Millimetre-Wave Astronomy

Study Notes

John Garrett

Contents

2 Radio Astronomy

2.1 Introduction

Radio astronomy is the study of celestial objects in the radio-frequency (RF) spectrum, which is traditionally defined as 3 kHz to 300 GHz. Practically however, ground-based radio astronomy extends from ∼10 MHz, below which the ionosphere blocks all radiation, to ∼1.5 THz where H_2O and N_2 molecules begin to absorb RF radiation [?].

Radio and optical astronomy differ in many regards, but the fundamental difference is the frequency of the light being observed. Through the **Planck–Einstein relation**, $E = h\nu$, this means that there is also a difference in the photon energy. For example, a photon at 400 MHz has less than 1/1,000,000$^{\mathrm{th}}$ of the energy of a red, visible-light photon. In optical astronomy, photons are typically used to generate charge through the photoelectric effect. In radio astronomy, the photons no longer have enough energy to overcome the binding energy in semiconductors. Instead, the radiation needs to be coupled to an electric circuit where the signal can be amplified, down-converted, filtered, and sampled. We therefore think of the radiation as EM waves instead of photons, as in the case of optical astronomy.

Despite radio astronomy merely being a different subset of the electromagnetic spectrum from optical astronomy, there are many differences in the physical properties that are observed and the technology that is used to make the observations. Many physical phenomena are only able to be understood in the radio spectrum, therefore radio astronomy is necessary for a comprehensive study of astrophysics.

2.2 Signals in Radio Astronomy

In traditional radio communication, information signals are modulated with carrier signals to shift the spectral content of the information to such a frequency that it can be transmitted wirelessly. The result is a strong, coherent signal where the signal at time $t + \Delta t$ is strongly correlated to the signal at time t. An example is amplitude modulated (AM) radio where the audio signal (20 Hz to 5 kHz) is modulated with a carrier (single tone between 500–1,700 kHz) so that it can be transmitted over the air. The result is a sine-wave carrier with a time-varying envelope that contains audio information (Figure 2.1).

The radio emission from celestial objects however is **incoherent**, meaning that the phase at time $t + \Delta t$ is unrelated to the phase at t. **Celestial signals therefore appear as random noise** (Figure 2.1) and they are indistinguishable from the thermal noise of a resistor[1]. From the source's measured power spectrum (P_ν, in SI units W Hz^{-1}), we can then define T_{source} as the noise temperature of the source:

$$\boxed{T_{source} \equiv \frac{P_{\nu,source}}{k}} \tag{2.1}$$

T_{source} will be indistinguishable from all other sources of thermal noise including thermal emission from the ground, atmosphere and telescope optics, and the thermal noise of the receiver circuit itself. The overall signal that is measured (T_{meas}) will then be the combination of all of these noise sources (in SI units K):

$$\boxed{\begin{aligned} T_{meas} &= T_{source} + T_{gnd} + T_{atm} + T_{receiver} + T_{cmb} + ... = \frac{P_{\nu,meas}}{k} \\ &= T_{source} + T_{sys} \end{aligned}} \tag{2.2}$$

[1] Also called Johnson, Nyquist, or Johnson-Nyquist noise.

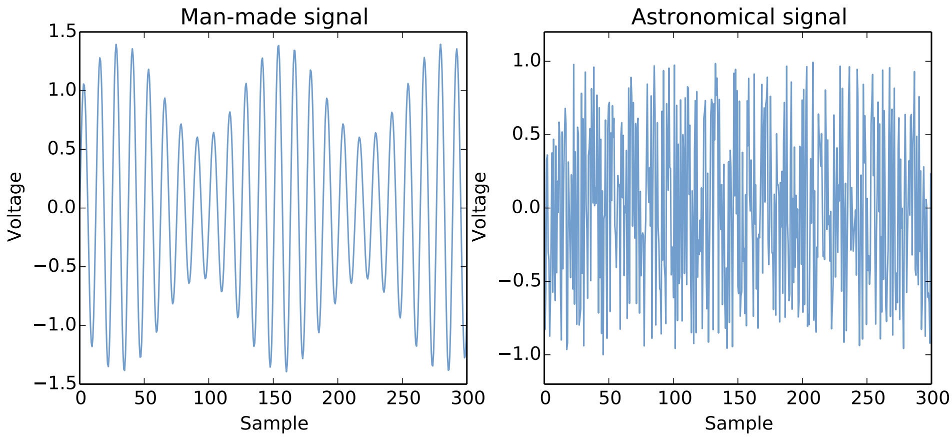


Figure 2.1: The difference between man-made and astronomical signals.

In Equation 2.2, T_{sys} represents the noise temperature of the entire system. Since $T_{source} + T_{sys}$ is the combination of many different noise sources, through the **central limit theorem** we know that this noise will have a **Gaussian distribution**. Note that T_{source} is typically several orders of magnitude weaker than the overall value for T_{meas}.

2.2.1 Measuring Thermal Noise

Thermal noise will have an average voltage of 0 (i.e., $\langle v \rangle = 0$), but a non-zero RMS value (v_{rms}) and therefore non-zero power. To measure the average noise power, a **radiometer** is used. The simplest radiometer consists of:

1. an **antenna** to couple freespace EM radiation to the electric circuit,
2. a **bandpass filter** to select a subset of the frequency spectrum ($\nu_{RF} - \Delta\nu/2$ to $\nu_{RF} + \Delta\nu/2$),
   - $\Delta\nu$ is the instantaneous bandwidth (the range of frequencies that will be measured simultaneously). In a heterodyne receiver, this is called the intermediate-frequency (IF) bandwidth.
   - After the bandpass filter, the signal will no longer look random. It will appear as a signal modulated at ν_{RF} with an envelope that varies on timescales of $\sim \Delta\nu^{-1}$.
3. a **square-law detector**[2] to generate a signal that is proportional to the input noise power,
   - $v_{out} \propto v_{in}^2 \ \therefore\ v_{out} \propto P_{in},\ \langle v_{out} \rangle \neq 0$
   - the frequency of the oscillating component $\sim 2\nu_{RF}$
4. an **integrator** to smooth out any fluctuations in v_{out}, and finally
   - Integrate for $\Delta t \gg (\Delta\nu)^{-1}$ (a low pass filter).
5. a **voltometer** to measure v_{out}.

After integrating for time Δt at the Nyquist rate[3], there will be $N = \Delta t / T_{sample} = 2\Delta\nu\Delta t$ independent samples of T_{sys}, each with $\sigma_T \approx \sqrt{2}T_{sys}$ RMS error. This error will reduce as more samples are gathered and averaged together:

$$\sigma_T = \frac{\sqrt{2}T_{sys}}{\sqrt{N}} \tag{2.3}$$

By plugging in the value for N,

[2]The simplest square-law detector is a diode.
[3]Nyquist sampling theorem: To fully sample a signal, samples must be taken at double the signal's frequency bandwidth. Note that it is double the instantaneous bandwidth ($\Delta\nu$), not just the highest frequency contained within the signal.

$$\boxed{\sigma_T \approx \frac{T_{sys}}{\sqrt{\Delta\nu \cdot \Delta t}}} \tag{2.4}$$

the **Dicke radiometer equation** is found. T_{source} should be $\gtrsim 5 \cdot \sigma_T$ for a well-defined detection (i.e., the signal-to-noise ratio should be > 5).

Note that Equation 2.4 implies that an arbitrarily large S/N can be achieved. In reality, systematic errors limit the ability to integrate over long periods of time. For example, if **gain instability** is included, equation 2.4 becomes:

$$\sigma_T \approx T_{sys} \left[\frac{1}{\Delta\nu \cdot \Delta t} + \left(\frac{\Delta G}{G} \right)^2 \right]^{1/2} \tag{2.5}$$

which puts a lower bound on σ_T. **Atmospheric fluctuations** will also cause a similar effect.

Both of these can be corrected through **Dicke switching**. This involves switching between the source and an empty patch of sky a few beam widths away (so that the beam is going through roughly the same atmosphere). Then when integrating, the power is multiplied by $+1$ when the beam is on the source, and by -1 when it is on the empty sky. The average power should then remain constant. Since the beam is only on the source for half the time however, σ_T will be twice as large.

2.2.2 Measuring the System Noise Temperature

The noise temperature of receiver systems is typically characterised by measuring the system's response to hot and cold loads, at T_H and T_C, respectively. The system's output power from these loads (P_H and P_C, respectively) will be proportional to the input temperature (with a zero offset due to noise) provided that the load temperatures are within the Rayleigh-Jeans approximation ($h\nu/kT$). The noise temperature of the system is then:

$$T_{sys} = \frac{T_H - T_C\, Y}{Y - 1} \tag{2.6}$$

where $Y = P_H/P_C$ and is often called the Y-factor.

2.3 Black Body Radiation

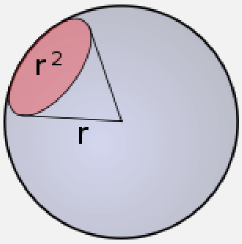
Planck's law gives the spectral density of electromagnetic radiation emitted by an ideal black body in thermodynamic equilibrium:

$$B_\nu\left(T\right) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kT}} - 1} \tag{2.7}$$

$$B_\lambda\left(T\right) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{k\lambda T}} - 1} \tag{2.8}$$

Aside

Note that I_ν and B_ν are used interchangeably for spectral radiance (also known as specific intensity, spectral intensity, spectral brightness, or just generally brightness), which is the energy per unit time per unit area per unit frequency per unit solid angle. In SI units, this is J $s^{-1}m^{-2}Hz^{-1}sr^{-1}$. A steradian (sr) is the SI unit for solid angle. One steradian is defined as:



An entire sphere is 4π steradians or roughly 12.5 sr.

The peak of the blackbody spectrum is given by **Wien's displacement law**,

$$\frac{\nu_{\max}}{\mathrm{GHz}} = 58.789\left(\frac{T}{K}\right). \tag{2.9}$$

For low frequencies (such as those in radio astronomy), $h\nu/kT \ll 1$. Then using

$$\frac{1}{e^{\frac{h\nu}{kT}} - 1} \approx \frac{kT}{h\nu}, \tag{2.10}$$

Equation 2.7 reduces to

$$\boxed{B_\nu(T) = \frac{2\nu^2}{c^2}kT} \tag{2.11}$$

which is **Rayleigh-Jeans law**. The difference between Planck's law and the Rayleigh-Jeans approximation becomes apparent as ν approaches the peak of the black body radiation curve (Figure 2.2).

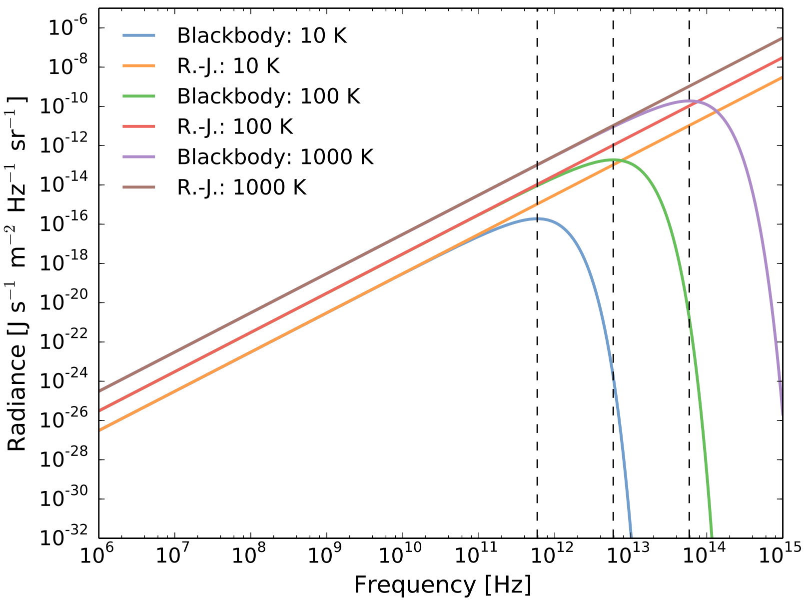


Figure 2.2: Black body spectral radiance for 10, 100, and 1000 K. The dashed lines mark the peaks of the black body curves.

Aside

In radio astronomy, we often pretend that sources are thermal (i.e., $I \propto \nu^2$) even when they're not. By rearranging Equation 2.11, we can define brightness temperature (T_b):

$$T_b(\nu) \equiv \frac{c^2}{2k\nu^2}\frac{1}{\nu^2}I_\nu = \frac{\lambda^2}{2k}I_\nu \tag{2.12}$$

If the signal is thermal and within the Rayleigh-Jeans limit, $T_b(\nu)$ will be constant. This will not be true for other emission (e.g., free-free, synchrotron).

Integrating the spectral radiance over all frequencies and solid angles gives the power per unit area (i.e., the radiant emittance), also known as **Stefan-Boltzman's law**:

$$j^* = P/A = \sigma T^4, \quad \sigma = \frac{2\pi^4 k^4}{15c^2h^3} \tag{2.13}$$

$$\therefore\ P \propto T^4 \tag{2.14}$$

In the case of a star (or any spherical object with radius R), j^* can be integrated over the surface to determine the total bolometric power (energy per unit time):

$$L_{bol} = 4\pi R^2 \sigma T^4 \tag{2.15}$$

2.4 Measuring Black Body Radiation

To figure out how much energy will be measured with a telescope, we can integrate the spectral radiance over solid angle, frequency and area:

$$P_{measured} = \iiint I_\nu \; d\Omega \; dA \; d\nu \tag{2.16}$$

It is important to note that the **spectral radiance (I_ν) is conserved (is constant) along any line of sight in empty space.** Therefore, I_ν is independent of distance, and it is the same at the source as it is at the detector. Spectral radiance can be thought of as energy flowing out of the source, or energy flowing into the detector [?] as shown in Figure 2.3.

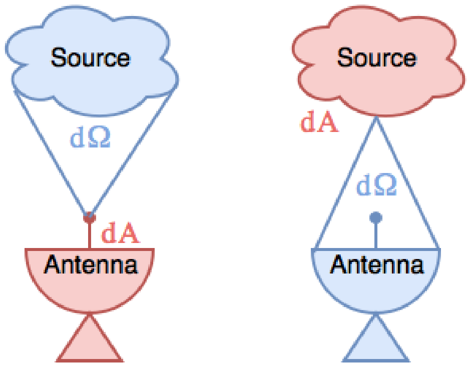


Figure 2.3: Two ways to think about spectral radiance.

If a source is discrete (meaning that it occupies a defined solid angle), the **flux density** (S_ν, in units W m^{-2} Hz^{-1}) at the telescope can be found:

$$\boxed{S_\nu = \int_{source} I_\nu \; d\Omega} \tag{2.17}$$

Aside

The SI unit for flux density [W m^{-2} Hz^{-1}] is far too large for any astronomical source. Instead, Janskys are used where 1 Jy = 10^{-26} W m^{-2} Hz^{-1}. Even then milli-Janskys are often needed!

The **flux**[4] can also be found (F, in SI units W m^{-2}):

$$F = \int I_\nu \; d\Omega \; d\nu = \int S_\nu \; d\nu \tag{2.18}$$

Note that flux and flux density *do* depend on distance through the inverse square law:

$$S_\nu \propto 1/d^2, \quad F \propto 1/d^2 \tag{2.19}$$

The total power measured by the telescope will then be $P = F \; A_{eff}$ where A_{eff} is the effective aperture area of the telescope. This ignores efficiency losses. Including polarisation and efficiency losses (η), Equation **??** becomes:

$$\boxed{P = \frac{1}{2}\eta F A} \tag{2.20}$$

To find the total power per unit frequency (units W Hz^{-1}) that a source is radiating, we can extend the flux density over the entire sphere to determine the **spectral luminosity** (L_ν, in units W Hz^{-1}):

$$L_\nu = \int S_\nu \; dA = 4\pi d^2 S_\nu \tag{2.21}$$

The **bolometric luminosity** (L_{bol}, in units W, sum of all radiation emitted from the source) is then defined as:

$$L_{bol} \equiv \int_0^\infty L_\nu \; d\nu \tag{2.22}$$

2.5 Radio Antennas

Most radio antennas use parabolic dishes to increase the directivity of the beam. However, even an ideal dish will be diffraction limited:

$$\Omega_A \approx \frac{\pi}{4}\theta^2 \approx \frac{\pi}{4}\left[\frac{4\lambda}{\pi D}\right]^2 = \frac{4}{\pi}\frac{\lambda^2}{D^2} \approx \frac{\lambda^2}{D^2} \tag{2.23}$$

[4]Radiative flux. This is different from other definitions of flux such as magnetic flux $\Phi_B = \iint \mathbf{B} \cdot d\mathbf{S}$.

Aside

The general equation to determine the beam solid angle of an antenna is:

$$\Omega_A = \int_{\text{sphere}} |F(\theta,\phi)|^2 \, d\Omega = \int_{\text{sphere}} P(\theta,\phi) \, d\Omega \tag{2.24}$$

where $F(\theta,\phi) = E/E_{\text{max}}$ is the normalised field pattern and P is the normalised power pattern (a.k.a., the beam pattern).
For an aperture antenna (e.g., waveguide horn or parabolic dish), the narrowest beam (minimum Ω_A/maximum directivity) is obtained when the E-field is uniform across the aperture. This gives:[a]

$$\Omega_A = \frac{\lambda^2}{A_p} = \frac{\lambda^2}{\pi(D/2)^2} = \frac{4}{\pi}\frac{\lambda^2}{D^2} \approx \frac{\lambda^2}{D^2} \tag{2.25}$$

The solid angle of a narrow cone (where the spherical cap's area is approximately the area of the base of the cone) is given by:

$$\Omega \approx \frac{\pi}{4}\theta^2 \tag{2.26}$$

The angular diameter θ of the uniform aperture antenna is then roughly $\theta \approx \frac{4\lambda}{\pi D}$ which is close to the well-known diffraction limit equation for circular apertures: $\theta \approx 1.22\frac{\lambda}{D}$.

[a] For proof see sec. 9.5 of "Antenna Theory and Design" by W. Stutzman and F. Thiele.

The measured flux density for an extended source[5] within the R.-J. limit will be:

$$S_\nu = \int I_\nu \, d\Omega = I_\nu \, \Omega_A \tag{2.27}$$

This can be integrated across the surface of the dish to find the **power spectral density (PSD):**

$$\begin{aligned} P_\nu &= \frac{1}{2} I_\nu \Omega_A A_p & (2.28) \\ &= \frac{1}{2}\frac{2\nu^2 kT}{c^2}\frac{4\lambda^2}{\pi D^2}\frac{\pi D^2}{4} & (2.29) \\ &= kT & (2.30) \end{aligned}$$

Aside

Johnson-Nyquist or thermal noise describes the noise in a resistor due to the random thermal motion of the electrons. It is typically written as:

$$v_n = \sqrt{4kTR\Delta f} \tag{2.31}$$

Or in terms of power spectral density as:

$$P_\nu = kT \tag{2.32}$$

The **power spectrum (P_ν) can therefore be described by a noisy resistor!** We can then set the power spectrum equal to

$$\boxed{T_a \equiv \frac{P_\nu}{k}} \tag{2.33}$$

where T_a is called the **antenna temperature**. Since T_a represents our signal, we can combine Equation 2.33 with Equation 2.4 to get the signal-to-noise ratio (S/N):

[5] Extended source: the source's beam solid angle Ω_s is larger than Ω_A.

$$\boxed{\frac{S}{N} = \frac{T_a}{T_{sys}} \sqrt{\Delta v \; \Delta t}} \tag{2.34}$$

In this ideal case, the antenna temperature will be equal to the brightness temperature (T_b). Real systems will have non-ideal beam patterns, losses, polarisation mismatch, and sources which can be smaller than the beam pattern. Therefore, $T_a \leq T_b$. For a source (Ω_S) that is smaller than the beam solid angle (Ω_A), the antenna temperature will be lower than the brightness temperature.

$$T_A \approx T_B \frac{\Omega_S}{\Omega_A} \tag{2.35}$$

The ratio Ω_S/Ω_A is known as the **beam filling factor.**

2.6 Heterodyne Receivers

Most radio astronomy telescopes use heterodyne receivers. This is because heterodyne receivers preserve phase information (important for interferometry) and provide extremely high spectral resolution ($\lambda/\Delta\lambda \approx 10^6$, limited by the backend). The block diagram for a millimetre-wave, heterodyne receiver is shown in Figure 2.4. Most radio receivers place filters and low-noise amplifiers (LNAs) as the first elements in the receiver chain, but LNAs don't exist yet in the millimetre-wave spectrum.

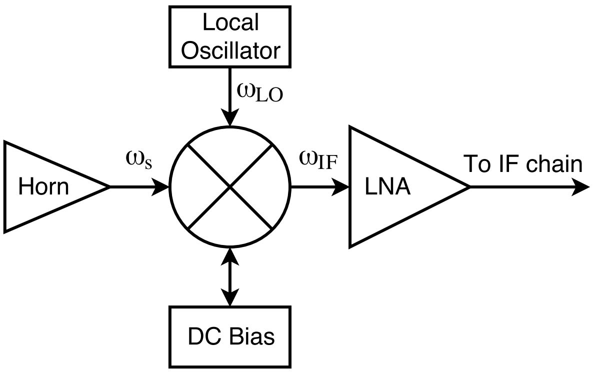


Figure 2.4: Block diagram of a heterodyne receiver. Mixers are traditionally represented by $\otimes$ in circuit diagrams.

Due to the way noisy components cascade,

$$T_{sys} = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} + \dots \tag{2.36}$$

$$T_{sys} \approx T_M + G_M^{-1} \; T_{IF} \tag{2.37}$$

the noise properties of the first in line components are the most important (the first components the signal encounters). Therefore, the noise temperature of the mixer (T_M) and the low-noise amplifier (T_{IF}) must be as low as possible, and the mixer's gain (G_M) should be kept close to 0 dB. Since the gain of the LNA is very high (~ 30 dB), T_{IF} is dominated by the noise from the LNA.

All heterodyne mixers are limited by Heisenberg's uncertainty principle,

$$\boxed{T_Q = \frac{hf}{k} = \frac{\hbar\omega}{k} \sim 11 \text{ K at } 230 \text{ GHz}} \tag{2.38}$$

since the phase is maintained through this process

Acknowledgements

In writing this chapter, I used several books and papers [**?**, **?**] as well as lecture notes from NRAO [**?**, **?**] and Mike Jones at the University of Oxford.