5. Some Basics of Radio Astronomy

This section contains some basic terms and concepts that are important in understanding radio emission measurement techniques.

5.1 Characteristics of the Signals

5.1.1 Electromagnetic Radiation: Rays, Waves, or Photons?

Wavelength and frequency ranges; what's λ, what's ν?

The radio-frequency range is arbitrarily defined to be from frequencies (ν) beginning at the super-sonic, meaning just above the range of sound waves that humans can hear, 20 kHz or so, up to about 600 GHz, where the far infrared begins. The corresponding wavelength (λ) range is from many kilometers down to about 0.5 mm or 500 microns. Frequency and wavelength are related, as for any conventional wave phenomena, by:

$$\nu = c/\lambda$$

where c is the speed of light, 299792.5 km/s.

Over this frequency range and at moderate equivalent temperatures, we can usually ignore photons; that is, we need not quantize the electromagnetic field. In most cases, this results in both simplification and insight that is lacking if one thinks photons instead of waves.

Many formulas used in radio astronomy (and electrical engineering) are based on the Rayleigh-Jeans approximation to the Planck black-body law. If, furthermore, the antenna structures are large compared to a wavelength, as is usually the case at centimeter and millimeter wavelengths, then we can often ignore waves also and use just ray-trace optics.

5.1.2 Law of large distances

Another helpful simplification involves the large distances and small angular sized of most astronomical objects. Angles in radians are, then, just the linear sizes divided by the distances.

5.1.3 Noise-like signals

A. Law of large numbers

Most astronomical sources are large in physical size, even though small in angular size, and radiation is emitted by a large number of statistically independent sources—atoms, molecules, or

electrons. The resulting signals are noise-like, that is, the electric fields are Gaussian random variables with spectra that depend on the details of the emission mechanisms. Laboratory laser and maser oscillators are usually coherent because of the cavity in which they oscillate. Astronomical masers have never been found to have non-Gaussian statistics.

For similar reasons, many radio-astronomical sources are unpolarized; that is, signals in one polarization are statistically independent of signals in the orthogonal polarization. But some sources, especially those involving magnetic fields that extend over a significant part of the spatial extent of the source, can be polarized, and studying such polarization sometimes leads to significant insights.

B. Law of large sizes

As a rough rule with some exceptions, an incoherent source that is not resolved in angle can be seen to vary only on time scales long compared with the light travel time through the length of the emitting region. A black sphere with cyclic temperature variations, for example, will have these variations smoothed over as seen from a large distance if the cycle time is comparable to or less than the light travel time across a radius of the sphere. Some astronomical sources vary, and the time scales of the variations can sometimes be used to infer maximum sizes.

C. Significance for measurement techniques

Some of the techniques used in radio astronomy depend on these characteristics of the sources. One-bit autocorrelation spectroscopy, for example, depends on the signal voltage being a Gaussian random variable. And many-hour-long interferometry depends on the source being stable over that time.

5.2 Spectra

5.2.1 Continuum sources, black body or ν^n?

A black body at temperature T in the radio-frequency range has a specific intensity given by the Raleigh-Jeans approximation to the Planck black-body law:

$$I = 2kT/\lambda^2$$

where I is the specific intensity in, for example, watts/(m²Hz ster), k is Boltzman's constant, 1.380 x 10^{-23} watts/Hz/K and λ is wavelength.

The proportionality to temperature allows us to talk about intensities and powers in temperature units, K or Kelvins. If a source really were thermal emission from a black or gray body, then this radiation temperature would be independent of wavelength. More typically, however, sources are

"colored", that is, radiation temperatures vary with wavelength and are not necessarily related to physical temperatures of the sources. A synchrotron-emission source, for example, has a spectral index n, usually defined by $I \propto \nu^{-n}$, that is related to the energy distribution of the emitting electrons. Regardless of the emission mechanism or spectrum, we can use the Rayleigh-Jeans equation above to define a brightness temperature, T_b, proportional to intensity, but then T_b will be, in general, a function of frequency. Even in the high-frequency low-temperature range where Rayleigh-Jeans is no longer useful, we can use this equation to define a convenient fake brightness temperature. Figure 1 is a cartoon example of such spectra.
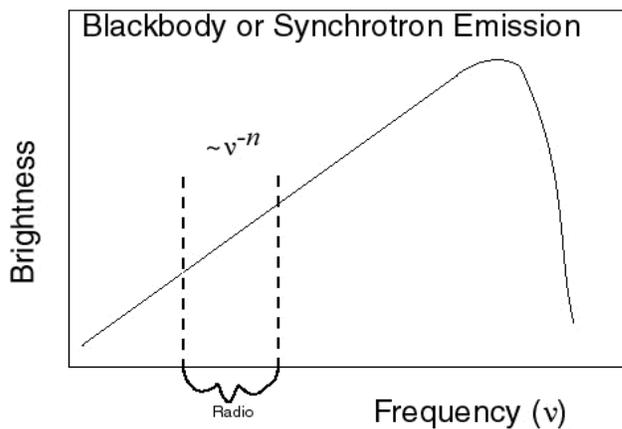

Figure 1

Electrical engineers implicitly use the Rayleigh-Jeans approximation also. The noise power per unit bandwidth from a warm resistor into a matched load is just kT provided that the frequency and temperature are in the range for which Rayleigh-Jeans is precise. To see this, imagine a resistor coupled to a feed looking at a black body. Even in the range where Rayleigh-Jeans is no longer useful; in thermal equilibrium, equal power must flow each way.

5.2.2 Atoms and molecules—emission and absorption lines

Emission and absorption lines in radio astronomy usually originate from atoms and small molecules or molecular ions in gaseous form, and molecular transitions at radio wavelengths are usually rotational. Emission lines result from warm gas overlying a cold background so that the intensity (or flux or radiation temperature) at the line frequency is sharply higher compared to nearby wavelengths. If such a gas cloud is optically thick (opaque), the specific intensity or "brightness" at the line frequency is given by the state temperature of the corresponding transition, which, at thermodynamic equilibrium, would be just the temperature. Thermodynamic equilibrium is, however, not very common in radio astronomy. Absorption lines result from cool gas overlying a hotter background source so that the intensity at the line frequency is sharply lower compared to nearby wavelengths. If such a gas cloud is optically thick, then the line center again gives the state temperature. Figure 2 is a cartoon of such spectra.
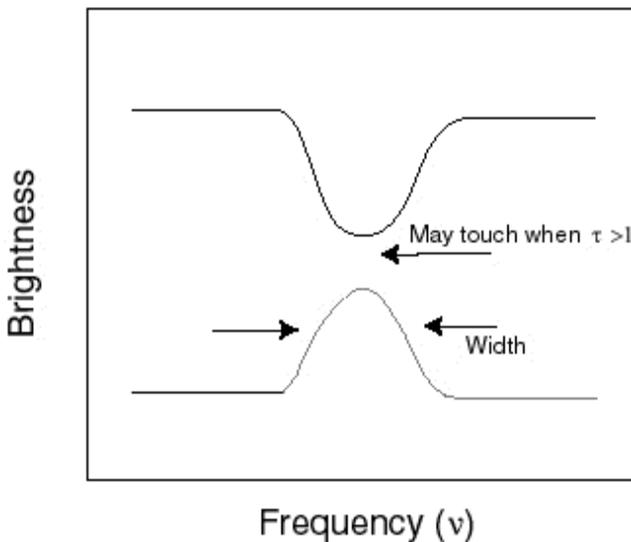

Figure 2

5.2.3 Doppler shifts and kinematics

Doppler shifts are very important in spectral-line radio astronomy. The non-relativistic form is usually written as

$$\Delta \nu / \nu = -v/c$$

where $\Delta \nu$ is the change in frequency ν due to the Doppler velocity v, defined as the rate of change of distance from source to observer (hence the minus sign), and c is the speed of light.

Even when speeds are relativistic, this non-relativistic formula is sometimes still used to define a convenient fake velocity.

Line widths and line shapes are influenced by several line-broadening mechanisms including *a*) natural line widths related to the lifetimes of the states involved in the transition, *b*) kinematic temperatures characterizing the small-scale random motions of the atoms or molecules, *c*) turbulence or larger-scale random motions, and *d*) kinematics, by which we mean large-scale ordered motions such as expansions, contractions, or rotations. A useful exercise is to estimate the spectrum that would be seen from some simple kinematic models such as a circumstellar shell or sphere that is expanding, contracting (infalling), or rotating and is unresolved in angle. In some cases the central star ionizes nearby gas, which makes a central continuum source. Such an object can then show both emission and absorption lines.

5.3 Antennas at Radio Wavelengths

5.3.1 Parabolic–why?

Radio-astronomical sources are far away, so incoming signals often look like plane waves from a specific direction (*point sources*), and the first goal of a radio telescope is to catch as much energy as possible from such a wave and avoid as much as possible any other signals, especially local interference. The signal at the antenna in this case can be characterized by a flux density in Janskys (1 Jy = 10^{-26} w/m^2/Hz), so the bigger the antenna, the more watts (well, pico-pico watts) we collect. A parabolic antenna (i.e., a parabola of revolution), which puts all this energy into a small spot where a *feed* can be placed, is usually an engineering optimum for centimeter and millimeter wavelengths.

5.3.2 Aperture efficiency and K/Jy

Consider a black sphere with diameter *d* and temperature *T* at a distance *r* from a circular receiving antenna with diameter *D*. Assume that *r* is much larger than either *d* or *D*. Figure 3 is a cartoon of this situation.
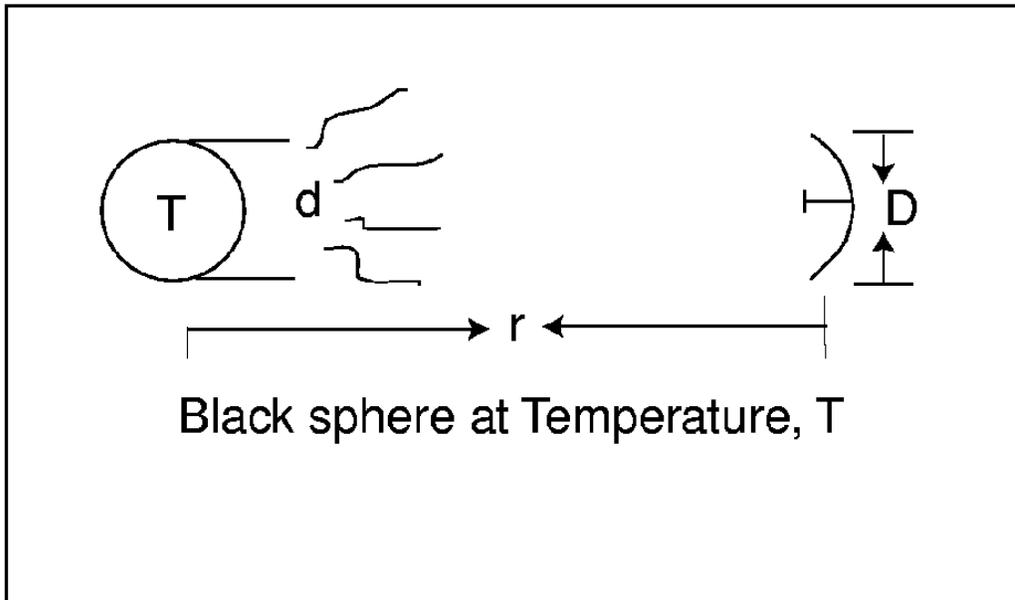
Black sphere at Temperature, T

Figure 3

Then the power density (power per unit frequency interval) received by the antenna, P, can be calculated as its collecting area times the flux from the source at the antenna or as the specific intensity of the source times the solid angle of the antenna as seen from the source. Either way gives the same formula, namely,

$$P = I(\rho d^2/4)(1/r^2)(\rho D^2/4)$$

The first three terms on the right are the source flux density, F, and $A = \pi D^2/4$ is the antenna collecting area, which is usually smaller than its physical area due to various losses. If we characterize P in temperature units, $P = 2kT_R$, as usual, then

$$T_R/F = A/2k$$

is a figure of merit, sometimes called *sensitivity*, in Kelvins per Jansky for the antenna. That extra 2 is because the flux density refers to the total in both polarizations, but a single receiver can only receive one polarization. The ratio of the antenna collecting area from this formula to its physical area is called *aperture efficiency* usually expressed as a percentage and usually 60% or less.

5.3.3 Beam efficiency and beam dilution

Another figure of merit, appropriate for sources extended in angle, is the *beam efficiency*, crudely defined as the ratio of T_R to the brightness temperature of the source. (There are more precise but less useful definitions.) Beam efficiency by this definition is, however, a function of the source angular size and shape, alas. A more useful parameter is the *beam dilution*, defined in the same way but for an assumed circular source with a specified diameter in beamwidths and as a function of this diameter. Planets with known brightness temperatures and angular sizes are candidate calibration sources for measuring aperture efficiency when they are small in angle compared with a beamwidth or points on the beam-dilution curve when they are larger in angle. Figure 4 is a cartoon of this situation.
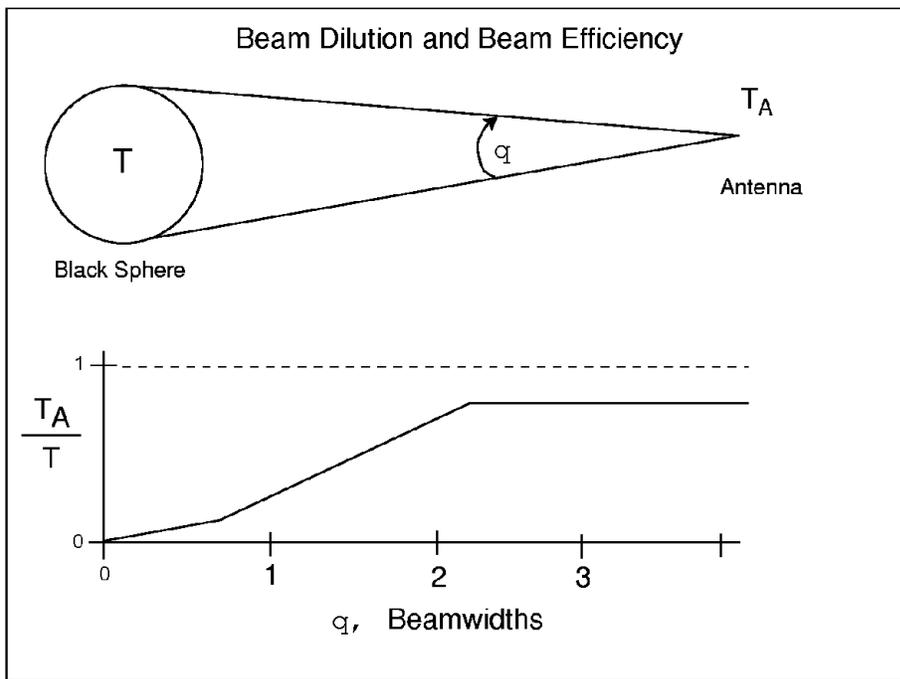

Figure 4

5.3.4 Cassegrain–why?

A Cassegrain antenna comprises a parabolical main reflector and a concave hyperbolical *subreflector* near the *prime focus* of the main reflector to reflect incoming signals back to a spot near the center of the main reflector, the *secondary focus*, where the feed is placed. This feed is usually mounted on the front of a receiver box that fits through a hole in the center of the main reflector. This arrangement trades a little additional aperture blockage (the subreflector is larger than a prime-focus feed would be), for the ability to place additional equipment, such as a cryogenic refrigerator, near the secondary focus.

5.3.5 Beamwidth: λD

The *beamwidth* (full width to half power) of the radiation pattern of a circular antenna is approximately 1.2 λ/D in radians, where λ is the wavelength and D is the diameter of the antenna in, of course, the same units. The 1.2 factor depends somewhat on the feed illumination pattern, that is, on the pattern of the feed as seen from the main reflector. A circular antenna with circular illumination gives a circular beam.

When a finite antenna is used to map extended sources over a range of angles one can show the resulting maps are *band limited* in the sense that they contain no angular frequencies above a maximum (called the Nyquist limit) that depends only on the wavelength and the antenna diameter. Band-limited maps are, then, smooth continuous functions of two angles on the sky, but they can be specified or measured at a finite grid of evenly spaced points provided that these points are no farther apart than a Nyquist step, which is $\lambda/(2D)$ in radians. A Nyquist is typically a little less than half a beamwidth. Smooth maps can be obtained from finite grids of points by convolution.

5.3.6 Requirements for surface precision

The effective area of an antenna is less than its physical area because of various losses, one of which is due to the departure of the surface from an ideal parabolic shape. An antenna with a surface that is rough on the scale of a wavelength will be almost useless because of low aperture efficiency and also susceptibility to interference scattered into the feed. An ideal antenna would have at least so-called 1/20-wave optics, meaning that the surface is within a 1/20 of a wavelength of perfect. We must sometimes make do with antennas less than ideal; all antennas have some short-wavelength limit based on this criterion.

5.4 Interferometers

5.4.1 Why? Resolution: λ/D

Two or more antennas looking at the same source at the same time can have their signals combined into an interferometer to give some of the information that would be obtained from a single antenna with a diameter equal the spacing between the interferometer antennas. The resolution of a two-antenna interferometer is approximately λ/D (no 1.2), where D is now the spacing between antennas, but this resolution is only in the direction parallel to a line connecting the antennas and only for sources in a plane perpendicular to this line. A single measurement with a single pair of antennas gives a single datum to be placed in the so-called *uv* plane, which is the Fourier-transform plane of the angular distribution of the source on the sky. The *uv* plane is roughly comparable to the plane of the surface of an equivalent single antenna. Even with only two antennas, because of Earth's rotation, data can be obtained over time along two elliptical arcs in the *uv* plane. More antennas give more elliptical arcs, often crossing, but unless the antennas actually touch or overlap, there will always be gaps with no data in the *uv* plane. The corresponding synthesized antenna beam has side-lobes because of the missing *uv* data. There are helpful data-reduction techniques that amount to model fitting or to interpolating or extrapolating across these gaps.

5.4.2 Connected vs VLBI

Interferometers with antennas spaced up to, say, a few kilometers usually have cables or fibers connecting the antennas to a central site where the cross correlations are done in real time. With wider spacing, including transcontinental and intercontinental interferometers, the techniques of tape-recorder or very-long-baseline interferometry (VLBI) are used. At each antenna, data and microsecond timing information, usually derived from atomic clocks, are recorded on magnetic tapes to be shipped to a central site for cross correlation at a time that may be weeks or months later. This delay is a disadvantage, especially for troubleshooting, but there is almost no other way to do large data rates on megameter baselines with their milliarcsecond resolutions. Figure 5 is a sample uv coverage plot of 3C84 for a realistic VLBI experiment.
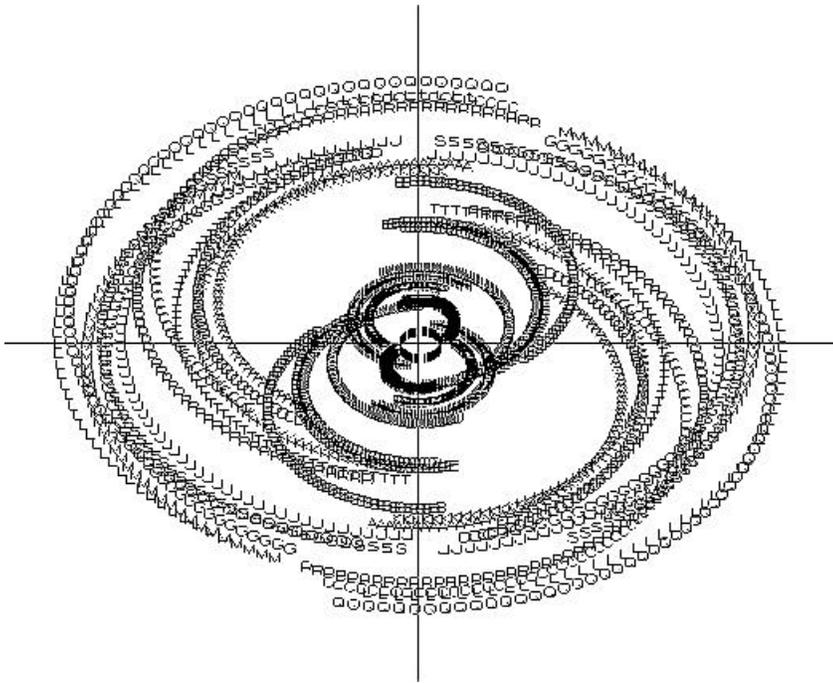
Figure 5.