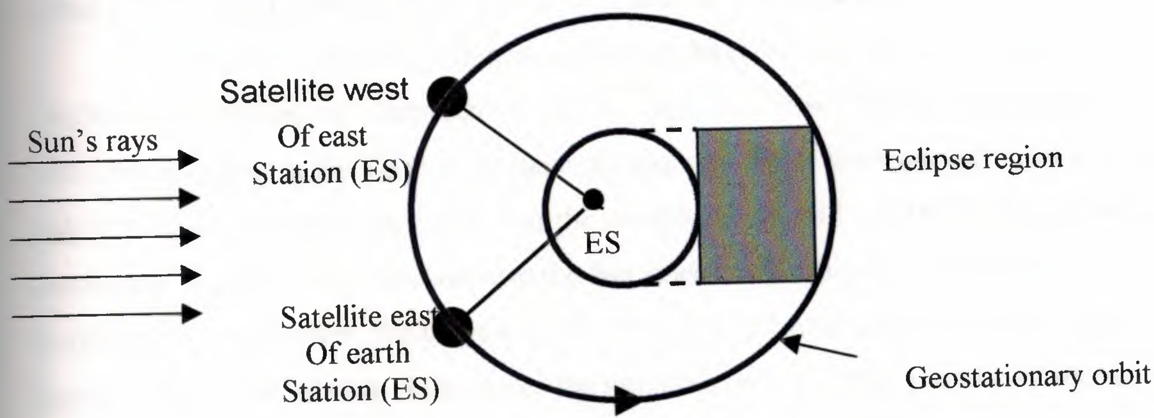
(From ***geo*** = Earth + ***synchronous*** = moving at the same rate).

A satellite in geosychonous equatorial orbit (GEO) is located directly above the equator, exactly 22,300 miles out in space. Satellites in these orbits circle the Earth at the same rate as the Earth spins. The Earth actually takes 23 hours, 56 minutes, and 4.09 seconds to make one full revolution. So based on Kepler's Laws of Planetary Motion, this would put the satellite at approximately 35,790 km above the Earth. The satellites are located near the equator since at this latitude, there is a constant force of gravity from all directions. At other latitudes, the bulge at the center of the Earth would pull on the satellite. The satellite and Earth move together. So, a satellite in GEO always stays directly over the same spot on Earth. (A geosynchronous orbit can also be called a GeoSTATIONARY Orbit.)

Geosynchronous orbits allow the satellite to observe almost a full hemisphere of the Earth. These satellites are used to study large-scale phenomenon such as hurricanes, or cyclones. These orbits are also used for communication satellites. The disadvantage of this type of orbit is that since these satellites are very far away, they have poor resolution. The other disadvantage is that these satellites have trouble monitoring activities near the poles.

2.1.1.2 Sun Transit Outage

The transit of the satellite between earth and sun such that the sun comes within the beam-width of the earth station antenna. When this happens, the sun appears as an extremely noisy source, which completely blanks out the signal from the satellite. This effect is termed sun transit outage. And it lasts for short periods each day for about 6 days around the equinoxes. The occurrence and duration of the sun transit outage depend on the latitude of the earth station, a maximum outage time of 10 minutes being typical.



-The geostationary orbit is now employed for most commercial satellites because of the following advantages:

1. The satellite remains stationary with respect to one point on earth; therefore the earth station antenna in not required to track the satellite periodically. Instead, the earth station antenna beam can be accurately aimed toward the satellite by using the elevation angle and the azimuth angle. This reduces the station's cost considerably.
2. With a 5° minimum elevation angle of the station antenna, the geostationary satellite can cover almost 38% of the surface of the earth.
3. Three geostationary satellites (120° apart) can cover the entire surface of the earth with some overlapping, except for the Polar Regions above latitudes 76°N and 76°S, assuming a 5° minimum elevation angle.

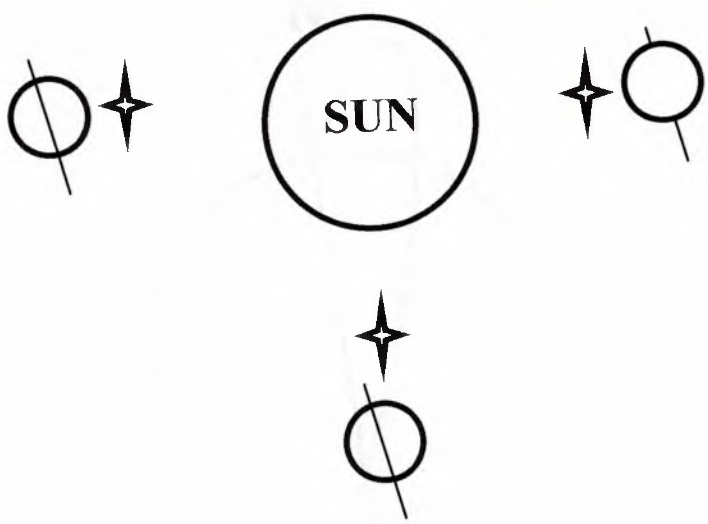
4. The Doppler shift caused by a satellite drifting in orbit (because of the gravitational attraction of the moon and the sun) is small for all the earth stations within the geostationary satellite coverage. This is desirable for many synchronous digital systems.

2.1.2 Sun-Synchronous Orbit

A satellite with a circular orbital period of one sidereal day (A sidereal day is defined as the time required for the earth to rotate once on its axis relative to the stars. Is called a synchronous satellite and has an orbit radius of **42,164.2 Km**.

This orbit is used by Earth observation satellites. Its altitude vary between 580 and 800km. And inclination is between 98 and 110 degrees. These orbits allows a satellite to pass over a section of the Earth at the same time of day. Since there are 365 days in a year and 360 degrees in a circle, it means that the satellite has to shift its orbit by approximately one degree per day. These satellites use the fact since the Earth is not perfectly round (the Earth bulges in the center, the bulge near the equator will cause additional gravitational forces to act on the satellite. This causes the satellite's orbit to either proceed or recede.

These orbits are used for satellites that need a constant amount of sunlight. Satellites that take pictures of the Earth would work best with bright sunlight, while satellites that measure longwave radiation would work best in complete darkness.

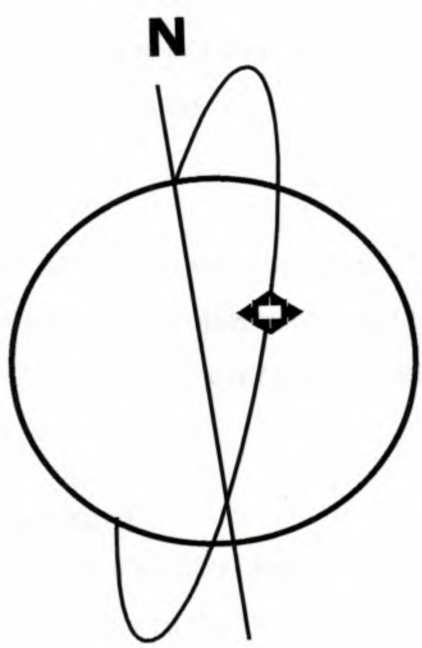


2.1.3 Molniya Orbit

This orbit is used by the first ever constellation of non-GEO communication satellites that was deployed by the USSR in 1960-th, when rocket performance was insufficient to put satellite in GEO. It has perigee 400km. And apogee 40,000km. And inclination 63deg. Constellation of sixteen orbiting satellites is arranged the way that guarantees constant visibility of at least one satellite from any point in CIS or North America. This satellite may be used similarly to GEO satellite, but its position changes with time and antenna have to turn. This orbit is broadly used by other satellites as well.

2.1.4 Polar Orbit

The more correct term would be near polar orbits. A Polar orbit is a particular type of low earth orbit. The only difference is that a satellite in polar orbit travels a north-south direction, rather than the more common east-west direction. These orbits have an inclination near 90 degrees. This allows the satellite to see virtually every part of the Earth as the Earth rotates underneath it. It takes approximately 90 minutes for the satellite to complete one orbit. These satellites have many uses such as measuring ozone concentrations in the stratosphere or measuring temperatures in the atmosphere



2.1.4.1 The Use of A Polar Orbit

Polar orbits are useful for viewing the planet's surface. As a satellite orbits in a north-south direction, Earth spins beneath it in an east-west direction. As a result, a satellite in polar orbit can eventually scan the entire surface. It's like pealing an orange in one piece. Around and around, one strip at a time, and finally you've got it all. For this reason, satellites that monitor the global environment, like remote sensing satellites and certain weather satellites, are almost always in polar orbit. No other orbit gives such thorough coverage of Earth.

2.1.4.2 Polar Coverage

While most communications satellites are in Geosynchronous orbit, the footprints of GEO satellites do not cover the Polar Regions of Earth. So communications satellites in elliptical orbits cover the areas in the high northern and southern hemispheres that are not covered by GEO satellites.

2.1.5 Low Earth Orbit

When a satellite circles close to Earth we say its in Low Earth Orbit (LEO). Satellites in LEO are just 200 - 500 miles (320 - 800 kilometers) high. Because they orbit so close to Earth, they must travel very fast so gravity won't pull them back into the atmosphere. Satellites in LEO speed along at 17,000 miles per hour (27,359 kilometers per hour)! They can circle Earth in about 90 minutes

2.1.6 Elliptical Orbit

A satellite in elliptical orbit follows an oval-shaped path. One part of the orbit is closest to the center of Earth (perigee) and the other part is farthest away (apogee). A satellite in this orbit takes about 12 hours to circle the planet. Like polar orbits, elliptical orbits move in a north-south direction.

2.2 Global Positioning System

This satellite is part of a group of satellites that can tell you your exact latitude, longitude, and altitude. The military developed the Global Positioning System (GPS), but now people everywhere can use these satellites to determine where in the world they are.

2.2.1 Other uses For GPS Satellite:

GPS satellites are used for navigation almost everywhere on Earth -- in an airplane, boat, or car, on foot, in a remote wilderness, or in a big city. Wherever you are, if you have a GPS receiver, you'll never be lost

2.3 Lunching Orbits

Satellite may be directly injected into low-altitude orbits, up to about 200 Km altitude, from a launch vehicle. Launch vehicles may be classified as expendable or reusable.

Where an orbital altitude greater than about 200 Km is required, it is not economical, in term of launch vehicle power, to perform direct injection, and the satellite must be placed into transfer orbit between the initial low earth orbit and the final high-altitude orbit. In most cases, the transfer orbit is selected to minimize the energy required for transfer, and such an orbit is known as a Hohmann transfer orbit. The time required for transfer is longest for this orbit compared to all other possible transfer orbits.

CHAPTER 3
THE SPACE SEGMANT

Overview

A satellite communication system can be broadly divided into tow segment, Aground segment and space segment. The space segment obviously includes the satellite, but it also includes the ground facilities needed to keep the satellite operational, these being referred to as *tracking, telemetry, and command* (TT&C) facilities. In many networks it is common practice to have aground station employed solely for the purpose of TT&C.

The equipment carried aboard the satellite can also be classified according to function. *Payload* refers to the equipment used to provide the service for which the satellite has been launched. *Bus* refers not only the vehicle that carries the payload, but includes the various subsystem that provide power, attitude control, orbital control, thermal control, and command and telemetry functions required to service the payload.

In a communication satellite, the equipment that provide the connecting link between transmit and receive antennas is the *transponder*. The transponder forms one of the main sections of the payload, the other being the antenna subsystem.

3.1 Power Supply

The primary electrical power for operating electronic equipment is obtained from solar cells. Individual cells can generate only small amounts of power, and there for arrays of cells in series parallel connection are required.

3.2 Attitude Control

The attitude of a satellite refers to its orientation in space. Much of the equipment Carried aboard a satellite is there for the space controlling its attitude. Attitude control is Necessary, for example, to ensure that direction antennas point in the proper direction.

Attitude control must not be confused with station keeping, which is the term used for maintaining a satellite in it's correct orbital position, although the tow are closely related.

To exercise attitude control, there must be available some measure of a satellite's orientation in space and of any tendency for this to shift. In one method, infrared sensors referred to as *horizon detectors*, are used to detect the rim of the earth against the background of space. With the use of four such sensors, one for each quadrant, the center of the earth can readily be established as a reference point. Any shift in orientation is detected by one or another of the sensors, and a corresponding control signal is generated which activates a restoring torque.

Usually, the attitude-control process takes place aboard the satellite, but it is also possible for control signal to be transmitted from earth, based on attitude data obtained from the satellite. Also, where a shift in attitude id desired, an attitude maneuver is executed. The control signal needed to achieve this maneuver may be transmitted from an earth station.

Controlling torque's may be generated in a number of ways. *Passive attitude* control refers to the use of mechanisms that stabilize without putting a drain on the satellite's energy supply.

The other form of attitude control is *active control.* With active attitude control there is no overall stabilizing torque present to resist the disturbance torque. Instead, corrective torque are applied as required in response to disturbance torque. Method used to generate active control torque includes momentum wheels, electromagnetic coils, and mass-expulsion devices such as gas jets and ion thrusters.

The three axes that define a satellite's attitude are it's *roll, pitch, and yaw (RPY)* axes, All three axes pass through the center of gravity of the satellite. For an equatorial orbit, movement of the satellite about the roll axis moves the antenna footprint north and south, movement about the pitch axis moves the footprint east and west, and movement about the yaw axis rotates the antenna footprint.

3.2.1 Spin Stabilization

Spin stabilization used with cylindrical satellite. The satellite is constructed so that it is mechanically balanced about one particular axis, and is then set spinning around this axis. For geostationary satellites the spin axis is adjusted to be parallel to the N-S axis of the earth

In the absence of disturbance torque, the spinning satellite would maintain it's correct attitude relative to the earth. Disturbance torque are generated in a number of ways, both external and internal to the satellite, solar radiation, gravitational gradients, and meteorite impacts are all examples of external forces that can give rise to disturbance torque. Motor bearing fraction and the movement of satellite elements such as the antennas can also give rise to disturbance torque. The over all effect is that the spin rate will decrease and direction of the angular spin axis will change. Impulse-type thrusters, or jet, can be used to increase the spin rate again and to shift the axis back to it's correct N-S orientation, *Nutation,* which is a form of wobbling, can occur as a result of the disturbance torque's, and/or from misalignment or imbalance of the control jets. This nutation must be damped out by means of energy absorbers known as nutation dampers.

Tow forms of spin stabilization are commonly employed. In what is referred to simply as *spin stabilization*, the entire satellite about an axis, which for earth-orbiting satellites is the pitch axis. Where an omnidirectional antenna is used, the antenna, which point along the pitch axis, also rotates with the satellite. Where a directional antenna is used, which is more common for communication satellite, the antenna must be despun, giving rise to a dual-spin construction.

3.2.2 *Three-axis stabilization*

In three-axis stabilization, there are stabilizing elements for each of the three axes, roll, pitch, and yaw. Because the body of the satellite remains fixed relative to earth, Three-axis stabilization is also known as body stabilization.

Active attitude control is used with three-axis stabilization. This may take the form of control jets (mass-expulsion controllers) fired to correct the attitude of the satellite. Reaction wheels also can be used. A reaction wheel is a flywheel that is normally stationary but reacts when a torque tends to shift the spacecraft orientation.

3.3 Station Keebing

In addition to attitude control, it is important that a geostationary satellite remain in its correct orbit slot. The equatorial bulge of the earth causes geostationary satellite to drift slowly along the orbit. To counter this drift, an oppositely directed velocity component is imparted to the satellite by means of jets, which is pulsed every 2 or 3 weeks, This result in the satellite drifting back through tit's nominal station position, coming to a stop, and recommending the drift along the orbit until the jets are pulsed once again. These maneuvers are termed east-west station keeping maneuvers. For satellite in the 6/4-GHz band, a satellite must be kept within ±0.1 deg of its designated longitude, and in the 14/12-GHz band, within ±0.05 deg.

To prevent the shift in inclination from exceeding specified limits, jets may be pulsed at the appropriate time to return the inclination to zero. Counteracting jets must be pulsed when the inclination is at zero to half the change in inclination. These maneuvers are termed northwest station-keeping maneuvers, and they are much more expensive in fuel than are east-west station keeping maneuvers. The north-south station keeping tolerances are the same as those east-west station keeping, ±0.1 deg in the C-band and ±0.05 deg in the Ku-band.

Orbital correction is carried out by command from the TT&C earth station, which monitors the satellite position. East-west and north-south station–keeping maneuvers are usually carried out using the same thrusters as are used for attitude control.

3.4 Thermal Control

The Thermal Control sub-system protects all the satellite's equipment from damage in the harsh space environment. In orbit, a satellite is exposed to extreme temperature changes -- from 120 degrees below zero in the shade, to 180 degrees above zero in the Sun.

The most important consideration is that the satellite equipment should operate as near as possible in a stable temperature environment. Various steps are taken to achieve this. Thermal blankets and shields may be used to provide insulation. Radiation mirrors are often used to remove heat from the communication payload.

To maintain constant-temperature conditions, heaters may be switched on (usually on command from the ground) to make up for the heat reduction that occurs when transponders are switched off.

3.5 Tt&C Subsystem

The telemetry, tracking, and command (TT&C) subsystem perform several Routine Functions aboard a spacecraft. The telemetry or ''telemetering'' function could be interpreted as ''measurement at a distance'' Specifically, it refers to overall operation of generating an electrical signal proportional to the quantity being measured, and encoding and transmitting this to a distant station, which for the satellite is one of the earth station. Data that are transmitted as telemetry signals include attitude information such as that obtained from sun and earth sensors; environmental information such as the magnetic field intensity and direction; the frequency of meteorite impact, and so on, and spacecraft information such as temperatures, power supply voltages, and stored-fuel pressure, Certain frequencies have been designated by international agreement for satellite telemetry transmission. During the transfer and drift orbital phases of the satellite launch, a special channel is used along with an omnidirectional antenna, Once the satellite in station, one of the normal communication transponders may be used along with its directional antenna, unless some emergency arises that makes it necessary to switch back to the special channel used during the transfer orbit.

Telemetry and command may be though of as complementary functions, The telemetry subsystem transmits information about the satellite to the earth station. While the command subsystem receives command signals from the earth station, often in response to telemetered information. The command subsystem demodulates, and if the necessary decodes, The command signals, and routes these to the appropriate equipment needed to execute the necessary action. Thus attitude changes may be made, communication transponders switched in and out of circuits, antenna redirection, and station-keeping maneuvers carried out on command. It is clearly important to prevent unauthorized command signals are often encrypted

Tracking of the satellite is accomplished by having the satellite transmit beacon Signals which are received at the TT&C earth station. Tracking is obviously important during the transfer and drift orbital phase of the satellite launch. When on-station, A geostationary satellite will tend to shifted as a result of the various disturbing force, therefore, it is necessary to be able to track the satellite movement and send correction signals as required. Tracking beacons may be transmitted in the telemetry channel, or by means of pilot carriers at frequencies in one of the main communication channels, or through special tracking antennas. Satellite range from the ground station is also required from time to time. This can be determined by measurement of the propagation delay of signals specially transmitted for ranging purposes.

3.6 Transponders

A transponder is the series of interconnected units which forms a single communication channel between the receive and transmit antennas in a communication satellite. Some of the units utilized by transponder in a given channel may be common to A number of transponder. Thus, although reference may be made to specific transponder, this must be though of as an equipment channel rather than a single item of equipment.

3.6.1 Wideband Receiver

In the wideband receiver a duplicate receiver is provided so that if one fails, the other is switched in automatically. The combination is referred to as a *redundant receiver*, meaning that although tow is provided, only one is in use at a given time.

The first step in the receiver is a low-noise amplifier (LNA). This amplifier adds little nose to the carrier being amplified, and at the same time it provides sufficient amplification for the carrier to override the higher noise level present in the following mixer stage.

The LNA feeds into a mixer stage, which also requires a local oscillator signal for the frequency-conversion process. The power drive from the local oscillator to the mixer input is about 10 dBm. The oscillator frequency must be highly stable and have low phase noise. A second amplifier follows the mixer stage to provide an overall receiver gain of about 60 dB. Splitting the gain between the preamplifier at 6 GHz and the second amplifier at 4 GHz prevents oscillation which might occur if all the gain were to be provided at the same frequency.

The wideband receiver utilizes all solid-state active devices. In some designs, tunnel-diode amplifiers have been used for the preamplifier at a 6 GHz in 6/4-GHz transponder, and for parametric amplifiers at 14 GHz in 14/12-GHz transponder. With advances in FET technology, FET amplifiers which offer equal or better performance are now available for both bands. Diode mixer stages are used. The amplifier following the mixer may utilize BJTs at 4 GHz, and FETs at 12 GHz or FETs may in face used in used in both bands.

3.6.2 Input Demultiplexer

The input demultiplexer separates the broadband input, covering the frequency range 3.7 to 4.2 GHz into the transponder frequency channels. The channels are usually arranged in even-numbered and odd-numbered groups. This provides greater frequency separation between adjacent channels in a group, which reduces adjacent channel interference. The output from the receiver is fed to a power splitter, which in turn feeds the two separate chains of circulators. The full broadband signal is transmitted to each circulator. Although there are considerable losses in the demultiplexer, these are easily made up in the overall gain for the transponder channels.

3.6.3 Power Amplifier

A separate power amplifier provides the output power for each transponder channel. An input attenuator precedes each power amplifier. This is necessary to permit the input drive to each power amplifier to be adjusted to the desired level. The attenuator has a fixed section and a variable section. The fixed attenuation is needed to balance out variations in the input attenuation so that each transponder channel has the same nominal attenuation, the necessary adjustment being made during assembly. The variable attenuation is needed to set the level as required for different types of service and to about 60 dB. Splitting the gain between the preamplifier at 6 GHz and the second amplifier at 4 GHz prevents oscillation which might occur if all the gain were to be provided at the same frequency.

The wideband receiver utilizes all solid-state active devices. In some designs, tunnel-diode amplifiers have been used for the preamplifier at a 6 GHz in 6/4-GHz transponder, and for parametric amplifiers at 14 GHz in 14/12-GHz transponder. With advances in FET technology, FET amplifiers which offer equal or better performance are now available for both bands. Diode mixer stages are used. The amplifier following the mixer may utilize BJTs at 4 GHz, and FETs at 12 GHz or FETs may in face used in used in both bands.

3.6.2 Input Demultiplexer

The input demultiplexer separates the broadband input, covering the frequency range 3.7 to 4.2 GHz into the transponder frequency channels. The channels are usually arranged in even-numbered and odd-numbered groups. This provides greater frequency separation between adjacent channels in a group, which reduces adjacent channel interference. The output from the receiver is fed to a power splitter, which in turn feeds the two separate chains of circulators. The full broadband signal is transmitted to each circulator. Although there are considerable losses in the demultiplexer, these are easily made up in the overall gain for the transponder channels.

3.6.3 Power Amplifier

A separate power amplifier provides the output power for each transponder channel. An input attenuator precedes each power amplifier. This is necessary to permit the input drive to each power amplifier to be adjusted to the desired level. The attenuator has a fixed section and a variable section. The fixed attenuation is needed to balance out variations in the input attenuation so that each transponder channel has the same nominal attenuation, the necessary adjustment being made during assembly. The variable attenuation is needed to set the level as required for different types of service and to compensate for equipment degradation in service. Because this attenuation adjustment is an operational requirement, it must be under the control of the ground TT&C station.

Traveling-wave-tube amplifiers (TWTAs) are widely used in transponder to provide the final output power required to the transmit antenna. In the TWT, an electron-beam gun assembly consisting of heater, a cathode, and focusing electrodes is used to form an electron beam. A magnetic filed can be provided by mean of solenoid and dc power supply, the comparatively large size and high power consumption of solenoids makes them unsuitable for use aboard satellites, and lower-power TWTs are used which employ permanent-magnet focusing.

The RF signal to be amplified is coupled into the helix at the end nearest the cathode, and sets up a traveling wave along the helix. The electric field of the wave will have a component along the axis of the helix. In some regions, this field will decelerate the electrons in the beam. The average beam velocity, which is determined by the dc potential on the tube collector, is kept slightly greater than the phase velocity of the wave along the helix. Under these conditions an energy transfer takes place, kinetic energy in the beam being converted to potential energy in the wave. The wave actually travel around the helical path at close to the speed of light, but it is the axial component of wave velocity that interacts with the electron beam. This component is less than the velocity of light approximately in the ratio of helix pitch to circumference. Because of this effective reduction in phase velocity, the helix is referred to as a slow-wave structure.

The advantage of the TWT over other types of tube amplifiers is that it can provide amplification over a very wide bandwidth. Input levels to the TWT must be carefully controlled, however, to minimize the effects of certain forms of distortion. The worst of these results from the nonlinear transfer characteristic of the TWT, At low input powers, the output-input power relationship is linear, that is, a given decibel change in input power will produce the same decibel change in output power. At higher powers inputs, the output power saturates, the point of maximum power output being known as the saturation point, The saturation point is a very convenient reference point, and input and output quantities are usually referred to it .The linear region of the TWT is defined as the region bound by the thermal noise limit at the low end. And by what is termed the 1-dB compression point at the upper end. This is the point where the actual transfer curves drops 1 dB below the extrapolated straight line.

The absolute time delay between input and output signals at a fixed input level is generally not significant. However, at higher input levels, where more of the beam energy is converted to output power, the average beam velocity is reduced, therefore the delay time is increased. Since phase delay is directly proportional to time delay, this results in a phase shift, which varies with input level.

Frequency modulation is usually employed in satellite communication circuits. However, unwanted amplitude can occur from the filtering that takes place prior to the TWT input. The AM/PM process convert the unwanted amplitude modulation to phase modulation, which appears as noise on the FM carrier. Where only a single carrier is present, it may be passed through a hard-limiter before being amplified in the TWT.the hard-limiter is a circuit that clips the carrier amplitude close to the zero baseline to remove any amplitude modulation. The frequency modulation is preserved in the zero-crossover points and is not affected by the limiting.

A TWT may also be called upon to amplify two or more carriers simultaneously, this being referred to as multicarrier operation. The AM/PM conversion is then a complicated function of carrier amplitudes, but in addition, the nonlinear transfer characteristic introduces a more serious form of distortion known as intermodulation distortion.

3.7 Antenna Subsystem

The antenna carried aboard a satellite provides the dual functions of receiving the uplink and transmitting the downlink signals. They range from dipole-types antennas. Where omindirectional characteristics are required, to the highly directional antenna required for telecommunication purposes and TV relay and broadcast.

Directional beams are usually produced by mean of reflector-type antennas, the paraboloidal antenna being the most common. The gain of the paraboloidal antenna can be increased and the beam-width made narrower by increasing the reflector size or decreasing the wavelength. The largest reflectors are those that are required for the 6/4-GHz band.

Wide beams for global coverage are produced by simple horn antenna at 6/4-GHz. These horns beam the signal directly to the earth without the use of reflectors. Also a simple biconical dipole antenna is used for the tracking and control signal. The complete antenna platform and the communication payload are despond to keep the antennas pointing to their correct location on earth. The same feed horn may be used to transmit or receive carriers with the same polarization. The transmit and receive signal are separated in a device known as a diplexer, and the separation is further aided by means of frequency filtering, Polarization discrimination may be also be used to separate the transmit and receive signals using the same feed horn. Separate horns may also be used for the transmit and receive functions, with both horns using the same reflector.

CHAPTER 4

TRAFFIC CAPACITY

AND ACCESS CONTROL

Overview

A transponder channel aboard a satellite may be fully loaded by a single transmission from an earth station. This is referred to as single access mode of operation. If is also possible, and more common, for a transponder to be loaded by a number of carriers. These may originate from a number of earth stations geographically separate, and each earth station may transmit one or more of the carriers. This mode of operation is termed multiple access. The need for multiple access arises because more than tow earth stations, in general, will be within the service area of a satellite. Even so called spot beams from satellite antennas cover areas several hundred miles across.

The two most commonly used methods of multiple access are frequency division multiple access (FDMA) and time division multiple access (TDMA). These are analogous to frequency division multiplexing (FDM) and time division multiplexing (TDM), multiple access and multiplexing are different concepts, modulation (and hence multiplexing 0 is essentially a transmission feature, whereas multiple access is essentially a traffic feature.

A third category of multiple access is code division multiple access (CDMA). In this method each signal is associated with a particular code that is used to spread the signal in frequency and or time. All such signals will be received simultaneously at an earth station, but by using the key to the code the station can recover the desired signal by means of correlation, The other signals occupying the transponder channel appear very much like random noise to the correlation decoder. The two subsets of CDMA are spread spectrum multiples access(SSMA) and Pulse address multiple access (PAMA).

Multiple access may also be classified by the way in which circuits are assigned to users ("circuits" in this context implies one communication channel through the multiple access transponder). Circuits may be preassigned, which means they are allocated on a fixed or partially fixed basis to certain users. Preassignment is simple to implement, but is efficient only for circuits with continuous heavy traffic.

An alternative to preassignment is demand-assigned multiple access (DMAA) In this method, all circuits are available to all users and are assigned according to the demand. DAMA results in more efficient overall use of the circuits but is more costly and complicated to implement.

Both FDMA and TDMA can be operated as preassigned or demand-assigned systems. CDMA is a random access system , there being no control over the timing of theaccess or of the frequency slots accessed.

Theses multiple access carries a number of refer to the way in which a single transponder channel is utilized. A satellite carriers a number of transponders, and normally each covers a different frequency channel. This provides a form of frequency division multiple access to the whole satellite. It is also possible for transponder to operate at the same frequency but to be connected to different spot beam antennas. These allow the satellite as whole to be accessed by earth stations widely separated geographically but transmitting on the same frequency. This is termed frequency reuse. This method of access is referred to as space division multiple access (SDMA).

4.1 Single Access

With single access, a single modulated carrier occupies the whole of the available bandwidth of a transponder. Single access operation is used on heavy traffic routes, and requires large earth station antennas.

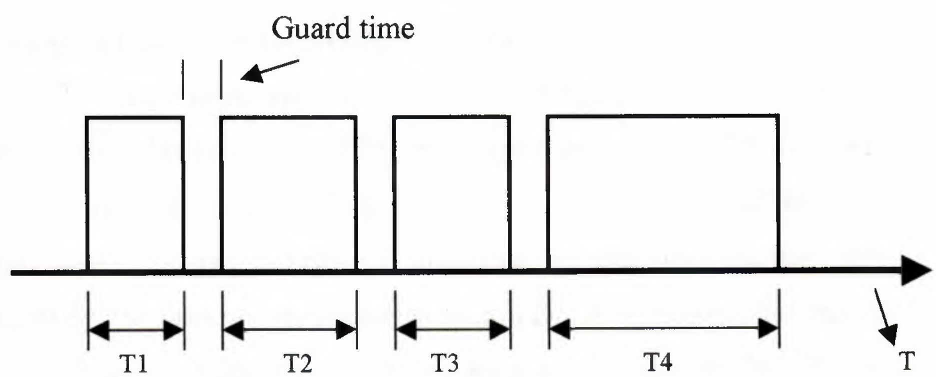
4.2 Fixed-assignment technique

Fixed-assignment techniques, such as time division multiple access (TDMA) and frequency division multiple access (FDMA), incorporate permanent subchannel assignments (in the time-frequency domain) for individual users. These schemes perform well when each user transmitting a steady flow of messages, however, in burst traffic applications, they are inefficient.

4.2.1 Time Division Multiple Access

In TDMA the earth stations that share the satellite transponder use a carrier at the same frequency for transmission on a time division basis. Earth stations are allowed to transmit traffic bursts in a periodic time frame called the TDMA frame. During the burst, an earth station has the entire transponder bandwidth available to it for transmission. The transmit timing of the bursts is carefully synchronized so that all the burst arriving at the satellite transponder are closely spaced in time but do not overlap. The satellite transp0onder receives one burst at a time, amplifiers it, and retransmits it back to earth. Thus every earth station in the satellite beam served by the transponder can receive the entire burst stream and extract the bursts intended for it.

In TDMA scheme each user will transmit over the full system bandwidth of B Hz but only during its own time slot in each frame. TDMA will require guard times to separate one users transmission from that of next one in the frame. Since the transmission from each user is essentially done in the burst mode, additional preambles may also be required to enable carrier synchronization and clock synchronization, these synchronization procedures are essential to allow the receiver to receive the information contained in the transmitted burst.



TDMA FRAME

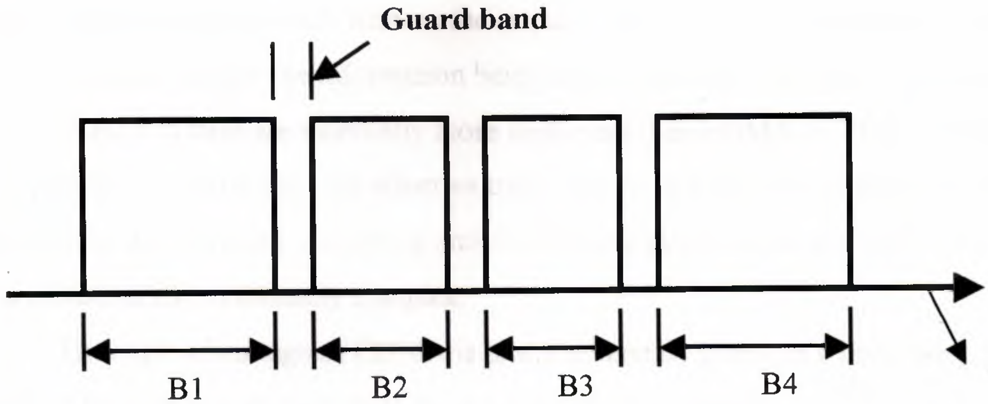
A TDMA system also requires overall system timing so that each user can unambiguously determine the frame boundaries and the position of its own slot within the frame. The receivers in a TDMA system must be capable of rapid burst synchronization, that is quickly acquiring carrier at the beginning of the burst they are intended to receive. In a TDMA system at any given instant only one user will be in the transmitting mode. This makes it easier to exercise power control in the system by requiring each user to adhere to a pre-specified transmission power limit. Because of the system may be operated such that the satellite transponder is close to its saturation limit. this ensures greater power efficiency in the system.

4.2.2 Frequency Division Multiple Access

In FDMA each earth station in the community of earth station that share the transponder capacity transmits one or more carriers to the satellite transponder at different center frequencies. Each carrier is assigned a frequency band in the transponder bandwidth. Along with a small guard band to avoid interference between adjacent carriers. The satellite transponder receives all the carriers in its bandwidth, amplifiers them, and retransmits them back to earth. The earth station in the satellite antenna beam served by the transponder can select the carrier that contains the messages intended for it.

In some circumstances, FDMA schemes may be comparatively inefficient since they require guard bands between neighbouring bands to prevent the transmission in one band from interfering with those in neighbouring bands, These guard bands cannot be used for information transfer and represent a waste of the system resources.

Another problem with the FDMA system is that they require the satellite transponder to be linear in nature. The high power amplifier (HPA) on board the satellite is usually constructed from a travelling wave tune (TWT) device. An HPA works most efficiently when it is operated close to saturation but then nonlinearities occur. Unfortunately, the linearity requirements of an FDMA system implies that the TWT has to be backed off substantially in order to operate it as a linear amplifier. This in true, leads to inefficient usage of the available satellite power. For this reason fixed assignment strategies have tended to increasingly shift from FDMA to TDMA as certain technical problems associated with the latter were overcome.



Transponder bandwidth
Concept of FDMA

4.2.3 Code Division Multiple Access

In CDMA each user employs a particular code address to spread the carrier bandwidth over a much larger band width so that the earth station community can transmit simultaneously without frequency or time separation and low interference.

For digital transmission, it is also possible to have number of users share the same bandwidth and transmission time, while keeping them separated by ensuring that they use different codes, which are orthogonal to each other.

For any given n, there exist special binary sequences, referred to as pseudo-random or pseudo-nose (PN) sequences. For large values of n, a number of such sequences will be available, The sequences have interesting property of being orthogonal to each other.

CDMA uses this property to separate the transmission of different users even though they may overlap each in frequency and/or time. By considering a situation where the choice of n is such that N such sequences are available where each sequence is of length M,

$$M = 2^n - 1$$

Each station is assigned a sequence for transmitting its message bits. The mode of transmission is to EX-OR (exclusive 'or') each data bit with the specified sequence and then transmit the resulting sequence. Known the sequence being used for transmission, the receivers can then extract the information being sent essentially a process of correlation.

CDMA system are inherently more inefficient than FDMA or TDMA schemes. This is primarily because the high effective transmission rate requires considerably more bandwidth for its operation. Acquiring and maintaining clock and carrier synchronization in such a system is also reasonably complex.

The main advantage of CDMA is that it inherently provides a processing gain of M and can be used in systems, which must transmit at low power.

In environments with nose, this processing gain may also be used to advantage. The inherent anti-jam (AJ) and anti-interference (AI) properties of the PN-sequences provide a certain amount of data transmission in a CDMA system when used in a military environment.

4.3 Random access protocol

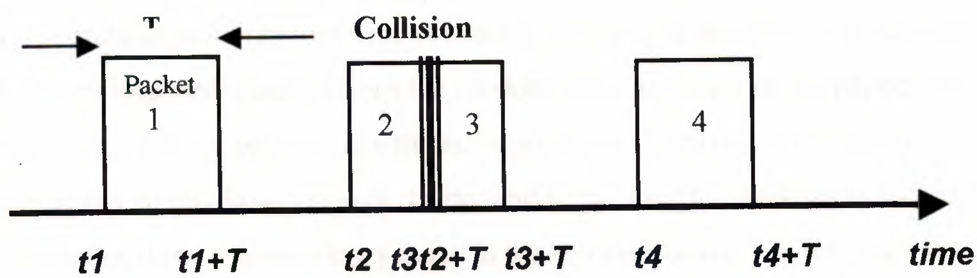
Burst traffic is serviced more efficiently by random access protocol. The ALOHA and carrier sense multiple access (CSMA) schemes are typical examples. Major problems with ALOHA are low achievable goodput and stability. In CSMA, stability problem persist, besides, it is difficult to sense collisions in a wireless channel, performance rapidly deteriorates as the maximum propagation delay increases. Details of ALOHA and its variants, CSMA and its variants can be found in csc6220 textbook chapter 7. For LANs, random access protocols take advantage of the short propagation delays between users.

4.3.1 Random Access Protocols

Random Access Protocols address the situation where the traffic is bursty. With bursty traffic a station can be busy for short amounts of time and idle for the remaining stretches. When the channel assignment is fixed, as is the case with FDMA and TDMA, those idle periods amount to wasted bandwidth even though other stations may have data to transmit. One solution would be to allow stations to transmit whenever they have data to send. The obvious result eventually would be collided data (i.e. two or more stations transmitting at the same time). Because there is no guarantee of successful transmission in advance individual stations account for and recover from corrupted data packets.

4.3.2 Pure Aloha Protocol

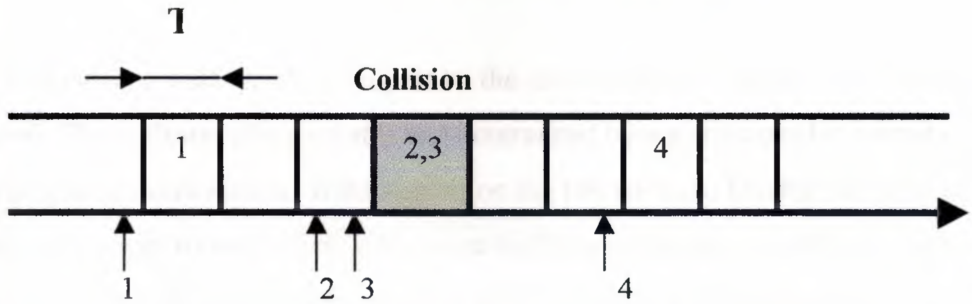
With Pure Aloha, stations are allowed access to the channel whenever they have data to transmit. Because the threat of data collision exists, each station must either monitor its transmission on the rebroadcast or await an acknowledgment from the destination station. By comparing the transmitted packet with the received packet or by the lack of an acknowledgement, the transmitting station can determine the success of the transmitted packet. If the transmission was unsuccessful it is resent after a random amount of time to reduce the probability of re-collision.



forward channel of a pure Aloha system

4.3.3 Slotted Aloha Protocol

By making a small restriction in the transmission freedom of the individual stations, the throughput of the Aloha protocol can be doubled. Assuming constant length packets, transmission time is broken into slots equivalent to the transmission time of a single packet. Stations are only allowed to transmit at slot boundaries. When packets collide they will overlap completely instead of partially. This has the effect of doubling the efficiency of the Aloha protocol and has come to be known as Slotted Aloha.



The forward channel of a slotted Aloha system

. The channel bandwidth is a continuous stream of slots whose length is the time necessary to transmit one packet. A station with a packet to send will transmit on the next available slot boundary. In the event of a collision, each station involved in the collision retransmits at some random time in order to reduce the possibility of recollision. Obviously the limits imposed which govern the random retransmission of the packet will have an effect on the delay associated with successful packet delivery. If the limit is too short, the probability of recollision is high. If the limit is too long the probability of recollision lessens but there is unnecessary delay in the retransmission. For the Mars regional network studied here, the resending of the packet will occur at some random time not greater than the burst factor times the propagation delay.

Another important simulation characteristic of the Slotted Aloha protocol is the action which takes place on transmission of the packet. Methods include blocking (i.e. prohibiting packet generation) until verification of successful transmission occurs. This is known as "stop-and-wait". Another method known as "go-back-n" allows continual transmission of queued packets, but on the detection of a collision, will retransmit all packets from the point of the collision. This is done to preserve the order of the packets. In this simulation model queued packets are continually sent and only the packets involved in a collision are retransmitted. This is called "selective-repeat" and allows out of order transmission of packets.

4.3.3.1 Simulation Results

The Slotted ALOHA simulation results Compared to GTDMA the utilization and delay results come up short, however the flexibility of Slotted ALOHA greatly exceeds that of GTDMA or TDMA. With respect to the station queue results the following should be noted. The collision of a packet is not determined by a station until it is read one propagation delay after its transmission on the rebroadcast. During this time the packet must be kept in some kind of hold queue buffer until the success of the transmission is known. This hold queue buffer must be of sufficient size to accommodate the number of packets which can be transmitted during one propagation delay. After the result of the transmission is determined, the packet is either purged from the hold queue or requeued in the transmission queue. Since the state of the hold queue is constant and the effects of collisions are seen in the size of the transmission queue, the packets in the hold queues are not represented in the graphs. With that said, the most remarkable statistic depicted in the queue buffer graphs is the almost non-existent average. This indicates that while these queue buffers do exhibit some peaks, mostly due to the burst factor, most of the time they are empty.

4.3.3.2 Slotted Aloha Network Stability Bhaviour

It is known through theoretical analysis of the Slotted ALOHA protocol that the maximum achievable throughput is or about 0.368 for a Poisson distributed network with uniform traffic. However, the excess capacity exists in a slotted ALOHA network when there is one large user and several small users. the network achieves stability at a channel speed of 900 Mb/sec (in fact it is something less than 900 Mb/sec for burst=1) which equates to a utilization greater than .40. This is excess capacity due to the disparity in station traffic (i.e. three large users and three small users). The results above are for a burst factor of 1. Higher burst factors require slightly higher channel speeds for stability so we will begin our simulation execution at 1000 Mb/sec.

4.4 Demand Assignment Access

The previous two categories spanned the spectrum of restrictive to unrestrictive access to the channel. On the one hand, restricting access to the channel limits a station's adaptability to bursty data. On the other hand, by giving stations the freedom to transmit whenever required, the channel efficiency is limited due to collisions. Protocols in which a station's intentions are announced in the form of some kind of explicit or implicit reservation are called Demand Assignment protocols.

4.4.1 Demand assignment techniques(centrally or distributed control)

Demand assignment techniques provide channel capacity to users on a demand basis. Demand assignment involves two stages: a reservation stage followed by a transmission stage. In the first stage, a portion of the channel capacity is required, short reservation packets are sent to request channel time in contention mode. Once channel time is reserved, information packets are transmitted conflict-free. Control of the reservation and transmission stages can be either centralized or distributed. A common example of demand assignment with central control is polling. The proper operation depends on the reliability of the controller. With distributed control, users base their actions on the broadcast information.

4.4.2 Reservation Aloha

This channel allocation scheme divides the channel bandwidth into slot sizes equal to the transmission time of a single packet. This again assumes that the packet sizes are of constant length. The slots are organized into frames of equal size whose length spans the length of one propagation delay. A station makes an implicit reservation by successfully transmitting in an available slot. After a successful transmission the station is guaranteed that slot in succeding frames until it is no longer required. A slot becomes unused either by going empty in the previous frame or by collision in the previous frame. The remaining stations can then compete for unused slots using Slotted Aloha. If the frame size did not span the propagation delay, a slot could go unused successive times before stations sensed its availability.

Because of Slotted ALOHA's limited utilization potential, various methods have been developed which improve upon its efficiency. One of these methods is known as Reservation Aloha.The main modification has to do with slot ownership after a successful packet transmission. With Slotted ALOHA any slot is available for use by any station regardless of its prior usage. With Reservation ALOHA the slot is considered owned temporarily by the station which used it successfully. When the station is through with the slot it simply stops sending. An idle slot is available to all stations on a contention basis.

Given the requirement that the frame size must span the propagation delay, the number of slots per frame varies with the channel speed. Table"A" shows the frame sizes in slots for a variety of channel speeds and packet sizes and a propagation delay of 0.113 seconds.

Table A: Frame sizes for various channel speeds.

Channel Speed (Mb/sec)	**400**	**500**	**600**	**700**	**800**	**900**	**1000**	**1100**	**1200**	**1300**
.5 Mb Slot	91	113	136	159	181	204	226	249	272	294
1.0 Mb Slot	46	57	68	80	91	102	113	125	136	147
1.5 Mb Slot	31	38	46	53	61	68	76	83	91	98

As we can see, the number of slots per frame for a 1 Mb slot greatly exceeds the number of stations in the Mars regional network. For this reason, a station is allowed to attempt to access more than one slot per frame as needed. Additionally, a station will attempt to acquire a slot with a probability proportional to the length of its transfer queue. For example, a station in which the current transfer queue length is 80% of it's maximum would attempt acquisition of the slot with a higher probability than a station whose queue length is 20% of it's maximum.

4.4.3 First-In First-Out (FIFO) Reservation

This scheme is unique in that it uses smaller reservation packets and a so called queue-in-the-sky to control access the channel. The channel bandwidth is once again divided into packet sized slots. Every so often a slot is divided into several smaller reservation packets. When a station has packets to transmit, it attempts to transmit a reservation packet in one of the reservation slots using Slotted ALOHA access. The reservation packet contains the request information necessary to reserve a number of data packets up to some maximum. The data slots are then assigned using a FIFO method. Upon a successful reservation, tranmission all stations become aware of the reservation request on the rebroadcast and each maintains a record of outstanding reservations and ultimately when its own reservations begin.

4.4.3.1 FIFO Protocol

In this protocol, channel access is gained by explicitly making a reservation transmission such that all other stations are aware of the intention to transmit at some agreed upon future time. This is done by dividing the channel bandwidth into slot sizes equal to the transmission time of a single slot and then organizing some number of those slots into a frame. After every M slots one slot is divided into V smaller reservation slots. These reservation slots are contended for using an Aloha technique. When a station wishes to transmit, it randomly picks a reservation slot and attempts the transmission of a reservation packet. A reservation packet contains the necessary information to reserve 1 to N packet slots. On the broadcast of the reservation packet, all stations are aware of the sending station's reservation request and consequently update their individual queue status. In this way all stations are aware of which station is to transmit in upcoming packet slots. When the global queue requests drop to zero, the channel reverts to only reservation slots. The value of M is adaptable to network traffic based upon an agreed to algorithm.New or out of synch stations can resynchronize themselves by reading other station's queue status, which is appended to each data packet.

4.4.4 Round Robin Reservation

Initially this protocol resembles TDMA in that the channel bandwidth is divided into equal size slots and are owned by individual stations. The difference here is that when a station is not using one of its "owned" slots, that slot can be borrowed and used by other stations as necessary. Whenever the owning station needs its slot back it causes a collision. On a collision, everyone but the owner must desist from using the slot.

4.5 Packet Delay vs. Utilization

This performance measurement focuses on the effect of increasing the channel utilization of a protocol on the delay associated with completing the packet transmission. It is a widely used performance benchmark from which much can be seen about the behavior of a MAC protocol. There does seem to be a lack of consensus as to the definition of throughput versus the definition of utilization.

Throughput is defined as the ratio of successful transfers to transfer attempts and is denoted by . In the case of fixed assignment protocols (no contention), this value is always 1 since, barring any transmission errors which we ignore, the slot is owned by the station and thus not susceptible to collisions. In contention based protocols, this definition of throughput is a descriptive representation of how well packets are being processed with respect to their transmission.

The amount of the channel bandwidth used by a station in relation to the amount of channel bandwidth available to the station is defined as *utilization* and denoted as . It is the author's opinion that the term utilization best describes this value since it is a measure of how well the channel is utilized.

The *delay* of a packet is defined as the amount of time between the point when a packet is inserted into the transfer queue to the point when the last bit has successfully been sent. Another way to look at this value is the amount of time the packet spent queued at the transmitting station. It is important to note that a packet remains in some queue until it has been determined that it has arrived at the receiving station successfully. This simulator does not deal with acknowledgments as they are inherent in broadcast channels since the sender receives a copy of its packet. With this in mind a successful transmission is one in which no packet collision took place.

Intuitively it is expected that the delay of a packet should increase with the channel utilization. This is the case with most protocols with differences lying in the position of the knee of the curve and the slope approaching the knee.

4.6 Protocol Performance Characteristics

In order to compare the results of the simulation of each protocol, we will focus on three key areas: packet delay at various channel utilization levels, transfer queue characteristics at various channel speeds and packet delay at various packet sizes. The network definition against which all simulation models are run is the six station Mars regional network.

In the evaluation of all protocols we make the following common assumptions:

1. Channel speed is represented in C bits per second.

2. Packet sizes of b bits are uniform and a slot is equal to the transmission time of one packet .
3. The channel is error-free except for collisions.
4. Messages are generated concurrently with transmission, thus queues are required to regulate transmission and generation discrepansies.
5. Each message generation consists of a group of packets whose number is geometrically distributed with a mean equal to the burst factor (L).
6. Simulation time is set at 1,000,000 milliseconds. This translates to 16.67 minutes. While longer time settings were tried, no perceptible difference was seen for the protocols tested.
7. The interarrival times are in units of slots and are geometrically distributed with a mean where is the probability that a message is generated during slot time.

4.7 Evaluation Methodology

The performance evaluations of the individual protocols will attempt to follow the following methodology:

1. Determine stability limit. A *stable system* is defined to be one in which the rate at which packets are serviced is greater than or equal to the rate at which they are generated by the station. This is sometimes called server utilization or traffic intensity and should always be less than 1 for a stable system. In this simulation environment an unstable system is made stable by increasing the channel speed within a protocol.
2. Validate the simulator with known theoretical models. As previously mentioned, theoretical models exist for some protocols while others involve complexities that are difficult to model mathematically. When possible we will attempt to run simulations which conform to the assumptions made in certain theoretical models in order to validate the simulation results.
3. Run simulation with burst settings of 1, 5 and 10. Once a stability limit is found we will run a number of simulation models at incremental channel speeds for each of the above burst settings. The intention is to track the performance of the protocol for varying message sizes.

4. Present results in a manner that best emphasizes the performance characteristic being compared. In most cases this will involve two or three dimensional plots. When necessary, actual measurement results will be listed in table fashion for comparison purposes with theoretical values.

CHAPTER 5

SPREAD SPECTRUM

Overview

A satellite communication system, which is inherently power limited, employing spread spectrum techniques in order to trade-off bandwidth for small ground station antennas. In a one-way system embodiment a central station transmits data to a satellite for relay to a large number of small antenna receiving stations, the transmissions being spread spectrum encoded with spreading code lengths selected to provide adequate data recovery at the least sensitive station to which the transmissions are directed. Spreading codes may also function to address particular stations. In a two-way system embodiment, the central station additionally functions as a terrestrial relay station. A plurality of small antenna transmitting stations, at least one of which may be at the same site as a receiving station, transmit code division multiplexed data via the satellite to the central relay station using sufficiently long and distinct spreading codes as to permit adequate data error rates and to distinguish the transmissions of the various stations. The central relay station reformats the received data for retransmission to the satellite for relay to the receiving stations

5.1 Spread Spectrum:

Spread spectrum is a modulation technique originally developed for military applications as a secure means of communication unencumbered by natural noise and other signal interference.

The term spread spectrum describes a modulation technique which makes the sacrifice of bandwidth in order to gain signal-to-noise performance. Basically, the SS system is a system in which the transmitted signal is spread over a frequency much wider than the minimum bandwidth required to send the signal. The fundamental premise is that, in channels with narrowband noise, increasing the transmitted signal bandwidth results in an increased probability that the received information will be correct. If total signal power is interpreted as the area under the spectral density curve then signals with equivalent total power may have either a large signal power concentrated in a small bandwidth or a small signal power spread over a large bandwidth.

The spread spectrum modulation signal is uniformly spread over a wide range of frequencies for transmission, allowing the signal to avoid or eliminate interference and noise from other signals. The two most popular techniques utilized for encoding spread spectrum signals are direct sequence and frequency hopping.

5.2 Spread Spectrum Techniques

A method and apparatus employing spread spectrum techniques in a wide bandwidth communications system is disclosed. A plurality of transmitting stations are each equipped to provide a transmission signal representing a pseudo-random coded, phase modulated, message signal.

The transmission signal is directed through a bandwidth which encompasses otherwise dedicated, relatively narrow bandwidth repeater channels, employed in connection with a communications satellite, to a generally fixed receiver station. At the receiving station, the incoming signal is:

(a) code acquired and tracked,
(b) carrier acquired and tracked,
(c) phase locked to the receiver local oscillator and
(d) coherently demodulated to extract the desired data.

The receiving station advantageously employs plural receiving elements each having a pseudo-random sequence code matched filter which significantly reduces code acquisition time by obviating the necessity of exhaustively correlating the incoming signal with a replica of the pseudo-random code word at the receiver station.

A spread spectrum link in which a transmitter section generates a different frequency signal for each unique data symbol accepted from the data source and phase shift modulates the frequency signals with a sequence of spread spectrum symbols, the frequency differences used being such that orthogonality is obtained over the receivers observation time of each spread spectrum symbol

5.3 Spread Spectrum Encoding

Spread spectrum encoding is a technique for encoding data which makes maximal use of channel bandwidth, which allows for multiple signals to use the same channel without collision, and which is highly resistant to both interference and jamming. When combined with strong encryption and wireless transmission, this technology can be use to make secure, robust, wireless networks. This paper focuses on an elegant technique for spread spectrum encoding and decoding which requires a minimal amount of hardware.

5.4 Digital Spread Spectrum Encoding

In digital spread spectrum encoding, a stream of data bits is combined with deterministic noise to produce stream of encoded bits. These encoded bits have the desirable feature of appearing very noisy -- up to the repeat length of the noise source there is no apparent structure in the encoded data stream. This is desirable, because this stream of bits can then be transmitted in such a way as to completely utilize the bandwidth of the transmission channel.

For instance, Let us assume we want to send eight bit values, and have 256 discrete frequencies in a channel. A simple transmission scheme would be to represent each byte as a pulse on its corresponding frequency. If we were sending ASCII text repressing this document, we would find that some frequencies were used much more than others (in fact, we would be using only one half of your channel since ASCII values are always between zero and 127). This would be a poor use of your spectral real-estate.

If on the other hand we were to spread our data using an LFSR, we would find that all of our frequencies were being used equally. Thus, we would be making maximal use of the channel, and the power spectrum across the discrete frequencies would be flat.

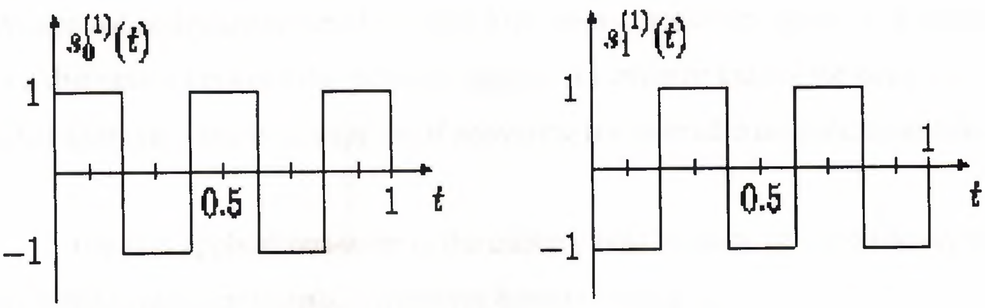
5.5 Spread Spectrum Multiple Access

In a spread spectrum communication system users employ signals which occupy a significantly larger bandwidth than the symbol rate. Such a signalling scheme provides some advantages which are primarily of interest in secure communication systems, e.g., low probability of intercept or robustness to jamming. In this problem we explore the inherent multiple access capability of spread spectrum signalling, i.e., the ability to support simultaneous transmissions in the same frequency band.

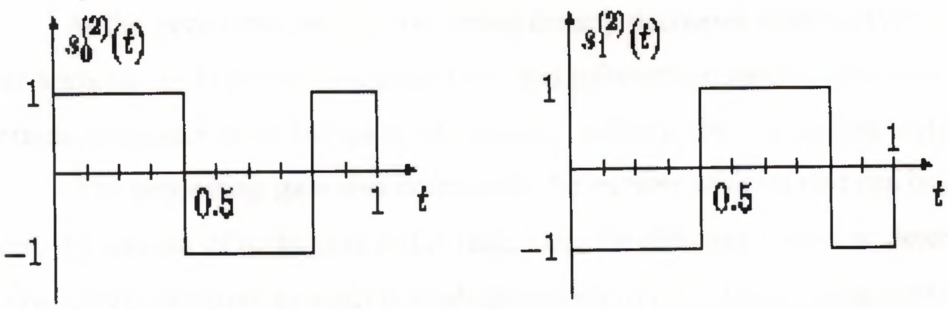
In the sequel, assume that the communication channel is an additive white Gaussian noise channel with spectral height .

$$\frac{N_0}{2}$$

1. One user employs the following signal set to transmit equally likely binary symbols



2. Now, a second users transmits one of the following signals with equal probability



Both signals are transmitted simultaneously, such that the received signal is given by

$$R_l = A_1 s_i^{(1)}(t) + A_2 s_j^{(2)}(t) + N_l,$$

where N1 is the noise process and i,j {0,1}indicate which symbol each of the users is transmitting. We are interested in receiving the first user's signal in the presence of the second (interfering) user.

5.6 CDMA Techniques

Code Division Multiple Access) is used in spread spectrum systems to enable multiple-access. It is a transmission technique in which the frequency spectrum of a data-signal is spread using a code uncorrelated with that signal and unique to every addressee. As the applied codes are selected for their low cross-correlation values, it is possible to make a distinction between the different signals. An initiator knows the code of the intended addressee and is so capable of activating the desired communication link.

The first applications were in the military field because of the difficulty to jam or detect spread spectrum signals. Nowadays however spread spectrum systems are gaining popularity also in commercial applications.

If a signal is combined with a code the bandwidth of the original signal increases. The spectrum is ``spread" which justifies the name ``spread spectrum".

At the same time the spectral power density decreases as the total transmitted power stays equal. The ratio of transmission and information bandwidth is therefore an important parameter in spread spectrum systems. which is the ``spreading factor".

The processing gain also determines the number of users that can be allowed in a system, the amount of multi-path effect reduction, the difficulty to jam or detect a signal etc. For spread spectrum systems it is advantageous to have a processing gain as high as possible.

Different spread spectrum techniques exist: Direct-Sequence (*ds*), Frequency-Hopping (*fh*), Time-Hopping (*th*) and Multi-Carrier CDMA. It is also possible to make use of combinations.

5.6.1 Direct Sequence

Direct Sequence is the most popular Spread Spectrum Technique. The data signal is multiplied with a pseudo random bit sequence, often referred to as pseudo random noise code (*pn-code*).

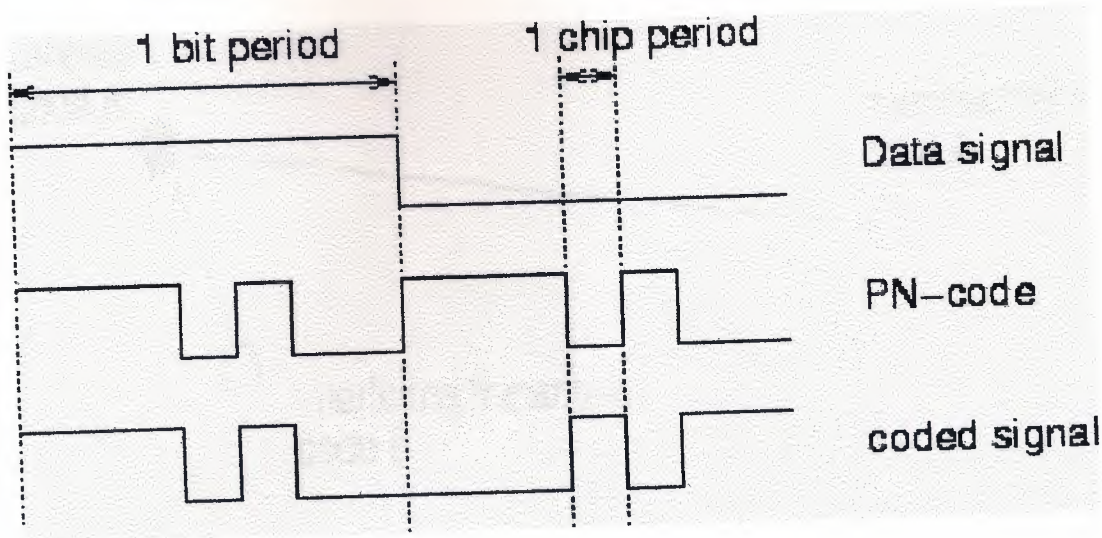
A *pn-code*is a sequence existing of chips, valued -1 and 1 (polar) or 0 and 1 (non-polar). Such bit-sequences have noise-like properties like spectral flatness and low cross and auto correlation values, and thus complicate jamming or detection by non-target receivers

Several families of binary *pn-codes* exist: *m*-sequences, Gold-codes and Kasami-codes where the latter two can be created by combining a number of selected *m*-sequences. An usual way to create a *pn-code*is by means of shift-registers with feed-back taps. By putting the feed-back taps at specific positions, the output sequence of a shift register is of ``maximum length". The above mentioned code-families have this property. When the length of a shift-register is *n*, the length of the resulting sequences is

$$N_{DS} = 2^n - 1.$$

In direct-sequence systems the length of the code is equal to the spreading-factor, so:

$$G_p(DS) = N_{DS}.$$

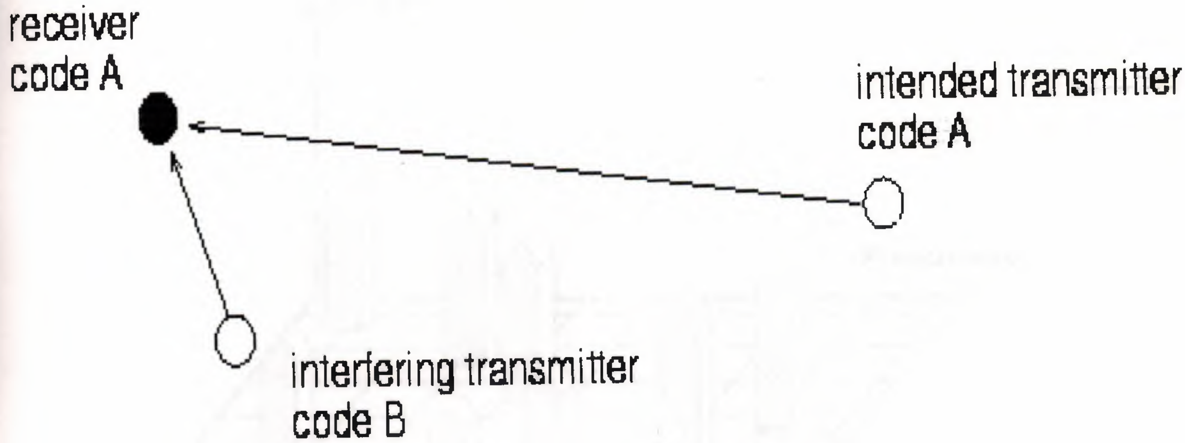


Direct-Sequence spreading

The generation of *pn-codes* is relatively easy. A number of shift-registers with feed-back taps is all that is required. For this reason it is easy to obtain a large processing-gain in Direct-Sequence systems.

In the receiver, the received signal is multiplied again with the same (synchronized) *pn-code*. Since a code exists of +1s and -1s, this operation completely removes the code from the signal and the original data-signal is left. Another observation is that the despread operation is the same as the spread operation. The consequence is that a possible jamming or interference signal in the radio channel will be spread before data-detection is performed. In this way jamming effects are reduced

A large problem with multi-access direct sequence spreading is the so-called near-far effect which is illustrated in figure below. This effect is present when a CDMA interfering transmitter is much closer to the receiver than the intended transmitter. Although the cross-correlation of ``code A" and ``code B" is low, the correlation of the received signal from the interfering transmitter with ``code A" in the receiver can exceed the correlation of the received signal from the intended transmitter and the correct code. As a result proper data detection is hardly possible.



Near-Far effect illustrated

5.6.2 Frequency Hopping

When applying frequency hopping, the carrier frequency is ``hopping" according to a unique sequence (an *fh-sequence* of length **NFH**). In this way the bandwidth is increased by a factor **NFH** (if the channels are non-overlapping):

$$G_p(FH) = N_{FH}.$$

The process of frequency hopping is illustrated in figure below. A disadvantage of frequency-hopping compared to direct-sequence is that it is hard to obtain a high processing gain. A frequency synthesizer is required that is capable of rapidly hopping over a set of carrier (*fh*) frequencies. The more *fh*-frequencies, the higher the processing gain and the more demanding the frequency synthesizer becomes.

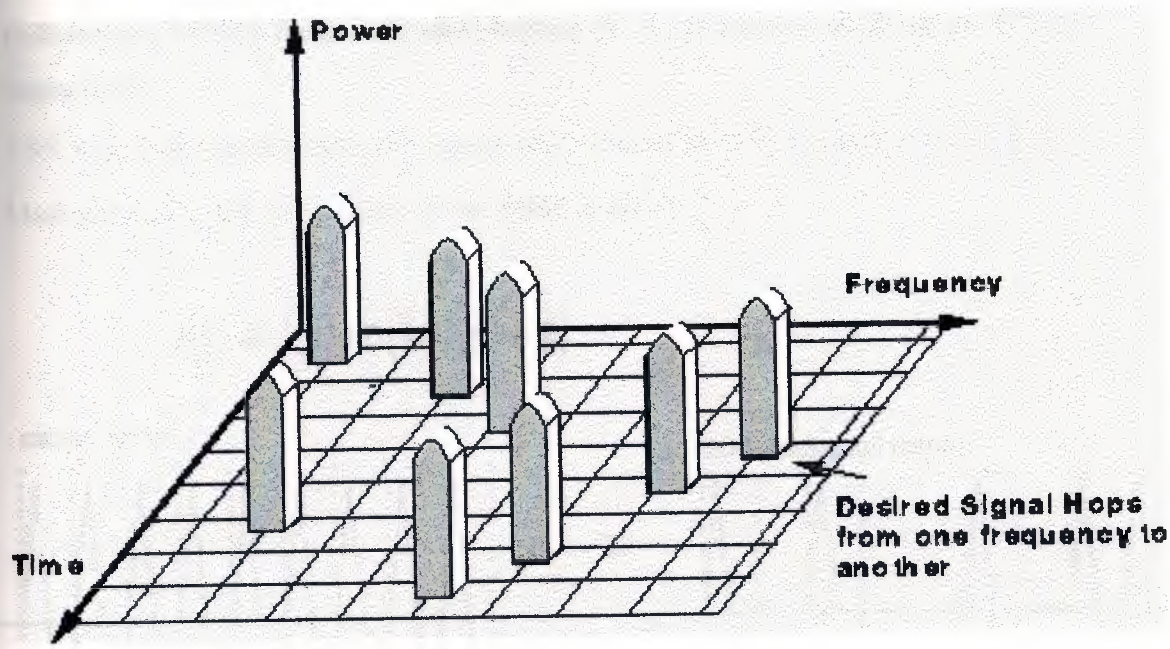


Illustration of the frequency hopping concept

On the contrary, frequency-hopping is less vulnerable to the near-far effect than direct-sequence. Frequency-hopping sequences have only a limited number of ``hits" with each other. This means that if a near-interferer is present, not the whole signal is blocked but only a limited number of ``frequency-hops". From the ``hops" that are not blocked it should be possible to recover the original data-message, for instance by applying error correcting techniques.

Two types of frequency-hopping techniques can be distinguished. In ``fast frequency hopping" the period of a ``frequency-hop" is smaller than a data symbol-period while in ``slow frequency hopping" the period of a ``frequency-hop" is larger than a data symbol-period. Choosing one of those techniques has consequences on the error correcting coding to be applied

5.7 Digital modulation, ASK, FSK and PSK

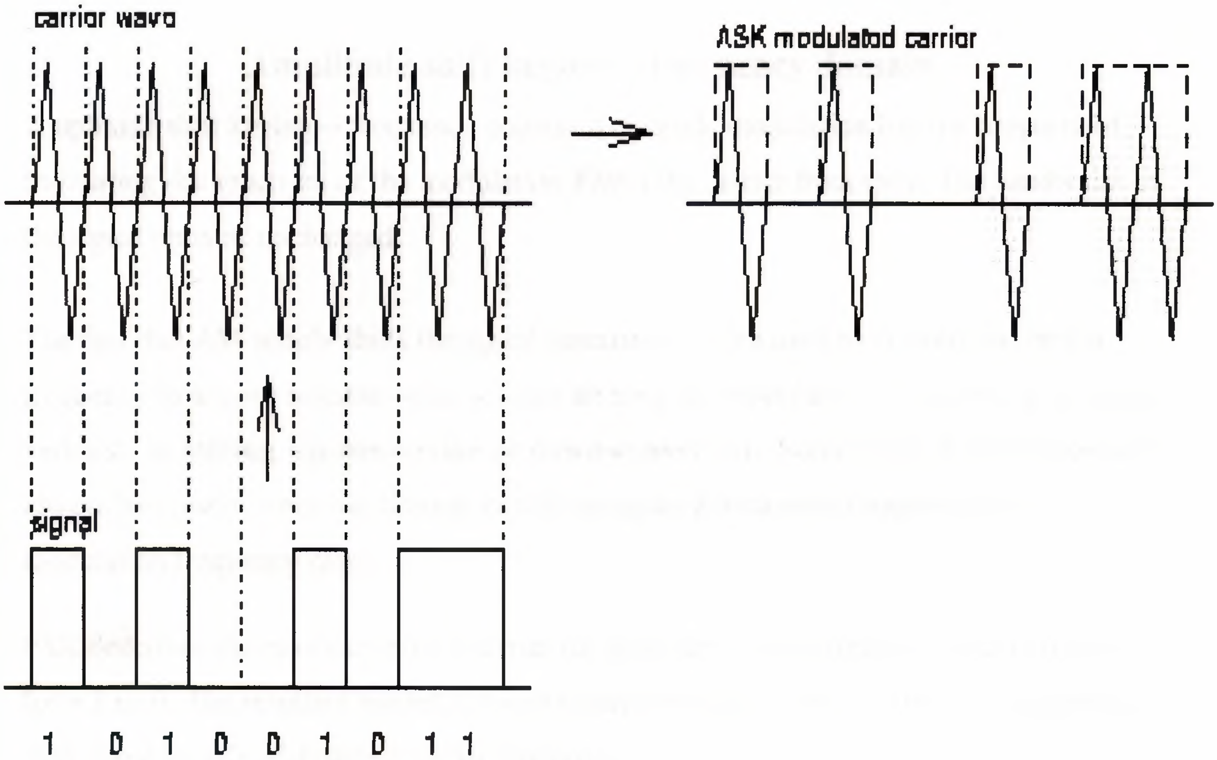
There are three ways in which the bandwidth of the channel carrier may be altered simply. It is worth emphasising that these methods are chosen because they are practically simple, not because they are theoretically desirable. These are the altering of the amplitude, frequency and phase of the carrier sine wave. These techniques give rise to **amplitude-**

shift-keying (ASK), **frequency-shift-keying (FSK)** and **phase-shift-keying (PSK)**, respectively.

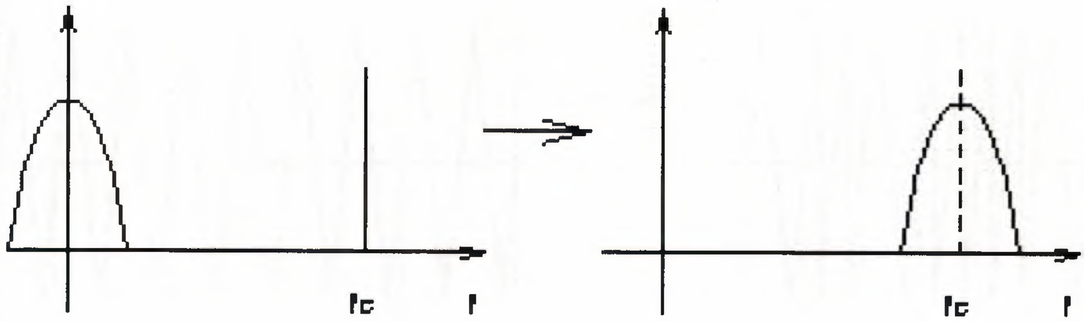
ASK describes the technique the carrier wave is multiplied by the digital signal ***F(t)*** .

Mathematically, the modulated carrier signal is ***s(t)***:

$$s(t) = f(t)\sin(2\pi f_C t + \phi)$$



Amplitude shift keying



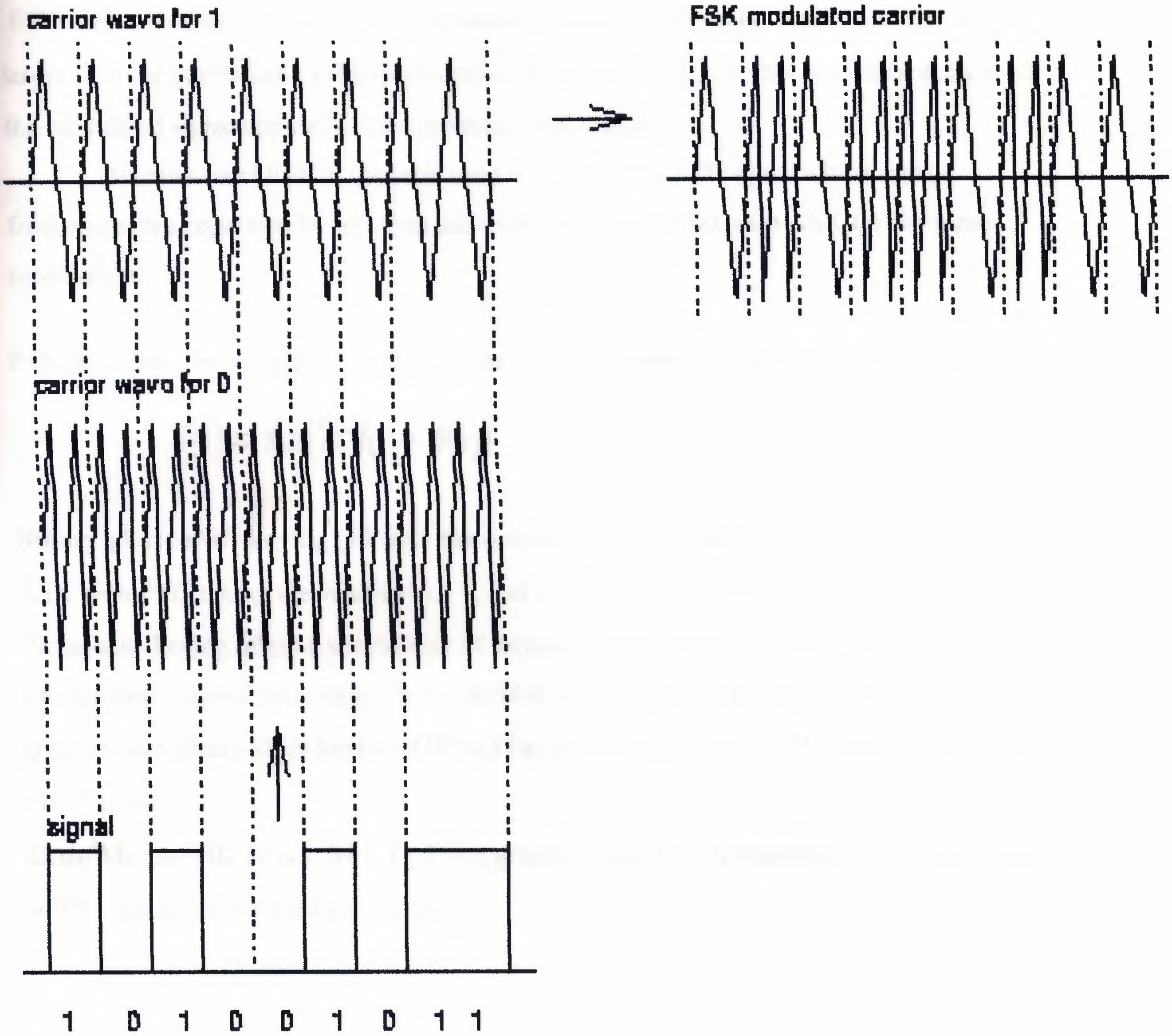
Amplitude shift keying -- frequency domain

Amplitude shift keying -- frequency domain Amplitude modulation has the property of translating the spectrum of the modulation ***F(t)***to the carrier frequency. The bandwidth of the signal remains unchanged

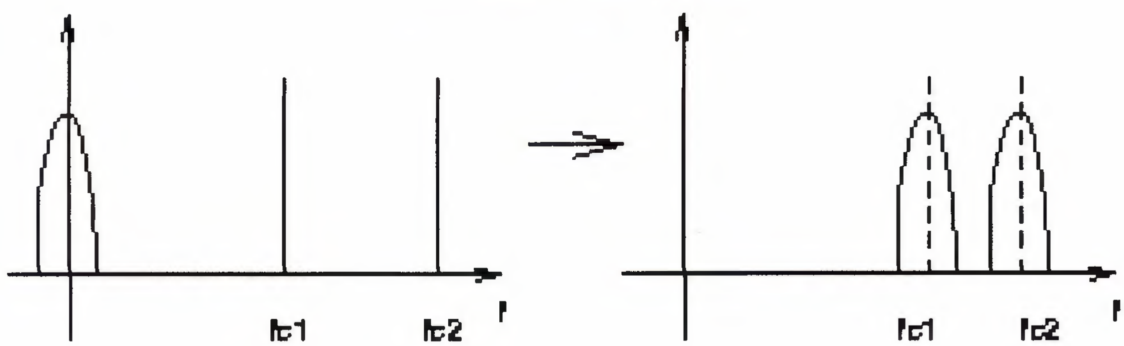
The fact that AM simply shifts the signal spectrum is often used to convert the carrier frequency to a more suitable value without altering the modulation. This process is known variously as **mixing, up-conversion** or **down-conversion**. Some form of conversion will always be present when the channel carrier occupies a frequency range outside the modulation frequency range.

FSK describes the modulation of a carrier (or two carriers) by using a different frequency for a 1 or 0. The resultant modulated signal may be regarded as the sum of two amplitude modulated signals of different carrier frequency

$$s(t) = f_1(t)\sin(2\pi f_{C1}t + \phi) + f_2(t)\sin(2\pi f_{C2}t + \phi)$$



Frequency shift keying



Frequency shift keying -- frequency domain

FSK is classified as wide-band if the separation between the two carrier frequencies is larger than the bandwidth of the spectrums of ***F1(t)*** and ***F2(t)***. In this case the spectrum of the modulated signal appears as two separate ASK signals

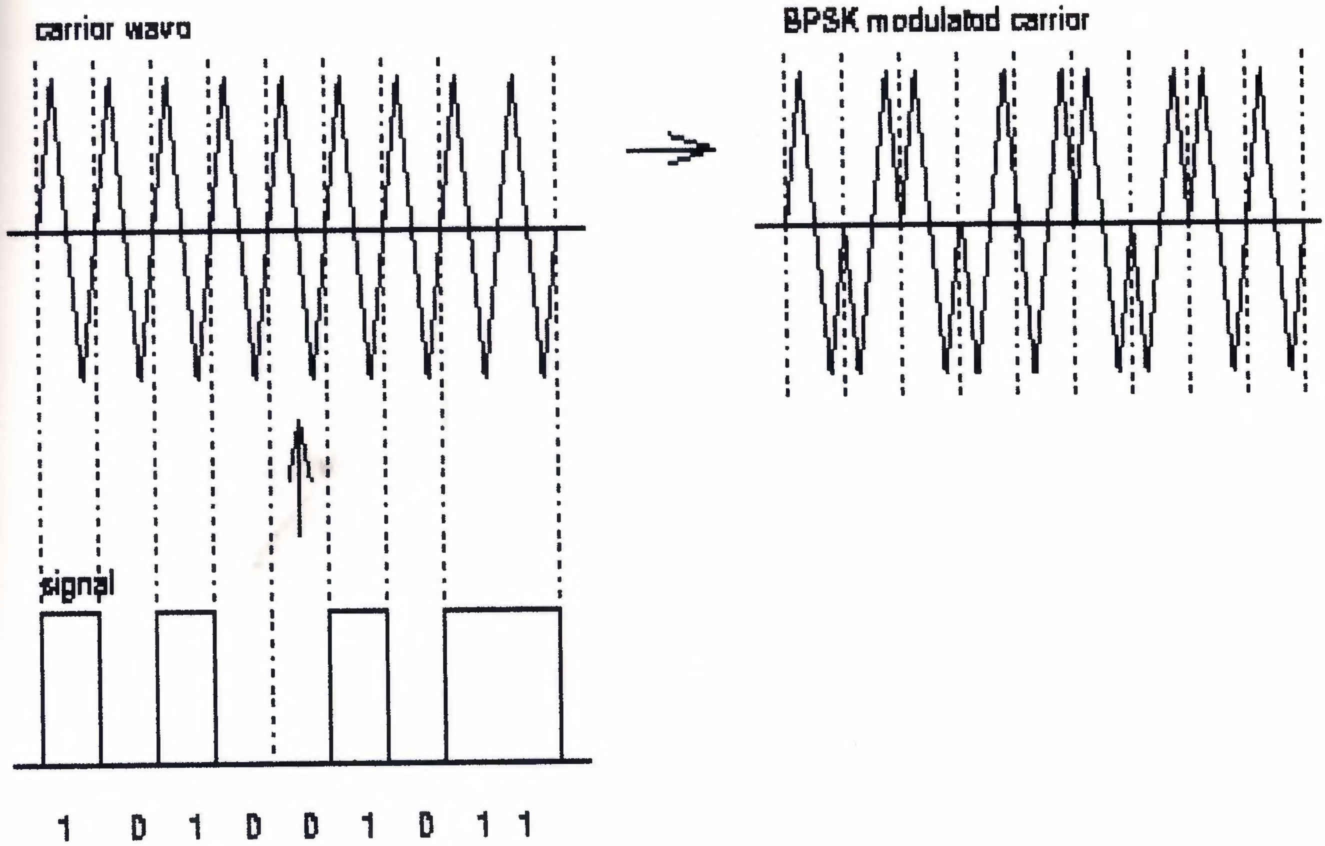
Narrow-band FSK is the term used to describe an FSK signal whose carrier frequencies are separated by less than the width of the spectrum than ASK for the same modulation.

PSK describes the modulation technique that alters the phase of the carrier. Mathematically

$$s(t) = \sin(2\pi f_c + \phi(t))$$

Binary phase-shift-keying (BPSK) has only two phases, 0 and π . It is therefore a type of ASK with ***F(t)*** taking the values -1 or 1, and its bandwidth is the same as that of ASK. Phase-shift-keying offers a simple way of increasing the number of levels in the transmission without increasing the bandwidth by introducing smaller phase shifts. **Quadrature phase-shift-keying (QPSK)** has four phases, **0, π/2, π, 3π/2** **M-ary PSK** has M phases,

2πm/M; m= 0, 1, ... M - 1 For a given bit-rate, QPSK requires half the bandwidth of PSK and Is widely used for this reason.



Binary phase shift keying

The number of times the signal parameter (amplitude, frequency, phase) is changed per second is called the **signaling rate**. It is measured in **baud**. 1 baud = 1 change per second. With binary modulations such as ASK, FSK and BPSK, the signaling rate equals the bit-rate. With QPSK and M-ary PSK, the bit-rate may exceed the baud rate.

CHAPTER 6

TRELLIS CODED MODULATION

Overview

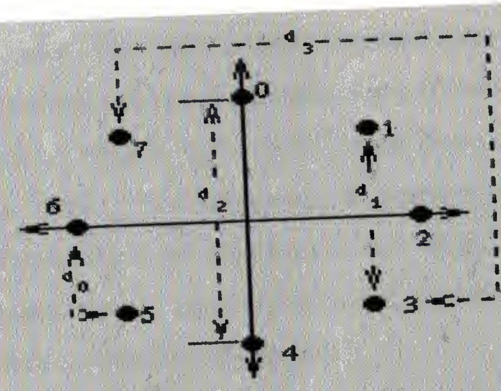
The treatment of coding and modulation as inseparable parts of a single system was first suggested by Massey 1974. Two years later Ungerboeck presented a feasible means for integrating coding and modulation through what is now called "trellis coded modulation" (TCM), At that point TCM began its phenomenally rapid move from theoretical concept to worldwide application.

TCM uses convolutional codes and multidimentional signal constellations to provide reliable, high-data-rate communication over bandwidth-limited channels. The most immediate application for TCM was in the area of digital data transmission over standard telephone lines. The telephone channel has a bandwidth of roughly 2700 Hz. In the early 1960s, many engineers felt that 2400 bps was the maximum data rate that would ever be supported on this channel. By 1970 the estimated ceiling on data rate had been raised to 9600 bps, with many exceedingly fine engineers believing that it would never go any higher [For6]. But of course, TCM has demolished this ceiling as surely as the 2400-bps ceiling was demolished.

The 19.2-Kbps modem is already a reality, and the corresponding standards are not far behind. We should also note that TCM has had a substantial impact on other bandwidth-limited applications, particularly in cellular mobile ratio and satellite communications.

6.1 Trellis Description

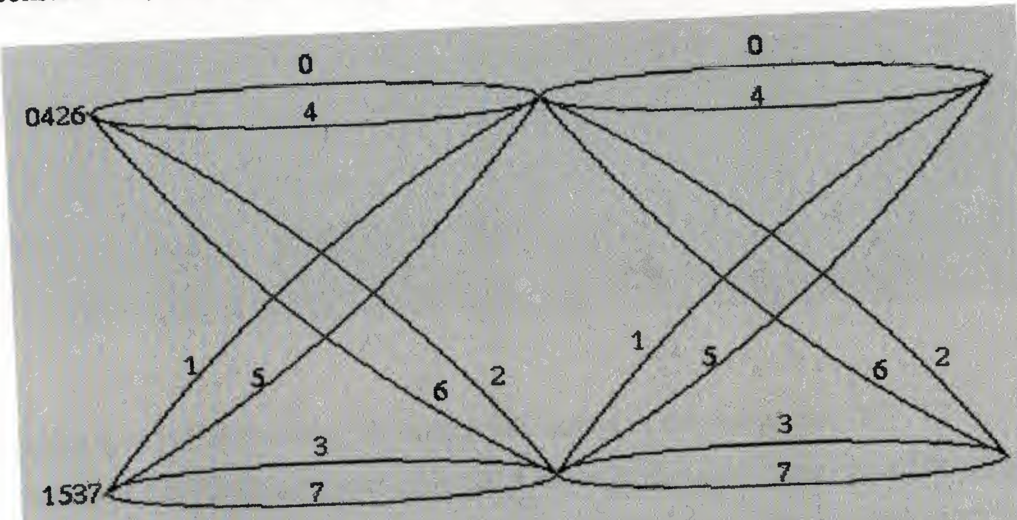
Trellis Coding provides controlled redudancy, by doubling the number of signal points. In addition, Trellis coding defines the way in which signal transitions are allowed to occur. (Signal transitions that do not follow this scheme will be detected as errors). This is best explained using the Trellis Coded 8-PSK



PSK signal constellation with M=8
(used with two state Trellis)
Figure 7

Without coding, the performance in 8-PSK depends on d_0 (d_0=2 sin (pi/8)=0.765), which corresponds to a higher bit error rate than QPSK (recall that d_1=1.414). By using Trellis Coding, it is possible to *improve* the performance by restricting the way in which signals are allowed to transition.

First the states of the trellis are defined. Lets label, one state as "0426", and the other state as "1537". Each digit refers to one of four permitted signal points in the state (hereafter refered to as "state points"), with each state by itself represting a QPSK constellation, with each state's constellation being offset by 45 degrees from the other.



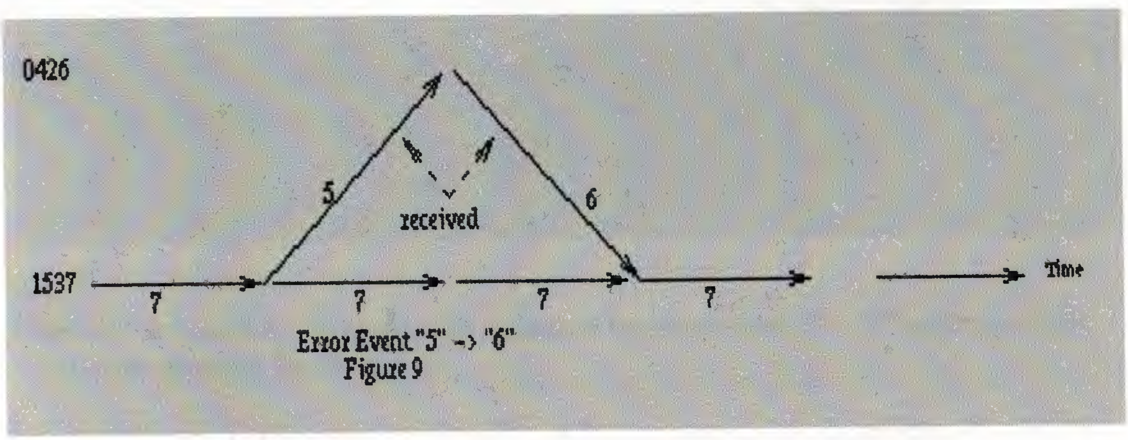
Two-state trellis, 8-PSK system
Figure 8

If the system is in state "0426" only one of these four state points is used. If a "0" or "4" is transmitted, the system remains in the same state. If however, a "2" or "6" is transmitted the system switches to the "1537" state. Now, only one of these four state points is used. If a "3" or "7" is transmitted, the system remains in this state, otherwise if a "1" or a "5" is transmitted it switches back to the "0426" state. Again, keep in mind that the system is each symbol represents two bits, so that when switching states, the "QPSK constellation is shifted by 45 degrees".

6.2 Error Analysis.

Assuming that all input signals are equaly likely, all signal paths are traced out over time. (For non-Trellis coding), the received signal includes noise and will tend to be located somewhere around the state points. The receiver again has to make a decision based on which signal point is closest and a mistaken output state value will be chosen if the receiver made an incorrect decision.

In order to illustrate error events, lets assume that the transmitter is sending continuous "7" symbols. The following graphs illustrate the possible error events:



Error Event "5" -> "6"
Figure 9

In this case "5" followed by "6" is received instead of the transmitted "7" - "7" sequence. The Euclidean mean-squared distance for this path is the sum of the squares of the distance of each interval (figure 7 for an illustration of the Euclidean distances and figure 8 for the Trellis Diagram):

$$\mathrm{sqrt}(d^2(7,5) + d^2(7,6)) = \mathrm{sqrt}(d_1^2 + d_0^2) = \mathrm{sqrt}(2 + (2 \sin(pi/8))^2) = 1.608$$

(where d(7,5) and d(7,6) are the Euclidean distances between these two the signals "7" and "5", and "7" and "6" respectively.)

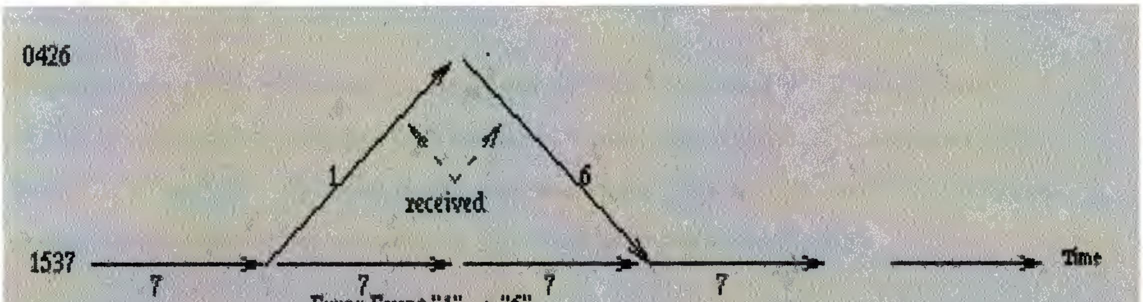


Figure 10

In this case "1" followed by "6" is received instead of the transmitted "7" - "7" sequence. The Euclidean distance for this path:

$$\mathbf{sqrt(d^2(7,1) + d^2(7,6)) = sqrt(d_1^2 + d_0^2) = sqrt(2 + (2\ sin(pi/8))^2) = 1.608}$$

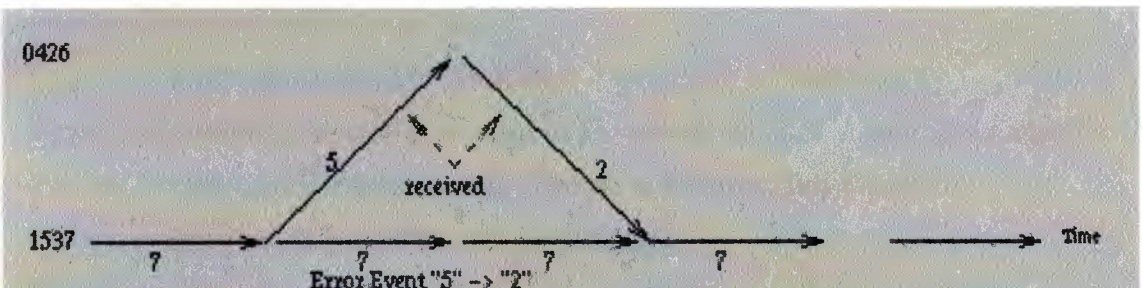


Figure 11

Here "5" followed by "2" is received instead of the transmitted "7" - "7" sequence. The Euclidean distance for this path is:

$$\mathbf{sqrt(d^2(7,5) + d^2(7,2)) = sqrt(d_1^2 + d_3^2) = sqrt(2 + (2\ sin(3\ pi/8))^2) = 2.33}$$

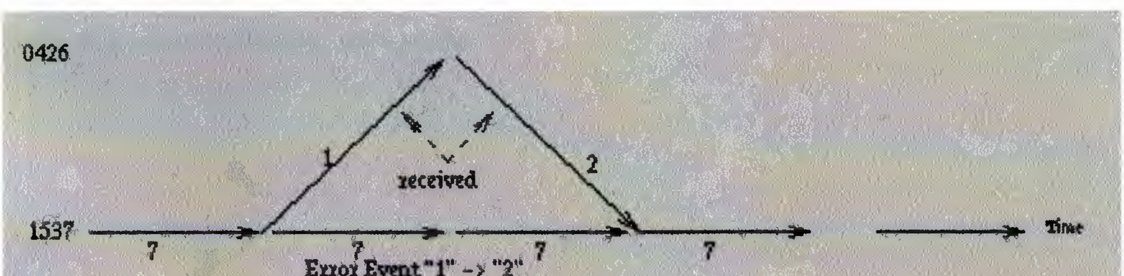


Figure 12

Here "1" followed by "2" is received instead of the transmitted "7" - "7" sequence. The Euclidean distance for this path is:

$$\mathbf{sqrt(d^2(7,1) + d^2(7,2)) = sqrt(d_1^2 + d_3^2) = sqrt(2 + (2\ sin(3\ pi/8))^2) = 2.33}$$

The only remaining error event is the single interval "3" instead of "7" error event, which has a Euclidean distance of 2 Because of their large Euclidean distances (2.33), the "5" - "2" and "1" - "2" error events are least likely. The "1" - "6" and "5" - "6" error events are most likely because of their low Euclidean distances (1.608).

6.3 Error Analysis

The minimum Euclidean distances for a trellis was defined by Ungerboeck as the minimum free Euclidean distance "d_E" (similar to the minimum free distance in convolusional coding). For the above example d_E=1.608.

Since d_E is a measure of the closest spacing between adjacent state points (and therefore also more likely to cause errors), it dictates the lower bound for probability of error for the entire Trellis in the following way:

$$\mathbf{P_e >= a*d_E*(1/2)\ erfc\ [(\ d_E * sqrt(E))/(\ 2 * sqrt(n_o))].}$$

Where, a(d_E) is the number of error paths at distance d_E. In the 2 state trellis example, there are 2 error paths at a distance of d_E. Therefore the probability of error is:

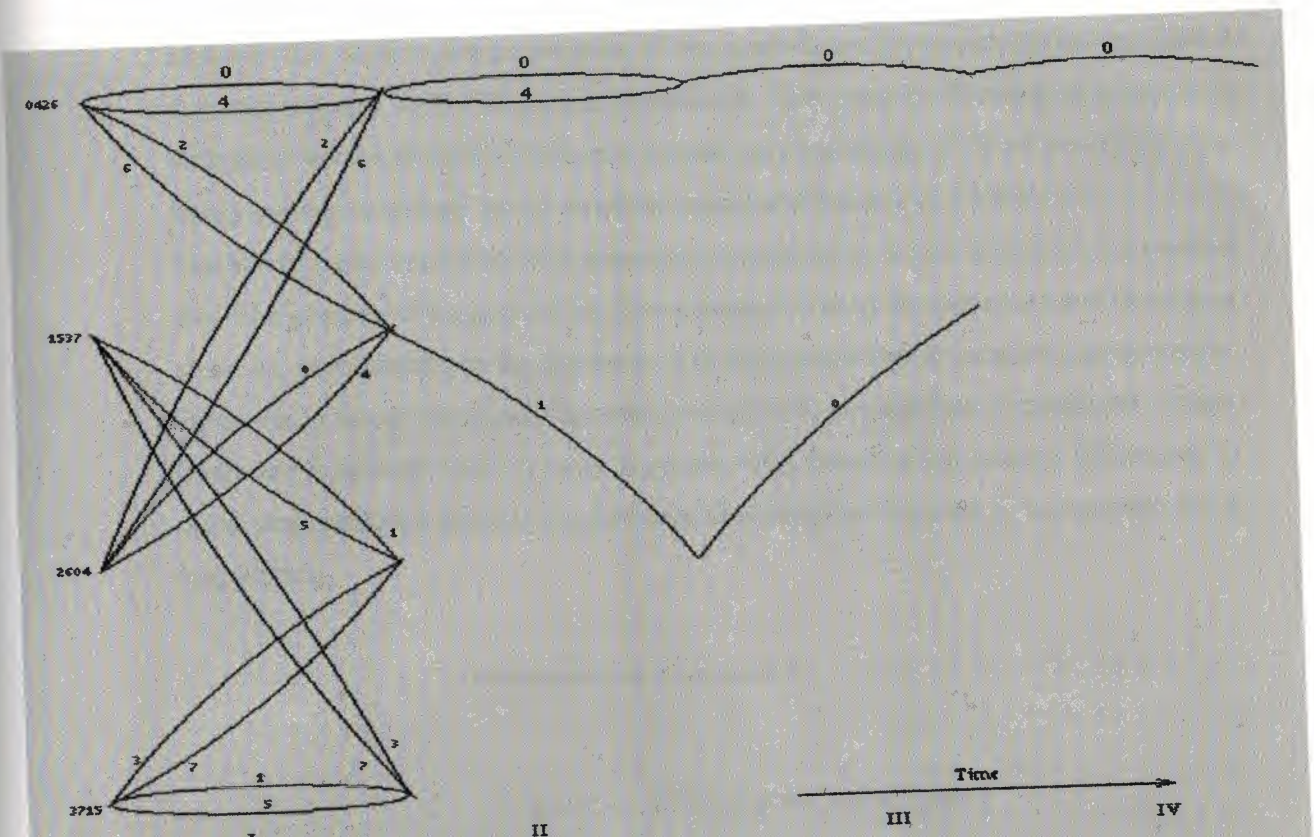
$$\mathbf{P_e >= erfc\ [(1.608 / 2)* sqrt(E/n_o)].}$$

The improvement of two state Trellis Coding over QPSK is therefore 1.608/1.414 or 1.1 dB

6.4 Improvements over 2 state Trellis Coding.

This is a very low coding gain for the amount of overhead required to handle Trellis coding. One might ask, is there a way to increase the coding gain obtainable with Trellis Coding?

There certainly is. To start out with, it is possible to increase the number of trellis states above 2, such as this four state trellis:



Four-state trellis, 8-PSK system
Figure 13

Describing error events for more than 2 states becomes overly complex. Suffice it to say, that by increasing the number of states, one also increases the Euclidean distances. For example, in the above four state trellis coding case, all error events have a Euclidean distance of more than 2 (single "4" error). This single " 4" error, is the only "lowest" error event with minimum Euclidean path distance d_E: Therefore, the lower bound for therefore 4 state Trellis Coding is:

$$P_e >= (1/2) \text{ erfc } [(2 / 2)* \text{sqrt}(E/n_o)].$$

6.5 M-ary Signaling

Digital systems are in the sense that they transmit information in the form of discrete symbols. In a binary system (e.g., binary phase-shift keying (BPSK) and binary frequency-shift keying (BFSK) each of the transmitted symbols is assigned one of two possible values. The Nyquist bandwidth for a communication signal is equal to the rate at which the symbols are transmitted. If the symbols are binary, an Rs bit-per-second data rate requires an Rs-Hz Nyquist bandwidth. The spectral efficiency of a system is defined to be the number of bits per second transmitted per 1 Hz of bandwidth. The binary system mentioned above thus has a spectral efficiency of 1 bit/sec/Hz.

The binary signal constellation is generally represented by a pair of points on a number line. The position of a signal on the line is proportional to its magnitude and its relative phase or, equivalently, to the square root of the transmitted or received signal energy. Figure (6-1) shows the signal diagram for the BPSK constellation. Normalized, integer labels are frequently used in such diagrams; they preserve the relative differences in signal magnitude but are less cumbersome than absolute received or transmitted signal magnitudes.

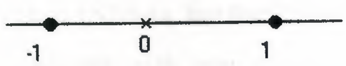


Figure 6-1. BPSK Signal Constellation

When we need to compute performance figures given some received noise energy, the square root of the average received signal energy. The symbol error rate is computed by determining the probability that the channel noise causes the received signal to be closer to one of the incorrect signals than to the signal that was actually transmitted.

For example, consider an additive white Gaussian noise channel with two-sided spectral density N0/2 at the input to the receiver. On such a channel the noise can be represented by a Gaussian random variable n that is added to the transmitted signal; if the signal x has been transmitted, the receiver will see a signal of the form r=x+n. If all signals are equally likely to be transmitted, the maximum likehood receiver selects the signal that is closest in Euclidean distance to the received signal.

In many applications the desired data rate (in bits per second) far exceeds the available bandwidth (in hertz). In such cases it is necessary to increase the spectral efficiency of the communication system. This is done by increasing the size of the signaling constellation; the symbol transmission rate (and thus the Nyquist bandwidth) remains the same, but the number of possible values taken on by each symbol is increased.

The most obvious expansion of the constellation in figure (6-1) is amplitude modulation (AM), as exemplified by the 8-AM constellation in figure (6-2). The 8-AM constellation provides 3 bits/sec/Hz (double-sideband AM), three times the spectral efficiency of the BPSK signal.

Figure 6-2. 8-AM Signal Constellation

At high signal-to-noise ratios (SNRs), the most likely error events are those in which the transmitted symbol is confused with one of its nearest neighbors. Let the minimum Euclidean d (min) for a constellation be the shortest distance between any pair of distinct signals. In figure (6-2) we have d (min)=2. If we assume that there are always two nearest neighbors, the symbol error rate for AM on an AWGN channel can be easily approximated.

The amount of energy required to transmit one of the 8-AM signals in figure (3-2) is proportional to the square of the distance of the signal from the origin. If the input as that in figure (6-1), then the average energy per transmitted symbol is significantly more for 8-AM than it is for BPSK. Let the normalized average received energy for the BPSK case be S1=1. The letter "S" denotes normalized average received energy, while the subscript denotes the log(to the base 2) of the cardinality of the signal constellation. 8-AM thus requires S3=21.0 times (13.22dB) more energy than BPSK to maintain the same minimum distance.

The AM constellations discussed above are one-dimensional. We can obtain better performance at the expense of a little additional complexity by moving to two-dimensional, or "quadrature," amplitude modulation (QAM). Figure (6-3) shows a set of commonly used rectangular QAM constellation, while figure (3-4) shows some M-ary phase-shift keying (MPSK) constellations. Both types have their respective advantages and disadvantages. The rectangular constellations provide better minimum-distance versus average-energy performance but are subject to distortion when they pass through nonlinear devices (e.g., traveling-wave tubes and other amplifiers that are operated in saturation). In MPSK constellations the minimum distances are relatively small, but the modulated signals have a constant envelope and are thus not distorted by channel nonlinearities.

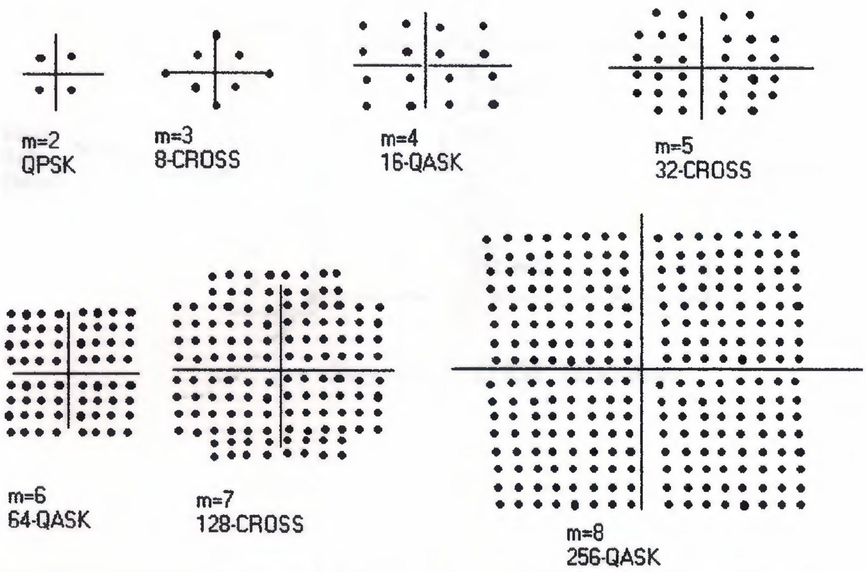


Figure 6-3. Several Rectangular Signal Constellations

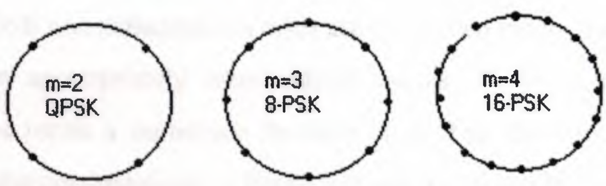


Figure 6-4. A Few MPSK Signal Constellations

Two-dimensional constellations can be implemented by modulating a pair of orthogonal carriers, as shown in figure (6-5) m bits are taken from the source and mapped onto an ordered pair (X,Y). These two are then used to modulate a pair of orthogonal carriers, which when added form a single complex signal z. The modulated sine and cosine are orthogonal and do not interfere with each other on a linear channel. Two-dimensional rectangular constellations can thus be viewed as a pair of orthogonal one-dimensional constellations.

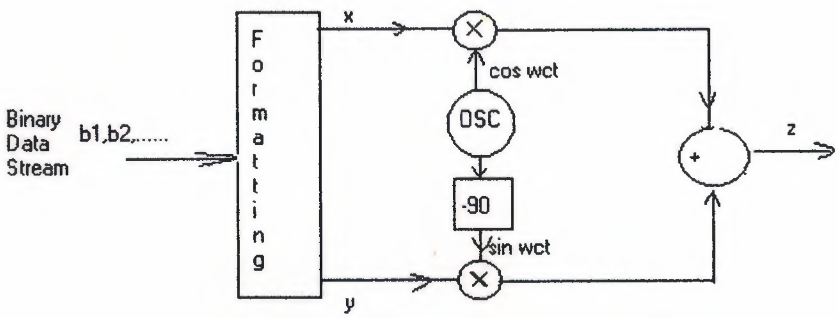


Figure 6-5. QAM Modulator

Now take a closer look at 64-QASK. Since there are 64 possible singles that can be sent in each transmission interval, each signal represents 6 bits, and the corresponding spectral density is 6 bites/sec/Hz. The constellation is perfectly square and can be viewed as a pair of orthogonal 8-AM constellations. We can thus also consider the constellation is achieving 3bits/sec/Hz/dimension.The average energy levels for the various CROSS constellations (m odd) are computed through more direct means but are found to be appropriately intermediate values. If we assume that all rectangular constellations retain a minimum distance of 2, then the normalized average energy S required by the constellations in figure 6-3 **are as shown in table (3-3).**

We can perform a similar analysis for the MPSK constellations. Assuming that signals have unity energy (i.e., the circle on which the signals rest has unit radius), If the average signal energy is adjusted to ensure that the minimum distance is always 2, then we obtain the **results in table (3-4).**

Clearly MPSK is a much less efficient means of obtaining higher spectral efficiency. Note, for example, that 64-PSK requires an additional 10 dB of energy to obtain the same dmin /average energy performance as rectangular 64-QASK.

CONCLUSION

As we have seen through this project the use of Digital satellite communication has became to be the most widely used field of communication through the world, the features of the of the satellite provides world wide communication at any time where ever the place, these features also provides high quality of the communication too.

As it becomes more and more common to the hole world the importance of the utilization of the satellite (digital satellite) in so many fields, it has become more important to the designers to make it possible for all the people to use it with the easiest way.

The use of digital satellite rather instead of analog satellite has provide high quality of communication with the least possible errors.

As the Digital satellite is become the most preferred method of providing world wide communication on all the levels in the community (with respect to the cost), digital satellite is made smaller, easier to lunch to the orbit, easier to control, easier to implement.

Also associated with a digital satellite communication are techniques such as demand assignment and digital speech interpolation to further increase the efficiency. With advanced satellite systems with onboard switching and processing, multiple spot beam, and beam hopping, a digital satellite can serve a mixture of large, medium, and small earth stations with high efficiency. Unlike an analog satellite system, a digital satellite system can employ error-correction coding to trade bandwidth for power.

REFERENCE

1. W.L. Pritched. (1977). Satellite Communication" –An overview of the history,"Proc IEEE.
2. W.L. Morgan. (1984). Satellite Location" Proc IEEE, Vol.72
3. Pascal Stephan C & Witrthers David W.(1997). Commercial Satellite Communications", Focal Press.
4. Richaria, M. (1995). Satellite Communication System, Design Principles", Mac Milan Pres LTD.
5. Meadows, R.G.& Parsons, A.J.(1989). "Satellite Communication", Hutchinson education,.
6. Balanis, Dennis. (1982). Antenna Theory, "Analysis and Design" John Wiley & Sons, Inc,.
7. Collins, Stephanie. (1999). Global Positioning System Overview", from internet and IEEE magasine,.
8. J.J. Spilker, Jr.(1997). Digital Communication by Satellite. Englewood Cliffs,NJ.:Prentice-Hall,.
9. S.J. Campanella and D. Schaefer. (1999). Time-Division Multiple-Access System (TDMA)," in K, feher (ed.) Digital Communication Satellite / Earth station Engineering. Englewood Cliffs, N.J.: Prentice-Hall.
10. Collins, Stephanie." INMARSAT System and Services, and INMARSAT's role in the GDMMS".