The Basics of Patch Antennas, Updated

By D. Orban and G.J.K. Moernaut,
Orban Microwave Products
www.orbanmicrowave.com

Introduction

This article introduces the basic concepts of patch antennas. We use a simple rectangular, half wave long, probe-fed patch operating in its fundamental mode as an example.

Topics include principles of operation, impedance matching, radiation patterns, circular polarization, bandwidth, efficiency, alternative feed types, stacked patches and higher mode behavior.

This article was originally published in September 2005. After it was published, we received substantial feedback from those who read it. The article has been revised based upon the feedback we received.

Properties Of A Basic Microstrip Patch

A microstrip or patch antenna is a low-profile antenna that has a number of advantages over other antennas: it is lightweight, inexpensive, and electronics like LNA's and SSPA's can be integrated with these antennas quite easily. While the antenna can be a 3-D structure (wrapped around a cylinder, for example), it is usually flat and that is why patch antennas are sometimes referred to as planar antennas.

Figure 1 shows a patch antenna in its basic form: a flat plate over a ground plane. This antenna is often built of printed circuit board material and the substrate makes up the patch antenna's dielectric. The distance between the patch and the ground plane – the substrate or dielectric height h – determines the bandwidth.

A thicker substrate increases the gain to some extent, but may lead to undesired effects like surface wave excitation: surface waves decrease efficiency and perturb the radiation pattern.

As seen in the September 29, 2009 edition of the RF Globalnet (www.rfglobalnet.com) newsletter.

The ground plane should extend beyond the edges of the patch by at least 2 to 3 times the board thickness for proper operation. A ground plane that is too small will result in a reduced front to back ratio.

Making the ground plane larger also increases the gain, but as the ground plane size increases, diffraction near the edges plays less of a role and increasing the size of an already "large" ground plane has very little effect on gain.

In the antenna in *Figure 1*, the center conductor of a coaxial line serves as the feed probe to couple electromagnetic energy in and/or out of the patch. A thicker substrate leads to a longer feed probe, a larger feed probe inductance and a degradation of impedance matching. This can be compensated by using a different feed type and we'll look at alternative feed methods further down.

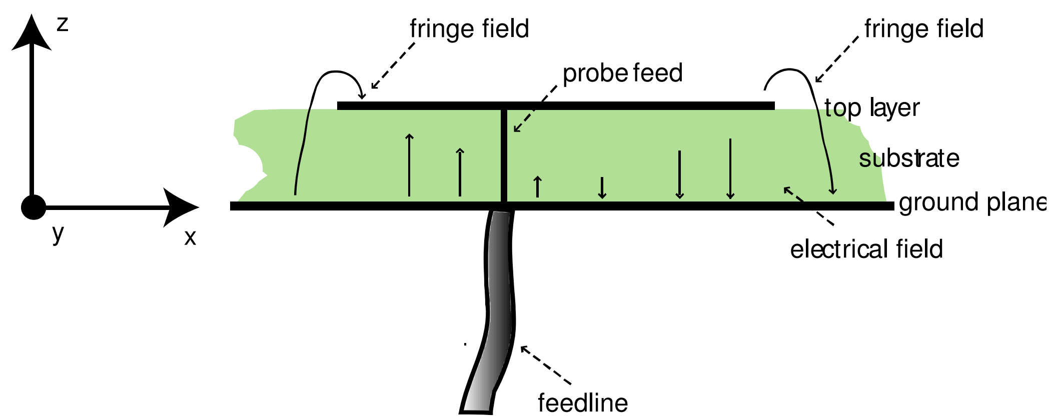

Figure 1: *Cross section of a patch antenna in its basic form*

A half wave long patch operates in what we call the fundamental mode: the electric field is zero at the center of the patch, maximum (positive) on one side, and minimum (negative) on the opposite side. These minima and maxima continuously change side like the phase of the RF signal.

The electric field does not stop abruptly near the patch's edges like it would in a cavity: the field extends beyond the outer periphery. These field extensions are known as fringing fields and cause the patch to radiate. Some popular analytic modeling techniques for patch antennas are based on this leaky-cavity concept and the fundamental mode of a rectangular patch is often denoted using cavity theory like the TM10 mode.

This TM notation often leads to confusion and here is an attempt to explain that:
Figure 1 uses a Cartesian coordinate system, where the x and y axes are parallel with the ground-plane and the z-axis is perpendicular to it.

TM stands for a magnetic field distribution –between patch and ground– that is transverse to the z-axis of the antenna shown in *Figure 1*. This 'transverse' with respect to the z-axis is usually dropped because the magnetic fields in patch antennas are always transverse to their z-axis.

So, we can simplify things and only consider three field components instead of six (magnetic and electric fields in each x, y and z axis): the electric field in the z direction, and the magnetic field components in x and y directions.

In general, modes are designated as TMnmp. The 'p' value is mostly omitted because the electric field variation is considered negligible in the z-axis since only a phase variation exists in the z axis. So, TMnm represents the field variations in the x and y directions. The field variation in the y direction (impedance width direction) is negligible and m is 0. The field has one minimum-to-maximum variation in the x direction (resonance length direction and a half wave long), n is 1 in this case and we say that this patch operates in the TM10 mode.

Dimensions

The resonant length (the x axis in *Figure 2*) determines the resonant frequency and is about $\lambda_d/2$ for a rectangular patch excited in its fundamental mode where λ_d is the wavelength in the PCB material. The patch is actually a bit larger electrically than its physical dimensions due to the fringing fields and the difference between electrical and physical size is mainly dependent on the PC board thickness and dielectric constant of the substrate.

A good approximation for the resonant length is:

$$L \approx 0.49\,\lambda_d = 0.49\,\frac{\lambda_0}{\sqrt{\varepsilon_r}}.$$

This formula includes a first order correction for the edge extension due to the fringing fields, with:

- L = resonant length
- λ_d = wavelength in PC board
- λ_0 = wavelength in free space
- ε_r = dielectric constant of the printed circuit board material

Other parameters that have less influence on the resonant frequency include:

- Ground plane size
- Metal (copper) and dielectric thickness
- Patch (impedance) width

Impedance Matching
The feed position of a patch antenna excited in its fundamental mode is typically located in the center of the patch width direction (y axis) and somewhere along the patch resonant length direction (x axis). The exact position along the resonant length is determined by the electromagnetic field distribution in the patch. Looking at the current (magnetic field) and voltage (electric field) variation along the patch, the current has a *maximum* at the center and a *minimum* near the left and right edges, while the electric field is *zero* in the center and *maximum* near the left and *minimum* near the right edges. Keep in mind that the field distribution constantly changes in amplitude and sign. *Figures 2 and 3* below clarify this:

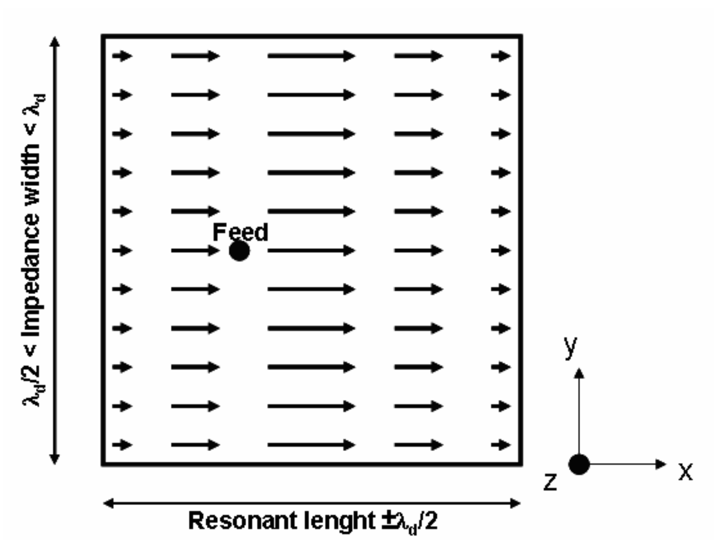

***Figure 2:** Current distribution on the patch surface*

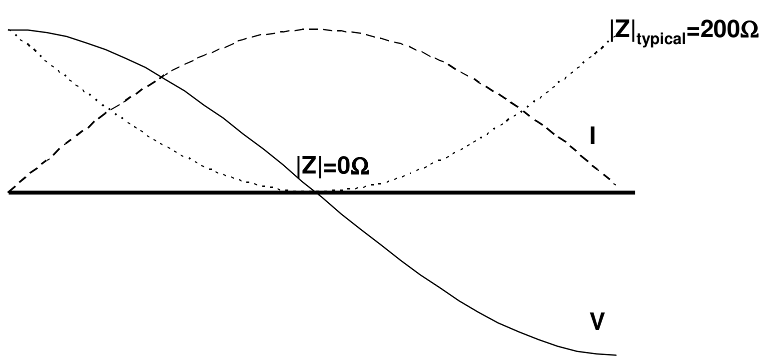

***Figure 3:** Voltage (V), current (I) and impedance (|Z|) distribution along the patch's resonant length*

From the magnitude of the current and the voltage, we can determine that the impedance is minimum (theoretically zero Ω) in the center of the patch and maximum (typically a couple hundred Ω) near the edges. This means that there are two points where the impedance is 50 Ω somewhere along the resonant length (x) axis of the element and this is where you would typically connect to the antenna.

The possibility to connect to the patch at other impedance points is quite useful and impedances up to 200 Ω are common. For example, a two element array can be fed with a simple parallel feed by matching the individual patch elements to 100 Ω and connecting them in parallel results in 50 Ω end impedance without the need for impedance transformers. The same woks for a 4 element array with the elements connected at their 200 Ω point.

If you wanted to connect to the edge of the patch and were looking for a specific impedance, you could modify the width of the patch to yield the impedance your are looking for. Increasing the width decreases the impedance.

Fundamental Specifications Of Patch Antennas

Radiation Pattern

A patch antenna radiates power in certain directions and we say that the antenna has directivity (usually expressed in dBi). If the antenna had a 100% radiation efficiency, all directivity would be converted to gain. Typical half wave patches have efficiencies well above 90%.

The directivity of a patch can be estimated quite easily:

The radiating edges of a patch can be seen as two radiating slots placed above a ground-plane and, assuming all radiation occurs in one half of the hemisphere (on the patch side of the ground), we get a 3 dB directivity increase. This would be an antenna with a perfect front-to-back ratio where all radiation occurs towards the front and no radiation towards the back. This front-to-back ratio is highly dependent on ground-plane size and shape in real life.

Another 3 dB can be added because there are 2 slots. The length of these slots typically equals the impedance width (length in the y-axis) of the patch and the width of these slots equals the substrate height. These slots typically have a directivity of 2 to 3 dB compared to an isotropic radiator and behave like a dipole.

All of this results in a total maximum directivity of 8 to 9 dBi.

The rectangular patch excited in its fundamental mode has a maximum directivity in the direction perpendicular to the patch (z-axis or broadside). The directivity decreases when moving away from broadside towards lower elevations. The 3 dB beamwidth is the width at which the gain of the beam decreases by 3 dB relative to the gain in broadside to either side of the main beam.

Figure 4 shows a typical radiation pattern for a square, half wave patch.

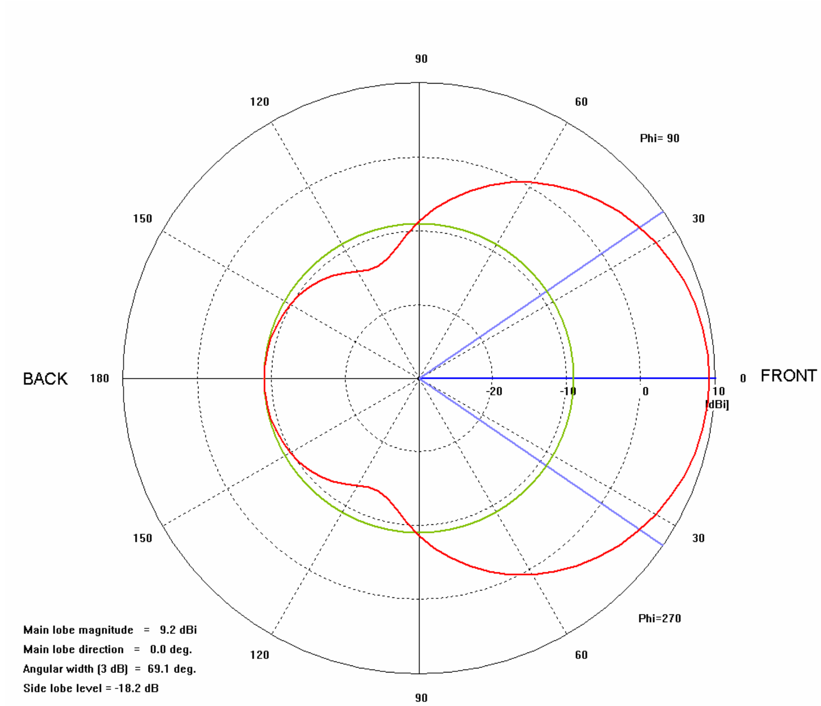

Figure 4: *Typical radiation pattern of a simple square patch*

So far, the directivity has been defined relative to an isotropic radiator and we use dB*i*. An isotropic radiator emits an equal amount of power in all directions and it has no directivity. Antenna directivity can also be specified relative to that of a dipole. A dipole has 2.15 dB*i* of directivity over an isotropic radiator. When we specify the directivity of an antenna relative to a dipole, we use dB*d*.

No antenna losses have been included so far and the integrated average of the directivity pattern over an entire sphere has to be 0 dBi. This implies that creating directivity in a certain direction reduces directivity in other directions.

Antenna Gain

Antennas do not have gain because they are passive structures. Antenna gain is defined as antenna directivity times a factor representing the radiation efficiency. Radiation efficiency is always lower than 100% so the antenna gain is always lower than antenna directivity. This efficiency quantifies the losses in the antenna and is defined as the ratio of radiated power (Pr) to input power (Pi). The input power is transformed into radiated power, surface wave power and a small portion is dissipated due to conductor and dielectric losses. Surface waves are guided waves captured within the substrate and partially radiated and reflected back at the substrate edges. Surface waves are more easily excited when materials with higher dielectric constants and/or thicker materials are used. Surface waves are not excited when air dielectric is used. Several techniques to prevent surface wave excitation exist, but this is beyond the scope of this article.

Antenna gain can also be specified using the total efficiency rather than just the radiation efficiency. This total efficiency is a combination of the radiation efficiency and efficiency linked to the impedance matching of the antenna.

Polarization

The plane in which the electric field varies is also known as the polarization plane. The basic patch covered so far is linearly polarized since the electric field varies in only one direction. This polarization can be anything between vertical and horizontal depending on the orientation of the patch. The polarization plane is the xz-plane in *Figure 1*. For optimum system performance, transmit and receive antennas must have the same polarization. The patch described above yields horizontal polarization and when rotated by 90°, the current flows in the vertical plane and the antenna is now vertically polarized.

A large number of applications like satellite communications, do not work well with linear polarization because the relative orientation of the antennas is unknown and because of Faraday rotation. In these applications, circular polarization is useful since it is not sensitive to antenna orientation. In a circularly polarized antenna, the electric field varies in two orthogonal planes (x and y direction) with the same magnitude and a 90° phase difference. The result is the simultaneous excitation of two modes, i.e. the TM10 mode (x direction) and the TM01 mode (y direction). One of the modes is excited with a 90° phase delay with respect to the other mode. A circularly polarized antenna can either be right-hand circular polarized (RHCP) or left-hand circular polarized (LHCP). The antenna is RHCP when the phases are 0° and -90° for the antenna in *Figure 5* when it radiates towards the reader, and it is LHCP when the phases are 0° and +90°.

To excite circular polarization in a patch we need to do three things:

- Split the signal in two equal parts.
- Feed one signal to the horizontal radiator (x axis) and the other to the vertical radiator (y axis). Each radiator behaves like a pair of radiating slots in the patch antenna as shown in *Figure 5*.
- Change the phase of one of the signals by 90°.

Splitting the signal in half can be done with a Wilkinson power divider or other splitter. If a square patch is fed with two feed points like in *Figure 5* and a 90° delay is added to one of the signal lines, a circularly polarized antenna is built.

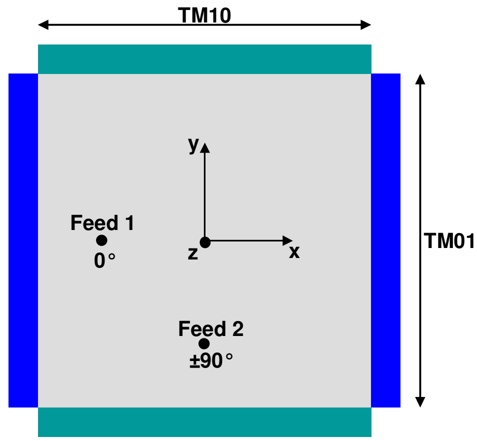

Figure 5: *The nearly square mechanism for circular polarization*

While this works well, the splitter and delay line take up valuable board space, introduce losses, tend to radiate and may degrade the radiation pattern.

A different approach is based on the fact that each patch mode (TM10, TM01) behaves like a parallel RLC resonant circuit. If we make the x and y dimensions of the patch slightly different, there will be two resonant frequencies, fa and fb and we will see two different RLC circuits. This creates a phase shift change versus frequency as shown in *Figure 6*:

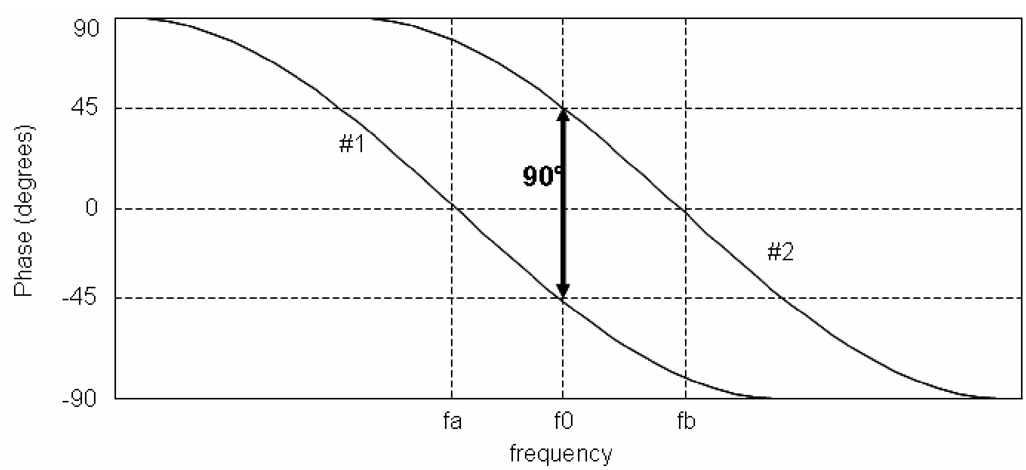

Figure 6: *The dual resonant circuit analogy of the nearly square mechanism*

If we pick these two resonance frequencies right, there will be a small frequency band where the phase difference between the two RLC circuits is 90°.

So, circular polarization can be achieved by building a patch with two resonance frequencies in the orthogonal directions and using the antenna right in between the two resonances at f0. The two modes must be excited with equal power and with a 90° phase difference.

One way to implement this is shown in *Figure 7*: this nearly square patch has slightly different lengths in the x and y axis. This causes the excitation of two orthogonal modes with the required 90° phase shift. Feeding the antenna close to the diagonal splits the power equally between the two orthogonal modes and all three requirements to generate circular polarization have been met.

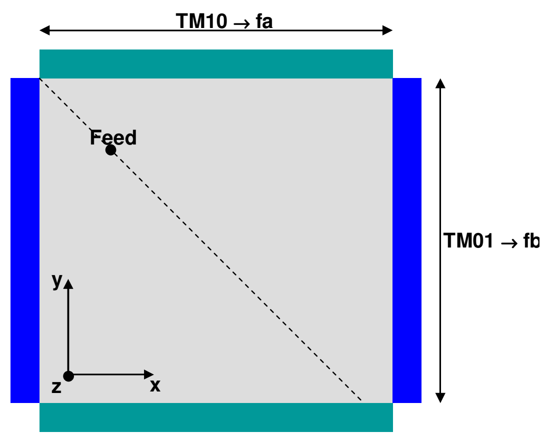

Figure 7: *Nearly square patch fed on the diagonal*

The corners truncated patch (often used in GPS L1 antennas), is another implementation of the same idea. This is shown in *Figure 8*.

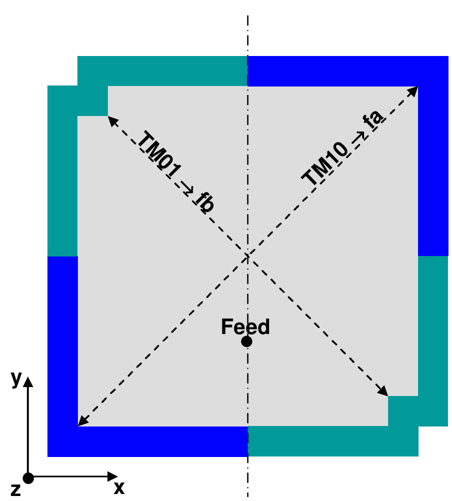

Figure 8: *The corners truncated patch antenna*

The nearly square and corners truncated patches have a smaller (1 to 3%) circular polarization bandwidth than the double fed patch (at least 3%): the polarization bandwidth of the antenna is mainly determined by the bandwidth of the splitter-phase shifter.

The quality of the circular polarization is commonly quantified as the axial ratio (AR) and is expressed in dB. A 3 dB axial ratio is considered sufficient for most applications. Since the phase difference between modes varies with frequency, it is clear that the axial ratio also varies with frequency: it has a theoretical optimum of 0 dB right in between the resonance frequencies of the two orthogonal modes. Not shown so far is that the axial ratio varies with elevation as well: the AR is optimal in broadside (z-axis), degrades towards lower elevations (the x-y plane) and the amount of degradation is highly dependent on the antenna geometry.

The quality of the circular polarization can also be expressed by showing the co- and cross-polar radiation patterns. The co-polar radiation pattern is the radiation pattern of the wanted polarization, and the cross-polar radiation pattern is the radiation pattern of the unwanted and opposite polarization. The co and cross polar behavior is often quantified using the co/cross polar ratio (20log(co/cross)) and expressed in dB.

The relationship between the axial ratio and the co/cross polar ratio is shown in *Figure 9*.

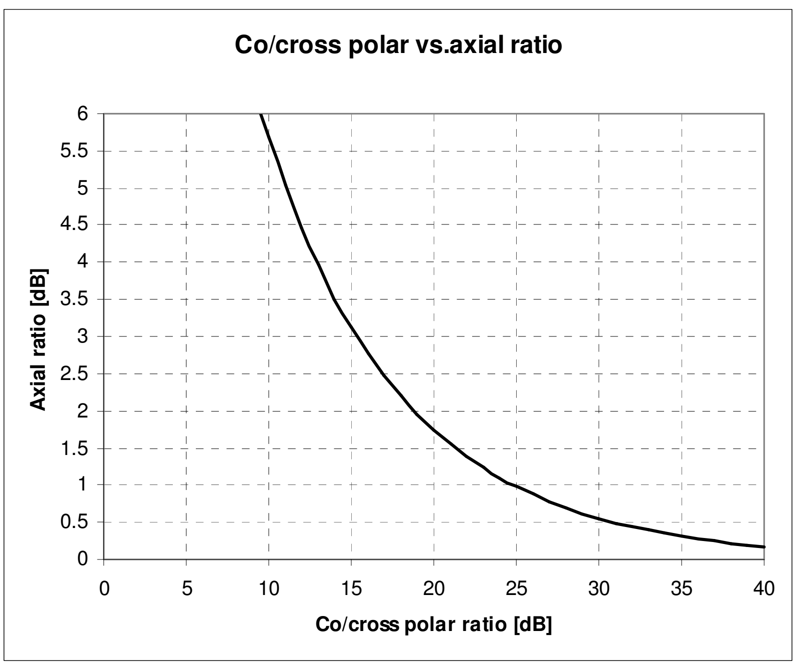

Figure 9: *Co/cross polar ratio versus axial ratio*

Bandwidth

Another important parameter of an antenna is the bandwidth it covers. Most of the time only impedance or return loss bandwidth is specified. However, it is important to realize that several other definitions of bandwidth exist: directivity bandwidth, polarization bandwidth and efficiency bandwidth. Directivity and efficiency are often combined to gain bandwidth.

Impedance/return loss bandwidth:

This is the frequency range over which the structure has a good impedance matching with respect to a certain reference impedance.

The impedance bandwidth depends on a large number of parameters related to the patch antenna element itself like quality factor Q and the type of feed used. The plot in *Figure 10* shows the return loss bandwidth of a patch antenna. The impedance bandwidth of a square, half wave patch antenna is typically limited to 1 to 3% and this is a major disadvantage of this type of patch antenna.

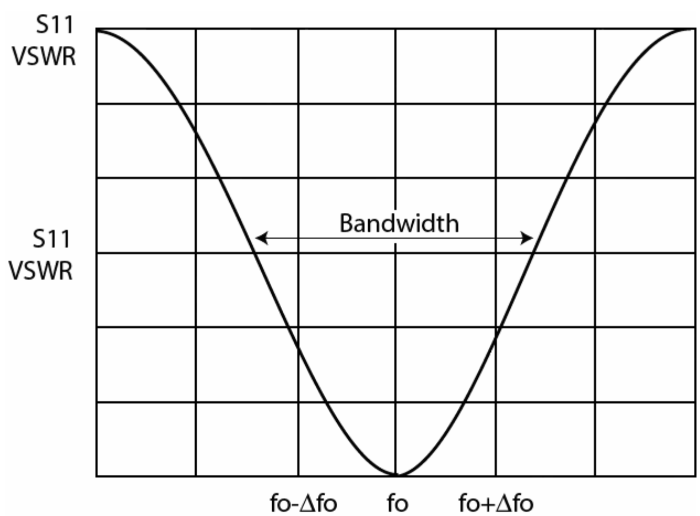

Figure 10: *Impedance bandwidth definition*

Keep in mind that impedance bandwidth can be defined in a couple of ways and that these do not forcibly mean the same thing: VSWR = 2:1, S11 <–10 dB or the maximum real impedance divided by the square root of two [Z(Re)/√2] bandwidth.

Directivity/gain bandwidth:

This is the frequency range where the antenna meets a certain directivity/gain requirement (like a 1 dB gain flatness).

Efficiency bandwidth:

This is the frequency range where the antenna has a defined efficiency.

Polarization bandwidth:
This is the frequency range where the antenna maintains a suitable co/cross ratio.

Axial ratio bandwidth:
This bandwidth is related to the polarization bandwidth and this number expresses the quality of the circular polarization of an antenna.

Alternative Feed Types

A classic way to feed a patch antenna is by using a coaxial probe or connector mounted at the appropriate impedance point and we will look at a couple of alternate ways to connect to a patch. Since in most patch antennas, the impedance bandwidth is much smaller than the radiation bandwidth, we will also look at some techniques to increase the bandwidth.

While the probe feed discussed so far has a bandwidth of only a couple %, it is still very useful for a lot of applications like GPS and WLAN.

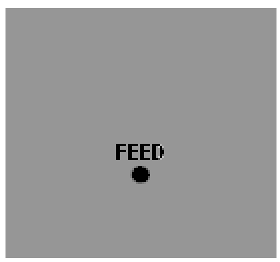

Figure 11: *Probe fed patch*

Another way to connect to the patch is a microstrip line at the edge of the element as shown in *Figure 12*.

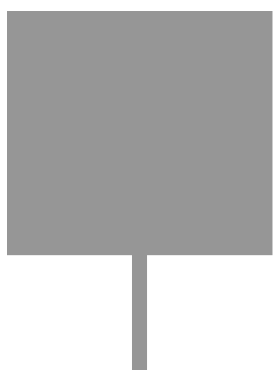

Figure 12: *Microstrip line fed patch antenna*

With this type of feed, impedance transformation to a useful value is needed as the impedance near the edge of the patch is quite high (a couple hundred Ω typical).

The advantage of this structure is the ability to place circuitry near the elements on the same substrate and avoid a multi-layer board. Also, elements can be connected in parallel in the same plane to form an array.

The impedance transformer itself tends to radiate and may distort the radiation pattern of the antenna.

A nice solution is a combination of the antennas from *Figures 11 and 12* and shown in *Figure 13*.

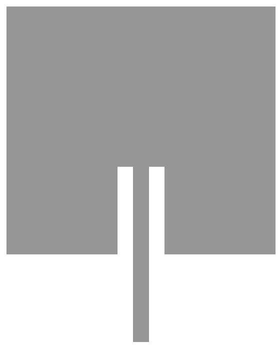

Figure 13: *Inset microstrip line fed patch antenna*

We keep the microstrip feed from *Figure 12* and create an inset to find the same impedance point from *Figure 11*. The gaps distort the radiation pattern of the element, but it is well worth the advantage of being able to connect a standard impedance circuitry straight to the element.

As discussed before, the impedance bandwidth of a basic patch is a couple % at best. The radiation bandwidth is usually much larger and can be up to 50%. Aperture coupling to a patch is a classic impedance bandwidth enhancement technique. The aperture coupled patch has a slot cut in the ground plane under the radiating element as shown in *Figures 14 and 15*. The slot is excited with a microstrip or stripline transmission line over the slot on the opposite side of the ground plane.

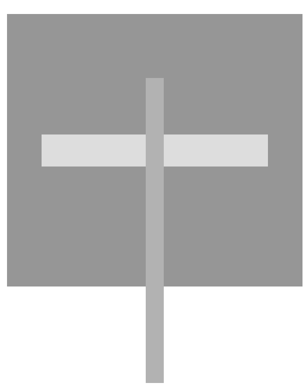

Figure 14: *Aperture coupled patch antenna*

The microstrip fed aperture coupled patch may suffer from a reduced front to back ratio. A stripline feed avoids this: the feed line is 'captured' between two ground layers. A cross section of a stripline fed, aperture coupled patch is shown in *Figure 15*. Impedance bandwidths of 10% can be readily achieved but this structure tends to be a bit more challenging to design.

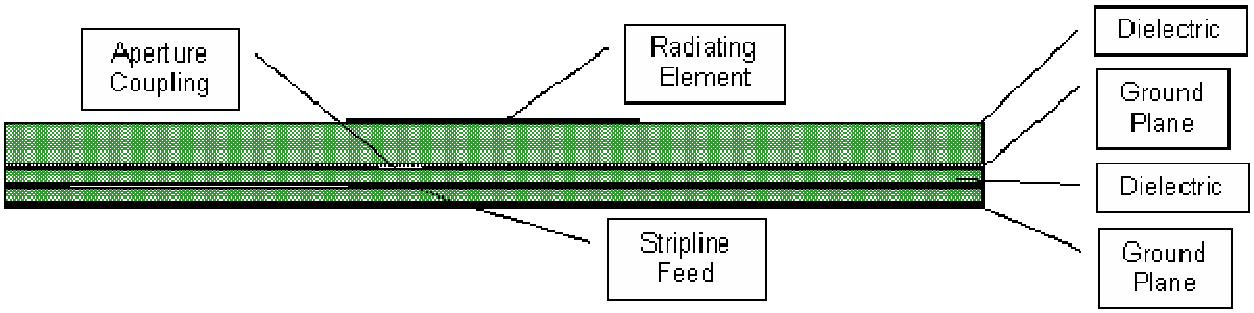

Figure 15: *Cross section of the aperture coupled patch antenna*

The Stacked Patch

Two patches can be vertically stacked and built like a multi layer printed circuit board. Vertically stacking two patches creates a dual band or a wide band patch antenna.

There are three ways to connect to this antenna:

- Connecting to both patches individually (dual band)
- Connecting to the upper patch only: series fed (dual band)
- Connecting to the lower patch only: shunt fed (wide band)

Connecting to both patches is a useful technique to create a dual band patch with separate feeds for each band. The required volume is considerable lower than when using two separate patch antennas for each band. The antenna itself acts as a diplexer, eliminating the need for an additional discrete diplexer. These features effectively reduce the overall system cost.

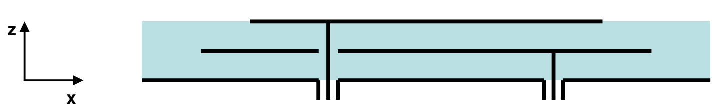

Figure 16: *Stacked patches with dual feed*

Feeding the upper patch only (series feed): this is a classic technique to create a dual band antenna with a single feed. Typically, the lower frequency band uses the lower (and larger) patch while the upper frequency band uses the upper (and smaller) patch. At the lower frequency, the leaky cavity between the lower patch and the ground plane resonates while the leaky cavity between upper patch and lower patch is electrically short circuited because it is off-resonance (the off-resonance impedance of patch antennas is effectively zero). At the higher frequency, the leaky cavity between the upper patch and lower patch resonates while the leaky cavity between lower patch and the ground plane is electrically short circuited because it is off-resonance. This is illustrated in the impedance versus frequency plot in *Figure 18* and the input reflection versus frequency plot in *Figure 19*. It is important that the upper patch is sufficiently smaller than the lower patch in order to avoid fringing field coupling between the two patches. This implies that the frequency bands should be separated by at least 5 to 10%.

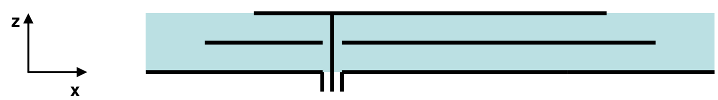

Figure 17: *Series fed stacked patches*

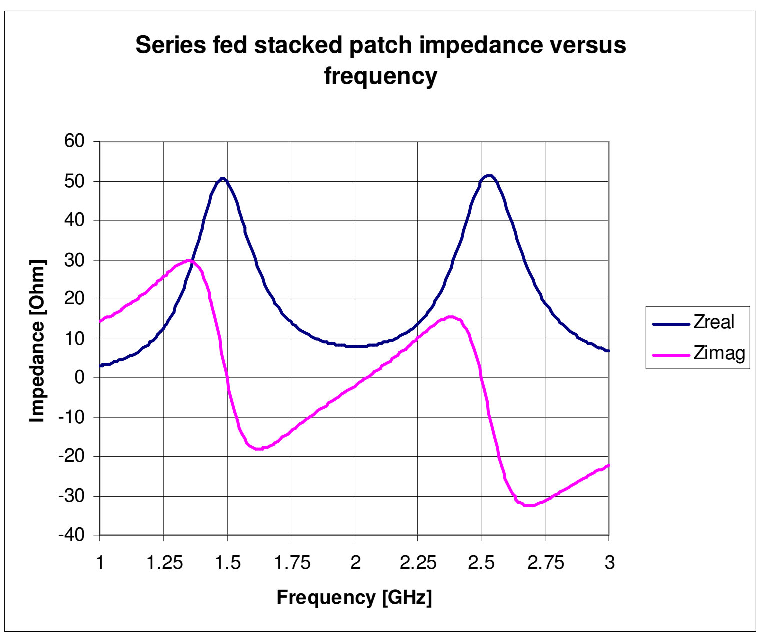

Figure 18: *Series fed stacked patch impedance behavior versus frequency*

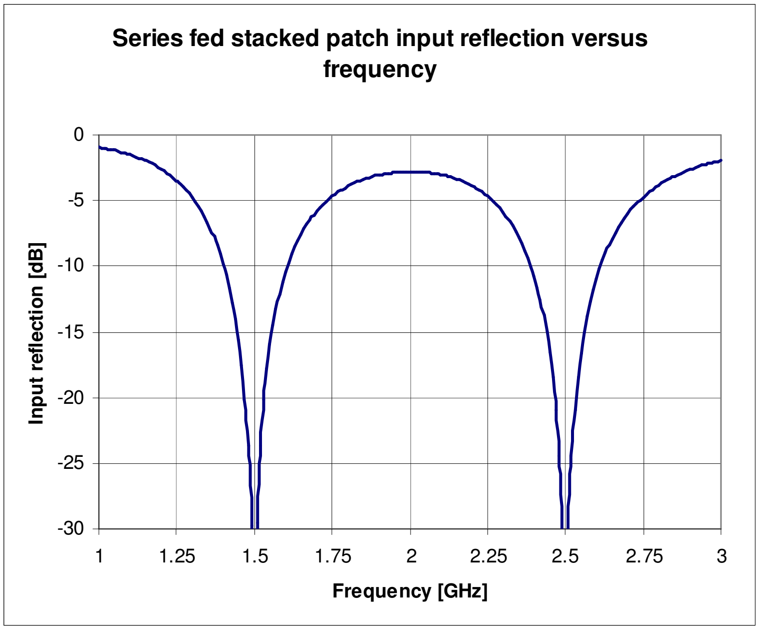

Figure 19: *Series fed stacked patch input reflection versus frequency behavior*

Feeding the lower patch only (shunt feed): this is a classic technique to create a wide band (3 to 10%) patch antenna where each of the two vertically stacked patches covers a sub-band. Dual frequency operation does not work very well with this configuration because the upper patch is electromagnetically coupled with the lower patch. This puts a constraint on the dimensions of the patches making this a wide band rather than dual band structure.

Figure 20: *Shunt fed stacked patches*

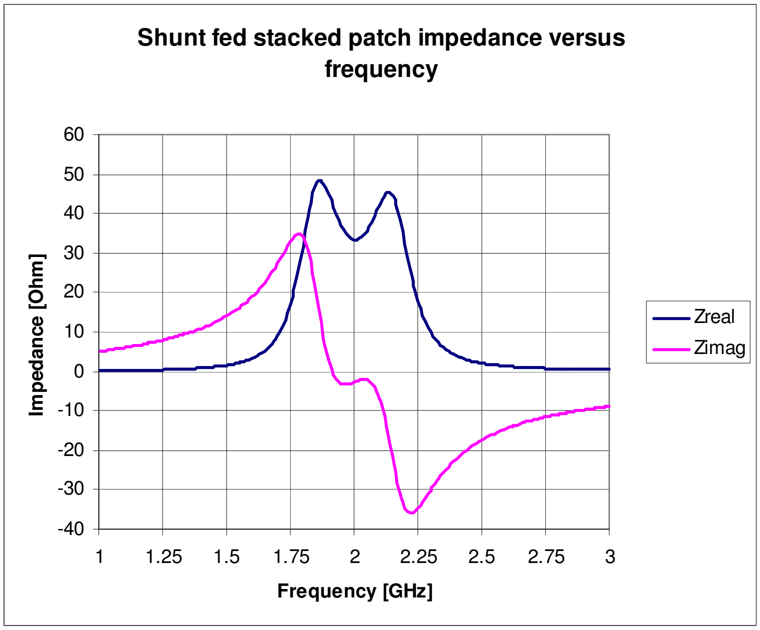

Figure 21: *Shunt fed stacked patch impedance behavior versus frequency*

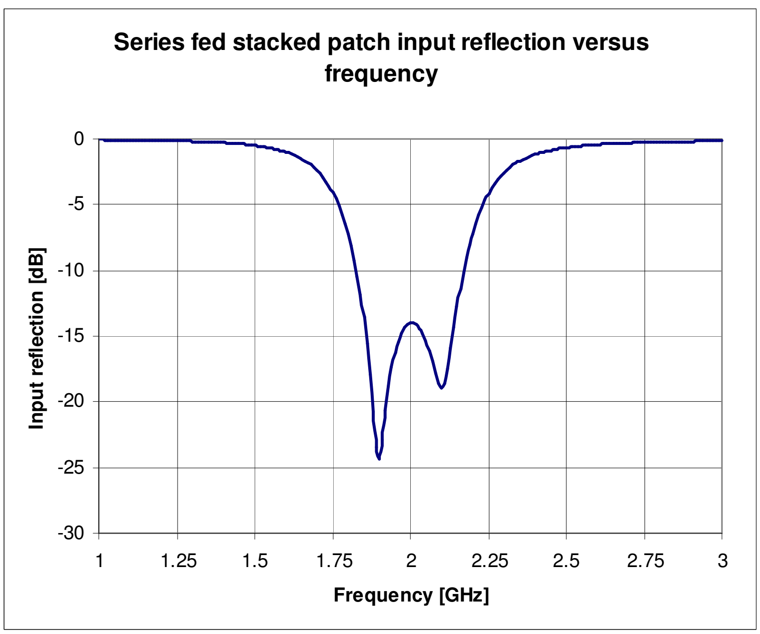

Figure 22: *Shunt fed stacked patch input reflection versus frequency behavior*

Series fed patches work very well (and only then) when the two bands are well separated and the antennas present an effective short at each other's bands of operation. Shunt fed patches work well (and only then) when both resonances are within a couple % of each other.

Higher Modes

So far, we've been talking about the fundamental TM10 mode. But, patch antennas are resonators and inherently exhibit higher modes.

Only a few TM modes are potentially useful and the other useful ones in a rectangular patch are TM20 and TM30. Other modes exist but result in exotic radiation patterns and have little -if any- practical use.

The radiation pattern cross sections of antennas operating in TM10, TM20 and TM30 modes have been plotted in *Figure 23*.

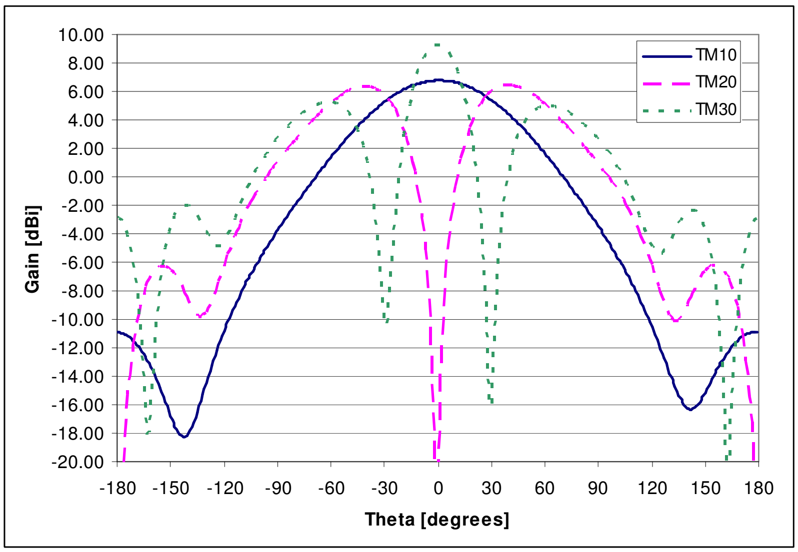

Figure 23: *Radiation patterns of higher rectangular patch modes*

This plot shows that the fundamental mode has one main lobe in broadside, the TM20 mode has two lobes symmetrically around broadside with a radiation null in broadside and the TM30 mode has three lobes: one in broadside and two symmetrically next to the broadside lobe.

These are for a rectangular patch only and differently shaped patch antennas exhibit other mode behavior and associated radiation patterns. Another excellent example is the round patch antenna shown in *Figure 24*:

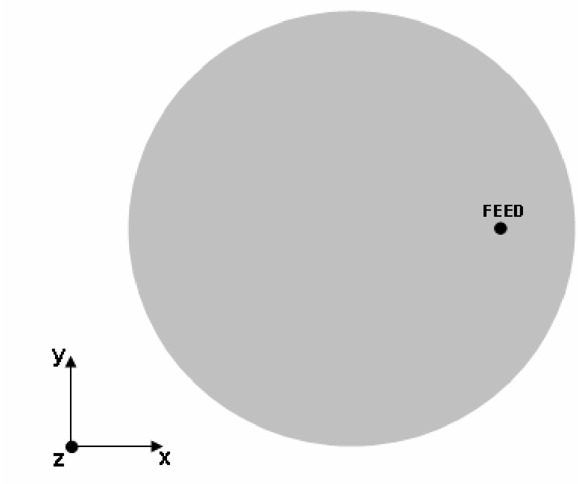

Figure 24: *Probe fed circular disk patch antenna*

The useful three modes of a round patch are TM11, TM21 and TM31. The radiation pattern cross sections are shown in *Figure 25*.

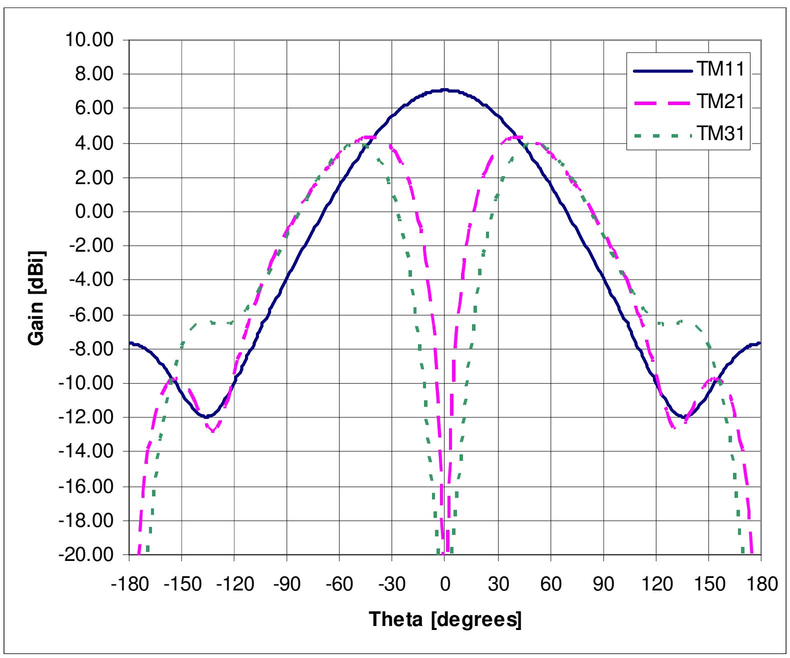

Figure 25: *Radiation patterns of higher circular disk patch modes*

The fundamental mode, TM11, radiates like a square patch operating in TM10. The second mode, TM21, radiates like a square patch operating in TM20. The third mode, TM31, radiates different from the rectangular patch: there is no broadside lobe.

Tuning higher modes tends to be a bit of a challenge relative to fundamental mode tuning because parasitic effects such as cross polar behavior and feed reactance are more problematic. The higher the mode, the more complex tuning is going to be.

Summary

In this article, some of the basic properties of linear and circular polarized patch antennas have been discussed. A basic set of specifications has been defined allowing the reader to understand and write a set of requirements for a specific application. A couple of patch antenna feeding techniques have been described. The stacked patch configuration has been introduced and the higher mode behavior of patch antennas has been explained.

Besides the patch antennas covered here, many more design options and different implementations are possible. Coverage of these alternatives is beyond the scope of this article, but they should be considered during the specification and development phases of the antenna.

If you have any questions, you can contact the authors at info@orbanmicrowave.com