**UNIT I**

**FUNDAMENTALS OF RADIATION**

An ***antenna*** is a device for converting electromagnetic radiation in space into electrical currents in conductors or vice-versa, depending on whether it is being used for receiving or for transmitting, respectively.  Passive radio telescopes are receiving antennas.  It is usually easier to calculate the properties of transmitting antennas. Fortunately, most characteristics of a transmitting antenna (e.g., its radiation pattern) are unchanged when the antenna is used for receiving, so we often use the analysis of a transmitting antenna to understand a receiving antenna used in radio astronomy.

**Radiation from a Short Dipole Antenna (Hertz Dipole)**

*The coordinate system used to describe the radiation from a short (total length lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char1C.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png) dipole driven by a current source of frequency http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png.*

The simplest antenna is a short (total length l much smaller than one wavelength http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png) ***dipole antenna***, which is shown above as two colinear conductors (e.g., wires or conducting rods).  Since they are driven at the small gap between them by a current source (a transmitter), the current in the bottom conductor is 180 deg out of phase with the current in the top conductor.  The radiation from a dipole depends on frequency, so we consider a driving current I varying sinusoidally with angular frequency http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.png=2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png:

I=I0cos(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngt) http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.png

where I0 is the peak current going into each half of the dipole.  It is computationally convenient to replace the trigonometric function cos(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngt) with its exponential equivalent, the real part of

e−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngt=cos(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngt)−isin(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngt)

so the driving current can be rewritten as

I=I0e−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngt

with the implicit understanding that only the real part of I represents this current. The driving current accelerates charges in the antenna conductors, so we can use Larmor's formula to calculate the radiation from the antenna by converting from the language of charges and accelerations to time-varying currents.

Recall that electric ***current*** is defined as the time derivative of electric charge:

Ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pngdtdq http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Along a wire, the current is the amount of charge flowing past any point per unit time. For a wire on the z-axis

I=dtdq=dzdqdtdz=dzdqv http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.png

where v is the instantaneous flow velocity of the charges.

It is a common misconception to believe that the velocities v of individual electrons in a wire are comparable with the speed of light c because electrical *signals* do travel down wires at nearly the speed of light.  A wire filled with electrons is like a garden hose already filled with an incompressible fluid—water.  When the faucet is turned on, water flows from the other end of a full hose almost immediately, even though individual water molecules are moving slowly along the hose.  As a specific example, consider a current of 1 ampere flowing through a copper wire of cross section http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1B.png=1 mm2=10−6 m2.  The number density of free electrons is about equal to the number density of copper atoms in the wire, nhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png1029 m−3.  In mks units, the charge of an electron is

−ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png80http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−12 statcoulhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1 coul3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png109 statcoulhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png60http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−19 coul

One Ampere is one Coulomb per second, so the number N of electrons flowing past any point along the wire in one second is

N=Ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.pngehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png1 coul s−11http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png60http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−19 coulhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png6http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png25http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1018 s−1

The average electron velocity is only

vhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.pngNhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1B.pngnhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png6http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png25http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1018 s−110−6 m2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1029 m−3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png6http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−5 m s−1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char1C.pngc

Thus the nonrelativistic Larmor equation may be used directly to calculate the radiation from a wire.

From the derivation of Larmor's formula, recall that

Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png=rc2qvhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char5F.pngsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png

so in the short dipole

Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png+lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2z=−lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2dzdqdzrc2vhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char5F.pngsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png

For a sinusoidal driving current,

vhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char5F.png=−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngv

and

Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png=rc2−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png−lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2+lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2dzdqvdz

Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png=rc2−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png−lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2+lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2Idz

That is, the radiated electric field strength Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png is proportional to the integral of the current distribution along the antenna. The current at the center is just the driving current I=I0e−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngt and the current must drop to zero at the ends of the antenna, where the conductivity goes to zero. For a short antenna, we can make the approximation that the current declines linearly from the driving current at the center to zero at the ends:

I(z)=I0e−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngthttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char14.png1−http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.pngzhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.png(lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2)http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char15.png

Then

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png−lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2+lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char3D.png2Idz=2I0le−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngt

and

Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png=rc2−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png2I0le−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngt

Substituting http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.png=2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngchttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png gives

Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngrc2−i2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngcsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png 2I0le−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngt

Ehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.png=c−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngI0lre−ihttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char21.pngt

Since http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.pngEhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char7E.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.pngHhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char7E.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.png (cgs), the time-averaged Poynting flux (power per unit area) is

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngShttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=c4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngE2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char3F.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.png

and

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngShttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=c4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngI0lchttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2r2sin2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.png21http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png

because http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngcos2(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngt)http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2. Note that the radiation from a short dipole has the same polarization and the same doughnut-shaped ***power pattern*** (the power pattern is the angular distribution of radiated power, often normalized to unity at the peak)

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char2F.pngsin2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png(3A1)

as Larmor radiation from an accelerated charge because all of the charges in the dipole are being accelerated along one line much shorter than one wavelength.  From the observer's point of view, the power received depends only on the *projected* (perpendicular to the line of sight) length (lsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png) of the dipole.  Thus the electric field received is proportional to the *apparent* length of the dipole. The time-averaged total power emitted is obtained by integrating the Poynting flux over the surface area of a sphere of any radius rhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char1D.pngl centered on the antenna:

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngShttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.pngdA=c4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngI0lchttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.png21http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char1E.png=02http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char12.png=0r2sin2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pngrsinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pngdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.pngrdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=c4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngI0lchttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.png21http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char12.png=0sin3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pngdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png

Recall that http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char52.png0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngsin3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pngdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png3 , so the time-averaged power radiated by a short dipole is

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=3chttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngI0lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2(3A2) http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.png

where I0cos(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngt) is the driving current, l is the total length of the dipole, and http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngis the wavelength.

Most practical dipoles are half-wave dipoles (lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2) because half-wave dipoles are resonant, meaning that they provide a nearly resistive load to the transmitter.  When each half of the dipole is http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png4 long, the standing-wave current is highest at the center and naturally falls as cos(2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngzhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png) to almost zero at the ends of the conductors.

The ***ground-plane vertical*** shown below is very similar to the dipole.  A ground-plane vertical is one half of a dipole above a conducting plane, which is called a "ground plane" because historically the conducting plane for vertical antennas was the surface of the Earth.  The transmitter is connected between the base of the vertical, which is insulated from the ground, and the ground plane near the base.  Many AM broadcast transmitting antennas are tall (at http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char18.png1 MHz, http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char18.png300 m and a http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png4 vertical antenna is about 75 m high) towers acting as quarter-wave verticals. The conducting ground plane is a mirror that creates the lower half of the dipole as the mirror image of the upper half.  Electric fields produced by the vertical must induce currents in the conducting plane that ensure that the horizontal component of the electric field goes to zero on the conductor.  This implies that the virtual electric fields from the image vertical must have the same amplitude but be 180 degrees out of phase, exactly as in a half-wave dipole.  Consequently the radiation field from a ground-plane vertical is identical to that of a dipole in the half space above the ground plane and zero below the ground plane.

*The ground-plane vertical is just half of a dipole above a conducting plane. The lower half of the dipole is the reflection of the vertical in the mirror provided by the conducting "ground plane."  The image vertical is 180 deg out of phase with the real vertical.  Above the ground plane, the radiation from the ground-plane vertical is exactly the same as the radiation from the dipole.*

According to the strict definition of an antenna as a device for converting between electromagnetic waves in space and currents in conductors, the only antennas in most radio telescopes are half-wave dipoles and their relatives, quarter-wave ground-plane verticals. The large parabolic reflector of a radio telescope serves only to focus plane waves onto the ***feed antenna***.  [The term "feed" comes from radar antennas used for transmitting; the "feed" feeds transmitter power to the main reflector.  Receiving antennas used in radio astronomy work the other way around, and the "feed" actually collects radiation from the reflector.]  
  
Actual half-wave dipoles, backed by small reflectors about http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png4 behind them to focus the dipole pattern in the direction of the main dish, are normally used as feeds at low frequencies (http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3C.png1 GHz) or long wavelengths (http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3E.png0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png3 m) because of their relatively small size.  However, the radiation patterns of  half-wave dipoles backed by small reflectors are not well matched to most parabolic dishes, so their performance is less than optimum.    
  
At shorter wavelengths, almost all radio-telescope feeds are quarter-wave ground-plane verticals inside ***waveguide horns***. Radiation entering the relatively large (size http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3E.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png) rectangular or circular aperture of the tapered horn is concentrated into a rectangular or circular waveguide with parallel conducting walls.  In the case of the rectangular waveguide shown below, the side walls are separated by slightly over http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2 so that vertical electric fields can travel down the waveguide with low loss.  The top and bottom walls are separated by somewhat less than http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2 so only one mode with vertical electric fields can propagate.  The http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png4 vertical inserted through a small hole in the bottom wall collects most of this vertically polarized radiation and converts it into an electric current that travels down the coaxial cable to the receiver.  The backshort wall about http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png4 behind the dipole ensures that the dipole sees only radiation coming from the direction of the horn opening.

*Most high-frequency feeds are quarter-wave ground-plane verticals inside waveguide horns. The only true antenna in this figure is the http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png4 ground-plane vertical, which converts electromagnetic waves in the waveguide to currents in the coaxial cable extending down from the waveguide.*

**Radiation Resistance**

The power flowing through a circuit is P=Vhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.pngI, where V is the voltage (defined as energy per unit charge) and I is the current (defined as charge flow per unit time), so P has dimensions of energy per unit time.  The physicist George Simon Ohm observed that the current flowing through most materials is proportional to the applied voltage, so many (but not all) objects have a well-defined *resistance* defined by R=Vhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngI (***Ohm's law***).  For them, P=Vhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.pngI=I2R=V2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngR.

From Ohm's law for time-varying currents,

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngI2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.pngR

If I=I0cos(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char21.pngt),

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png =2I02R

The ***radiation resistance*** of an antenna is defined by

Rhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pngI022http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png(3A3)

For our short dipole, the radiation resistance is

R=3c2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.pnglhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2

Example: A "half-wave" dipole has length l=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2. This is the length of a *resonant* dipole antenna.  Resonant antennas are used in most real applications because the impedance of a resonant antenna is resistive; nonresonant antennas have large capacitive or inductive reactances as well. A half-wave dipole doesn't strictly satisfy our criterion lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char1C.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png for being "short," so the current distribution along the dipole is actually closer to sinusoidal than linear and our calculated radiation resistance will not be exact. Proceeding nonetheless to estimate the radiation resistance of a half-wave dipole,

Rhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png23http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1010 cm s−1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.png21http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png5http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png5http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−11 s cm−1

Engineers and real test instruments use the mks "Ohm" (symbol http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png) as the unit of resistance. The conversion factor is 1 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png= (10−11http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png9) s cm−1, so

Rhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png5http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png5http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−11scm9 cm10−11 s http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png50 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png

[The actual radiation resistance of a half-wave dipole in free space is about 73 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png.]

For a given driving current, a ground-plane vertical of height lhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2 emits exactly like a dipole of length l above the ground plane and zero below the ground plane.  Thus the total power emitted by the vertical is half the power emitted by the dipole, and the radiation resistance of the vertical is half the radiation resistance of the dipole.

The radiation resistance R0 of free space can be obtained from the relations

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.pngShttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char7E.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.png=c4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngE2 and P=RV2 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Since the electric field E is just the voltage per unit length Vhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngl and the flux is the power per unit area l2,

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.pngShttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char7E.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char6A.png=cV24http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngl2=V22R0l2

R0=c4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1010 cm s−1

Converting to mks units yields the radiation resistance of space in Ohms:

R0=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1010 cm s−1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png9http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−11 sec cm−1 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png−1=120http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png377 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Since a black hole is a perfect absorber of radiation, its impedance must also be 120http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png to match that of free space.  A black hole spinning in an external magnetic field can generate electrical power with a voltage/current ratio of 120http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png, and this process may be important in powering quasars (see Blandford & Znajek 1977, MNRAS, 179, 433).

**The Power Gain of a Transmitting Antenna**

The ***power gain*** G(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png) of a transmitting antenna is defined as the power transmitted per unit solid angle in direction (http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png) divided by the power transmitted per unit solid angle from an isotropic antenna driven by a transmitter supplying the same total power. Frequently the value of G is expressed logarithmically in units of decibels (dB):

G(dB)http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.png10http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.pnglog10(G)

For a lossless antenna, energy conservation requires that the gain averaged over all directions be

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngGhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char52.pngspheredhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char52.pngsphereGdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png

or

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngGhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=1(3A4)

Consequently

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.pngsphereGdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.pngsphere1dhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png

for *any* lossless antenna. Different lossless antennas may radiate with different directional patterns, but they cannot alter the total amount of power radiated. Consequently, the gain of a lossless antenna depends only on the angular distribution of radiation from that antenna. In general, an antenna having peak gain Gmax must beam most of its power into a solid angle http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngsuch that

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngGmax http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Thus the higher the gain, the narrower the beam or power pattern.

**The Effective Area of a Receiving Antenna**

How can we characterize antennas used for receiving, as in radio astronomy, rather than for transmitting? The receiving counterpart of transmitting power gain is the ***effective area*** or ***effective collecting area*** of an antenna.

Imagine an ideal antenna that collects all of the radiation falling on it from a distant point source and converts it to electrical power—a "rain gauge" for collecting photons.  The total spectral power that it collects will be the product of its geometric area A and the incident spectral power per unit area, or flux density S.  By analogy, if any real antenna collects spectral power Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png, its *effective area* Ae is defined by

Ae http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pngPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.pngS(matched)(3A5) http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.png

where S(matched) is the flux density in the "matched" polarization.

What does ***matched polarization*** mean? Any electromagnetic wave can be decomposed into two orthogonal polarized components. For example, the transverse electric field can be resolved into horizontal and vertical components, or horizontal and vertical ***linear polarizations***. If the horizontal and vertical electric fields are equal in amplitude and 90http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char0E.png out of phase, the radiation is ***circularly polarized****.* Any radio wave can also be decomposed into left- and right-circular polarizations. If the wave is essentially random (noise generated by blackbody radiation for example), the two orthogonal components will vary rapidly in intensity but have equal powers when averaged over long times. Such radiation is called ***unpolarized***. Thus for an *unpolarized* source,

S(matched)=2S http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Blackbody radiation is unpolarized.  Most radio astronomical sources are unpolarized or nearly so.

Any antenna with a single output collects only one of the two polarizations from an electromagnetic wave. For example, a linear dipole antenna collects radiation only from the linear polarization whose electric field is parallel to the antenna wires. Electric fields perpendicular to the dipole antenna do not produce currents in the antenna, so the linear dipole is completely insensitive to the linear polarization perpendicular to its wires.  A pair of crossed dipoles is needed to collect power from both orthogonal polarizations simultaneously.

Just as energy conservation implies that all lossless transmitting antennas have the same average power gain, all lossless receiving antennas have the same average collecting area.  This average collecting area can be calculated via another thermodynamic thought experiment.

*A cavity in thermodynamic equilibrium at temperature T containing a resistor R is coupled to an antenna, also at temperature T, through a filter passing frequencies in the range http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.pngto http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png+dhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png.*

Imagine an antenna inside a cavity in full thermodynamic equilibrium at temperature T connected through a transmission line to a matched resistor in a second cavity at the same temperature. Suppose further that the connection contains a filter that passes only a narrow range of frequencies between http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.pngand http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png+dhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png. Since this entire system is in thermodynamic equilibrium, no net power can flow between the antenna and the resistor. Otherwise, one cavity would heat up and the other would cool down, in violation of the second law of thermodynamics. Thus the total spectral power from all directions collected by the antenna

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=AeS(matched)=Ae2S=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngAe(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)2Bhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.pngdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png

must equal the Nyquist spectral power Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=kT produced by the resistor. Using the Rayleigh-Jeans approximation

Bhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png22kT

gives

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png22kThttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngAedhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png=kT

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngAedhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2

The average collecting area is defined by

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngAehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char52.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char52.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngAedhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

The effective collecting area of a receiving antenna is independent of its radiation environment, so this result applies for any type of radiation, not just blackbody radiation.  Without using Maxwell's equations we have obtained the remarkable result true for *all* lossless antennas:

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngAehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png24http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png(3A6)

* Any antenna, from a short dipole to the 100-m diameter Green Bank Telescope, has the same average collecting area http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngAehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png that depends only on wavelength.

In the case of an isotropic antenna, the effective collecting area in any direction equals the average collecting area:

Ae(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngAehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png24http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png

Consequently, a nondirectional antenna operating at a very short wavelength http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngwill have a very small effective collecting area and poor sensitivity for reception. For this reason, most satellite broadcast services, GPS or satellite FM radio for example, transmit at relatively long wavelengths (10 to 20 cm). Likewise, practical radio telescopes can be constructed from arrays of dipoles at long wavelengths (http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3E.png1 m) but not at short wavelengths because the number of dipoles needed to produce useful collecting areas would be too large.

**Reciprocity Theorems**

Many antenna properties are the same for both transmitting and receiving. It is often easier to calculate the gain of a transmitting antenna than the collecting area of a receiving antenna, and it is often easier to measure the receiving power pattern than to measure the transmitting power pattern of a large radio telescope. Thus this receiving/transmitting "reciprocity" greatly simplifies antenna calculations and measurements. Reciprocity can be understood via Maxwell's equations or by thermodynamic arguments.

Burke & Smith (1997) state the electromagnetic case for ***reciprocity*** clearly: "An antenna can be treated either as a receiving device, gathering the incoming radiation field and conducting electrical signals to the output terminals, or as a transmitting system, launching electromagnetic waves outward. These two cases are equivalent because of time reversibility: the solutions of Maxwell's equations are valid when time is reversed."

The ***strong reciprocity theorem***:

*If a voltage is applied to the terminals of an antenna A and the current is measured at the terminals of another antenna B, then an equal current (in both amplitude and phase) will appear at the terminals of A if the same voltage is applied to B.*

can be formally derived from Maxwell's equations [see a partial derivation in Rohlfs & Wilson Section 5.4] or by network analysis [see Kraus "Antennas", p. 252].

*The strong reciprocity theorem implies that the transmitter voltages VA and VB are related to the receiver currents IA and IB by*

*IBVA=IAVB*

*for any pair of antennas A and B.*

For most radio astronomical applications, we are not concerned with the detailed phase relationships of voltages and currents, and we can use a ***weak reciprocity theorem*** that relates the angular dependences of the transmitting power pattern and the receiving collecting area of any antenna:

*The power pattern of an antenna is the same for transmitting and receiving.*

That is:

G(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char2F.pngAe(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)(3A7)

The weak reciprocity theorem can be proven by another simple thermodynamic thought experiment: An antenna is connected to a matched load inside a cavity initially in equilibrium at temperature T. The antenna simultaneously receives power from the cavity walls and transmits power generated by the resistor. The total power transmitted in all directions must equal the total power received from all directions since no net power can be transferred between the antenna and the resistor; otherwise the resistor would not remain at temperature T. Moreover, in any direction, the power received and transmitted by the antenna must be the same, else the cavity wall in directions where the transmitted power was greater than the received power would rise in temperature and the cavity wall in directions of lower transmitted/received power ratio would cool, leading to a violation of the second law of thermodynamics.

*The weak reciprocity theorem states that the transmitting and receiving power patterns of an antenna cannot differ as shown here, without violating the second law of thermodynamics.*

The constant of proportionality relating G and Ae can be derived from our earlier results for an isotropic antenna

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngAehttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png24http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngandhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char68.pngGhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char69.png=1

Thus energy conservation and the weak reciprocity theorem imply

Ae(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2G(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)(3A8)

for any antenna. This extremely useful equation lets us compute the receiving power pattern from the transmitting power pattern and vice versa.

Example: We can use our calculation of the transmitting power pattern of a short dipole to calculate its effective collecting area when used as a receiving antenna:

Ae(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2G(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png24http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png23sin2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png

Ae=8http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2sin2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png

Example: What is the power per unit bandwidth Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png collected by a short dipole at http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png=10 GHz broadside to (http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png=90http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/120/char0E.png) the Sun, a thermal source whose flux density is Shttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png106 Jy?

The broadside (sinhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.png=1) collecting area of the short dipole is

Ae=8http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png=chttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png=1010 Hz3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png108 m s−1=0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png03 m

so

Ae=8http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png3http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png(0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png03 m)2=1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png07http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−4 m2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png1 cm2 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

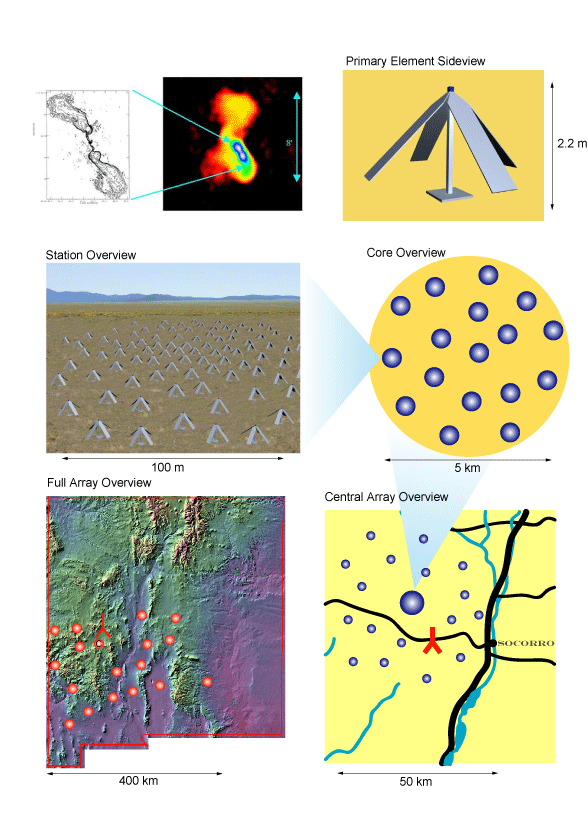
Clearly, dipoles or other nearly isotropic antennas have very small collecting areas at short wavelengths.  For an unpolarized source,

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=AeS(matched)=2AeS

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png07http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−4 m2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png21http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png106 Jyhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1 Jy10−26 W m−2 Hz−1

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=6http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−25 W Hz−1

Arrays of dipoles do make sense at long wavelengths.  For example, the Long Wavelength Array ([LWA](http://www.phys.unm.edu/%7Elwa/index.shtml)) for http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png= 20 to 80 MHz (http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png 4 to 15 m) will consist of 53 stations, each a 100 m http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png100 m array of crossed dipoles.  The effective collecting area of the LWA will be up to Ae=4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png106 m2 at http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png=15 m.

  
*The proposed Long-Wavelength Array of crossed dipoles.* [*Image credit*](http://www.phys.unm.edu/%7Elwa/design.shtml)

LWDA first 8 dipoles
      near VLA  
  
*Eight test dipoles of the LWDA (Long-Wavelength Development Array) on the VLA site.* [*Image credit*](http://www.phys.unm.edu/%7Elwa/lwda.shtml)

**Antenna Temperature**

A convenient practical unit for the power output per unit frequency from a receiving antenna is the ***antenna temperature*** TA. Antenna temperature has nothing to do with the physical temperature of the antenna as measured by a thermometer; it is only the temperature of a matched resistor whose thermally generated power per unit frequency equals that produced by the antenna. It is widely used because:

1. 1 K of antenna temperature is a conveniently small power. TA=1 K corresponds to Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=kTA=1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png38http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−23 J K−1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png1 K=1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png38http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−23 W Hz−1.
2. It can be calibrated by a direct comparison with hot and cold *loads* (another word for matched resistors) connected to the receiver input.
3. The units of receiver noise are also K, so comparing the signal in K with the receiver noise in K makes it easy to decide if a signal will be detectable.

TAhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pngkPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png(3A9)

An unpolarized point source of flux density S increases the antenna temperature by

TA=kPhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=2kAeS(3A10)

where Ae is the effective collecting area.  It is often convenient to express the point-source sensitivity of a radio telescope in units of "Kelvins per Jansky" rather than in units of area (m2).  The effective area corresponding to a sensitivity of 1 K Jy−1 is

Ae=S2kT=10−26 W m−2 Hz−12http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char01.png1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png38065http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−23 J K−1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char01.png1 K=2761 m2 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Example: What is the increase in the antenna temperature of our short dipole produced by thermal emission from the Sun at http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char17.png=10 GHz?

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=6http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−25 W Hz−1

so

TA=1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png38http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−23 J K−16http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png10−25 W Hz−1http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png046 K

**Beam Solid Angle**

The ***beam solid angle*** http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngA of a lossless antenna is defined as

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngAhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngPn(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)dhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png(3A11)

where Pn(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png) is the power pattern normalized to unity maximum:

Pn=GmaxG(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png) http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngA=1Gmaxhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngGdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png =4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngGmax

The beam solid angle is a useful parameter for estimating the antenna temperature produced by a compact source covering solid angle http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngs and having uniform brightness temperature TB.  "Compact" means that the source is much smaller than the beam so that the variation of Pn is small across the source.  The power per unit bandwidth received from the source by the antenna pointing at it is

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char19.pngAe(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)2Bhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.pngdhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png

Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2G(max)kTBhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngs=kTBhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngshttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngA http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Thus the antenna temperature TA=Phttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/120/char17.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngk is

TAhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.pngTBhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngshttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngA http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Stated in words, the antenna temperature equals the source brightness temperature multiplied by the fraction of the beam solid angle filled by the source.  A TB=104 K source covering 1% of the beam solid angle will add 100 K to the antenna temperature.

**Main Beam Solid Angle**

The ***main beam***of an antenna is defined as the region containing the principal response out to the first zero; responses outside this region are called ***sidelobes*** or, very far from the main beam, ***stray radiation***.  The ***main beam solid angle***http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngMB is defined as

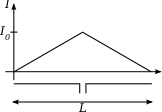
http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngMBhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char5A.pngMBPn(http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char12.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3B.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char1E.png)dhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.png(3A12)

The fraction of the total beam solid angle inside the main beam is called the **main beam efficiency** or, loosely, the ***beam efficiency****.*

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char11.pngBhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char11.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngAhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char0A.pngMB(3A13)

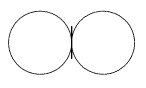
**UNIT II**

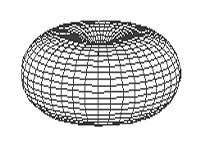
**APERTURE AND SLOT ANTENNA**

[](http://en.wikipedia.org/wiki/File:Short-dipole.svg)

A short dipole is a physically feasible dipole formed by two conductors with a total length \scriptstyle{L}very small compared with the wavelength \scriptstyle{\lambda}. The two conducting wires are fed at the centre of the dipole. We assume the hypothesis that the current is maximal at the centre (where the dipole is fed) and that it decreases linearly to be zero at the ends of the wires. Note that the direction of the current is the same in both the dipole branches - to the right in both or to the left in both. The far field \scriptstyle{E_\theta}of the electromagnetic wave radiated by this dipole is:

E_\theta={-iI_\circ\sin\theta\over 4\varepsilon_\circ c r}{L\over\lambda}e^{i\left(\omega t-kr\right)}.

[](http://en.wikipedia.org/wiki/File:Elem-doubl-rad-pat.jpg)

[](http://en.wikipedia.org/wiki/File:Elem-doub-rad-pat-pers.jpg)

Emission is maximal in the plane perpendicular to the dipole and zero in the direction of wires which is the direction of the current. The emission diagram is circular section [torus](http://en.wikipedia.org/wiki/Torus) shaped (right image) with zero inner diameter. In the left image the doublet is vertical in the [torus](http://en.wikipedia.org/wiki/Torus) centre.

Knowing this electric field, we can compute the total emitted power and then compute the resistive part of the series impedance of this dipole due to the radiated field, known as the [radiation resistance](http://en.wikipedia.org/wiki/Radiation_resistance):

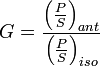
R_{series}={\pi\over6}Z_0 \left({L\over\lambda}\right)^2 (for \scriptstyle{L \ll \lambda}).

where *Z*0 is the [impedance of free space](http://en.wikipedia.org/wiki/Impedance_of_free_space). Using a common approximation of Z_0 \approx 120 \piohms, we get:

R_{series}\approx 20\pi^2\left({L\over\lambda}\right)^2 ohms

**Antenna gain**

[Antenna gain](http://en.wikipedia.org/wiki/Antenna_gain) is the ratio of surface power radiated by the antenna to the surface power radiated by a hypothetical [isotropic antenna](http://en.wikipedia.org/wiki/Isotropic_antenna):



The surface power carried by an electromagnetic wave is:

\textstyle{\left({P\over S}\right)_{ant}}=\textstyle{1\over2}c\varepsilon_\circ E_\theta^2\simeq\textstyle{{1\over120\pi}}E_\theta^2

The surface power radiated by an isotropic antenna feed with the same power is:

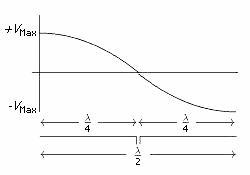
\textstyle{\left({P\over S}\right)_{iso}}=\textstyle{{1\over2} R_{series}I_\circ^2\over4\pi r^2}

Substituting values for the case of a short dipole, final result is:

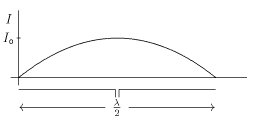
G=\textstyle{{\pi\left({L\over\lambda}\right)^2\over \varepsilon_\circ c{2\pi\over3\varepsilon_\circ c}\left({L\over\lambda}\right)^2}}= 1.5 = 1.76 dBi

**dBi** simply means [decibels](http://en.wikipedia.org/wiki/Decibels) gain, relative to an **i**sotropic antenna.

**Half-wave antenna**

[](http://en.wikipedia.org/wiki/File:Dipole_Antenna.jpg)

Typically a dipole antenna is formed by two quarter wavelength conductors or elements placed back to back for a total length of \scriptstyle{\lambda/2}. A standing wave on an element of a length ~\scriptstyle{\lambda/4} yields the greatest voltage differential, as one end of the element is at a node while the other is at an antinode of the wave. The larger the differential voltage, the greater the current between the elements.

[](http://en.wikipedia.org/wiki/File:Lambdaover2-antenna.jpg)

Assuming a sinusoidal distribution, the current impressed by this voltage differential is given by:

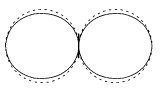
\textstyle{I=I_\circ e^{i\omega t}\cos{k\ell}}

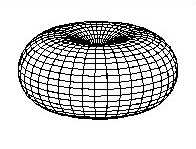
For the far-field case, the formula for the electric field of a radiating electromagnetic wave is somewhat more complex:

\textstyle{E_\theta={-iI_\circ\over 2\pi\varepsilon_\circ c r}}{\cos\left(\scriptstyle{\pi\over 2}\cos\theta\right)\over\sin\theta}e^{i\left(\omega t-kr\right)}

But the fraction \textstyle{{\cos\left(\scriptstyle{\pi\over 2}\cos\theta\right)\over\sin\theta}}is not very different from \scriptstyle{\sin\theta}.

The resulting emission diagram is a slightly flattened [torus](http://en.wikipedia.org/wiki/Torus).

[](http://en.wikipedia.org/wiki/File:L-over2-rad-pat.jpg)

[](http://en.wikipedia.org/wiki/File:L-over2-rad-pat-per.jpg)

The image on the left shows the section of the emission pattern. We have drawn, in dotted lines, the emission pattern of a short dipole. We can see that the two patterns are very similar. The image at right shows the perspective view of the same emission pattern.

This time it is not possible to compute analytically the total power emitted by the antenna (the last formula does not allow), though a simple numerical integration or series expansion leads to the more precise, actual value of the half-wave resistance:

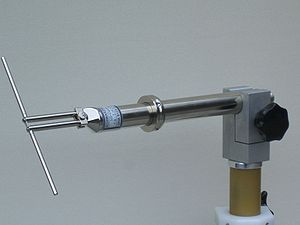
\begin{align}R_{\frac{\lambda}{2}}
&= \frac{Z_0}{2\pi}\left[\ln(2\pi\gamma)-\operatorname{Ci}(2\pi)\right]\approx 60\operatorname{Cin}(2\pi)=  60\left[\ln(2\pi\gamma)-\operatorname{Ci}(2\pi)\right]=120\int_{0}^{\frac{\pi}{2}}\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)^2}{\sin\theta}d\theta,\\
&=15\left[2\pi^2-\frac{1}{3}\pi^4+\frac{4}{135}\pi^6-\frac{1}{630}\pi^8+\frac{4}{70875}\pi^{10}\ldots-(-1)^n\frac{(2\pi)^{2n}}{n(2n)!}\right],\\
&\approx 73.1 \Omega;
\end{align}\,\!

This leads to the gain of a dipole antenna, G_{\frac{\lambda}{2}}\,\!:

\begin{align}G_{\frac{\lambda}{2}}
&=\frac{60^2}{30R_{\frac{\lambda}{2}}}=\frac{3600}{30R_{\frac{\lambda}{2}}}=\frac{120}{R_{\frac{\lambda}{2}}}=\frac{1}{{}^{\int_{0}^{\frac{\pi}{2}}\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)^2}{\sin\theta}d\theta}},\\
&\approx\frac{120}{73.1296}\approx 1.64\approx 2.15\,\mathrm{dBi};\end{align}\,\!

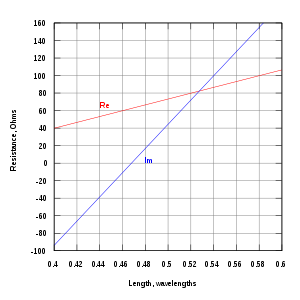
The resistance, however, is not enough to characterize the dipole impedance, as there is also an imaginary part——it is better to measure the impedance.

In the image below, the real and imaginary parts of a dipole's impedance are drawn for lengths going from \scriptstyle{0.4\,\lambda}\,\!to \scriptstyle{0.6\,\lambda}\,\!, accompanied by a chart comparing the gains of dipole antennas of other lengths, both as a number and in dBi:

[](http://en.wikipedia.org/wiki/File:Half_%E2%80%93_Wave_Dipole.jpg)

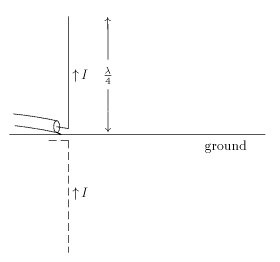
[http://bits.wikimedia.org/skins-1.18/common/images/magnify-clip.png](http://en.wikipedia.org/wiki/File:Half_%E2%80%93_Wave_Dipole.jpg)

UHF–Half–Wave Dipole, 1.0–4 GHz

[](http://en.wikipedia.org/wiki/File:Dipole_antenna_impedance.svg)

|  |  |  |
| --- | --- | --- |
| Gain of dipole antennas | | |
| length **L** in \scriptstyle{\lambda} | Gain | Gain(dBi) |
| \scriptstyle{\ll}0.1 | 1.50 | 1.76 |
| **0.5** | **1.64** | **2.15** |

**Quarter-wave antenna**

[](http://en.wikipedia.org/wiki/File:A6-3EN.jpg)

The antenna and its image form a \scriptstyle{{\lambda\over 2}}dipole that radiates only upward.

The quarter wave [monopole](http://en.wikipedia.org/wiki/Monopole_antenna) antenna is a single element antenna fed at one end, that behaves as a dipole antenna. It is formed by a conductor \scriptstyle{{\lambda\over 4}}in length. It is fed in the lower end, which is near a conductive surface which works as a reflector (see [Effect of ground](http://en.wikipedia.org/wiki/Antenna_%28radio%29#Effect_of_ground)). The current in the reflected image has the same direction and phase as the current in the real antenna. The quarter-wave conductor and its image together form a half-wave dipole that radiates only in the upper half of space.

In this upper side of space the emitted field has the same amplitude of the field radiated by a half-wave dipole fed with the same current. Therefore, the total emitted power is one-half the emitted power of a half-wave dipole fed with the same current. As the current is the same, the radiation resistance (real part of series impedance) will be one-half of the series impedance of a half-wave dipole. As the reactive part is also divided by 2, the impedance of a quarter wave antenna is \scriptstyle{{73+i43\over 2}=36+i21}ohms. Since the fields above ground are the same as for the dipole, but only half the power is applied, the gain is twice (3dB over) that for a half-wave dipole (\scriptstyle{{\lambda\over 2}}), that is 5.14 dBi.

The earth can be used as ground plane, but it is a poor conductor: the reflected antenna image is only clear at glancing angles (far from the antenna). At these glancing angles, electromagnetic fields and radiation patterns are thus the same as for a half-wave dipole.

Naturally, the impedance of the earth is far inferior to that of a good conductor ground plane -- this can be improved (at cost) by laying a copper mesh.

When ground is not available (such as in a vehicle) other metallic surfaces can serve as a ground plane (typically the vehicle's roof). Alternatively, radial wires placed at the base of the antenna can simulate a ground plane. For VHF bands, the radiating and ground-plane elements can be constructed from rigid rods or tubes.

**Antenna Array Notes**

For a dipole array antenna with the arrays lined up on the x-axis, the following field pattern characteristics are valid.

1. *Main-beam direction*

The maximum value of the array pattern factor, *F*(, ), occurs when

*kd*sinsin + = 0, which leads to

 .

Two special cases are of particular importance.

(a) *Broadside Array*

For a broadside array, maximum radiation occurs at a direction perpendicular to the line of the array: that is, = /2. This requires that the antennas are excited in phase, or = 0.

(b) *Endfire Array*

For an endfire array, maximum radiation occurs at = 0. The antennas are excited out of phase

**Array definitions.**

An array of antenna elements is a spatially extended collection of N similar radiators or elements, where N is a countable number bigger than 1, and the term "similar radiators" means that all the elements have the same polar radiation patterns, orientated in the same direction in 3-d space. The elements don't have to be spaced on a regular grid, neither do they have to have the same terminal voltages, but it is assumed that they are all fed with the same frequency and that one can define a fixed amplitude and phase angle for the drive voltage of each element.

**Element pattern, Array pattern.**

The polar radiation pattern of a single element is called the "element pattern". It is possible for the array to be built recursively; for example the element may itself be an array, as would be the case if we had an array of Yagi-Uda antennas. A [Yagi-Uda antenna](http://personal.ee.surrey.ac.uk/Personal/D.Jefferies/yagiuda.html) may be thought of as an array of dipoles with different amplitudes and phases of the dipole currents.

The array pattern is the polar radiation pattern which would result if the elements were replaced by isotropic radiators, having the same amplitude and phase of excitation as the actual elements, and spaced at points on a grid corresponding to the far field phase centres of the elements.

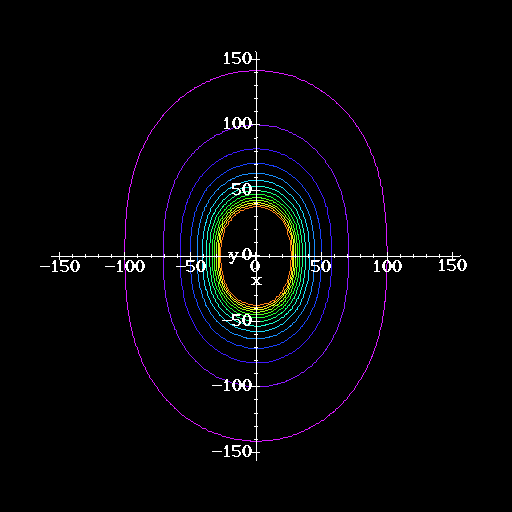
**Pattern multiplication.**

If we assume that all the polar radiation patterns of the elements taken individually are identical (within a certain tolerance) and that the patterns are all aligned in the same direction in azimuth and elevation, then the total array antenna pattern is got by multiplying the array pattern by the element pattern. It does not matter if we consider the patterns in question to be both patterns of radiated power, or both patterns of amplitudes having modulus and phase angle, provided the resulting pattern is interpreted as being of the same type as the original patterns.

**Calculation of array patterns**

The radiated field strength at a certain point in space, assumed to be in the far field, is calculated by adding the contributions of each element to the total radiated fields. The field strengths fall off as 1/r where r is the distance from the isotrope to the field point. We must take into account any phase angle of the isotrope excitation, and also the phase delay which is due to the time it takes the signal to get from the source to the field point. This phase delay is expressed as 2 Pi radians times (r/lambda) where lambda is the free space wavelength of the radiation. Contours of equal field strength may be interpreted as an amplitude polar radiation pattern. Contours of the squared modulus of the field strength may be interpreted as a power polar radiation pattern.

Here is an example of a power polar radiation pattern for two isotropes spaced 1/4 wavelength apart along the x axis (horizontally on your screen or paper) and fed with equal amplitudes and phases......-->



TWO ISOTROPES 1/4 WAVELENGTH APART FED IN PHASE

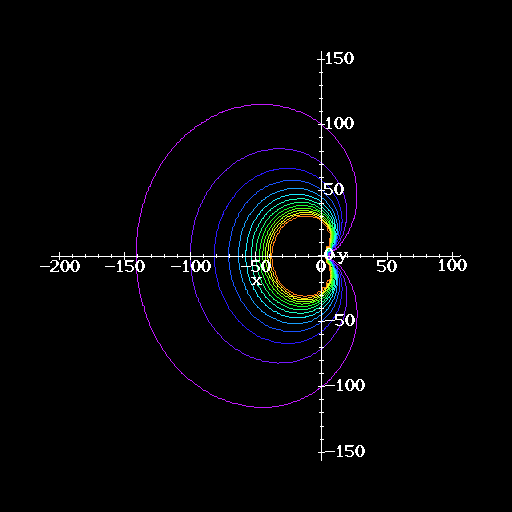
If we increase the spacing to 1/2 wavelength, but still keep the excitation in phase and equal amplitudes, we see deep nulls developing.......-->



TWO ISOTROPES 1/2 WAVELENGTH APART FED IN PHASE

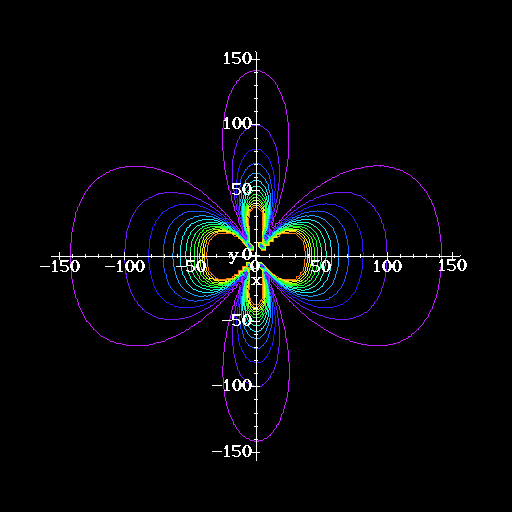
If we restore the original 1/4 wavelength spacing and feed the isotropes in phase quadrature (that is, there is Pi/2 phase shift between the excitations), we see a single lobe develop. This pattern is close to a Cardioid......-->

(To make the lobe face the other direction we would have to reverse the relative phase, and feed the second isotrope at -Pi/2 phase angle.)



TWO ISOTROPES 1/4 WAVELENGTH APART 90 DEGREES PHASE SHIFT

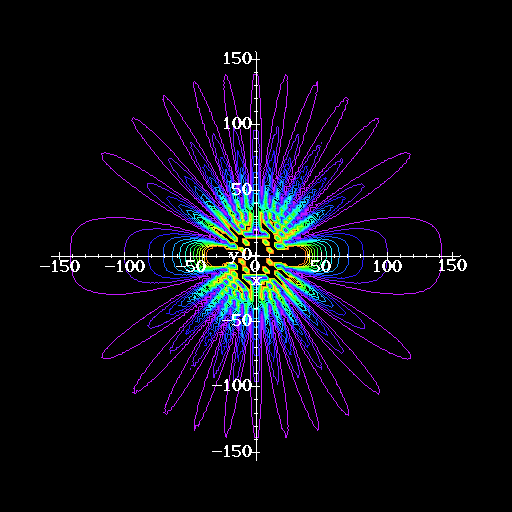
Returning now to feeding the two isotropes in phase, as we increase the spacing we see more sub-lobes or sidelobes develop. Here is a picture for spacing of a whole wavelength....-->



TWO ISOTROPES 1 WAVELENGTH APART, IN PHASE

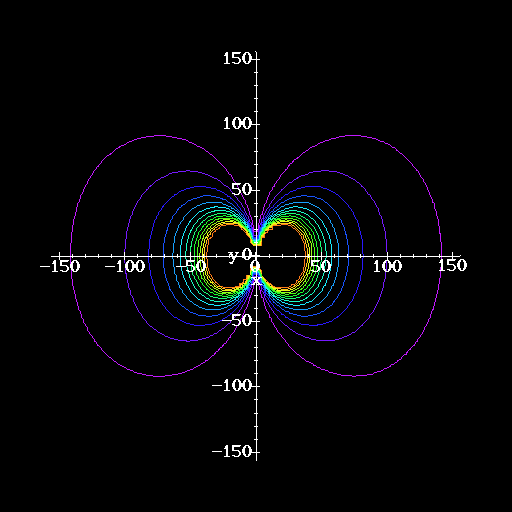
And if we increase the spacing to six whole wavelengths, there are large numbers of lobes developing......-->

This is a method used to make sharp beams for radio telescopes using ["Very Long Baseline Interferometry".](http://www.atnf.csiro.au/people/atzioumi/vlbi/) In this method two elements are spaced thousands of kilometres apart, which makes the individual lobes very narrow indeed.



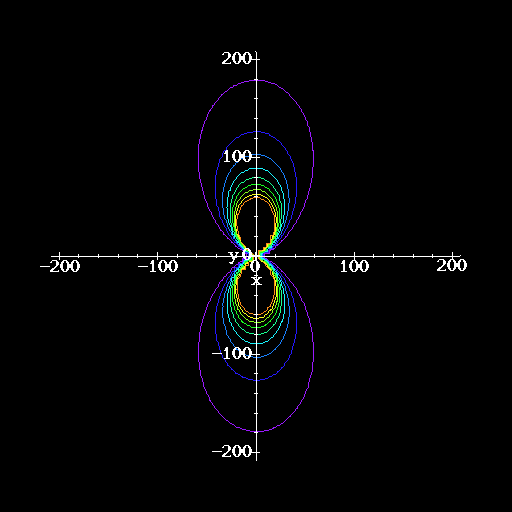
TWO ISOTROPES 6 WAVELENGTHS APART IN PHASE

If we now take two isotropes spaced by a half a wavelength and feed the elements in antiphase (Pi phase difference) we see the pattern with horizontal lobes rather than vertical lobes. This radiation pattern is similar to that of a vertical dipole, which may roughly be regarded as a "doublet" consisting of two isotropes fed in antiphase. The shape of the radiation pattern is not very dependent on the separation.....-->



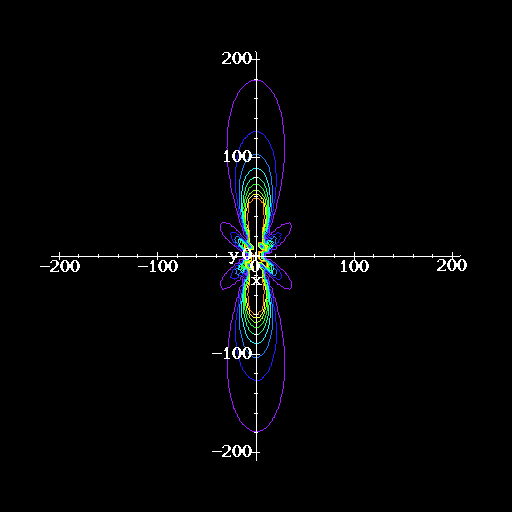
TWO ISOTROPES 1/2 WAVELENGTH APART ANTIPHASE

Now we turn to arrays of four elements, just to illustrate the ideas further. Equally excited (same amplitudes and phases) and spaced along the x axis at intervals of 1/4 wavelength, we see......-->



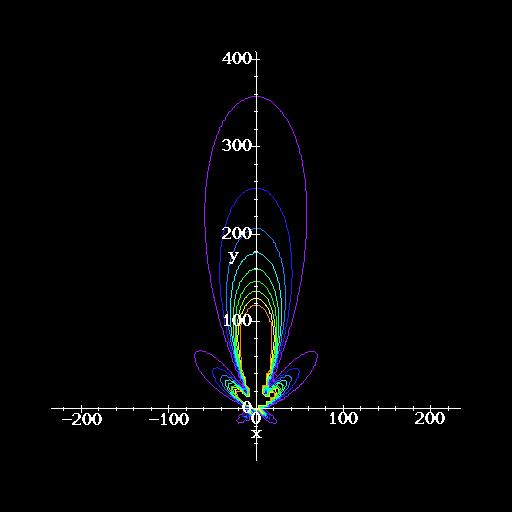
FOUR ISOTROPES SPACED 1/4 WAVELENGTH IN PHASE

and again, if we increase the spacing to 1/2 wavelength we get more sidelobes....-->



FOUR ISOTROPES SPACED 1/2 WAVELENGTH, IN PHASE.

To see the pattern multiplication, we add a row of 4 extra isotropes, spaced by 1/4 wavelength and fed in phase quadrature, below the original row of 4 sources. This has the effect of multiplying the pattern seen above by the cardioid pattern seen earlier. Of course, the cardioid has the main lobe facing upwards because the quadrature array is spaced out along the y axis.......-->

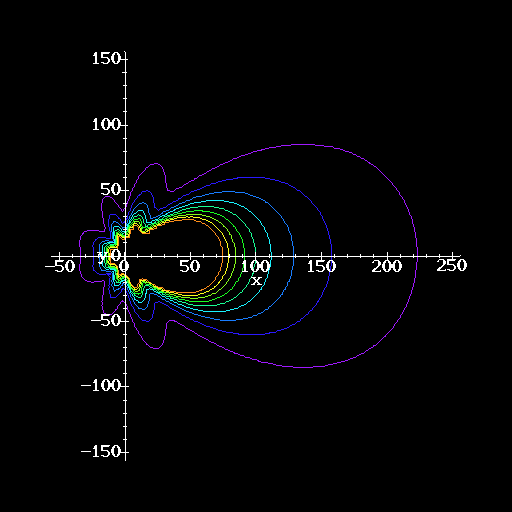


PATTERN MULTIPLICATION. 8 ISOTROPES 4x2.   
1/2 WAVELENGTH SPACING FOR 4 ELEMENTS ALONG X (IN PHASE)   
1/4 WAVELENGTH SPACING TO 4 ELEMENTS ALONG Y (IN QUADRATURE)

x x x x

x x x x

Finally, let us take a guess at the amplitudes and phases of the elements in a 4-element Yagi-Uda antenna. The pattern we arrive at looks like this........



**UNIT III**

**ANTENNA ARRAYS**

In [electromagnetics](http://en.wikipedia.org/wiki/Electromagnetics) and [antenna](http://en.wikipedia.org/wiki/Antenna_%28radio%29) theory, **antenna aperture** or **effective area** is a measure of how effective an antenna is at receiving the power of [radio waves](http://en.wikipedia.org/wiki/Radio_wave). The aperture is defined as the area, oriented perpendicular to the direction of an incoming radio wave, which would intercept the same amount of power from that wave as is produced by the antenna receiving it. At any point, a beam of radio waves has an [*irradiance*](http://en.wikipedia.org/wiki/Irradiance) or *power flux density* *(PFD)* which is the amount of radio power passing through a unit area. If an antenna delivers an output power of *Po* [watts](http://en.wikipedia.org/wiki/Watt) to the load connected to its output terminals when irradiated by a uniform field of power density *PFD* watts per square metre, the antenna's aperture *Aeff* in square metres is given by:[[1]](http://en.wikipedia.org/wiki/Antenna_aperture#cite_note-Bakshi-0)

A_{eff} = \frac {P_o}{PFD} \,.

So the power output of an antenna in watts is equal to the power density of the radio waves in watts per square metre, multiplied by its aperture in square metres. The larger an antenna's aperture is, the more power it can collect from a given field of radio waves. To actually obtain the predicted power available *Po*, the [polarization](http://en.wikipedia.org/wiki/Polarization_%28waves%29) of the incoming waves must match the polarization of the antenna, and the load (receiver) must be [impedance matched](http://en.wikipedia.org/wiki/Impedance_match) to the antenna's feedpoint impedance.

Although this concept is based on an antenna receiving a radio frequency wave, knowing *Aeff* directly supplies the [(power) gain](http://en.wikipedia.org/wiki/Antenna_gain) of that antenna. Due to [reciprocity](http://en.wikipedia.org/wiki/Reciprocity_%28electromagnetism%29), an antenna's gain in receiving and transmitting are identical. Therefore *Aeff* can just as well be used to compute the performance of a transmitting antenna. Note that *Aeff* is a function of the direction of the radio wave relative to the orientation of the antenna, since the gain of an antenna varies according to its [radiation pattern](http://en.wikipedia.org/wiki/Radiation_pattern). When no direction is specified, *Aeff* is understood to refer to its maximum value, that is, the direction that the antenna would be pointed to take advantage of its maximum gain.

|  |
| --- |
|  |

**Aperture efficiency**

In general, the aperture of an antenna is not directly related to its physical size.[[2]](http://en.wikipedia.org/wiki/Antenna_aperture#cite_note-Narayan-1) However some types of antennas, for example [parabolic dishes](http://en.wikipedia.org/wiki/Parabolic_antenna) and [horns](http://en.wikipedia.org/wiki/Horn_antenna), have a physical aperture (opening) which collects the radio waves. In these *aperture antennas*, the effective aperture *Aeff* is always less than the area of the physical aperture *Aphys* of the antenna. An antenna's *aperture efficiency*, ea is defined as the ratio of these two areas:

e_a = \frac {A_{eff}} {A_{phys}} \,

The aperture efficiency is a dimensionless parameter between 0 and 1.0 that measures how far the antenna falls short of using all the radio power entering its physical aperture. If the antenna were perfectly efficient, all the radio power falling within its physical aperture would be converted to electrical power delivered to the load attached to its output terminals, so these two areas would be equal *Aeff* = *Aphys* and the aperture efficiency would be 1.0. But all antennas have losses, such as power dissipated as heat in the resistance of its elements, nonuniform illumination by its [feed](http://en.wikipedia.org/wiki/Antenna_feed), and radio waves scattered by structural supports and diffraction at the aperture edge, which reduce the power output. Aperture efficiencies of typical antennas vary from 0.35 to 0.70 but can range up to 0.90.

**Aperture and gain**

The [directivity](http://en.wikipedia.org/wiki/Directivity) of an antenna, its ability to direct radio waves in one direction or receive from a single direction, is measured by a parameter called its [gain](http://en.wikipedia.org/wiki/Antenna_gain), which is the ratio of the power received by the antenna to the power that would be received by a hypothetical [isotropic antenna](http://en.wikipedia.org/wiki/Isotropic_radiator), which receives power equally well from all directions.

It can be shown that the aperture of a lossless isotropic antenna, which by definition has unity gain, is:

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

where *λ* is the [wavelength](http://en.wikipedia.org/wiki/Wavelength) of the radio waves. So the gain of any antenna is proportional to its aperture:

G = \frac {4 \pi A_{eff} } { \lambda^2 }  = \frac {4 \pi A_{phys} e_a } { \lambda^2 }   \,

So antennas with large effective apertures are [high gain antennas](http://en.wikipedia.org/wiki/High_gain_antenna), which have small angular [beam widths](http://en.wikipedia.org/wiki/Beam_width). Most of their power is radiated in a narrow beam in one direction, and little in other directions. As receiving antennas, they are most sensitive to radio waves coming from one direction, and are much less sensitive to waves coming from other directions. Although these terms can be used as a function of direction, when no direction is specified, the gain and aperture are understood to refer to the antenna's axis of maximum gain, or [boresight](http://en.wikipedia.org/wiki/Boresight).

**Friis transmission equation**

Main article: [Friis transmission equation](http://en.wikipedia.org/wiki/Friis_transmission_equation)

The fraction of the power delivered to a transmitting antenna that is received by a receiving antenna is proportional to the product of the apertures of both the antennas. This is given by a form of the [Friis transmission equation](http://en.wikipedia.org/wiki/Friis_transmission_equation):.[[2]](http://en.wikipedia.org/wiki/Antenna_aperture#cite_note-Narayan-1)

P_r = \frac { A_t A_r }{ r^2 \lambda^2 } P_t \,

where

*Pr* is the power delivered by the receiving antenna in watts

*Pt* is the power applied to the transmitting antenna in watts

*Ar* is the aperture of the receiving antenna in m2

*At* is the aperture of the transmitting antenna in m2

*r* is the distance between the antennas in m

*λ* is the wavelength of the radio waves in m

**Effective length**

For antennas which are not defined by a physical area, such as [monopoles](http://en.wikipedia.org/wiki/Monopole_antenna) and [dipoles](http://en.wikipedia.org/wiki/Dipole_antenna) consisting of thin rod [conductors](http://en.wikipedia.org/wiki/Electrical_conductor), the aperture bears no obvious relation to the size or area of the antenna. An alternate measure of antenna gain that has a greater relationship to the physical structure of such antennas is *effective length* *leff* measured in metres, which is defined for a receiving antenna as:[[3]](http://en.wikipedia.org/wiki/Antenna_aperture#cite_note-Rudge-2)

l_{eff} = V_0 / E_s \, 

where

*V0* is the open circuit voltage appearing across the antenna's terminals

*Es* is the electric [field strength](http://en.wikipedia.org/wiki/Field_strength) of the radio signal, in [volts](http://en.wikipedia.org/wiki/Volt) per metre, at the antenna.

The longer the effective length the more voltage and therefore the more power the antenna will receive. Note, however, that an antenna's gain or *Aeff* increases according to the *square* of *leff*, and that this proportionality also involves the antenna's [radiation resistance](http://en.wikipedia.org/wiki/Radiation_resistance). Therefore this measure is of more theoretical than practical value and is not, by itself, a useful figure of merit relating to an antenna's directivity.

**The method of images**

Suppose that we have a point charge $q$held a distance $d$from an infinite, grounded, conducting plate. Let the plate lie in the $x$-$y$ plane, and suppose that the point charge is located at coordinates (0, 0, $d$). What is the scalar potential above the plane? This is not a simple question because the point charge induces surface charges on the plate, and we do not know how these are distributed.

What do we know in this problem? We know that the conducting plate is an equipotential surface. In fact, the potential of the plate is zero, since it is grounded. We also know that the potential at infinity is zero (this is our usual boundary condition for the scalar potential). Thus, we need to solve Poisson's equation in the region $z>0$, for a single point charge $q$at position (0, 0, $d$), subject to the boundary conditions

|  |  |
| --- | --- |
| \begin{displaymath} \phi(z=0) = 0, \end{displaymath} | (710) |

and

|  |  |
| --- | --- |
| \begin{displaymath} \phi\rightarrow 0 \end{displaymath} | (711) |

as $x^2+y^2+z^2\rightarrow\infty$. Let us forget about the real problem, for a moment, and concentrate on a slightly different one. We refer to this as the *analogue problem*. In the analogue problem, we have a charge $q$located at (0, 0, $d$) and a charge $-q$located at (0, 0, -$d$), with no conductors present. We can easily find the scalar potential for this problem, since we know where all the charges are located. We get

|  |  |
| --- | --- |
| \begin{displaymath} \phi_{\rm analogue} (x, y, z) = \frac{1}{4\pi \epsilon_0} \... ...{x^2+y^2+(z-d)^2}}- \frac{q}{\sqrt{x^2+y^2+ (z+d)^2}}\right\}. \end{displaymath} | (712) |

Note, however, that

|  |  |
| --- | --- |
| \begin{displaymath} \phi_{\rm analogue}(z=0) = 0, \end{displaymath} | (713) |

and

|  |  |
| --- | --- |
| \begin{displaymath} \phi_{\rm analogue}\rightarrow 0 \end{displaymath} | (714) |

as $x^2+y^2+z^2\rightarrow\infty$. In addition, $\phi_{\rm analogue}$satisfies Poisson's equation for a charge at (0, 0, $d$), in the region $z>0$. Thus, $\phi_{\rm analogue}$is a solution to the problem posed earlier, in the region $z>0$. Now, the uniqueness theorem tells us that there is only *one* solution to Poisson's equation which satisfies a given, well-posed set of boundary conditions. So, $\phi_{\rm analogue}$must be the correct potential in the region $z>0$. Of course, $\phi_{\rm analogue}$is completely wrong in the region $z<0$. We know this because the grounded plate shields the region $z<0$from the point charge, so that $\phi=0$in this region. Note that we are leaning pretty heavily on the uniqueness theorem here! Without this theorem, it would be hard to convince a skeptical person that $\phi = \phi_{\rm analogue}$is the correct solution in the region $z>0$.

Now that we know the potential in the region $z>0$, we can easily work out the distribution of charges induced on the conducting plate. We already know that the relation between the electric field immediately above a conducting surface and the density of charge on the surface is

|  |  |
| --- | --- |
| \begin{displaymath} E_\perp = \frac{\sigma}{\epsilon_0}. \end{displaymath} | (715) |

In this case,

|  |  |
| --- | --- |
| \begin{displaymath} E_\perp = E_z(z=0_+) = - \frac{\partial \phi(z=0_+)}{\partial z} = - \frac{\partial \phi_{\rm analogue}(z=0_+)}{\partial z}, \end{displaymath} | (716) |

so

|  |  |
| --- | --- |
| \begin{displaymath} \sigma = - \epsilon_0 \frac{\partial\phi_{\rm analogue}(z=0_+)}{\partial z}. \end{displaymath} | (717) |

It follows from Eq. ([712](http://farside.ph.utexas.edu/teaching/em/lectures/node64.html#e5.114)) that

|  |  |
| --- | --- |
| \begin{displaymath} \frac{\partial\phi}{\partial z} = \frac{q}{4\pi  \epsilon_0... ...-d)^2]^{3/2}} + \frac{(z+d)}{[x^2+y^2+(z+d)^2]^{3/2}}\right\}, \end{displaymath} | (718) |

so

|  |  |
| --- | --- |
| \begin{displaymath} \sigma(x,y) = - \frac{q d}{2\pi  (x^2+y^2+d^2)^{3/2}}. \end{displaymath} | (719) |

Clearly, the charge induced on the plate has the opposite sign to the point charge. The charge density on the plate is also symmetric about the $z$-axis, and is largest where the plate is closest to the point charge. The total charge induced on the plate is

|  |  |
| --- | --- |
| \begin{displaymath} Q = \int_{x-y  \rm plane} \sigma dS, \end{displaymath} | (720) |

which yields

|  |  |
| --- | --- |
| \begin{displaymath} Q = - \frac{q d}{2\pi} \int_0^\infty \frac{2\pi  r dr}{(r^2+d^2)^{3/2}}, \end{displaymath} | (721) |

where $r^2 = x^2+y^2$. Thus,

|  |  |
| --- | --- |
| \begin{displaymath} Q = - \frac{q d}{2} \int_0^\infty \frac{dk}{(k+d^2)^{3/2}} = q d\left[ \frac{1}{(k+d^2)^{1/2}}\right]_0^\infty = - q. \end{displaymath} | (722) |

So, the total charge induced on the plate is equal and opposite to the point charge which induces it.

Our point charge induces charges of the opposite sign on the conducting plate. This, presumably, gives rise to a force of attraction between the charge and the plate. What is this force? Well, since the potential, and, hence, the electric field, in the vicinity of the point charge is the same as in the analogue problem, then the force on the charge must be the same as well. In the analogue problem, there are two charges $\pm q$a net distance $2 d$apart. The force on the charge at position (0, 0, $d$) (*i.e.*, the real charge) is

|  |  |
| --- | --- |
| \begin{displaymath} {\bf F} = - \frac{1}{4\pi \epsilon_0} \frac{q^2}{(2 d)^2}  \hat{\bf z}. \end{displaymath} | (723) |

What, finally, is the potential energy of the system. For the analogue problem this is just

|  |  |
| --- | --- |
| \begin{displaymath} W_{\rm analogue} = - \frac{1}{4\pi \epsilon_0} \frac{q^2}{2 d}. \end{displaymath} | (724) |

Note that the fields on opposite sides of the conducting plate are mirror images of one another in the analogue problem. So are the charges (apart from the change in sign). This is why the technique of replacing conducting surfaces by imaginary charges is called the *method of images*. We know that the potential energy of a set of charges is equivalent to the energy stored in the electric field. Thus,

|  |  |
| --- | --- |
| \begin{displaymath} W = \frac{\epsilon_0}{2} \int_{\rm all space} E^{2} dV. \end{displaymath} | (725) |

In the analogue problem, the fields on either side of the $x$-$y$ plane are mirror images of one another, so $E^2(x, y, z) = E^2(x, y, -z)$. It follows that

|  |  |
| --- | --- |
| \begin{displaymath} W_{\rm analogue} = 2  \frac{\epsilon_0}{2} \int_{z>0} E^2_{\rm analogue}  dV. \end{displaymath} | (726) |

In the real problem

|  |  |  |  |
| --- | --- | --- | --- |
| $\displaystyle {\bf E} (z>0)$ | $\textstyle =$ | $\displaystyle {\bf E}_{\rm analogue}(z>0),$ | (727) |
| $\displaystyle {\bf E}(z<0)$ | $\textstyle =$ | $\displaystyle {\bf0}.$ | (728) |

So,

|  |  |
| --- | --- |
| \begin{displaymath} W = \frac{\epsilon_0}{2} \int_{z>0} E^2 dV= \frac{\epsilon_... ...>0} E_{\rm analogue}^2  dV = \frac{1}{2}  W_{\rm analogue}, \end{displaymath} | (729) |

giving

|  |  |
| --- | --- |
| \begin{displaymath} W = - \frac{1}{4\pi  \epsilon_0} \frac{q^2}{4 d}. \end{displaymath} | (730) |

There is another method by which we can obtain the above result. Suppose that the charge is gradually moved towards the plate along the $z$-axis from infinity until it reaches position (0, 0, $d$). How much work is required to achieve this? We know that the force of attraction acting on the charge is

|  |  |
| --- | --- |
| \begin{displaymath} F_z = - \frac{1}{4\pi  \epsilon_0} \frac{q^2}{4  z^2}. \end{displaymath} | (731) |

Thus, the work required to move this charge by $dz$is

|  |  |
| --- | --- |
| \begin{displaymath} d W = - F_z dz=\frac{1}{4\pi \epsilon_0} \frac{q^2}{4 z^2} dz. \end{displaymath} | (732) |

The total work needed to move the charge from $z=\infty$to $z= d$is

|  |  |
| --- | --- |
| \begin{displaymath} W = \frac{1}{4\pi  \epsilon_0}\int_{\infty}^d \frac{q^2}{4\... ...t]_{\infty}^d = - \frac{1}{4\pi \epsilon_0} \frac{q^2}{4 d}. \end{displaymath} | (733) |

Of course, this work is equivalent to the potential energy we evaluated earlier, and is, in turn, the same as the energy contained in the electric field.

As a second example of the method of images, consider a grounded spherical conductor of radius $a$placed at the origin. Suppose that a charge $q$is placed outside the sphere at $(b, 0,  0)$, where $b>a$. What is the force of attraction between the sphere and the charge? In this case, we proceed by considering an analogue problem in which the sphere is replaced by an image charge $-q'$placed somewhere on the $x$-axis at $(c, 0, 0)$. The electric potential throughout space in the analogue problem is simply

|  |  |
| --- | --- |
| \begin{displaymath} \phi = \frac{q}{4\pi \epsilon_0} \frac{1}{[(x-b)^2+y^2+z^2... ...\frac{q'}{4\pi \epsilon_0} \frac{1}{[(x-c)^2+y^2+z^2]^{1/2}}. \end{displaymath} | (734) |

The image charge is chosen so as to make the surface $\phi=0$correspond to the surface of the sphere. Setting the above expresion to zero, and performing a little algebra, we find that the $\phi=0$surface satisfies

|  |  |
| --- | --- |
| \begin{displaymath} x^2 + \frac{2 (c-\lambda b)}{\lambda-1}  x + y^2 + z^2 = \frac{c^2-\lambda b^2}{\lambda-1}, \end{displaymath} | (735) |

where $\lambda=q'^{ 2}/q^2$. Of course, the surface of the sphere satisfies

|  |  |
| --- | --- |
| \begin{displaymath} x^2+y^2+z^2 = a^2. \end{displaymath} | (736) |

The above two equations can be made identical by setting $\lambda = c/b$and $a^2=\lambda b^2$, or

|  |  |
| --- | --- |
| \begin{displaymath} q' = \frac{a}{b} q, \end{displaymath} | (737) |

and

|  |  |
| --- | --- |
| \begin{displaymath} c = \frac{a^2}{b}. \end{displaymath} | (738) |

According to the uniqueness theorem, the potential in the analogue problem is now identical with that in the real problem, outside the sphere. (Of course, in the real problem, the potential inside the sphere is zero.) Hence, the force of attraction between the sphere and the original charge in the real problem is the same as the force of attraction between the two charges in the analogue problem. It follows that

|  |  |
| --- | --- |
| \begin{displaymath} f = \frac{q^2}{4\pi \epsilon_0 (b-c)^2} = \frac{q q'}{4\pi} \frac{a b}{(b^2-a^2)^2}. \end{displaymath} | (739) |

There are many other image problems, each of which involves replacing a conductor with an imaginary charge (or charges) which mimics the electric field in some region (but not everywhere). Unfortunately, we do not have time to discuss any more of these problems.

**Slot antennas** are used typically at frequencies between 300 MHz and 24 GHz. The slot antenna is popular because they can be cut out of whatever surface they are to be mounted on, and have [radiation patterns](http://www.antenna-theory.com/basics/radPattern.html) that are roughly omnidirectional (similar to a linear wire antenna, as we'll see). The polarization of the slot antenna is linear. The slot size, shape and what is behind it (the cavity) offer design variables that can be used to tune performance.

Consider an infinite conducting sheet, with a rectangular slot cut out of dimensions *a* and *b*, as shown in Figure 1. If we can excite some reasonable fields in the slot (often called the aperture), we have a slot antenna.

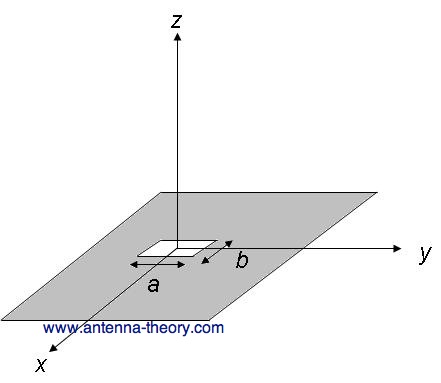


Figure 1. Rectangular Slot antenna with dimensions *a* and *b*.

To gain an intuition about slot antennas, first we'll learn Babinet's principle (put into antenna terms by H. G. Booker in 1946). This principle relates the radiated fields and impedance of an aperture or slot antenna to that of the field of its dual antenna. The dual of a slot antenna would be if the conductive material and air were interchanged - that is, the slot antenna became a metal slab in space. An example of dual antennas is shown in Figure 2:

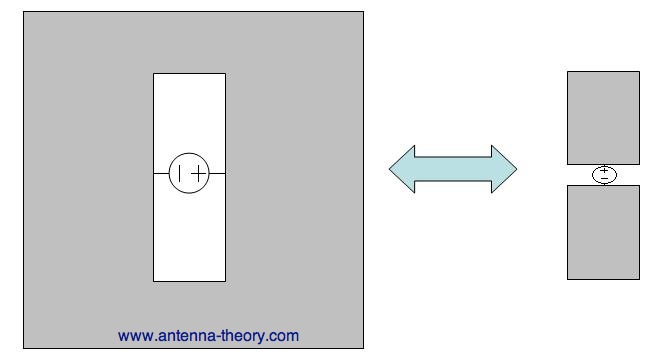
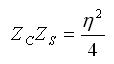


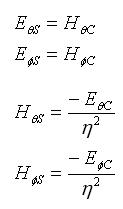
Figure 2. Dual antennas - (left) the slot antenna, (right) the dipole antenna.

Note that a voltage source is applied across the short end of the slot antenna. This induces an E-field distribution within the slot, and currents that travel around the slot perimeter, both contributed to radiation. The dual antenna is similar to a [dipole antenna](http://www.antenna-theory.com/antennas/dipole.php). The voltage source is applied at the center of the dipole, so that the voltage source is rotated.

Babinet's principle relates these two antennas. The first result states that the [impedance](http://www.antenna-theory.com/basics/impedance.php) of the slot antenna (impedance of slot) is related to the impedance of its dual antenna (antenna impedance of slot's dual) by the relation:



In the above, impedance of free spaceis the [intrinsic impedance](http://www.antenna-theory.com/definitions/intrinsicimpedance.php) of free space. The second major result of Babinet's/Booker's principle is that the fields of the dual antenna are almost the same as the slot antenna (the fields components are interchanged, and called "duals"). That is, the fields of the slot antenna (given with a subscript *S*) are related to the fields of it's complement (given with a subscript *C*) by:



Hence, if we know the fields from one antenna we know the fields of the other antenna. Hence, since it is easy to visualize the fields from a [dipole antenna](http://www.antenna-theory.com/antennas/dipole.php), the fields and impedance from a slot antenna can become intuitive if Babinet's principle is understood.

### Introduction to Horn Antennas

**Horn antennas** are very popular at UHF (300 MHz-3 GHz) and higher frequencies (I've heard of horn antennas operating as high as 140 GHz). Horn antennas often have a directional [radiation pattern](http://www.antenna-theory.com/basics/radPattern.html) with a high [antenna gain](http://www.antenna-theory.com/basics/gain.php), which can range up to 25 dB in some cases, with 10-20 dB being typical. Horn antennas have a wide impedance [bandwidth](http://www.antenna-theory.com/basics/bandwidth.php), implying that the [input impedance](http://www.antenna-theory.com/basics/impedance.php) is slowly varying over a wide frequency range (which also implies low values for [S11](http://www.antenna-theory.com/definitions/sparameters.php) or [VSWR](http://www.antenna-theory.com/definitions/vswr.php)). The bandwidth for practical horn antennas can be on the order of 20:1 (for instance, operating from 1 GHz-20 GHz), with a 10:1 bandwidth not being uncommon.

The gain of horn antennas often increases (and the [beamwidth](http://www.antenna-theory.com/basics/radPatDefs.php) decreases) as the frequency of operation is increased. This is because the size of the horn aperture is always measured in wavelengths; at higher frequencies the horn antenna is "electrically larger"; this is because a higher frequency has a smaller wavelength. Since the horn antenna has a fixed physical size (say a square aperture of 20 cm across, for instance), the aperture is more wavelengths across at higher frequencies. And, a recurring theme in antenna theory is that larger antennas (in terms of wavelengths in size) have higher directivities.

Horn antennas have very little loss, so the [directivity](http://www.antenna-theory.com/basics/directivity.php) of a horn is roughly equal to its gain.

Horn antennas are somewhat intuitive and relatively simple to manufacture. In addition, acoustic horn antennas are also used in transmitting sound waves (for example, with a megaphone). Horn antennas are also often used to feed a dish antenna, or as a "standard gain" antenna in measurements.

Popular versions of the horn antenna include the E-plane horn, shown in Figure 1. This horn antenna is flared in the E-plane, giving the name. The horizontal dimension is constant at **w**.

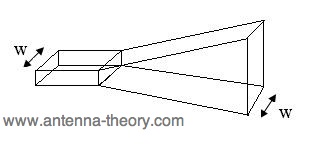


Figure 1. E-plane horn antenna.

Another example of a horn antenna is the H-plane horn, shown in Figure 2. This horn is flared in the H-plane, with a constant height for the waveguide and horn of *h*.

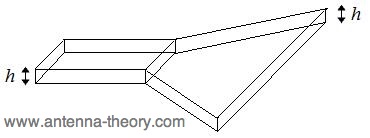


Figure 2. H-Plane horn antenna.

The most popular horn antenna is flared in both planes as shown in Figure 3. This is a pyramidal horn, and has a width *B* and height *A* at the end of the horn.

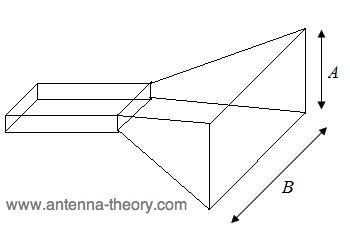


Figure 3. Pyramidal horn antenna.

Horn antennas are typically fed by a section of a waveguide, as shown in Figure 4. The waveguide itself is often fed with a [short dipole](http://www.antenna-theory.com/antennas/shortdipole.php), which is shown in red in Figure 4. A waveguide is simply a hollow, metal cavity (see [the waveguide tutorial](http://www.antenna-theory.com/tutorial/waveguides/waveguide.php)). Waveguides are used to guide electromagnetic energy from one place to another. The waveguide in Figure 4 is a rectangular waveguide of width *b* and height *a*, with *b*>*a*. The E-field distribution for the dominant mode is shown in the lower part of Figure 1.

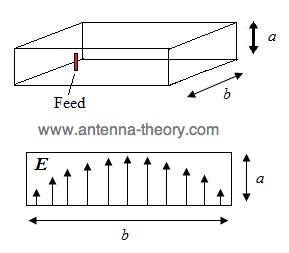


Figure 4. Waveguide used as a feed to horn antennas

### Fields and Geometrical Parameters for Horn Antennas

Antenna texts typically derive very complicated functions for the radiation patterns of horn antennas. To do this, first the E-field across the aperture of the horn antenna is assumed to be known, and the far-field radiation pattern is calculated using the radiation equations. While this is conceptually straight forward, the resulting field functions end up being extremely complex, and personally I don't feel add a whole lot of value. If you would like to see these derivations, pick up any antenna textbook that has a section on horn antennas. (Also, as a practicing antenna engineer, I can assure you that we never use radiation integrals to estimate patterns. We always go on previous experience, computer simulations and measurements.)

Instead of the traditional academic derivation approach, I'll state some results for the horn antenna and show some typical radiation patterns, and attempt to provide a feel for the design parameters of horn antennas. Since the pyramidal horn antenna is the most popular, we'll analyze that. The E-field distribution across the aperture of the horn antenna is what is responsible for the radiation.

The radiation pattern of a horn antenna will depend on *B* and *A* (the dimensions of the horn at the opening) and *R* (the length of the horn, which also affects the flare angles of the horn), along with *b* and *a* (the dimensions of the waveguide). These parameters are optimized in order to taylor the performance of the horn antenna, and are illustrated in the following Figures.

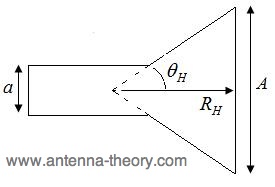


Figure 5. Cross section of waveguide, cut in the H-plane.

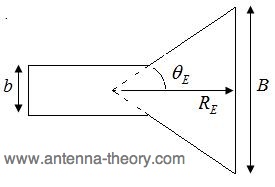


Figure 6. Cross section of waveguide, cut in the E-plane.

Observe that the flare angles (flare angle in the E-plane and flare angle for horn antenna) depend on the height, width and length of the horn antenna.

Given the coordinate system of Figure 6 (which is centered at the opening of the horn), the radiation will be maximum in the +z-direction (out of the screen).

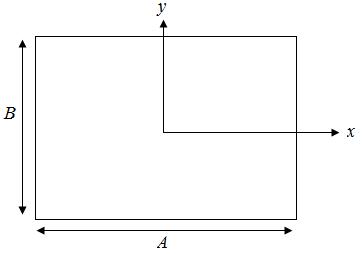
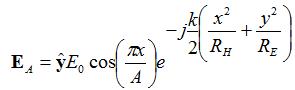


Figure 7. Coordinate system used, centered on the horn antenna opening.

The E-field distribution across the opening of the horn antenna can be approximated by:



The E-field in the far-field will be linearly polarized, and the magnitude will be given by:

far fields radiated from a horn antenna

The above equation states that the far-fields of the horn antenna is [the Fourier Transform](http://www.thefouriertransform.com) of the fields at the opening of the horn. Many textbooks evaluate this integral, and end up with supremely complicated functions, that I don't feel sheds a whole lot of light on the patterns.

# Reflector Antennas

**Paraboloidal Reflectors**

Antennas useful for radio astronomy at short wavelengths must have collecting areas much larger than the collecting area http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png(4http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.png) of an isotropic antenna and much higher angular resolution than a short dipole provides. Since arrays of dipoles are impractical at wavelengths http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3C.png1 m or so, most radio telescopes use large ***reflectors*** to collect and focus power onto the simple ***feed antennas***, such as waveguide horns or dipoles backed by small reflectors, that are connected to receivers. The most common reflector shape is a paraboloid of revolution because it can focus the plane wave from a distant point source onto a single ***focal point***.

The reflector shape that can focus plane waves onto a single point must keep all parts of an on-axis plane wavefront in phase at the focal point. Thus the total path lengths to the focus must all be the same, and this requirement is sufficient to determine the shape of the desired reflecting surface. Clearly the surface must be rotationally symmetric about its axis. In any plane containing the axis, the surface looks like the curve below.

*A plane containing the axis of a* ***paraboloidal reflector*** *with focal length f. Plane wave fronts from a distant point source are shown as dotted lines perpendicular to the z axis.  From a wavefront at height h above the vertex, the ray path (dashed line) lengths at all radial offsets r down to the reflector and up to the focal point at z=f must be the same.*

The requirement of constant path length can be written by equating the on-axis path length (f+h) from any height h to the reflector and then back to the focus at height f with the off-axis path length:

(f+h)=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char71.pngr2+(f−z)2+(h−z) http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

We need to extract the reflector height z as a function of radius r:

http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char71.pngr2+(f−z)2+h−z=f+h

r2+f2+z2−2fz=f2+z2+2fz

  The result is

z=r24f(3B1)

This is the equation of a paraboloid with ***focal length*** f.  The ratio of the focal length f to the diameter D of the reflector is called the ***fhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngD ratio*** or ***focal ratio***.  In principle it is a free parameter for the telescope designer.  In practice it is constrained.  If fhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngD is too high, the support structure needed to hold the feed or subreflector at the focus becomes unwieldy.  Thus most large radio telescopes have fhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngDhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png4, an unusually low focal ratio by optical standards.  The drawback of a low fhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngD is a small field of view.  The ***focal ellipsoid*** is the volume around the exact focal point that remains in reasonably good focus.  Only a small number (about seven) feeds can fit inside the focal ellipsoid of an fhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngDhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png4 paraboloid.  Large arrays of feeds or imaging cameras require larger fhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.pngD ratios, obtained either by flattening the paraboloid or by using magnifying subreflectors to increase the effective focal length.

Consequently, the primary mirrors of most radio telescopes are circular paraboloids or sections thereof. Their advantages are:

* The effective collecting area Ae of a reflector antenna can approach its projected geometric area A=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char19.pngD2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png4.
* Electrical simplicity (compared with a phased array of dipoles, for example).
* A single reflector works over a wide range of frequencies. Changing frequencies only requires changing the feed antenna and receiver located at the focal point, not building a whole new radio telescope.

**The Far-field Distance**

How far away must a point source be for the received waves to satisfy our assumption that they are nearly planar? The answer depends on both the wavelength http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pngand the reflector diameter D. Consider the spherical wave emitted by a point source a finite distance R from the reflector.

*The spherical wave emitted by a point source at distance R deviates from a plane by http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.pngat the edge of an aperture whose size is D.*

The maximum departure http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.pngfrom a plane wave occurs at the edge of the reflector. The ***far-field distance***Rff is somewhat arbitrarily defined by requiring that http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3C.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png16. Using the Pythagorean theorem, at the aperture edge we get

R2=(R−http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.png)2+http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char12.png2Dhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmex10/alpha/173/char13.png2 http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Thus

R=2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.png+D28http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

Since http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char1C.pngD, http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char1C.pngD2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png(8http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.png) and

Rhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.pngD28http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png

For http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmr10/alpha/173/char01.png=http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.pnghttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3D.png16,

Rffhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char19.png http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png2D2(3B2)

Unless Rhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3E.pngRff, the path-length errors will introduce significant phase errors in the waves coming from the off-axis portions of the reflector, reducing the effective collecting area and degrading the antenna pattern.

Example: What is the far-field distance of the Green Bank Telescope (D=100 m) observing at http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png=1 cm?

Rff=1cm2(100m)2=0http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3A.png01m2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png104m2=2http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmsy10/alpha/173/char02.png106m=2000km

Such a large far-field distance makes ground-based measurements of the GBT antenna pattern impractical. To measure the shape of the GBT reflector surface using radio holography, we can observe a geostationary satellite having an orbital altitude Rhttp://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char3E.png2000 km. Similarly, to determine the transmitting power pattern for a large radar antenna such as the D=305 m Arecibo reflector, we can passively observe a celestial point source in the far field and use the reciprocity theorem to equate the transmitting and receiving patterns.

**Patterns of Aperture Antennas**

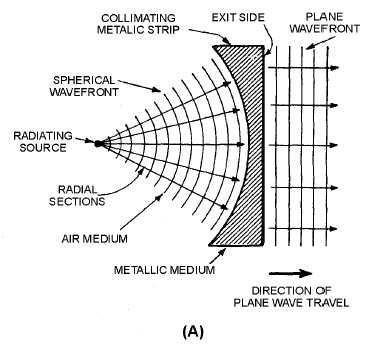
In optics, the term ***aperture*** refers to the opening through which all rays pass. For example, the aperture of a paraboloidal reflector antenna would be the plane circle, normal to the rays from a distant point source, that just covers the paraboloid. The phase of the plane wave from a distant point source would be constant across the aperture plane when the aperture is perpendicular to the line-of-sight.

*The aperture plane associated with a paraboloidal dish of diameter D.*

Another example of an aperture is the mouth of a waveguide horn antenna.

  
  
*"Doc" Ewen looking into the rectangular aperture of the horn antenna used to discover the http://www.cv.nrao.edu/course/astr534/jsMath/fonts/cmmi10/alpha/173/char15.png=21 cm line of  neutral hydrogen.*

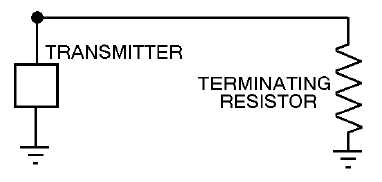
LENS ANTENNAS With a LENS ANTENNA you can convert spherically radiated microwave energy into a plane wave (in a given direction) by using a point source (open end of the waveguide) with a COLLIMATING LENS. A collimating lens forces all radial segments of the spherical wavefront into parallel paths. The point source can be regarded as a gun which shoots the microwave energy toward the lens. The point source is often a horn radiator or a simple dipole antenna. Waveguide Type The WAVEGUIDE-TYPE LENS is sometimes referred to as a conducting-type. It consists of several parallel concave metallic strips which are placed parallel to the electric field of the radiated energy fed to the lens, as shown in figure 3-10A and 3-10B. These strips act as waveguides in parallel for the incident (radiated) wave. The strips are placed slightly more than a half wavelength apart



UNIT IV

SPECIAL ANTENNA

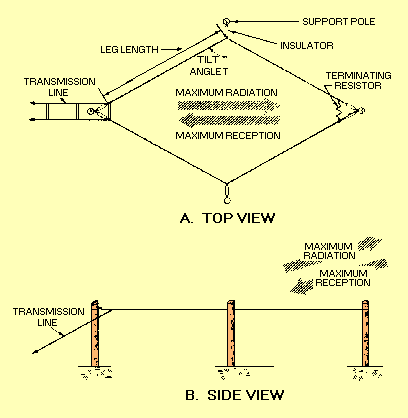
A LONG-WIRE ANTENNA is an antenna that is a wavelength or longer at the operating frequency. In general, the gain achieved with long-wire antennas is not as great as the gain obtained from the multielement arrays studied in the previous section. But the long-wire antenna has advantages of its own. The construction of long-wire antennas is simple, both electrically and mechanically, with no particularly critical dimensions or adjustments. The long-wire antenna will work well and give satisfactory gain and directivity over a frequency range up to twice the value for which it was cut. In addition, it will accept power and radiate it efficiently on any frequency for which its overall length is not less than approximately 1/2 wavelength. Another factor is that long-wire antennas have directional patterns that are sharp in both the horizontal and vertical planes. Also, they tend to concentrate the radiation at the low vertical angles. Another type of long-wire antenna is the BEVERAGE ANTENNA, also called a WAVE ANTENNA. It is a horizontal, long-wire antenna designed especially for the reception and transmission of low-frequency, vertically polarized ground waves. It consists of a single wire, two or more wavelengths long, supported 3 to 6 meters above the ground, and terminated in its characteristic impedance, as shown in figure 4-34. Figure 4-34.—Beverage antenna. Q44.   To radiate power efficiently, a long-wire antenna must have what minimum overall length? Q45.   What is another name for the Beverage antenna? V ANTENNA A V ANTENNA is a bi-directional antenna used widely in military and commercial communications. It consists of two conductors arranged to form a V. Each conductor is fed with currents of opposite polarity. The V is formed at such an angle that the main lobes reinforce along the line bisecting the V and make a very effective directional antenna (see figure 4-35). Connecting the two-wire feed line to the apex of the V and exciting the two sides of the V 180 degrees out of phase cause the lobes to add along the line of the bisector and to cancel in other directions, as shown in figure 4-36. The lobes are designated 1, 2, 3, and 4 on leg AA', and 5, 6, 7, and 8 on leg BB'. When the proper angle between AA' and BB' is chosen, lobes 1 and 4 have the same direction and combine with lobes 7 and 6, respectively. This combination of two major lobes from each leg results in the formation of two stronger lobes, which lie along an imaginary line bisecting the enclosed angle. Lobes 2, 3, 5, and 8 tend to cancel each other, as do the smaller lobes, which are approximately at right angles to the wire legs of the V. The resultant waveform pattern is shown at the right of the V antenna in figure 4-36



RHOMBIC ANTENNA

The highest development of the long-wire antenna is the RHOMBIC ANTENNA (see figure 4-37). It consists of four conductors joined to form a rhombus, or diamond shape. The antenna is placed end to end and terminated by a noninductive resistor to produce a uni-directional pattern. A rhombic antenna can be made of two obtuse-angle V antennas that are placed side by side, erected in a horizontal plane, and terminated so the antenna is nonresonant and unidirectional.

Figure 4-37. - Basic rhombic antenna.



The rhombic antenna is WIDELY used for long-distance, high-frequency transmission and reception. It is one of the most popular fixed-station antennas because it is very useful in point-to-point communications.

Advantages

The rhombic antenna is useful over a wide frequency range. Although some changes in gain, directivity, and characteristic impedance do occur with a change in operating frequency, these changes are small enough to be neglected.

The rhombic antenna is much easier to construct and maintain than other antennas of comparable gain and directivity. Only four supporting poles of common heights from 15 to 20 meters are needed for the antenna.

The rhombic antenna also has the advantage of being noncritical as far as operation and adjustment are concerned. This is because of the broad frequency characteristics of the antenna.

Still another advantage is that the voltages present on the antenna are much lower than those produced by the same input power on a resonant antenna. This is particularly important when high transmitter powers are used or when high-altitude operation is required.

YAGI-UDA ANTENNA

A Yagi-Uda array, commonly known simply as a Yagi antenna, is a [directional antenna](http://en.wikipedia.org/wiki/Directional_antenna) consisting of a driven element (typically a [dipole](http://en.wikipedia.org/wiki/Dipole_antenna) or [folded dipole](http://en.wikipedia.org/wiki/Folded_dipole)) and additional [parasitic elements](http://en.wikipedia.org/wiki/Parasitic_element) (usually a so-called *reflector* and one or more *directors*). The reflector element is slightly longer (typically 5% longer) than the driven dipole, whereas the so-called directors are a little bit shorter. This design achieves a very substantial increase in the antenna's [directionality](http://en.wikipedia.org/wiki/Directional_antenna) and [gain](http://en.wikipedia.org/wiki/Antenna_gain) compared to a simple dipole.[[1]](http://en.wikipedia.org/wiki/Yagi-Uda_antenna#cite_note-0)

Highly directional antennas such as the Yagi-Uda are commonly referred to as "beam antennas" due to their high gain. However the Yagi-Uda design only achieves this high gain over a rather narrow bandwidth, making it more useful for various communications bands (including amateur radio) but less suitable for traditional radio and television broadcast bands. [Amateur radio](http://en.wikipedia.org/wiki/Amateur_radio) operators ("hams") frequently employ these for communication on HF, [VHF](http://en.wikipedia.org/wiki/VHF), and [UHF](http://en.wikipedia.org/wiki/UHF) bands, often constructing such antennas themselves ("[homebrewing](http://en.wikipedia.org/wiki/Amateur_radio_homebrew)"), leading to a quantity of technical papers and software. Wideband antennas used for VHF/UHF broadcast bands include the lower-gain [log-periodic dipole array](http://en.wikipedia.org/wiki/Log-periodic_dipole_array), which is often confused with the Yagi-Uda array due to its superficially similar appearance. That design along with other [phased arrays](http://en.wikipedia.org/wiki/Phased_array) have electrical connections on each element, whereas the Yagi-Uda design operates on the basis of electromagnetic interaction between the "parasitic" elements and the one driven (dipole) element.



**ANTENNA MEASUREMENTS**

**Antenna measurement** techniques refers to the testing of [antennas](http://en.wikipedia.org/wiki/Antenna_%28radio%29) to ensure that the antenna meets specifications or simply to characterize it. Typical parameters of antennas are gain, radiation pattern, beamwidth, polarization, and impedance.

The [antenna pattern](http://en.wikipedia.org/wiki/Antenna_pattern) is the response of the antenna to a plane wave incident from a given direction or the relative power density of the wave transmitted by the antenna in a given direction. For a reciprocal antenna, these two patterns are identical. A multitude of antenna pattern measurement techniques have been developed. The first technique developed was the far-field range, where the antenna under test (AUT) is placed in the far-field of a range antenna. Due to the size required to create a far-field range for large antennas, near-field techniques were developed, which allow the measurement of the field on a surface close to the antenna (typically 3 to 10 times its [wavelength](http://en.wikipedia.org/wiki/Wavelength)). This measurement is then predicted to be the same at [infinity](http://en.wikipedia.org/wiki/Infinity). A third common method is the compact range, which uses a [reflector](http://en.wikipedia.org/wiki/Reflector_%28antenna%29) to create a field near the AUT that looks approximately like a plane-wave

**Far-field range (FF)**

The [far-field](http://en.wikipedia.org/wiki/Far-field) range was the original antenna measurement technique, and consists of placing the AUT a long distance away from the [instrumentation](http://en.wikipedia.org/wiki/Instrumentation) antenna. Generally, the far-field distance or [Fraunhofer distance](http://en.wikipedia.org/wiki/Fraunhofer_distance), *d*, is considered to be

d = {{2D^2}\over{\lambda}},

where D is the maximum dimension of the antenna and λ is the wavelength of the radio wave.[[1]](http://en.wikipedia.org/wiki/Antenna_measurement#cite_note-0) Separating the AUT and the instrumentation antenna by this distance reduces the phase variation across the AUT enough to obtain a reasonably good antenna pattern.

IEEE suggests the use of their antenna measurement standard, document number IEEE-Std-149-1979 for far-field ranges and measurement set-up for various techniques including ground-bounce type ranges.

**[**[**edit**](http://en.wikipedia.org/w/index.php?title=Antenna_measurement&action=edit&section=2)**] Near-field range (NF)**

Main article: [Electromagnetic near-field scanner](http://en.wikipedia.org/wiki/Electromagnetic_near-field_scanner)

**[**[**edit**](http://en.wikipedia.org/w/index.php?title=Antenna_measurement&action=edit&section=3)**] Planar near-field range**

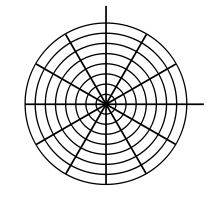
Planar [near-field](http://en.wikipedia.org/wiki/Near-field) measurements are conducted by scanning a small probe antenna over a planar surface. These measurements are then transformed to the far-field by use of a [Fourier transform](http://en.wikipedia.org/wiki/Fourier_transform), or more specifically by applying a method known as stationary phase[[2]](http://en.wikipedia.org/wiki/Antenna_measurement#cite_note-1) to the [Laplace transform](http://en.wikipedia.org/wiki/Laplace_transform) . Three basic types of planar scans exist in near field measurements.

**[**[**edit**](http://en.wikipedia.org/w/index.php?title=Antenna_measurement&action=edit&section=4)**] Rectangular planar scanning**

The probe moves in the [Cartesian coordinate system](http://en.wikipedia.org/wiki/Cartesian_coordinate_system) and its linear movement creates a regular rectangular sampling grid with a maximum near-field sample spacing of Δx = Δy = λ /2.

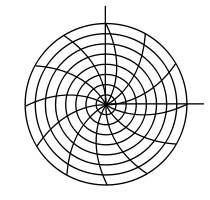
**[**[**edit**](http://en.wikipedia.org/w/index.php?title=Antenna_measurement&action=edit&section=5)**] Polar planar scanning**

More complicated solution to the rectangular scanning method is the plane polar scanning method.

[](http://en.wikipedia.org/wiki/File:PolarGrid.svg)

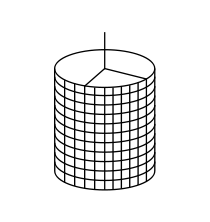
**[**[**edit**](http://en.wikipedia.org/w/index.php?title=Antenna_measurement&action=edit&section=6)**] Bi-polar planar scanning**

The bi-polar technique is very similar to the plane polar configuration.

[](http://en.wikipedia.org/wiki/File:BipolarGrid.svg)

**Cylindrical near-field range**

Cylindrical near-field ranges measure the electric field on a cylindrical surface close to the [AUT](http://en.wikipedia.org/wiki/AUT). Cylindrical [harmonics](http://en.wikipedia.org/wiki/Harmonics) are used transform these measurements to the far-field.

[](http://en.wikipedia.org/wiki/File:CylindricalGrid.svg)

**Spherical near-field range**

Spherical near-field ranges measure the electric field on a spherical surface close to the AUT. Spherical harmonics are used transform these measurements to the far-field

**Free-space ranges**

|  |  |
| --- | --- |
| [[icon]](http://en.wikipedia.org/wiki/File:Wiki_letter_w_cropped.svg) | This section requires [expansion](http://en.wikipedia.org/w/index.php?title=Antenna_measurement&action=edit). |

The formula for electromagnetic radiation dispersion and information is:

\displaystyle D^2=\frac{P}{S} \propto \!\, 3dB

Where D=Distance, P=Power, and S=Speed

What this means is that double the communication distance requires four times the power. It also means double power allows double communication speed (bit rate). Double power is approx. 3dB (10 log(.5) to be exact) increase. Of course in the real world there are all sorts of other phenomena which enter in, such as Fresnel canceling, path loss, background noise, etc.

**Compact range**

A Compact Antenna Test Range (CATR) is a facility which is used to provide convenient testing of antenna systems at frequencies where obtaining far-field spacing to the AUT would be infeasible using traditional [free space](http://en.wikipedia.org/wiki/Free_space) methods. The CATR uses a source antenna which radiates a spherical wavefront and one or more secondary reflectors to collimate the radiated spherical wavefront into a planar wavefront within the desired test zone. One typical embodiment uses a horn feed antenna and a [parabolic reflector](http://en.wikipedia.org/wiki/Parabolic_reflector) to accomplish this.

The CATR is used for [microwave](http://en.wikipedia.org/wiki/Microwave) and [millimeter wave](http://en.wikipedia.org/wiki/Millimeter_wave) frequencies where the 2 D2/λ far-field distance is large, such as with high-gain reflector antennas. The size of the range that is required can be much less than the size required for a full-size far-field anechoic chamber, although the cost of fabrication of the specially-designed CATR reflector can be expensive due to the need to ensure precision of the reflecting surface (typically less than λ/100 RMS surface accuracy) and to specially treat the edge of the reflector to avoid diffracted waves which can interfere with the desired beam pattern.

**Elevated range**

A means of reducing reflection from waves bouncing off the ground.

**Slant range**

A means of eliminating symmetrical wave reflection.

**Antenna parameters**

Except for polarization, the SWR is the most easily measured of the parameters above. Impedance can be measured with specialized equipment, as it relates to the [complex](http://en.wikipedia.org/wiki/Complex_number) SWR. Measuring radiation pattern requires a sophisticated setup including significant clear space (enough to put the sensor into the antenna's [far field](http://en.wikipedia.org/wiki/Far_field), or an anechoic chamber designed for antenna measurements), careful study of experiment geometry, and specialized measurement equipment that rotates the antenna during the measurements.

**Radiation pattern**

The [radiation pattern](http://en.wikipedia.org/wiki/Radiation_pattern) is a graphical depiction of the relative field strength transmitted from or received by the antenna, and shows [sidelobes](http://en.wikipedia.org/wiki/Sidelobe) and backlobes. As antennas radiate in space often several curves are necessary to describe the antenna. If the radiation of the antenna is symmetrical about an axis (as is the case in dipole, [helical](http://en.wikipedia.org/wiki/Helical_antenna) and some [parabolic](http://en.wikipedia.org/wiki/Parabolic_antenna) antennas) a unique graph is sufficient.

Each antenna supplier/user has different standards as well as plotting formats. Each format has its own advantages and disadvantages. Radiation pattern of an antenna can be defined as the locus of all points where the emitted power per unit surface is the same. The radiated power per unit surface is proportional to the squared electrical field of the electromagnetic wave. The radiation pattern is the locus of points with the same electrical field. In this representation, the reference is usually the best angle of emission. It is also possible to depict the directive gain of the antenna as a function of the direction. Often the gain is given in [decibels](http://en.wikipedia.org/wiki/Decibels).

The graphs can be drawn using [cartesian (rectangular) coordinates](http://en.wikipedia.org/wiki/Cartesian_coordinate_system) or a [polar plot](http://en.wikipedia.org/wiki/Polar_coordinate_system). This last one is useful to measure the beamwidth, which is, by convention, the angle at the -3dB points around the max gain. The shape of curves can be very different in cartesian or polar coordinates and with the choice of the limits of the logarithmic scale. The four drawings below are the radiation patterns of a same [half-wave antenna](http://en.wikipedia.org/wiki/Dipole_antenna#Half-wave_dipole_or_dipole_.28lambda_over_2.29).

|  |  |
| --- | --- |
| [http://upload.wikimedia.org/wikipedia/commons/f/fd/RadPatt-lin.png](http://en.wikipedia.org/wiki/File:RadPatt-lin.png)  [http://bits.wikimedia.org/skins-1.18/common/images/magnify-clip.png](http://en.wikipedia.org/wiki/File:RadPatt-lin.png)  Radiation pattern of a half-wave dipole antenna. Linear scale. | [http://upload.wikimedia.org/wikipedia/commons/thumb/e/e6/HWDipoleGain.svg/340px-HWDipoleGain.svg.png](http://en.wikipedia.org/wiki/File:HWDipoleGain.svg)  [http://bits.wikimedia.org/skins-1.18/common/images/magnify-clip.png](http://en.wikipedia.org/wiki/File:HWDipoleGain.svg)  Gain of a half-wave dipole. The scale is in dBi. |
| [http://upload.wikimedia.org/wikipedia/commons/f/f8/RadPatt-Cart.png](http://en.wikipedia.org/wiki/File:RadPatt-Cart.png)  [http://bits.wikimedia.org/skins-1.18/common/images/magnify-clip.png](http://en.wikipedia.org/wiki/File:RadPatt-Cart.png)  Gain of a half-wave dipole. Cartesian representation. |  |

**UNIT V**

**PROPAGATION OF RADIO WAVES**

**Radio propagation** is the behavior of [radio waves](http://en.wikipedia.org/wiki/Radio_wave) when they are [transmitted](http://en.wikipedia.org/wiki/Transmitted), or [propagated](http://en.wikipedia.org/wiki/Wave_propagation) from one point on the [Earth](http://en.wikipedia.org/wiki/Earth) to another, or into various parts of the [atmosphere](http://en.wikipedia.org/wiki/Atmosphere).[[1]](http://en.wikipedia.org/wiki/Radio_propagation#cite_note-0) As a form of [electromagnetic radiation](http://en.wikipedia.org/wiki/Electromagnetic_radiation), like light waves, radio waves are affected by the phenomena of [reflection](http://en.wikipedia.org/wiki/Reflection_%28physics%29), [refraction](http://en.wikipedia.org/wiki/Refraction), [diffraction](http://en.wikipedia.org/wiki/Diffraction), [absorption](http://en.wikipedia.org/wiki/Absorption_%28electromagnetic_radiation%29), [polarization](http://en.wikipedia.org/wiki/Polarization_%28waves%29) and [scattering](http://en.wikipedia.org/wiki/Scattering).[[2]](http://en.wikipedia.org/wiki/Radio_propagation#cite_note-1)

Radio propagation is affected by the daily changes of [water vapor](http://en.wikipedia.org/wiki/Water_vapor) in the [troposphere](http://en.wikipedia.org/wiki/Troposphere) and ionization in the [upper atmosphere](http://en.wikipedia.org/wiki/Upper_atmosphere), due to the [Sun](http://en.wikipedia.org/wiki/Sun). Understanding the effects of varying conditions on radio propagation has many practical applications, from choosing frequencies for international [shortwave](http://en.wikipedia.org/wiki/Shortwave) [broadcasters](http://en.wikipedia.org/wiki/Broadcasting), to designing reliable [mobile telephone](http://en.wikipedia.org/wiki/Mobile_phone) systems, to [radio navigation](http://en.wikipedia.org/wiki/Radio_navigation), to operation of [radar](http://en.wikipedia.org/wiki/Radar) systems.

Radio propagation is also affected by several other factors determined by its path from point to point. This path can be a direct [line of sight](http://en.wikipedia.org/wiki/Line-of-sight_propagation) path or an over-the-[horizon](http://en.wikipedia.org/wiki/Horizon) path aided by [refraction](http://en.wikipedia.org/wiki/Refraction) in the [ionosphere](http://en.wikipedia.org/wiki/Ionosphere), which is a region between approximately 60 and 600 km.[[3]](http://en.wikipedia.org/wiki/Radio_propagation#cite_note-2) Factors influencing ionospheric radio signal propagation can include [sporadic-E](http://en.wikipedia.org/wiki/Sporadic_E_propagation), spread-F, [solar flares](http://en.wikipedia.org/wiki/Solar_flare), [geomagnetic storms](http://en.wikipedia.org/wiki/Geomagnetic_storm), ionospheric layer tilts, and [solar proton events](http://en.wikipedia.org/wiki/Solar_proton_event).

Radio waves at different frequencies propagate in different ways. At extra low frequencies (ELF) and very low frequencies the wavelength is very much larger than the separation between the earth's surface and the D layer of the ionosphere, so electromagnetic waves may propagate in this region as a waveguide. Indeed, for frequencies below 20 kHz, the wave propagates as a single waveguide mode with a horizontal magnetic field and vertical electric field.[[4]](http://en.wikipedia.org/wiki/Radio_propagation#cite_note-3) The interaction of radio waves with the ionized regions of the atmosphere makes radio propagation more complex to predict and analyze than in free space. Ionospheric radio propagation has a strong connection to [space weather](http://en.wikipedia.org/wiki/Space_weather). A [sudden ionospheric disturbance](http://en.wikipedia.org/wiki/Sudden_ionospheric_disturbance) or shortwave fadeout is observed when the x-rays associated with a [solar flare](http://en.wikipedia.org/wiki/Solar_flares) ionize the ionospheric D-region.[[*citation needed*](http://en.wikipedia.org/wiki/Wikipedia:Citation_needed)] Enhanced ionization in that region increases the absorption of radio signals passing through it. During the strongest solar x-ray flares, complete absorption of virtually all ionospherically propagated radio signals in the sunlit hemisphere can occur.[[*citation needed*](http://en.wikipedia.org/wiki/Wikipedia:Citation_needed)] These solar flares can disrupt [HF radio](http://en.wikipedia.org/wiki/HF_radio) propagation and affect [GPS](http://en.wikipedia.org/wiki/Global_Positioning_System) accuracy.[[*citation needed*](http://en.wikipedia.org/wiki/Wikipedia:Citation_needed)]

Predictions of the average propagation conditions were needed and made during the [Second world war](http://en.wikipedia.org/wiki/Second_world_war). A most detailed code developed by [Karl Rawer](http://en.wikipedia.org/wiki/Karl_Rawer) was applied in the German [Wehrmacht](http://en.wikipedia.org/wiki/Wehrmacht), and after the war by the [French Navy](http://en.wikipedia.org/wiki/French_Navy).[[*citation needed*](http://en.wikipedia.org/wiki/Wikipedia:Citation_needed)]

Since radio propagation is not fully predictable, such services as [emergency locator transmitters](http://en.wikipedia.org/wiki/Distress_radiobeacon), in-flight communication with ocean-crossing aircraft, and some [television](http://en.wikipedia.org/wiki/Television) broadcasting have been moved to [communications satellites](http://en.wikipedia.org/wiki/Communications_satellite). A satellite link, though expensive, can offer highly predictable and stable line of sight coverage of a given area.

## Free space propagation

In [free space](http://en.wikipedia.org/wiki/Free_space), all [electromagnetic waves](http://en.wikipedia.org/wiki/Electromagnetic_wave) (radio, light, X-rays, etc.) obey the [inverse-square law](http://en.wikipedia.org/wiki/Inverse-square_law) which states that the power density of an electromagnetic wave is proportional to the inverse of the square of the distance from a [point source](http://en.wikipedia.org/wiki/Point_source) [[5]](http://en.wikipedia.org/wiki/Radio_propagation#cite_note-4) or:

\rho_P \propto \frac{1}{r^2}.

Doubling the distance from a transmitter means that the power density of the radiated wave at that new location is reduced to one-quarter of its previous value.

The power density per surface unit is proportional to the product of the electric and magnetic field strengths. Thus, doubling the propagation path distance from the transmitter reduces each of their received field strengths over a free-space path by one-half.

## [[edit](http://en.wikipedia.org/w/index.php?title=Radio_propagation&action=edit&section=2)] Modes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Radio frequencies and their primary mode of propagation** | | | | |
| **Band** | | **Frequency** | **Wavelength** | **Propagation via** |
| [ELF](http://en.wikipedia.org/wiki/Extremely_low_frequency) | Extremely Low Frequency | 3–300 [Hz](http://en.wikipedia.org/wiki/Hertz) | 1000-100,000 km |  |
| [VLF](http://en.wikipedia.org/wiki/Very_low_frequency) | Very Low Frequency | 3–30 [kHz](http://en.wikipedia.org/wiki/Kilohertz) | 100–10 km | Guided between the earth and the [ionosphere](http://en.wikipedia.org/wiki/Ionosphere). |
| [LF](http://en.wikipedia.org/wiki/Low_frequency) | Low Frequency | 30–300 [kHz](http://en.wikipedia.org/wiki/Kilohertz) | 10–1 km | Guided between the earth and the [D layer](http://en.wikipedia.org/wiki/D_layer) of the ionosphere.  [Surface waves](http://en.wikipedia.org/wiki/Surface_wave). |
| [MF](http://en.wikipedia.org/wiki/Medium_frequency) | Medium Frequency | 300–3000 [kHz](http://en.wikipedia.org/wiki/Kilohertz) | 1000–100 m | Surface waves.  E, [F layer](http://en.wikipedia.org/wiki/F_layer) ionospheric refraction at night, when D layer absorption weakens. |
| [HF](http://en.wikipedia.org/wiki/High_frequency) | High Frequency ([Short Wave](http://en.wikipedia.org/wiki/Shortwave)) | 3–30 [MHz](http://en.wikipedia.org/wiki/Megahertz) | 100–10 m | [E layer](http://en.wikipedia.org/wiki/E_layer) ionospheric refraction.  F1, [F2](http://en.wikipedia.org/wiki/F2_propagation) layer ionospheric refraction. |
| [VHF](http://en.wikipedia.org/wiki/Very_high_frequency) | Very High Frequency | 30–300 [MHz](http://en.wikipedia.org/wiki/Megahertz) | 10–1 m | [Infrequent E ionospheric refraction](http://en.wikipedia.org/wiki/Sporadic_E_propagation). Extremely rare F1, [F2](http://en.wikipedia.org/wiki/F2_propagation) layer ionospheric refraction during high sunspot activity up to 80 MHz. Generally direct wave. Sometimes [tropospheric ducting](http://en.wikipedia.org/wiki/Tropospheric_ducting). |
| [UHF](http://en.wikipedia.org/wiki/Ultra_high_frequency) | Ultra High Frequency | 300–3000 [MHz](http://en.wikipedia.org/wiki/Megahertz) | 100–10 cm | [Direct wave](http://en.wikipedia.org/wiki/Line-of-sight_propagation). Sometimes [tropospheric ducting](http://en.wikipedia.org/wiki/Tropospheric_ducting). |
| [SHF](http://en.wikipedia.org/wiki/Super_high_frequency) | Super High Frequency | 3–30 [GHz](http://en.wikipedia.org/wiki/Gigahertz) | 10–1 cm | Direct wave. |
| [EHF](http://en.wikipedia.org/wiki/Extremely_high_frequency) | Extremely High Frequency | 30–300 [GHz](http://en.wikipedia.org/wiki/Gigahertz) | 10–1 mm | Direct wave limited by absorption. |

### Surface modes

Main article: [Surface wave](http://en.wikipedia.org/wiki/Surface_wave)

Lower frequencies (between 30 and 3,000 kHz) have the property of following the curvature of the earth via [groundwave](http://en.wikipedia.org/wiki/Groundwave) propagation in the majority of occurrences.

In this mode the radio wave propagates by interacting with the semi-conductive surface of the earth. The wave "clings" to the surface and thus follows the curvature of the earth. Vertical [polarization](http://en.wikipedia.org/wiki/Polarization_%28waves%29) is used to alleviate short circuiting the electric field through the conductivity of the ground. Since the ground is not a perfect electrical conductor, ground waves are attenuated rapidly as they follow the earth’s surface. [Attenuation](http://en.wikipedia.org/wiki/Attenuation) is proportional to the frequency making this mode mainly useful for [LF](http://en.wikipedia.org/wiki/Low_frequency) and [VLF](http://en.wikipedia.org/wiki/VLF) frequencies.

Today LF and VLF are mostly used for [time signals](http://en.wikipedia.org/wiki/Time_signal), and for [military communications](http://en.wikipedia.org/wiki/Military_communications), especially with ships and submarines, although radio amateurs have an allocation at 137 kHz in some parts of the world. Early commercial and professional radio services relied exclusively on [long wave](http://en.wikipedia.org/wiki/Long_wave), low frequencies and ground-wave propagation. To prevent interference with these services, amateur and experimental transmitters were restricted to the higher (HF) frequencies, felt to be useless since their ground-wave range was limited. Upon discovery of the other propagation modes possible at [medium wave](http://en.wikipedia.org/wiki/Medium_wave) and [short wave](http://en.wikipedia.org/wiki/Short_wave) frequencies, the advantages of HF for commercial and military purposes became apparent. Amateur experimentation was then confined only to authorized frequency segments in that range. Although the exact frequencies vary by region, the most common HF wavelengths used are 160 m, 80 m, 40 m, 30 m, 20 m, 17 m, 15 m, 12 m and 10 m. The 30 m, 17 m and 12 m are [WARC bands](http://en.wikipedia.org/wiki/WARC_bands), which were made available later than the other bands.

### Direct modes (line-of-sight)

[Line-of-sight](http://en.wikipedia.org/wiki/Line-of-sight_propagation) is the direct propagation of radio waves between antennas that are visible to each other. This is probably the most common of the radio propagation modes at [VHF](http://en.wikipedia.org/wiki/VHF) and higher frequencies. Because radio signals can travel through many non-metallic objects, radio can be picked up through walls. This is still line-of-sight propagation. Examples would include propagation between a satellite and a ground antenna or reception of television signals from a local TV transmitter.

[Ground plane](http://en.wikipedia.org/wiki/Ground_plane) [reflection](http://en.wikipedia.org/wiki/Reflection_%28physics%29) effects are an important factor in VHF line of sight propagation. The interference between the direct beam line-of-sight and the ground reflected beam often leads to an effective inverse-fourth-power i.e. (1/lamda)^4 law for ground-plane limited radiation. [Need reference to inverse-fourth-power law + ground plane. Drawings may clarify]

### Ionospheric modes (skywave)

Main article: [Skywave](http://en.wikipedia.org/wiki/Skywave)

[Skywave](http://en.wikipedia.org/wiki/Skywave) propagation, also referred to as [skip](http://en.wikipedia.org/wiki/Skip_%28radio%29), is any of the modes that rely on [refraction](http://en.wikipedia.org/wiki/Refraction) of radio waves in the [ionosphere](http://en.wikipedia.org/wiki/Ionosphere), which is made up of one or more ionized layers in the upper [atmosphere](http://en.wikipedia.org/wiki/Earth%27s_atmosphere). F2-layer is the most important ionospheric layer for HF propagation, though F1, E, and D-layers also play some role. These layers are directly affected by the sun on a daily cycle, the seasons and the 11-year [sunspot cycle](http://en.wikipedia.org/wiki/Sunspot_cycle) determines the utility of these modes. During solar maxima, the whole HF range up to 30 MHz can be used and F2 propagation up to 50 MHz are observed frequently depending upon daily solar flux values. During [solar minima](http://en.wikipedia.org/wiki/Solar_minima), propagation of higher frequencies is generally worse.

Forecasting of skywave modes is of considerable interest to [amateur radio](http://en.wikipedia.org/wiki/Amateur_radio) operators and commercial [marine](http://en.wikipedia.org/wiki/Ocean) and [aircraft](http://en.wikipedia.org/wiki/Aircraft) communications, and also to [shortwave](http://en.wikipedia.org/wiki/Shortwave) broadcasters.

Meteor scattering

Meteor scattering relies on reflecting radio waves off the intensely ionized columns of air generated by [meteors](http://en.wikipedia.org/wiki/Meteor). While this mode is very short duration, often only from a fraction of second to couple of seconds per event, digital [Meteor burst communications](http://en.wikipedia.org/wiki/Meteor_burst_communications) allows remote stations to communicate to a station that may be hundreds of miles up to over 1,000 miles (1,600 km) away, without the expense required for a satellite link. This mode is most generally useful on VHF frequencies between 30 and 250 MHz.

#### Auroral backscatter

Intense columns of [Auroral](http://en.wikipedia.org/wiki/Aurora_%28astronomy%29) ionization at 100 km altitudes within the auroral oval [backscatter](http://en.wikipedia.org/wiki/Backscatter) radio waves, perhaps most notably on HF and VHF. Backscatter is angle-sensitive—incident ray vs. magnetic field line of the column must be very close to right-angle. Random motions of electrons spiraling around the field lines create a Doppler-spread that broadens the spectra of the emission to more or less noise-like—depending on how high radio frequency is used. The radio-auroras are observed mostly at high latitudes and rarely extend down to middle latitudes. The occurrence of radio-auroras depends on solar activity ([flares](http://en.wikipedia.org/wiki/Solar_flare), [coronal holes](http://en.wikipedia.org/wiki/Coronal_holes), [CMEs](http://en.wikipedia.org/wiki/Coronal_mass_ejection)) and annually the events are more numerous during solar cycle maximas. Radio aurora includes the so-called afternoon radio aurora which produces stronger but more distorted signals and after the Harang-minima, the late-night radio aurora (sub-storming phase) returns with variable signal strength and lesser doppler spread. The propagation range for this predominantly back-scatter mode extends up to about 2000 km in east-west plane, but strongest signals are observed most frequently from the north at nearby sites on same latitudes.

Rarely, a strong radio-aurora is followed by Auroral-E, which resembles both propagation types in some ways.

#### Sporadic-E propagation

[Sporadic E](http://en.wikipedia.org/wiki/Sporadic_E) (Es) propagation can be observed on HF and VHF bands. It must not be confused with ordinary HF E-layer propagation. Sporadic-E at mid-latitudes occurs mostly during summer season, from May to August in the northern hemisphere and from November to February in the southern hemisphere. There is no single cause for this mysterious propagation mode. The reflection takes place in a thin sheet of ionisation around 90 km height. The ionisation patches drift westwards at speeds of few hundred km per hour. There is a weak periodicity noted during the season and typically Es is observed on 1 to 3 successive days and remains absent for a few days to reoccur again. Es do not occur during small hours; the events usually begin at dawn, and there is a peak in the afternoon and a second peak in the evening.[[6]](http://en.wikipedia.org/wiki/Radio_propagation#cite_note-5) Es propagation is usually gone by local midnight.

Maximum observed frequency (MOF) for Es is found to be lurking around 30 MHz on most days during the summer season, but sometimes MOF may shoot up to 100 MHz or even more in ten minutes to decline slowly during the next few hours. The peak-phase includes oscillation of MOF with periodicity of approximately 5...10 minutes. The propagation range for Es single-hop is typically 1000 to 2000 km, but with multi-hop, double range is observed. The signals are very strong but also with slow deep fading.

Thomas F. Giella, retired [meteorologist](http://en.wikipedia.org/wiki/Meteorologist), space plasma physicist and an Amateur Radio Operator (NZ4O), cites the following from his professional research:[[*citation needed*](http://en.wikipedia.org/wiki/Wikipedia:Citation_needed)][[*original research?*](http://en.wikipedia.org/wiki/Wikipedia:No_original_research)]

Just as the E layer is the main refraction medium for medium frequency (300–3000 kHz) signal propagation within approximately 5000 km (3000 mi), so is a Sporadic-E (Es) cloud. Sporadic-E (Es) clouds occur at approximately 100 km (60 mi) in altitude and generally move from ESE to WNW. Like [Stratosphere](http://en.wikipedia.org/wiki/Stratosphere) level warming and [Troposphere](http://en.wikipedia.org/wiki/Troposphere) level temperature and moisture discontinuities, Sporadic-E (Es) clouds can depending on the circumstances absorb, block or refract medium, high and very high frequency [RF](http://en.wikipedia.org/wiki/Radio_frequency) signals in an unpredictable manner.

The main source for "high latitude" Sporadic E (Es) clouds is [geomagnetic storming](http://en.wikipedia.org/wiki/Geomagnetic_storm) induced radio aurora activity.

The main source for "mid latitude" Sporadic-E (Es) clouds is [wind shear](http://en.wikipedia.org/wiki/Wind_shear) produced by internal buoyancy/gravity waves (IBGW's), that create traveling ionosphere disturbances (TID's), most of which are produced by severe [thunderstorm](http://en.wikipedia.org/wiki/Thunderstorm) cell complexes with overshooting tops that penetrate into the Stratosphere. Another tie in between Sporadic-E (Es) and a severe thunderstorm is the [Elve](http://en.wikipedia.org/wiki/Lightning#Sprites.2C_Elves.2C_Jets_and_other_Upper_Atmospheric_Lightning).

The main sources for "low latitude" Sporadic-E (Es) clouds is wind shear produced by internal buoyancy/gravity waves (IBGW's), that create traveling ionosphere disturbances, most of which are produced by severe thunderstorm cell complexes tied to tropical cyclones. High electron content in the Equatorial Ring Current also plays a role.

The forecasting of Sporadic-E (Es) clouds has long been considered to be impossible. However it is possible to identify certain troposphere level meteorological conditions that can lead to the formation of Sporadic E (Es) clouds. One is as mentioned above the severe thunderstorm cell complex.

Sporadic-E (Es) clouds have been observed to initially occur within approximately 150 km (90 mi) to the right of a severe thunderstorm cell complex in the northern hemisphere, with the opposite being observed in the southern hemisphere. To complicate matters is the fact that Sporadic-E (Es) clouds that initially form to the right of a severe thunderstorm complex in the northern hemisphere, then move from ESE-WNW and end up to the left of the severe thunderstorm complex in the northern hemisphere. So one has to look for Sporadic-E (Es) clouds on either side of a severe thunderstorm cell complex. Things get even more complicated when two severe thunderstorm cell complexes exist approximately 1000–2000 miles apart.

Not all thunderstorm cell complexes reach severe levels and not all severe thunderstorm cell complexes produce Sporadic-E (Es). This is where knowledge in tropospheric physics and weather analyses/forecasting is necessary.

Some of the key elements in identifying which severe thunderstorm cell complexes have the potential to produce Sporadic-E (Es) via wind shear, from internal buoyancy/gravity waves, that produce traveling ionosphere disturbances include:

1.) Negative tilted mid and upper level long wave troughs.

2.) Approximate 150 knot (170 mph, 280 km/h) [jet stream](http://en.wikipedia.org/wiki/Jet_stream) jet maxes that produce divergence and therefore create a sucking vacuum effect above thunderstorm cells, that assist thunderstorm cells in reaching and penetrating the [tropopause](http://en.wikipedia.org/wiki/Tropopause) into the stratosphere.

3.) 500 mb (50 kPa) temperatures of −20 °C or colder, which produce numerous positive and negative lightning bolts and inter-related [Sprites](http://en.wikipedia.org/wiki/Sprite_%28lightning%29) and Elves.

4.) Approximate 150–175 knot (170–200 mph) updrafts within thunderstorm cells complexes that create overshooting tops that penetrate the Tropopause into the Stratosphere (See definition #20 on Stratospheric Warming), launching upwardly propagating internal buoyancy/gravity waves, which create traveling ionosphere disturbances and then wind shear.

### Tropospheric modes

#### Tropospheric scattering

At [VHF](http://en.wikipedia.org/wiki/VHF) and higher frequencies, small variations (turbulence) in the density of the [atmosphere](http://en.wikipedia.org/wiki/Troposphere) at a height of around 6 miles (10 km) can scatter some of the normally line-of-sight beam of radio frequency energy back toward the ground, allowing over-the-horizon communication between stations as far as 500 miles (800 km) apart. The military developed the [White Alice Communications System](http://en.wikipedia.org/wiki/White_Alice_Communications_System) covering all of Alaska, using this [tropospheric scattering](http://en.wikipedia.org/wiki/Tropospheric_scatter) principle.

#### ] Tropospheric ducting

Main article: [Tropospheric ducting](http://en.wikipedia.org/wiki/Tropospheric_ducting)

Sudden changes in the atmosphere's vertical moisture content and temperature profiles can on random occasions make [microwave](http://en.wikipedia.org/wiki/Microwave) and [UHF](http://en.wikipedia.org/wiki/UHF) & [VHF](http://en.wikipedia.org/wiki/VHF) signals propagate hundreds of kilometers up to about 2,000 kilometers (1,300 mi)—and for ducting mode even farther—beyond the normal radio-horizon. The [inversion layer](http://en.wikipedia.org/wiki/Inversion_%28meteorology%29) is mostly observed over high pressure regions, but there are several tropospheric weather conditions which create these randomly occurring propagation modes. Inversion layer's altitude for non-ducting is typically found between 100 meters (300 ft) to about 1 kilometer (3,000 ft) and for ducting about 500 meters to 3 kilometers (1,600 to 10,000 ft), and the duration of the events are typically from several hours up to several days. Higher frequencies experience the most dramatic increase of signal strengths, while on low-VHF and HF the effect is negligible. Propagation path attenuation may be below free-space loss. Some of the lesser inversion types related to warm ground and cooler air moisture content occur regularly at certain times of the year and time of day. A typical example could be the late summer, early morning tropospheric enhancements that bring in signals from distances up to few hundred kilometers for a couple of hours, until undone by the Sun's warming effect.

#### Tropospheric delay

This is a source of error in radio ranging techniques, such as the [Global Positioning System](http://en.wikipedia.org/wiki/GPS#Atmospheric_effects) (GPS). See also the page of [GPS meteorology](http://en.wikipedia.org/wiki/GPS_meteorology).

#### Rain scattering

Rain scattering is purely a microwave propagation mode and is best observed around 10 GHz, but extends down to a few [gigahertz](http://en.wikipedia.org/wiki/Gigahertz)—the limit being the size of the scattering particle size vs. [wavelength](http://en.wikipedia.org/wiki/Wavelength). This mode scatters signals mostly forwards and backwards when using [horizontal polarization](http://en.wikipedia.org/wiki/Horizontal_polarization) and side-scattering with [vertical polarization](http://en.wikipedia.org/wiki/Vertical_polarization). Forward-scattering typically yields propagation ranges of 800 km. Scattering from snowflakes and ice pellets also occurs, but scattering from ice without watery surface is less effective. The most common application for this phenomenon is microwave rain radar, but rain scatter propagation can be a nuisance causing unwanted signals to intermittently propagate where they are not anticipated or desired. Similar reflections may also occur from insects though at lower altitudes and shorter range. Rain also causes attenuation of point-to-point and satellite microwave links. Attenuation values up to 30 dB have been observed on 30 GHz during heavy tropical rain.

#### Airplane scattering

Airplane scattering (or most often reflection) is observed on VHF through microwaves and, besides back-scattering, yields momentary propagation up to 500 km even in mountainous terrain. The most common back-scatter applications are air-traffic radar, bistatic forward-scatter guided-missile and airplane-detecting trip-wire radar, and the US space radar.

#### Lightning scattering

Lightning scattering has sometimes been observed on VHF and UHF over distance of about 500 km. The hot lightning channel scatters radiowaves for a fraction of a second. The RF noise burst from the lightning makes the initial part of the open channel unusable and the ionisation disappears soon because of combination at low altitude high atmospheric pressure. Although the hot lightning channel is briefly observable with microwave radar, this mode has no practical use for communications.

### Other effects

#### Diffraction

[Knife-Edge diffraction](http://en.wikipedia.org/wiki/Knife-Edge_diffraction) is the propagation mode where radio waves are bent around sharp edges. For example, this mode is used to send radio signals over a mountain range when a [line-of-sight](http://en.wikipedia.org/wiki/Line-of-sight_propagation) path is not available. However, the angle cannot be too sharp or the signal will not diffract. The diffraction mode requires increased signal strength, so higher power or better antennas will be needed than for an equivalent line-of-sight path.

Diffraction depends on the relationship between the wavelength and the size of the obstacle. In other words, the size of the obstacle in wavelengths. Lower frequencies diffract around large smooth obstacles such as hills more easily. For example, in many cases where VHF (or higher frequency) communication is not possible due to shadowing by a hill, one finds that it is still possible to communicate using the upper part of the HF band where the surface wave is of little use.

Diffraction phenomena by small obstacles are also important at high frequencies. Signals for urban [cellular telephony](http://en.wikipedia.org/wiki/Cellular_telephony) tend to be dominated by ground-plane effects as they travel over the rooftops of the urban environment. They then diffract over roof edges into the street, where [multipath propagation](http://en.wikipedia.org/wiki/Multipath_propagation), absorption and diffraction phenomena dominate.

#### Absorption

Low-frequency radio waves travel easily through brick and stone and VLF even penetrates sea-water. As the frequency rises, absorption effects become more important. At [microwave](http://en.wikipedia.org/wiki/Microwave) or higher frequencies, absorption by molecular resonance in the atmosphere (mostly water, H2O and oxygen, O2) is a major factor in radio propagation. For example, in the 58–60 GHz band, there is a major absorption peak which makes this band useless for long-distance use. This phenomenon was first discovered during [radar](http://en.wikipedia.org/wiki/Radar) research in [World War II](http://en.wikipedia.org/wiki/World_War_II). Beyond around 400 GHz, the Earth's atmosphere blocks some segments of spectra while still passes some—this is true up to UV light, which is blocked by ozone, but visible light and some of the near-infrared is transmitted.

Heavy rain and snow also affect microwave reception.

**Tropospheric propagation** describes [electromagnetic propagation](http://en.wikipedia.org/wiki/Electromagnetic_propagation) in relation to the [troposphere](http://en.wikipedia.org/wiki/Troposphere).

The service area from a television (TV) or [frequency modulated](http://en.wikipedia.org/wiki/Frequency_modulated) (FM) radio transmitter extends to just beyond the [optical horizon](http://en.wikipedia.org/wiki/Horizon), at which point signals start to rapidly reduce in strength. Viewers living in such a "deep fringe" reception area will notice that during certain conditions, weak signals normally masked by noise increase in signal strength to allow quality reception. Such conditions are related to the current state of the [troposphere](http://en.wikipedia.org/wiki/Troposphere).

Tropospheric propagated signals travel in the part of the [atmosphere](http://en.wikipedia.org/wiki/Earth%27s_atmosphere) adjacent to the surface and extending to some 25,000 feet (7,620 m). Such signals are thus directly affected by [weather](http://en.wikipedia.org/wiki/Weather) conditions extending over some hundreds of miles. During very settled, warm [anticyclonic](http://en.wikipedia.org/wiki/Cyclone) weather (i.e., high [pressure](http://en.wikipedia.org/wiki/Atmospheric_pressure)), usually weak signals from distant transmitters improve in strength. Another symptom during such conditions may be [interference](http://en.wikipedia.org/wiki/Interference_%28communication%29) to the local transmitter resulting in [co-channel interference](http://en.wikipedia.org/wiki/Co-channel_interference), usually horizontal lines or an extra [floating picture](http://en.wikipedia.org/wiki/Ghosting_%28television%29) with analog broadcasts and break-up with digital broadcasts. A settled [high-pressure system](http://en.wikipedia.org/wiki/High-pressure_system) gives the characteristic conditions for enhanced tropospheric propagation, in particular favouring signals which travel along the prevailing [isobar](http://en.wikipedia.org/wiki/Isobar_%28meteorology%29) pattern (rather than across it). Such weather conditions can occur at any time, but generally the summer and autumn months are the best periods. In certain favourable locations, enhanced tropospheric propagation may enable reception of [ultra high frequency](http://en.wikipedia.org/wiki/Ultra_high_frequency) (UHF) TV signals up to 1,000 miles (1,600 km) or more.

The observable characteristics of such high-pressure systems are usually clear, cloudless days with little or no wind. At sunset the upper air cools, as does the surface temperature, but at different rates. This produces a boundary or [temperature gradient](http://en.wikipedia.org/wiki/Temperature_gradient), which allows an [inversion](http://en.wikipedia.org/wiki/Temperature_inversion) level to form – a similar effect occurs at sunrise. The inversion is capable of allowing [very high frequency](http://en.wikipedia.org/wiki/Very_high_frequency) (VHF) and UHF signal propagation well beyond the normal [radio horizon](http://en.wikipedia.org/wiki/Radio_horizon) distance.

The inversion effectively reduces [sky wave](http://en.wikipedia.org/wiki/Sky_wave) [radiation](http://en.wikipedia.org/wiki/Electromagnetic_radiation) from a transmitter – normally VHF and UHF signals travel on into space when they reach the horizon, the [refractive index](http://en.wikipedia.org/wiki/Refractive_index) of the ionosphere preventing signal return. With temperature inversion, however, the signal is to a large extent [refracted](http://en.wikipedia.org/wiki/Refraction) over the horizon rather than continuing along a direct path into [outer space](http://en.wikipedia.org/wiki/Outer_space).

[Fog](http://en.wikipedia.org/wiki/Fog) also produces good tropospheric results, again due to inversion effects. Fog occurs during high-pressure weather, and if such conditions result in a large belt of fog with clear sky above, there will be heating of the upper fog level and thus an inversion. This situation often arises towards night fall, continues overnight and clears with the sunrise over a period of around 4 – 5 hours.

|  |
| --- |
|  |

**Tropospheric scatter** (known as "troposcatter" among practitioners) is a method of transmitting and receiving [microwave radio](http://en.wikipedia.org/wiki/Microwave_radio) signals over considerable distances – often up to 300 km. This method of propagation uses the tropospheric scatter phenomenon, where radio waves at particular frequencies are randomly scattered as they pass through the upper layers of the [troposphere](http://en.wikipedia.org/wiki/Troposphere) (hence troposcatter). Radio signals are transmitted in a tight beam aimed at the [tropopause](http://en.wikipedia.org/wiki/Tropopause), midway between the transmitter and receiver sites; as the signals pass through the troposphere they are scattered, allowing the receiver station to pick up the signal.[[1]](http://en.wikipedia.org/wiki/Tropospheric_scatter#cite_note-0)

Normally, microwave signals, transmitted at various frequencies, usually around 12 [Gigahertz](http://en.wikipedia.org/wiki/Gigahertz) (GHz) or 19 GHz, are only used for ‘[line of sight](http://en.wikipedia.org/wiki/Line-of-sight_propagation)’ applications, where the receiver can be ‘seen’ from the transmitter. However, tropospheric scatter signals use a frequency of around 2 GHz.

Because the troposphere is turbulent and has a high proportion of moisture the tropospheric scatter radio signals are [refracted](http://en.wikipedia.org/wiki/Refracted) and consequently only a proportion of the radio energy is collected by the receiving antennae. Frequencies of transmission around 2 GHz are best suited for tropospheric scatter systems as at this frequency the wavelength of the signal interacts well with the moist, turbulent areas of the troposphere, improving [signal to noise ratios](http://en.wikipedia.org/wiki/Signal_to_noise_ratio).

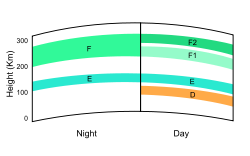
[High gain](http://en.wikipedia.org/wiki/High-gain_antenna) [dish](http://en.wikipedia.org/wiki/Dish_antenna) or [billboard antennae](http://en.wikipedia.org/wiki/Billboard_antenna) are required for tropospheric scatter systems as the [propagation losses](http://en.wikipedia.org/w/index.php?title=Propagation_loss&action=edit&redlink=1) are very high; only about one billion-billionth (1 x 10−12) of the transmit power is available at the receiver. Typically, dish antennae with [isotropic gains](http://en.wikipedia.org/wiki/Antenna_gain) of between 40 [decibels](http://en.wikipedia.org/wiki/Decibels)(dB) and 60dB are used with transmitter powers of 1 [Kilowatt](http://en.wikipedia.org/wiki/Kilowatt)(kW) to 10 kW.

Tropospheric scatter is a fairly secure method of propagation as dish alignment is critical, making it extremely difficult to intercept the signals, especially if transmitted across open water, making them highly attractive to military users. Military systems have tended to be ‘thin-line’ tropo – so called because only a narrow [bandwidth](http://en.wikipedia.org/wiki/Bandwidth_%28signal_processing%29) ‘information’ channel was carried on the tropo system; generally up to 32 analogue (4 kHz bandwidth) channels.

Civilian troposcatter systems, such as the [British Telecom](http://en.wikipedia.org/wiki/British_Telecom)(BT) [North Sea](http://en.wikipedia.org/wiki/North_Sea) oil communications network required higher capacity ‘information’ channels than were available using hf (high frequency – 3 to 30 MHz) radio signals, before satellite technology was available. The BT systems, based at Scousburgh in the Shetland Islands, Mormond Hill in Aberdeenshire and Row Brow near Scarborough, were capable of transmitting and receiving 156 analogue (4 kHz bandwidth) channels of data and telephony to / from North Sea oil production platforms, using frequency division multiplexing (FDMX) to combine the channels.

Because of the nature of the turbulence in the troposphere, quadruple [diversity propagation](http://en.wikipedia.org/wiki/Antenna_diversity) paths were used to ensure 99.98% reliability of the service, equating to about 3 minutes of downtime due to propagation drop out per month. The quadruple space and polarisation diversity systems needed two separate dish antenna (spaced several metres apart) and two differently [polarised](http://en.wikipedia.org/wiki/Polarization_%28waves%29) [feed horns](http://en.wikipedia.org/wiki/Feed_horn) – one using vertical polarisation, the other using horizontal polarisation. This ensured that at least one signal path was open at any one time. The signals from the four different paths were recombined in the receiver where a phase corrector removed the [phase differences](http://en.wikipedia.org/wiki/Phase_difference) of each signal. Phase differences were caused by the different path lengths of each signal from transmitter to receiver. Once phase corrected, the four signals could be combined additively.

## The ionospheric layers

[](http://en.wikipedia.org/wiki/File:Ionosphere_Layers_en.svg)

[http://bits.wikimedia.org/skins-1.18/common/images/magnify-clip.png](http://en.wikipedia.org/wiki/File:Ionosphere_Layers_en.svg)

Ionospheric layers.

At night the F layer is the only layer of significant ionization present, while the ionization in the E and D layers is extremely low. During the day, the D and E layers become much more heavily ionized, as does the F layer, which develops an additional, weaker region of ionisation known as the F1 layer. The F2 layer persists by day and night and is the region mainly responsible for the refraction of radio waves.

### [[edit](http://en.wikipedia.org/w/index.php?title=Ionosphere&action=edit&section=3)] D layer

The [D layer](http://en.wikipedia.org/wiki/D_region) is the innermost layer, 60 km to 90 km above the surface of the Earth. Ionization here is due to [Lyman series](http://en.wikipedia.org/wiki/Lyman_series)-alpha hydrogen radiation at a [wavelength](http://en.wikipedia.org/wiki/Wavelength) of 121.5 [nanometre](http://en.wikipedia.org/wiki/Nanometre) (nm) ionizing [nitric oxide](http://en.wikipedia.org/wiki/Nitric_oxide) (NO). In addition, with high [Solar activity](http://en.wikipedia.org/wiki/Solar_activity) hard [X-rays](http://en.wikipedia.org/wiki/X-ray) (wavelength < 1 nm) may ionize (N2, O2). During the night [cosmic rays](http://en.wikipedia.org/wiki/Cosmic_rays) produce a residual amount of ionization. Recombination is high in the D layer, the net ionization effect is low, but loss of wave energy is great due to frequent collisions of the electrons (about ten collisions every msec). As a result high-frequency (HF) [radio waves](http://en.wikipedia.org/wiki/Radio_wave) are not reflected by the D layer but suffer loss of energy therein. This is the main reason for [absorption of HF radio waves](http://en.wikipedia.org/wiki/Ionospheric_absorption), particularly at 10 MHz and below, with progressively smaller absorption as the frequency gets higher. The absorption is small at night and greatest about midday. The layer reduces greatly after sunset; a small part remains due to [galactic cosmic rays](http://en.wikipedia.org/wiki/Galactic_cosmic_ray). A common example of the D layer in action is the disappearance of distant AM [broadcast band](http://en.wikipedia.org/wiki/Broadcast_band) stations in the daytime.

During [solar proton events](http://en.wikipedia.org/wiki/Solar_proton_event), ionization can reach unusually high levels in the D-region over high and polar latitudes. Such very rare events are known as Polar Cap Absorption (or PCA) events, because the increased ionization significantly enhances the absorption of radio signals passing through the region. In fact, absorption levels can increase by many tens of dB during intense events, which is enough to absorb most (if not all) transpolar HF radio signal transmissions. Such events typically last less than 24 to 48 hours.

### E layer

The [E layer](http://en.wikipedia.org/wiki/Kennelly-Heaviside_Layer) is the middle layer, 90 km to 120 km above the surface of the Earth. Ionization is due to soft X-ray (1-10 nm) and far ultraviolet (UV) solar radiation ionization of molecular [oxygen](http://en.wikipedia.org/wiki/Oxygen) (O2). Normally, at oblique incidence, this layer can only reflect radio waves having frequencies lower than about 10 MHz and may contribute a bit to absorption on frequencies above. However, during intense [Sporadic E](http://en.wikipedia.org/wiki/Sporadic_E) events, the Es layer can reflect frequencies up to 50 MHz and higher. The vertical structure of the E layer is primarily determined by the competing effects of ionization and recombination. At night the E layer rapidly disappears because the primary source of ionization is no longer present. After sunset an increase in the height of the E layer maximum increases the range to which radio waves can travel by reflection from the layer.

This region is also known as the [Kennelly-Heaviside Layer](http://en.wikipedia.org/wiki/Kennelly-Heaviside_Layer) or simply the Heaviside layer. Its existence was predicted in 1902 independently and almost simultaneously by the American electrical engineer [Arthur Edwin Kennelly](http://en.wikipedia.org/wiki/Arthur_Edwin_Kennelly) (1861–1939) and the British physicist [Oliver Heaviside](http://en.wikipedia.org/wiki/Oliver_Heaviside) (1850–1925). However, it was not until 1924 that its existence was detected by [Edward V. Appleton](http://en.wikipedia.org/wiki/Edward_V._Appleton).

### Es

The Es layer ([sporadic](http://en.wiktionary.org/wiki/sporadic#Adjective) E-layer) is characterized by small, thin clouds of intense ionization, which can support reflection of radio waves, rarely up to 225 MHz. Sporadic-E events may last for just a few minutes to several hours. [Sporadic E propagation](http://en.wikipedia.org/wiki/Sporadic_E_propagation) makes [radio amateurs](http://en.wikipedia.org/wiki/Amateur_radio_high_bands) very excited, as propagation paths that are generally unreachable can open up. There are multiple causes of sporadic-E that are still being pursued by researchers. This propagation occurs most frequently during the summer months when high signal levels may be reached. The skip distances are generally around 1,000 km (620 mi). Distances for one hop propagation can be as close as 900 km [500 miles] or up to 2,500 km (1,600 mi). Double-hop reception over 3,500 km (2,200 mi) is possible.

### F layer

The [F layer](http://en.wikipedia.org/wiki/F_region) or region, also known as the [Appleton](http://en.wikipedia.org/wiki/Edward_Victor_Appleton) layer extends from about 200 km to more than 500 km above the surface of Earth. It is the densest point of the ionosphere, which implies signals penetrating this layer will escape into space. Beyond this layer is the [topside ionosphere](http://en.wikipedia.org/w/index.php?title=Topside_ionosphere&action=edit&redlink=1). Here extreme ultraviolet (UV, 10–100 nm) solar radiation ionizes atomic [oxygen](http://en.wikipedia.org/wiki/Oxygen). The F layer consists of one layer at night, but during the day, a deformation often forms in the profile that is labeled F1. The F2 layer remains by day and night responsible for most [skywave](http://en.wikipedia.org/wiki/Skywave) propagation of [radio](http://en.wikipedia.org/wiki/Radio) waves, facilitating high frequency (HF, or [shortwave](http://en.wikipedia.org/wiki/Shortwave)) radio communications over long distances.

From 1972 to 1975 [NASA](http://en.wikipedia.org/wiki/NASA) launched the [AEROS and AEROS B](http://en.wikipedia.org/wiki/AEROS_%28satellite%29) satellites to study the F region.[[2]](http://en.wikipedia.org/wiki/Ionosphere#cite_note-Yenne-1)

### Ionospheric model

An ionospheric model is a mathematical description of the ionosphere as a function of location, altitude, day of year, phase of the sunspot cycle and geomagnetic activity. Geophysically, the state of the ionospheric [plasma](http://en.wikipedia.org/wiki/Plasma_%28physics%29) may be described by four parameters: *electron density, electron and ion* [*temperature*](http://en.wikipedia.org/wiki/Temperature) and, since several species of ions are present, *ionic composition*. [Radio propagation](http://en.wikipedia.org/wiki/Radio_propagation) depends uniquely on electron density.

Models are usually expressed as computer programs. The model may be based on basic physics of the interactions of the ions and electrons with the neutral atmosphere and sunlight, or it may be a statistical description based on a large number of observations or a combination of physics and observations. One of the most widely used models is the [International Reference Ionosphere](http://en.wikipedia.org/wiki/International_Reference_Ionosphere) (IRI)[[3]](http://en.wikipedia.org/wiki/Ionosphere#cite_note-2) [(IRI 2007)](http://en.wikipedia.org/wiki/Ionosphere#References), which is based on data and specifies the four parameters just mentioned. The IRI is an international project sponsored by the [Committee on Space Research](http://en.wikipedia.org/wiki/Committee_on_Space_Research) (COSPAR) and the [International Union of Radio Science](http://en.wikipedia.org/wiki/International_Union_of_Radio_Science) (URSI).[[4]](http://en.wikipedia.org/wiki/Ionosphere#cite_note-3) The major data sources are the worldwide network of ionosondes, the powerful incoherent scatter radars (Jicamarca, Arecibo, Millstone Hill, Malvern, St. Santin), the ISIS and Alouette topside sounders, and in situ instruments on several satellites and rockets. IRI is updated yearly. IRI will be established in 2009 by the International Organization for Standardization (ISO) as standard TS16457. IRI is accurate in describing the variation of the electron density from bottom of the ionosphere to the altitude of maximum density than in describing the total electron content (TEC).Since 1999 the IRI is "International Standard" for the terrestrial ionosphere.

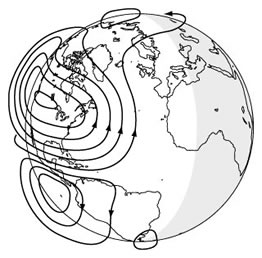
## Anomalies to the ideal model

Ionograms allow deducing not only the shape of the different layers but also the structure of the [electron](http://en.wikipedia.org/wiki/Electron)/[ion](http://en.wikipedia.org/wiki/Ion)-[plasma](http://en.wikipedia.org/wiki/Plasma_%28physics%29). Rough traces, indicating nonhomogeneity, are seen predominantly at night and at higher latitudes, and during disturbed conditions.

### Winter anomaly

At mid-latitudes, the F2 layer daytime ion production is higher in the summer, as expected, since the Sun shines more directly on the Earth. However, there are seasonal changes in the molecular-to-atomic ratio of the neutral atmosphere that cause the summer ion loss rate to be even higher. The result is that the increase in the summertime loss overwhelms the increase in summertime production, and total F2 ionization is actually lower in the local summer months. This effect is known as the winter anomaly. The anomaly is always present in the northern hemisphere, but is usually absent in the southern hemisphere during periods of low solar activity.

### [[edit](http://en.wikipedia.org/w/index.php?title=Ionosphere&action=edit&section=10)] Equatorial anomaly

[](http://en.wikipedia.org/wiki/File:Diurnal_ionospheric_current.jpg)

Electric currents created in sunward ionosphere.

Within approximately ± 20 degrees of the *magnetic equator*, is the [*equatorial*](http://en.wikipedia.org/wiki/Equator) *anomaly*. It is the occurrence of a trough of concentrated ionization in the F2 layer. The Earth's [magnetic field](http://en.wikipedia.org/wiki/Magnetic_field) lines are horizontal at the magnetic equator. Solar heating and [tidal](http://en.wikipedia.org/wiki/Tidal) oscillations in the lower ionosphere move plasma up and across the magnetic field lines. This sets up a sheet of electric current in the E region which, with the [horizontal](http://en.wikipedia.org/wiki/Horizontal_plane) magnetic field, forces ionization up into the F layer, concentrating at ± 20 degrees from the magnetic equator. This phenomenon is known as the *equatorial fountain*.

### Equatorial electrojet

The worldwide solar-driven wind results in the so-called Sq (solar quiet) current system in the E region of the Earth's ionosphere (100–130 km altitude). Resulting from this current is an electrostatic field directed E-W (dawn-dusk) in the equatorial day side of the ionosphere. At the magnetic dip equator, where the geomagnetic field is horizontal, this electric field results in an enhanced eastward current flow within ± 3 degrees of the magnetic equator, known as the [equatorial electrojet](http://en.wikipedia.org/wiki/Equatorial_electrojet).

## Ionospheric perturbations

### X-rays: sudden ionospheric disturbances (SID)

When the Sun is active, strong [solar flares](http://en.wikipedia.org/wiki/Solar_flare) can occur that will hit the Earth with hard X-rays on the sunlit side of the Earth. They will penetrate to the D-region, release electrons which will rapidly increase absorption causing a High Frequency (3 - 30 MHz) radio blackout. During this time Very Low Frequency (3 – 30 kHz) signals will become reflected by the D layer instead of the E layer, where the increased atmospheric density will usually increase the absorption of the wave, and thus dampen it. As soon as the X-rays end, the [sudden ionospheric disturbance](http://en.wikipedia.org/wiki/Sudden_ionospheric_disturbance) (SID) or radio black-out ends as the electrons in the D-region recombine rapidly and signal strengths return to normal.

### Protons: polar cap absorption (PCA)

Associated with solar flares is a release of high-energy protons. These particles can hit the Earth within 15 minutes to 2 hours of the solar flare. The protons spiral around and down the magnetic field lines of the Earth and penetrate into the atmosphere near the magnetic poles increasing the ionization of the D and E layers. PCA's typically last anywhere from about an hour to several days, with an average of around 24 to 36 hours.

### Geomagnetic storms

A [geomagnetic storm](http://en.wikipedia.org/wiki/Geomagnetic_storm) is a temporary intense disturbance of the Earth's [magnetosphere](http://en.wikipedia.org/wiki/Magnetosphere).

* During a geomagnetic storm the F2 layer will become unstable, fragment, and may even disappear completely.
* In the Northern and Southern pole regions of the Earth [aurora](http://en.wikipedia.org/wiki/Polar_aurora) will be observable in the sky.

### Lightning

[Lightning](http://en.wikipedia.org/wiki/Lightning) can cause ionospheric perturbations in the D-region in one of two ways. The first is through VLF frequency radio waves launched into the [magnetosphere](http://en.wikipedia.org/wiki/Magnetosphere). These so-called "whistler" mode waves can interact with radiation belt particles and cause them to precipitate onto the ionosphere, adding ionization to the D-region. These disturbances are called Lightning-induced Electron Precipitation (LEP) events.

Additional ionization can also occur from direct heating/ionization as a result of huge motions of charge in lightning strikes. These events are called Early/Fast.

In 1925, C. F. Wilson proposed a mechanism by which electrical discharge from lightning storms could propagate upwards from clouds to the ionosphere. Around the same time, Robert Watson-Watt, working at the Radio Research Station in Slough, UK, suggested that the ionospheric sporadic E layer (Es) appeared to be enhanced as a result of lightning but that more work was needed. In 2005, C. Davis and C. Johnson, working at the Rutherford Appleton Laboratory in Oxfordshire, UK, demonstrated that the Es layer was indeed enhanced as a result of lightning activity. Their subsequent research has focussed on the mechanism by which this process can occur.

## Radio application

[DX communication](http://en.wikipedia.org/wiki/DX_communication), popular among [amateur radio](http://en.wikipedia.org/wiki/Amateur_radio) enthusiasts, is a term given to communication over great distances. Thanks to the property of ionized atmospheric gases to [refract](http://en.wikipedia.org/wiki/Refract) high frequency (HF, or [shortwave](http://en.wikipedia.org/wiki/Shortwave)) radio waves, the ionosphere can be utilized to "bounce" a transmitted signal down to ground. Transcontinental HF-connections rely on up to 5 bounces, or [hops](http://en.wikipedia.org/wiki/Hop_%28telecommunications%29). Such communications played an important role during [World War II](http://en.wikipedia.org/wiki/World_War_II). [Karl Rawer](http://en.wikipedia.org/wiki/Karl_Rawer)'s most sophisticated prediction method[[1]](http://en.wikipedia.org/wiki/Ionosphere#cite_note-rawer-0) took account of several (zig-zag) paths, attenuation in the D-region and predicted the 11-years [solar cycle](http://en.wikipedia.org/wiki/Sun) by a method due to [Wolfgang Gleißberg](http://en.wikipedia.org/wiki/Solar_variation).

### Mechanism of refraction

When a radio wave reaches the ionosphere, the [electric field](http://en.wikipedia.org/wiki/Electric_field) in the wave forces the electrons in the ionosphere into [oscillation](http://en.wikipedia.org/wiki/Oscillation) at the same frequency as the radio wave. Some of the radio-frequency energy is given up to this resonant oscillation. The oscillating electrons will then either be lost to recombination or will re-radiate the original wave energy. Total refraction can occur when the collision frequency of the ionosphere is less than the radio frequency, and if the electron density in the ionosphere is great enough.

The [critical frequency](http://en.wikipedia.org/wiki/Critical_frequency) is the limiting frequency at or below which a radio wave is reflected by an ionospheric layer at vertical [incidence](http://en.wikipedia.org/wiki/Angle_of_incidence). If the transmitted frequency is higher than the [plasma frequency](http://en.wikipedia.org/wiki/Plasma_frequency) of the ionosphere, then the electrons cannot respond fast enough, and they are not able to re-radiate the signal. It is calculated as shown below:

f_{critical} = 9 \times 10^{-3} \sqrt{N}

where N = electron density per cm3 and fcritical is in MHz.

The Maximum Usable Frequency (MUF) is defined as the upper frequency limit that can be used for transmission between two points at a specified time.

f_{muf} = \frac{f_{critical}}{ \sin \alpha} 

where α = [angle of attack](http://en.wikipedia.org/wiki/Angle_of_attack), the angle of the wave relative to the [horizon](http://en.wikipedia.org/wiki/Horizon), and sin is the [sine](http://en.wikipedia.org/wiki/Sine) function.

The [cutoff frequency](http://en.wikipedia.org/wiki/Cutoff_frequency) is the frequency below which a radio wave fails to penetrate a layer of the ionosphere at the incidence angle required for transmission between two specified points by refraction from the layer.

## Other applications

The [open system](http://en.wikipedia.org/wiki/Open_system_%28systems_theory%29) [electrodynamic tether](http://en.wikipedia.org/wiki/Electrodynamic_tether), which uses the ionosphere, is being researched. The [space tether](http://en.wikipedia.org/wiki/Space_tether) uses plasma contactors and the ionosphere as parts of a circuit to extract energy from the Earth's magnetic field by [electromagnetic induction](http://en.wikipedia.org/wiki/Electromagnetic_induction).

## Measurements

### Ionograms

Ionograms show the virtual heights and **critical frequencies** of the ionospheric layers and which are measured by an [ionosonde](http://en.wikipedia.org/wiki/Ionosonde). An ionosonde sweeps a range of frequencies, usually from 0.1 to 30 MHz, transmitting at vertical incidence to the ionosphere. As the frequency increases, each wave is refracted less by the ionization in the layer, and so each penetrates further before it is reflected. Eventually, a frequency is reached that enables the wave to penetrate the layer without being reflected. For ordinary mode waves, this occurs when the transmitted frequency just exceeds the peak plasma, or critical, frequency of the layer. Tracings of the reflected high frequency radio pulses are known as ionograms. Reduction rules are given in: "URSI Handbook of Ionogram Interpretation and Reduction", edited by [William Roy Piggott](http://en.wikipedia.org/wiki/William_Roy_Piggott) and [Karl Rawer](http://en.wikipedia.org/wiki/Karl_Rawer), Elsevier Amsterdam, 1961.

### Incoherent scatter radars

[Incoherent scatter](http://en.wikipedia.org/wiki/Incoherent_scatter) radars operate above the critical frequencies. Therefore the technique allows to probe the ionosphere, unlike ionosondes, also above the electron density peaks. The thermal fluctuations of the electron density scattering the transmitted signals lack [coherence](http://en.wikipedia.org/wiki/Coherence_%28physics%29), which gave the technique its name. Their power spectrum contains information not only on the density, but also on the ion and electron temperatures, ion masses and drift velocities.

### Solar flux

Solar flux is a measurement of the intensity of solar radio emissions at a frequency of 2800 MHz made using a [radio telescope](http://en.wikipedia.org/wiki/Radio_telescope) located in [Dominion Radio Astrophysical Observatory](http://en.wikipedia.org/wiki/Dominion_Radio_Astrophysical_Observatory), Penticton, British Columbia, Canada.[[5]](http://en.wikipedia.org/wiki/Ionosphere#cite_note-4) Known also as the 10.7 cm flux (the wavelength of the radio signals at 2800 MHz), this solar radio emission has been shown to be proportional to sunspot activity. However, the level of the Sun's ultraviolet and X-ray emissions is primarily responsible for causing ionization in the Earth's upper atmosphere. We now have data from the [GOES](http://en.wikipedia.org/wiki/GOES) spacecraft that measures the background **X-ray flux** from the Sun, a parameter more closely related to the ionization levels in the ionosphere.

* The [*A*](http://en.wikipedia.org/wiki/A-index) and [*K*](http://en.wikipedia.org/wiki/K-index) indices are a measurement of the behavior of the horizontal component of the **geomagnetic field**. The *K* index uses a scale from 0 to 9 to measure the change in the horizontal component of the geomagnetic field. A new *K* index is determined at the [Table Mountain Observatory](http://en.wikipedia.org/wiki/Table_Mountain_Observatory), north of [Boulder](http://en.wikipedia.org/wiki/Boulder,_Colorado), [Colorado](http://en.wikipedia.org/wiki/Colorado).
* The geomagnetic activity levels of the Earth are measured by the fluctuation of the Earth's magnetic field in [SI](http://en.wikipedia.org/wiki/SI) units called [teslas](http://en.wikipedia.org/wiki/Tesla_%28unit%29) (or in non-SI [gauss](http://en.wikipedia.org/wiki/Gauss_%28unit%29), especially in older literature). The Earth's magnetic field is measured around the planet by many observatories. The data retrieved is processed and turned into measurement indices. Daily measurements for the entire planet are made available through an estimate of the *ap* index, called the *planetary A-index* (PAI).

### [[edit](http://en.wikipedia.org/w/index.php?title=Ionosphere&action=edit&section=24)] Scientific research on ionospheric propagation

Scientists also are exploring the structure of the ionosphere by a wide variety of methods, including passive observations of optical and radio emissions generated in the ionosphere, bouncing radio waves of different frequencies from it, [incoherent scatter](http://en.wikipedia.org/wiki/Incoherent_scatter) radars such as the [EISCAT](http://en.wikipedia.org/wiki/EISCAT), Sondre Stromfjord, [Millstone Hill](http://en.wikipedia.org/wiki/Millstone_Hill_Observatory), [Arecibo](http://en.wikipedia.org/wiki/Arecibo_Observatory), and [Jicamarca](http://en.wikipedia.org/wiki/Jicamarca_Radio_Observatory) radars, coherent scatter radars such as the [Super Dual Auroral Radar Network (SuperDARN)](http://en.wikipedia.org/wiki/Super_Dual_Auroral_Radar_Network) radars, and using special receivers to detect how the reflected waves have changed from the transmitted waves.

A variety of experiments, such as HAARP ([High Frequency Active Auroral Research Program](http://en.wikipedia.org/wiki/High_Frequency_Active_Auroral_Research_Program)), involve high power radio transmitters to modify the properties of the ionosphere. These investigations focus on studying the properties and behavior of ionospheric plasma, with particular emphasis on being able to understand and use it to enhance communications and surveillance systems for both civilian and military purposes. HAARP was started in 1993 as a proposed twenty year experiment, and is currently active near Gakona, Alaska.

The SuperDARN radar project researches the high- and mid-latitudes using coherent backscatter of radio waves in the 8 to 20 MHz range. Coherent backscatter is similar to Bragg scattering in crystals and involves the constructive interference of scattering from ionospheric density irregularities. The project involves more than 11 different countries and multiple radars in both hemispheres.

Scientists are also examining the ionosphere by the changes to radio waves from satellites and stars passing through it. The [Arecibo radio telescope](http://en.wikipedia.org/wiki/Arecibo_radio_telescope) located in [Puerto Rico](http://en.wikipedia.org/wiki/Puerto_Rico), was originally intended to study Earth's ionosphere.

## [[edit](http://en.wikipedia.org/w/index.php?title=Ionosphere&action=edit&section=25)] Ionospheres on other planets and Titan

The [atmosphere of Titan](http://en.wikipedia.org/wiki/Atmosphere_of_Titan), the only moon known to have one, includes an ionosphere.[[6]](http://en.wikipedia.org/wiki/Ionosphere#cite_note-5) It ranges from about 1100 to 1300 km in altitude, and contains carbon compounds.

Planets with ionospheres (incomplete list):

[Ionosphere of Venus](http://en.wikipedia.org/wiki/Atmosphere_of_Venus#Upper_atmosphere_and_ionosphere)

[Ionosphere of Uranus](http://en.wikipedia.org/wiki/Atmosphere_of_Uranus#Thermosphere_and_ionosphere)

## [[edit](http://en.wikipedia.org/w/index.php?title=Ionosphere&action=edit&section=26)] History

[Guglielmo Marconi](http://en.wikipedia.org/wiki/Guglielmo_Marconi) received the first trans-Atlantic radio signal on December 12, 1901, in [St. John's, Newfoundland](http://en.wikipedia.org/wiki/St._John%27s,_Newfoundland) (now in [Canada](http://en.wikipedia.org/wiki/Canada)) using a 152.4 m (500 ft) kite-supported antenna for reception. The transmitting station in [Poldhu](http://en.wikipedia.org/wiki/Poldhu), Cornwall used a spark-gap transmitter to produce a signal with a frequency of approximately 500 [kHz](http://en.wikipedia.org/wiki/Kilohertz) and a power of 100 times more than any radio signal previously produced. The message received was three dits, the [Morse code](http://en.wikipedia.org/wiki/Morse_code) for the letter **S**. To reach Newfoundland the signal would have to bounce off the ionosphere twice. Dr. Jack Belrose has recently contested this, however, based on theoretical and experimental work.[[7]](http://en.wikipedia.org/wiki/Ionosphere#cite_note-6) However, Marconi did achieve transatlantic wireless communications beyond a shadow of doubt in [Glace Bay, Nova Scotia](http://en.wikipedia.org/wiki/Glace_Bay,_Nova_Scotia) one year later.

In 1902, [Oliver Heaviside](http://en.wikipedia.org/wiki/Oliver_Heaviside) proposed the existence of the *Kennelly-Heaviside Layer* of the ionosphere which bears his name. Heaviside's proposal included means by which radio signals are transmitted around the Earth's curvature. Heaviside's proposal, coupled with Planck's law of black body radiation, may have hampered the growth of radio astronomy for the detection of electromagnetic waves from celestial bodies until 1932 (and the development of high frequency radio transceivers). Also in 1902, [Arthur Edwin Kennelly](http://en.wikipedia.org/wiki/Arthur_Edwin_Kennelly) discovered some of the ionosphere's radio-electrical properties.

In 1912, the [U.S. Congress](http://en.wikipedia.org/wiki/U.S._Congress) imposed the [Radio Act of 1912](http://en.wikipedia.org/wiki/Radio_Act_of_1912) on [amateur radio operators](http://en.wikipedia.org/wiki/Amateur_radio_operators), limiting their operations to frequencies above 1.5 MHz (wavelength 200 meters or smaller). The government thought those frequencies were useless. This led to the discovery of HF radio propagation via the ionosphere in 1923.

**Faraday rotation** is a [Magneto-optical](http://en.wikipedia.org/wiki/Magneto-optic) phenomenon, that is, an interaction between [light](http://en.wikipedia.org/wiki/Light) and a [magnetic](http://en.wikipedia.org/wiki/Magnetic) field in a medium. The Faraday effect causes a rotation of the plane of [polarization](http://en.wikipedia.org/wiki/Polarization_%28waves%29) which is linearly proportional to the component of the magnetic field in the direction of propagation.

Discovered by [Michael Faraday](http://en.wikipedia.org/wiki/Michael_Faraday) in 1845, the Faraday effect was the first experimental evidence that light and electromagnetism are related. The theoretical basis of [electromagnetic radiation](http://en.wikipedia.org/wiki/Electromagnetic_radiation) (which includes visible light) was completed by [James Clerk Maxwell](http://en.wikipedia.org/wiki/James_Clerk_Maxwell) in the 1860s and 1870s. This effect occurs in most optically [transparent](http://en.wikipedia.org/wiki/Transparency_%28optics%29) [dielectric](http://en.wikipedia.org/wiki/Dielectric) materials (including liquids) under the influence of [magnetic fields](http://en.wikipedia.org/wiki/Magnetic_field).

The Faraday effect causes left and right circularly polarized waves to propagate at slightly different speeds, a property known as [circular birefringence](http://en.wikipedia.org/wiki/Optical_activity#Theory). Since a linear polarization can be decomposed into the [superposition](http://en.wikipedia.org/wiki/Superposition_principle) of two equal-amplitude circularly polarized components of opposite handedness and different phase, the effect of a relative [phase](http://en.wikipedia.org/wiki/Phase_%28waves%29) shift, induced by the Faraday effect, is to rotate the orientation of a wave's linear polarization.

The Faraday effect has a few applications in measuring instruments. For instance, the Faraday effect has been used to measure optical rotatory power and for remote sensing of magnetic fields. The Faraday effect is used in [spintronics](http://en.wikipedia.org/wiki/Spintronics) research to study the polarization of electron spins in semiconductors. [Faraday rotators](http://en.wikipedia.org/wiki/Faraday_rotator) can be used for amplitude modulation of light, and are the basis of [optical isolators](http://en.wikipedia.org/wiki/Optical_isolator) and [optical circulators](http://en.wikipedia.org/wiki/Optical_circulators); such components are required in optical telecommunications and other laser applications.

**WHISTLER**

**whistler** is a [very low frequency](http://en.wikipedia.org/wiki/Very_low_frequency) [electromagnetic (radio) wave](http://en.wikipedia.org/wiki/Electromagnetic_radiation) which can be generated, for example, by [lightning](http://en.wikipedia.org/wiki/Lightning). Frequencies of terrestrial whistlers are 1 to 30 kHz, with maximum usually at 3 to 5 kHz. Although they are electromagnetic waves, they occur at [audio frequencies](http://en.wikipedia.org/wiki/Audio_frequency), and can be converted to audio using a suitable receiver. They are produced by lightning strikes (mostly intracloud and return-path) where the impulse travels away from the earth and returns to the earth traveling along magnetic field lines. They undergo [dispersion](http://en.wikipedia.org/wiki/Dispersion) of several thousand kHz due to the slower velocity of the lower frequencies through the [plasma](http://en.wikipedia.org/wiki/Plasma_%28physics%29) environments of the [ionosphere](http://en.wikipedia.org/wiki/Ionosphere) and [magnetosphere](http://en.wikipedia.org/wiki/Magnetosphere). Thus they are perceived as a descending tone which can last for a few seconds. The study of whistlers allows categorization into Pure Note Whistlers, Diffuse, 2-hop, and Echo Train types.