Intuitive Guide to Principles of Communications
www.complextoreal.com

Link Budgets

You are planning a vacation. You estimate that you will need $1000 dollars to pay for the hotels, restaurants, food etc.. You start your vacation and watch the money get spent at each stop. When you get home, you pat yourself on the back for a job well done because you still have $50 left in your wallet.

We do something similar with communication links, called creating a link budget. The traveler is the signal and instead of dollars it starts out with "power". It spends its power (or attenuates, in engineering terminology) as it travels, be it wired or wireless.

Just as you can use a credit card along the way for extra money infusion, the signal can get extra power infusion along the way from intermediate amplifiers such as microwave repeaters for telephone links or from satellite transponders for satellite links. The designer hopes that the signal will complete its trip with just enough power to be decoded at the receiver with the desired signal quality.

In our example, we started our trip with $1000 because we wanted a budget vacation. But what if our goal was a first-class vacation with stays at five-star hotels, best shows and travel by QE2? A $1000 budget would not be enough and possibly we will need instead $5000. The quality of the trip desired determines how much money we need to take along.

With signals, the quality is measured by the Bit Error Rate (BER). If we want our signal to have a low BER, we would start it out with higher power and then make sure that along the way it has enough power available at every stop to maintain this BER.

How we specify the quality of signal transmission

The BER, as a measure of the signal quality, is the most important figure of merits in all link budgets. The BER is a function of a quantity called E_b/N_o, the bit energy per noise-density of the signal. For a QPSK signal in an additive white-gaussian-noise (AWGN) channel, the BER is given by

$$BER = \frac{1}{2} erfc\left(\sqrt{E_b / N_0}\right)$$

This formula says that the BER of any signal is related to its E_b/N_o by the function, *erfc*. The function *erfc,* called the complimentary error function describes the cumulative probability curve of a gaussian distribution. It is found tabulated in most communications textbooks and is available as a built-in function in most math programs.

The above equation when plotted has a classic waterfall shape when plotted on a log-log scale. The BER is inversely related to E_b/N_o. Higher E_b/N_o means better quality.

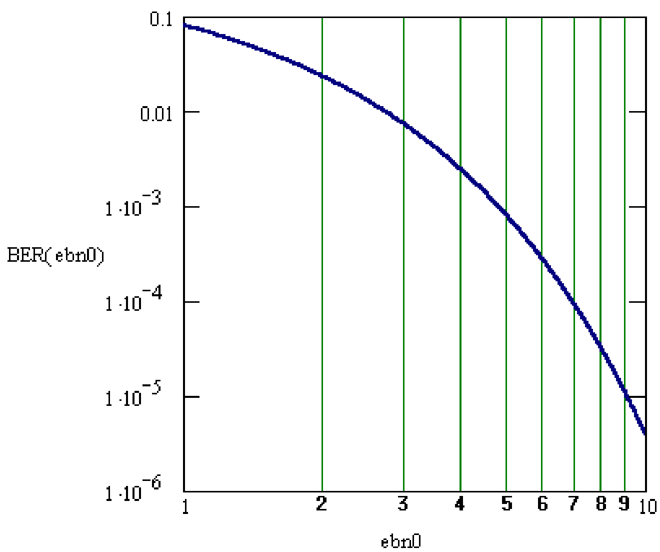


Figure 1 - The Bit Error rate of a signal is a function of its E_b/N_o

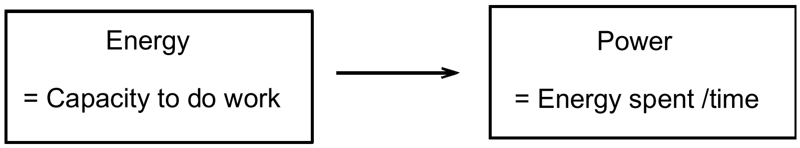
Although I started out talking about signal power as a measure of signal quality, you may have noticed that I have suddenly shifted to E_b/N_0. Instead of monitoring just the signal power as I said on the first page, we will actually monitor a ratio of the bit power to the noise power injected along the way.

This means that in doing link budgets, we will keep track of not just the signal power and how it is getting attenuated, but also where the noise is entering into the link and how much. This dual role makes things complicated and messy because there are numerous sources of noise.

E_b/N_o - a measure for digital inks

E_b/N_o is the most common parameter used to compare communication systems even when they have differing bit rates, modulations, and even media. Let's take a closer look at the E_b/N_o.

The quantity E_b is a measure of the Bit Energy. What is energy? The energy is the capacity to do work and energy expended per time is called power.



To compute E_b, we divide the average signal power by its bit rate.

$$E_b = \frac{P_{avg.}}{R_b}$$

This makes sense because the average power is the energy per unit time, and the bit rate is the number of bits per unit time. The division removes the units of time leaving energy per bit.

We can also write the above equation in an alternate form with the amplitude-squared representing the P^{avg}.

$$E_b = \frac{A^2}{R_b}$$

Example 1: A signal has power of 10 watts. Its bit rate is 200 bits per second. What is its E_b?

E_b = 10 Log (10) - 10 Log (200) = -13 dB

In the denominator of E_b/N_0, the quantity N_0 is called the noise density. It is the total noise power in the frequency band of the signal divided by the bandwidth of the signal. It is measured as Watts/Hz and is the noise power in one Hz of bandwidth.

$$N_0 = \frac{P_N}{B_N}$$

where P_N = noise power and B_N = noise bandwidth. The units are Joules.

Example 2: A signal experiences a noise power of 2 watts. The signal bandwidth is 500 Hz.
(a) What is its noise power density? (b) What is its E_b/N_o?
(c) What would be its E_b/N_o if its bandwidth is 300 Hz instead of 500 Hz? Is it more or less?

(a) N_0 = 10 Log (2) - 10 Log (500) = -23.9 dB

Case 1, Bandwidth = 500 Hz
(b) E_b/N_0 = -13 - (-23.9) = 10.9 dB
Case 2, Bandwidth = 300 Hz
(d) E_b/N_0 = -13 -10 Log (2) + 10 Log (300) = 9.7 dB

Note: The reason E_b/N_o went down in case 2 in the above example is that when we decreased the bandwidth is that the noise density has increased. The same amount of noise now occupies a smaller signal space. E_b/N_o decreases as the bandwidth reduced. It also decreases when we increase the bit rate. This is because E_b is inversely proportional to the bit rate. So one of the simplistic ways to improve the E_b/N_o of a link is to reduce its bit rate.

C/N and C/N_o - a measure of analog inks

For analog signals, we use a quantity called C/N_0 in the same way as E_b/N_o, where C is the signal power. C and E_b are related by the bit rate. So you will typically see C/N_0 specified for the analog portions (or the passband signals) of the link and E_b/N_o for the digital (or the baseband) portions.

C/N is simply the carrier power in the whole useable bandwidth, where C/N_0 is carrier power per unit bandwidth.

Let's relate E_b/N_o to C/N_0 and C/N. From Eq 1, we know that

C = Energy per bit x bit rate = Eb x Rb, from which we get

$$\frac{C}{N_0} = \frac{E_b}{N_0} \times R_b$$

Both C/N_0 and E_b/N_0 are densities so we do not need to specify the bandwidth of the signal. But to convert C/N_0 to C/N, need to divide by the signal bandwidth.

$$\frac{C}{N} = \frac{E_b}{N_0} \times \frac{R_b}{B}$$

In dB, we would write the above equation as

$$\frac{C}{N} = \frac{E_b}{N_0} + R_b - B$$

and C/N_0 similarly is

$$\frac{C}{N_0} = \frac{E_b}{N_0} + R_b$$

The difference between C/N and C/N_0 is then only the bandwidth of the signal. And E_b/N_0 is related to these quantities by the bit rate.

Since we are using either E_b/N_o or C/N_0 as our budgeting quantity, it helps to know how these quantities are impacted by some of the common parameters. We can summarize the effect on E_b/N_o of its various components as

Variable	*Action*	*What happens to signal* E_b/N_o	*What happens to signal* C/N_0
Signal power	increase	increases ⇑	increases ⇑
Total noise	increase	decreases ⇓	decreases ⇓
Bandwidth	increase	increases ⇑	increases ⇑
Bit rate	increase	decreases ⇓	no effect

What is a link?

A link consists of three parts.

1. Transmitter
2. Receiver
3. Media

The very simplest form of a link equation is written as

$P_{received}$ = Power of the transmitter + Gain of the transmitting antenna + Gain of the receiving antenna - Sum of all losses

This equation of course only talks about the signal power. We have not accounted for noise yet. Now let's talk about each of these three items.

Important things about a transmitter

A transmitter receives baseband data, modulates onto a higher frequency carrier, amplifies it and broadcasts it via an antenna. The two main items that are associated with transmitters are

1. Flux Density
2. 2. EIRP

Flux Density

Flux Density is a measure of energy that is available for gathering from a particular source. It is called the *Radio Power of a Source* in Astronomy. The Sun, Moon, and stars all emit Radio Power (Flux Density). The Sun bathes us with Flux density at the rate of 10^{-19} Watts per square ft per unit bandwidth. (ITU however defines this unit bandwidth to be equal to 4 MHz.)

The Flux Density is defined by

$$\psi = \frac{G\,P}{4\pi\,r^2}$$

where G = gain of the transmitting antenna and P = transmitter power in watts.

Figure 2 - A Transmitter consists of an amplifier and an antenna

The amplifier puts out a certain amount of power and the antenna is said to have a particular gain that further amplifies this power. The combination is called the Transmitter. Usually lossy elements such as wires connect these two components in the preferred direction of radiation of the antenna. These losses are included in the quoted EIRP figure for the Transmitter.

EIRP is closely related to the Flux Density. Where Flux Density is energy as measured a distance away from the source, EIRP is a measure only of the transmitted power, sort of like a Wattage rating of appliances which allows you to compare one with another. For a transmitter, this "power rating" called EIRP, is defined as the combination of

EIRP = Power of transmitter x Gain of the antenna

$= P_{amp} \times G_{antenna}$ or in dB,

$EIRP_{ES} = P_{amp} + G_{antenna}$

If you look at the equation for the Flux Density, you will see that EIRP is the numerator. EIRP is an important number for transmitters of all sorts. Spacecraft too are characterized by their EIRP which is usually in the range of about 50 dB.

The Flux density is a measure of the amount of energy that is received at a distance *r* from a transmitter of gain G and transmit power P watts. Just as the power received is a function of the square of the amplitude of the signal, the flux density is a function of the square of the distance.

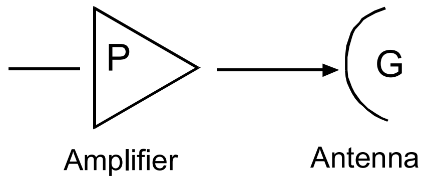
Ex. 3 What is the Flux Density received at a distance of 22,000 km away from this Earth station? (EIRP = 60 dB)

$\psi = 60 - 10 \log (4\pi \ (22{,}000)^2) = -50.2$ dB

What is EIRP?

EIRP is a term closely associated with a radiating source or a transmitter and is a subset of Flux Density.

A very basic transmitter consists of an amplifier and an antenna.



There are two hidden assumption in EIRP. First is that the transmitter is putting out the maximum power that it can, and second, that the EIRP figure is delivered at the antenna's boresight. So if you happen to have your antenna pointed not quite straight into the boresight of the transmitting antenna then you will not get the quoted EIRP.

Ex 4. An Earth station transmits with 10 watts. Its antenna has a gain of 50 dB. (a) What is its EIRP?

EIRP = 10 log 10 + 50 = 60 dBw

(b). Another Earth station transmits with 8 Watts but has an antenna gain of 52 dB. Which of these would you prefer?

EIRP = 10 log 8 + 52 = 61 dBw

Other things being equal the second Earth Station is preferable. It has a better "rating" or the EIRP.

Important thing about Receivers is …

Received Power

EIRP and Flux density both tells us something about a transmitter but nothing about what is actually received. Like two people talking, the listener has to be able to hear well before communication can take place, no matter how loudly the talker talks.

To compute power received by a receiver at a distance *r* from the source, we need to multiply the flux density with the receiving antenna's area. Why? Because, flux density is energy per unit area per unit time. The only useable part of this energy is what is accepted by the receiving antenna.

So the power received is equal to the flux density times the receiving area. We write this as

$$P_{received} = \psi A_{eff}$$

The effective receiving area (not actually a physical area but strongly related to it) of any antenna is defined by

$$A_{eff} = \frac{\lambda^2 G_R}{4\pi}$$

where G_R is the gain of the receiving antenna and λ is the wavelength. Now we can write the expression for computing the received power as

$$P_{received} = \psi A_{eff} = \ = \frac{G_{ES}\ P_{ES}}{4\pi\ r^2} \frac{\lambda^2 G_R}{4\pi}$$

We can rewrite the above in dB as

$$P_{received} = EIRP_{ES} + G_R - 10 \text{ Log}\left(\frac{4\pi r}{\lambda}\right)^2$$

This equation says that if we know the gain of the receiving antenna, the EIRP of the transmitter, the operating frequency, and the distance between the two, then we can calculate the received power. The last portion of the expression above containing the ratio of the distance *r* to the wavelength λ, i.e. the number of wavelengths in the distance, is called the Free Space Loss (FSL).

Noise, noise everywhere

So far we have been talking about signal powers, but now we must jump into a topic that causes a lot of confusion, particularly when tackling link budgets and dealing with noise figures etc. As we can see, our important parameters Eb/N_0, C/N, C/N_0 all have this pesky noise term on the bottom. Let's discuss it in some detail so we can combine all the different ways of defining noise.

All objects not at absolute zero emit electromagnetic radiation. The band of frequencies emitted are a function of the temperature of the object. A light bulb emits many different frequencies owing to the fact that the temperature of the filament is not uniform. However most of its radiation is in the range of infra-red light and ultraviolet frequencies which we can see and feel. The light coming from a light bulb, a jumble of frequencies, is noise that can actually be seen and appreciated.

The sun puts out visible noise in the light wave frequencies among of course many others that we can not see such as X-rays and infra-red. The noise coming to us from the galaxies is typically in microwave frequencies. The moon similarly also bombards us with microwaves.

The statistics of this noise is well described by quantum physics. The black body radiation problem was first solved by Max Planck in 1901.

Max Planck
Born in Germany, Planck is considered the originator of quantum mechanics. Planck studied at Berlin where his teachers included Helmholtz and Kirchhoff. He received his doctorate at the age of 21 in the field of thermodynamics. He taught at the University of Berlin for 38 years until he retired in 1927.

He studied the distribution of energy according to wavelength. By combining the formulas of Wien and Rayleigh, Planck announced in 1900 a formula, now known as Planck's radiation formula. Very quickly he renounced classical physics by introducing the quantum of energy. At first the theory met resistance but due to the successful work of Niels Bohr in 1913, quantizing the angular momentum of the electron and calculating positions of spectral lines using the theory, it became generally accepted.

Planck received the Nobel Prize for Physics in 1918. He remained in Germany during World War II through a time of the deepest personal difficulties. He lost his eldest son during World War I. In World War II, his house in Berlin was burned down in an air raid. In 1945 his other son was executed when declared guilty of complicity in a plot to kill Hitler.

The system containing the noise is modeled as a radiator of energy quanta by Max Planck. One obtains the energy radiated as a function of frequency and temperature, given by the following formula

$$E = hf\left[\frac{1}{e^{\frac{(hf)}{kT}} - 1} + \frac{1}{2}\right]$$

where h is Planck's constant, f is the frequency in Hz, k is Boltzmann's constant, and T is the temperature in degrees Kelvin. For radio, radar, and general microwave frequencies, the factor hf is quite small relative to the factor kT in the nominal range of room temperatures, say 290 Degrees Kelvin, and even down to the range of liquid nitrogen, say 77 degrees Kelvin. Thus the exponential function in the expression can be approximated by the first two terms. When this approximation is made, the denominator of the first term in the energy equation simplifies considerably, resulting in

$$E = kT + \frac{hf}{2}$$

Again applying the approximation $hf \ll kT$, one obtains the well-known result

$$E = kT$$

This is the energy at frequency f, provided f is small enough such that $hf \ll kT$. When f becomes large enough that this approximation no longer holds, the frequencies are in the generalized optical (i.e., infrared, visual optical, ultraviolet, ..., and X-ray) range. In this frequency range, the $hf/2$ term eventually dominates the frequency-dependent noise energy. This term is called by various names, but ***quantum noise*** is probably the most popular. Thus it seems appropriate to refer to the low-frequency noise as thermal noise (it is proportional to the temperature) and the high-frequency noise, from somewhere sub-optical on up, as the quantum noise, since it is proportional to Planck's constant h.

Physicists have shown mathematically that this energy is equivalent in magnitude to the power spectral density of the radiation noise. Thus the thermal noise power in a bandwidth B is just kTB, another of the well-known results from noise theory.

Definition 1:

$$N_0 = \frac{P_N}{B_N}$$

We can define noise density simply as the power of the noise signal divided by its bandwidth. We will worry about how to measure the noise power later.

Definition 2:

$$N_0 = k\,T$$

This definition is consistent with widespread usage, where T is considered to be the system temperature, not the ambient or room temperature (more will be said about this later). In the radio, radar, and microwave bands, the spectral density is taken as N_o for a one-sided spectrum, and as $N_o/2$ for a two-sided spectrum. The noise power will be N_oB in all cases where this convention is applied.

When the system temperature is taken as 290 degrees, the product *kT* (the power spectral density) in dB will be -204 dBW/Hz, or -174 dBm/Hz, or -144 dBW/MHz, or -114 dBm/MHz. Since 290 degrees is considered to be the reference temperature (instead of the system temperature), the value of -114 dBm/MHz is often considered to be a reference level for the power spectral density of systems in general. It is an lower bound for the system temperature, which is set by the amplifier characteristics at some value nominally higher than the reference temperature.

In practice, the system temperature may be (and normally is) entirely fictional in the sense that it cannot be measured with a thermometer. For instance, an amplifier tube may have a system temperature of 580 degrees Kelvin while the room environment for the tube is at 290 degrees Kelvin and the electron-emitting oxide-coated cathode is at 1100 degrees Kelvin. In this case, where would one measure this temperature of 580 degrees? It is not even the average temperature.

The physicist will respond to this question by first admitting that the system temperature is difficult to measure, then explaining that the mean fluctuation kinetic energy of the electrons in the beam where the electrons are coupled to the RF circuit is *kT*, or 580k. This compares to the mean fluctuation kinetic energy of 290k (or so) of the electrons when they are in the conductors of the dc beam circuit, and to the mean fluctuation kinetic energy of 1100k for the electrons as they are emitted from the cathode surface. (Incidentally, this correspondence of temperature with electron kinetic energy can lead to some very interesting observations. We note that the mean free velocity of electrons in space is very large due to reduced electron density; thus their temperature is very high; yet a thermometer determines that space is quite “cold” -- simply because there are so few collisions of electrons with a typical-profile thermometer placed in space.)

This increase in noise level, due to the increased fluctuation kinetic energy of the electrons in the RF circuit, is often quantified by defining a new parameter called the noise figure.

Definition 3:

Noise Figure in dB = System noise-power spectral density in dBW/Hz + 204.

This definition of the noise figure in dB is equivalent to the formula

Noise Figure in dB = 10 log [(System temperature)/290],

which simply shows that system temperature and noise power spectral density in the system are the same number of dB away from the reference temperature and the reference noise-power spectral density, respectively.

Let's now set the first two definitions for noise density, N_0 equal

$$kT = \frac{P_N}{B_N}$$

or we can write

$$P_N = kTB_N$$

here k = Boltzmann's constant = 1.38 x 10^{-23} Joules/Kelvin, T in Kelvins and B_N in Hertz, and P_N in Watts.

Ludwig Boltzmann

Boltzmann was awarded a doctorate from the University of Vienna in 1866 for a thesis on the kinetic theory of gases. After obtaining his doctorate, he became an assistant to his teacher Josef Stefan. Boltzmann taught at Graz, moved to Heidelberg and then to Berlin. In these places he studied under Bunsen, Mach, Kirchhoff and Helmholtz.

Boltzmann it seems was unable to stay at one place very long. He suffered from what is now known as manic-depressive disorder and this effected his work and his career greatly.

Boltzmann's fame is based on his invention of statistical mechanics. This he did independently of Willard Gibbs. Their theories connected the properties and behaviour of atoms and molecules with the large scale properties and behavior of the substances of which they were the building blocks.

Boltzmann worked on statistical mechanics using probability to describe how the properties of atoms determine the properties of matter. In particular his work relates to the Second Law of Thermodynamics which he derived from the principles of mechanics in the 1890s. However, his ideas were generally not accepted by his contemporaries.

In 1904 Boltzmann visited the World's Fair in St Louis, USA. He lectured on applied mathematics and then went on to visit Berkeley and Stanford. Unfortunately he failed to realize that the new discoveries concerning radiation that he learnt about on this visit were about to prove his theories correct.

Attacks on his work continued and he began to feel that his life's work was about to collapse despite his defense of his theories. Depressed and in bad health, Boltzmann committed suicide just before experiments verified his work.

The Bandwidth

What is the noise bandwidth in the equation for noise. In simple terms, it is the noise that is allowed to enter into the system by the receive filter. Is it the same as the 3 dB bandwidth for the signal? It is not. The idea of noise bandwidth is to cover all power allowed-in by the filter, so when we set the two equal we discover that generally the noise bandwidth is about 1.12 times the 3 dB bandwidth of the signal.

There is no confusion when talking about bit rates but unfortunately there are many different ways of defining bandwidth. The most common definition of bandwidth is the distance from one passband edge to another where the edge is defined as the point where the amplitude is 3 dB below the maximum. For the signal below, the 3 dB bandwidth is 36 Hz.

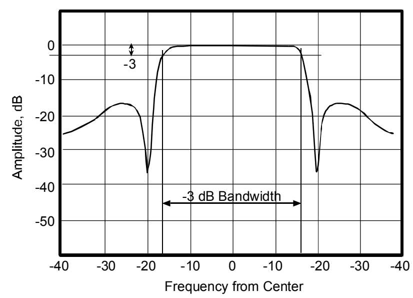


Figure 3 - Bandwidth is often defined as the 3 dB point

Example 4
An antenna has a noise temperature of 70^0K. What is the noise power density? (b) What is the noise power if we assume that the bandwidth is 24 MHz.

(a) $N_0 = k\,T_N = 1.38 \text{ x } 10_{-23} \text{ x } 70 = 9.66 \text{ x } 10\text{-}^{22}$ Joules

(b) $P_N = N_0\,B_N = 9.66 \text{ x } 10\text{-}^{22} \text{ x } 24 \text{ x } 10^6 = 2.3 \text{ x } 10\text{-}^{14}$ Watts

Back to Receivers

Previously we only mentioned how to compute the received power at a receiver. Now that we know about noise, we will discuss a very important parameter of receivers called G/T. The T is the thermal noise temperature of the receiver and impacts the ability of the receiver to "see" a signal in the noise.

G/T of a Receiver

Just as we characterize a transmitter by its EIRP, we use G/T in a similar way to specify receivers. In dB, G/T is the difference in the gain of the receiving antenna gain and its noise temperature.

$$\frac{G}{T} = G_R - T \quad \text{dB K}^{-1}$$

G_R = Gain of the Receiving Antenna

T = Thermal Noise temperature of the receiver

This handy variable allows us to compare receiving systems of all kinds. A G/T of 20 dB is better than 15 dB. A positive number is preferred but some weak ground Earth Stations often have G/T that is negative.

This parameter is usually given for earth stations as well as for satellite receivers, and does not need to be calculated. In doing link budgets, we will assume that it is given. Actually calculating the G/T or the Noise Figure of a receiver is a topic in itself worthy of another ten pages.

Example 5
A satellite receiver has a gain of 30 dB and its antenna temperature is 700K. What is its G/T?

G/T = 30 dB - 10 log 700 = 1.5 dB

The Media

There are basically two types of media, wired or wireless. We are going to discuss only the wireless medium here. Electromagnetic waves travel through the earth's atmosphere in the following four ways.

1. **Ground Wave propagation**

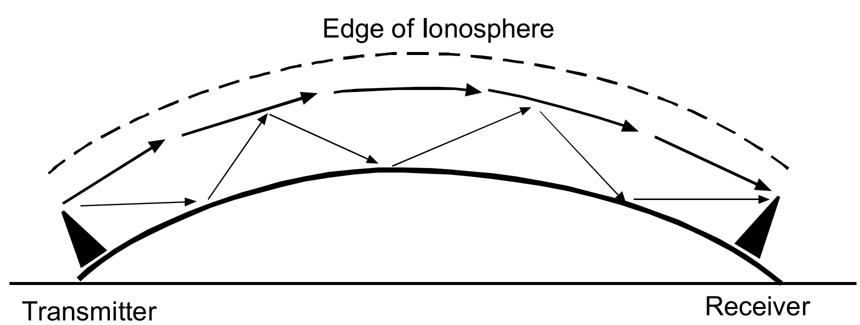


Figure 4 - Ground waves travel parallel to the surface of the earth

Frequencies below 30 MHz propagate along the earth's curvature guided by the surface and are called ground waves. This guided wave has two main components,

1. The direct wave, and
2. The reflected component

AM broadcasting and much of the mobile communications fall in this category. This mode is also called ducting.

2. **Ionospheric Waves**

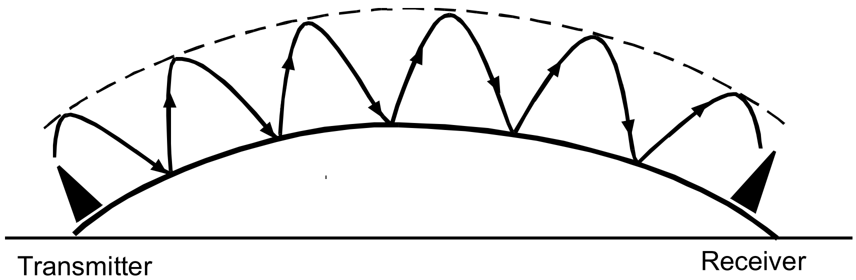


Figure 5 - Waves bounce back and forth between the edge of ionosphere and the earth

Frequencies between 30 and 300 MHz are reflected by the ionosphere and travel much further than ground waves. Examples of this types of propagation are commercial FM broadcasting and VHF TV signal and short wave radio.

3. **Troposheric Scattering**

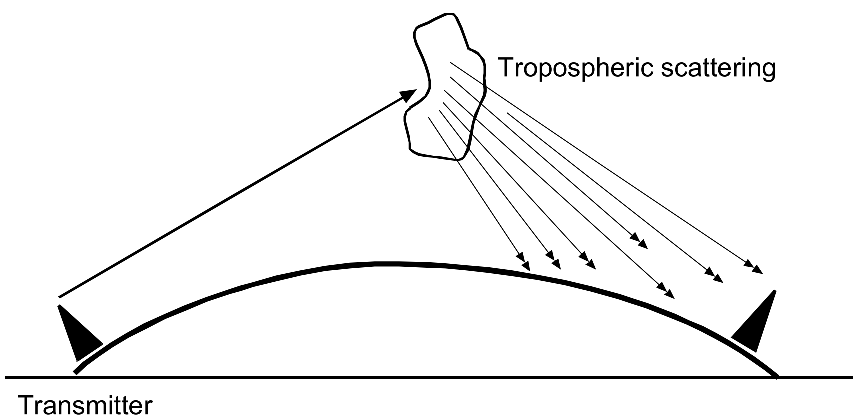


Figure 6 - Waves are scattered by the troposphere

In the range of frequencies above 300 MHz and less than 3 GHz, we see a phenomena where the signals cannot cross the troposphere and are scattered by it. The scattered waves, which are much weaker, can be received and demodulated. This mode of media behavior is called troposhperic scattering.

4. **Line of Sight**

This is happily the mode for satellite communications. For frequencies above 3 GHz to about 12 GHz, the earth's atmosphere offers practically no resistance and degradation. Frequencies above 12 GHz once again suffer from oxygen and water vapor absorption and cause degradation.

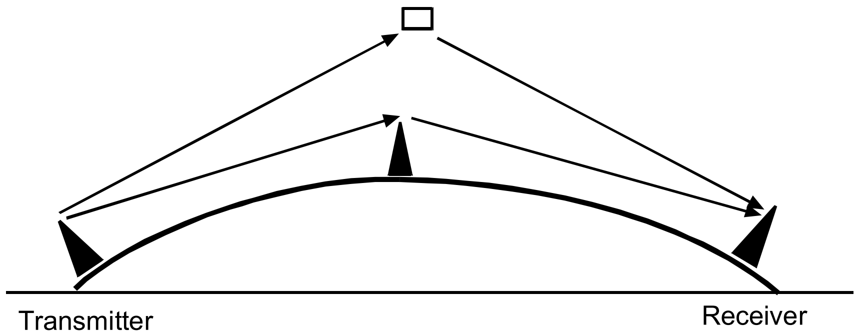


Figure 7 - Line of sight means straight through

Losses experienced in Line of Sight links

The losses experienced by the signal fall into these categories.

- Free Space Loss
- Rain
- Antenna Misalignment
- Gaseous Absorption

Free Space Loss

This is the largest signal energy attenuation as a function of the distance traveled. For line of sight links, this loss is a function of the square of the distance. Radar signals which also fall partly in the line of sight category typically suffer a free space loss which is a function of the cube of the distance traveled.

Computing Free Space Loss

For a signal going from ground to the satellite, the free space loss is largest of all other types of losses. It can be simplified and written as

$$FSL = \left(\frac{4\pi r}{\lambda} \right)^2$$

Simplifying, we can write this in dB form as

Free Space Loss (FSL) = 32.4 + 20 Log (r) + 20 Log (f) where r is the distance and f the frequency.

Ex.6 The Earth Station has an EIRP of 60 dB and the satellite antenna has a gain of 52 dB at 12 GHz. What is the received power at the satellite?

$P_{received} = EIRP_{ES} + G_R - FSL$

$P_{received} = 60\text{ dBW} + 52 - FSL$

$FSL = -92.4 + 20\text{ Log }(22000) + 20\text{ Log }(12 \times 10^9) = 196\text{ dB}$

$P_{received} = 60\text{ dBW} + 52 - FSL = -80\text{ dBW}$

Rain

Signal attenuation due to rain is the second most significant after free space loss. It is particularly significant for frequencies in the Ku and Ka bands. We have to deal with rain losses for both uplinks and downlinks.

It also varies a great deal from location to location since it is a function of the rain rate. The attenuation can vary from .1 dB in California to app. 12 dB in Seattle. Providing for this large attenuation in satellite links results in over design of the system for areas which have little rain.

Accommodations are instead made by providing ground diversity, which just means that there are two receivers instead of one which may or may not be geographically separated. Ground Diversity such as having another ground station located a few miles away in a rainy region can improve the rain attenuation by more than half. Other ways to accommodate for location-specific rain attenuation is to allow higher power for the transmitters and variable error correcting codes and variable data rates.

There many popular rain models that help us compute the rain loss. Some of these are

1. NASA Rain Attenuation Model
2. Crane Rain Attenuation Model
3. CCIR Rain Attenuation Model

The result from each can vary by 1-2 dB depending on how the regions are defined and estimate of rain rate is made.

Antenna misalignment

The antenna gain has played an important role in the above calculations and we have assumed that the receiving antenna and the transmitting antenna are oriented perfectly so that the maximum gain of the receiving antenna is aligned with the uplink. The gain varies a great deal off the bore sight and as shown in the figure below, unless we have perfect alignment, we are going to have losses associated with this.

There are two parts to this loss. One is at the transmitter, if its antenna is not pointed to deliver maximum gain and the second is at the receiving antenna is not pointed to receive the maximum gain. Antenna pointing is a serious business and calibrations are performed when the system is set up.

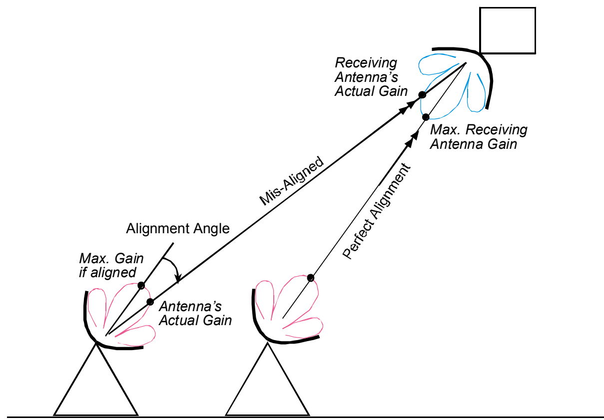


Figure 8 - To receive maximum EIRP, both receiving and transmitting antenna have to be aligned perfectly

Gaseous Absorption

The attenuation caused by clouds and fog due to the phenomena of gaseous absorption and has to be considered particularly for frequencies above 10 GHz. These effects are primarily due to amplitude reduction which reduces signal power, but water can also affect the phase of dual-polarized signals.

For Ku-band and below, these losses are small and can be ignored.

Putting together the link power and noise

Now let's write the link equation in terms of C/N

$C = EIRP + G - Losses$

In above, first write out the carrier power, which is just the sum of the EIRP of the transmitter, the gain of the receiver and any associated implementation losses. So thats all the power that is available to the signal.

Remember we said that

$N = k\ T\ B_N$,

Now divide the above expression for C, with expression for N and write it out in dB form

$$C/N = EIRP + G/T - Losses - k - B_n$$

Now let's convert C/N to C/N_0, we do that by dividing the above by Noise Bandwidth B_n.

$$C/N_0 = C/N \times B_n$$

$$C/N_0 = EIRP_{ES} + G/T_s - \text{Losses} - k - B_n + B_n$$

The bandwidth has dropped out as we said earlier, all terms containing the N_0 term are independent of signal bandwidth. C/N_0 is then a figure that is independent of signal bandwidth.

However, you will hear much more often about C/N (also referred to as SNR) that C/N_0. What is nice about C/N as opposed to C/N_0 is that we can say things like, "I want my CNR to be 15 dB for the uplink and 10 dB for the downlink." Since has a range that applies universally, it makes it easy to compare links.

The figure C/N allows us to compare systems and determine if we have met certain specs and as such is a much more important number than C/N_0.

The downlink is really no different than the uplink except now the role is reversed, the satellite is the transmitter and ground the receiver. So by our previous convention, we now need the EIRP of the satellite and the G/T of the receiving earth. Whereas we were dealing with C/N_0 in the uplink, we now shift to either C/N or E_b/N_o depending on the type of signal. C/N is used for analog signals such as TV-FM, and E_b/N_o for digital data signals. Both of these are derivable from the C/N_0 and we actually continue with C/N_0 until the last calculation before the receiver, when we convert it to E_b/N_o or C/N.

Conversion to E_b/N_o requires knowledge of the symbol/bit rate knowledge and C/N the knowledge of the bandwidth.

Summarizing these conversions below

$$E_b/N_o = C/N_0 - R_b$$

$$C/N = C/N_0 + B_n$$

Symbol rate and Bit rate

An often-missed aspect of link budgets is the distinction between bit and symbol rates. Nowadays nearly all the BER curves are given as a function of the bit energy or E_b/N_0. However, the modem part of the communications systems operate on the basis of symbol rates and not bit rates. Bits are quite transparent to the system. This happens because, we use a symbol as a proxy for a pre-determined number of bits.

In a binary case, a symbol consists of just one bit, so the bit rate of a binary system is equal to the symbol rate. In a QPSK system, one symbol represents two bits, so here the bit rate is twice symbol rate. In 8-PSK system, one symbol stands for three bits, so now the bit rate is three times the symbol rate. When several bits, say three or more are contained in a symbol, the modulation technique is generically referred to as "higher order modulation". The advantage is of these modulations is that higher bit rates can be contained in the same bandwidth as the binary signal.

We can make a symbol stand for as many bits as we want but since we find that the BER increases faster than the bit rate, which requires more power to overcome. If the system is bandwidth-limited and but has plenty of power, then these higher order modulations make sense but for satellites which are power limited (due to their being weight-limited), we do not use them.

The following table summarizes the bit-rate and symbol-rate considerations of the preceding paragraphs. It also shows the degradation in Eb/N0 and C/N which occur for these signals when the target error rate is 10^{-5}.

Signaling method (Modulation)	One symbol represents this many bits	Symbol Rate	Bit rate	E_b/N_0 Degradation	C/N Degradation
BPSK	1	1	1	0	0
QPSK	2	1	2	0	3.01
8-PSK	3	1	3	-4.0	6.01

The important thing to keep in mind is that same number of symbols are going through the communications system in each of these cases and that the communication system is not "working any harder" in a higher modulation.

QPSK is currently the predominant method of satellite transmission. We need to remember that here the bit rate is twice the symbol rate, conversely, the symbol rate is half the bit rate, whichever is easier to remember.

Coding gain

When one bank transmits money over the line to an another bank, they want an "error-free" transmission. Typically a data transmission at about 10^{-11} BER is considered *nearly* error free. To get a picture of 10^{-11} BER, this is equivalent to about one error per ten million pages of text transmitted. Rightfully so, we consider this "quasi error-free" using the official term.

A typical satellite link can provide, for a reasonable power level, only about 10^{-2} BER. This is equal to about ten errors per page. This may be OK if this a voice signal or music but data at this rate is quite unacceptable.

Error correction coding (ERC) is nearly always used on all types of links now. There are many different varieties of ERC suitable for all different types of link. In all cases, the use of coding is a trade-off between power and important real estate, which is bits. The bits used to convey coding can not be used for information, effectively reducing the useable bit rate for a given bandwidth and power level.

Consider this very simple code. Say a group of 7 bits. Of these we assign the last bit as a parity bit. Which means that when the blocks of bits is checked at the receiver, if the first six bits are even (equal number of 0 and 1) then, the parity bit is set to 0, and if odd then it is set to 1. In simplest terms, now we can tell if a block was received erred. So by using a parity bit, we now have a way to tell if the block has a bit error. If the first six bit of the received block have an even number of 1s, as in Figure 9, we know there has been an error somewhere.

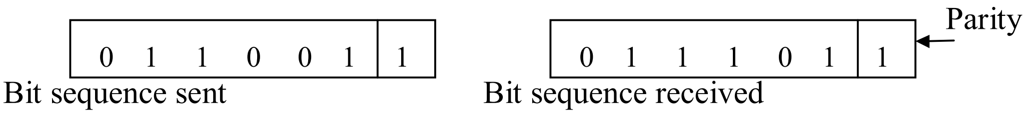


Figure 9 - Example of a simple Linear Block Code

This type of code is called a linear block code because it deals with a group of bits at one time. It is said to have a rate of 6/7, as the code uses 1 out of 7 bits for coding purposes and only 6 out 7 bits can be used for information. In return for an Eb/N_0 of 10 dB, it gives us an approximate BER of 10^{-9} instead of the 10^{-5} we would have gotten otherwise. So in return for a loss of app. 15% of the bit rate to coding overhead, we get an improvement of approximately 4 dB. How did we get this improvement in E_b/N_o?

Let's take a look at the following curve. Three bit error rate curves are shown. The right most curve is for an uncoded QPSK signal, the other two for coded links. First thing we see is that the use of coding moves the BER curve to the left. This is always so or otherwise there would be no point in using coding. What this says is that we can get the same BER but for a lesser E_b/N_0. Of course what it does not say is how this has also reduced our effective data rate. But we won't worry about the reduced data rate, we can live with it.

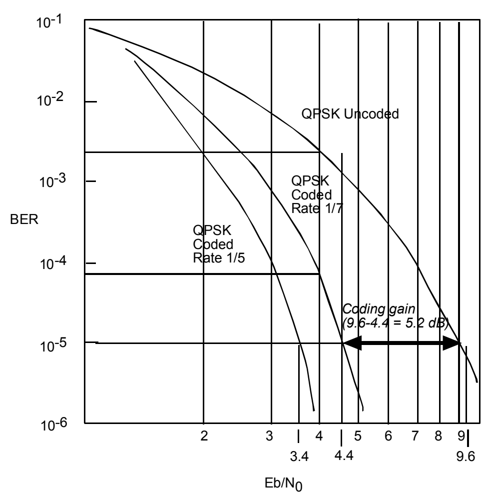


Figure 10 - Coding Gain is the difference between E_b/N_0 needed for an uncoded systems vs. that needed for a coded system

We see that for a desired BER of 10^{-5}, an uncoded QPSK signal requires an E_b/N_0 of 9.6 dB. Now if we add coding, for the first code we need only 4.4 dB and the second code requires only 3.4 dB. We define the coding gain as

Coding Gain = Uncoded E_b/N_o - Coded E_b/N_o for any particular BER.

The coding gain is a function of the BER level chosen. So before we calculate the coding gain we need to know what BER level is desired for the link. Most commonly a BER of 10^{-5} is chosen for voice and 10^{-11} for data links.

Also we always need to be clear about whether the BER quoted is for the uncoded channel or coded channel.

The coded BER curves shown below are usually produced by testing or simulation. There are many published sources for E_b/N_o vs. BER for various types of codes and are used when determining the coding gain.

So here is the tradeoff in using a code

Type of Code	E_b/N_o for 10^{-5}	Coding gain	Loss in Bit rate
Uncoded	9.6	0	0
Code 1	4.4	5.2	15%
Code 2	3.4	6.2	20%

In most cases we sacrifice this loss in useful bit rate for reduction in transmit power. This is true now for all types of links including satellite and wireless phones.

Output and Input Backoff

The Uplink Transmitter has a rated EIRP of a certain level, but if the uplink transmitter transmits instead at a lower power level, then this reduction in power is called the input backoff for the satellite. But if the input power is backed off then obviously the output power would also be backed off, and the maximum EIRP can not be delivered. The reduction in the output power due to the reduction in the input power or input-backoff is called the Output Backoff. The relationship between Input Backoff (IBO) and Outback off (OBO) is not linear and generally looks like the curve of Figure 11.

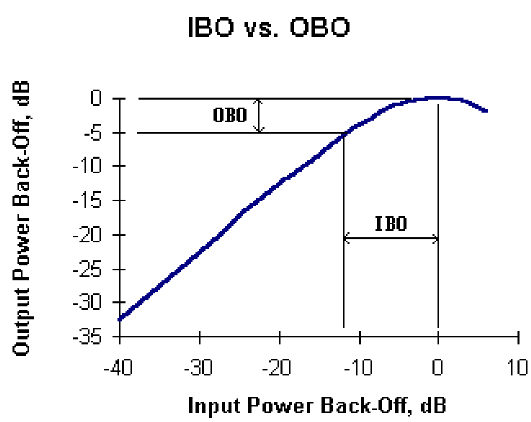


Figure 11 - Operating the transmitter at less than max power on earth or on satellite requires adjustment to the EIRP, called OBO

In the Figure above, we see that if we backoff the input power of the link, then its output power from the TWTA will be reduced by 5 dB. This relationship is important for our link budgets because we do not always operate the links at their maximum power. Backoff is necessary to avoid signal degradation when multi-carriers are present.

The downlink equation has the same components as the uplink.

$C/N_0 = EIRP_{sat} + G/T_{ES}$ - $Losses_d + k$ - OBO

The losses we include in the downlink are also pretty much the same. There is the free-space path loss, there are atmospheric losses due to rain and fog, misalignment losses and antenna noise. And there is OBO depending on the IBO.

In satellite links, the we can have an OBO without having an associated IBO, the reason is that often gain controls are used onboard the satellites and power to the TWTA is always set at a constant. But we can still have an OBO without any IBO.

In summary, then let's write the complete set of link equations.

The uplink

$$\frac{C}{N_0} = EIRP_{ES} + \frac{G}{T}_{Satelliterceiver} - FreeSpaceLos - k - Losses_{rain,misalign} - IBO$$

The downlink

$$\frac{C}{N_0} = EIRP_{Satellit} + \frac{G}{T}_{Ground\,Receiver} - FreeSpaceLos - k - Losses_{rain,misalign} - OBO$$

Conversion to E_b/N_0

$$\frac{E_b}{N_0} = \frac{C}{N_0} - R_b$$

Link Margin

$$M\arg in = \left(\frac{E_b}{N_0}\right)_{available} - \left(\frac{E_b}{N_0}\right)_{requred}$$

$$\left(\frac{E_b}{N_0}\right)_{Required} = \left(\frac{E_b}{N_0}\right)_{FromTheory} - CodingGain \text{ - } Receiver\ implementation\ losses$$

The ultimate goal of our link budget exercise is to have a reasonable margin for the chosen data rate, bandwidth and EIRP and the G/T figures. Often some shifting and adjusting is needed to get the desired link margin.

Other factors such as interference from adjacent channels and adjacent satellites will affect the link budgets, and in another Tutorial we will talk about dealing with C/I and other interferences.

Thanks to Dr. Dave Watson who wrote the noise section and provided most valuable edits. Without him these Tutorials would not read so well (which I hope they do!)

I can be contacted at
mntcastle@earthlink.net

Other tutorials at
www.complextoreal.com