- The ***window scale*** option for supporting large TCP window sizes (larger than 64KB).
- ***Selective acknowledgment*** (***SACK***) to enable faster recovery from transmission errors.
- Timestamps for calculating appropriate RTT and retransmission timeout values for the link in use.

Long round-trip time (RTT)

Satellite links have an average RTT of around 520ms to the first hop. TCP uses the slow-start mechanism at the start of a connection to find the appropriate TCP/IP parameters for that connection. Time spent in the slow-start stage is proportional to the RTT, and for a satellite link it means that TCP stays in slow-start mode for a longer time than would otherwise be the case. This drastically decreases the throughput of short-duration TCP connections. This is can be seen in the way that a small website might take surprisingly long to load, but when a large file is transferred acceptable data rates are achieved after a while.

Furthermore, when packets are lost, TCP enters the congestion-control phase, and owing to the higher RTT, remains in this phase for a longer time, thus reducing the throughput of both short- and long-duration TCP connections.

Large bandwidth-delay product

The amount of data in transit on a link at any point of time is the product of bandwidth and the RTT. Because of the high latency of the satellite link, the bandwidth-delay product is large. TCP/IP allows the remote host to send a certain amount of data in advance without acknowledgment. An acknowledgment is usually required for all incoming data on a TCP/IP connection. However, the remote host is always allowed to send a certain amount of data without acknowledgment, which is important to achieve a good transfer rate on large bandwidth-delay product connections. This amount of data is called the ***TCP window size***. The window size is usually 64KB in modern TCP/IP implementations.

On satellite networks, the value of the bandwidth-delay product is important. To utilize the link fully, the window size of the connection should be equal to the bandwidth-delay product. If the largest window size allowed is 64KB, the maximum theoretical throughput achievable via satellite is (window size) / RTT, or 64KB / 520 ms. This gives a maximum data rate of 123 KB/s, which is 984 kbps, regardless of the fact that the capacity of the link may be much greater.

Each TCP segment header contains a field called ***advertised window***, which specifies how many additional bytes of data the receiver is prepared to accept. The advertised window is the receiver's current available buffer size.

The sender is not allowed to send more bytes than the advertised window. To maximize performance, the sender should set its send buffer size and the receiver should set its receive buffer size to no less than the bandwidth-delay product. This buffer size has a maximum value of 64KB in most modern TCP/IP implementations.

To overcome the problem of TCP/IP stacks from operating systems that don't increase the window size beyond 64KB, a technique known as ***TCP acknowledgment spoofing*** can be used (see Performance Enhancing Proxy, below).

Transmission errors

In older TCP/IP implementations, packet loss is always considered to have been caused by congestion (as opposed to link errors). When this happens, TCP performs congestion avoidance, requiring three duplicate ACKs or slow start in the case of a timeout. Because of the long RTT value, once this congestion-control phase is started, TCP/IP on satellite links will take a longer time to return to the previous throughput level. Therefore errors on a satellite link have a more serious effect on the performance of TCP than over low latency links. To overcome this limitation, mechanisms such as ***Selective Acknowledgment*** (***SACK***) have been developed. SACK specifies exactly those packets that have been received, allowing the sender to retransmit only those segments that are missing because of link errors.

The Microsoft Windows 2000 TCP/IP Implementation Details White Paper states

> "*Windows 2000 introduces support for an important performance feature known as Selective Acknowledgment (SACK). SACK is especially important for connections using large TCP window sizes.*"

SACK has been a standard feature in Linux and BSD kernels for quite some time. Be sure that your Internet router and your ISP's remote side both support SACK.

Implications for universities

If a site has a 512 kbps connection to the Internet, the default TCP/IP settings are likely sufficient, because a 64 KB window size can fill up to 984 kbps. But if the university has more than 984 kbps, it might in some cases not get the full bandwidth of the available link due to the "long fat pipe network" factors discussed above. What these factors really imply is that they prevent a single machine from filling the entire bandwidth. This is not a bad thing during the day, because many people are using the bandwidth. But if, for example, there are large scheduled downloads at night, the administrator might want those downloads to make use of the full bandwidth, and the "long fat pipe network" factors might be an obstacle. This may also become critical if a significant amount of your network traffic routes through a single tunnel or VPN connection to the other end of the VSAT link.

Administrators might consider taking steps to ensure that the full bandwidth can be achieved by tuning their TCP/IP settings. If a university has implemented a network where all traffic has to go through the proxy (enforced by network layout), then the only machines that make connections to the Internet will be the proxy and mail servers.

For more information, see *http://www.psc.edu/networking/perf_tune.html* .

Performance-enhancing proxy (PEP)

The idea of a Performance-enhancing proxy is described in RFC 3135 (see *http://www.ietf.org/rfc/rfc3135*), and would be a proxy server with a large disk cache that has RFC 1323 extensions, among other features. A laptop has a TCP session with the PEP at the ISP. That PEP, and the one at the satellite provider, communicate using a different TCP session or even their own proprietary protocol. The PEP at the satellite provider gets the files from the web server. In this way, the TCP session is split, and thus the link characteristics that affect protocol performance (long fat pipe factors) are overcome (by TCP acknowledgment spoofing, for example). Additionally, the PEP makes use of proxying and pre-fetching to accelerate web access further.

Such a system can be built from scratch using Squid, for example, or purchased "off the shelf" from a number of vendors.

More information

While bandwidth optimization is a complex and often difficult subject, the techniques in this chapter should help reduce obvious sources of wasted bandwidth. To make the best possible use of available bandwidth, you will need to define a good access policy, set up comprehensive monitoring and analysis tools, and implement a network architecture that enforces desired usage limits.

For more information about bandwidth optimization, see the free book *How to Accelerate Your Internet* (*http://bwmo.net/*).

4

Antennas & Transmission Lines

The transmitter that generates the RF[1] power to drive the antenna is usually located at some distance from the antenna terminals. The connecting link between the two is the ***RF transmission line***. Its purpose is to carry RF power from one place to another, and to do this as efficiently as possible. From the receiver side, the antenna is responsible for picking up any radio signals in the air and passing them to the receiver with the minimum amount of distortion, so that the radio has its best chance to decode the signal. For these reasons, the RF cable has a very important role in radio systems: it must maintain the integrity of the signals in both directions.

There are two main categories of transmission lines: cables and waveguides. Both types work well for efficiently carrying RF power at 2.4 GHz.

Cables

RF cables are, for frequencies higher than HF, almost exclusively coaxial cables (or ***coax*** for short, derived from the words "of common axis"). Coax cables have a core ***conductor*** wire surrounded by a non-conductive material called ***dielectric,*** or simply ***insulation***. The dielectric is then surrounded by an encompassing shielding which is often made of braided wires. The dielectric prevents an electrical connection between the core and the shielding. Finally, the coax is protected by an outer casing which is generally made

---

[1] Radio Frequency. See chapter two for discussion of electromagnetic waves.

from a PVC material. The inner conductor carries the RF signal, and the outer shield prevents the RF signal from radiating to the atmosphere, and also prevents outside signals from interfering with the signal carried by the core. Another interesting fact is that high frequency electrical signal always travels along the outer layer of a conductor: the larger the central conductor, the better signal will flow. This is called the "skin effect".

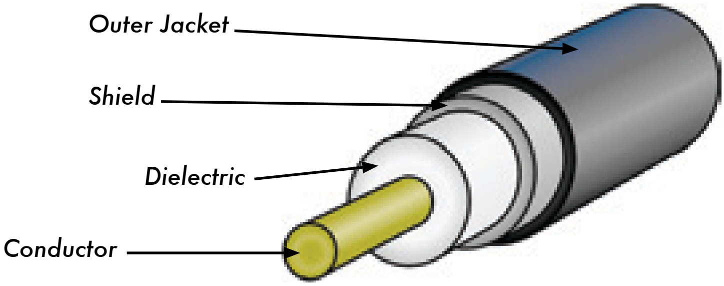

Figure 4.1: Coaxial cable with jacket, shield, dielectric, and core conductor.

Even though the coaxial construction is good at containing the signal on the core wire, there is some resistance to the electrical flow: as the signal travels down the core, it will fade away. This fading is known as ***attenuation***, and for transmission lines it is measured in decibels per meter (***dB/m***). The rate of attenuation is a function of the signal frequency and the physical construction of the cable itself. As the signal frequency increases, so does its attenuation. Obviously, we need to minimize the cable attenuation as much as possible by keeping the cable very short and using high quality cables.

Here are some points to consider when choosing a cable for use with microwave devices:

1. "The shorter the better!" The first rule when you install a piece of cable is to try to keep it as short as possible. The power loss is not linear, so doubling the cable length means that you are going to lose much more than twice the power. In the same way, reducing the cable length by half gives you more than twice the power at the antenna. The best solution is to place the transmitter as close as possible to the antenna, even when this means placing it on a tower.
2. "The cheaper the worse!" The second golden rule is that any money you invest in buying a **good quality** cable is a bargain. Cheap cables are intended to be used at low frequencies, such as VHF. Microwaves require the highest quality cables available. All other options are nothing but a dummy load[2].

---

[2] **2.** A dummy load is a device that dissipates RF energy without radiating it. Think of it as a heat sink that works at radio frequencies.

3. Always avoid RG-58. It is intended for thin Ethernet networking, CB or VHF radio, not for microwave.
4. Always avoid RG-213. It is intended for CB and HF radio. In this case the cable diameter does not imply a high quality, or low attenuation.
5. Whenever possible, use ***Heliax*** (also called "Foam") cables for connecting the transmitter to the antenna. When Heliax is unavailable, use the best rated LMR cable you can find. Heliax cables have a solid or tubular center conductor with a corrugated solid outer conductor to enable them to flex. Heliax can be built in two ways, using either air or foam as a dielectric. Air dielectric Heliax is the most expensive and guarantees the minimum loss, but it is very difficult to handle. Foam dielectric Heliax is slightly more lossy, but is less expensive and easier to install. A special procedure is required when soldering connectors in order to keep the foam dielectric dry and uncorrupted. LMR is a brand of coax cable available in various diameters that works well at microwave frequencies. LMR-400 and LMR-600 are a commonly used alternative to Heliax.
6. Whenever possible, use cables that are pre-crimped and tested in a proper lab. Installing connectors to cable is a tricky business, and is difficult to do properly even with the proper tools. Unless you have access to equipment that can verify a cable you make yourself (such as a spectrum analyzer and signal generator, or time domain reflectometer), troubleshooting a network that uses homemade cable can be difficult.
7. Don't abuse your transmission line. Never step over a cable, bend it too much, or try to unplug a connector by pulling directly the cable. All of those behaviors may change the mechanical characteristic of the cable and therefore its impedance, short the inner conductor to the shield, or even break the line. Those problems are difficult to track and recognize and can lead to unpredictable behavior on the radio link.

Waveguides

Above 2 GHz, the wavelength is short enough to allow practical, efficient energy transfer by different means. A waveguide is a conducting tube through which energy is transmitted in the form of electromagnetic waves. The tube acts as a boundary that confines the waves in the enclosed space. The Faraday cage effect prevents electromagnetic effects from being evident outside the guide. The electromagnetic fields are propagated through the waveguide by means of reflections against its inner walls, which are considered perfect conductors. The intensity of the fields is greatest at the center along the X dimension, and must diminish to zero at the end walls because the existence of any field parallel to the walls at the surface would cause an infinite current to flow in a perfect conductor. Waveguides, of course, cannot carry RF in this fashion.

The X, Y and Z dimensions of a rectangular waveguide can be seen in the following figure:

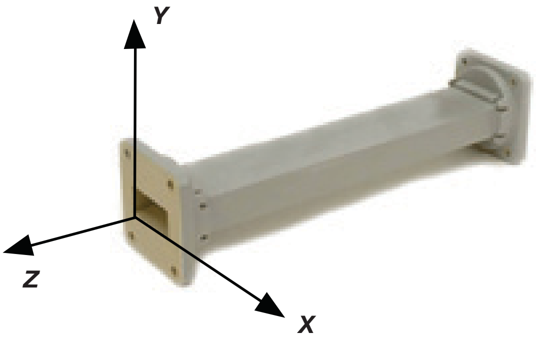

Figure 4.2: The X, Y, and Z dimensions of a rectangular waveguide.

There are an infinite number of ways in which the electric and magnetic fields can arrange themselves in a waveguide for frequencies above the low cutoff frequency. Each of these field configurations is called a ***mode***. The modes may be separated into two general groups. One group, designated ***TM*** (Transverse Magnetic), has the magnetic field entirely transverse to the direction of propagation, but has a component of the electric field in the direction of propagation. The other type, designated ***TE*** (Transverse Electric) has the electric field entirely transverse, but has a component of magnetic field in the direction of propagation.

The mode of propagation is identified by the group letters followed by two subscript numerals. For example, TE 10, TM 11, etc. The number of possible modes increases with the frequency for a given size of guide, and there is only one possible mode, called the ***dominant mode***, for the lowest frequency that can be transmitted. In a rectangular guide, the critical dimension is X. This dimension must be more than 0.5 λ at the lowest frequency to be transmitted. In practice, the Y dimension usually is made about equal to 0.5 X to avoid the possibility of operation in other than the dominant mode. Cross-sectional shapes other than the rectangle can be used, the most important being the circular pipe. Much the same considerations apply as in the rectangular case. Wavelength dimensions for rectangular and circular guides are given in the following table, where X is the width of a rectangular guide and r is the radius of a circular guide. All figures apply to the dominant mode.

Type of guide	Rectangular	Circular
Cutoff wavelength	**2X**	**3.41r**
Longest wavelength transmitted with little attenuation	**1.6X**	**3.2r**
Shortest wavelength before next mode becomes possible	**1.1X**	**2.8r**

Energy may be introduced into or extracted from a waveguide by means of either an electric or magnetic field. The energy transfer typically happens through a coaxial line. Two possible methods for coupling to a coaxial line are using the inner conductor of the coaxial line, or through a loop. A probe which is simply a short extension of the inner conductor of the coaxial line can be oriented so that it is parallel to the electric lines of force. A loop can be arranged so that it encloses some of the magnetic lines of force. The point at which maximum coupling is obtained depends upon the mode of propagation in the guide or cavity. Coupling is maximum when the coupling device is in the most intense field.

If a waveguide is left open at one end, it will radiate energy (that is, it can be used as an antenna rather than as a transmission line). This radiation can be enhanced by flaring the waveguide to form a pyramidal horn antenna. We will see an example of a practical waveguide antenna for WiFi later in this chapter.

Cable Type	Core	Dielectric	Shield	Jacket
RG-58	0.9 mm	2.95 mm	3.8 mm	4.95 mm
RG-213	2.26 mm	7.24 mm	8.64 mm	10.29 mm
LMR-400	2.74 mm	7.24 mm	8.13 mm	10.29 mm
3/8” LDF	3.1 mm	8.12 mm	9.7 mm	11 mm

Here is a table contrasting the sizes of various common transmission lines. Choose the best cable you can afford with the lowest possible attenuation at the frequency you intend to use for your wireless link.

Connectors and adapters

Connectors allow a cable to be connected to another cable or to a component of the RF chain. There are a wide variety of fittings and connectors designed to go with various sizes and types of coaxial lines. We will describe some of the most popular ones.

BNC connectors were developed in the late 40s. BNC stands for Bayonet Neill Concelman, named after the men who invented it: Paul Neill and Carl Concelman. The BNC product line is a miniature quick connect / disconnect connector. It features two bayonet lugs on the female connector, and mating is achieved with only a quarter turn of the coupling nut. BNC's are ideally suited for cable termination for miniature to subminiature coaxial cable (RG-58 to RG-179, RG-316, etc.) They have acceptable performance up to few GHz. They are most commonly found on test equipment and 10base2 coaxial Ethernet cables.

TNC connectors were also invented by Neill and Concelman, and are a threaded variation of the BNC. Due to the better interconnect provided by the threaded connector, TNC connectors work well through about 12 GHz. TNC stands for Threaded Neill Concelman.

Type N (again for Neill, although sometimes attributed to "Navy") connectors were originally developed during the Second World War. They are usable up to 18 Ghz, and very commonly used for microwave applications. They are available for almost all types of cable. Both the plug / cable and plug / socket joints are waterproof, providing an effective cable clamp.

SMA is an acronym for SubMiniature version A, and was developed in the 60s. SMA connectors are precision, subminiature units that provide excellent electrical performance up to 18 GHz. These high-performance connectors are compact in size and mechanically have outstanding durability.

The ***SMB*** name derives from SubMiniature B, and it is the second subminiature design. The SMB is a smaller version of the SMA with snap-on coupling. It provides broadband capability through 4 GHz with a snap-on connector design.

MCX connectors were introduced in the 80s. While the MCX uses identical inner contact and insulator dimensions as the SMB, the outer diameter of the plug is 30% smaller than the SMB. This series provides designers with options where weight and physical space are limited. MCX provides broadband capability though 6 GHz with a snap-on connector design.

In addition to these standard connectors, most WiFi devices use a variety of proprietary connectors. Often, these are simply standard microwave connectors with the center conductor parts reversed, or the thread cut in the opposite direction. These parts are often integrated into a microwave system using a short jumper called a ***pigtail*** that converts the non-standard connector into something more robust and commonly available. Some of these connectors include:

RP-TNC. This is a TNC connector with the genders reversed. These are most commonly found on Linksys equipment, such as the WRT54G.

U.FL (also known as ***MHF***). The U.FL is a patented connector made by Hirose, while the MHF is a mechanically equivalent connector. This is possibly the smallest microwave connector currently in wide use. The U.FL / MHF is typically used to connect a mini-PCI radio card to an antenna or larger connector (such as an N or TNC).

The ***MMCX*** series, which is also called a MicroMate, is one of the smallest RF connector line and was developed in the 90s. MMCX is a micro-miniature connector series with a lock-snap mechanism allowing for 360 degrees rotation enabling flexibility. MMCX connectors are commonly found on PCMCIA radio cards, such as those manufactured by Senao and Cisco.

MC-Card connectors are even smaller and more fragile than MMCX. They have a split outer connector that breaks easily after just a few interconnects. These are commonly found on Lucent / Orinoco / Avaya equipment.

Adapters, which are also called coaxial adapters, are short, two-sided connectors which are used to join two cables or components which cannot be connected directly. Adapters can be used to interconnect devices or cables with different types. For example, an adapter can be used to connect an SMA connector to a BNC. Adapters may also be used to fit together connectors of the same type, but which cannot be directly joined because of their gender.

Figure 4.3: An N female barrel adapter.

For example a very useful adapter is the one which enables to join two Type N connectors, having socket (female) connectors on both sides.

Choosing the proper connector

1. "The gender question." Virtually all connectors have a well defined gender consisting of either a pin (the "male" end) or a socket (the "female" end). Usually cables have male connectors on both ends, while RF devices (i.e. transmitters and antennas) have female connectors. Devices such as directional couplers and line-through measuring devices may have both male and female connectors. Be sure that every male connector in your system mates with a female connector.
2. "Less is best!" Try to minimize the number of connectors and adapters in the RF chain. Each connector introduces some additional loss (up to a few dB for each connection, depending on the connector!)
3. "Buy, don't build!" As mentioned earlier, buy cables that are already terminated with the connectors you need whenever possible. Soldering connectors is not an easy task, and to do this job properly is almost impossible for small connectors as U.FL and MMCX. Even terminating "Foam" cables is not an easy task.
4. Don't use BNC for 2.4 GHz or higher. Use N type connectors (or SMA, SMB, TNC, etc.)
5. Microwave connectors are precision-made parts, and can be easily damaged by mistreatment. As a general rule, you should rotate the outer sleeve to tighten the connector, leaving the rest of the connector (and cable) stationary. If other parts of the connector are twisted while tightening or loosening, damage can easily occur.
6. Never step over connectors, or drop connectors on the floor when disconnecting cables (this happens more often than what you may imagine, especially when working on a mast over a roof).
7. Never use tools like pliers to tighten connectors. Always use your hands. When working outside, remember that metals expand at high temperatures and reduce their size at low temperatures: a very tightened connector in the summer can bind or even break in winter.

Antennas & radiation patterns

Antennas are a very important component of communication systems. By definition, an antenna is a device used to transform an RF signal traveling on a conductor into an electromagnetic wave in free space. Antennas demonstrate a property known as ***reciprocity***, which means that an antenna will maintain the same characteristics regardless if whether it is transmitting or receiving. Most antennas are resonant devices, which operate efficiently over a relatively narrow frequency band. An antenna must be tuned to the same frequency band of the radio system to which it is connected, otherwise the reception and the transmission will be impaired. When a signal is fed into an antenna, the antenna will emit radiation distributed in space in a certain way. A graphical representation of the relative distribution of the radiated power in space is called a ***radiation pattern***.

Antenna term glossary

Before we talk about specific antennas, there are a few common terms that must be defined and explained:

Input Impedance

For an efficient transfer of energy, the ***impedance*** of the radio, antenna, and transmission cable connecting them must be the same. Transceivers and their transmission lines are typically designed for 50Ω impedance. If the antenna has an impedance different than 50Ω, then there is a mismatch and an impedance matching circuit is required. When any of these components are mismatched, transmission efficiency suffers.

Return loss

Return loss is another way of expressing mismatch. It is a logarithmic ratio measured in dB that compares the power reflected by the antenna to the power that is fed into the antenna from the transmission line. The relationship between SWR and return loss is the following:

$$\text{Return Loss (in dB)} = 20\log_{10} \frac{\text{SWR}}{\text{SWR-1}}$$

While some energy will always be reflected back into the system, a high return loss will yield unacceptable antenna performance.

Bandwidth

The ***bandwidth*** of an antenna refers to the range of frequencies over which the antenna can operate correctly. The antenna's bandwidth is the number of Hz for which the antenna will exhibit an SWR less than 2:1.

The bandwidth can also be described in terms of percentage of the center frequency of the band.

$$\text{Bandwidth} = 100 \times \frac{F_H - F_L}{F_C}$$

...where F_H is the highest frequency in the band, F_L is the lowest frequency in the band, and F_C is the center frequency in the band.

In this way, bandwidth is constant relative to frequency. If bandwidth was expressed in absolute units of frequency, it would be different depending upon the center frequency. Different types of antennas have different bandwidth limitations.

Directivity and Gain

Directivity is the ability of an antenna to focus energy in a particular direction when transmitting, or to receive energy from a particular direction when receiving. If a wireless link uses fixed locations for both ends, it is possible to use antenna directivity to concentrate the radiation beam in the wanted direction. In a mobile application where the transceiver is not fixed, it may be impossible to predict where the transceiver will be, and so the antenna should ideally radiate as well as possible in all directions. An omnidirectional antenna is used in these applications.

Gain is not a quantity which can be defined in terms of a physical quantity such as the Watt or the Ohm, but it is a dimensionless ratio. Gain is given in reference to a standard antenna. The two most common reference antennas are the ***isotropic antenna*** and the ***resonant half-wave dipole antenna***. The isotropic antenna radiates equally well in all directions. Real isotropic antennas do not exist, but they provide useful and simple theoretical antenna patterns with which to compare real antennas. Any real antenna will radiate more energy in some directions than in others. Since antennas cannot create energy, the total power radiated is the same as an isotropic antenna. Any additional energy radiated in the directions it favors is offset by equally less energy radiated in all other directions.

The gain of an antenna in a given direction is the amount of energy radiated in that direction compared to the energy an isotropic antenna would radiate in the same direction when driven with the same input power. Usually we are only interested in the maximum gain, which is the gain in the direction in which the antenna is radiating most of the power. An antenna gain of 3 dB compared to an isotropic antenna would be written as **3 dBi**. The resonant half-wave dipole can be a useful standard for comparing to other antennas at one frequency or over a very narrow band of frequencies. To compare the dipole to an antenna over a range of frequencies requires a number of dipoles of different lengths. An antenna gain of 3 dB compared to a dipole antenna would be written as **3 dBd**.

The method of measuring gain by comparing the antenna under test against a known standard antenna, which has a calibrated gain, is technically known as a ***gain transfer*** technique. Another method for measuring gain is the 3 antennas method, where the transmitted and received power at the antenna terminals is measured between three arbitrary antennas at a known fixed distance.

Radiation Pattern

The ***radiation pattern*** or ***antenna pattern*** describes the relative strength of the radiated field in various directions from the antenna, at a constant distance. The radiation pattern is a reception pattern as well, since it also describes the receiving properties of the antenna. The radiation pattern is three-dimensional, but usually the measured radiation patterns are a two-dimensional slice of the three-dimensional pattern, in the horizontal or vertical planes. These pattern measurements are presented in either a ***rectangular*** or a ***polar*** format. The following figure shows a rectangular plot presentation of a typical ten-element Yagi. The detail is good but it is difficult to visualize the antenna behavior in different directions.

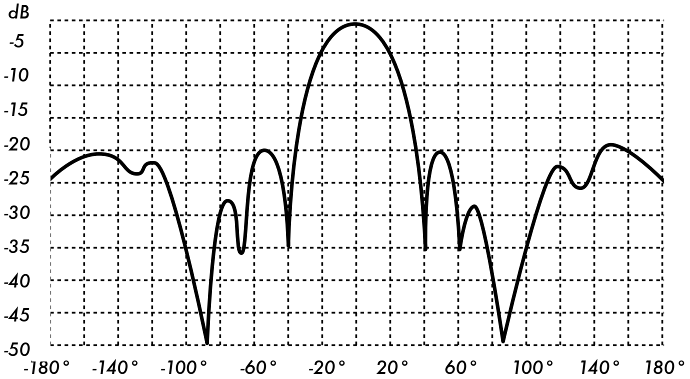

Figure 4.4: A rectangular plot of a yagi radiation pattern.

Polar coordinate systems are used almost universally. In the polar-coordinate graph, points are located by projection along a rotating axis (radius) to an intersection with one of several concentric circles. The following is a polar plot of the same 10 element Yagi antenna.

Polar coordinate systems may be divided generally in two classes: ***linear*** and ***logarithmic***. In the linear coordinate system, the concentric circles are equally spaced, and are graduated. Such a grid may be used to prepare a linear plot of the power contained in the signal. For ease of comparison, the equally spaced concentric circles may be replaced with appropriately placed circles representing the decibel response, referenced to 0 dB at the outer edge of the plot. In this kind of plot the minor lobes are suppressed. Lobes with peaks more than 15 dB or so below the main lobe disappear because of their small size. This grid enhances plots in which the antenna has a high directivity and small minor lobes. The voltage of the signal, rather than the power, can also be plotted on a linear coordinate system. In this case, too, the directivity is enhanced and the minor lobes suppressed, but not in the same degree as in the linear power grid.

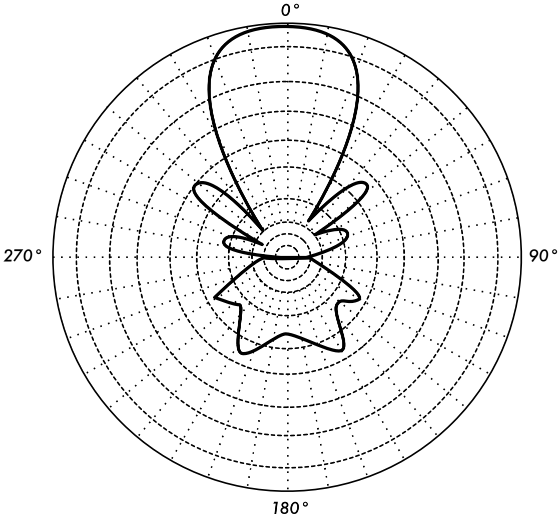

Figure 4.5: A linear polar plot of the same yagi.

In the logarithmic polar coordinate system the concentric grid lines are spaced periodically according to the logarithm of the voltage in the signal. Different values may be used for the logarithmic constant of periodicity, and this choice will have an effect on the appearance of the plotted patterns. Generally the 0 dB reference for the outer edge of the chart is used. With this type of grid, lobes that are 30 or 40 dB below the main lobe are still distinguishable. The spacing between points at 0 dB and at -3 dB is greater than the spacing between -20 dB and -23 dB, which is greater than the spacing between -50 dB and -53 dB. The spacing thus correspond to the relative significance of such changes in antenna performance.

A modified logarithmic scale emphasizes the shape of the major beam while compressing very low-level (>30 dB) sidelobes towards the center of the pattern. This is shown in **Figure 4.6**.

There are two kinds of radiation pattern: ***absolute*** and ***relative***. Absolute radiation patterns are presented in absolute units of field strength or power. Relative radiation patterns are referenced in relative units of field strength or power. Most radiation pattern measurements are relative to the isotropic antenna, and the gain transfer method is then used to establish the absolute gain of the antenna.

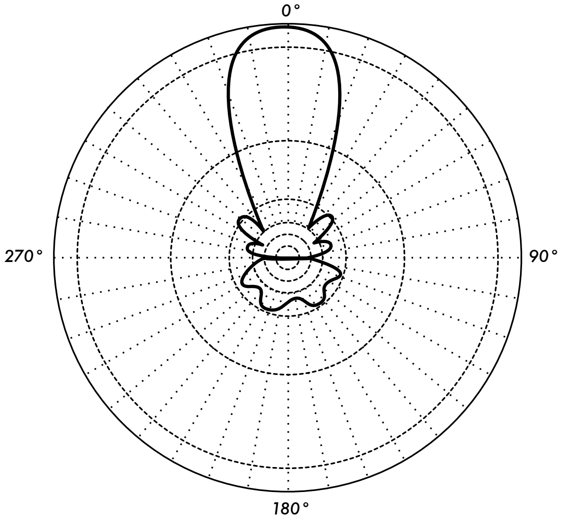

Figure 4.6: The logarithmic polar plot

The radiation pattern in the region close to the antenna is not the same as the pattern at large distances. The term near-field refers to the field pattern that exists close to the antenna, while the term far-field refers to the field pattern at large distances. The far-field is also called the radiation field, and is what is most commonly of interest. Ordinarily, it is the radiated power that is of interest, and so antenna patterns are usually measured in the far-field region. For pattern measurement it is important to choose a distance sufficiently large to be in the far-field, well out of the near-field. The minimum permissible distance depends on the dimensions of the antenna in relation to the wavelength. The accepted formula for this distance is:

$$r_{min} = \frac{2d^2}{\lambda}$$

where r_{min} is the minimum distance from the antenna, d is the largest dimension of the antenna, and λ is the wavelength.

Beamwidth

An antenna's ***beamwidth*** is usually understood to mean the half-power beamwidth. The peak radiation intensity is found, and then the points on either side of the peak which represent half the power of the peak intensity are located. The angular distance between the half power points is defined as the beamwidth. Half the power expressed in decibels is -3dB, so the half power beamwidth is sometimes referred to as the 3dB beamwidth. Both horizontal and vertical beamwidth are usually considered.

Assuming that most of the radiated power is not divided into sidelobes, then the directive gain is inversely proportional to the beamwidth: as the beamwidth decreases, the directive gain increases.

Sidelobes

No antenna is able to radiate all the energy in one preferred direction. Some is inevitably radiated in other directions. These smaller peaks are referred to as ***sidelobes***, commonly specified in dB down from the main lobe.

Nulls

In an antenna radiation pattern, a ***null*** is a zone in which the effective radiated power is at a minimum. A null often has a narrow directivity angle compared to that of the main beam. Thus, the null is useful for several purposes, such as suppression of interfering signals in a given direction.

Polarization

Polarization is defined as the orientation of the electric field of an electromagnetic wave. Polarization is in general described by an ellipse. Two special cases of elliptical polarization are ***linear polarization*** and ***circular polarization***. The initial polarization of a radio wave is determined by the antenna.

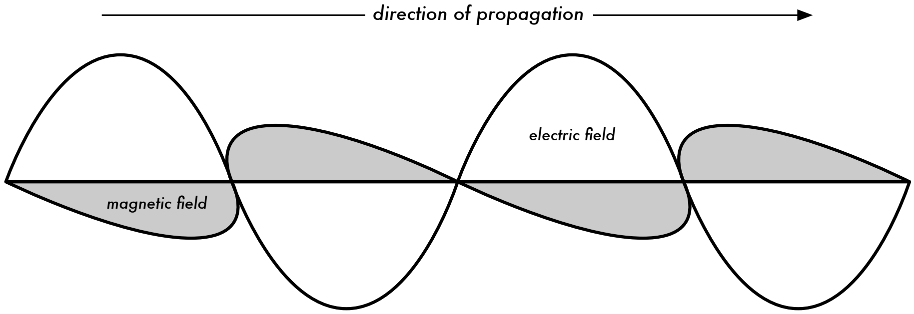

Figure 4.7: The electrical wave is perpendicular to magnetic wave, both of which are perpendicular to the direction of propagation.

With linear polarization, the electric field vector stays in the same plane all the time. The electric field may leave the antenna in a vertical orientation, a horizontal orientation, or at some angle between the two. ***Vertically polarized*** radiation is somewhat less affected by reflections over the transmission path. Omnidirectional antennas always have vertical polarization. With ***horizontal polarization***, such reflections cause variations in received signal strength. Horizontal antennas are less likely to pick up man-made interference, which ordinarily is vertically polarized.

In circular polarization the electric field vector appears to be rotating with circular motion about the direction of propagation, making one full turn for each RF cycle. This rotation may be right-hand or left-hand. Choice of polarization is one of the design choices available to the RF system designer.

Polarization Mismatch

In order to transfer maximum power between a transmit and a receive antenna, both antennas must have the same spatial orientation, the same polarization sense, and the same axial ratio.

When the antennas are not aligned or do not have the same polarization, there will be a reduction in power transfer between the two antennas. This reduction in power transfer will reduce the overall system efficiency and performance.

When the transmit and receive antennas are both linearly polarized, physical antenna misalignment will result in a polarization mismatch loss, which can be determined using the following formula:

$$\text{Loss (dB)} = 20 \log (\cos \theta)$$

...where θ is the difference in alignment angle between the two antennas. For 15° the loss is approximately 0.3dB, for 30° we lose 1.25dB, for 45° we lose 3dB and for 90° we have an infinite loss.

In short, the greater the mismatch in polarization between a transmitting and receiving antenna, the greater the apparent loss. In the real world, a 90° mismatch in polarization is quite large but not infinite. Some antennas, such as yagis or can antennas, can be simply rotated 90° to match the polarization of the other end of the link. You can use the polarization effect to your advantage on a point-to-point link. Use a monitoring tool to observe interference from adjacent networks, and rotate one antenna until you see the lowest received signal. Then bring your link online and orient the other end to match polarization. This technique can sometimes be used to build stable links, even in noisy radio environments.

Front-to-back ratio

It is often useful to compare the ***front-to-back ratio*** of directional antennas. This is the ratio of the maximum directivity of an antenna to its directivity in the opposite direction. For example, when the radiation pattern is plotted on a relative dB scale, the front-to-back ratio is the difference in dB between the level of the maximum radiation in the forward direction and the level of radiation at 180 degrees.

This number is meaningless for an omnidirectional antenna, but it gives you an idea of the amount of power directed forward on a very directional antenna.

Types of Antennas

A classification of antennas can be based on:

- **Frequency and size**. Antennas used for HF are different from antennas used for VHF, which in turn are different from antennas for microwave. The wavelength is different at different frequencies, so the antennas must be different in size to radiate signals at the correct wavelength. We are particularly interested in antennas working in the microwave range, especially in the 2.4 GHz and 5 GHz frequencies. At 2.4 GHz the wavelength is 12.5 cm, while at 5 GHz it is 6 cm.
- **Directivity.** Antennas can be omnidirectional, sectorial or directive. ***Omnidirectional antennas*** radiate roughly the same pattern all around the antenna in a complete 360° pattern. The most popular types of omnidirectional antennas are the ***dipole*** and the ***ground plane***. ***Sectorial antennas*** radiate primarily in a specific area. The beam can be as wide as 180 degrees, or as narrow as 60 degrees. ***Directional*** or ***directive antennas*** are antennas in which the beamwidth is much narrower than in sectorial antennas. They have the highest gain and are therefore used for long distance links. Types of directive antennas are the Yagi, the biquad, the horn, the helicoidal, the patch antenna, the parabolic dish, and many others.
- **Physical construction.** Antennas can be constructed in many different ways, ranging from simple wires, to parabolic dishes, to coffee cans.

When considering antennas suitable for 2.4 GHz WLAN use, another classification can be used:

- **Application**. Access points tend to make point-to-multipoint networks, while remote links are point-to-point. Each of these suggest different types of antennas for their purpose. Nodes that are used for multipoint access will likely use omni antennas which radiate equally in all directions, or sectorial antennas which focus into a small area. In the point-to-point case, antennas are used to connect two single locations together. Directive antennas are the primary choice for this application.

A brief list of common type of antennas for the 2.4 GHz frequency is presented now, with a short description and basic information about their characteristics.