Antenna Theory 101

From our discussion last time, you should have been able to take away simple but conservative operating rules to improve your chances of being heard on the air and enjoy long service life from your antenna and transmitting equipment. We discussed the role of antenna system components like the balun and the role it plays in getting our signal heard.

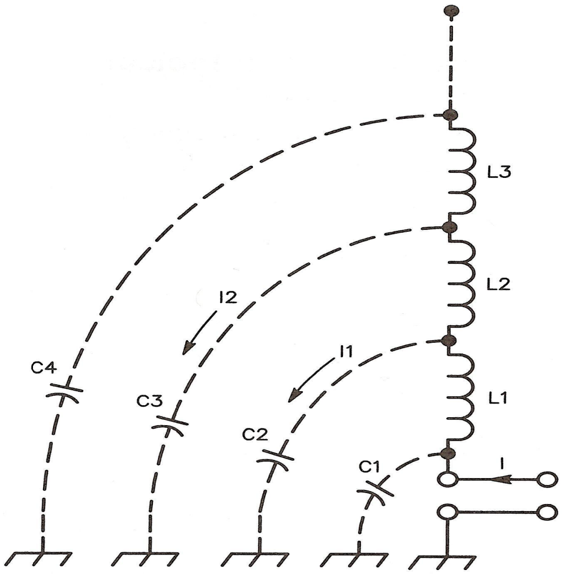
In our discussion this time we will look at the theory behind some simple antennas. Valuable information and understanding are possible with a little effort and study of the presented material. The stated goal of this series is to arm you with the knowledge and understanding that will enable you to do things for yourself rather than always relying on antenna manufacturers to supply what you need, and to make troubleshooting problems easier. We begin our discussion with the simplest of antenna types.

The vertical antenna

Let's start by illustrating a common simple antenna and the basic theory involved. The included diagram shown is a discrete component equivalence diagram. In other words the diagram shows in discrete terms the factors that influence RF as it travels on your vertical antenna.

We said from the beginning that transmission lines and antennas have complex characteristics. This diagram illustrates the reason for the complex nature of an antenna of any type or size.

By definition, there is series inductance along the length of the antenna conductor that will influence RF current in the conductor of the antenna element. Inductance increases in value with the length of the element. You may remember from studying basic electronic theory, inductors tend to oppose the flow of changing currents in a conductor. This is true of the antenna as a conductor of radio frequency currents as well.



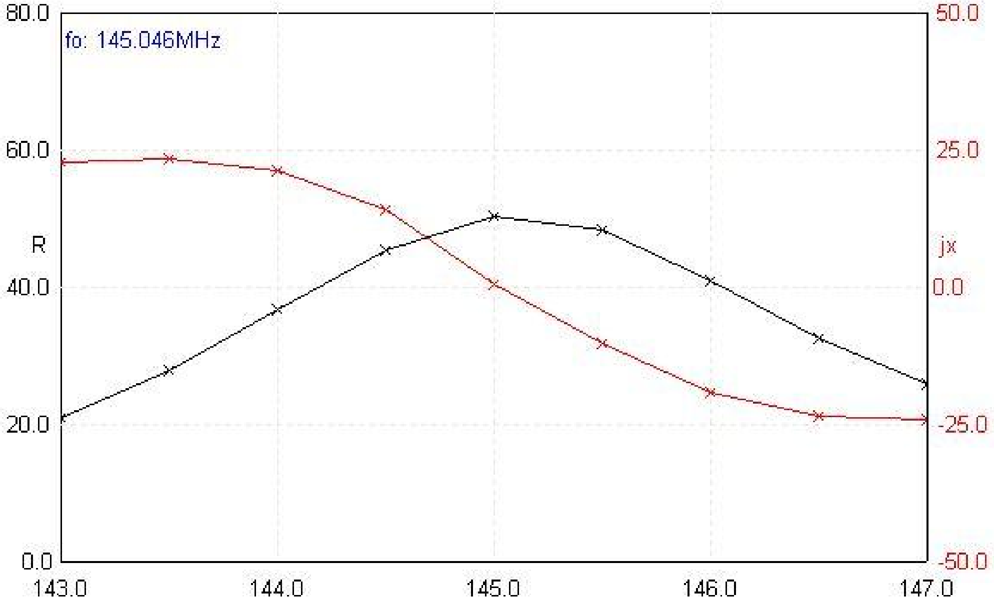
The vertical antenna as a discrete model
©ARRL Antenna Book 21st edition

Because of the proximity to ground, the antenna element will also have a capacitive component. From study of basic electronic components, you may remember that the definition of a capacitor is two conductors separated by a dielectric, carrying electrical charges of opposite voltage, and that oppose changes in voltage. That surely fits our vertical antenna (the ground being a conductor, however poor, and the other being the vertical antenna element separated by air as a dielectric). But, also notice that only part of the antenna is very close to the ground. The base section of the antenna will (by definition) have more capacitive effect than the tip. Also by definition, any point near the base will have less series inductive quality than the entire length of the element. By way of cumulative reasoning we can say that the reactive component values at different points along the length of the antenna element will have different values even though the frequency and element length may not change.

It may be apparent from the illustration that the series inductance and the shunt capacitance form a parallel LC circuit. It is, in fact, a parallel tuned circuit at the frequency we wish to transmit and receive on if the antenna is resonant.

Resonance

We now have to deal with the concept of resonance.



A real world graph of resonant antenna impedances

Basic electronic theory for AC circuits explained resonance of LC circuits to mean that the reactance value of one component is exactly the same magnitude but opposite in phase as the reactance value of the other reactive component in the circuit at the resonant frequency. This equal but opposite action cancels the reactive values. In this case, for our antenna to be resonant, the series inductance of the element conductor and the total shunt capacitance of the element proximity to ground would be equal and opposite at the resonant frequency. Here we are talking about capacitive reactance in ohms and inductive reactance in ohms not capacitance or inductance values specifically. We must be very clear on the difference between the two to properly understand the implications of antenna resonance and the rest of antenna theory.

How impedance is notated

We can illustrate this in a couple of ways. First lets recall the way we express the feed point impedance of the antenna.

We learned that it is written as:

R + jX

In this expression, a reactive component carries a "+" or "-" expression to indicate inductive or capacitive reactance values in ohms. The minus sign expresses the capacitive reactance value and the plus the inductive reactance. The "jX" is the unmatched (capactive vs. inductive) reactance value in ohms of the reactive components. A resonant antenna would have no reactive component at the resonate frequency because they are equal and opposite – canceling each other (the jX value is zero). *Editors note: It is not valid to say there is inductive and capacitive reactace at the same time. The imaginary notation is the sum of the two vectors.* Therefore the only part of the expression that has meaning, at resonance, is the “real”(“R” in the expression) value due to the canceling effect of the inductive and capacitive reactances.

Graphing Resonance

The above illustration involves looking at reactance in a more dynamic way. That is, we must see the effect that frequency has on reactance near resonance.

In the above illustration reactance is graphed against a frequency range, with the resistive value (the “R” or “real” value scale on the left), capacitive, and inductive reactance of a resonant antenna on appropriate scales in separate but overlaid lines - the “jX” or “imaginary value” scale on the right. At the frequency where capacitive reactance becomes zero and inductive reactance becomes zero simultaneously (i.e. where the reactance plot crosses the right hand scale zero line), the antenna becomes resonant. At this frequency only the “real” value of the expressed impedance plays a part in antenna performance. If this “R” value is at or near our coax characteristic impedance (usually 50 ohms), we have a very good match and consequently low VSWR values (near 1.00:1).

At frequencies above and below resonance, the impedance at the antenna feed point becomes more complex. At lower frequencies, the antenna is more capacitive and at higher frequencies than resonance, it appears more inductive (illustrated by the reactance curve making a positive then negative excursion as frequency is changed either side of the resonate frequency).

It is this dynamic and complex nature of the antenna that is at the root of most SWR readings measured on the coax. In our antenna resonance graph, the resonant frequency has a convenient “real” value of 50 ohms. However, off-resonance “real” values are much lower with an added reactive value. One could expect the SWR to be approaching 2:1 long before going off scale in either direction of our graph.

So how do we apply all this knowledge to normal practice? In the case of vertical antennas, we can see from the material presented that, the length of the antenna is a major determining factor in any antenna being ***self-resonant*** (a term indicating no other components like external capacitors or loading coils are needed to resonate).

Wavelength

Antenna physical length is measured in terms of how far RF will travel during one RF cycle (called a wavelength). Remember, this length will depend on frequency more than any other factor. The higher the frequency, the shorter the wavelength, and the shorter the resonant antenna for the same wavelength.

When the resonant antenna is long enough to allow a complete current cycle to reach the end of the antenna exactly, it is called a full-wavelength. Resonant lengths may occur at certain fractional wavelengths as well. The most common fractional wavelength vertical antenna is the ¼ wave. Quarter wave means that the antenna length allows only 90º of the full current cycle to flow on the antenna. That’s OK, because we know that resonant circuits only need a small excitation to sustain maximum current. We will examine current along our fractional wavelength vertical antenna later.

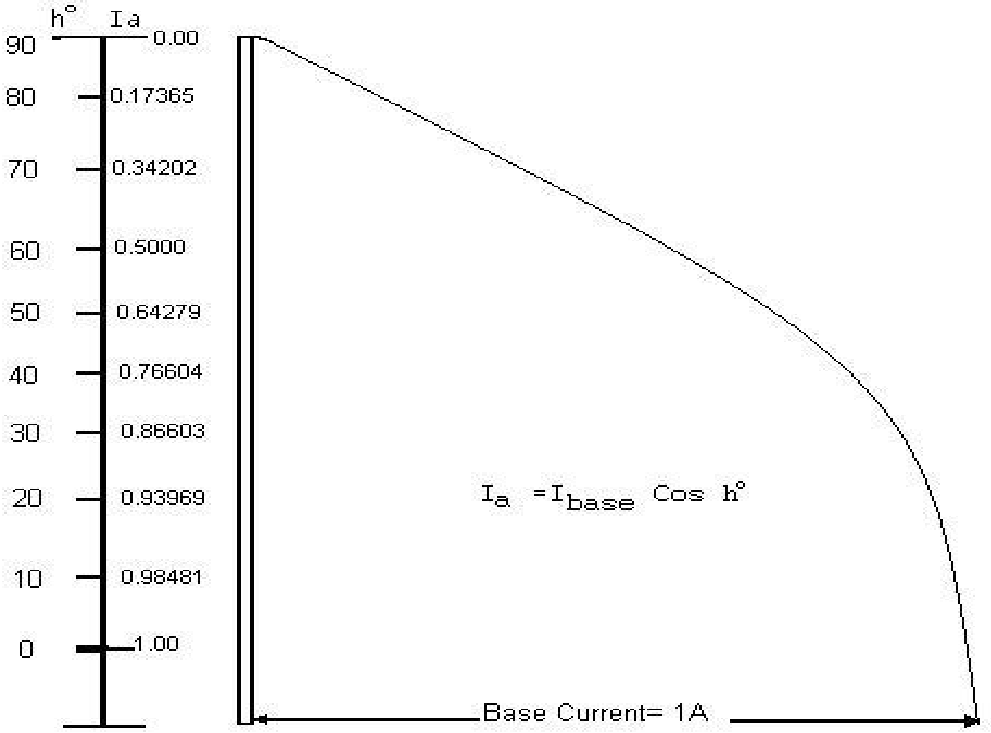
The most fundamental of all antennas is a wire whose length is half the transmitting wavelength. It is the unit from which many more complex forms of antennas are constructed and, if fed from the center, is known as a dipole antenna. To simplify things we can consider the half wave dipole in free space to be of a length expressed in this formula:

$$\text{Length(ft)} = \frac{492}{f\,(Mhz)}$$

The exact length in real world measurements will be slightly different than this simple formula will yield, due to variations like wire diameter and insulation (if any), as well as proximity to ground. One must also consider the additional length needed for attachment to insulators at the end and center. This formula is a theoretical guide for actual antenna element sizing that takes all factors into account in practice. Often an additional factor multiplied into the 492 numerator value to account for materials used, e.g. aluminum tubing instead of wire. This factor is noted with a distinctive “K” in the formula and is appropriately referenced as the “K” factor. For simplicity we will not use the “K” factor when referencing this formula when wire antennas are considered. *Note: more detail on antenna length and other detail information is in the ARRL Antenna Book 85th Edition Chapter 22.*

Dipoles

Dipoles are usually ½ to one full wavelength long. The feed point connection is usually somewhere near the center (recall our mention of the Windom dipole as being one of several exceptions). Scientists and experimenters often use the ½ wavelength, center-fed, horizontal dipole as a reference for gain and feed point characteristics. In free space, the ½ wave dipole exhibits predictable impedance and radiation pattern characteristics that are practically and mathematically convenient. To that end, the reference is often used by designation “dbi”. The comparison is made in the “db” (short for decibels) difference. The “i” indicates a comparison to the ½ wavelength dipole in free space (scientists call this an isotropic dipole – “i” for isotropic). Actually, it does not matter whether the isotropic dipole is horizontal or vertical, seeing that there is no influence from ground in free space. The only difference would be the polarization of the electric and magnetic waves radiated.



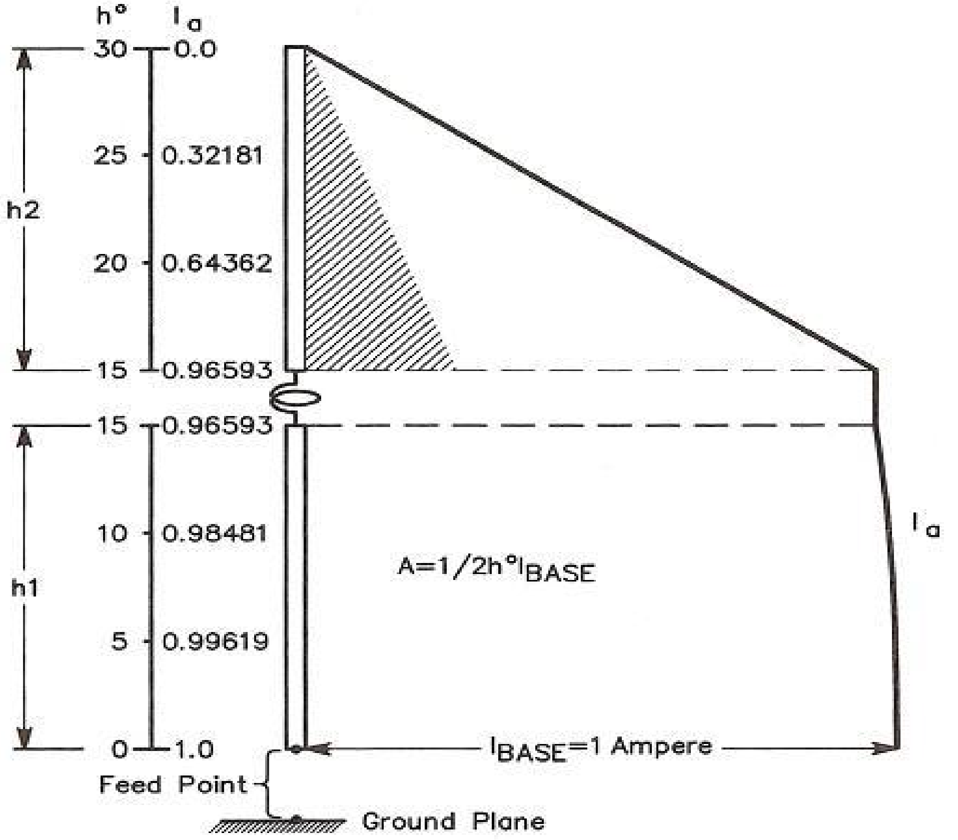
Suggested from material in the ARRL Handbook

Current on the vertical antenna

This illustration shows the current magnitude of RF as it flows on a ¼ wave vertical antenna. The curve traced on the right side of the diagram shows a non-linear current curve almost all the way to the tip. This type of curve is called a cosine curve, because it follows a mathematical calculation based on the sinusoidal nature of RF.

More importantly to us, there is a dramatic fall off of current for most of the lengt as RF reaches the tip of the vertical radiator and current flow is not a constant magnitude over the entire length. Ideally, we would like to have the maximum current flow all the way to the end in order to produce a maximum radiated signal. It is a physically impractical goal where HF mobile antenna installations are concerned. We can, however, dramatically increase current flow in most of the antenna length. To accomplish this, we can place reactive components in series with the antenna element, part of the way from the base to the tip (most conveniently 50-60% of the length).

By doing so, we alter the normal current flow along the element to resemble the next diagram below. This type of antenna is called a center-loaded vertical. The following illustration depicts current flow in the center-loaded antenna.



The center load is provided by the coil inductor, placed about 50-60% of the length from the base. This is a very short antenna – sometimes only 30º (it can be from 1/12 to 7/8 wavelength). You can see from the current flow trace at the right of the diagram, there is a more or less linear current flow all the way to the coil near the center. The inductive load altered the current flow in the antenna element to provide maximum current for more than half the length of the element. Allowing the maximum current to flow for more of the length of the antenna produces the possibility of more radiated signal. As can be seen from this illustration, the current flow in the top half of the antenna falls off in a more or less linear fashion becoming zero at the tip. This diagram is true for fractional wavelength vertical antennas only. The same indications would be true for any non-resonant fractional wavelength vertical antenna. However, the longer the antenna element overall, the more distance the maximum current flow is allowed – theoretically, more radiated signal will result. It is logical to assume that a 5/8 wave center loaded antenna will outperform a 3/8 wave center loaded antenna in signal strength (given all other factors are the same - e. g. power, environment, ground plane, etc.). In this case "bigger" is "better" - size always matters where antennas are concerned.

Shortened antenna performance

Can shortened antennas perform adequately on the air? By all means. But to do so, all sources of inefficiency must be minimized. The greatest source of signal loss in shortened antennas is what is called I^2R loss. This term expresses the theory that power that is dissipated in passive components like coils as heat and in the grounding system, known as ground loop losses. Loss in loading components is due to incorrect diameter and length choices the antenna maker employs and the size and/or material of the conductor used to make the coil inductor and in ground systems from conductivity of the grounding materials. Smaller wire will have higher series DC resistance and more I^2R loss than larger wire or tubing. A good rule of thumb for HF ham radio antenna inductors is “short, fat and made with big wire” is good. Whereas “long skinny and made with small wire” is not so good.

Consider the time honored “Bug Catcher” mobile antenna. It is an extremely short mobile vertical for HF frequencies, but one that has been a good performer because of the wise use of the design guidelines given here. On the opposite extreme is the Hamstick® style antenna. Moderate success has been enjoyed using the Hamstick® or Outbacker® style antenna despite the design inefficiencies. They are popular due in large part for the convenience of the small, lightweight design and relatively low wind resistance – not their radiation efficiency.

Somewhere in the middle of the extremes is a compromise design that proves to be both convenient and efficient. The Hustler™ style large-diameter resonators on a mobile vertical, prove to be quite efficient and are more practical than the much bulkier “Bug Catcher” style antenna for HF. Using the Hustler™ method of antenna construction, a single or multiple band vertical can be constructed by choosing screw-together radiating and resonating components in combination to construct a useful shortened vertical for mobile or stationary use.

Newer mobile designs include motor-driven spiral-grooved shaft tuning, variable inductor, base loaded mobile whips (whew! What a mouthful). This type of antenna is affectionately known as the “screwdriver” mobile antenna. The motor tuned variable inductor is of large diameter and uses a large gage wire, thus providing a good compromise of size and efficiency. This design is almost always a base loaded vertical whip antenna. In practical form, the best designed simple vertical mobile antenna has a maximum efficiency of no more than 30% or 3.25 dbi gain.

When considering the vertical for base station use, look for design and construction techniques as described above. A good practical example is the Bencher Butternut™ HF6V. In this model, coils are wound with very large diameter aluminum tubing and the length of the coils in normal operation is about 2:1 over diameter. The HF6V is a base loaded antenna on 80 meters and a linear loaded ¼ wave antenna on 15 meters. The radiation resistance tends to be much higher than other vertical designs. Bencher provides a considerable amount of advice and documentation on their web site and theory explanations with the product. It is a fine example of good engineering practice.

Do not be persuaded by advertising hype that expounds the virtues of some new “whiz-bang” design that will allow you to work all bands with 1.00:1 SWR or has 75% efficiency being no bigger than a bread box. They still operate under the same laws of physics as everyone else. To coin a phrase “It ain’t gonna happen”.

The base station vertical

Base station verticals have three basic types:

1. **The single or multi-band monopole.** This type uses one vertical radiator to transmit on all designed frequencies and depends on earth ground as a reflector or counterpoise (not an accurate term but a common one). This type of antenna may require ground wires to improve the effect of ground radiation at lower HF frequencies in some soil conditions and to improve overall efficiency. If good construction is used, as with the Hustler™ 5BTV and 6BTV (a multi-band trap loaded vertical) or Butternut™ HF6V (a base loaded multi-band

vertical), it can be a very useful DX or non-directional HF base antenna.

2. **The vertical dipole or vertical offset-fed dipole.** This type of vertical antenna is unique in that it has the feedpoint somewhere near the center of the vertical element. The top half being one conductor and the bottom half being the other. The advantage here is there is no need for a radial system. The antenna contains it's own ground radiator as part of the antenna construction. Feedpoint impedances are usually higher than ground-mounted verticals and are an easier match in some cases. This does not necessarily translate to better signal radiation.
3. **The vertical dipole with complex loading.** In this style of antenna, the feed point is again elevated to somewhere near the center and external reactive elements are added to cause parts of the vertical to be resonant at different frequencies by providing proximity loading. You may notice this in the form of wires, coils, or cylindrical (trombone) capacitors at certain points on the antenna elements.

Verticals that exhibit an overall length of more than ¼ wavelength at the lowest operating frequency will perform better overall. Radically shortened antennas are a compromise at best and at worst are a total waste of your hard earned cash.

It would be an error in judgment, if the importance of the ground in the efficiency of vertical antennas was not mentioned. This subject is provided in a more thorough discussion of the next chapter.