Satellite Radio Communications Fundamentals and Link Budgets

Daniel R. Glover

Contents

Abstract

Satellite communications makes use of radiofrequency links. Particular frequencies are allocated for satellite communications through international regulatory registration and coordination processes which prevents interference between systems. In typical operation, a satellite's transponder receives an uplinked signal from Earth, changes its frequency slightly to avoid self-interference, and retransmits it on a downlink to Earth. Antennas provide gain by focusing the transmitted energy. Path loss describes a natural spreading out of the transmitted wave front as it travels through space. A link budget is an accounting of gains and losses throughout a system that is used as a design tool to provide sufficient power (or gain) to allow a satellite connection to be established. The link margin is the

D.R. Glover (✉)
International Space University, Rocky River, OH, USA
e-mail: danglover@gmail.com

J.N. Pelton et al. (eds.), *Handbook of Satellite Applications*,
DOI 10.1007/978-3-319-23386-4_17

excess amount of received signal power above what is required. Shannon's law implies that there are trade-offs possible in a communications system design between power, bandwidth, and complexity.

Keywords
Bandwidth efficiency • Bit error rate • Coding decibel • Effective isotropic radiated power (EIRP) • Forward error correction • Free space loss • Gain line-of-sight • Link budget • Link margin • Modulation path loss • Satellite slot Shannon's law • Signal-to-noise ratio • Spectrum • Trade-off • Transponder • Wavelength

Introduction

This chapter provides an overview of radio communications as they apply to commercial satellite communications systems. This chapter is not intended to be a manual for designing satellite communications systems but rather provides the basic concepts needed to understand a system design and link budget. The general concepts are presented briefly and simply and may only hint at the actual complexity of a satellite communications system. As in any system design, satellite communications must trade-off several parameters where a choice for a given parameter results in constraining the choices for the other parameters. The major trade-offs are between power, bandwidth, and system complexity.

Basic System Concepts

One of two images may spring to mind when the term "satellite communications" is encountered. The image may be that of a "satellite dish" antenna on the ground. Or one might picture a spacecraft in space with its antennas and solar arrays deployed. These two images represent two major subdivisions of a satellite communications system: the flight segment (or satellite) and the ground segment (or the ground station with its associated antenna being the predominant feature). The flight segment is the part of the system that makes it a satellite communications system, but it is the ground segment that interfaces with users and with terrestrial communications systems. The ground segment is becoming a more familiar sight, with large dishes at "satellite farms" or gateways, medium-sized dishes atop retail stores, and small dishes for satellite TV becoming a ubiquitous sight. To provide useful communications capabilities, a system must have at least two ground stations (one providing a transmission and another receiving the transmission) and one satellite that relays the information between the two ground stations. In addition, there is a control segment (which is usually transparent to the user) that controls the spacecraft operations. Of course, a particular system may have many satellites and many ground stations (or user terminals), but for our purposes it is best to start with the simplest picture.

A satellite is an object (in our case, a spacecraft) that orbits around another object (in our case, the Earth). The communications spacecraft has several design constraints placed upon it (and thus, on the overall communications system) because it must be placed in orbit. Spacecraft designs are limited in their mass and volume in order to fit on the launch vehicle that places them into orbit. The mass and volume limits affect the size of the power system on the spacecraft, and so the power available is also constrained. In addition, the space environment (thermal, radiation, atomic oxygen, space debris, micrometeoroids, etc.) imposes constraints on the design (such as parts and material selection).

A spacecraft may be considered as consisting of two parts: the spacecraft "bus" and the payload. The spacecraft bus provides support services to the payload, while the payload provides the useful, or moneymaking, part of the satellite. Examples of payloads are scientific instruments, remote sensing instruments, navigation service transmitters, or (in our case) communications equipment. A satellite may have one type of payload or a combination of payload types.

The spacecraft bus provides services to the communications payload including power, structural support, attitude control and pointing, propulsion and station keeping, thermal control, commands, and telemetry. These services are provided by spacecraft bus subsystems which may vary slightly in name and content among different organizations or spacecraft manufacturers. The spacecraft bus typically has its own communications system separate from the communications payload which is used to control the spacecraft bus and the payload operations. The bus communications system is spacecraft part of the control segment. The spacecraft bus communications system receives commands from the control segment ground station and also transmits data concerning the state and status of the spacecraft bus and payload to the control segment ground station.

A typical communications payload consists of antennas and electronics designed for the reception and transmission of radio signals. Some processing of these signals is done on board which may be simple or complex depending on the system. In effect, the communications payload is a radio relay station in space with the satellite bus analogous to a tower.

Let us take a look at generic radio communications, then see where satellites fit in. Two radio stations are able to communicate when they are in "line-of-sight" or have no obstructions or barriers between them. Line-of-sight is a term that is used to describe the path between two stations along which an electromagnetic wave may travel without obstruction. As long as electromagnetic waves can propagate from one station's antenna to the other and can be distinguished from background noise or other interfering signals, communications is possible. Due to the curvature and topographic features of the Earth, as the separation distance on the ground between two locations increases, eventually the two stations will not be within line-of-sight and communications will not be possible. For two ground stations that are not within line-of-sight of each other, a relay station may be used to enable communications if the relay station is within line-of-sight of each of the two ground stations. The relay station receives a signal from one ground station and retransmits it to the other ground station. Additional relay stations can be used to extend the distance between the two ground stations.

A relay station might be placed on a tower to allow the line-of-sight to reach farther over the horizon. The higher the tower, the farther the horizon is extended (or the line-of-sight can reach) and so the farther apart the two ground stations can be and still communicate. For two ground stations on either side of the Atlantic Ocean, a relay tower would have to be over 600 km high (above sea level) in the middle of the ocean to have a line-of-sight to each shore. Since such a large tower greatly exceeds the height of the largest structure ever built, it makes sense to use a satellite as a platform for a radio relay station. Even a series of small relay towers across an ocean would be prohibitively expensive, although undersea cables (originally telegraph cables but now using optical fibers and repeaters) have been used since the nineteenth century and are still an effective means of bridging the oceans.

An example of a passive relay is a phenomenon that occurs at certain frequencies at night. Radio waves are affected by the ionosphere, an electrically conductive layer of the atmosphere that contains charged particles from ionization by solar radiation. The ionosphere is dynamic, as illustrated by the aurora, and changes from day to night. Certain frequencies of radio waves have their paths bent by the ionosphere to the point of being reflected when conditions are right at night and so can propagate much farther around the world than normal. Another example of a reflective relay was the experimental Echo satellite (1960) which was a metalized balloon 30 m in diameter placed in low Earth orbit. Other frequency bands can penetrate the ionosphere with minimal effect and can be used for communication with spacecraft.

One parameter that describes electromagnetic waves (such as radio waves) is the frequency. An alternate parameter is the inverse of the frequency, wavelength. The frequency (f) of a wave in a vacuum is related to its wavelength (λ) by the relationship $f\lambda = c$ where c is the speed of light. Communications engineers tend to use frequency except when a physical representation is useful, as in antenna design, where wavelength is typically employed.

So, a satellite in orbit can act as a relay by receiving signals from one ground station and transmitting them back to another ground station that may not be in the line-of-sight of the first ground station. A satellite in geosynchronous Earth orbit (GEO) can "see" a little less than half the Earth's surface, so one satellite is not enough to provide global coverage. The path from one ground station to a satellite then on to another ground station is known as a "hop." It may take more than one hop through more than one satellite to communicate from a ground station to another on the other side of the world.

Three satellites spaced 120° around the equator at GEO can provide almost total coverage of the globe (with the poles being just out of sight). This idea was first put forward by Arthur C. Clarke in 1945 in an article in the magazine *Wireless World* entitled "Extra Terrestrial Relays." An example implementation of this concept is the NASA Tracking and Data Relay Satellite System (TDRSS) which, among other uses, provides global coverage for International Space Station and Shuttle operations.

A ground station or terminal may be a building full of electronics with a large parabolic dish antenna or as simple as a mobile telephone handset or a satellite television receiver in a home. A ground terminal might transmit, receive, or both transmit and receive. A transmitting ground station requires attention to detail in the ground terminal setup and operation as improper transmissions can interfere with other systems. Terminals may be set up as nodes for point-to-point communications between two stations, broadcast, point-to-multipoint, or in a mesh. Small ground stations that are used in satellite communications networks were named VSATs (very small aperture terminals) when they were introduced. The name comes from their relatively small antennas, although more recent ground terminals (such as satellite television receivers) have even smaller antennas.

Two communications systems that try to use the same frequency at the same time in the same location may interfere with one another such that reliable communications is not possible. Sharing frequencies requires careful coordination to prevent interference. Purposeful interference is known as jamming.

The radiofrequency (RF) spectrum is a portion of the electromagnetic spectrum that is in great demand for uses including communications. Many different uses (applications) require some amount of spectrum in order to operate. To prevent interference and to allow efficient use of spectrum, national and international regulations have been put in place to coordinate the assignment of frequencies to applications. The international body facilitating this coordination is the International Telecommunications Union (ITU), a specialized agency of the United Nations (UN). National regulations control the use of spectrum within the borders of a country and supersede international regulations as long as other countries' spectrum use is not affected.

Figure 1 shows the frequency allocations for spectrum in the USA between 10 and 30 GHz. The purpose of this illustration is to indicate the incredible intricacy of the allocations within just a single country. Applications that use a portion of the RF spectrum include television broadcast, radio broadcast, mobile telephones, wireless telephones, consumer goods, microwave ovens, wireless networks, aeronautical and maritime communications, radio navigation, meteorological radar, radio astronomy, space research, microwave relays, multipoint distribution systems, fixed satellite, mobile satellite, direct broadcast television by satellite, and satellite radio broadcast. Frequency allocations are negotiated among the administrations of the world via a process established by the International Telecommunication Union (ITU) that is described separately in the chapter on this subject by Dr. Ram Jakhu. Satellite communications operators, in short, are in competition for spectrum with many other applications.

There are other types of communications besides commercial. The military may use commercial systems or their own satellite systems. There are search and rescue locating systems that use communications technology. Global navigation satellite systems are essentially highly specialized communications systems. There are satellite communications systems for collecting scientific data on Earth and deep space communications systems for scientific spacecraft exploring the solar system.

On a deep space mission, the communications system is somewhat different from that of a commercial spacecraft. There is typically a "low-gain" and a "high-gain" system named after the characteristics of their respective antennas. The low-gain

Fig. 1 US frequency allocations from 10 to 30 GHz (From US Department of Commerce chart "United States Frequency Allocations, The Radio Spectrum" available at http://www.ntia.doc.gov/osmhome/allochrt.pdf)

Table 1 Radar band designations for some frequencies used for satellite communications

Letter designation	Frequency band (approx.)
L	0.4–1.55 GHz
C	6/4 GHz
Ku	14/12 GHz
Ka	30/20 GHz

system performs the functions that the spacecraft bus communications subsystem does on a commercial satellite, namely, command and data handling, tracking functions for navigation, and housekeeping telemetry. The high-gain system provides a wide bandwidth "pipe" for sending back large amounts of scientific data (e.g., image data, video, radar data, etc.). A deep space probe is like a television station in space that needs to transmit a large amount of data back to Earth. The radio system is also sometimes used as a science instrument itself as is planned with the European Space Agency's Mars Radio Science Experiment (MaRS). MaRS will use the radio signals that convey data and instructions between the spacecraft and Earth to probe the planet's ionosphere, atmosphere, surface, and even the interior. The low-gain antenna also can serve as a backup to the high-gain antenna as with NASA's Galileo spacecraft when a failure in the high-gain antenna resulted in the science data being sent back by the low-gain system (allowing much less data to be returned than would have been possible with the high-gain system but much better than nothing).

Satellite communications engineers frequently make use of old radar band designations when discussing frequency bands. Table 1 lists a few common letter designations for bands useful for commercial communications. Another convention is to list the uplink frequency first, then the downlink frequency when a band is used for both uplink and downlink. Two different frequencies are used to keep the uplink and downlink from interfering with each other. Thus, in C band, a frequency band (of, say, 500 MHz) around 6 GHz is typically used for the uplink and 4 GHz for the downlink (Gagliardi 1984).

Transponders

A *transponder* is the electronic portion of the communications relay station that receives a signal, changes the frequency, and retransmits it. The word transponder is a contraction of the words transmitter and responder. Sometimes one might hear the word "repeater" used, but that is something of a misnomer in this case because a transponder intentionally changes the frequency of the received signal, while a repeater usually refers to electronics that receive and retransmit a signal without any intended changes to it other than amplification.

The transponder is the basic unit of the communications payload. Another way of thinking of a transponder is as a communications channel of the satellite. A communications payload may have from 1 to over 100 transponders, with a typical number being a couple of dozen. Each transponder operates in a portion of the total frequency band available for use by the satellite. The bandwidth of an individual transponder might be designed to be 20 MHz or as much as 500 MHz. For many years, the typical bandwidth of a transponder was set at 36 MHz; because of this history, one may still speak of 36 MHz equivalent transponders as a unit of comparison when discussing satellite capacity. The bandwidth required to transmit an analog, studio quality video signal was 36 MHz which was a major driver in satellite communications design in the early days of the business when analog television distribution was an important application. Today, transponders can be found with bandwidths of 26, 27, 36, 54, 72 MHz, or even larger bandwidths.

A transponder that receives and retransmits a signal without affecting it (other than amplifying it and changing its carrier frequency) is said to be operating in "bent-pipe mode." This term is an analogy that compares information flowing in a communications channel to fluid flowing in a pipe. A "bent pipe" merely changes the direction of the flow without performing any processes on the "fluid." This is not a perfect analogy (since we know the signal is amplified and the carrier frequency changed) but is good enough to be widely used. A satellite in "bent-pipe mode" does not process the signals in the channels but is only a conduit that provides a little boost (perhaps a better analogy would include a pump at the bend, but that might be making the analogy a little too clumsy to use and still does not incorporate an analogy for the frequency translation).

Originally, satellite transponders were analog devices. With the advent of digital signals, the use of error correction coding was introduced. A regenerative transponder demodulates the signal and provides error correction on board the satellite before modulating and transmitting back to the ground. This has the advantage of correcting errors introduced on the uplink, thus regenerating the original signals before exposing them to potential errors on the downlink.

In contrast to a bent-pipe satellite is the satellite that includes onboard processing of signals. An early example of the use of onboard processing in a commercial communications satellite is the Iridium system. Iridium uses a set of satellites (known as a "constellation") in low Earth orbit (LEO) to enable personal communications through satellite telephones. Telephone calls are switched on board the satellites and relayed through the constellation and the ground system. Iridium uses inter-satellite links between the satellites in the constellation to route an individual call to its destination.

A signal is a representation of information; that information could be a voice signal, video, data, or a combination. A channel (in our case, a particular transponder's bandwidth) is spaced around a carrier frequency and has a capacity to contain a signal or collection of signals with a certain amount of information within that channel's frequency bandwidth.

Antennas

An antenna radiates and/or captures radio frequency waves to and from free space. It is the interface that couples our electronics to free space and enables telecommunications without wires. Antenna design is a specialized topic within telecommunications engineering.

Different antennas have different patterns of radiation. An isotropic antenna radiates equally in all directions (a pattern described as "omnidirectional"). Other antennas concentrate their radiation in patterns known as "lobes" or "beams." A design may attempt to concentrate all of the radiated energy into one main beam, but there are typically many additional smaller beams, known as side lobes, radiating in unwanted directions. The main lobe of the pattern has a beamwidth that is usually measured as the angle between points on that pattern that are at half the power of the peak of the main lobe. To simplify matters, we will ignore the side lobes and concentrate on the main lobe of the pattern when we talk about the antenna beam.

The antenna pattern characterizes the radiated energy (transmission) of a given antenna as well as its ability to collect energy (receive signals). If you are familiar with optics, it may help to think of an antenna as the radio equivalent of a telescope at optical wavelengths. There are similarities such as diffraction limits, diameter versus field of view, etc., since light and radio waves are both electromagnetic waves (just at different wavelengths). Telecommunications engineers tend to use frequency except in a few instances such as antenna design. With antenna engineers, wavelength is a more pertinent view since they are dealing with physical dimensions of shapes and sizes.

There are many different types of antennas, but the one most commonly associated with satellite communications is the parabolic dish antenna. These dish antennas have a narrow beamwidth, concentrating the energy of the radiated main beam into a smaller solid angle. This means more of the radiated energy reaches, or "illuminates," the satellite when using a dish antenna as compared to an omnidirectional (or "omni" for short) antenna. A useful analogy is: a dish antenna is to an omni as a searchlight is to a bare light bulb. This advantage of the dish relative to the omni (that is due to focusing the energy) is known as "antenna gain." The antenna gain of an omni is 1 (or 0 dB), while the antenna gain of a dish antenna is greater than one and related to the diameter (or aperture) of the dish in wavelengths.

Antenna beams allow us to reuse spectrum without interfering with other satellite systems. By using a narrow enough beam, we can communicate with one satellite and not illuminate its neighbors with unwanted electromagnetic energy. The beamwidth of ground antenna systems is the key parameter in defining the separation of spacecraft placed in the geosynchronous Earth orbit (GEO) arc 36,000 km above the equator at sea level. The spacing supporting a single communicating spacecraft at a particular frequency is commonly referred to as a *satellite slot*.

A crude analogy may help to illustrate the concept: Imagine communicating by Aldis lamp (a flashing signal lamp used for naval communications in the twentieth century) using Morse code. In flashing the lamp, the sender illuminates an area (at a given distance) that is a cross section of the beamwidth of the lamp. That illuminated area hopefully includes the intended receiver but may include other unintended receivers. Now imagine being surrounded by a ring of signalers, only one of which is the intended communications partner (or node). Now there are two problems with communicating: first is illuminating only the desired node in the ring (corresponding to a satellite in a GEO slot) and the other problem is filtering out the signal flashes coming from non-desired nodes in the ring and only receiving flashes from the desired node. The first problem is easily solved by using a lamp with a narrow enough beam such that only the desired node is illuminated (perhaps using a flashing laser if necessary). The second, reception, problem can be solved by using a mask with an aperture that emphasizes the desired node (say by cutting a hole in a piece of cardboard that masks the neighbors of the desired node).

If you were being indiscriminant in your illumination and inadvertently included neighbor nodes that you were not trying to communicate with, you would be a bad neighbor. By illuminating non-desired nodes, you are introducing distracting flashes (or interfering noise) into your neighbor's systems. To be a good neighbor and avoid jamming the communications of others, it is required to use a narrow enough beam to illuminate the desired node and not its neighbors.

In a satellite system, the antenna beamwidth solves both problems by illuminating only the desired satellite in the main lobe and masking the neighbors (since they are not in the main lobe of the antenna pattern) in the uplink. In the downlink, the ground station receives strong signals from the desired satellite and only small amounts of energy (which count as noise for our system) from the neighbors because the desired satellite is located within the main lobe (beam) and the neighbors are not. The separation between adjacent satellites needs to be at least the same amount that corresponds to the beamwidth of the antennas that are practical at that frequency (say 3° for C band).

Polarization can be used to cram a few more channels into a given amount of bandwidth. There are two types of polarization: linear and circular. Linear polarization is the same effect that is used in sunglasses to reduce glare. A filter can be used to allow a horizontally polarized wave through while blocking a vertically polarized one. In communications, polarized antennas can be used to capture or reject polarized radiofrequency carriers.

There are two bands of frequencies in the electromagnetic spectrum that are said to be windows through the atmosphere because electromagnetic waves can pass through relatively unchanged. One window is the radio window used for satellite communications (the other is the visible window). At low frequencies, the ionosphere bends the paths of electromagnetic waves to the point of reflection, while at higher radio frequencies, the atmosphere absorbs the waves energy. Water vapor absorbs radio waves, and at high frequencies, such as at Ka band, clouds and rain can produce significant signal loss (see Fig. 2) known as rain fade.

An antenna located at the equator pointing straight up at a satellite in the GEO arc will have a shorter path and look through less atmosphere (optical depth) than an antenna located at a higher geographic latitude. This distance from the ground station to the satellite is the slant range. Another parameter of ground station antenna pointing is the look angle (azimuth and elevation) by which the ground station antenna must be pointed to view a particular satellite. An antenna pointed at a GEO satellite will also be pointed at the sun for a short period of time around the equinox twice a year. The sun is a significant noise source at all frequencies and will cause interference (known as a sun outage) at those two short periods (typically 15 min) each year.

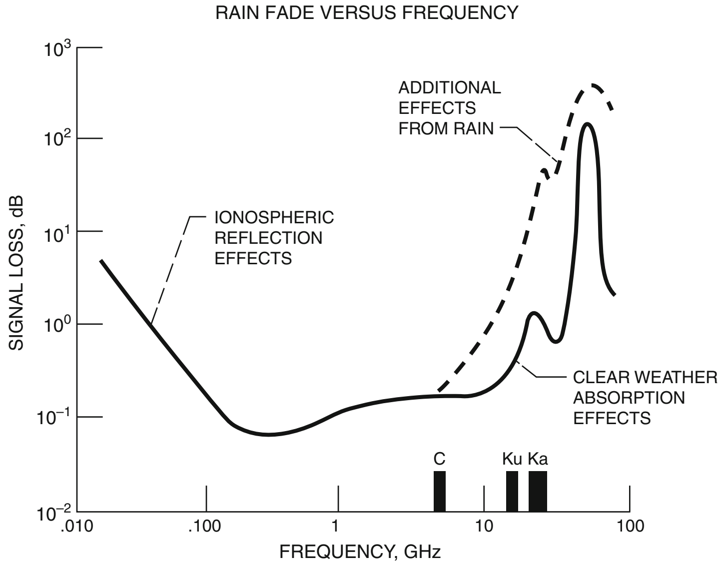

Fig. 2 Rain fade versus frequency (Courtesy NASA)

Digital Communications

Initially, satellite systems started out as analog systems and over the years have increasingly become digital systems. Digital communications allows flexibility in design through the use of compression and error correction coding. Digital signals can be perfectly reconstructed when a system is designed properly, while analog signals tend to become more corrupted in each stage of processing. Converting an analog signal to a digital signal does introduce quantization noise, but the amount of noise added can be selected (by finer quantization and sampling) and is only added once (at the time of conversion).

A signal can be converted from analog to digital through a process called pulse code modulation (PCM) that includes sampling (Jayant and Noll 1984). The analog signal is measured periodically at some sample rate. The samples are quantized to numerical values. The analog signal is reconstructed by essentially connecting the dots of the values at the sampling times. The reconstruction will be poor unless enough samples have been taken and the samples were quantized to a suitable range of values. The sample values can be represented by binary digital values (bits) and strung together into a bitstream. The bitstream (with values of 1 or 0) represents the signal.

The analog signal at its original frequency location and bandwidth, typically from 0 Hz up to some frequency that includes most of the signal energy, is the baseband signal. When digitized, the signal will typically take up much more bandwidth than the original analog signal, but that can be reduced through digital data compression. For example, a typical voice signal for telephony would have a baseband bandwidth of

4 kHz and would require 64 kbps in digital form after PCM at 8 bits/sample and 8 k samples/s. Using compression, the 64 kbps PCM data rate could be reduced to 10 kbps or even 4 kbps (with a quality vs. data rate trade-off). In the digital world, data rate is sometimes referred to as bandwidth and they are closely related. Approximate data rates for some types of signals range from 32 to 320 kbps for MP3 (Motion Picture Experts Group 1 Layer III) compressed audio, 1,411 kbps for CD audio, and 2–40 Mbps for MPEG-2 compressed video (typically ~5 Mbps for standard video resolutions, ~19 Mbps for high-definition TV, ~36 Mbps (1x) for Blu-ray video).

Modulation and Coding

As presented in the previous chapter, modulation is the technique of imposing a signal (say, a baseband signal) onto a radiofrequency carrier (or carriers) for transmission by varying the carrier's amplitude, frequency, or phase in accordance with the signal. A few modulation schemes are demonstrated in the figures below. These highly simplified representations are not to scale; the carrier frequency is typically much higher than the baseband signal.

Figure 3 depicts analog frequency modulation (FM). The carrier wave's frequency is varied in accordance with the change in the information signal's amplitude. Figure 4 depicts digital amplitude-shift keying (ASK or on-off keying) modulation. In this case, a digital string of data is represented by a digital waveform. The carrier amplitude is varied according to the digital information bit stream, on for a 1 and off for a 0. The envelope for this modulated carrier is obviously not constant which would present a problem for most transponders. Figure 5 depicts the digital version of frequency modulation, frequency-shift keying (FSK). Figure 6 depicts

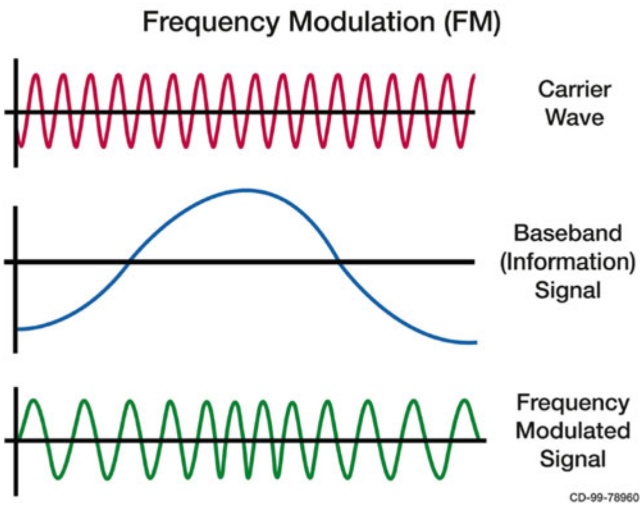

Fig. 3 A cartoon representation of analog frequency modulation, not to scale (Courtesy NASA)

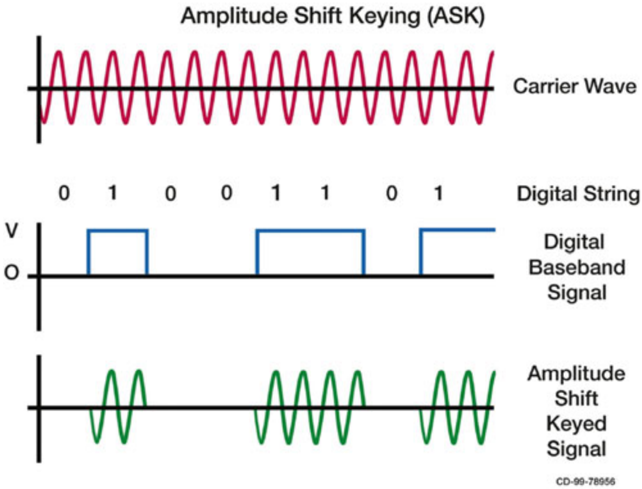

Fig. 4 A cartoon representation of digital amplitude-shift keying modulation, not to scale (Courtesy NASA)

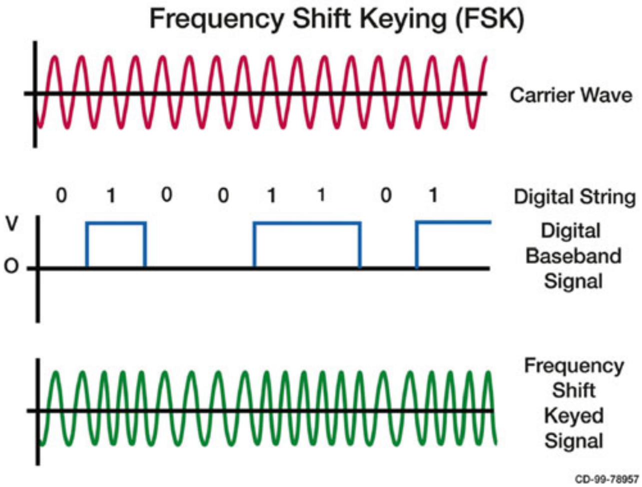

Fig. 5 A cartoon representation of digital frequency-shift keying modulation, not to scale (Courtesy NASA)

digital phase-shift keying (PSK). The digital waveform used is typically a non-return-to-zero (NRZ) representation where a bit value of 1 is represented by a positive value and a bit value of 0 is represented by a negative value, but here a bit value of 0 is represented as 0 for illustrative purposes only. There are many other methods of representing a bit stream as a waveform such as Manchester (or biphase) encoding.

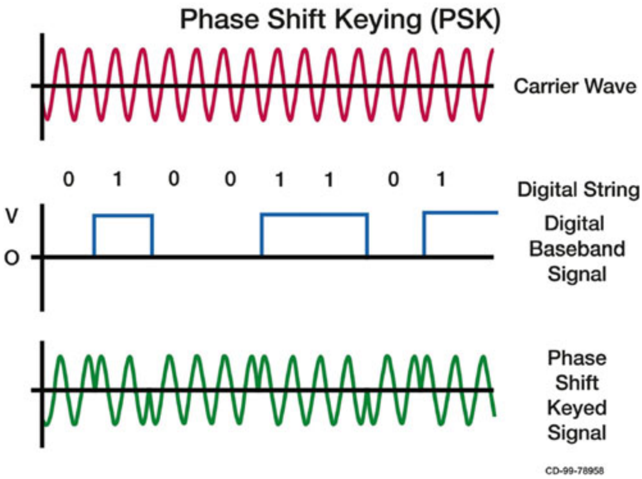

Fig. 6 A cartoon representation of digital phase-shift keying modulation, not to scale. V represents a positive voltage and O is usually a negative voltage but shown here as 0 for illustrative purposes (Courtesy NASA)

Table 2 Bandwidth efficiency of various digital modulation schemes

Modulation scheme	Theoretical bandwidth efficiency
BPSK	1 bit per second (bps)/Hz
QPSK	2 bps/Hz
8-PSK	3 bps/Hz
16-QAM	4 bps/Hz

A version of PSK, quadrature phase-shift keying (QPSK), is widely found in satellite systems. Biphase PSK (BPSK) is sometimes encountered. QPSK is better from the standpoint of bandwidth efficiency as it is equivalent to two independent (orthogonal) BPSK schemes combined. Although BPSK is a simpler method, QPSK chipsets have become widely available making it a cost-effective choice. Bandwidth efficiency refers to the amount of bandwidth required to accommodate a given amount of data. The more bandwidth efficient a scheme is, the more data can be accommodated in a given channel bandwidth, but generally this results in a more complicated scheme. In other words, there is usually a complexity versus bandwidth trade-off to be made in selecting a modulation scheme, among other considerations.

Advanced modulation schemes include 8-PSK, quadrature amplitude modulation (QAM) including 16-QAM, 64-QAM, and 256-QAM where the numbers refer to the number of symbol states in the scheme. Table 2 shows the theoretical bandwidth efficiency for various modulation schemes. A system's bandwidth efficiency is affected by more than just the modulation scheme and will be less than that shown in the table.

System bandwidth efficiency will be less than that shown in Table 7 due to things like pulse shaping (typically raised cosine filtering (Hayken 1983)) and coding. A

QPSK system might have a bandwidth efficiency of 1.5 bps/Hz or less. A 16-phase coding system using the latest technology, however, might use this level of complexity to achieve an efficiency of 5 bps/Hz or better.

An important measure of quality in a digital system is the bit error rate (BER). The BER is the number of bits in error divided by the total number of bits sent. Bit errors can occur randomly spaced or in blocks. The BER is usually assumed to be an average over some period of time. For example, if 1 bit is in error on average for every million bits sent, the BER would be $1/10^6$ or 10^{-6}. Error correction coding, also known as channel coding, can be used to correct some errors. Coding requires that additional bits are inserted which adds information redundancy to the bitstream at the cost of increasing the number of bits that must be sent.

In a simple example of error correction coding, each original data bit could be sent three times. The original data bit is combined with two parity bits that provide a 3-bit representation in the transmission data stream. One bit error can be detected and corrected since it will conflict with two unchanged bits. This triples the number of bits that must be sent compared with the original data, however. If the original bit was only sent once, then a bit error could be detected but not corrected since the receiver will not know which bit is the correct value. It turns out that there are more complicated codes that allow various levels of error detection and correction that use fewer than 3-bit symbols. The amount of redundancy that a particular code adds is described by the code rate where the rate is the number of original data bits divided by the total number of bits (data and parity) that are sent. For example, the rate of our inefficient tripling code above is 1/3.

There are several types of error correction coding including block codes, convolutional codes, concatenated codes, turbo codes, low-density parity-check (LDPC) codes, etc. (Peterson and Weldon 1991). They add varying amounts of complexity and provide varying levels of protection. Source coding refers to data compression coding. Channel coding adds redundancy to the bitstream while source coding removes redundancy. A bit error in a compressed data stream may cause damage to more than one data bit when uncompressed.

Coding gain is a measure of the power savings that a coding scheme achieves compared to an uncoded signal in the same system. The power savings come at the cost of increased bandwidth and complexity as coding adds redundancy to the original bitstream.

Multiple access (Bhargava et al. 1981) refers to the sharing of a communications channel by more than one signal. It is similar to multiplexing but is a bit more complicated because of the interactions between the signals. There are three main types of multiple access: time division multiple access (TDMA), frequency division multiple access (FDMA), and code division multiple access (CDMA). In TDMA, each signal (which may itself be a collection of signals) uses the channel for a specified period of time. This requires complex time synchronization. In FDMA, each signal gets its own frequency band subchannel. In CDMA, each signal is encoded with a pseudorandom code which spreads the signal over a larger bandwidth (Dixon 1976) (CDMA is a spread spectrum technique) and is extracted using the same code. Multiple access usually takes place at an intermediate radio frequency, and there are complex interactions between signals that can affect them.

Intermodulation distortion can occur when several carriers share a nonlinear device, such as some power amplifiers. An amplifier may have a linear transfer function in a portion of its operating curve but may be nonlinear at the high power output portion of its operating curve. The amplifier may be operated at a back off point compared to its maximum power output to ensure that it is operating as a linear device to minimize intermodulation distortion effects if several carriers are sharing it.

Shannon's Law

Claude Shannon was one of the seminal thinkers of the twentieth Century. He worked as a researcher at Bell Labs in its heyday. By applying Boolean algebra to telephone switching circuits, he set the stage for digital electronic design. He developed a mathematical theory of communication that included the foundation of information theory (Hamming 1986). One result of the theory is a description of the limit of the amount of information that can be transmitted on a given channel (that is sometimes known as "Shannon's law" or the "Shannon–Hartley theorem") commonly called the channel capacity. The "Shannon limit" refers the theoretical best rate of transmission for a given channel. The channel capacity equation is

$$R \leq W \log2 \left(P_R / P_N + \right)$$

where

R = theoretical maximum transmission rate, bits per second
W = usable bandwidth, Hertz
P_R = power of received signal
P_N = power of noise

The Shannon limit implies that a system can be designed (i.e., with sufficient complexity, e.g., using error-correcting codes) that can allow transmission on a channel that approaches the rate of the Shannon limit. For error-free communications, E_b/N_0 has to be at least -1.6 dB, even with infinite bandwidth. Error correction coding can reduce the power required but uses more bandwidth. By adding coding to the system, complexity and bandwidth are increased, but power required is decreased (i.e., complexity and bandwidth are traded for power).

Link Budgets

Communications is achieved by introducing into free space a modulated radio wave (transmitting) which propagates over a distance to a point where a receiving station captures a portion of the transmitted energy and extracts information from it. As the radio wave propagates through space, it spreads out and the energy transmitted is spread across an increasingly larger cross-sectional area; less energy is available to be captured by a given antenna as the distance increases. Although we have been talking about beamwidth as a two-dimensional measurement characterizing an antenna, the transmitted wave actually propagates in a three-dimensional solid angle. A solid angle can be visualized as a cone whose apex is at the transmitting antenna having a base, or cross section of the beam, at any given distance from the antenna (this approximation is an illustrative aid, since waves are usually assumed to propagate spherically which would mean that the base bulges out as a spherical section). The area of the cross section of the beam gets larger as the distance from the antenna increases and the energy of the transmitted wave is spread over that increasingly larger cross section. The received power flux density drops off inversely proportional to the square of the distance.

As the distance between the transmitter and receiver increases, the signal power decreases (all else being equal). At some point the received signal power will be less than the noise that is received/generated by the receiver, and reliable communications will not be possible. The greatest contribution to the decrease in signal power received is the loss due to the propagation distance known as the path loss or free space loss. The free space loss, L_s, is given by

$$L_s = (4\pi d)^2/\lambda^2$$

where

d is the distance in meters (or whatever unit is used for λ)
λ is the wavelength in meters

The wavelength term is not part of the physical path loss but is related to the receiving antenna's aperture (and related gain) and is included with the physical path loss.

The propagation distance from the ground to a geostationary satellite is around 36,000 km (the distance varies depending on the latitude and longitude of the ground station and the location of the satellite in the geostationary arc). At a frequency of 6 GHz, the free space loss is approximately 1×10^{20}. Because this is a loss, this means that the signal will be 10^{-20} times weaker when the path loss is accounted for.

Although it is usually the largest factor by far, the path loss is not the only contribution to the received power; there are many factors that can boost or decrease the power of the transmitted wave or that can introduce noise to the system. The ability to communicate depends on being able to receive the desired signal at a level above the received and generated noise. Factors that decrease signal (loss) or increase noise are detrimental to reliable communications, whereas factors that increase signal (gain) or decrease noise are beneficial.

The system *link budget* is used to account for the various gains and losses across a communications link between two nodes. Part of the link budget is calculated according to a conceptually simple equation accounting for gains and losses:

$$T_p * G_{link}/L_{link} = R_p$$

where

T_p is the transmitted power
G_{link} is the product of all the gains in the link
L_{link} is the product of all the losses in the link
R_p is the received power

Once we have the received power, we can calculate the signal-to-noise ratio (SNR):

$$SNR = S/N$$

where, in the simplest case, the received signal power is $S = R_P$ and N is the received noise (including noise generated by the receiver). The signal-to-noise ratio is the key parameter in determining whether or not a communications link will be successful (or in the vernacular, whether it is possible to "close the link"). The higher the SNR, the better the link. The signal is modulated onto a radiofrequency (RF) wave known as a *carrier*. Sometimes the carrier-to-noise ratio, C/N or CNR, is used in satellite system design to characterize a link at the RF level before demodulation, whereas the SNR is a measure of the signal strength after demodulation. An RF carrier will usually carry many baseband signals. In a digital system, the signal-to-noise quantity we are looking for is the ratio of the energy per bit to noise density, E_b/N_0.

In a satellite relay system, there are two links that need to be analyzed: (1) the uplink, or the link from the ground to the satellite, and (2) the downlink or the link from the satellite to a second ground station. The uplink and the downlink are analyzed separately and the results combined to get the overall link budget.

The key result from a link budget is a quantity known as the link margin. This is the difference between the actual received power and the required received power in decibels (see below for a discussion of decibels). Hopefully the link margin is a positive quantity; a negative quantity indicates that the link will not be closed, and communications will not be successful. Generally, the higher the link budget, the better it is from an engineering standpoint (too high and a cost accountant somewhere should be complaining).

Understanding Decibels

The range in values of the various factors involved in calculating the link budget is very large. Engineers use a dimensionless unit, the *decibel* or dB (named after Alexander Graham Bell), to simplify the calculations. A result in decibels, R, is defined as a ratio of two powers, P_1 and P_2, such that

$$R\ dB = 10 \log_{10}(P_1/P_2)$$

or ten times the logarithm (to the base 10) of the ratio of the two power quantities. (The logarithm of a number is the exponent to which the base is raised to get that number.) Some examples of the conversion to decibels:

Let the ratio $P_1/P_2 = 10$ (i.e., $P_1 = 10 \times P_2$). Then

$$R = 10\log_{10}(10) = 10 \text{ dB}$$

since $\log_{10}(10) = 1$. Admittedly, this example is not a very exciting transformation. Let us try some more.

Let the ratio $P_1/P_2 = 1$ (i.e., $P_1 = P_2$). Then

$$R = 10\log_{10}(10) = 10 \text{ dB}$$

$$\text{since } \log_{10}(1) = 0.$$

Let the ratio $P_1/P_2 = 1/10$ (that is $10 \times P_1 = P_2$). Then

$$R = 10\log_{10}(1/10) = -10 \text{ dB}$$

since $\log_{10}\left(10^{-1}\right) = -1$. That is still not very impressive. Let us look at something a little bigger.

Let the ratio $P_1/P_2 = 1{,}000{,}000$ (that is $P_1 = 1{,}000{,}000\ P_2$). Then

$$R = 10\log_{10}(1{,}000{,}000) = 60 \text{ dB}$$

since $\log_{10}\left(10^6\right) = 6$. Ah, now we are getting somewhere. Looking at the result (60 dB), we can see that there should be six zeros in the original ratio because of the logarithmic nature of the dB. Table 3 shows that working with decibels makes it easier to work with power ratios having large dynamic ranges.

Dealing with large ranges of ratios cannot be the only reason for using decibels since scientific notation also makes it easier to work with large ranges of numbers

Table 3 Decibels provide compact representations of large dynamic ranges

Ratio	Exponent	Decibels
1	10E0 or 10^0	0 dB
10	10E1	10 dB
100	10E2	20 dB
1,000	10E3	30 dB
1,000,000	10E6	60 dB
1,000,000,000	10E9	90 dB
1,000,000,000,000	10E12	120 dB
100,000,000,000,000,000,000	10E20	200 dB

and does not require calculations. Another reason for using decibels is that the equivalent function to multiplying ratios is adding decibels and the equivalent function to dividing ratios is subtracting decibels. This is because of the logarithm in the definition. That means that instead of multiplying gains and dividing losses to get a link budget, we can convert to decibels and add gains and subtract losses. For example:

$$1/10 = 0 \text{ dB} - 10 \text{ dB} = -10 \text{ dB}$$

as we saw before. We will come back to link budgets, but first a few more examples.

Let the ratio $P_1/P_2 = 2$ (i.e., $P_1 = 2P_2$). Then

$$R = 10\log_{10}(2) = 3.0\,1 \text{ dB}$$

since $\log_{10}(2) = 0.30102999\ldots$. Most engineers round off and use 3 dB to mean a ratio of 2. The power ratio equivalent of 3 dB is calculated:

$$10\text{E}\ (R/10) = P_1/P_2$$
$$10^{(3/10)} = 1.99526\ldots$$

which is approximately 2 as we would expect. Some other useful values to remember are given in Table 4.

Knowing the conversions for 2, 3, 7, and 10, we can get the rest from easy manipulation of the decibel values (the examples below use various levels of rounding off):

$$4 = 2 \times 2 = 3 \text{ dB} + 3 \text{ dB} = 6 \text{ dB}$$
$$5 = 10/2 = 10 \text{ dB} - 3 \text{ dB} = 7 \text{ dB}$$
$$6 = 2 \times 3 = 3.0 \text{ dB} + 4.8 \text{ dB} = 7.8 \text{ dB}$$
$$8 = 2 \times 4 = 3 \text{ dB} + 6 \text{ dB} = 9 \text{ dB}$$
$$9 = 3 \times 3 = 4.77 \text{ dB} + 4.77 \text{ dB} = 9.54 \text{ dB}$$

Table 4 Converting numbers from 1 to 10 to decibels (rounded to two decimal places)

Ratio	dB
1	0
2	3.01
3	4.77
4	6.02
5	6.99
6	7.78
7	8.45
8	9.03
9	9.54
10	10.00

Table 5 Decibel values from 0 to 10 and their corresponding ratio values to two decimal places

Decibels	Ratio
0	1
1	1.26
2	1.58
3	1.99
4	2.51
5	3.16
6	3.98
7	5.01
8	6.31
9	7.94
10	10

With these values, it is easy to put together whatever number we need in decibels. For example:

$$2{,}000{,}000 = 2 \times 1{,}000{,}000 = 3\text{ dB} + 60\text{ dB} = 63\text{ dB}$$
$$400 = 4 \times 100 = 6\text{ dB} + 20\text{ dB} = 26\text{ dB}$$

Table 5 shows the case for the inverse conversion from decibels to power ratios rounded off to two places. As we have already seen, 3 dB is typically rounded off to a value of 2, which would have happened if we rounded these values off to one decimal place.

To convert, say, 26 dB to a ratio, one could use the definition and calculate:

$$10\text{E}(26/10) = 398$$

or use some simple arithmetic and some memorized conversions (i.e., that 6 dB is approximately a ratio of 4):

$$26\text{ dB} = 20\text{ dB} + 6\text{ dB} = 100 \times 4 = 400$$

which is hopefully close enough for the desired application.

Up to this point, we have been talking about converting power ratios to decibels because that is the way decibels are defined. Of course, if one needs to convert a number, the special case where $P_2 = 1$ can be used, so that the ratio P_1/P_2 becomes P_1. At the risk of confusing the reader, there are cases where dB are used for amplitude ratios where the definition is $20 \log_{10}(A_1/A_2)$. This definition is not used in the link budget discussion below but is provided for completeness in the decibel discussion. The amplitude definition of decibels, dBV (V for volts in this case), can be encountered by replacing the power quantities in the power ratio by voltages:

$$P = V^2/Z$$

where P = power, V = voltage, and Z = impedance, so that

$$
\begin{aligned}
R &= 10\log10(P_1/P_2)\\
&= 10\log_{10}\left(\left(V_1{}^2/Z\right)/\left(V_2{}^2/Z\right)\right)\\
&= 10\log_{10}(V_1/V_2)^2\\
&= 20\log_{10}(V_1/V_2)
\end{aligned}
$$

because $\log_{10}X^2 = 2\log_{10}X$.

The decibel is not an official unit of the International System of Units (SI); however, it is approved by the International Committee for Weights and Measures (CIPM, *Comité International des Poids et Mesures*) for use with the SI. A common practice, which is not in accordance with official SI rules, is to attach a reference unit indicator to the dB symbol, such as dBW, dBm, dBHz, etc., as a shorthand way of specifying a reference level (Ambler and Taylor 2008). These units indicate a reference value in order to use the decibel definition in a specialized way as an absolute unit rather than a ratio (relative) unit. For example, dBW indicates a measurement relative to a value of 1 W (or $P_2 = 1$ W). Thus

$$R\ \mathrm{dBW} = 10\log_{10}(P_1/1\mathrm{W})$$

so that, if $P_1 = 1$ W, then $R = 0$ dBW. In SI units, the reference value of 1 W would have to be specified explicitly each time, i.e., R (re 1 W) = 0 dB or $R_{(1\ W)} = 0$ dB. Common reference values that may be encountered in non-SI literature are shown in Table 6.

Link Budget Calculation

The simplified power balance equation (see above)

$$T_p * G_{link}/L_{link} = R_p$$

using decibels becomes

$$T_p\mathrm{dB} + G_{link}\mathrm{dB} - L_{link}\mathrm{dB} = R_p\mathrm{dB}$$

Table 6 Informal decibel reference indications

Informal reference unit	Reference value
dBW	1 W
dBm	0.001 W or 1 mW
dBi	Isotropic antenna gain
dBHz	1 Hz

where

T_p dB is the transmitted power in decibels (referenced to W or mW)
G_{link}dB is the sum of all the gains in the link in decibels
L_{link}dB is the sum of all the losses in the link in decibels
R_pdB is the received power in decibels (referenced to W or mW as in T_p)

Although this equation is very simple, the challenge is in finding all of the significant gains and losses and keeping their signs correct. The definition of a loss, for example, could be as a positive quantity to be subtracted or as a negative quantity to be added. For example, the free space loss to GEO at 6 GHz (which we saw above is 10^{20}) is 200 dB which is subtracted in the link budget equation.

The transmitted power and transmitter antenna gain are usually combined in a quantity called the effective isotropic radiated power (EIRP). As the name implies, EIRP is the power that would have to be radiated from an isotropic antenna (an antenna that radiates equally in all directions) to provide the same power flux density at the receiver:

$$\text{EIRP} = P_T - L_T + G_T \text{dBW}$$

where

P_T = transmitter power
L_T = transmitter losses
G_T = transmit antenna gain

The important parameter to the receiver is the received power. Different combinations of transmitter hardware can provide the same received power, and EIRP is a way of representing the transmitted power in a standard way. For example, if we ignore the losses, an EIRP of 50 dBW could be obtained by any number of combinations of transmit power and antenna gain (e.g., $P_T = 15$ dBW and $G_T = 35$ dB, $P_T = 20$ dBW and $G_T = 30$ dB, etc.). Higher EIRP results in higher receiver power, all else being equal, but there are regulatory limits on the amount of power that can be radiated onto the Earth. These regulatory limits are for the purpose of limiting interference to other communications systems.

Another concept, G/T, is used to characterize the receiver. The receiver figure of merit is the ratio of the receiving antenna gain, G, to the receiver system noise temperature, T (in Kelvins). The higher the G/T, the better the receiver. The noise temperature is a way of lumping together various noise contributions, many of which are thermal in origin and the remainder of which are modeled as thermal. The received noise power is related to the noise temperature by Boltzmann's constant and the bandwidth of interest.

The simplified link budget equation (Miller et al. 1993) can be written using EIRP and G/T:

$$C/N = EIRP - L_s + 10\log_{10}(G/T) - 10\log_{10}(kB)$$

where k is Boltzmann's constant and B is the equivalent noise bandwidth of the receiver.

For the purposes of this chapter, it is sufficient to realize that there are many quantities that need to be determined for entry into a link budget calculator, some of which are combinations of lower-level details of the system design. The level of detail in the link budget depends on its purpose. As a design tool, the system engineer will have a link budget that may have dozens of separate items that allow trade-offs to be made. Once a system is built and losses become more determined and tightly range bound, some quantities can be combined into a simplified link budget for use in, say, determining a particular ground station design.

A detailed design link budget may include gains and losses due to items such as circuits (may include cables, connectors, waveguides, switches, etc.), antenna size, antenna polarization, antenna alignment, weather, atmospheric gasses, thermal noise, active device noise, interference, back off or attenuation, coding, etc.

There are many free and commercial automated tools for calculating radiofrequency link budgets and especially link budgets for satellite systems. Many link budget tools for educational purposes require the user to enter all parameters but may provide a limited number of variable parameters by lumping detailed parameters together. This provides a nice platform for getting the feel of a link budget without getting bogged down in the details of a design (Fig. 7).

A system engineer doing a detailed design will probably use a spreadsheet to build their own tool so that they can introduce parameters as needed. Some of the parameters in the link budget are related to the satellite design (e.g., transponder bandwidth, transmitted power, antenna gain), and some are related to the ground station design (e.g., terminal location, antenna gain, receiver sensitivity). To add services with an existing satellite, the satellite design parameters are already determined since it is already deployed in space and can be obtained from the satellite service provider or from a database.

Commercial tools allow potential users to investigate system options by including a database of existing satellite and ground systems with parameters needed to conduct link budget analyses in addition to a tool for doing the calculations.

Some software packages provide a useful representation of a satellite transponder's coverage called a footprint map (see Fig. 8). A footprint is the intersection of a satellite's beam with the surface of the Earth. The contours of the map show the areas of signal strength due to the antenna pattern. Note in Fig. 4 that the designer of the spacecraft antenna shaped the antenna coverage contour to evenly spread the maximum EIRP output over the US northern continental area and the populous areas of Canada. Presumably, this is the area of the primary intended customer base.

Table 7 shows sample output from a commercial software package called SatFinder. This table indicates the considerable complexity found in an actual link

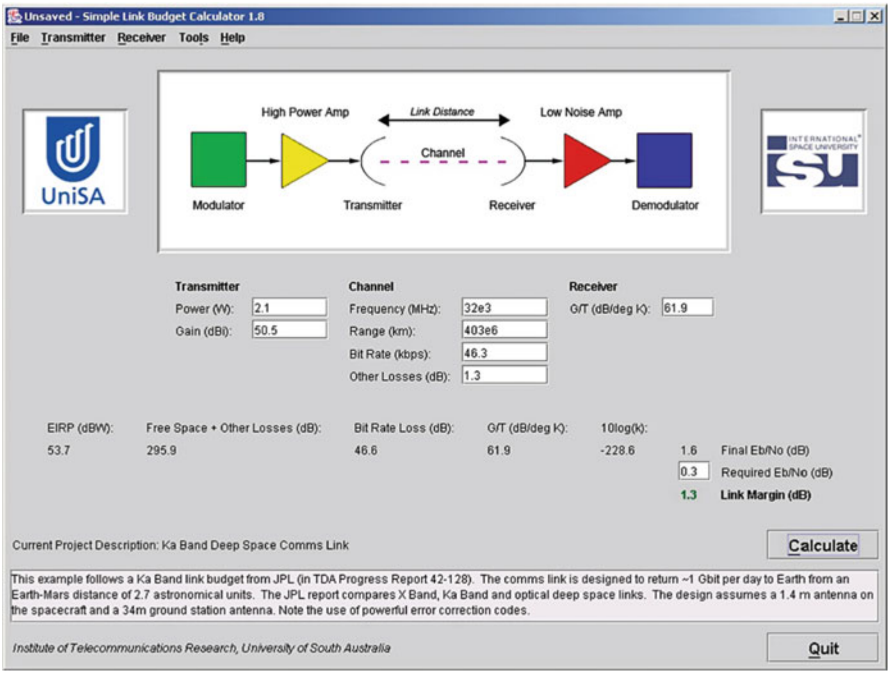

Fig. 7 Example of a free tool for demonstrating link budget concepts (Courtesy UniSA, from http://www.itr.unisa.edu.au/itrusers/bill/public_html/software/)

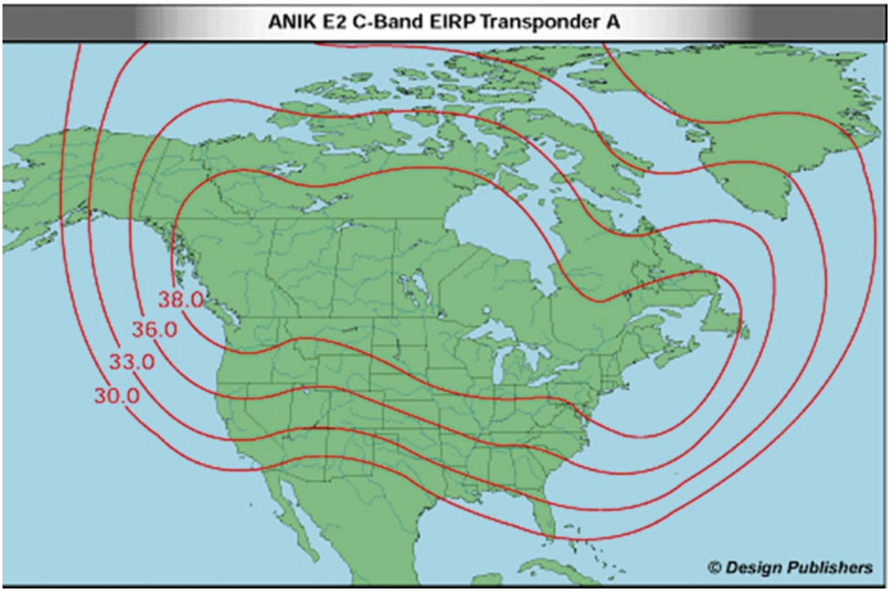

Fig. 8 Example footprint map from SatFinder (Courtesy SatNews.com)

Table 7 Example output from SatFinder link budget software

Digital link budget produced using SatFinder (http://www.satnews.com/linksample.htm)				
Service name	Unspecified			
Coverage	Unspecified			
Uplink earth station	Denver			
Downlink earth station	Los Angeles			
Satellite name	Telstar 402R			
Link input parameters	*Up*	*Down*	*Units*	
Site latitude	39.73 N	34.05 N	degrees	
Site longitude	105.00 W	118.25 W	degrees	
Site altitude	1	0.1	km	
Frequency	14.472	12.172	GHz	
Polarization	Vertical	Vertical	–	
Rain model	ITU (30.3)	ITU (19.7)	(mm/h or zone)	
Availability (average year)	99.9	99.9	%	
Water vapor density	3	10	gm/m^3	
Surface temperature	10	20	°C	
Antenna aperture	10	0.9	meters	
Antenna efficiency/gain	65	70	% (+ prefix dBi)	
Coupling loss	0.2	0	dB	
Antenna tracking/mispoint error	0	0.3	dB	
LNB noise figure/temp	–	0.75	dB (+ prefix K)	
Antenna noise	–	27	K	
Adjacent carrier interference	30	30	dB	
Adjacent satellite interference	30	12	dB	
Cross polarization interference	200	200	dB	
Uplink station HPA output back off	3	–	dB	
Number of carriers/HPA	1	–	–	
HPA C/IM (up)	200	–	dB	
Uplink power control	0	–	dB	
Uplink filter truncation loss	0	–	dB	
Required HPA power capability	MAX	–	W	
Satellite input parameters	*Value*	*Units*		
Satellite longitude	89.00 W	degrees		
Transponder type	TWTA	–		
Receive G/T	1	dB/K		
Saturation flux density	−86.05	dBW/m^2		
Satellite attenuator pad	0	dB		
Satellite ALC	0	dB		
EIRP (saturation)	47	dBW		
Transponder bandwidth	27	MHz		

(*continued*)

Table 7 (continued)

Digital link budget produced using SatFinder (http://www.satnews.com/linksample.htm)				
Input back off total	1	dB		
Output back off total	0.3	dB		
Intermodulation interference	30	dB		
Number of transponder carriers	AUTO	–		
Carrier/link input parameters	*Value*	*Units*		
Modulation	4-PSK	–		
Required bit error rate performance	10^{-9}	–		
Required Eb/No without FEC coding	20	dB		
Required Eb/No with FEC coding	5.1	dB		
Information rate	23.6	Mbps		
Overhead	0	%		
FEC code rate	0.59	–		
Spreading gain	0	dB		
Reed–Solomon code	1	–		
(1 + Roll-off factor)	1.2	–		
Carrier spacing factor	1.3	–		
Bandwidth allocation step size	0.1	MHz		
System margin	0.9	dB		
Calculations at saturation	*Value*	*Units*		
Gain 1 m^2	44.67	dB/m^2		
Uplink C/No	98.88	dBHz		
Downlink C/No	89.72	dBHz		
Total C/No	79.59	dBHz		
Uplink EIRP for saturation	76.65	dBW		
General calculations	*Up*	*Down*	*Units*	
Elevation	41.14	39.61	degrees	
True azimuth	155.84	134.99	degrees	
Compass bearing	145.77	121.46	degrees	
Path distance to satellite	37692.47	37809.06	km	
Propagation time delay	0.125728	0.126117	seconds	
Antenna efficiency	65	70	%	
Antenna gain	61.75	39.65	dBi	
Availability (average year)	99.9	99.9	%	
Link downtime (average year)	8.766	8.766	hours	
Availability (worst month)	99.615	99.615	%	
Link downtime (worst month)	2.809	2.809	hours	
Spectral power density	−59.11	−26.31	dBW/Hz	
Uplink calculation	*Clear*	*Rain up*	*Rain dn*	*Units*
Uplink transmit EIRP	75.65	75.65	75.65	dBW
Transponder input back off (total)	1	1	1	dB
Input back off per carrier	1	4.32	1	dB
Mispoint loss	0	0	0	dB
Free space loss	207.18	207.18	207.18	dB

(*continued*)

Table 7 (continued)

Digital link budget produced using SatFinder (http://www.satnews.com/linksample.htm)				
Atmospheric absorption	0.1	0.1	0.1	dB
Tropospheric scintillation fading	0.08	0.08	0.08	dB
Atmospheric losses total	0.18	0.18	0.18	dB
Total path loss (excluding rain)	207.36	207.36	207.36	dB
Rain attenuation	0	3.32	0	dB
Uplink power control	0	0	0	dB
Uncompensated rain fade	0	3.32	0	dB
C/No (thermal)	97.88	94.56	97.88	dB. Hz
C/N (thermal)	24.08	20.76	24.08	dB
C/ACI	30	26.68	30	dB
C/ASI	30	26.68	30	dB
C/XPI	200	196.68	200	dB
C/IM	200	200	200	dB
Eb/(No + Io)	13.26	13.26	13.26	dB
Downlink calculation	*Clear*	*Rain up*	*Rain dn*	*Units*
Satellite EIRP total	47	47	47	dBW
Transponder output back off (total)	0.3	0.3	0.3	dB
Output back off per carrier	0.3	3.62	0.3	dB
Satellite EIRP per carrier	46.7	43.38	46.7	dBW
Mispoint loss	0.3	0.3	0.3	dB
Free space loss	205.71	205.71	205.71	dB
Atmospheric absorption	0.1	0.1	0.1	dB
Tropospheric scintillation fading	0.3	0.3	0.3	dB
Atmospheric losses total	0.4	0.4	0.4	dB
Total path loss (excluding rain)	206.11	206.11	206.11	dB
Rain attenuation	0	0	2.74	dB
Noise increase due to precipitation	0	0	4.14	dB
Downlink degradation (DND)	0	0	6.89	dB
Total system noise	81.67	81.67	212.02	K
Figure of merit (G/T)	20.23	20.23	16.09	dB/K
C/No (thermal)	89.42	86.1	82.53	dB. Hz
C/N (thermal)	15.61	12.29	8.73	dB
C/ACI	30	26.68	30	dB
C/ASI	12	8.68	12	dB
C/XPI	200	196.68	200	dB
C/IM	30	26.68	30	dB
Eb/(No + Io)	−4.15	−4.15	−4.15	dB
Totals per carrier (end-to-end)	*Clear*	*Rain up*	*Rain dn*	*Units*
C/No (thermal)	79.59	79.59	79.59	dB. Hz
C/N (thermal)	5.78	5.78	5.78	dB

(*continued*)

Table 7 (continued)

Digital link budget produced using SatFinder (http://www.satnews.com/linksample.htm)				
C/ACI	23.19	23.19	23.19	dB
C/ASI	6.15	6.15	6.15	dB
C/XPI	193.19	193.19	193.19	dB
C/IM	26.2	26.2	26.2	dB
C/(No + Io)	59.96	59.96	59.96	dB. Hz
C/(N + I)	−13.85	−13.85	−13.85	dB
Eb/(No + Io)	−13.77	−13.77	−13.77	dB
System margin	0.9	0.9	0.9	dB
Net Eb/(No + Io)	−14.67	−14.67	−14.67	dB
Required Eb/(No + Io)	5.1	5.1	5.1	dB
Excess margin	−19.77	−19.77	−19.77	dB
Earth station power requirements	*Value*	*Units*		
EIRP per carrier	75.65	dBW		
Antenna gain	61.75	dBi		
Antenna feed flange power per carrier	13.9	dBW		
Uplink power control	0	dB		
HPA output back off	3	dB		
Waveguide loss	0.2	dB		
Filter truncation loss	0	dB		
Number of HPA carriers	1	–		
Total HPA power required	17.1016	dBW		
Required HPA power capability	10	W		
Spectral power density	−59.11	dBW/Hz		
Space segment utilization	*Value*	*Units*		
Overall link availability	99.8	%		
Information rate (inc overhead)	23.6	Mbps		
Transmit rate	40	Mbps		
Symbol rate	20	Mbaud		
Occupied bandwidth	24	MHz		
Noise bandwidth	73.8	dB.Hz		
Minimum allocated bandwidth required	26	MHz		
Allocated transponder bandwidth	26.1	MHz		
Percentage transponder bandwidth used	96.67	%		
Used transponder power	46.7	dBW		
Percentage transponder power used	100	%		
Max carriers by transponder bandwidth	1.03	–		
Max carriers by transponder power	1	–		
Max transponder carriers limited by	Power	[1.00]		
Power equivalent bandwidth usage	27	MHz		

budget calculation for a particular satellite. SatFinder has a database with information on over 500 satellites (including orbital location, frequencies, EIRP, G/T, bandwidth, etc.). Similar satellite link budget tools include products that can be found online such as Satmaster or Customizable Link Budget Tool.

Some commercial network modeling tools, such as OpNet, have satellite link modeling capability and can integrate with other tools, such as Analytical Graphics, Inc.'s Satellite Tool Kit (STK), to provide analysis of satellite systems. STK has several nice communications functions and tutorials that can help in understanding link budgets and system trade-offs. There are tools from other companies that are specialized for regulatory coordination to avoid interference between systems. Examples include Visualyse (from Transfinite) and Sat-Coord (from RPC Telecommunications).

The main result from a link budget analysis is the link margin. This is the amount of received power over and above the minimum amount required to close the link and meet the system requirements for quality. A good value for a link margin in a fixed satellite system (where the ground stations are well characterized and design parameters do not change radically) is around 3 dB. For high-frequency systems (Ka band), rain fade is an important design consideration. From Fig. 2 we can see that rain fade may require 10 dB or more of margin, but there are techniques for dealing with rain including site diversity and adaptive power control.

For a mobile satellite system, at least one ground terminal is a mobile device, perhaps handheld. In that case, the design parameters may be more variable depending on how and where the handset is being used. Handheld devices generally need to be low power (to minimize radiation of the user) and have low-gain antennas (to avoid having to point the antenna at the satellite). In an urban environment, there is shadowing from buildings as well as reflections that cause multipath distortions. Mobile systems may make use of LEO constellations or a GEO satellite. In a LEO system, there may be a need for some margin for handoffs between satellites. In general, a link margin of 10–15 dB might be appropriate for mobile systems. However, one must examine the link budget to see where various effects are accounted for. Assumptions and confidence in models used to arrive at parameters used in the link budget calculation (such as rain fade) will affect the amount of link margin required.

Adaptive power control can be used to vary the transmitted power to overcome short-term increases of the link path loss (known as fading). When conditions are good, the transmitted power can be reduced, and when conditions deteriorate, transmitted power levels can be increased. Some link margin is still required to accommodate dynamic changes in losses that occur faster than the adaptive feedback can respond.

Availability and signal quality are two requirements that will determine an acceptable link margin. Availability refers to the percentage of time that a satellite link can be closed during a certain time period (a year is typically used). An availability of 99.95 % implies that a link is down for a total of around 4 h per year. Signal quality depends on the type of signal being transmitted. For digital signals, the signal quality parameter of interest is the bit error rate. For other signals it might be the SNR or C/N.

Conclusion

Link budgets are an accounting of the gains and losses in a communications system. The link margin is a measure of the robustness of the link. The channel capacity equation implies that a communications engineer can trade off power, bandwidth, and complexity to achieve various system requirements in a design. Establishing link budget and link margins for different types of satellite systems represents a variety of challenges. As one moves to higher frequencies and one moves from fixed to mobile satellite systems, the more difficult the challenge becomes and the higher the link margin must be set to provide reliable service.

Cross-References

- ▶ An Examination of the Governmental Use of Military and Commercial Satellite Communications
- ▶ Global Communications Satellite Systems
- ▶ US Domestic Communications Satellite Systems
- ▶ Common Elements Versus Unique Requirements in Various Types of Satellite Application Systems
- ▶ Economics and Financing of Communications Satellites
- ▶ Fixed Satellite Communications: Market Dynamics and Trends
- ▶ Geographic Information Systems and Geomatics
- ▶ History of Satellite Communications
- ▶ Introduction to Satellite Navigation Systems
- ▶ Mobile Satellite Communications Markets: Dynamics and Trends
- ▶ Overview of the Spacecraft Bus
- ▶ Regulatory Process for Communications Satellite Frequency Allocations
- ▶ Satellite Antenna Systems Design and Implementation Around the World
- ▶ Satellite Communications and Space Telecommunication Frequencies
- ▶ Satellite Communications Antenna Concepts and Engineering
- ▶ Satellite Communications Modulation and Multiplexing
- ▶ Satellite Communications Overview
- ▶ Satellite Communications: Regulatory, Legal, and Trade Issues
- ▶ Satellite Communications Video Markets: Dynamics and Trends
- ▶ Satellite Earth Station Antenna Systems and System Design
- ▶ Satellite Orbits for Communications Satellites
- ▶ Satellite Transmission, Reception, and Onboard Processing, Signaling, and Switching
- ▶ Space Telecommunications Services and Applications
- ▶ Telemetry, Tracking, and Command (TT&C)
- ▶ Trends and Future of Satellite Communications

References

T. Ambler, B.N. Taylor, *Guide for the Use of the International System of Units (SI)*. NIST Special Publication 811, 2008th edn. (U.S. Department of Commerce, Technology Administration, National Institute of Standards and Technology, Gaithersburg, 2008), p. 30

V.K. Bhargava, D. Haccoun, R. Matyas, P.P. Nuspl, *Digital Communications by Satellite* (Wiley, New York, 1981), p. 24

R.C. Dixon, *Spread Spectrum Systems* (Wiley, New York, 1976), p. 13

R.M. Gagliardi, *Satellite Communications* (Van Nostrand Reinhold, New York, 1984), pp. 17–20

R.W. Hamming, *Coding and Information Theory*, 2nd edn. (Prentice-Hall, Englewood Cliffs, 1986), p. 2

S. Hayken, *Communications Systems*, 2nd edn. (Wiley, Chichester, 1983), p. 469

N.S. Jayant, P. Noll, *Digital Coding of Waveforms* (Prentice-Hall, Englewood Cliffs, 1984), p. 221

M.J. Miller, B. Vucetic, L. Berry (eds.), *Satellite Communications Mobile and Fixed Services* (Kluwer, Boston, 1993), p. 51

W.W. Peterson, E.J. Weldon Jr., *Error-Correcting Codes*, 2nd edn. (MIT Press, Cambridge, 1991), p. 6