FY1003 ELEKTRISITET OG MAGNETISME

Project paper on:

Radio Astronomy

- Radio Telescopes -

Kristin Borck
Andrea Klubicka

23/04-2004

Abstract

We introduce the most general types of radio telescopes and describe how these operate; starting from the simplest type, the single aperture telescope, then moving on to the function of a radio telescope consisting of two apertures and finally presenting the most complex combinations. The theory behind electromagnetic waves and how they radiate, is occupying a rather big part of this paper, as it is essential in understanding how radio telescopes function. Examples and a description of the origins of radio waves from space are also presented, and a detailed explanation of radio signals and noise is included to bring even more clarity on the subject of radio telescopes.

Introduction [1-3]

Almost everything we know about the universe, about stars and stellar systems, their distribution, kinematics and dynamics, has been obtained from information brought to us by electromagnetic radiation. Only a small part of our knowledge stems form material sources of information, such as meteorites that hit the surface of the Earth, cosmic particles of radiation or samples of material collected by manned or unmanned space probes.

For thousands of years, all the information that astronomers gathered about the universe was based on visible light. In the twentieth century, however, the wavelength range was slightly expanded. Radio waves were the first part of the electromagnetic spectrum beyond the visible light to be exploited for astronomy. Karl Jansky, a young electrical engineer at Bell Telephone Laboratories, was the first to discover these waves. This happened in the early 1930s, as he was trying to locate what was causing interference with the then-new transatlantic radio link at a wavelength of 14.6 m. He realized that one kind of radio noise is strongest when the constellation Sagittarius is high in the sky. Since the center of our galaxy is located in the direction of Sagittarius, he concluded that he was detecting radio waves from a source beyond the Earth. As Jansky continued his observations, he showed that the principal sources of radiation were distributed throughout the Milky Way, and that the radiation could not be from stellar sources similar to the Sun.

Jansky's observations were taken up and improved by the radio engineer Grote Reber. In 1936 he built the first radio telescope, a radio-wave detector dedicated to astronomy. In the years from 1938 to 1944 Reber did his measurements at wavelengths of 1.9 m and 0.63 m. He detected radio waves emitted from the entire Milky Way, with the greatest emission from the center of the galaxy. These observations were published in the *Astrophysical Journal* in 1944.

Radio physics made great progress during World War II, mainly due to the development of sensitive and efficient radar equipment. After the end of the war, new receivers were instrumental in opening up the new radio window in the Earth's atmosphere. The radio window reaching from $\lambda \cong$ 10-15 m to about $\lambda \cong$ 0.1 mm or less was the first new spectral range that became available to astronomy outside the slightly expanded optical window. Soon different new objects could be discovered, and the new astronomical discipline of radio astronomy contributed to changing our views on many subjects, requiring mechanisms for their explanation that differed considerably from those commonly used in “optical” astrophysics.

In the years since 1945 technological progress permitted the opening up of several additional spectral windows of widely different wavelength ranges, such as infrared, ultraviolet and X-ray. Each spectral window requires its own technology, and the ways of doing measurements differ for each. Therefore astronomers view these different techniques as different sections of astronomy: radio astronomy, infrared astronomy, X-ray astronomy and so on.

This paper concentrates mostly on the principals behind the simplest type of radio telescope, explaining the concepts of angular resolution, sensitivity and such. An application to a two element telescope is also presented, and finally we mention the generalization to N element telescopes.

Electromagnetic Radiation [1, 4]

Maxwell’s equations are today the foundation of most physical phenomena around us, and every physical law in electromagnetism can be deduced from these four equations. They are distinct because they effectively describe both static and time dependent fields.

Maxwell’s equations in free space:

$$\nabla \cdot \vec{E} = 0 \tag{1}$$

$$\nabla \cdot \vec{B} = 0 \tag{2}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{3}$$

$$\nabla \times \vec{B} = m_0 e_0 \frac{\partial \vec{E}}{\partial t} \tag{4}$$

The theory of EM waves and the associated laws and properties are a natural consequence of these equations. When not considering propagation through empty space, Gauss law for electric fields (1),states that the total electric flux through a closed surface equals the net charge inside that surface (divided by ε_0). This law relates the electric field to the charge distribution, where electric field lines originate from positive charges and end up on negative charges. Gauss’ law for magnetism, (2), shows that the net magnetic flux through a closed surface is zero, meaning that the number of field lines entering a surface, equals the number leaving it. This proves that magnetic field lines cannot begin or end at any point. Maxwell-Ampères law,(4), predicts how a time dependent electric field will produce a magnetic field, while Faraday’s law,(3) indicates the opposite. By applying the curl and combining these two latter equations, the wave equation is obtained both for the electric and magnetic field, respectively:

$$\nabla\times(\nabla\times\vec{E})=\nabla(\nabla\cdot\vec{E})-\nabla^2\vec{E}=\nabla\times(-\frac{\partial\vec{B}}{\partial t})=-\frac{\partial\vec{B}}{\partial t}(\nabla\times\vec{B})=-m_0e_0\frac{\partial^2\vec{E}}{\partial t^2}, \quad (5)$$

$$\nabla\times(\nabla\times\vec{B})=\nabla(\nabla\cdot\vec{B})-\nabla^2\vec{B}=\nabla\times(m_0e_0\frac{\partial\vec{E}}{\partial t})=m_0e_0\frac{\partial}{\partial t}(\nabla\times\vec{E})=-m_0e_0\frac{\partial^2\vec{B}}{\partial t^2}, \quad (6)$$

and since $\nabla\cdot\vec{E}=0$ and $\nabla\cdot\vec{B}=0$,

$$\nabla^2\vec{E}=m_0e_0\frac{\partial^2\vec{E}}{\partial t^2} \text{ and } \nabla^2\vec{B}=m_0e_0\frac{\partial^2\vec{B}}{\partial t^2} \quad \text{(7a) and (7b)}$$

Still considering propagation through empty space, where Q = 0 and I = 0, Maxwell's equations provide an important result, namely that the wave velocity is equal to c, the speed of light. We can also deduce the fact that at every instant, the ratio of the electric field to the magnetic field of an electromagnetic wave equals this same constant.

From electrostatics we know that stationary electric charges generate electric fields, while accelerating charges generate propagating electric and magnetic fields, i.e. EM waves. These fields are positioned at right angles to each other and to the direction of the wave propagation. Thus, EM waves are transverse. When an electromagnetic wave leaves its source, it spreads out in straight lines and its oscillating fields weaken with distance, according to the relation 1/r. Thus, the concentration of electromagnetic waves gets smaller as the distance from the source gets bigger. This accounts for the loss of signal strength by the time radiation from outer space reaches us.

Which type of electromagnetic wave we are dealing with is determined by the frequency of the wave. For higher frequencies than our eyes can detect, we have ultraviolet radiation, x-rays and gamma rays. On the other side of the spectrum there is infrared radiation and radio waves. Associated with the frequency is the concept of wavelength. Since all the electromagnetic waves travel through vacuum with the speed of light, their frequency f and wavelength λ are related by the equation f λ = c. There is really no upper limit to the value of the radiation frequency. Traveling through different media, different wavelengths do not behave in the same way. This is why looking at a star with our eyes and through a telescope will give different images.

.

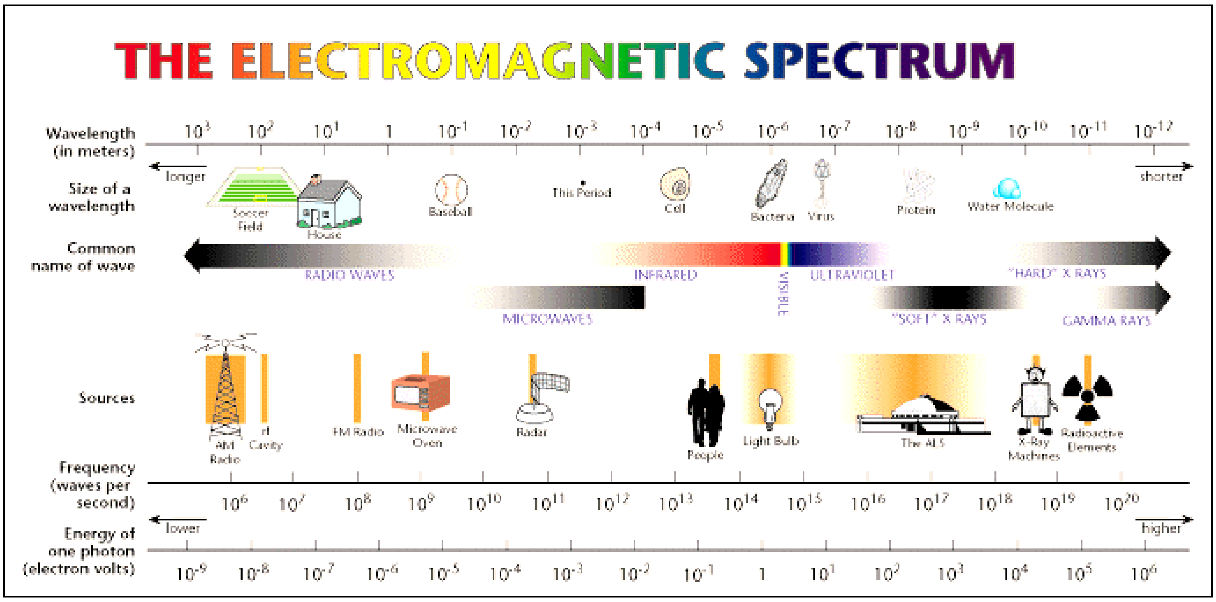


Figure 1 The EM spectrum covers a wide range of wavelengths and photon energies. [8]

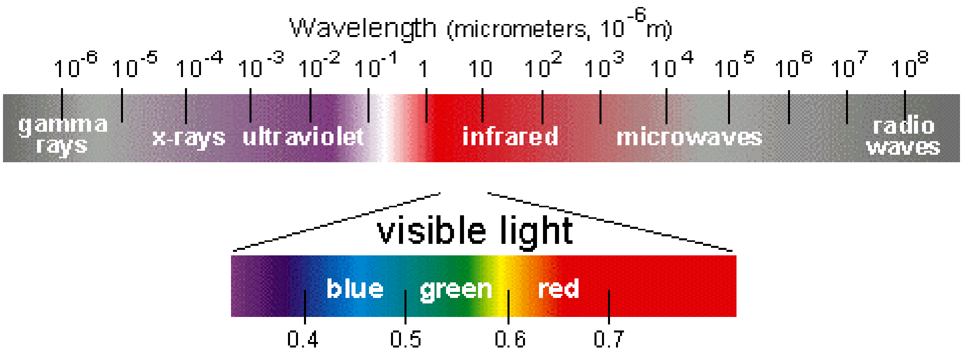


Figure 2 Radio waves have the longest wavelengths in the spectrum. [8]

Another important property of the waves is the polarization, which is defined by the direction of the electric vector. If this vector is constant with respect to the horizon, the wave is linearly polarized. In radio wave transmission, a wave is said to be horizontally polarized if it is parallel to the Earth's surface. When radiated vertically, we call it vertically polarized. A wave is circularly polarized when the electric and magnetic vectors rotate around the direction of wave propagation. Radio waves may be polarized in any of these ways, or not at all.

Electromagnetic waves carry both energy and momentum, so they can exert forces on surfaces they reach. The rate of flow of energy by an electromagnetic wave is described by the Poynting vector, $\vec{S}$ with direction along the wave propagation; $\vec{S} = \frac{1}{m_0}(\vec{E}\times\vec{B})$. The average of $\vec{S}$ taken over one or more cycles, gives the intensity I of the wave; $I = \langle \vec{S} \rangle = \frac{1}{2} c e_0 E_0^2$. In perfect blackbodies, all the radiation impinged will be absorbed, thus the total momentum delivered by the wave will be $\vec{p} = U/c$ and the radiation pressure will be given by $P = \vec{S}/c$. Similarly, perfect reflectors will deliver a momentum value of $\vec{p} = 2U/c$, which corresponds to a radiation pressure of $P = 2\vec{S}/c$.

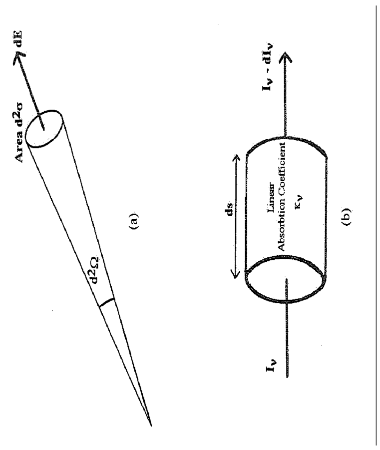


Figure 3 (a) specific intensity. (b) Emissivity and absorption

Specific intensity $I_n(\hat{n})$ of radiation (see Figure 3) at a point in space where energy dE is passing through an area $d^2\sigma$ in a solid angle $d^2\Omega$ is defined as:

$$I_n(\hat{n}) \equiv \frac{dE}{d^2\Omega\hat{n}\cdot d^2s\, dn\, dt} \quad , \tag{8}$$

where the units of power are per steridian, per unit area, per frequency interval. When matter comes in the path of the ray, energy can be absorbed or emitted and so the specific intensity will not be conserved. *Specific emissivity* j_ν is the power emitted per unit volume, per frequency, per steridian, while the reduction of specific intensity along a ray path depends on the linear absorption coefficient κ_ν. Traveling through a media the specific intensity may be reduced by absorption or scattering to a value dI_n.

The absorption coefficient is defined by

$$dI_n = -k_n I_n ds \,, \tag{9}$$

provided that the system is in a steady-state equilibrium. How specific intensity behaves along a ray path is described by the *equation of radiative matter*:

$$\frac{dI_n}{ds} = j_n - k_n I_n \,, \tag{10}$$

From this equation, some solutions can be given immediately, like Kirchhoff's law:

$$\frac{j_n}{k_n} = B_n(T) \,, \tag{11}$$

where $B_n(T)$ is Planck's function (34). Other simple solutions are obtained for emission and absorption only. A ray with specific intensity I_f^0 at the origin has,

$$I_f(s) = I_f^0 + \int_0^S j_f(s')ds' \qquad \text{(emission only)} \tag{12}$$

$$I_f(s) = I_f^0 e^{-\int_0^S k_f(s')ds'} \qquad \text{(absorption only)} \tag{13}$$

Electromagnetic waves can traverse the same space independently of one another and they have the property of superposition and interference. Another way of describing EM radiation is by photons. These can be thought of as small packets of energy with no mass that travel at the speed of light. When considered this way, the waves are characterized by the energy of each photon.

Thermal processes

As mentioned above, by regular changes in the electric and magnetic fields, electromagnetic waves are generated and they transport energy from point to point. This can be either a *thermal* or a *non-thermal* process. For energy to be transported thermally, it is required that the body emitting the radiation has thermal energy (temperature). This includes all bodies with a temperature above the absolute zero. Common examples of *thermal* radiation are blackbody radiation, free-free emission ("*bremsstrahlung*") in an ionized gas and spectral line emission. A blackbody is a hypothetical object that absorbs one hundred percent of the energy that reaches it, and reflects nothing. When it reaches an equilibrium temperature, it starts radiating energy. The characteristic wavelength of this radiation is maintained by the temperature. Free-free emission comes from ionized gas. Ionization happens when electrons are removed from atoms. When charged particles later move around in this plasma, they cause acceleration of electrons and hence the gas cloud emits radiation. Spectral line emission involves the transition of electrons in atoms from a high energy level to a lower energy level. When this happens, a photon is emitted with the same energy as the energy difference between the two levels. The emission of this photon at a certain discrete energy shows up as a discrete "line" or wavelength in the electromagnetic spectrum.

Non-thermal processes

In conclusion, all objects that radiate *thermally* will send out end receive energy of all frequencies. Some will emit mostly at ultraviolet frequencies, while others will mostly emit at infrared and the decisive factor is the temperature. Not many *thermal* stellar objects can be observed by the radio waves they release, but the ones that do are the Sun and some stars covered by stellar winds. Radio frequencies are best discovered from objects that emit *non-thermal* radiation. This mechanism is related to the interaction of charged particles with magnetic fields. A charged particle entering a magnetic field will begin moving in a circular or spiral path, an since it then accelerates it will give off energy. When the speed of this particle is comparable to the speed of light, it will emit synchrotron radiation. Most commonly, these particles are electrons. The frequency of the radiation is directly related to how fast the particle is traveling. The longer the particle stays in the magnetic field, the more energy it loses. As a result, the particle makes a wider spiral around the magnetic field, and emits electromagnetic radiation at a longer wavelength. To maintain synchrotron radiation, a continuous supply of relativistic particles is necessary. Typically, these are supplied by very powerful energy sources such as supernova remnants, quasars, or other forms of active galactic nuclei.

Masers (micro-wave-amplified stimulated emission of radiation) display another type of non-thermal radiation. They can be compared to lasers (which amplify radiation at or near visible wavelengths). This interstellar medium contains a small number of certain molecules, which would normally be very hard to detect, but due to “masing”, these clouds can even be detected in different galaxies, because maser action will amplify faint emission lines at a specific frequency.

A third type of *non-thermal* radiation is gyrosynchrotron emission from pulsars. This process is actually just a special form of synchrotron emission. Pulsars are the result of the death of massive stars. As a massive star runs out of “fuel”, its core begins to collapse, the outer layers of the star fall down onto the core and a shock wave is produced that results in a supernova explosion. After the explosion, an extremely dense neutron star is left behind. A rapidly rotating neutron star is known as a pulsar. A typical pulsar has a magnetic field a trillion times stronger than the Earth's, which accelerates particles to nearly the speed of light, causing them to emit radiation, including radio waves. When this radiation reaches Earth, we see a "pulse" of radiation from the pulsar.

Radio Waves [1,6]

Radio waves generally obey all the same laws as other electromagnetic waves including reflection, refraction and diffraction. These three properties are especially important when it comes to understanding how radio telescopes function. The angle at which a radio wave is reflected from a dish will equal the angle at which it approached the surface. However when it passes through regions of varying pressure, temperature and water content, it will diffract and take different paths through the atmosphere. Propagation of radio waves in space is free, because they do not experience a strong influence of a medium, unfortunately the Earth’s atmosphere is a solid barrier to much of the electromagnetic radiation reaching us from space. It absorbs most of the wavelengths smaller than the ultraviolet, most of the wavelengths between the infrared and microwaves, and most of the longest radio waves. This only leaves visible light, some ultraviolet and infrared, and the shortest radio waves to penetrate the atmosphere and bring information about the universe. The radio window ranges from about 30MHz to over 300 GHz, which corresponds to wavelengths of almost 100 m down to 1 mm. In general, when radio frequencies travel through a gas, certain wavelengths are absorbed. In the case of traveling through our atmosphere, the low frequency end is limited by signal absorption in the ionosphere, while the upper limit is determined by signal reduction caused by water vapor and carbon dioxide in the atmosphere. The radio window at higher frequencies is rather unaffected by weather, but clouds and rain can cause signals to weaken (this is why telescopes designed for studying sub-millimeter wavelengths are often built high up in the mountains).

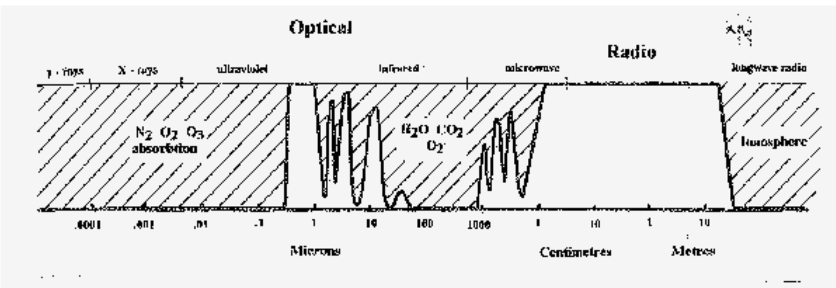


Figure 4 Energy flux of EM radiation arriving at the Earths surface. The radio window is limited by the ionosphere at wavelengths greater than a few meters and atmosphere absorption at wavelengths shorter than about 2 cm.

The refractive index for water vapor is almost twenty times greater at radio than optical wavelengths. This is due to the permanent dipole moment of water molecules. Refractivity N of the atmosphere at radio wavelengths and temperatures encountered in the atmosphere is given by:

$$N = 77.6\ T^{-1}(P_d + 4810 P_v T^{-1}) \qquad (14)$$

This is called the Smith and Weitraub formula, where P_d denotes the partial pressure of dry air and P_v the partial pressure of water vapour (both in millibars). When refraction of radio waves occurs in the atmosphere, the arrival of the wavefront at a radio telescope is affected in much the same way as for an optical telescope. The apparent increase Δz in elevation at zenith angle z is

$$\Delta z = (n-1)\tan z\,, \qquad (15)$$

n being the refractive index of air at radio wavelengths.

Radio waves are categorized into bands depending on the value of their frequency. Each band is ten times higher in frequency than the band immediately below it. Radio telescopes can be adjusted to tune into frequencies of one particular band. These are the bands in which astronomers use radio telescopes to observe the radio waves emitted by astronomical objects. The most common radio band names and their corresponding wavelengths/frequencies are:

Band	**Wavelength**	**Frequency**
P-band	90 cm	327 MHz
L-band	20 cm	1.4 GHz
C-band	6.0 cm	5.0 GHz
X-band	3.6 cm	8.5 GHz
U-band	2.0 cm	15 GHz
K-band	1.3 cm	23 GHz
Q-band	7 mm	45 GHz

Signal detection and noise [1-2]

According to the laws of quantum mechanics and the second law of thermodynamics the accuracy of all observations will be limited by the fundamental fluctuations that are generally known as noise. For the radio astronomer, the detected signal is the sum of both the cosmic signal and the noise generated by the environment, both with a Gaussian character. A *radiometer* is used to distinguish the astronomical signal from the interfering noise. The signal-to-noise ratio is limited by a system noise consisting of sky background noise and receiver noise, and it depends on the duration t of the observation and the bandwidth B of the detecting system. The estimate of the system noise gets better as the averaging time increases, but an uncertainty always remains. In this part we deal with the waveform and spectrum of noise, and we mention some principles of radiometers, which measure the noise power.

Gaussian noise

The detected signal is a superposition of an infinite assemblage of oscillators with random frequency and phase. The signal amplitude $V(t)$ is a stationary random variable, and the properties of the signal can only be described statistically. Gaussian random noise, for which the probability density function is a Gaussian function, is the most common form met with in practice. At each instant of time, the probability that the signal has an amplitude V is given by

$$p(V) = \frac{1}{s\sqrt{2p}} e^{-V^2/2s^2} \tag{16}$$

where s is the standard deviation of the probability density. The mean value of the signal amplitude is zero, and the noise power is the mean value of $V(t)^2$ which, for the given probability density, is s. Another interesting quantity is the autocorrelation function $R(t)$ of the signal amplitude, which is defined as

$$R_T(t) = \int_{-T/2}^{T/2} V(t)V(t+t)\,dt \tag{17}$$

where the subscript T indicates that the integration extends over an interval T.

The power spectrum of a random noise signal is the Fourier transform of the autocorrelation function. This also has a Gaussian character. In the real world, only an estimate of the signal power, or its autocorrelation function can be obtained by evaluation over a finite timespan T, so one arrives at an estimate of the power spectrum. This estimate becomes more precise as T increases. For a given time interval T the estimated power spectrum is the Fourier transform of $R_T(t)$:

$$S_T(n) = \int_{-T/2}^{T/2} R_T(t) e^{-iwt} dt \qquad (18)$$

which, in the limit of T going to infinity, becomes

$$S(n) = \int_{-\infty}^{\infty} R(t) e^{-iwt} dt \qquad (19)$$

The spectrum of Gaussian noise is evaluated easily by inspecting equation (3) for the cases of zero and non-zero time lag. When t is non-zero, the integral goes to zero since $V(t)$ is completely uncorrelated from one instant to the next. When $t = 0$, however, the integral tends towards the mean square deviation, $\langle s^2 \rangle$. For a sufficiently long integration time, therefore, the autocorrelation function is a d-function for a Gaussian random signal:

$$\lim_{T \to \infty} R(t) = s^2 d(t) \qquad (20)$$

The Fourier transform of the d-function is constant (unity), from which it follows that Gaussian noise has a flat power spectrum with a power spectral density which equals s^2. The flat power spectrum contains all frequencies and is therefore often called a white spectrum, from the colour analogy.

Band-limited noise

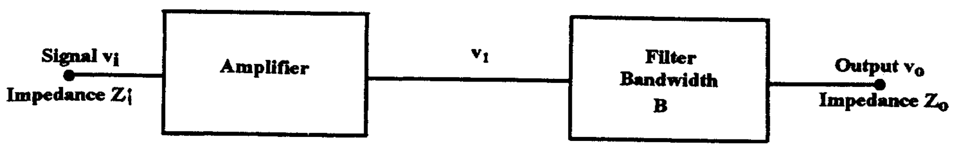


Figure 5 A linear receiver [1]

Equation (3) no longer holds when the signal has gone through a band-limiting filter of a radio receiver. A linear receiver, as illustrated in Figure 1, is a good receiver since it amplifies signals without distortion. The input signal amplitude, $V_i(t)$, is amplified, and the new signal amplitude, $V_1(t)$, goes through the filter, whose task is to modify the spectrum of the signal. Further amplification of the output signal gives an output amplitude $V_o(t)$. When the signal is passed through such a linear device, whose time response to a unit impulse $d(t)$ is $h(t)$, the output is the convolution of $V_i(t)$ and $h(t)$ (Figure 2):

$$V_o(t) = \int_{-\infty}^{\infty} V_i(t) h(t - t) dt \qquad (21)$$

The convolution theorem states that the Fourier transform of $V_o(t)$ is simply the product of the Fourier transforms of the two convolved functions:

$$V_o(n) = V_i(n) H(n) \tag{22}$$

From this it follows that the spectral distribution of the output is the product of the spectra of the input signal and of the device.

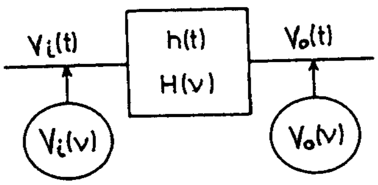


Figure 6 A schematic diagram to illustrate the analysis of noise in a linear system. The symbols above represent the time behaviour, those below the frequency behaviour [2]

When the impedance of the source of the input voltage $V_i(t)$ (a radio antenna or a signal generator) matches the input impedance of the amplifier, the maximum transfer of power from source to amplifier will occur. Matching impedances means that one is the complex conjugate of the other. The power gain G is generally given by

$$G = \frac{P_o}{P_i} = \frac{V_o^2}{V_i^2}\frac{Z_i}{Z_o} \tag{23}$$

where P_i and P_o are the powers of the input and output signals respectively, and Z_i and Z_o are the input and output impedances respectively. The usual unit of this quantity is the decibel (db), defined as ten times the $\log_{10}$ of the power ratio.

Detection and integration

Two examples of filtered white noise are shown in Figure 3. These are the outputs for a broad-band and a narrow-band filter respectively. As is easily seen, there is a greater correlation in time for the narrow-band case.

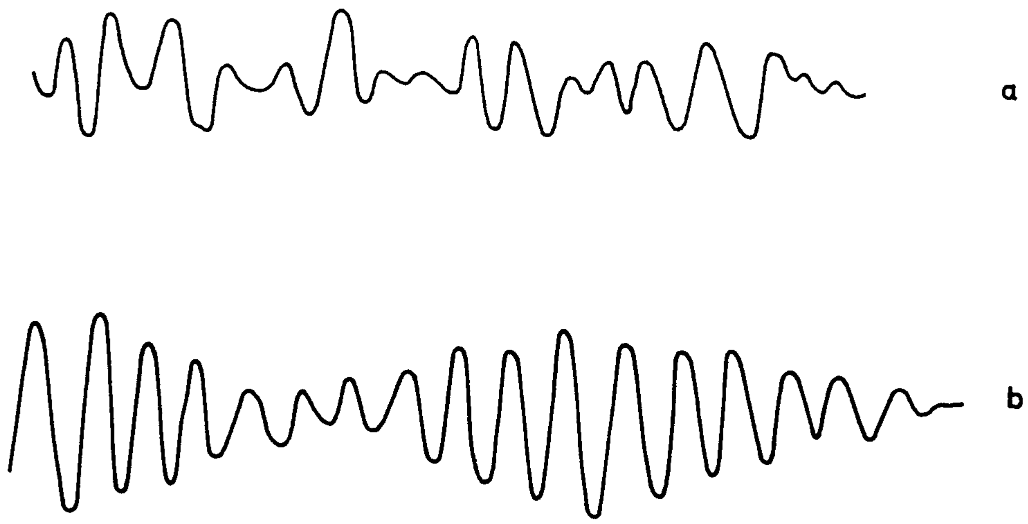


Figure 7 Filtered white noise: (a) broad-band and (b) narrow-band [1]

A radio receiver is a device that measures the spectral power density. The basic units of a *heterodyne receiver* are shown in Figure 4. A receiver of this type contains a *mixer* or *frequency converter* that shifts the initial frequency band (centred at n_0) up or down in frequency by letting a *local oscillator* inject a pure sinusoid at frequency n_1. The resulting *intermediate frequency* band is at either the sum or difference frequency.

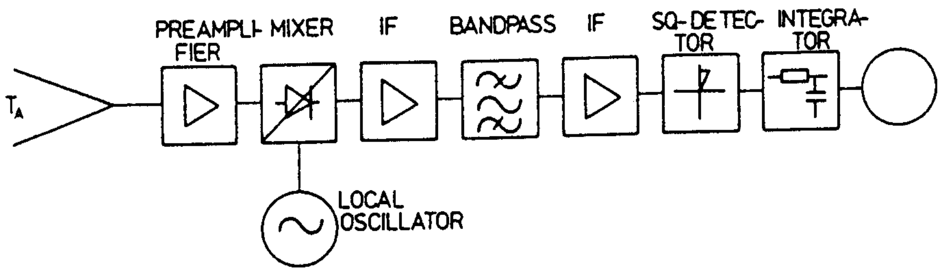


Figure 8 The principal parts of a heterodyne receiver [2]

As stated earlier, the noise power is proportional to the square of the amplitude. Therefore, to convert the output amplitude to power, the signal must be multiplied by itself. This is the task of the *square-law detector*. The output of the detector is still a fluctuating quantity, and so a time average of the power is taken, usually by an *integrator*. Integration of the signal power for a time t gives the average power readout $\langle P\rangle$, which is an estimate of the signal power across the band:

$$\langle P\rangle = \int_{-T/2}^{T/2} p_d(t)\,dt = \int_{-T/2}^{T/2} \left[V_o(t)\right]^2 dt \tag{24}$$

where $p_d(t)$ is the instantaneous power measured by the detector and equals the square of the output amplitude $V_o(t)$. The resulting power spectrum observed at the detector output is[1]

$$S_d(n) = R_o^2(0)d(n) + 2\int_{-\infty}^{\infty} S_o(n')H_o(n-n')\,dn' \tag{25}$$

The first term on the right-hand side would give the average power for a white noise signal, and the second term, the convolution, expresses the fact that the white noise signal has gone through a filter that induces correlation in the output signal.

There is an uncertainty in this final measurement. Since the power density is almost always measured with respect to the equivalent temperature T_{eq} of an input load, T_{eq} is a natural unit of power density. In the pre-detection (post-filtered signal) there is a correlation for a time of about $1/B$, from which it follows that there are about B independent measurements per unit time. If the detected signal is integrated for a time τ, there will be a total of about Bt independent measurements of the filtered noise signal. Since the input noise is random, the relative uncertainty, ΔT, in the measurement of the noise temperature, T_s, at the input of the detector, will be

$$\Delta T = \frac{T_s}{\sqrt{Bt}} \tag{26}$$

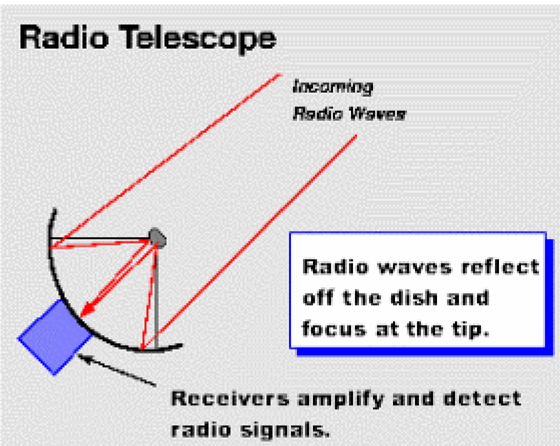
[1] A detailed treatment of the squaring process, and the resulting character of the detected signal can be found in [2].

Radio Telescopes [1, 4-6]

Radio telescopes are used to study radio emissions from stars, galaxies, quasars, and other astronomical objects. The results of these studies are presented as measurements of intensity and state of polarization as functions of frequency, angular position and time. Most telescopes are able to observe emissions at frequencies ranging from about 30MHz to 300GHz and they work best with wavelengths between 1 and 20cm. Since small structures can be built with greater precision than larger ones, radio telescopes designed for operation at millimeter wavelength are typically only a few tens of meters across, whereas those designed for operation at centimeter wavelengths range up to 100 meters in diameter.

The typical radio telescope consists of a radio receiver and an antenna system. The antenna or aperture collects radiation which is then transformed to an electric signal by a receiver, called the radiometer. This signal is then amplified, detected and integrated, and the output is registered on a recording device.

Radio telescopes are used as single operating apparatuses, in combinations of two, or arrangements of many can be used to study just one phenomenon. Some telescopes are also designed for studying milimetre and sub-milimeter wavelengths. Usually, the typical radio telescope will measure broad bandwidth continuum radiation, but also the study of spectroscopic features is common. Modern types observe simultaneously at a large number of frequencies by dividing the signals up into several thousand separate frequency channels that may range over a total bandwidth of tens to hundreds of megahertz.



The most common type of antenna is the parabolic reflector. It works by focusing the incoming radiation onto a small antenna called the feed. The feed is typically a waveguide horn connected to a radio receiver. In the simplest form, the receiver is placed directly at the focal point of the parabolic reflector, and the detected signal is carried through a cable along the feed support structure to a point near the ground where it is recorded and analyzed.

Figure 9 Radio waves reflect of the dish to the focus and are sent to a radio receiver. [7]

The receiver of a radio telescope should be as sensitive as possible because the incoming signals are so weak and often masked by noise both from space and noise generated by the telescope itself. For optimal sensitivity, amplifiers with very low internal noise are used, but sensitivity also depends on the area and efficiency of the antenna and the duration of the observation. The angular resolution of the telescope is mostly limited by its size, however the performance of a radio telescope can be limited by various other factors as well: the shape of the reflecting surface may not be ideal because of manufacturing faults, winds may have a powerful effect, there may be thermal deformations (expansion and contraction) and deflections due to changes in gravitational forces as the antenna is pointed to different parts of the sky. The largest effects are gravitational deformations, but these can be minimized by structural design, for example by allowing the reflector to deform. For instance, if the elevation of the telescope changes, the feed system can be moved to compensate for any change in axis and focal length.

Single aperture telescopes

The two basic components of a single aperture telescope are a large radio antenna and a sensitive radiometer. The radiometer is a device consisting of a radio receiver and a power measuring tool. Gathered intensity is dependant on several factors including man made interferences and measuring it consists of estimating the average value of all these quantities. The radiometer is the apparatus that does this. Distinguishing the astronomical noise signal from interfering noise is also the task of the radiometer. The noise signal, either from the antenna or dummy load passes to the receiver through a network that allows the radiometer to be calibrated. The receiver may incorporate a *mixer* or *frequency converter* that shifts the initial frequency band up or down in frequency, but otherwise preserves the signal. A *local oscillator* injects a pure sinusoid frequency and the resulting *intermediate frequency* band is either the sum or difference frequency. There are different types of receivers; the one that translates the input signal to an intermediate frequency like this is called a *heterodyne receiver*.

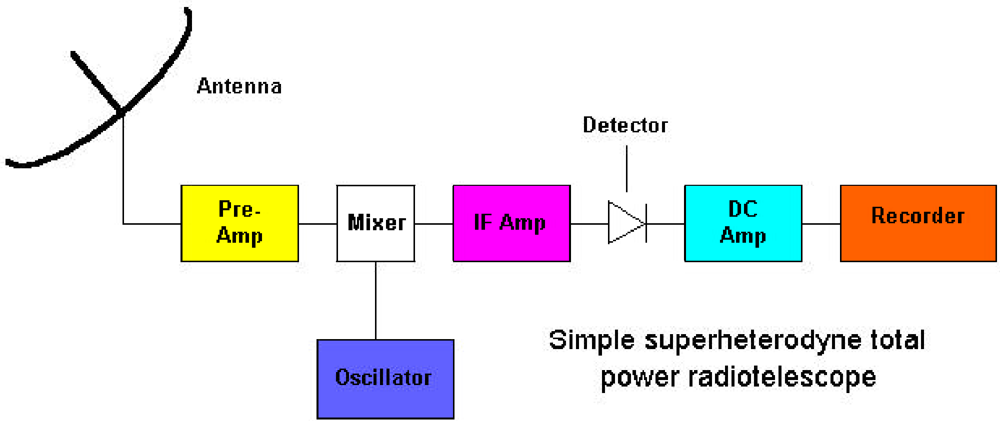


Figure 10 The basic components of a radiometer; the preamplifier, mixer, oscillator, IF amplifier, detector and a DC processor

The input signal $V_i(t)$ will first pass through a low noise preamplifier. This necessarily adds noise. After further amplification, the signal goes to a mixer where it is multiplied by a local oscillator signal. The product creates upper and lower sidebands and one of these is chosen as the intermediate frequency band. The output signal $V_o(t)$ passes to the detector which should at least have an output that has a known relation to the power of the input signal.

The average at time t_o lasting for integration time τ, is read out to the recorder. The random noise appearing at the output of an ideal radiometer will necessarily fluctuate with an rms uncertainty given by equation (26). The power density received by an antenna of effective area A_{eff}, observing an unpolarized source of flux S, will be SA/2. When this power is expressed in units of antenna temperature the limiting rms flux sensitivity for point source ΔS is

$$\Delta S = \frac{2kT_n}{A_{eff}\sqrt{Bt}} \tag{27}$$

One sees therefore that the quotient T_n/A_{eff} is a measure of the sensitivity of a radio telescope system.

The antenna is analogous to the lens of an optical telescope. It gathers the radiation, transmits it to an electrical current and after processing, the radiation is measured. In the case of radio telescopes, the antenna operates as a receiving device as opposed to a transmitting device. These two cases are actually equivalent because of time reversibility; solutions of Maxwell's equations are valid even when time is reversed. An antenna used for the purpose of receiving is considered for its receiving area, the so called *effective area* A_{eff}, an intercepted flux S and a yielded received power P_{rec} . *Effective area* is directionally dependent, and is a function of direction $\hat{n}$, measured with respect to the antenna axis so that

$$P_{rec} = A_{eff} S \tag{28}$$

An antennas *beamwidth* is the range of directions over which A_{eff} operates. From laws of diffraction; an antenna of size D will have a beamwidth of order λ/D. As a transmitter, that same antenna would have a power gain G(k) in a direction $\hat{n}$, as opposed to an *effective area*. The power gain is the ratio of the radiated power flux S(k) that would be measured at some large distance, to the power flux from a hypothetical isotropic radiator, measured at the same distance. For a transmitted power P_{tr}

$$S(\hat{n}) = \frac{G(\hat{n})P_{tr}}{4p\, r^2} \tag{29}$$

From the law of conservation of energy,

$$\int_{4p} G(\hat{n})\, d^2\Omega = 4p \tag{30}$$

The antenna will concentrate the radiation into a principal beam of solid angle Ω_0. For order of magnitude approximations this condition can be approximated by

$$G = 4p\ /\ \Omega_0 \tag{31}$$

And since the beam will have a width λ/D, A_{eff} will be proportional to the power gain:

$$A_{eff} = l^2 G / 4p \tag{32}$$

Imagine an antenna enclosed by a blackbody with temperature T. The antenna is connected to a transmission line, which is terminated by a matched load. This is a resistor having the same impedance as the line. Isolated from the outside world, the line and resistor will reach an equilibrium temperature T, identical to that of the blackbody. The same radio frequency power density P_ν for any narrow band $d\nu$ at frequency ν must flow in both directions along the transmission line. Planck's distribution describes the radio density u_ν

$$u_\nu d\nu = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/kT}-1}d\nu \tag{33}$$

In terms of specific intensity B_ν (flux per frequency interval per solid angle):

$$B_\nu d\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1}d\nu \tag{34}$$

Using Rayleigh-Jeans approximation, that frequency lies below the Planck maximum,

$$B_\nu d\nu = (2kT/\lambda^2)d\nu \tag{35}$$

The corresponding derivation for noise power flowing in a single-mode transmission line connected to a blackbody at temperature T leads to the one-dimensional analogue of Planck's law

$$P_\nu d\nu = \frac{h\nu}{e^{h\nu/kT}-1}d\nu \tag{36}$$

Again with Rayleigh-Jeans approximation this reduces to

$$P_\nu d\nu = kTd\nu\ , \tag{37}$$

showing that within a given narrow band, noise power is proportional to temperature.

To relate the telescope aperture to the size and shape of the beam, we start with considering the antenna as a transmitter, bearing in mind the reciprocity between antenna characteristics in reception and transmission. Imagine the aperture is a line distribution of excitation currents $i(\xi)$ at a single wavelength λ. At a large distance, in direction θ to the normal, the contribution of each element $i(\xi)d\xi$ to the radiation field depends on the phase introduced by the path $\xi \sin \theta$, the radiation pattern is $F(\theta)$:

$$F(\theta) = \int e^{(-i(2\pi\xi\theta/\lambda))} i(\xi)d\xi \quad , \tag{38}$$

where normalizing factors have been omitted and the approximation $\sin\theta=0$ has been made. ξ is measured in wavelengths and angles are measured as direction cosines l, m. The generalization to a two dimensional aperture, where the distribution of current density is referred to as grating, $g(\xi,\eta)$ gives:

$$F(q,f) = F_e(q,f)\int\int_{4p} g(x,h)e^{(-i2p(xq+hf))}dx\,dh\ , \tag{39}$$

where $F_e(\theta,\varphi)$ is the radiation pattern of a current element in the surface, taking due account of polarization. This equation sets out the basic relationship: *the radiation pattern is the Fourier transform of the aperture distribution.* A corresponding corollary yields: *the power gain pattern is the Fourier transform of the autocorrelation of the aperture current density distribution.* Fourier analysis actually provides a simple key to the theory of telescope beam shapes, and it is also important in understanding interferometers and aperture synthesis.

Narrowed aperture is common in radio astronomy because it implies that effective area in direction of maximum gain is less than the geometric area of the aperture. The ratio is the efficiency. Denoting effective are A(l,m) and sky brightness distribution B(l,m), the power output from sky is given by:

$$P_n = \int_{4p} B(l,m)\cdot A(l,m)d\Omega\,, \tag{40}$$

For a uniform temperature:

$$\int_{4p} A(l,m)d\Omega = l^2\ , \tag{41}$$

so that the all sky integral of the effective area is one square wavelength.

The angular resolution, or ability of a radio telescope to distinguish fine detail in the sky, depends on the wavelength of observations divided by the size of the instrument. Consider figure 11b. As the wavefront arrives at the telescope, all of its parts reach the focal point F in phase. At points all around F, they also arrive nearly in phase. A wavefront arriving at an angle less then λ/D to the plane of the dish will not be distinguished from the wavefront arriving on the axis of the dish. This angle is the angular resolution of the telescope.

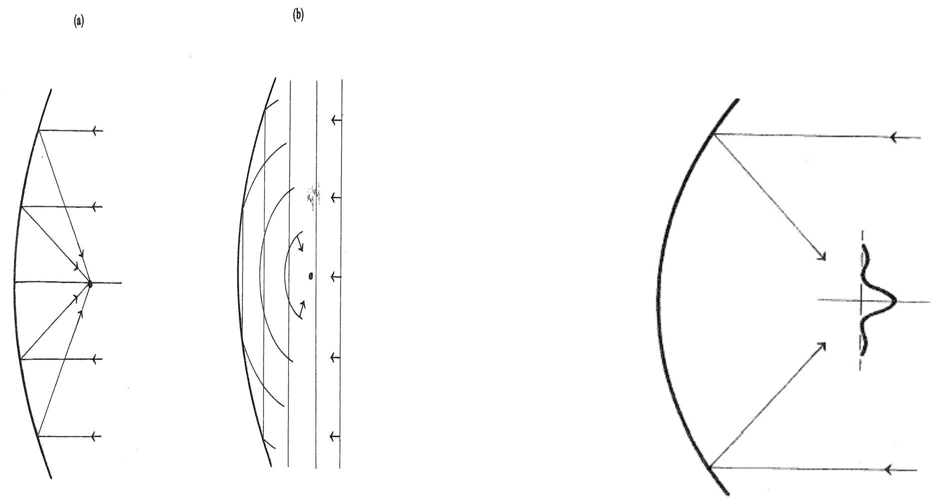


Figure 11.A paraboloidal reflector in terms of ray optics (a); (b) wavefronts (left) Wave amplitude at the focus. The width of the focal spot is approximately Fλ/D, where F is the focal length and D the diameter of the aperture [1]

To scan the sky with a single aperture telescope cannot reveal finer detail than an angular size of around λ/D, D being the diameter of the telescope. Antenna temperature T_a caused by a sky brightness distribution T_b(l,m) depends on the effective area A(l,m) of the telescope as a function of direction:

$$T_a = l^{-2}\int_{4p} T_b(l,m)A(l,m)d\Omega \tag{42}$$

Pointed in the direction l_0,m_0, the effective area is simply A(l - l_0,m - m_0). For small angles the projection factor can be neglected and the antenna temperature then becomes,

$$T_a = l^{-2}\int\int T_b(l,m)A(l-l_{0,}m-m_0)dldm \tag{43}$$

This is called the *smoothing function* of the sky brightness and the antenna beam.
The loss of detail in scanning the sky can now be expressed in terms of Fourier components, since the Fourier transform of T_a is the product of the Fourier transform of T_b(l,m) and A(l,m). These two are expressed as **t**(u,v) and **c**(u,v), functions of coordinate distances in the plane of the telescope aperture measured in wavelengths.

$$T_a(l,m) \leftrightarrow l^{-2}\left[\vec{t}(u,v)\cdot\vec{c}(u,v)\right] \tag{44}$$

$\vec{c}(u,v)$ is the telescope transfer function. Only non-zero components of $\vec{c}(u,v)$ are recorded.

As mentioned, the efficiency of a radio telescope is can be limited by many factors and imperfections. Some example are deformations and irregularities on the surface which can cause errors in phase across the aperture. These phase imperfection transfer some power from the main beam in to the sidelobes. This represents a loss of efficiency. For a random Gaussian distribution of phase error this loss is easily estimated. A portion of the wavefront with a small phase error of φ radians makes a reduced concentration to the power in the main beam by a fraction 1-φ^2, or more precisely $e^{-\varphi^2}$,and the contributions from the whole surface add randomly. The phase error at a reflector with a surface error ε is $4\pi\varepsilon/\lambda$ for normal reflection. Whole error is quoted as a single rms error ε, related to the surface efficiency η_{surf} by the Ruze formula:

$$\eta_{surf} = e^{-(4\pi\varepsilon/\lambda)^2} \tag{45}$$

Telescopes with a small angular resolutions are difficult to build, even the largest antennas, when used at their shortest operating wavelength, have an angular resolution only a little better than one arc minute, which is almost the same as that of the human eye at optical wavelengths. Because radio telescopes operate at much longer wavelengths than optical telescopes, they must be much larger than optical telescopes to achieve the same angular resolution. The angular resolution of a radio telescope used at given wavelength may be improved by extending the dimensions of the aperture by adding extra elements in the form of various types of interferometer.

Two-element interferometers

Interference is the net effect of the combinations of two or more wavefronts moving on coincident paths. The effect is that of the addition of the amplitudes of the individual waves at each point of intersection. This can cause constructive or destructive interference. Two element interferometers consist of two widely separated antennas connected by transmission lines. With their greatly increased resolving power, they can be used to determine the position or diameter of a radio source or to separate two closely spaced sources. As the Earth rotates first one and then the other will pass through antenna patterns. Their result is a superposition of the total power from the source, modulated by a fringe pattern. Near the maximum of the antenna response it is sufficiently close to sinusoidal form to be characterized by its frequency, amplitude and phase. In such interferometers, the correlation between signal amplitudes received by two antenna elements is measured, in contrast to total power systems associated with single apertures.

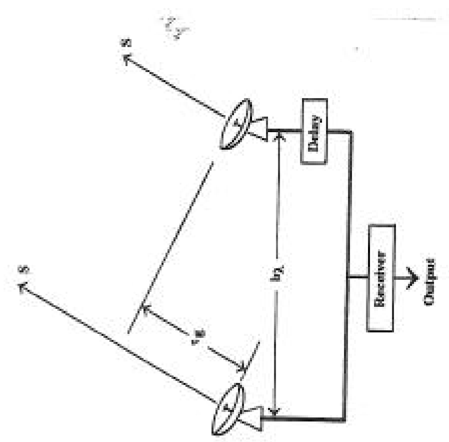


Figure 12 A two element interferometer. The path delay t g is compensated by the delay circuit in the receiver. [1]

Consider figure 6. The source is assumed to be monochromatic point source with frequency ν. The baseline vector **b** connects the phase centers of the two antennas; if these two are identical, any reference point will do. The radio source under observation is given by the unit vector **s**. The two antennas track the source as it moves and one of the antennas is designated the reference antenna. The signal arrives at the second antenna by geometric time delay τ_g:

$$t_g = \frac{\vec{b} \cdot \vec{s}}{c} \tag{46}$$

The two signals at frequency ν are fed to a voltage multiplier with an additional time delay τ_i, that can be inserted to equalize the signal delays. The cross correlation $R_{xy}(t)$ of two amplitudes (voltages) x(t) and y(t) is defined as the time averaged product of the two amplitudes, with one delayed by time τ:

$$R_{xy} \equiv \langle x(t)y(t-t)\rangle \tag{47}$$

This is called the *cross-power product.* The amplitudes of the antenna output terminals are $x(t) = v_1\cos 2\pi ft$ and $y(t) = v_2\cos 2\pi f(t-\tau_g)$. The rf voltage product, being the power received from the source will be proportional to the effective antenna area A(**s**) and the source flux S. Taking the time average, the cross correlation is:

$$R_{xy}(t_g) = A(\vec{s})S\cos 2p\, f t_g = A(\vec{s})S\cos 2p\, f\vec{b}\cdot\vec{s}\,/\,c \tag{48}$$

By measuring the baseline vector $\vec{b}$ in wavelengths, $\vec{b}$ can be designated by $\vec{b}_\lambda = \vec{b}/\lambda$, and the cross-correlation can be given by the expression:

$$R_{xy}(\vec{s}) = A(\vec{s})S\cos 2p\,\vec{b}_l \cdot \vec{s} \tag{49}$$

In either representation, the sinusoidal fringe variation is apparent. As the source direction changes, the fringe amplitude oscillates and for spacing of many wavelengths, when the source is close to transit, the variations is nearly sinusoidal since $\vec{b}_l \cdot \vec{s} \approx |\vec{b}_l| q$. In this approximation, with an angle θ between source and transit, and also making the assumption that the baseline is nearly perpendicular to the direction of observation:

$$R_{xy}(q) = A(s)S\cos 2p\,\vec{b}_l \cdot q \qquad (50)$$

The angle between fringes in this small angle limit is $1/|\vec{b}_l|$.

Extending the single frequency assumption to the case of finite bandwidth, leads to to realization that radio sources emit noise over a wide range of frequencies and all receiving systems have finite bandwidth. Radio noises are quasi-Gaussian, so the signal at one frequency is uncorrelated with other signals at adjacent frequencies. The radio spectra of continuum sources change slowly with frequency, and the interferometer has usually a small fractional bandwidth. The radio spectrum across the band is then effectively flat and this approximation simplifies the analysis. The effect of finite bandwidth is a diminished angular range over which fringes appear. Consider figure 13. If the source is normal to the interferometer baseline, constructive interference will occur at all frequencies. If displaced by a small amount from this direction, there will be phase differences across the band, and at long displacements, there may be destructive interference at one end, and constructive interferences at the other. As a result, the net fringe amplitude will be reduced and if error is large enough, fringes will disappear. This is known as the *delay beam* effect.

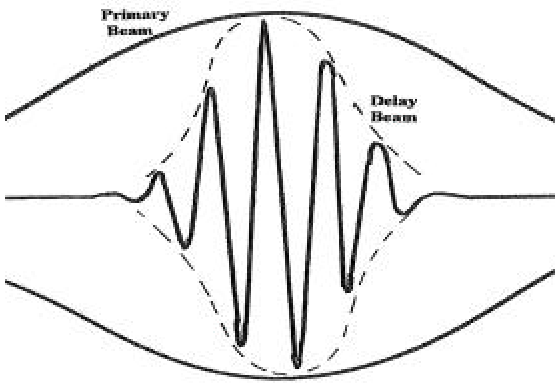


Figure 13 The delay beam effect in an interferometer. The width of the delay beam depends on the bandwidth of the receiver. [1]

Usually radio sources have finite angular sizes, and this has to be taken into account as well. The difference in path length to the elements of the interferometer will then vary across the source. Thus, the measured interference fringes from each interferometer pair depend on the radio brightness distribution in the sky. Each interferometer pair measures one Fourier component of the brightness distribution of the radio source. Movable antenna elements combined with the rotation of the Earth can collect enough Fourier components with which to synthesize the effect of a large opening and then reconstruct high-resolution images of the radio sky. The difficult computational task of doing Fourier transforms to obtain images from the interferometer data is accomplished with high-speed computers and the fast Fourier transform (FFT), a mathematical technique especially designed for computing discrete Fourier transforms.

Aperture Synthesis

It is possible to combine the data from two or more telescopes in such a way as to produce an image whose detail is equivalent to a telescope whose diameter was equal to the separation of the telescopes. Many antennas are linked electronically and the signals are recorded. A computer then takes the data and synthesizes a map with as high resolution as we would obtain if we were able to obtain a much larger dish.

Interferometer systems of almost unlimited element separation are formed by using the technique of very long baseline interferometry, or VLBI. In a VLBI system the signals received at each element are recorded by broad-bandwidth videotape recorders located at each element. The recorded tapes are then transported to a common location where they are replayed and the signals combined to form interference fringes. The successful operation of a VLBI system requires that the tape recordings be synchronized within a few millionths of a second and that the local oscillator reference signal is stable to more than one part in a trillion. Normally, the analysis of the two-element interferometer can be generalized to the case of N radio telescopes forming an aperture-synthesis array. For example, each pair of elements is combined as an interferometer, and the correlator will evaluate the cross power product R_{ij} for the two voltage amplitudes as in (47). The geometric time delay τ_g must be compensated by an instrumental time delay.

Conclusion

To fully understand how radio telescopes operate, complete knowledge of the theory on EM radiation and waves is required, as well as a deep insight into Fourier analysis. The latter is especially important when it comes to understanding how the telescopes detect and processes signals. Once the basics of the single aperture radio telescope are clear, it is easy to apply them to radio telescopes of many apertures. Finally, it is important to understand some basic concepts behind the electronics of a radio telescope so that one can know how to construct and use them to obtain the best results possible.

Appendix 1 Fourier transforms [1]

A1.1 Definitions

A harmonically oscillating amplitude $a(t)$ can be represented by the complex quantity

$$a(t) = a_0 e^{(i2\pi\nu t)} \tag{A.1}$$

The complex modulus a_0 gives the absolute value of the amplitude $| a_\circ |$ and the phase offset ϕ:

$$a_\circ = | a_\circ | e^{i\phi t} \tag{A.2}$$

The *phasor* $a_\circ$ rotates with angular frequency $\omega = 2\pi\nu$, and its projection on the real axis gives the physical amplitude $a(t)$. The linear addition of various components of a quantity such as the electric field in a radio wave is conveniently done by adding the phasors as vectors. Adding two components with identical phasors but with opposite signs of ω gives the real quantity directly.

The amplitudes a_ν associated with components at various frequencies ν can be superposed to give general amplitudes that can represent any physically realizable form $f(t)$. Fourier theory relates the function $f(t)$ to the spectral distribution function $F(\nu)$ of these components by the dual transformation theorem

$$\text{if} \quad f(t) = \int_{-\infty}^{\infty} F(\nu) e^{i2\pi\nu t} d\nu, \tag{A.3}$$

$$\text{then} \quad F(\nu) = \int_{-\infty}^{\infty} f(t) e^{(-i2\pi\nu t)} dt \tag{A.4}$$

The *spectral distribution function* $F(\nu)$ is the *Fourier transform* of $f(t)$. The frequency ν and the time t are *Fourier dual* coordinates. Another example of such dual coordinates is the relation between the radiation pattern of an antenna and the current distribution in its aperture (Chapter 4). The reciprocal relation between such pairs of functions may be written $f(t) \rightleftharpoons F(\nu)$. Figure A.1 shows some particularly useful examples.

The Dirac impulse function $\delta(t)$ has an especially simple Fourier transform. Since by definition it is zero everywhere except at the origin, but has an integral of unity, its transform is

$$F(\nu) = \int_{-\infty}^{\infty} \delta(t) e^{(i2\pi\nu t)} dt = 1 \tag{A.5}$$

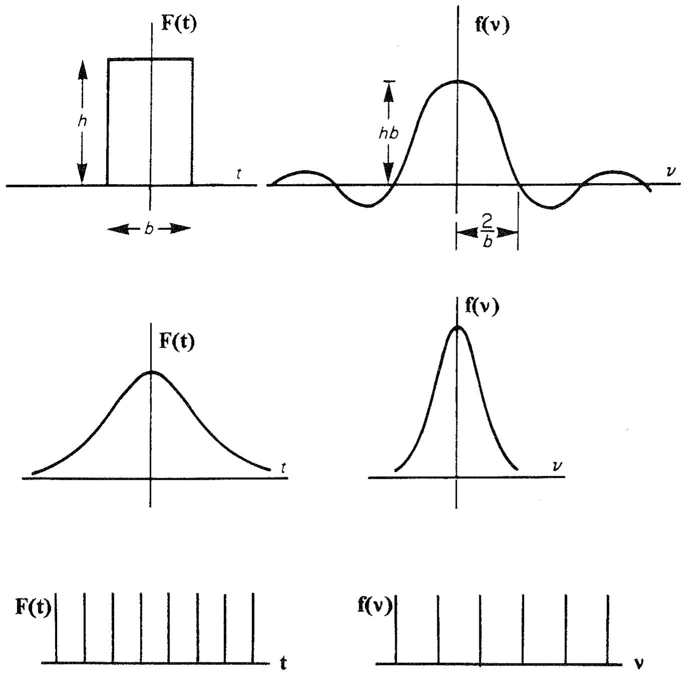


Fig. A1.1. Useful examples of Fourier transforms.

Thus, a unit impulse has a flat spectrum of amplitude unity. If such an impulse occurs at time T other than zero, the spectral distribution still has all frequencies present with unit amplitude, but with a frequency-dependent phase $2\pi\nu T$.

The square step, or gating function, with unit amplitude from $-T/2$ to $T/2$ and zero elsewhere, transforms into the 'sinc' function, also known as the sampling function

$$F(\nu) = T\sin(\pi\nu T)/\pi\nu \equiv T\mathrm{Sinc}(\pi\nu T) \tag{A.6}$$

The Gaussian function has the interesting property that its Fourier transform is also a Gaussian:

$$f(t) = \mathrm{e}^{-(t/T)^2} \rightleftharpoons \mathrm{e}^{-(\pi\nu T)^2}/(2\pi T) = F(\nu) \tag{A.7}$$

Note that when $T = 1/\pi^{1/2}$ the Gaussian is its own transform. Another self-similar pair is the periodic sampling function III(t), named after the cyrillic character shah by Bracewell and Roberts (1954).

The Gaussian function illustrates a general principle: when the characteristic time T of a function is short, its spectrum is wide, and vice versa. Similarly, if a sinusoidal function is switched on and off at times $-T/2$ and $T/2$, its spectrum has a significant value over a bandwidth $\Delta\nu$ that is of order $1/T$.

Modulated waves, familiar to radio engineers, are encountered throughout physics. Typically, the amplitude of a cosinusoidal wave varies with time (Figure A.2), either periodically as in the beat pattern (a), or aperiodically as in an isolated group of waves.

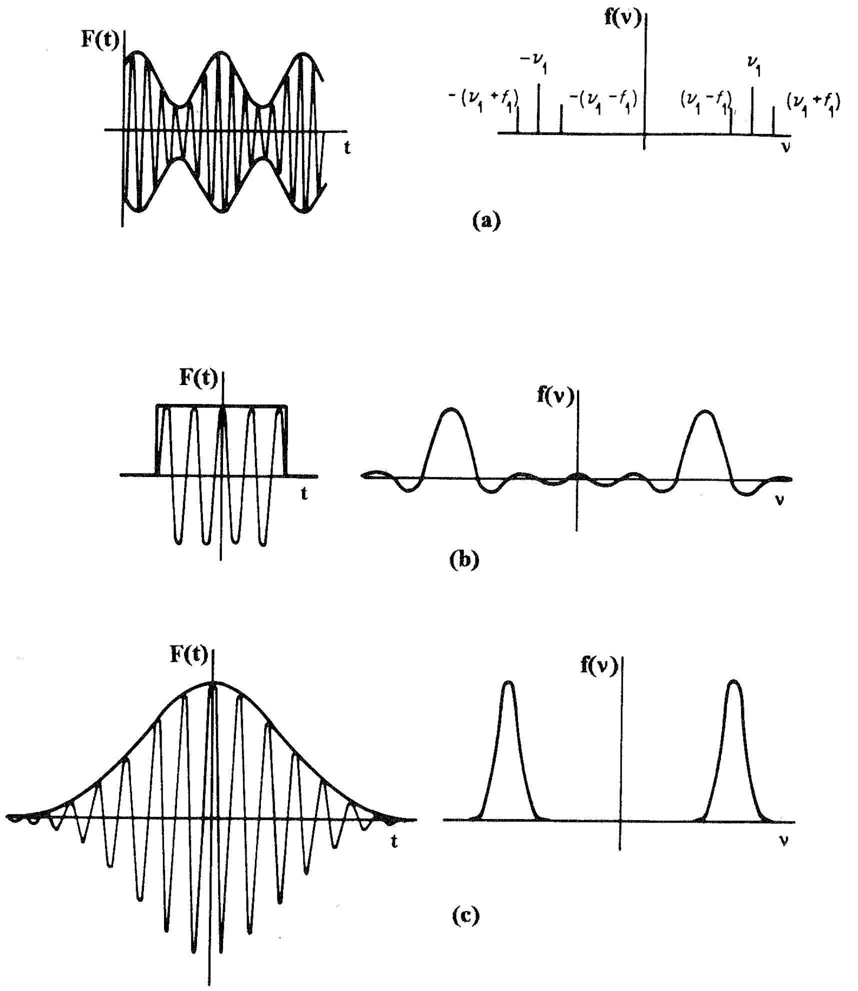


Fig. A1.2. Fourier transforms of modulated waves.

If the modulating function is $g(t)$, so that the wave is $g(t)\cos(2\pi\nu_1 t)$, then its spectrum is

$$\begin{aligned} F(\nu) &= \int g(t)\cos(2\pi\nu_1 t)e^{i2\pi\nu t}dt \\ &= \int g(t)\tfrac{1}{2}\{e^{i2\pi(\nu-\nu_1)t} + e^{i2\pi(\nu+\nu_1)t}dt \\ &= \tfrac{1}{2}G(\nu-\nu_1) + \tfrac{1}{2}G(\nu+\nu_1) \end{aligned} \tag{A.8}$$

where G(ν) is the transform of the modulating function $g(t)$.

A cosine modulation, as in Figure A.2(a), has a spectrum with 'sidebands' on either side of the components at $\pm\nu_1$. The top-hat modulation of (b) broadens the line components into sine functions covering a frequency band inversely proportional to the length of the wave train. The Gaussian modulation of (c) has Gaussian spectral line components; this is the spectrum of a wave group, such as the spectrum of matter waves associated with a single particle.

There are alternative definitions of the Fourier transform. One form that is met with frequently uses the angular frequency convention

$$F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{A.9}$$

$$f(t) = \int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega \tag{A.10}$$

The ν-convention has the advantage of symmetry, while the ω-convention is more compact.

The evaluation of a Fourier transform must frequently be done numerically. One can see that if a set of N discrete values $f_i(t_i)$ is chosen, the numerical evaluation of the Fourier transform *A*.4 over the N measurement points would appear to require N^2 multiplication operations to compute N values of $F_i(\nu_i)$. Fortunately this is not the case, since there exist a number of *Fast Fourier Transform* (FFT) algorithms that reduce the number of multiplication operations to the order of $N \ln N$. One of the most commonly used, the 'Cooley-Tukey' algorithm (previously developed by a number of authors, with Gauss as a forerunner) requires that the sample points be uniformly spaced; since in many practical situations, such as the aperture-synthesis technique (Chapter 6), the original data points are usually unequally spaced, an interpolation procedure is needed to obtain a regular grid of values. There are other FFT algorithms that do not require gridding, usually at the cost of greater demands on memory space.

A1.2 Convolution and cross-correlation

The *convolution theorem* plays an important role in practical applications of Fourier methods. For a pair of functions $g(t)$ and $h(t)$ with Fourier transforms $G(\nu)$ and $H(\nu)$, the convolution is defined as

$$f(t) \equiv \int_{-\infty}^{\infty} g(t')h(t-t')dt' \tag{A.11}$$

A convenient notation is $f(t) = g * h$, where the asterisk implies the convolution process. The convolution theorem states that the Fourier transform of $f(t)$ is simply the product of the Fourier transforms of the two convolved functions:

$$F(\nu) = G(\nu)H(\nu) \tag{A.12}$$

The theorem is easily proven and has many practical applications. For example, a signal $g(t)$ may be an input signal to an amplifier or other linear device that has a response $h(t)$ to an impulsive input; the spectral distribution of the output is then the product of the spectra of the input signal and of the device. When a device is being considered in the frequency domain, $H(\nu)$ is known as the *transfer function.*

The convolution theorem may be used to facilitate the rapid calculation of Fourier transforms in many instances. For example, if one starts with the gating function $\sqcap(t/T)$ and its dual sinc νT, it follows immediately that its autocorrelation function,

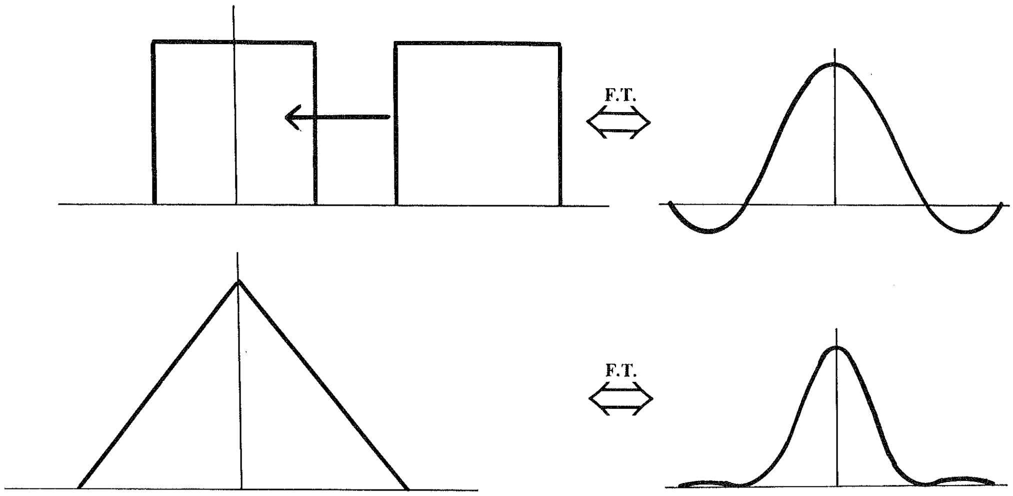


Fig. A1.3. Autocorrelation of the gating function.

the unit triangle of base $2T$ illustrated in Figure A1.3, must have the Fourier transform $\mathrm{sinc}^2 \nu T$.

The *cross-correlation* of a pair of functions bears a close relation to convolution. For a pair of functions $f(t)$ and $g(t)$ the operation is often designated $f \otimes g$, with

$$f \otimes g \equiv \int f(t')g(t'-t)\mathrm{d}t' \qquad \text{(A.13)}$$

and thus

$$f \otimes g = f(t) * g(-t) \qquad \text{(A.14)}$$

from which it follows that

$$f \otimes g \rightleftharpoons F(\nu)G^*(\nu) \qquad \text{(A.15)}$$

Note the time reversal in Equation (A.14); cross-correlation is not a simple commutative operation.

Cross-correlation has a particular significance in radio interferometry, where two signals $f(t)$ and $g(t)$ may be obtained from spaced antennas. The product fg^*, which has the dimensions of power, is then referred to as the *cross-power*, and the product FG^* is known as the *cross-spectral power density*.

The *autocorrelation function* $C(t)$, defined by

$$C(t) \equiv \int_{-\infty}^{\infty} f(t')f(t'+t)\mathrm{d}t' \qquad \text{(A.16)}$$

has a particular relevance in signals and systems, whose behaviour is often described in terms of *amplitudes* $e(t)$ (e.g. voltages, field strengths and currents), while the measurement process involves the *power* $p(t)$. Using a convenient choice of units we can write

$$p(t) = e^2(t) \tag{A.17}$$

The spectrum of a signal can be represented in Fourier terminology; the *spectral power density* $S(\nu)$ gives the power density in a given infinitesimal bandwidth $d\nu$. The total energy E_{em} emitted over time is then

$$E_{em} = \int e^2(t)dt = \int_0^\infty S(\nu)d(\nu) \tag{A.18}$$

Note that the distinction between positive and negative frequencies is no longer meaningful.

$S(\nu)$ is evidently proportional to $E(\nu)E^*(\nu)$, the product of the Fourier transforms of $e(t)$ and $e(-t)$. We explore this relation by writing the self-convolution $e(t') * e(-t')$, so that the emitted energy E_{em} may be expressed as

$$E_{em} = \langle e(t) * e(-t) \rangle_{t=0} = \int_{-\infty}^{\infty} e(t')e(t' - t)dt' \tag{A.19}$$

It follows from the definition of convolution (Equation (A.11)) that

$$\int_{-\infty}^{\infty} E(\nu)E^*(\nu)e^{-i2\pi\nu t}d\nu = e(t) * e(-t) \tag{A.20}$$

so that for $t = 0$

$$E_{em} = \int_{-\infty}^{\infty} e^2(t)dt = \int_{-\infty}^{\infty} | E(\nu) |^2 d\nu \tag{A.21}$$

This is known as *Rayleigh's theorem*, and is a generalization of the Parseval theorem for Fourier series. The integral over frequencies needs be taken over positive frequencies only, since $| E(\nu) |$ is symmetric in ν, and so, for this physical case,

$$S(\nu) = 2E(\nu)E^*(\nu) = 2 | E(\nu) |^2 \tag{A.22}$$

The self-convolution $e(t') * e(-t')$ is symmetric in the time offset t, and when written with the reverse sign becomes the autocorrelation function $C(t)$ as in (A.14) above, that is,

$$e \otimes e \equiv C(t) \tag{A.23}$$

The autocorrelation function has a most important property, known as the Wiener–Khinchin Theorem, which states that the Fourier transform of the amplitude autocorrelation is the power spectral density (with correction by a factor of 2 if only positive frequencies are treated, as in Equation (A.21)):

$$E(\nu) \otimes E^*(\nu) \rightleftharpoons C(t) \tag{A.24}$$

This relation provides the basis for autocorrelation spectrometry. In radio astronomy the autocorrelation function is obtained from the product of a digitized signal and the same signal subjected to a variable delay; in optical spectrometers light reaches a detector via two paths of variable path difference. In practice, many different delays may be used simultaneously, especially in radio astronomy, where the signal may be amplified coherently and split into many channels.

A1.3 Two or more dimensions

The definitions and examples in this appendix have all been presented in the one-dimensional case, but since the Fourier transform is a linear operation it can be generalized to Cartesian coordinates in many dimensions. Given a function $f(\mathbf{x})$ of the n-dimensional variable $\mathbf{x}$, it will have a Fourier transform $F(\mathbf{k})$. The Fourier dual coordinate $\mathbf{k}$ is called the *spatial frequency* by analogy with t and ν in the time/frequency case. A straightforward calculation shows that the fundamental Fourier inversion theorem, Equation (A.3), becomes

$$f(\mathbf{x}) = \int_{-\infty}^{\infty} f(\mathbf{k}) \mathrm{e}^{\mathrm{i}2\pi \mathbf{k}\cdot\mathbf{x}} \mathrm{d}^{\mathrm{n}}\mathbf{k} \tag{A.25}$$

$$F(\mathbf{k}) = \int_{-\infty}^{\infty} f(\mathbf{x}) \mathrm{e}^{-\mathrm{i}2\pi \mathbf{k}\cdot\mathbf{x}} \mathrm{d}^{\mathrm{n}}\mathbf{x}, \tag{A.26}$$

thus defining

$$f(\mathbf{x}) \rightleftharpoons F(\mathbf{k}). \tag{A.27}$$

References

[1] Burke B. F. and Graham-Smith F., An Introduction to Radio Astronomy, Cambridge University Press, 1997, chapters 1-7 and appendices 1 and 3

[2] Rohlfs K. and Wilson T. L., Tools of Radio Astronomy, 4th ed., Springer, 2004, chapter 4

[3] Freedman R. A. and Kaufmann W. J., Universe, 6th ed., Freeman and Company, 2002, chapter 6

[4] Griffiths David J., Introduction to electrodynamics, 3rd ed., Prentice Hall 1999, Chapters 9 and 11

[5] http://www.radiosky.com/simple.html

[6] http://www2.jpl.nasa.gov/radioastronomy/index.htm

[7] http://www.nrao.edu/

[8] http://www.lbl.gov/MicroWorlds/ALSTool/EMSpec/EMSpec2.html